

1 / 138

SEQ ID NO:	Nucleotide/ Amino Acid	Description
1.	Nucleotide	Nucleotide sequence of Fv3A, a GH3 enzyme from <i>F. verticillioides</i>
2.	Amino acid	Protein sequence of Fv3A
3.	Nucleotide	Nucleotide sequence of Pf43A, a GH43 enzyme from <i>P. funiculosum</i>
4.	Amino acid	Protein sequence of Pf43A
5.	Nucleotide	Nucleotide sequence of Fv43E, a GH43 enzyme from <i>F. verticillioides</i>
6.	Amino acid	Protein sequence of Fv43E
7.	Nucleotide	Nucleotide sequence of Fv39A, a GH39 enzyme from <i>F. verticillioides</i>
8.	Amino acid	Protein sequence of Fv39A
9.	Nucleotide	Nucleotide sequence of Fv43A, a GH43 enzyme from <i>F. verticillioides</i>
10.	Amino acid	Protein sequence of Fv43A
11.	Nucleotide	Nucleotide sequence of Fv43B, a GH43 enzyme from <i>F. verticillioides</i>
12.	Amino acid	Protein sequence of Fv43B
13.	Nucleotide	Nucleotide sequence of Pa51A, a GH51 enzyme from <i>P. anserina</i>
14.	Amino acid	Protein sequence of Pa51A
15.	Nucleotide	Nucleotide sequence of Gz43A, a GH43 enzyme from <i>G. zeae</i>
16.	Amino acid	Protein sequence of Gz43A
17.	Nucleotide	Nucleotide sequence of Fo43A, a GH43 enzyme from <i>F. oxysporum</i>
18.	Amino acid	Protein sequence of Fo43A
19.	Nucleotide	Nucleotide sequence of Af43A, a GH43 enzyme from <i>A. fumigatus</i>
20.	Amino acid	Protein sequence of Af43A
21.	Nucleotide	Nucleotide sequence of Pf51A, a GH51 enzyme from <i>P. funiculosum</i>
22.	Amino acid	Protein sequence of Pf51A
23.	Nucleotide	Nucleotide sequence of AfuXyn2, a GH11 enzyme from <i>A. fumigatus</i>
24.	Amino acid	Protein sequence of AfuXyn2
25.	Nucleotide	Nucleotide sequence of AfuXyn5, a GH11 enzyme from <i>A. fumigatus</i>
26.	Amino acid	Protein sequence of AfuXyn5
27.	Nucleotide	Nucleotide sequence of Fv43D, a GH43 enzyme from <i>F. verticillioides</i>
28.	Amino acid	Protein sequence of Fv43D
29.	Nucleotide	Nucleotide sequence Pf43B, a GH43 enzyme from <i>P. funiculosum</i>
30.	Amino acid	Protein sequence of Pf43B
31.	Nucleotide	Nucleotide sequence of Fv51A, a GH51 enzyme <i>F. verticillioides</i>
32.	Amino acid	Protein sequence of Fv51A
33.	Nucleotide	Nucleotide sequence of Cg51B, a GH51 enzyme from <i>C. globosum</i>
34.	Amino acid	Protein sequence of Cg51B
35.	Nucleotide	Nucleotide sequence of Fv43C, a GH43 enzyme from <i>F. verticillioides</i>

FIG. 1A

2 / 138

SEQ ID NO:	Nucleotide/ Amino Acid	Description
36.	Amino acid	Fv43C protein sequence
37.	Nucleotide	Nucleotide sequence of Fv30A, a GH30 enzyme from <i>F. verticillioides</i>
38.	Amino acid	Fv30A protein sequence
39.	Nucleotide	Nucleotide sequence of Fv43F, a GH43 enzyme from <i>F. verticillioides</i>
40.	Amino acid	Fv43F protein sequence
41.	Nucleotide	Nucleotide sequence of Xyn3, a GH10 xylanase from <i>T. reesei</i>
42.	Amino acid	Xyn3 protein sequence
43.	Amino acid	Protein sequence of Xyn2, a GH11 xylanase from <i>T. reesei</i>
44.	Amino acid	Protein sequence of Pa3C
45.	Amino acid	Protein sequence of Bxl1, a GH3 β -xylosidase from <i>T. reesei</i>
46.	Nucleotide	Deduced cDNA for Pa51A.
47.	Nucleotide	Codon optimized cDNA for Pa51A.
48.	Nucleotide	Coding sequence for CBH1 signal sequence upstream of genomic DNA encoding mature Gz43A.
49.	Nucleotide	Coding sequence for CBH1 signal sequence upstream of genomic DNA encoding mature Fo43A.
50.	Nucleotide	Nucleotide sequence for CBH1 signal sequence upstream of codon optimized DNA encoding Pf51A
51.	Nucleotide	Nucleotide sequence of Eg4, an endoglucanase from <i>T. reesei</i>
52.	Amino acid	Protein sequence of <i>T. reesei</i> Eg4
53.	Nucleotide	Nucleotide sequence of Pa3D, a GH3 β -glucosidase from <i>P. anserina</i>
54.	Amino acid	Protein sequence of Pa3D
55.	Nucleotide	Nucleotide sequence of Fv3G, a GH3 β -glucosidase from <i>F. verticillioides</i>
56.	Amino acid	Protein sequence of Fv3G
57.	Nucleotide	Nucleotide sequence of Fv3D, a GH3 β -glucosidase from <i>F. verticillioides</i>
58.	Amino acid	Protein sequence of Fv3D
59.	Nucleotide	Nucleotide sequence of Fv3C, a GH3 β -glucosidase from <i>F. verticillioides</i>
60.	Amino acid	Protein sequence of Fv3C
61.	Nucleotide	Nucleotide sequence of Tr3A, a GH3 β -glucosidase from <i>T. reesei</i>
62.	Amino acid	Protein sequence of Tr3A
63.	Nucleotide	Nucleotide sequence of Tr3B, a GH3 β -glucosidase from <i>T. reesei</i>
64.	Amino acid	Protein sequence of Tr3B
65.	Nucleotide	Nucleotide sequence of Te3A, a GH3 β -glucosidase from <i>T. emersonii</i> , optimized for expression in <i>T. reesei</i>
66.	Amino acid	Protein sequence of Te3A

FIG. 1B

SUBSTITUTE SHEET (RULE 26)

3 / 138

SEQ ID NO:	Nucleotide/ Amino Acid	Description
67.	Nucleotide	Nucleotide sequence of An3A, a GH3 β -glucosidase from <i>A. niger</i>
68.	Amino acid	Protein sequence of An3A
69.	Nucleotide	Nucleotide sequence of Fo3A, a GH3 β -glucosidase from <i>F. oxysporum</i>
70.	Amino acid	Protein sequence of Fo3A
71.	Nucleotide	Nucleotide sequence of Gz3A, a GH3 β -glucosidase from <i>G. zeae</i>
72.	Amino acid	Protein sequence of Gz3A
73.	Nucleotide	Nucleotide sequence of Nh3A, a GH3 β -glucosidase from <i>N. haematococca</i>
74.	Amino acid	Protein sequence of Nh3A
75.	Nucleotide	Nucleotide sequence of Vd3A, a GH3 β -glucosidase from <i>V. dahliae</i>
76.	Amino acid	Protein sequence of Vd3A
77.	Nucleotide	Nucleotide sequence of Pa3G, a GH3 β -glucosidase from <i>P. anserina</i>
78.	Amino acid	Protein sequence of Pa3G
79.	Amino acid	Protein sequence of Tn3B, a GH3 β -glucosidase from <i>T. neapolitana</i>
80.	Amino acid	Protein sequence of TrEGb or Eg7, from <i>T. reesei</i>
81.	Amino acid	Protein sequence of TIEG, from <i>Thielavia terrestris</i>
82.	Amino acid	Protein sequence of Tr6A from <i>T. reesei</i>
83.	Amino acid	Protein sequence of Tr7A from <i>T. reesei</i>
84.	Amino acid	Protein sequence motif 1 of GH61 family endoglucanases
85.	Amino acid	Protein sequence motif 2 of GH61 family endoglucanases
86.	Amino acid	Protein sequence motif 3 of GH61 family endoglucanases
87.	Amino acid	Protein sequence motif 4 of GH61 family endoglucanases
88.	Amino acid	Protein sequence motif 5 of GH61 family endoglucanases
89.	Amino acid	Protein sequence motif 6 of GH61 family endoglucanases
90.	Amino acid	Protein sequence motif 7 of GH61 family endoglucanases
91.	Amino acid	Protein sequence motif 8 of GH61 family endoglucanases
92.	Nucleotide	Nucleotide sequence of an Fv3C/Bgl3 chimeric β -glucosidase
93.	Amino acid	Protein sequence of an Fv3C/Bgl3 chimeric β -glucosidase
94.	Nucleotide	Nucleotide sequence of an Fv3C/Te3A/Bgl3 chimeric β -glucosidase
95.	Amino acid	Protein sequence of an Fv3C/Te3A/Bgl3 chimeric β -glucosidase
96.	Amino acid	Protein sequence motif 1 of a polypeptide of a chimeric β -glucosidase
97.	Amino acid	Protein sequence motif 2 of a polypeptide of a chimeric β -glucosidase
98.	Amino acid	Protein sequence motif 3 of a polypeptide of a chimeric β -glucosidase
99.	Amino acid	Protein sequence motif 4 of a polypeptide of a chimeric β -glucosidase
100.	Amino acid	Protein sequence motif 5 of a polypeptide of a chimeric β -glucosidase
101.	Amino acid	Protein sequence motif 6 of a polypeptide of a chimeric β -glucosidase

FIG. 1C

4 / 138

SEQ ID NO:	Nucleotide/ Amino Acid	Description
102.	Amino acid	Protein sequence motif 7 of a polypeptide of a chimeric β -glucosidase
103.	Amino acid	Protein sequence motif 8 of a polypeptide of a chimeric β -glucosidase
104.	Amino acid	Protein sequence motif 9 of a polypeptide of a chimeric β -glucosidase
105.	Amino acid	Protein sequence motif 10 of a polypeptide of a chimeric β -glucosidase
106.	Amino acid	Protein sequence motif 11 of a polypeptide of a chimeric β -glucosidase
107.	Amino acid	Protein sequence motif 12 of a polypeptide of a chimeric β -glucosidase
108.	Amino acid	Protein sequence motif 13 of a polypeptide of a chimeric β -glucosidase
109.	Amino acid	Protein sequence motif 14 of a polypeptide of a chimeric β -glucosidase
110.	Amino acid	Protein sequence motif 15 of a polypeptide of a chimeric β -glucosidase
111.	Amino acid	Protein sequence motif 16 of a polypeptide of a chimeric β -glucosidase
112.	Amino acid	Protein sequence motif 17 of a polypeptide of a chimeric β -glucosidase
113.	Amino acid	Protein sequence motif 18 of a polypeptide of a chimeric β -glucosidase
114.	Amino acid	Protein sequence motif 19 of a polypeptide of a chimeric β -glucosidase
115.	Amino acid	Protein sequence motif 20 of a polypeptide of a chimeric β -glucosidase
116.	Amino acid	Protein sequence motif 21 of a polypeptide of a chimeric β -glucosidase
206.	Amino acid	Protein sequence of a GH61/endoglucanase from <i>T. aurantiacus</i>
207.	Amino acid	Protein sequence of a GH61/endoglucanase from <i>T. aurantiacus</i>

FIG. 1D

5 / 138

Conserved residues inferred from alignment and structures
of TrEGb (or *T. reesei* Eg7, or TrEG7) (pdb: 2vtc) and TtEG (pdb: 3EII)

Protein	TtEG	TrEGb	TrEg4
Metal coordination	H19	H20	H22
Conserved surface patch	D42	D62	D61
Conserved surface patch	G44	G64	G63
Metal coordination	H86	H108	H107
Buried salt bridge	R153	R177	R177
Buried salt bridge	E155	E179	E179
Metal coordination	H160	H184	H184
Metal coordination	Q169	Q193	Q193
Metal coordination	Y171	Y195	Y195
Involved in activity	Y210	Y233	Y232
Disulfide	C56	C77	C77
Disulfide	C174	C197	C198

FIG. 2A

Conserved amino acids of CBM1 domains of TrEg4, Tr6A and Tr7A
inferred from alignment (Full length numbering)

CBM1	TrEg4	Tr6A	Tr7A
	G313	G32	G483
	Q314	Q33	Q484
	C315	C34	C485
	G316	G35	G486
	G317	G36	G487
	S321	S40	S491
	G322	G41	G492
	P323	P42	P493
	T324	T43	T494
	C326	C45	C496
	A327	A46	A497
	T331	T50	T501
	C332	C51	C502
	N336	N55	N506
	Y338	Y57	Y508
	Y339	Y58	Y509
	Q341	Q60	Q511
	C342	C61	C512
	L343	L62	L513

FIG. 2B

6 / 138

Conserved active site residues in Fv3C homologs. Selection based on crystal structure of *Thermotoga neapolitana* Bgl3B complexed with glucose in -1 subsite (pdb: 2X41).

Enzyme	Tn3B	Fv3G	Fv3D	Tr3A	Pa3D	Te3A	An3A	Tr3B	Nh3A	Gz3A	Fv3C	Fo3A	Pa3G	Vd3A
Substrate interaction	D58	D101	D111	D92	D87	D92	D92	D99	D106	D106	D119	D119	D101	D107
	R64	R107	R117	R98	R93	R98	R98	R105	R112	R112	R125	R125	R107	R113
	L116	L150	L160	L141	L136	L141	L141	L148	L155	L155	L168	L168	L150	L156
	R130	R165	R175	R156	R151	R156	R156	R163	R170	R170	R183	R183	R165	R171
	K163	K198	K208	K189	K184	K189	K189	K196	K203	K203	K216	K216	K198	K204
	H164	H199	H209	H190	H185	H190	H190	H197	H204	H204	H217	H217	K199	H205
	R174	R209	R219	R200	R195	R200	R200	R207	R214	R214	R227	R227	R209	R215
	M207	M237	M266	M232	M227	M242	M245	M252	M259	M259	M272	M272	M254	M260
	Y210	Y240	Y269	Y235	Y230	Y245	Y248	Y255	Y262	Y262	Y275	Y275	Y257	Y263
Nucleophile	D242	D272	D301	D267	D262	D277	D277	D287	D294	D294	D307	D307	D289	D295
Substrate interaction	W243	W273	W302	W268	W263	W278	W278	W288	W295	W295	W308	W308	W290	W296
	S370	S455	S472	S415	S406	S447	S451	S457	S464	S464	S477	S477	S458	S465
Acid/Base	E458	E509	E534	E472	E463	E505	E509	E516	E523	E523	E536	E536	E517	E524

FIG. 3

7 / 138

Protein composition of <i>T.reesei</i> Integrated strain H3A	
Protein	% of Total Area
Fv3A	9.6
Fv51A+Fv43D	14.8
Xyn 3	12.6
Bgl 1	7.5
CBH1	36.4
EGLs	5.6
CBH2	9.5
Other	4.0

FIG. 4

Proteins added to <i>T. reesei</i> integrated strain H3A		
	Protein	Stock Protein Concentration (mg/ml)
1	Purified <i>T. reesei</i> CBH1	7.4
2	Purified <i>T. reesei</i> CBH2	3.0
3	Purified <i>T. reesei</i> EGI	3.9
4	Unpurified Fv3C	2.1
5	Water	
6	Purified <i>T. reesei</i> EG4	1.1
7	H3A UF concentrate	102.8
8	Purified <i>T. reesei</i> Bgl1	3.9
9	Purified <i>T. reesei</i> Xyn2	2.6
10	Purified <i>T. reesei</i> Xyn3	4.6
11	Purified <i>F. verticillioides</i> Fv43D	6.8
12	Purified <i>F. verticillioides</i> Fv51A	7.8

FIG. 5

8 / 138

Dosing chart for testing range of EG4 concentrations for improved saccharification of dilute ammonia pretreated corncob							
Vial	Water mL	6N Sulfuric Acid (mL)	Substrate (g)	H3A or #27 (mL)	Purified EG4 (mL)	Volume (mL)	Sample Description
1	2.931	0.026	1.866	0.177		5	100% #27
2	2.982	0.026	1.866	0.127		5	100% H3A
3	2.874	0.026	1.866	0.114	0.120	5	90% H3A 10% EG4
4	2.766	0.026	1.866	0.101	0.241	5	80% H3A 20% EG4
5	2.551	0.026	1.866	0.076	0.482	5	60% H3A 40% EG4
6	2.335	0.026	1.866	0.051	0.723	5	40% H3A 60% EG4
7	2.119	0.026	1.866	0.025	0.963	5	20% H3A 80% EG4
8	2.896	0.026	1.866	0.127	0.086	5	14 H3A + 1 EG4
9	2.724	0.026	1.866	0.127	0.258	5	14 H3A + 3 EG4
10	2.551	0.026	1.866	0.127	0.430	5	14 H3A + 5 EG4
11	2.121	0.026	1.866	0.127	0.860	5	14 H3A + 10 EG4

FIG. 6

Dosing chart for testing range of EG4 concentrations (0.05 to 1.0 mg/g) for improved saccharification of dilute ammonia pretreated corn cob							
Vial	Water mL	6N Sulfuric Acid (mL)	Substrate (g)	H3A or #27 (mL)	Purified EG4 (mL)	Volume (mL)	Sample Description
1	2.9	0.0261	1.87	0.177		5.0	14 mg/g H3A/EG4#27
2	2.8	0.0261	1.87	0.177	0.086	5.0	14mg/g H3A/EG4#27 + 1 mg/g EG4
3	3.0	0.0261	1.87	0.127		5.0	14mg/g H3A
4	3.0	0.0261	1.87	0.127	0.004	5.0	14mg/g H3A + .05 mg/g EG4
5	3.0	0.0261	1.87	0.127	0.009	5.0	14mg/g H3A + 0.1 mg/g EG4
6	2.9	0.0261	1.87	0.127	0.043	5.0	14mg/g H3A + 0.5 mg/g EG4
7	2.9	0.0261	1.87	0.127	0.086	5.0	14mg/g H3A + 1.0 mg/g EG4

FIG. 7A

10 / 138

Dosing chart for testing range of EG4 concentrations (0.1 to 0.5 mg/g) for improved saccharification of dilute ammonia pretreated corn cob							
Vial	Water mL	6N Sulfuric Acid (mL)	Substrate (g)	H3A or #27 (mL)	Purified EG4 (mL)	Volume (mL)	Sample Description
1	2.9	0.0261	1.87	0.177		5.0	14mg/g #27
2	3.0	0.0261	1.87	0.127		5.0	14mg/g H3A
3	3.0	0.0261	1.87	0.127	0.009	5.0	14 mg/g H3A + 0.1 mg/g EG4
4	3.0	0.0261	1.87	0.127	0.017	5.0	14 mg/g H3A + 0.2 mg/g EG4
5	3.0	0.0261	1.87	0.127	0.026	5.0	14 mg/g H3A + 0.3 mg/g EG4
6	2.9	0.0261	1.87	0.127	0.034	5.0	14 mg/g H3A + 0.4 mg/g EG4
7	2.9	0.0261	1.87	0.127	0.043	5.0	14 mg/g H3A + 0.5 mg/g EG4

FIG. 7B

	CBH1-CBH2	CBH1-CBH2	CBH1-CBH2	CBH1-EG2	CBH2-EG2	CBH1-CBH2-EG2
%	80-20	50-50	20-80	90-10	90-10	70-20-10

FIG. 8A

11 / 138

CBH1 (mg/g glucan)	CBH2 (mg/g glucan)	EG2 (mg/g glucan)	EG4 (mg/g glucan)	Glucan Conversion (%)
			20	18.6
1.25			18.75	42.7
2.5			17.5	46.9
5			15	59.1
10			10	74.8
20				68.7
			20	18.3
	1.25		18.75	32.1
	2.5		17.5	38.2
	5		15	35.9
	10		10	41.7
	20			24.9
			20	17.6
		1.25	18.75	24.3
		2.5	17.5	26.3
		5	15	24.3
		10	10	29.2
		20		23.1
				12.4
1.25				28.1
2.5				34.1
5				40.0
10				52.9
20				68.2
				12.5
	1.25			15.9
	2.5			17.3
	5			19.9
	10			22.1
	20			26.2
				12.4
		1.25		15.0
		2.5		16.6
		5		17.0
		10		19.8
		20		22.1
			20	16.3
			18.75	17.4
			17.5	17.4
			15	16.2
			10	15.4
				11.1

FIG. 8B-1
SUBSTITUTE SHEET (RULE 26)

12 / 138

CBH1 (mg/g glucan)	CBH2 (mg/g glucan)	EG2 (mg/g glucan)	EG4 (mg/g glucan)	Glucan Conversion (%)
			20	22.8
1	0.25		18.75	56.6
2	0.5		17.5	67.0
4	1		15	77.4
8	2		10	102.0
16	4			65.5
			20	23.1
0.625	0.625		18.75	51.5
1.25	1.25		17.5	73.8
2.5	2.5		15	82.5
5	5		10	100.7
10	10			76.1
			20	30.5
0.25	1		18.75	58.0
0.5	2		17.5	69.7
1	4		15	74.5
2	8		10	85.6
4	16			60.4
			20	29.5
1.125		0.125	18.75	55.1
2.25		0.25	17.5	71.1
4.5		0.5	15	86.3
9		1	10	90.3
18		2		54.2
			20	30.3
	1.125	0.125	18.75	51.7
	2.25	0.25	17.5	66.4
	4.5	0.5	15	73.1
	9	1	10	72.6

FIG. 8B-2

13 / 138

Yield of Xylose Monomer released by EG4 on hydrolysis of dilute ammonia pretreated corncob			
mg/g Enzyme added per gram Glucan + Xylan	H3A	EG4 + 1.12 mg/g Xyn3	H3A/EG4 #27
1.7	30.1%	21.9%	36.1%
6.0	65.7%	23.0%	73.9%
8.0	70.1%	24.1%	79.9%
14.0	76.1%	23.5%	88.1%
21.0	80.5%	25.7%	92.0%

FIG. 9

Percent Yield of Glucose Monomer released by EG4 on hydrolysis of dilute ammonia pretreated corncob			
mg/g Enzyme added per gram Glucan + Xylan	H3A	EG4 + 1.12 mg/g Xyn3	H3A/EG4 #27
1.7	22.4%	11.0%	25.0%
6.0	45.7%	12.7%	67.6%
8.0	52.7%	13.2%	75.5%
14.0	65.4%	14.1%	90.4%
21.0	74.2%	15.4%	97.9%

FIG. 10

14 / 138

Total Fermentable Monomers (mg/ml) released by EG4 on hydrolysis of dilute ammonia pretreated corncob.			
mg/g Enzyme added per gram Glucan + Xylan	H3A	EG4 + 1.12 mg/g Xyn3	H3A/EG4 #27
1.7	45	27	52
6.0	95	30	122
8.0	105	31	134
14.0	12	132	155
21.0	132	35	164

FIG. 11

Table 6-1: Effect of addition of purified EG4 on glucose release from dilute ammonia pretreated corncob		
EG4 added (mg/g)	Xyn3 added (mg/g)	Mg/mL Glucose monomer released
0.53	0.53	3.4
0	0.53	0.77

FIG. 12

15 / 138

Beta-glucosidase activities of *T. reesei* Bgl1, *A. niger* Bglu, Fv3C, Fv3D, and Pa3C on cellobiose and CNPG

Enzyme	Concentration (mg/ml)	Cellobiase Activity		CNPG nM CNP/sec/mg Protein
		U/ml	U/mg	
<i>T. reesei</i> Bglu1, Purified	2.3	19.4	8.4	1242
Fv3C Shake Flask	2.4	42.7	18	1156
Fv3D Shake Flask	2.9	0.0	0.0	6221
Pa3C Shake Flask	1.9	0.0	0.0	2
<i>A. niger</i> Bglu1, Purified	2.4	244	102	168

FIG. 13

**Comparison of Enzyme Mixes in Ammonia
Pre-Treated Corncob Saccharification.**

Relative weight of
enzymes in a purified
hemicellulase mix as
used in Example 19.

Enzyme	w/w%
Xyn 3	45.0
Fv3A	15.0
Fv43D	5.0
Fv51A	35.0

FIG. 14

Enzyme Mix	Dose (mg protein /g cellulose) for Glucan Conversion:		
	70%	80%	90%
Accellerase 1500 + Multifect Xylanase	38	-	-
whole cellulase from <i>T. reesei</i> integrated strain (H3A)	21	28	36
75% whole cellulase from <i>T. reesei</i> integrated strain (H3A) / 25% Fv3C	12	15	19

FIG. 15

16 / 138

SEQ ID NO:1**Nucleotide sequence for Fv3A, a GH3 family enzyme from *Fusarium verticillioides***

atgctgctcaatcttcaggctcgtgcccagcgctttgtcgctttctcttttaggtggattggctgaggctg
 ctacgccatatacccttcggactgtaccaaaggacctttgagcaagaatggaatctgcgatacttcgtt
 atctccagctaaaagagcggctgctctagttgctgctctgacgcccgaagagaagggtgggcaatctggc
 aggtaaaatataccccccccccataatcactattcggagattggagctgacttaacgcagcaatgcaactg
 gtgcaccaagaatcggacttccaaggtacaactgggtggaacgaagcccttcattggcctcgttgatctcc
 aggtggctcgctttgccgacactcctccctacgaacggcgccacatcatttcccatgcctcttctcatggcc
 gctgctttcgacgatgatctgatccacgatatcggaacgctcgctggcaccgaagcggctgcttcaacta
 acggcgggttggcgcgaggctgacttctggacacccaacgctcaacccttttaagatcctcgtggggctcg
 tggctccgaaactccaggtgaagatgcccttcattgtcagccgggtatgctcgtatatcgtcaggggtctc
 gaaggcgataaggagcaacgacgtattgttgctacctgcaagcactatgctggaaacgactttgaggact
 ggggaggcttcacgcgtcagcactttgatgccaagattactcctcaggacttggctgagtactacgtcag
 gcccttccaggagtgcaccgctgatgcaaagggttgggtccatcatgtgcccctacaatgccgtgaacggc
 attcccgcatgcgcaaactcgtatctgcaggagacgatcctcagagggcactggaactggacgcgcgata
 acaactggatcactagtgtattgtggcgccatgcaggatatctggcagaatcacaagtatgtcaagaccaa
 cgctgaagggtgccaggttagcttttgagaacggcatggattctagctgcgagtatactactaccagcgat
 gtctccgattcgtacaagcaaggcctcttgactgagaagctcatggatcggttcgttgaagcgccctttcg
 aagggttggctcactactgggtttctttgacgggtgccaagcgcaatggaactcgtcagttttgcggtatgt
 caacaccaaggaagctcaggatcttgactcagatctgctgtggagggtgctgttcttcttaagaatgac
 ggcactttgcctctgaagctcaagaagaaggatagtgttgcaatgatcggattctgggccaacgatactt
 ccaagctgcagggtggttacagtggacgtgctccgttctccacagcccgttttatgcagctgagaagct
 tggctctgacaccaacgtggcttgggttccgacactgcagaacagctcatctcatgataactggaccacc
 aatgctgttgctgcggcggaagaagtctgattacattctctactttgggtggtcttgacgcctctgctgctg
 gcgaggacagagatcgtgagaaccttgactggcctgagagccagctgacccctcttcagaagctctctag
 tctcggaagccactgggttgggtatccagcttgggtgatcaagtcgatgacaccgctcttttgaagaacaag
 aagattaacagttattctttgggtcaattacctgggtcaggatggcgccactgcagtcattggacctgctca
 ctggacgaaagagtctgctggccgactaccgctcacgcaatatccagtaataactgagcagattgg
 catgactgacatggacctcagacctaccaagtcggttgccaggggagaacttatcgtggtactcaactcca
 gttcttccctacggctttggcctccactacaccaagttccaagccaagttcaagtccaacaagttgacgt
 ttgacatccagaagcttctcaagggtgcagtgctcaatactccgatactttgcgcgctgccccccatcca
 agttagtgtcaagaacacccggccgattacctccgactttgtctctctggtctttatcaagagtgaagtt
 ggacctaaagccttaccctctcaagacccttgcggttatggctcgttgcatgatgtcgcgccttcatcga
 cgaaggatatctcactggagtggacgttggataacattgcccgcaggggagagaatgggtgatttgggtgt
 ttatcctgggacttacactctgttgctggatgagcctacgcaagccaagatccagggttacgctgactgga
 aagaaggctattttggataagtgccctcaagaccaccaagctctcgtaa

FIG. 16A**SEQ ID NO:2****Protein sequence of Fv3A**

mlnlqvaasalslsllgqlaeaatpytlpdctkgplskngicdtslspakraaalvaaltpeekvgnlv
 snATGAPRIGLPRYNWWNEALHGLAGSPGGRFADTPPYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEA
 RAFTNGGWRGVDFWTPNVNPFKDPRWGRGSETPGEDALHVSRYARYIVRGLEGDKQRRIVATCKHYAGN
 DFEDWGGFTRHDFDAKITPQDLAEYYVRPFQECTRDQKVGSI MCAYNAVNGIPACANSYLQETILRGHWN
 WTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDssceytttsdvdsykgglltekldmrsl
 krlfeglvhtgffdgakagwnslsfadvntkeaqdlalrsavegAVLLKNDGTLPLKLLKKDSVAMIGFW
 ANDTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVAWGPTLQNSSSHDNWTTNAVA AAKKSDYILYFGGLD
 ASAAGEDRDRENLDWPESQLTLLQKLSSLGKPLVVIQLGDQVDDTALLKNKKINSILWVNYPGQDGGTAV
 MDLLTGRKSPAGRLPVTQYPSKYTEQIGMTDMDLRPTKSLPGRTYRWYSTPVLPGFGLHYTKfqakfks
 nkltfdiqklkkgcsaqysdtcalppiqvsvkntgritsdfvslvfiksevgpkpyplktlaaygrlhdv
 apsstkdislewtdlniarrgengdlvvypgtytllldeptgakiqvtltgkkaildkwpqdpksa

FIG. 16B

SUBSTITUTE SHEET (RULE 26)

17 / 138

SEQ ID NO:3**Nucleotide sequence for Pf43A, a GH43 family enzyme from *Penicillium funiculosum***

atgcttcagcgatttgccttatatatttaccactggctctattgagtggttgagtgaaagccgacaaaccct
ttgtgcagagcatctacaccgctgatccggcacccgatggtatacaatgaccgcgtttatgtcttcattgga
ccatgacaacacccggagctacctactacaacatgacagactggcatctgtttctcgtcagcagatatggcg
aattggcaagatcatggcattccaatgagcctggccaatttcacctgggccaacgcgaatgcgtgggccc
cgcaagtcacccctcgcaacgggccaattctacttttatgtctctgtccgacacaacgatggttctatggc
tateggtgtgggagtgagcagcaccatcacagggtccataccatgatgctatcggcaaaccgctagtagag
aacaacgagattgatccaccgctgttcacgcagatgacggtcaggcataacctgtactggggaaatccag
acctgtgggtacgtcaaattgaaccaagatatgatatcgtagacgggagccctactcagattccactcac
cacggctggatttggtagctgaacgggcaatgctcaacggccgaccacttttgaagaagctccatgggta
tacaacgcaacggcatctactatatcgccatgacagccgattgttgttctgaggatattcgctactcca
cggaaccagtgccactgggtccgtggacttatcgaggcgtcatcatgccgaccaaggtagcagcttcac
caatcacgagggattatcgacttccagaacaactcctactttttctatcacaacggcgctcttcccggc
ggaggcggtaccaacgatctgtatgtgtggagcaattcaaatacaatgcagatggaaccattccgacga
tcgaaatgaccaccgcccgtccagctcaaattgggactctcaacccttacgtgcgacaggaagccgaaac
ggcgcatggtcttcaggcatcactacggaggtttgtagcgaaggcggaattgacgtcggttttatcaac
aatggcgattacatcaaagttaaaggcgtagctttcggttcaggagcccattctttctcagcgcggttg
cttctgcaaatacgcgcgccactattgcaatacacctcggaagcacaactgggtacgctcgtgggcacttg
tactgtccccagcactggcggttggcagaacttggaactaccgttacctgttctgtcagtgggcgcatctggg
accaggatgtgtattttgttttcggtggttagcggaacaggataacctgttcaactttgattattggcagt
tcgcataa

FIG. 17A**SEQ ID NO:4****Protein sequence of Pf43A**

mlqrfayilplallsvqvkaadnpfvqsiytadpapmvynndrvyvfmdhdntgatyyntdwhlfssadma
nwqdhgipmslanftwananawapqviprngqfyfyapvrhndgsmagvgvsstitgpyhdaigkplve
nneidptvfiddgqaylywgnpdlwyvklngdmisysgsptqiplttagfgtrtgnagrpttfeeapwv
ykrngiyyiayaadccsedirystgtsatgpwtyrgvimptqgssftnhegiidfqnnsyffyhngalpg
gggyqrsvcveqfkynadgtiptiemttagpaqigtlnpyvrqEAETAAWSSGITTEVCSEGGIDVGFIN
NGDYIKVKGVAFGSGAHSFSARVASANSGGTIAIHLGSTTGTLVGTCTVPSTGGWQWTWTTVTCVSGASG
TQDVYFVFGSGGTGYLFNFDYWQFa

FIG. 17B

18 / 138

SEQ ID NO:5**Nucleotide sequence for Fv43E, a GH43 family enzyme from *Fusarium verticillioides***

atgaagggtataactggctcgtggcgtggggccacttcttttgacgcgcggcactggctggccttgattggacacc
 gtcgcgccaccaccttcaacaatcctatcatctactcagactttccagataacgatgtattcctcgggtcc
 agataactactactactttctctgcttccaacttccacttcagcccaggagcaccgcgttttgaagtctaaa
 gatctgctaaaactgggatctcatcgggccattcaattccccgcctgaactttggcgacgggotatgatcttc
 ctctggtcactcagcttattaccgtggagggtacttgggcacatccctcagatacagaaagagcaatggaca
 gtgggtactggatcggctgcacacttctggcagacactgggtatacactgcctcatcgccggaagggtcca
 tgggtacaacaagggaaacttcgggtgataacaattgctactacgacaatggcatactgatcgatgacgatg
 ataccatgtatgtcgtatacgggttcgggtgaggtcaaagtatctcaactatctcaggacggattcagcca
 gggtcaaatctcaggtagttttcaagaacactgatattgggggtccaagacttggagggtaaccgcgatgtac
 aagatcaacggggtctactatatactaaacgatagcccaagtggcagtcagacctggatttggaggtcga
 aatcacccctggggcccttatgagtctaaaggtcctcgccgacaaagtcaccccgctatctctgggtggtaa
 ctgcgcgcacacagggtagtctcataaagactcccaatgggtggctgggtacttcatgtcattcacttgggcc
 tatectgcccggccgtcttccgggttcttgcaccgattacgtggggtagcgatggtttccccattcttgtca
 aggggtgctaattggcgggatggggatcatcttaccacaacacttccctggcacggatgggtgtgacaaagaattg
 gacaaggactgataccttcgcgggaacctcacttgcctcgtcctgggagtggaaccataatccggacgtc
 aactccttccactgtcaacaacggcctgactctccgcactgctagcattacgaaggatatttaccaggcgga
 ggaacacgctatctcacgaactcatgggtgatcatccaacaggaatagtgaagattgatttctctccgat
 gaaggacgggcgaccgggcccgggctttcagcgtttcagaccaaagtgcatacatcggtattcatcgagat
 aacggaaagttcacaaatcgctacgaagcatgggatgaatatggatgagtggaacggaacaacaacagacc
 tgggacaaataaaaagccacagctaattgtgccttctggaaggaccaagatctggctgagacttcaacttga
 taccaaccacagcaggaactggcaacactatcttttcttacagttgggatggagtcaggtatgaaacactg
 ggtcccaacttcaaactgtacaatgggttgggcattcttttattgottaccgattcggcatcttcaacttcg
 ccgagacggcctttaggagggctcgatcaagggttgagtctttcacagctgcatag

FIG. 18A**SEQ ID NO:6****Protein sequence of Fv43E**

mkvywlvawatsltpalaglighrrattfnnpiiysdfpdndvflgpdnyyyfsasnfhfspgapvlksk
 dllnwdlighsiprlnfgdgydlppgsryyrggtwasslryrksnggwywigcinfwqtwvytasspegp
 wynkgnfgdnncyydngiliddddtmyvvysgevkvsqslsqdgfsqvksqvfkntdigvqdlegnrm
 kinglyyilndspsgsqtwiwskspswgyeskladkvtppisggnspggsliktpnggwyfmsftwa
 ypagrlpvlapitwgsdgfpilvkganggwssyptlpgt dgvtknwtrtdtfrgtslapswewnhnpdv
 nsftvnnlglrtasitkdiygarnltlshrthgdhptgivkidfspmkgddraglsafrdqsayigihrd
 ngkftiatkhgmnmdewngtttldlgqikatanvpsgrtkiwlrlqldtnpagtgntifsyswdgvkyetl
 gpnfklyngwaffiayrfgifnfaetalggsikvesftaa

FIG. 18B

19 / 138

SEQ ID NO:7**Nucleotide sequence for Fv39A, a GH39 family enzyme from *Fusarium verticillioides***

atgcactacgctaccctcaccacttttggtgctggctctgaccaccaacgctcgctgcacagcaaggcacag
caactgtcgacctctccaaaaatcatggaccggcggaaggcccttggttcaggcttcatatacggctggcc
tgacaacggaacaagcgctcgacacctccataccagatttcttggttaactgacatcaaattcaactcaaac
cgcggcggtggcgcccaaatcccatactgggttggggccagaggtggctatgaaggatacctcggccgct
tcaactcaaccttatccaaactatcgccaccacgcgcgaagtataacgctgactttatcttggttgcctcatga
cctctggggtgcggtatggcgggcaggggttcaaaactccccgtttcctggcgacaatggcaattggactgag
atggagttattctggaatcagcttgtgtctgacttgaaggctcataatatgctggaaggctcttggtgattg
atgtttggaatgagcctgatattgatatcttttgggatcgcccggtggcgcagtttcttgagtattacaa
tcgcgcgaccaaactacttcggtgagtctactactgatccatacgtattttacagtgaagctgactggctga
attagaaaaacacttcccaaaactcttctcagtgggcccagccatggcacattctcccattctgtccgatg
ataaatggcatacctggcttcaatcagtagcgggtaacaagacagtccttgatatttactcctggcatca
gattggcgcttggaacgtgagccggacagcactatccccgactttaccacottgcgggcgcaatatggc
gttcccgagaagccaattgacgtcaatgagtaecgtgcacgcgatgagcaaaatccagccaactccgtct
actacctctctcaactagagcgtcataacottagaggctcttcgcgcaaaactggggtagcggatctgacct
ccacaactggatgggcaacttgattttacagcactaccgggtacctcggaggggaacttactaccctaattggt
gaatggcaggttacaagtactatgcggccatggcagggcagagacttgtgaccaaaagcatcgctcggact
tgaagtttgatgtctttgcccactaagcaaggccgtaagattaagattatagccggcacgaggaccgttca
agcaaagtataacatcaaaatcagcgggtttggaagtagcaggacttcctaagatgggtaacggtaaaggctc
cggacttatcgggttcgactgggctgggcccgaatggaaagggttgacgggacctgttgatttggggggagaaga
agtatacttattcggccaatacgggtgagcagccctctacttga

FIG. 19A**SEQ ID NO:8****Protein sequence of Fv39A**

mhyatllttlvlalttnvaaqqgtatvdlsknhgpakalgsgfiygwpdngtsvdtssipdfvltdikfn
rgggagipslgwarggyegylgrfnstlsnyrttrkynadfillphdlwgadggggsnspfpdngn
wtemlfwnqlvsdlkahnmlleglvidvwnepdidifwdrpwsqfleyynratkllrktl
pktilslsgpamahspilsddkwhtwlqsvagnktvpdiyswhqigawerepdstipdf
ttlraqygvpekpidevneyaardeqnpansvyylsqlerhnlrglranwgsgsdlh
nwmgnliysttgtsegtyypngewqaykyyaamagqrlvtkassdlkfdvfatkqgr
kikilagtrtvqakynikisglevaglpkmgvtkvrtyrfdwagpngkvdgpvdl
gekkytysantvsspst

FIG. 19B

20 / 138

SEQ ID NO:9**Nucleotide sequence for Fv43A, a GH43 family enzyme from *Fusarium verticillioides***

atgtggctgacotccccattgctgttcgccagcaccctcctgggcctcactggcggttgcctctagcagaca
 accccatcgtccaagacatctacacccgagaccagcaccaatggctctacaatggccgcgtctacctctt
 cacaggccatgacaacgacggctctaccgacttcaacatgacagactggcggtctctctcgtcagcagac
 atgggtcaactggcagcaccatgggtgtccccatgagcttaaagaccttcagctgggccaacagcagagcct
 gggctgggtcaagtgcgttgcccgaacggaaagttttacttctatgttcctgtccgtaatgccaaagacggg
 tggaatggctattgggtgtcggtgttagtaccacatccttgggcccctacactgatgccttggaaagcca
 ttgggtcgagaacaatgagatcgacccaactgtctacatcgacactgatggccaggcctatctctaactggg
 gcaaccctggattgtactacgtcaagctcaaccaagacatgctctcctacagtggtagcatcaacaaagt
 atcgctcacaacagctggattcggcagccgcccgaacaacgcgcagcgtcctactacttttcgaggaagga
 ccgtggctgtacaagcgtggaaatctctactacatgatctacgcagccaactgctgttccgaggacattc
 gctactcaactggaccagcgccactggaccttggacttacccgcgggtgctcgtgatgaacaaggcgggtcg
 aagcttcaccaaccatcctggcatcatcgactttgagaacaactcgtacttcttttaaccacaatggcgct
 cttgatggaggttagcgggttatactcgggtctgtggctgtcgagagcttcaagtatgggttcggacgggtctga
 tccccgagatcaagatgactacgcaaggccagcgcagctcaagtctctgaaccatattgtcaagcagga
 ggccgagactatcgcttgggtctgaggggtatcgagactgaggtctgcagcgaaggtgggtctcaacgttgct
 ttcacgcacaatgggtgactacatcaagggtcaagggtgagtcgactttggcagcaccgggtgcaagacgttca
 ggcgccgtgtttgcttccaacagcagcgggaggaagattgagcttcgacttggtagcaagaccggtaagtt
 gggttggtacctgcacggtaacgactacgggaaactggcagacttataagactgtggattgccccgtcagt
 ggtgctactggtaacgagcagatctattctttgtcttcacgggctctggtctctctgttcaacttca
 actggtggcagtttagctaa

FIG. 20A**SEQ ID NO:10****Protein sequence of Fv43A**

mwltspllfastllqltgvaladnpivqdiytadpapmvyngrvylftghdndgstdfnmtdwrlfssad
 mvnwqhhgvpmslktfswansrawagqvvarngkfyfyvpvrnaktggmaigvgvstnilgpytdalgkp
 lvenneidptvyidtdgqaylywgnpglyyvklndmlsysgsinkvslttagfgsrpnnagrpttfeeg
 pwlykrnlyymiyaanccsedirystgpsatgpwtyrgvmmnkagrsftnhpgiidfennsyffyhnga
 ldggsgytrsvavesfkygsdglipeikmttqgpaqlksINPYVKQEAETIAWSEGIETEVCSEGLNVA
 FIDNGDYIKVRGVDFGSTGAKTF SARVASNSSGGKIELRLGSKTGKLVGTCTVTTTGNWQTYKTVDCPVS
 GATGTSDLFFVFTGSGSGSLEFNENWWQF's

FIG. 20B

21 / 138

SEQ ID NO:11**Nucleotide sequence for Fv43B, a GH43 family enzyme from *Fusarium verticillioides***

atgogcttctcttggctattgtgcccccttctagcgatgggaagtgccttctcctgaaacgaagacggatg
tttcgacatacaccaaccctgtccttccaggatggcactcggatccatcgtgtatccagaaagatggcct
ctttctctgogtcacttcaacattcatctccttcccagggtcttcccgtctatgcctcaagggatctagtc
aactggcgtctcatcagccatgtctggaaccgcgagaaacagttgcctggcattagctggaagacggcag
gacagcaacagggaatgtatgcaccaaccattcgataccacaaggggaacatactacgtcatctgcgaata
cctgggcgttggagatattatttgggtgtcatcttcaagaccaccaatccgtgggacgagagtagctggagt
gacctgtttacottcaagccaaatcacatcgaccccgatctgttctgggatgatgacggaaagggtttatt
gtgctacccatggcatcactctgcaggagattgatttggaaactggagagcttagcccgagcttaatat
ctggaacggcacaggaggtgtatggcctgaggggtcccatatctacaagcgcgacggttactactatctc
atgattggcaggggtggaactgccgaagaccacgctatcacaaatcgctcgggcccgaagatcacgggcc
cotatgaagcctacaataacaaccaatcttgaccaaccgcgggacatctgagtacttccagactgtcgg
tcacgggtgatctgttccaagataccaagggcaactggtggggctctttgtcttctactcgcatcacagca
caggaggtttcacccatgggcccgtgaagctgttttgttcaatggcacatggaacaagggcgaatggccca
agttgcaaccagtagcaggtcgcatgcctggaaacctctcccaaagccgacgcgaaacggttcccggaga
tgggccccttcaacgctgacccagacaactacaactgaagaagactaagaagatccctcctcactttgtg
caccatagagtcccaagagacgggtgccttctcttcttccaaagggtctgcacatcgtgcctagtgcga
acaacggttaccggtagtgtgttgccaggagatgagattgagctatcaggacagcaggtctagctttcat
cggacgcgcgcaaaactcacactctgttcaaataatagtggtgatatcgacttcaagcccaagtcgatgat
caggaagctggaatcacggttttccgcacgcagttcgaccatatcgatcttggcattgttcgtcttcccta
caaaccaaggcagcaacaagaaatctaagcttgccctccgattccggggccacaggagctcagaatgttcc
tgcaccgaaggtagtacgggtcccgatggctgggagaagggcgtaatcagttctacatatcgaggcagcc
aacgcgacgcactacaaccttgagcttcgagccacagaggcaagactctcgacatcgcgacagcatcag
caagtcttgtgagtggaggcacgggttcatttgttggtagtttgcttgaccttatgctacctgcaacgg
caaaggatctggagtggaatgtcccaaggaggtgatgtctatgtgacccaatggacttataagcccgtg
gcacaagagattgatcatgggtgtttttgtgaaatcagaattgtag

FIG. 21A**SEQ ID NO:12****Protein sequence of Fv43B**

mrfswllcp1lamgsalpetktdvstytnpvlpgwhsdpsciqkdglflevtstfifspglpvyasrdlv
nwrlishvwnrekqlpgiswktaggqqgmyaptiryhkgtyyviceylgvvgdiigvifktnpwdessws
dpvtfkpnhidpdlfwdddgvycathgitlqeidlletgelspelniwngtggvwpegphiykrdgyyyyl
miaeggtaedhaitiararkitgpyeaynnnpiltnrgtseyfqtvgdglfqdtkgnwwglclatrita
qgvspmgreavlfngtwnkgewpklqpvrgrmpgnllpkptrnvpdgpfnadpdnynlkktkkipphfv
hhrvprdgafslsskglihivpsrnnvtgsvlpgdeielsgqrglafigrqrqthtlfkysvdidfkpsdd
qeagitvfrtqfdhidlgivrlptnqgsnkksklafrfratgaqnvppkvvpdpdwekgvislhieaa
nathynlgasshrkgtldiatasaslvsaggtgsfvgsllgpyatcngksgsvecpkggdvyy
tqwtypvageidhgvfvksel

FIG. 21B

atgatccacctcaagccagccctcgccgcggttggtggcgctgtcgacgcaatgtgtggctattgattgtgt
ttgtcaagtcttcgggggggaataagacgactgatatcatgtatgggtcattatgcacgaggtatgtgtttt
gcgagatctcccttttgtttttgcgcaactgctgacatggagactgcaaacaggatatcaacaactccggc
gacggcggcacatacgcgcgagctaatactccaaccgcgcgttccaagggagtgagaagttccctccaacc
tcgacaactgggagccccgtcggtggcgctacccttacccttcagaagcttgccaagcccccttctctgc
gttgcccttactcogtcaatgttgccaaccccaaggaggggcaagggcaagggcaaggacaccaaggggaag
aaggttggttgcccaatgctgggttttgggtatggatgtcaagaggcagaagtacactggtagcttcc
acgttaactgggtgagtacaagggtgactttgaggttagcttgcgacgcgcgattaccgggggagacctttgg
caagaaggtggtgaaggggtgggagtaagaaggggaagtggaccgagaaggagtttgagttggtgccttcc
aaggatgcgcccaacagcaacaacacctttgttgtgcagtgggatgcccaggtatgtgcttctttgat
tggttgagatagaagttgggttgacatgatgtggtgcagggcgcaaaggacggatctttggatctcaact
tgatcagcttggtccctccgacattcaagggaagggaagaatgggctgagaattgatcttgcgacagcat
ggttgagctcaagccggttaagtcctctctagtcagaaaagtagagcctttgttaacgcttgacagacct
cttgcgcttcccggtggcaacatgctcgagggtaacaccttggaaccttggtggaagtggtaacgagacc
attggccctctgaaggatcgccccggcatggctggtgtctgggagtaccagcaaaccttggcttgggtc
tggtcgagtacatggagtgggcccgatgacatgaacttgagcccagtatgtgatcccatttctgagagt
acttctcttgctaacgtatccacagttgtcggtgtcttcgctggtcttgccctcgatggctcggttcgttc
ccgaatccgagatgggatgggtcatccaacaggctctcgacgaaatcgagttcctcactggcgatgctaa
gaccaccaaatgggggtgccgtccgcgcgaagcttgggtcaccccaagccttggaaggtcaagtgggttgag
atcggttaacgaggattggcttgcggacgccctgctggcttcgagtcgtacatcaactaccgcttcccca
tgatgatgaaggccttcaacgaaaagtaccccgacatcaagatcatcgctcgcctccatcttcgacaa
catgacaatcccgcgggtgctgccggtgatcaccacccgtacctgaactcccgatgagttcggttgagcga
ttcgccaagttcgataacttgagcaaggataacgtgacgctcatcggcgaggctgcgtcgacgcaccta
acggtgggtatcgcttgggagggagatctcatgcccttgccttgggtggggcggcagtggtgctgaggctat
cttcttgatcagcaactgagagaaaacggtgacaagatcatcggtgctacttacgcgcctggtcttcgcagc
ttggaccgctggcaatggagcatgacctgggtgcagcatgcgcgcgacccggccctcaccactcgctcga
ccagttgggtatgtctggagaatcctcgcccaccacatcatccgtgagacgctcccggtcgatgccccggc
cggcaagcccaactttgacctctgttctacgttgcgggaaagagcgcgagagtggcacccggtatcttcaag
gctgcogtctacaactcgactgaatcgatcccggtgtcggtgaagtttgatgggtctcaacgagggagcgg
ttgccaacttgacgggtgcttactgggcccggaggatccgtatggatacaacgaccccttcaactgggtatcaa
tgttgtcaaggagaagaccaccttcatcaaggccggaaaggcggaagtacacctcaccttgcggggc
ttgagtggtgctgtgttggtgagacggccgacgcggtcaaggggtggcaagggaagggaagggaaggga
agggttaactga

FIG. 22A

23 / 138

SEQ ID NO:14**Protein sequence of Pa51A**

mihlkpalaallalstqcvaidl fvkssggnktt dimyglmhedinnsgdgggiya elisnrafqgsekfp
snldnwspvggatitlqlakplssalpysv nvanpkegkgkgtkkgkkgkvglanagfwgmdvkrqkytg
sfhvtgeykgdfevslrsaitgetfgkkvkvkggskkgkwtekefelvpfk dapsnntf vvgwdaegakd
gsldlnlislfpptfkg rkngrlridlaqtmvelkptflrfpggnmlegntldtwwkwyetigplkdrpgm
agvweyqgtlglglveymewaddm nlepivgvfaglal dgsfvpesemgwviqqal deiefltgdakttk
wgavraklghpkpwkvkwveignedwlagrpagfes yinyrfpmmmkafnekypdikii aspsifdnmti
pagaagdhhpyltpdefverfakfdnl skdnvtligeaasthpnnggiawegdlmplpwwggsvaeaifli
sterngdkiigatyapglrsl drwqwsmtwvqhaadpalttrstswyvvrilahhiiretlpvdapagkp
nfdplfyvagksesgtgifkaavynstesipvslkfdglnegavanltvltgpedpygyndpftginvvk
ekttfikagkggkftftlpglsvavletadavkggkgkgkgkgkgn

FIG. 22B**SEQ ID NO:15****Nucleotide sequence for Gz43A, a GH43 family enzyme from *Gibberella zeae***

atgaagtccaagttgttattcc cactcctctctttcg ttgggtcaaagtcttgc caccacaacgacgactgtc
ctctcatcactagtagatgg actgcggatccttcgg gtcctgtctttaacgacaccttgtggctctaccc
gtctcatgacatcgatgctgg atttgagaatgatcct gatggaggccagta cgcctatgagagattacat
gtctactctatcgacaagat ctacggttccttgcc ggtcgatcacggta cggccctgtcagtgaggatg
tcccctgggcctctcgacag atgtgggctcctgac gctgcccacaagaac ggcaaatactacctatactt
cctgccc aaagacaaggatg atatcttcagaatcg gcgttgctgtctcac caacccccggcg gaccattc
gtccccgacaagagttggat ccctcacactttcag catcgaccccgccag tttcgtcgatgatgaca
gagcctacttggcatgggg tgggtatcatgggtgg ccagcttcaac gatggcaggataagaacaagtaca
cgaatctggcactgagccag gaaacggcaccgctgc ccttgagccctcag attgccaagctgagcaaggac
atgcacactctggcagaga agcctcgcgacatgct cattcttgacccca agactggcaagccgctcctt
ctgaggatgaagaccgacg cttcttcgaaggacct ggattcacaaagcg caacaagatttactacctac
ctactctactggcacaaccc actatcttgtctatgc gacttcaaagaccc cctatgggtccttacacctac
cagggcagaattctggagcc agttgatggctggacta ctactctagtatc gtcaagtaccaggggtcagt
gggtggctattttatcacg atgccaagacatctgg caaggactatcttc gcccaggtaaaggctaagaagat
ttggtacgatagcaaaggaa gatccttgacaaagaagc cttga

FIG. 23A

24 / 138

SEQ ID NO:16**Protein sequence of Gz43A**

mksklflfpllsfvqgslatnddcplitsrwtadpsahvfndtlwlypshdidagfendpdgggyamrdyh
vysidkiygslpvdhgtalsvedvpwasrqmwapdaahkngkyylyfpakdkddifrigvavsptpggpf
vpdkswiphtfsidpasfvddddraylawggimggqlqrwqdknkynesgtepgngtaalspqiaklskd
mhtlaekprdmliildpktgkpllsededrrffegpwihkrnkiiyltystgtthylvyatsktpygypty
qgrilepvdgwtthssivkyqggqwwlfyhdaktsgkdylrqvkakkiwydskgkiltkkip

FIG. 23B**SEQ ID NO:17****Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum***

atgcagctcaagtttctgtcttcagcattgctgttctctctgaccagcaaattgcgctgcgcaagacacta
atgacattcctcccctgatcaccgacctctggctcgcagatccctcggctcatgttttcgaaggcaagct
ctgggtttaccatctcagcagatcgaagccaatgttgtaacggcacaggagcgctcaatacgccatg
agggattaccatacctactccatgaagagcatctatggtaaagatcccggttgctcgaccacggcgctcgtc
tctcagtcgatgacgttccctgggccaagcagcaaattgtgggctcctgaacgagctcataagaacggcaa
atattatctgtacttccccgccaaggacaaggatgagatcttcagaattggagttgctgtctccaacaag
cccagcggctcctttcaaggccgacaagagctggatccctggcagctacagtatcgatcctgctagctacg
tcgacactgataacgaggcctacctcatctggggcggtatctggggcgggccagctccaagcctggcagga
taaaaagaactttaacgagtcgtggattggagacaaggctgctcctaacggcaccatgcctatctcct
cagatcgccaagctaagcaaggacatgcacaagatcacgaaacaccccgcatctcgtcattctcgccc
ccgagacaggcaagcctcttcaggctgaggacaacaagcgaacgattcttcgaggggcccttggtaccaca
gcggcggaagctttactacctcatgtactccaccggtgataccacttcttgctacgctacttccaag
aacatctacggctccttatacctaccggggcaagattcttgatcctggtgatgggtggactactcatggaa
gtattggtgagtataagggacagtggtggcttttctttgctgatgcgcatacgtctggttaaggattacct
tcgacaggtgaaggcgaggaagatctggtatgacaagaacggcaagattcttgcttcaccgctccttag

FIG. 24A**SEQ ID NO:18****Protein sequence of Fo43A**

mqlkflssallfsltskcaaqdtndipplitdlwsadpsahvfegklwvypshdieanvvngtggagqyam
rdyhtysmksiygkdpvvdhgvalsvddvpwakqqmwapdaahkngkyylyfpakdkdeifrigvavsnk
psgpfkadmkipgtysidpasyvdtneayliwgggiwggqlqawqdkknfneswigdkaapngtnalsp
qiaklskdmhkitetprdlvilapetgkplqaednkrrffegpwihkrnklyylmystgdthflvyatsk
niygyptyrgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkkiwydkngkillhrp

FIG. 24B

25 / 138

SEQ ID NO:19**Nucleotide sequence for Af43A, a GH43 family enzyme from *Aspergillus fumigatus***

atggcagctccaagtttatcctacccacaggtatccaatcgtataccaatcctctcttccctgggtggc
actccgatcccagctgtgcctacgtagcggagcaagacaccttttctgcgtgacgtccactttcattgc
cttccccgggtotctcctctttatgcaagccgagatctgcagaactggaaaactggcaagcaatattttcaat
cggcccagccagatccctgatcttcgcgtcacggatggacagcagtcgggtatctatgcgccactctgc
gctatcatgagggccagttctacttgatcgtttcgtacctgggcccgcagactaagggcttgctgttcac
ctcgtctgatcogtacgacgatgccgcgtggagcgatccgctcgaattcgcggtacatggcatcgacccg
gatatcttctgggatcacgacgggaagggtctatgtcacgtccgcccaggaccagatgattaagcagtaca
cactcgatctgaagacggggggcgattggcccgggttgactacctctggaacggcaccggaggagtctggcc
cgagggcccgcacatttacaagagagacggatactactacctcatgatcgcagagggaggtaaccgagctc
ggccactcggagaccatggcgcgatctagaacccggacaggctccctgggagccatacccgcacaaatccgc
tcttgtcgaacaaggggcacctcggagtacttccagactgtgggcatgcggacttggtccaggatgggaa
cggcaactggtgggcccgtggcggttgagcaccgatcagggcctgcattggaagaactatcccatgggtcgg
gagacgggtgctcgccccgcgcgttgggagaaggggtgagtggcctgtcattcagcctgtgagaggccaaa
tgcaggggcccgtttccaccaccaataagcgagttcctcgcggcgagggcggtgatcaagcaaccgca
caaagtggatttcaggcccggatcgaagataccggcgccacttccagtaactggcgatatcccaagacagag
gattttaccgtctccctcggggccaccogaatactcttcgggtcacaccctccttttacaacctcaccg
gaactgcggacttcaagccggatgatggcctgtcgttgttatgcgcaaacagaccgacaccttggtcac
gtacactgtggacgtgtcttttgacccccaaaggttgccgatgaagaggcgggtgtgactgttttcccttacc
cagcagcagcacatcgatcttggtattgtccttctccagacaaccgagggggtgtcgttgtccttccggt
tccgcgtggaaggccgcggtaactacgaaggtcctcttcagaagccaccgtgctgttcccaaggaatg
gtgtggacagaccatccggttgagattcaggccgtgagtgacaccgagtatgtctttgcggtgccccg
gctcggcacccctgcacagaggcaaatacatcagccgcgccaactcgttgattgtcagtggtgatacgggac
ggtttactggctcgttgttggcggtgatgccacgtcgaacgggggtgcccgatccacgcccgcataatat
cagcagatggagatacgaaggacggggccagatgattgattttggtcagtggtcccgagctactga

FIG. 25A**SEQ ID NO:20****Protein sequence of Af43A**

maapslsyptgiqsytnplfpgwhsdpscayvaeqdtffcvststfiafpglplyasrdlqnwklasnifn
rpsqipdlrvtdgqqsgiyaptlryheggfyilivsyqlgpgtkgllftssdpyddaawsdplefavhgidp
difwdhdgtvyvtsaedqmikqytlldlktgaigpvdylwngtggvwpegphiykrdgyyylmiaeggtel
ghsetmarsrtrtgpwepyphnpllsnkgstseyfqtvgadlfdqngnwwavalstrsgpawknypmgr
etvlapaawekgewpviqpvrqgmqgpfpppnkrvprgeggwikqpdkvdfprgskipahfqywrypkte
dftvsprghpntlrltspfyntgtadfkpddglslvmrkqtdtlftytdvvsfdpkvadecagvtvflt
qqqhldlgivllqtteglslsfrfrvegrgnyegplpeatvpvpkewcgqtirleiqavsdteyvfaaap
arhpaqrqiisranslivsgdtgrftgslvgvyatsnggagstpayisrwrvegrgqmidfgrvpsy

FIG. 25B

26 / 138

SEQ ID NO:21**Nucleotide sequence for Pf51A, a GH51 family enzyme from
*Penicillium funiculosum***

atgggaaagatgtggcattcgatcttggttggttggtggttattgtctgtcgggcatgccatcactatca
acgtgtcccaaagtggcgggaataagaccagtcctttgcaatatggtctgatgttcgaggtaatccttct
cttataccacatataaaaagttgcgtcatttctaagacaagtcaaggacataaatcacggcgggtgatggcg
gtctgtatgcagagcttggttcgaaaccgagcattccaaggtagcaccgtctatccagcaaacctcgatgg
atacgactcgggtcaatggagcaatcctagcgttcagaatttgacaaaacctctatcacccctccatgcct
agctctctcaacgtcggcaagggtccaacaatggaagcatcgggttcgcaaatgaaggctgggtggggga
tagaagtcaagccgcaagatacgcgggctcattctacgtccagggggactatcaaggagatttcgacat
ctctcttcagtcgaaattgacacaagaagtcttcgcaacggcaaaagtcagggtcctcgggcaaacacgag
gactgggttcaatacaagtacgagttgggtgccccaaaaggcagcatcaaacaccaataaacactctgacca
ttacttttgactcaaaggtatgttaaattttgggttttagttcgatgtctggcaattgtcttacgagaaac
gtagggttgaaagacggatccttgaacttcaacttgatcagcctatttcccccaacttacaacaatcgg
cccaatggcctaagaatcgacctgggtgaagctatgggtgaactagagggggtaagctcttacaatatcaa
ctttatctttacgaagactaatgtgaaaacttagaaaatttctgcggtttccaggcgggtagcgatgtggaa
ggtgtacaagctccttactggtataagtggatgaaacggtaggagatctcaaggaccggttatagtaggc
ccagtgcatggacgtacgaagaaagcaatggaattggcttgattgagtacatgaattgggtgtgatgacat
ggggcttgagccgagtgagtgatttccattcagcgtcaaateccagtggttctaatcatcacatcagttct
tgccgtatgggatggacattaccttccgaacgaagtgatatcggaacagatttgacgccatatatcgac
gacacctcaaccaactggaattcctgatgggtgcccagataacgccatatggtagttggcgtgcgtctc
tgggctatccgaagccgtggacgattaactacgtcgagattggaaacgaagacaatctatacgggggact
agaaacatacatcgctaccggtttcaggcatattacgacgtataacagctaaatatccccatatgacg
gtcatggaatctttgacggagatgcctgggtccggcggccgctgcaagcgattaccatcaatattctactc
ctgatgggtttgtttccagttcaactactttgatcagatgccagtcactaatagaacactgaacgggtat
gaaaaccccccttttttaaatatgcttttaattggtattaaccatctttcataggagagattgcaaccgt
ttatccaaataatcctagtaattcgggtggcctggggaagcccattcccccttgatatccttgggtggattggg
tcggttgacagaagctgttttcctaattggtgaagagaggaattcgccaaagataatcgggtgctagctacg
tacggaattctacttttcgagatttttaacattggataagaaggactaacctcaatacagggtccaatggt
cagaaatatcaacaattggcagtggtctccaacactcatcgcttttgacgctgactcgtcgcgtacaagt
cgttcaacaagctggcatgtgatcaagggtatgctaattttctcctcattcaaaccgcgagatgtgagct
aactttccgaagcttctctcgacaaaacaaaatcacgcaaaatttaccacgacttgagtggtgggtgaca
taggtccattatactgggtagctggacgaaacgacaatacaggatcgaacatattcaaggccgctgttta
caacagcacctcagacgtccctgtcaccgttcaatttgaggatgcaacgcaagagcgcgcaatttgacc
atcttgtcatccgacgatccgaacgcacgaactaccctggggggcccgaggttggaagactgagatcc
agtctgtcactgcaaatgctcatggagcatttgagttcagttctccgaacctaaagtgtgggtgttctcaa
aacggagtaa

FIG. 26A

27 / 138

SEQ ID NO:22**Protein sequence of Pf51A**

mqkmwhsilvvlgllsvqhaitinvssqsggnktsplqyglmfedinhggdgglyaelvrnrafqgstvyp
 anldgydsvngailalqnltnplspsslnvakgsnngsigfanegwwgievqpqryagsfyvqgdyq
 gdfdislqskltqevfatakvrssgkhedwvqykyelpkkaasntnntltitfdskglkdgslnfnlis
 lfpptynnrpnglridlveamaelegkflrfpggsdvegvpqpywykwnetvgdlkdrysrsaswtyees
 ngiglieymnwcdmglepilavwdghylsnevisendlqpyiddtlnqleflmgapdtpygswraslgy
 pkpwtinyveignednlyggletyiayrfqayydaitakyphtvmesltempgpaaaasdyhqystpdg
 fvsqfnyfdqmpvtntlngeiatvypnnpsnsavawgspfpfplypwwigsvaeavfligeernspkiigas
 yapmfrninnwqwsptliafdadsrtrsrswhvikllstnkitqnlpttwsggdigplywvagrndnt
 gsnifkaavynstsdvpvtvqfagcnaksanltlssddpnasnypggpevvkteiqsvtanahgafefs
 lpnlsvavlkte

FIG. 26B**SEQ ID NO:23****Nucleotide sequence for AfuXyn2, a GH11 family enzyme from *Aspergillus fumigatus***

atggttttctttctcctacctgctgctggcgtgctcggccattggagctctggctgcccccgctcgaacccg
 agaccacctcggttcaatgagactgctcttcatgagttcgctgagcgcgcggcaccaccaagctccaccgg
 ctggaacaacggctactactactccttctggactgatggcggcgcgacgtgacctacaccaatggcgcc
 ggtggctcgtaactccgtcaactggaggaacgtgggcaactttgtcggtggaaagggtggaaccctggaa
 gcgctaggtaccgagctttgtcaacgtcggatgtgcagacctgtggctgacagaagtagaaccatcaact
 acggaggcagcttcaacccagcggcaatggctacctggctgtctacggctggaccaccaacccttgat
 tgagtactacgttgttgagtcgtatggtacatacaaccccggcagcggcggtaccttcaggggcactgtc
 aacaccgacgggtggcacttacaacatctacacggccggttcgctacaatgctccctccatcgaaggcacca
 agacottcaccagtaactggtctgtgcgcacctccaagcgtacggcgggcaactgtcaccatggccaacca
 cttaacgcctggagcagaactgggcataaacctgggaactcacaactaccagattgtcgccactgaggggt
 taccagagcagcggatctgcttccatcactgtctactag

FIG. 27A**SEQ ID NO:24****Protein sequence of AfuXyn2**

mvsfsylllacsailaapvepettsfnetalhefaeragtpsstgwnngyyysfwtdgggdvtytnga
 ggsysvnrnrvgnfvvgkgwnpgsartinyggsfnpsngylavvgwttnplieyyvvesygtynpgsgg
 tfrgtvntdggtyniytavrynapsiegtktftqywsvrtskrtggvtmanhfnawsrlgmnlgthnyq
 ivategyqssgsasitvy

FIG. 27B

28 / 138

SEQ ID NO:25**Nucleotide sequence for AfuXyn5, a GH11 family enzyme from *Aspergillus fumigatus***

atgatctccatttcctcgcctcagctttggactcgcgcctatcgccggcgcatatgctcttccgagtgaca
aatccgtcagcttagcgggaacgtcagacgatcacgaccagccagacaggcacaacaatggctactacta
ttccttctggaccaacggtgccggatcagtgcaatatacaaatggtgctggtggcgaatatagtgtgacg
tgggcgaaccagaaacggtggtgactttacctgtgggaagggctggaatccagggagtgaccagtaggcaa
cgcccgagaaactatagaagaggacgcaaagaaagcactaaactctctactagtgcattaccttctctgg
cagcttcaatccttccggaaatgcttacctgtccgtgtatggatggactaccaaccccttagtcgaatac
tacatcctcgagaactatggcagttacaatcctggctcgggcacacgacaagggcaccgtcaccagcg
atggatccacctacgacatctatgagcaccaacaggtcaaccagccttcgatcgtcggcaccggccacctt
caaccaatactggtccatccgccaaaacaagcgatccagcggcacagtcaccaccgcgaatcacttcaag
gectgggctagtctggggatgaacctgggtacccataactatcagattgtttccactgagggatatgaga
gcagcgggtacctcgaccatcactgtctcgtctggtggttcttcttctggtggaagtgggtggcagctcgtc
tactacttctcaggcagctccctactggtggctccggcagtgtaagtcttcttccatatggttgtggc
tttatgtgtattctgactgtgatagtgtctgtcttgtggggccagtgcggtggaattggctggtctggt
cctacttgcctgtcttccggcacttgccaggtttcgaaactcgtactactcaccagtgctttagtagtacctc
ttgcagggttatatccaagtga

FIG. 28A**SEQ ID NO:26****Protein sequence of AfuXyn5**

MISISSLSFGLAAIAGAYALPSDKSVSLAERQTITTSQTGTNNGYIYSFWTNGAGSVQYTNGAGGEYSVT
WANQNGGDFTCGKGWNPFGSDHDITFSGSFNPSGNAYLSVYGWTTNPLVEYYILENYGSYNPGSGMTHKGT
VTSDGSTYDIYEHQQVNQPSIVGTATFNQYWSIRQNKRRSGTVTTANHFKAWASLGMNLGTHNYQIVSTE
GYESSGTSTITVSSGGSSSGSGGSSSTTSSGSSPTGGSGSCSALWGQCGGIGWSGPTCCSSGTCQVSNS
YYSQCL

FIG. 28B

29 / 138

SEQ ID NO:27**Nucleotide sequence for Fv43D, GH43 family enzyme from *Fusarium verticillioides***

atgcagctcaagtttctgtcttcagcattggttgcgtgtctttgaccggcaattgcgctgcgcaagacacta
atgatatccctcctctgatcaccgaacctctgggtctgcggatccctcggtcatgttttcgagggcaaaact
ctgggtttacccatctcagcagatcgaagccaatgtcgtcaacggcaccggaggcgctcagtaacgccatg
agagattatcacacctattccatgaagaccatctatggaaaagatcccgttatcgaccatggcgctcgctc
tgtcagtcgatgatgtcccatgggccaagcagcaaatgtgggctcctgacgcagcttacaagaacggcaa
atattatctctacttccccgcccaaggataaagatgagatcttcagaattggagttgctgtctccaacaag
cccagcggtcctttcaaggccgacaagagctggatccccggtaacttacagtatcgatcctgctagctatg
tcgacactaatggcgaggcatacctcatctggggcggtatctggggcgggccagcttcaggcctggcagga
tcacaagacctttaatgagtcgtggctcgggcgacaaagctgctcccaacggcaccaacgccttatctct
cagatcgccaagctaagcaaggacatgcacaagatcaccgagacacccccgcgatctcgtcatectggccc
ccgagacaggcaagcccccttcaagcagaggacaataagcgacgatttttcgaggggcccctgggttcacaa
gcgcggcaagctgtactacctcatgtaactctaccggcgacacgcacttcctcgctacgcgacttccaag
aacatctacggctccttatacctatcagggcaagattctcgacctggttgatgggtggactacgcattggaa
gtattggttgagtacaagggacagtggtggttgttctttgcggtatgcgcatacttctggaaaggattatct
gagacagggttaaggcgaggaagatctggtatgacaaggatggcaagattttgcttactcgtcctaagatt
tag

FIG. 29A**SEQ ID NO:28****Protein sequence of Fv43D**

mqlkflssallsltncaaqdtn dipplitdlwsadpsahvfegklwvypshdieanvvngtggaaqyam
rdyhtysmktiygkdpvidhgvalsvddvpwakqqmwapdaaykngkyylyfpakdkdeifrigvavsnk
psgpfkadkswipgtysidpasyvdtngayliwggiwggqlqawqdhtkfneswlgdkaapngtnalsp
qiaklskdmhkietetprdlvilapetgkplqaednkrrffegpwvhkrklyylmystgdthflvyatsk
niygpytyqgkildpvdgwtthgsiveykgqwwlffadahtsgkdylrqvkarkiywdkdgkillltpki

FIG. 29B

30 / 138

SEQ ID NO:29**Nucleotide sequence for Pf43B, GH43 family enzyme from *Penicillium funiculosum***

atgagtcgcagcacccttcgtacgcctctgttttcgcctcctggcggggctatcgccgaaccgtttt
tggttctcaatagcgattttcccgatccaggtctcatagagacatccagcggatactatgcattcggtac
cacoggaaacggagtcgaatgcgaggttgcttcttcaccagaactttaataacctggactttgctttccggc
acagatgcctcccgggaccatttcggtcatgggtagcttcgtctccacaaatctggcgccagatgttt
tggttaagggtatgttcttatggaataacagtttttaggagtaggtcagccaggatattgacaaaattataa
taggcgatgggtacctatgtcatgtacttttcggcatctgctgcgagtgactcgggcaaactgcgttg
gtgcgcgaactgcgacctcacoggaaaggaccttacaccccggtcgatagecgtgttgectgtccattaga
ccaggaggagctattgatgccaatggatttattgacaccgacggcactatatacgttgatatacaaaatt
gatggaaacagtcctagacgggtgatggaaccacacatcctacccccatcatgcttcaacaaatggaggcag
acggaacaaccccaacoggcagcccaatccaactcattgacogatccgacctcgacggacctttgatcga
ggctcctagtttgctcctctccaatggaatctactacctcagtttctcttccaactactacaacactaat
tactacgacacttcatacgcctatgcctcgtcgattactgggtccttggaacaaacaatctgcgccttatg
caccttggttggttactggaaccgagactagcaatgacggcgcatgtgagcgccctgggtggtgccgattt
ctccgtcgatggcaccagatgttggtccacgcaaacctcaatggacaagatatctcggcgcgacgcgc
ttatttgctgcgtcaattactgaggccagcgaatgttggttacattgcagtag

FIG. 30A**SEQ ID NO:30****Protein sequence of Pf43B**

msrsilpyasvfallqqaiaepflvlnsdfpdpslietssgyyafggtgngvnaqvasspdfntwtllsg
tdalpgpfpswvasspqiwapdvlvkadgtyvmyfsasaasdsgkhcvgaatatspegpytpvdsavacp
ldqggaidangfidtdgtiyvvykidgnsldgdgtthptpimlqqmeadgttptgspiqldrsdldgpl
ieapslllsngiyylsfssnyntnyydtayassitgpwtkqsapyapllvtgtetsndgalsapgga
dfsvdgtkmlfhanlngqdisggralfaasiteasdvvtlq

FIG. 30B

31 / 138

SEQ ID NO:31**Nucleotide sequence of Fv51A, a GH51 family enzyme from *Fusarium verticillioides***

atgggttcgcttcagttcaatcctagcggctgcggccttgcttcgtggctggttgagtcagtcacatcaagg
tcgacagcaaggcggaacgctactagcggtcaccaatatggcttccttcacgaggttggtattgacac
accactggcgatgattgggatgctaacttgagctaggatatacaacaattccgggtgatgggtggcatctac
gctgagctcatccgcaatcgtgctttccagtacagcaagaaataccctgtttctctatctggctggagac
coatcaacgatgctaagctctccctcaaccgtctcgacactcctctctccgaacgctctccccgtttccat
gaacgtgaagcctggaaagggaaggccaaggagattggtttcctcaacgaggggttaactggggaatggat
gtcaagaagcaaaagtacactggctctttctgggttaaggggcgcttacaaagggccactttacagcttctt
tgogatctaacccttacggacgatgtctttggcagcgtcaagggtcaagtccaaggccaacaagaagcagtg
ggttgagcatgagtttgtgcttactcctaacaagaatgccctaacagcaacaacacttttgctatcacc
tacgatcccaagggtgagtaacaatcaaaactgggacgtgatgtatactgacaatttgtagggcgctgatg
gagctcttgacttcaacctcattagcttgttccctcccacotacaaggggcgcgaagaacgggtcttcgagt
tgatcttgccgagggctctcgaagggtctccaccccgtaagggtttacogtctcacgtgtatcgtgaacagtc
gctgacttgtagaaaagagcctgctgcgcttccccgggtggtaacatgctcgaggggaacaccaacaagac
ctgggtgggactggaaggataccctcggacctctccgcaaccgtcctgggttcgaggggtgtctggaactac
cagcagaccatggtcttggaaatcttgagtagctccagtgggctgaggacatgaaccttgaaatcagta
ggttctataaaattcagtgacgggttatgtgcatgctaacagatttcagttgtcgggtgtctacgctggcct
ctccctcgacgggtccgtcaccccccaaggaccaactccagccctcatcgacgacgcgctcgacgagatc
gaattcatccgaggtccccgtcaacttcaaagtggggaaagaagcgcgctgagctcggccacccccaaagcctt
tcagaactctcctacgttgaagtcggaaacgaggactggctcgtggttatcccactgggtggaactctta
caaggagtaccgcttccccatgttcctcgaggctatcaagaaagctcacccogactctacccgtcatctcc
cttggtgcttctattgaccccggttggttaagaaggatgctgggttcgatattcctgctcctggaatcgggtg
actaccaccccttaccgagagcctgatgttcttgttgaggagttcaacctgtttgataacaataagtatgg
tcacatcattgggtgaggttgcttctacccacccccaaagggtggaactggctggagtggtaaccttatgct
tacccttggtggatctctgggtgttggcgaggccgtcgtctctctcggttatgagcgcaacgcgcgatcgta
ttcccggaacattctacgctcctatcctcaagaacgagaaccggttggcagtgggctatcaccatgatcca
attcgccgcgactccgccatgaccacccgctccaccagctgggtatgtctgggtcactcttcgcaggccac
cccatgaccataactctccccaccaccgcccgaacttcgacccccctctactacgtcgctggtaagaacgagg
acaagggaactcttatctggaagggtgctgctgataaacaccaccaagggtgctgacgttcccggtgtctct
gtccttcaagggtgtcaagcccggtgctcaagctgagcttactcttctgaccaacaaggagaaggatcct
tttgcttcaatgatcctcacaagggaacaatgttgttgatactaagaagactgttctcaaggccgatg
gaaagggtgcttcaacttcaagcttccctaacctgagcgtcgtgttcttgagaccctcaagaaggga
gccttactctagctag

FIG. 31A

32 / 138

SEQ ID NO:32**Protein sequence of Fv51A**

mvrfssilaaaacfvavesvnikvdskggnatsghqygflhedinnsgdggiyaelirrafqyskkypv
slsgwrpindaklslnrldtplsdalpvsnmvkgpgkgakeigflnegywgmdvkkqkytgsfwvkgayk
ghftaslrslntddvfgsvkvkskankkqwvehefvltpnknapnsntfaitydpkgadgaldfnlisl
fpptykgrkngrlrvdlaealeglhpsllrfpgggnmlegntnktwwdwkdtlgplrnrrpgfegvwnyqqth
glgileylqwaedmnleiiivgvyaglsldgsvtpkdqlqpliddaldeiefirgpvtswgkkrakelghp
kpfrlsyvevgnedwlagypgtwnsykeyrfpmfleaiikkahpdltvissgasidpvgkkdagfdipapg
igdyhpyrepdvolveefnlfdnkkyghiigevasthpnggtgwsgnlmpypwwisgvgeavalcgyerna
dripgtfyapilknenrwqwaitmiqfaadsamttrstswyvwslfaghpmthtlpttadfdplyyvagk
nedkgtliwkgaaynttkgadvpslsfkgvkggaqaeltlltnkekdpfafndphkgnnvdtkkktvlk
adgkgafnfklpnlsavvletlkkgkpyss

FIG. 31B

33 / 138

SEQ ID NO:33, Nucleotide sequence for Cg51B, a GH51 family enzyme from *Chaetomium globosum*

atggcgccccctttcgcttcgggcccctctcgctgctcgcgctcacaggagccgcagccgcggtgacctat
cggtcgcgaactctggcggtaatgatacgtctcgtacatgtatggcatcatgttcgaggacatcaatca
gagcgggtgacggcgggctgtaagttctgtcgcggttccccctgacaagcttgcatgatgcttaactaaag
tccttaggtacgcccagctgattcgcaaccgagccttccataatagctccctccaggcctggaccgcggt
gggggacagcactctcgaggctgtaacctctgcaccggttatcggtatgccctgcctcgctcggtcaaggtc
acgagtggaaagggcaaggcgggcttgaagaatgcgggctaactggggaatggacgtccagaagaccgaca
agtatagcggcagcttctactcgtacggcgccctacgacggaaagtttaccctctctctggtgtcggacat
cacaaatgagaccctggccaccaccaagatcaagtcagggtcggtggagcatgcctggaccgagcacaag
ttcgagcttctcccgaccaagagcgcggcggaacagcaacaacagcttcgtgctggagttccgccccctgcc
accagacgggagctccagttcaacctcatcagcttgttcccgccgacgtataagaacaggcccaacggcat
ggcgccgagagctcatggagaagctcgacacctcaagcccagtttcccttcggattccaggaggcaacaac
ctgtaagtgttccggcgaaactagcagtagttgcctgagagacactaatctcagcgaacaacagcgagg
gcaactatgctggcaactactggaactggtcaagcacacttggccccgctgaccgaccggccccggtcgtga
cggcgtgtggacgtacgccaaacacggacggcatcgggctggctgagtacatgcactgggcccaggacctc
gacgtggaggttgtgctggcggtcgccgcaggcctgtacctgaacggcgatgtggtcccgaggaggagc
tgcacgtcttcgtggaggatgcgctgaacgagctcgagttccctcatgggcgacgtctcgaccccccttgggg
cgcgcgccgcgctaagctcggctacccccaaagcgcgtggaacatcaagttcgtcgagggtcggaacgaggac
aacctgtggggcggcctcgactcgtacaagagctaccggctgaagactttctacgacgccatcaaggcga
agtacccccgacatctccatcttttcgtcgaccgacgagtttgtgtacaaggagtcggggccaggactacca
caagtacacccggccggactactcgtgtcccagttcgacctgtttgacaactgggcccgacggccacccc
atcatcatcgagagtgagtgaaacggcgacccccacctccccctaacgcgggatcgcgagctgatagatc
accccaggtatgcgaccatccagaacaacacgggcaagctcgaggacacggactgggacgcgcgcccaagaa
caagtgggtccaaactggatcggtcgcgcgaggcgcgtcttcacctcctcgagccgagcgcaacggcgac
cgggtctggggcaccacctttgcgcgcgatacctccagaacctcaacagctaccaatgggctgtaagtacat
acatacataccgcaccccccaaccccccaaaagcgcacctccacccacccacccaaacacaccac
aactacctagctaaccgcgcacacaaaacaaacagccccgacctaatctccttcacgcgcaacccggccgac
accacgcccagcgtctcgtacccgatcatccagctgctcgcctcgcacccgcatcacgcacaccctccccg
tcagcagcgccgacgccttcggccccggcctactgggtggccgggtcgcggcgcgcgacgacggctcgtacat
cetcaaggcggccggtgtacaacagcaacgggggggtgcggatgtaccggtgaggggtgcagtttgaggcgggg
ggtggtggtggtggtggtggtggtggtggtggtggtggtgatgggaaggggaagggtaaaggggaagg
gaggggaggggtggtgaggggtgtgaagaaggggtgaccgcgcgcagttgacctgtttgacggcgccggagg
gccttgggcgcataataacgcgcgagaataagggggcggtcaagacgacagtgacgacggttgaaggccggg
aggggtgggggtgtttgagtttagtctgcgggatttgcgggtggcggtgtttggtggtggaggggggagaagt
ga

FIG. 32A

34 / 138

SEQ ID NO:34, Protein sequence of Cg51B

maplsrlalsllaltgaaaavtlsvansggndtspymygimfedinqsgdgglyaelirnrafhnsslqa
wtavgdstlevvtsaplsdalprsvkvtsqgkgkaglkknagywgmdivqktdkysgsfysygaydgkftlsl
vsditnetlattkiksrsvehawtehkfellptksaansnnsfvlefrpchtelqfnlislfpptyknr
pngmrrelmekladlkpsflripggnnlegnyagnywnwsstlgpltdrpgrdgvwtyantdgiglvveym
hwaedldvevvlavaaglylmgdvvpееelhvfvedalneleflmgdvstpwgarraklgypkpwnikfv
evgnednlwggldsyksyrktfydaikakypdisifsstdefvykesggdyhkytrpdysvsqfdlfdn
wadghpiiigeyatiqnntgkledtdwdapknkwsnwigsvaeavfilgaerngdrvwgttfapilqnl
syqwapdlisftanpadtppsypiiqlashrithtlpvssadafgpaywvagraddgsyilkaavy
nstggadvprvqfeaggggggggggggggdkgkgkgkggeggegkvkkgdraqltvltapegpwahnt
penkgavkttvttkagrggvfefsldlsvavlvvegek

FIG. 32B**SEQ ID NO:35, Nucleotide sequence for Fv43C, a GH43 family enzyme from *Fusarium verticillioides***

atgcgctcttctatcgctttcccagccatctcctcgctggccttccctaaccctcaaagaggcttcateccctcg
ccctcagcaaacgggatagccctgtcctccccggcctctggggcggaacccaacatcgccatcgctcgacaa
gacatactacatcttccctaccaccgaacgggtttcgaaggctggggcggaacacgtcttctactgggtggaaa
tcaaaagatctcgtatcatggacaaagagcgacaagccattccttactctcaatggtacgaatggcaacg
ttccctgggctacaggtaatgcctgggctcctgctttcgtctgctcgcgaggaggaagtattacttctacca
tagtgggaataatccctctgtgagtgatgggcataagagtattggtgcggcggtggctgatcctctgag
gggccgtggaaggcacaggataagccgatgatcaagggaacttctgatgaggagattgtcagcaaccagg
ctatcgatcccgctgcctttgaagaccctgagactggaaagtggatatctactggggaaacgggtgtccc
cattgtcgcagagctcaacgacgacatggtctctctcaaagcaggctggcacaaaatcacagggtcttcag
aattttccgcgaggggtcttttcgtcaactatcgcgatggaacatatcatctgacatactctatcgacgata
cgggctcagagaactatcgcggttggttacgctacggcgataaccccattggaccttgacatatcggtg
tgttcttctggagaaggacgaatcgaagggcattcttgctacgggacataaactccatcatcaacattcct
ggaacgggatgagtggatatcgcgatcatcgcttccatattcccgatggaaatgggtataatagggaga
ctacgattgatagggtacccatcgacaaggatacggggttggttgaaagggttacgcgcgactttgcagag
tgttgatcctaggcctttgtag

FIG. 33A**SEQ ID NO:36, Protein sequence for Fv43C**

mrllsfpslllvafltlkeasslalskrdspvlpglwadpniaivdktyyifpttdgfegwgggnv
fywwkskdlvswtksdkpfltlngtngnvpwatgnawapafaarggkyyfyhsgnnpvsvdghks
igaavadhpegpwkaqdkpmikgtsdeeivsnqaidpaafedpetgkwyiywngnvpivaeindd
mvslkagwhkitglqnfreglfvnyrdgtyhltysiddtgsenyrvgyatadnpigpwttyrgvll
ekdeskgilatghnsiinipgtdewyiahyrfhipdgngynrettidrvpidkdtglfgkvtptl
qsvdprpl

FIG. 33B

35 / 138

SEQ ID NO:37, Nucleotide sequence for Fv30A, a GH30 family enzyme from *Fusarium verticillioides*

atgctcttctcgcgtctcttcttacccttgcctttcaagccagcctggcgctcggcgatacatccgtta
ctgtcgacaccagccagaaactccaggtcatcgatggctttgggtgtctcagaagcctacggccacgcaa
acaattccaaaacctcgggtcctggaccacagaaagagggcctcgatcttctcttcaacactacaacggc
gcaggcttatccatcatccgaaacaagatcggtgcgacgcctccaactccatcaccagcaccaacaccg
acaaccagataagcaggctgttttaccattttgacggcgatgatgatgggtcagggtatgggttagcaaa
ggccatgagctatgggtgtagatactatctacgctaattgcttgggtctgcgcctgtatacatgaagtcagcc
cagagtatgggcccgtctctgcggtacacctgggtgtgtcgtgctcctctggagattggagacatcgttacg
ttgagatgatagctgagtagctctctactacaagcaggctggcatcccagtgctgcacggttgattctt
caatgaggggtgacggctcggactttatgctctcaactgcgaacaggctgcagatgtcattcctcttcta
cacagcgttttgagtcacaagggccttggcgatatcaagatgacgtgctgtgataacatcggttggaagt
cacagatggactataccgccaagctggctgagcttgagggtggagaagtatctatctgtcatcacatcca
cgagtactccagcagccccaaccagcctatgaacactacattgccaacctggatgtccgagggagctgcc
aatgaccaggcatttgccacagcgtggtagctcaacggcggttccaacgaagggttccatgggcagtc
agatcgcacaaggcctcgtcaatgccgacctctcagcgtatatctactgggagggcggttgagaccaaaa
caaggggtctctatctcagctcatcgacacggacggtagccaagtttaccatatcctcgattctctgggccc
attgctcactggctcgcgccatattcgccttgggtgcgcatagactttcgaacttcagggtgttggtgcaagata
cgattgttggtgctgttgagaacggttgatggcagtgctcgtcatgggtgctcaccaactctggcactgctgc
tcagactgtggacctgggtgtttcgggaagtagcttctcaacagctcaggctttcacttcggatgctgag
gcgcagatgggtcgataccaaggtgactctgtccgacggctcgtgtcaaggttacgggtcccggtgcacggtg
tcgtcactgtgaagctcacaacagcaaaaagctccaaaccgggtctcaactgctgtttctgcgcaatctgc
ccccactccaactagtggttaagcacaccttgactcaccagaagacttcttcaacaacactctcgacccgcc
aaggccccaacctccactcagactacctctgtagttgagtcagccaaggcggtgaaataacctgtccccc
ctgtagcatccaagggtatcctcgaagagtgtcccaagaagggtaccaagaagaccactacgaagaaggg
ctcccaccaatcgcaagaaggcgcatagtgctactcatcgctcgatgccgccatggaagttaccgtcgtggc
cactgcaccaactaa

FIG. 34A**SEQ ID NO:38, Protein sequence of Fv30A**

mlfslvlptlafgaslaldtsvtvdtssqklqvidgfgvseayghakqfqnlpgppqkeglldllfntttg
aglsiiirnkigcdasnsitstntdnpdkqavyhfdgdddggsaqsmgrlctpgvscssgdwrhryvemi
aeylsyykqagipvshvgflnegdgsdfmlstaeqaadvipllhsalqskglgdikmtccdnigwksqmd
ytaklaelevekylsvitsheyssspnqpmnttlptwmsegaandqafatawyvnggsnegftwavkiaq
givnadlsayiywegvetnnkgslshvidtdgktftissilwaiahwsrhirpgahrlstsgvvqdtivg
afenvdgsvvmvltnsqtaaqtvdlgvsgssfstagaftsdaeaqmvdtkvtlsdgrvkvtvpvhgvvtv
klttaksskpvtavsaqsaptptsvkhltlhqktssttlstakaptstqttsvvesakavkypvppvas
kgssksapkkgtkktttkkqshqshkahsathrrcrhgsyrrghctn

FIG. 34B

36 / 138

SEQ ID NO:39, Nucleotide sequence for Fv43F, a GH43 family enzyme from *Fusarium verticillioides*

atgtggaaactcctcgtcagcgggtcttggtcgcgcgtcgcgtccctcagcggcgtgaacgctgcttataccta
accctgggtcccggtcacgcggcgatactcgtgttcacgaccctacggttgtcaagactcccagcgggtggata
cttgctggctcatactggcgataacggttccgctcaagacttcttctgatcgaactgcttggaaggatgca
ggtgctgttttccccaacgggtgcgccttggactacgcagtagacaccaagggcgacaagaacctctgggccc
ctgatatctcctaccacaacgggccagtagctatctgtactaactccgcctcttcccttcgggtcagcgtacctc
tgccatttttctcgtcaccagcaagaccgggtgcacccggctcgtggaccaaccaaggcgtcgtcgtcgag
tccaacaacaacaacgactacaatgccattgacggaaatctctttgtcgactctgatggaaaatgggtggc
tctccttcgggtcttttctgggtccggcatcaagctcatccaactcgaccccaagaccgggaagcgcaccgg
ctcaagcatgtactccctcgccaaaacgcgacgcctccgtcgaaggcgcgcgtcgaggctccgttcacaccc
aaacgcggaagcacctactacctctgggtgctcgttcgacaagtgttgccagggcgcgtgctagcacgtacc
gtgtcatgggttgacgggtcgagcagcattactgggtccttatgttgacaaggctggtaagcagatgatgtc
tggtggaggaacggagattatggctagtcacggatctattcatggaccggggacataatgctgttttccact
gataacgatgcggacgttcttctgtctatcattactacgataacgctggcacagcgcgtgttgggcatcaact
tgctcagatatgacaatggctggcctgttgcttattag

FIG. 35A**SEQ ID NO:40, Protein sequence for Fv43F**

mwkllvsglvavaslsqvnaaypnpgpvtgdtrvhdpvvtkspggyllahtgdnvslktssdrtawkda
gavfpngapwttqytkgdknlwapdisyhngqyylyysassfgqrtsaiflatsktgasgswtnqgvvve
snnndynaiddgnlfvdsdgkwlsfsgsfwsgiklikldpktgkrtgssmyslakrdasvegaveapfit
krgstyyllwvsfdkccqgaastyrvmvgrsssitgpyvdkagkqmmssgggteimashgsihgpghnavft
dndadvlvvyhydnagtallginllrydngwpvay

FIG. 35B

37 / 138

SEQ ID NO:41**Nucleotide sequence for Xyn3, a GH10 xylanase from *Trichoderma reesei***

atgaaagcaaacgtcatcttgtgctctcctggcccccctggctgcgcgtctctccccaccgaaaccatccacc
tcgaccccgagctcgccgctctccgcgccaacctcaccgagcgaacagccgacctctgggaccgccaagc
ctctcaaagcatcgaccagctcatcaagagaaaaggcaagctctactttggcaccgccaaccgaccgoggc
ctctccaacgggaaaagaacgcggccatcatccaggcagacctcgggccaggtgacgcggagaaacagca
tgaagtggcagtcgctcgagaacaaccaaggccagctgaactggggagacgcgcgactatctcgtaactt
tgcccagcaaacggcaagtcgatacgcggccacactctgatctggcactcgcagctgcctgcgtgggtg
aacaatatcaacaacgcggatactctgcggcaagtcacccgacccatgtctctactgtggttgggcggt
acaagggcaagattcgtgcttgggtgagttttgaacaccacatgccccctttcttagtccgctcctcctc
ctcttggaacttctcacagttatagccgtatacaacattcgacaggaaatttaggatgacaactactgac
tgacttgtgtgtgtgatggcgataggacgtggtcaatgaaatcttcaacgaggatggaaacgctgcgctct
tcagctcttttccaggctcctcggcgaggagtttgtctcgattgcctttcgtgctgctcgagatgctgacc
cttctgcccgtctttacatcaacgactacaatctcgaccgcgccaaactatggcaagggtcaacgggttgaa
gaactacgtctccaagtggatctctcaaggagttccattgacgggtattgggtgagccacgacccctaaat
gtccccattagagtcctctttctagagccaaggcttgaagccattcagggaactgacacgagagccttctc
tacaggaagccagtcctcatctcagcggcggcggaggctctggtacgctgggtgcgctccagcagctggca
acgggtacccgtcaccgagctggccattaccgagctggacattcagggggcaccgacgacggattacacce
aagttgttcaagcatgcctgagcgtctccaagtgcgtcgccatcaccgtgtggggcatcagtgacaagggt
aagttgcttcccctgtctgtgcttatcaactgtaagcagcaacaactgatgctgtctgtctttacctagg
actcgtggcgtgccagcaccaacctcttctgtttgacgcaaaccttcaaccccaagccggcatataacag
cattggtggcatcttacaatag

FIG. 36A**SEQ ID NO:42****Protein sequence for Xyn3**

mkanvilcllaplvaaip~~teti~~hldpelaalranlt~~ertad~~lwdrqasqsidqlikrkgklyfgtatdrq
llqreknaaiiqadlgqvtpensmkwqslennqgqlnwgdadylvnfaqqngksirghtliwhsqlpawv
nninnadtlrqvirthvstvvgrykgkirawdvvneifnedgtlrssvfsrllgeefvsiafraardap
sarlyindynldranygkvnglktyvskwisqgvpidgigsqshlsgggsgtlgallqqlatvpvtelai
teldiqgapttdytqvvaqlsvskcvgitvwigsdkswrastnpllfdanfnpkpaynsivgilq

FIG. 36B

38 / 138

SEQ ID NO:43**Protein sequence of Xyn2, a GH11 family xylanase from *Trichoderma reesei***

mvsftsllaasppsrascrpaaevesvavekrqtiqpgtgynngyfysywndghggvtytngpgggqfsvn
 wnsngnfvggkgwqpgtknkvinfsgsynpngnsylsvygwsrnplieyyivenfgtynpstgatklegev
 tsdgsvydiyrtqrvnqpsiigtatfyqywsrrnrhrrssgsvntanhfnawaqqgltlgtmdyqivaveg
 yfssgsasitvs

FIG. 37A**SEQ ID NO:44****Protein sequence of Pa3C, a GH3 family enzyme from *P. anserina***

MAYRSLVLGAFASTSLAASVVTPRDPVPPGFVAAPYYPAPHGGWVASWEEAYSKAELVSQMTLAEKTNI
 TSGIGIFMGNTGSAERLGFPRMCLQDSALGVSSADNVTAFFAGITTGATFDKKLIYARGVAIGEHRGKG
 TNVYLGPSVGPLGRKPLGGRNWEFGSDPVLQAKAAALTIKGVQEQGIIATIKHLIGNEQEMYRMYNPFQ
 PGYSANIDDRTLHELYLWPFSAESVHAGVGSAMTAYNAVNGSACSQHSYLINGILKDELGFQGFVMSDWLS
 HISGVDSALAGLDMNMPGDTNIPFGFSNWHYELSRSVLNGSVPLDRLNDMVTRIVATWYKFGQDRDHPR
 PNFSSNTRDRDGLLYPAALFSPKGQVNWVNVQADHYLIAREVAQDAITLLKNNGSFLPLTTSQSLHVFG
 TAAQVNPDPGNACMNACNKGTLGMGWGSGVADYPYLDLDPISAIRKRVDPVKFFNTDGFPWFHPTSPDD
 VAIVFITS DAGENSFTVEGNGDRNSAKLAAWHNGDELVRKTAEKYNNVIVVAQTVGPLDLESWIDNPRV
 KGVLFQHLPGQEAGESLANILFGDVSPSGHLPYSITKRANDEPDSEANLRGFAFGQVQDITYSEGLYIDYR
 WLNKEKIRPREAFGHGLSYTNFSFDATIESVTPLSLVPPARAPKGSTPVYSTEIPPASEAYWPEGFNRIW
 RYLYSWLNKNDADNAYAVGIAGVKKYNYPAGYSTAQKPGPAAGGGEGGNPALWDIAFRVPVTVKNTGDTF
 SGRASVQAYVQYPEGIPYDTPVVQLRDFEKTRVLAPGEEETVTVELTRKDLVWDTELQNWVVPVGVGKR
 YTVWIGEASDRLFTACYTDTGVCEGGRVPPV

FIG. 37B**SEQ ID NO:45****Protein sequence of Bxl1, a GH3 family β -xylosidase from *Trichoderma reesei***

mvnnaallaalsallptalaqnnqtyanysagggpdlypetlatltlsfpdcehgplknnlvcdssagvy
 eraqalisiftleeliilntqnsqpgvprlglpnyqvwnéalhgldranfatkggqfewatsfpmptlta
 alnrtlihqiaadiistqarafsnsgrygldvyapnvngfrsplwgrggetpgedafflssaytyeyitgi
 qggvdpehlkvaatvkhfagydlennwnnqsrldgldaiitqqdlseyytpqflaaaryaksrslmcaynsv
 ngvpsscansfflqtllreswgfpewgyvssdcdavynvfnphdyasnqssaaasslragtdidcgqtypw
 hlnefsvagevsrgeiersvtrliyanlvrlgyfdkknqyrslgwkdvvktdawnisyeaavegivilknd
 gtllplskkvrsialigpwanattqmqqnyygpapylispleaakkagyhvnfelgteiagnsttgfakai
 aaakksdaiiylggidntieqegadrtdiawpgnqldlikqlsevgkplvvlqmaggqvdssslksnkkv
 nslvwggypggsggvalfdilsgkrapagrlvttygpaeyvhqfpqndmnlrpdgksnpgqtiwytgkp
 vyefgsglfyttfketlashpkslkfntssilsaphpgytyseqipvftfeaniknsgktespytamlfv
 rtasnagpapypnkwlvgfdrladikpghsskl sipipvsalarvdshgnrivypgkyelalntdesvkle
 felvgeevtienwpleeqqikdatpda

FIG. 38

atgatccacctcaagccagccctcgcgcgcttgttggcgctgtgcagcgaatgtgtggctattgatttgt
ttgtcaagtcttcgggggggaataagacgactgatatcatgtatgggtcttatgcacgaggatatcaacaa
ctccggcgacggcgccatctacgccgagctaatactccaacccgcgcgttccaaggagggtgagaagttcccc
tccaacctcgacaaactggagccccgcgtgggtggcgctacccttacccttcagaagcttgccaagccccctt
cctctgcgttgcccttactccgtcaatgttgccaaccccaaggaggggcaagggcaagggcaaggacaccaa
ggggaagaagggttggcttggccaatgctgggttttggggataggatgtcaagaggcagaagtacactgggt
agcttccacgttactgggtgagtacaagggtgactttgaggttagcttgcgcagcgcgattaccggggaga
cctttggcaagaagggtggtgaagggtgggagtaagaaggggaagtggaccgagaaggagtttgagttgggt
gcctttcaaggatgcgcccacagcaacaacacctttgttgtgcagtgggatgcgcgagggcgcaaaggac
ggatctttggatctcaacttgatcagcttgttccctccgacattcaagggaaggagaatgggctgagaa
ttgatcttgcgcagacgatgggttgagctcaagccgaccttcttgcgcttccccgggtggcaacatgctcga
gggtaacaccttggacacttgggtggaagtggtaacgagaccattggccctctgaaggatcgccccgggcatg
gctgggtgtctgggagtaaccagcaaaccttggcttgggtctggtcgagtacatggagtgggccgatgaca
tgaacttggagcccattgtcgggtgtcttcgctgggtcttgcctctgatggctcgcttcggttccgaatccga
gatgggatgggtcatccaacaggctctcgacgaaatcgagttcctcactggcgatgctaagaccacaaaa
tggggtgccgtccgcgcgaagcttggtcaccccaagccttgggaagggtcaagtgggttgagatcggtaacg
aggattggcttgcgggacgcccgtgctggcttcgagtcgtacatcaactaccgcttccccatgatgatgaa
ggccttcaacgaaaagtaccccgacatcaagatcatcgccctcgccctccatcttcgacaacatgacaatc
cccgcgggtgctgcccgtgatcaccacccgtacctgactcccgatgagttcggtgagcgattcgccaagt
tcgataaacttgagcaaggataacgtgacgctcatcggcgaggctgcgtcgacgcacatcctaacgggtgggtat
cgcttgggaggggagatctcatgccccttgcccttgggtggggcggcagtggttgcgtgaggctatcttcttgatc
agcaactgagagaaaacgggtgacaagatcatcggtgctacttaacgcgcctgggtcttcgcagcttggaccgct
ggcaatggagcatgacctgggtgcagcatgcgcgcgacccggccctcaccactcgctcgaccagttggta
tgtctggagaatcctcgcccaccacatcatcgtgagacgctcccggtcgatgccccggcgccggaagccc
aactttgacctctgttctacgttgccggaaaagagcgagagtggcaccgggtatcttcaaggctgccgtct
acaactcgactgaatcgatccccgggtgtcgttgaagtttgatgggtctcaacgagggagcgggttgccaactt
gacgggtgcttactggggccggaggatccgtatggatacaacgaccccttcaactgggtatcaatgttgtcaag
gagaagaccaccttcatcaaggccggaaagggcggaagttcaccttcaacctgcccgggcttgagtggtg
ctgtgttgagacgqccgacgcggtcaagggttggcaagggaaggqcaaggqcaaggqcaaggqtaactga

FIG. 39A

40 / 138

SEQ ID NO:47**Codon optimized cDNA for Pa51A, a GH51 family enzyme from *Podospira anserina***

atgatccacctcaagcccgccctcgccgcccctcctcgccctcagcaccacaatgcgctcgccatcgacctct
tcgtcaagagcagcggcggaacaagaccaccgacatcatgtacggcctcatgcacgaggacatcaacaa
cagcggcgacggcgccatctacgcgagctgatcagcaaccgcgccttcagggcagcgagaagttcccc
agcaacctcgacaactgggtccccgctcggcggcgccaccctcaccctccagaagctcgccaagccccgt
cctctgcccctcccctactccgtcaacgtcgccaaccccccaaggagggttaagggttaagggtcaaggacaccaa
gggcaagaaggctcgccctcgccaacgcgcgctttttggggcatggacgtcaagcgccagaaatacaccggc
agcttcacgtcaccggcgagtacaaggcgacttcgaggtcagcctccgcagcgccattaccggcgaga
ccttcggcaagaaggctcgtcaaggcgcgagcaagaagggtgaagcgagagaaggagttcgagctgggt
ccccttcaaggacgcccccaacagcaacaacaccttcgtcgtccagtgggacgcccagggcgccaaggac
ggcagcctcgacctcaacctcatcagcctcttcccgcacccttcaaggcgccaagaacggcctccgca
tcgacctcgcccagaccatggctcgagctgaagcccaccttccctccgctttcccgggcggaacatgctcga
gggcaacaccctcgacacctgggtggaagtggtaacgagaccatcgccccctgaaggacgcctggcatg
gcccggctctgggagtaaccagcagacgctgggcctcgccctggctcgagtacatggagtgggcccagcaca
tgaacctcgagcccacatcgctcggcgtctttgctggcctggcctggatggcagctttgtccccgagagcga
gatgggctgggtcatccagcaggtctcgtatgagatcgagttcctcaccggcgacgccaagaccaccaag
tgggggcgccgtccgcgccaagctcgggccaccctaagccctggaagggtcaaatgggtcgagatcggaacg
aggactggctcgccggccgacctgcgcgcttcgagagctacatcaactaccgcttcccctatgatgatgaa
ggccttcaacgagaaataccccgacatcaagatcattgccagccccctccatcttcgacaacatgaccatt
ccagccggtgctgccggtgaccaccaccctacctcacccccgaacgaatttgctcgagcgcttcgccaagt
tcgacaacctcagcaaggacaacgtcaccctcattggcgaggcgccagcaccaccccaacggcgggcat
tgccctgggagggcgacctcatgcccctgcctgggtggggcggcagcgtcgccgaggccatcttctcctc
agcaccgagcgcaacggcgacaagatcatcggcgccacctacgcccctggcctccgatctctcgaccgct
ggcagtgagcatgacctgggtccagcagccgcgcgacctgcctcaccacccgcagcaccagctggta
cgtctggcgcatcctcgcccaccacatcattcgcgagacctccccgtcgacgcccccgccggcaagccc
aacttcgacccccctcttctacgtcgctggcaagtcggagagcgggcaccggcatcttcaaggccgcgctct
acaacagcacogagagcatccccgtcagcctcaagttcgacggcctcaacgagggcgccgtcgccaacct
caccgtcctcaccggccccgaggaccctacggctacaacgaccccttcaccggcatcaacgtcgtcaag
gaaaagaccaccttcacgaaggccggcaaggggcggaaggttcacctttaccctccccggcctctctgtcg
ccgtcctcgagaccgcgcgacgccgtgaagggtggcaagggaagggaagggaagggttaagggttaacta
a

FIG. 39B

41 / 138

SEQ ID NO:48**Nucleotide sequence for Gz43A, a GH43 family enzyme from *Gibberella zeae***

atgtatcggaaagttggccgtcatctcggccttctttggccacaagctcgtgctaccaacgacgactgtcctc
tcatcactagtagatggactgcggtaccttcgggtcatgtctttaacgacaccttgtgggtctacccgtc
tcatgacatcgatgctggatttgagaatgatcctgatggaggccagtaacgccatgagagattaccatgtc
tactctatcgacaagatctacgggttccttgccggtcgatcacggtacggccctgtcagtgaggatgtcc
cctgggcctctcgacagatgtgggctcctgacgctgccacaagaacggcaaatactacctatacttccc
tgccaaagacaaggatgatatacttcagaatcggcggtgtgtgtctaccaacccccggcggaccattcgtc
cccgacaagagttggatccctcacactttcagcatcgaccccgccagtttcgtcgatgatgatgacagag
cctacttggcatggggtgggtatcatgggtggccagcttcaacgatggcaggataagaacaagtacaacga
atctggcactgagccaggaaacggcacccgtgccttgagccctcagattgccaaagctgagcaaggacatg
cacactctggcagagaagcctcgcgacatgctcattcttgaccccaagactggcaagccgctcctttctg
aggatgaagaccgacgcttcttcgaaggaccctggattcacaagcgcaacaagatttactacctaccta
ctctactggcacaacccactatcttgtctatgogacttcaaagacccctatgggtccttacacctaccag
ggcagaattctggagccagttgatggctggactactcactctagtatcgtaagtaaccagggtcagtggt
ggctattttatcacgatgccaaagacatctggcaaggactatcttcgccaggtaaaggctaagaagatttg
gtacgatagcaaaggaaagatcttgacaaagaagccttga

FIG. 39C**SEQ ID NO:49****Nucleotide sequence for Fo43A, a GH43 family enzyme from *Fusarium oxysporum***

atgtatcggaaagttggccgtcatctcggccttctttggccacaagctcgtgctcaagacactaatgacattc
ctcccctgatcaccgacctctgggtccgcagatccctcgggtcatgttttcgaaggcaagctctgggttta
cccattctcacgacatcgaagccaatgttggtcaacggcacaggaggcgtcaatacggcatgagggattac
catacctactccatgaagagcatctatggtaaagatcccgttgctcgaccacggcgctcgctctctcagtcg
atgacgttccctgggccaagcagcaaatgtgggctcctgacgcagctcataagaacggcaaatattatct
gtacttccccgccaaaggacaaggatgagatcttcagaattggagttgctgtctccaacaagcccagcgggt
cctttcaaggccgacaagagctggatccctggcaagtaacagtatcgatcctgctagctacgtcgacactg
ataacgaggccctacctcatctggggcggtatctggggcgccagctccaagcctggcaggataaaaagaa
ctttaacgagtcgtggattggagacaaggctgctcctaacggcaccaatgccctatctcctcagatcgcc
aagctaagcaaggacatgcacaagatcacccgaacacccccgcgatctcgtcattctcgcccccgagacag
gcaagcctcttcaggctgaggacaacaagcgacgattcttcgagggcccttggatccacaagcgcgga
gctttactacctcatgtactccacgggtgatacccacttcccttgtctacgctacttccaagaacatctac
ggctccttatacctaccggggcaagattcttgatcctgttgatgggtggactactcatggaagtattgttg
agtataagggacagtggtggcttttctttgctgatgcgcatacgtctggttaaggattaccttcgacaggt
gaaggcgaggaagatctggatgacaagaacggcaagatcttgcttcaccgtccttag

FIG. 39D

42 / 138

SEQ ID NO:50**Nucleotide sequence for Pf51A, a GH51 family enzyme from *Penicillium funiculosum***

atgtaccggaagctcgccgtgatcagccgcttctctggcgactgctcgccgcatcaccatcaacgtcagcc
agagcggcggcaacaagaccagcccgctccagtaaggcctcatgttcgaggacatcaaccacggcggcga
cggcggcctctacgcccagagctgggtccggaaccgggcttccagggcagcaccgtctacccggccaacctc
gaaggctacgactcgggtgaacggcgcgattctcgcgctccagaacctcaccaccccgctcagcccagagca
tgccctcgtcgctgaacgtcgccaagggtcgaacaacggcagcatcggcttcgccaacgaggggtggtg
gggcatcgagggtcaagccgcagcggtaacgcccgcagcttctacgtccagggcgactaccagggcgacttc
gacatcagcctccagagcaagctcaccacaggaggtcttcgcgacggcggaaggtccgggtcgagcggcaagc
acgaggactgggtccagtacaagtaacgagctgggtcccggaagaaggccgccaacaccaacaacacctc
caccatcaccttcgacagcaagggcctcaaggacggcagcctcaacttcaacctcatcagcctcttcccg
ccgacctacaacaaccggccgaacggcctccggatcgacctcgtcgaggccatggcggagctggagggca
agttcctccgcttccccggcggtcggacgtggagggtccagggccccgtactgggtacaagtggaaacga
gaccgtcggcgacctcaaggaccgctactcgcgcccagcgcctggacctacgaggagagcaacggcatc
ggcctcatcgagtacatgaactggtgcgacgacatgggcctcgagccgatactcgcgctctgggacggcc
actacctcagcaacgagggtcatcagcgagaacgacctccagccgtacatcgacgacacctcaaccagct
cgagttcctcatgggcgccccggacaactccctacgggtcttggagggttagcctcggtacccgaagccg
tggaccatcaactacgtcgagatcggcaacgaggacaacctctacggcggcctcgagacctacatcgct
accggttccaggcctaactacgacgccatcaccgccaaagtaaccgcacatgaccgtcatggagagcctcac
cgagatgcccggcccccgctgccgcggtcggactaccaccagtaactcgacgcccagcggcttcgtcagc
cagttcaactacttcgaccagatgccggtcaccaaccgcacgctgaacggcgagatcgccaccgtctacc
ccaacaacccgagcaactcgggtggcggtggggcagcccggttcccgctctaccggtgggtggatcgggtccgt
ggctgaggccgctcttctcatcggcgaggagcggaaacagcccgaagatcatcggcgccagctacgcccc
atgttcgcaacattaacaactggcagtgaggcccgaccctgatcgcttcgacgcgcagcagcggga
cgtcgcgctctacttctggcacgtcatcaagctcctcagcaccaacaagatcaccagaacctgccac
gaogtgggtctgggggggacatcgggccgctctactgggtcgccggccggaacgacaacccggcagcaac
atcttcaaggccgctctacaacagcaccagcgacgtcccggtcaccgtccagttcgccgggtgcaacg
ccaagagcgccaacctcaccatcctctcgtcggacgaccccaacgccagcaactaccggggcgccccga
ggtcgtcaagaccgagatccagagcgtcacccgcaacgcccacggcgcccttcgagttcagcctcccgaa
ctgtcgggtggctgtgctgaagacggagtag

FIG. 39E

43 / 138

SEQ ID NO:51

Nucleic acid sequence of Eg4, a GH61 family endoglucanase from *Trichoderma reesei*

atgatccagaagctttccaaccttcttctcaccgcactagcgggtggcaaccgggtgttgttggaacaggac
 acatcaacaacattgtcgtcaacggagtgactaccagggatgatgcctacatcgttcccatatgaatc
 tgacccgcccatagtgggtgggctggacgggtgcgatcttgacaacgggttcgtctcaccgcagcgcata
 cagagcccggaacatcatctgccacaagaatgccaccaacgccaaaggacacgcgtccgtcaaggccggag
 acaactattccccctccagtgggtgccagttccttgggcgcacccaggccccatcgtcgactacctggccaa
 ctgcaacgggcgactgcgagaccgtggacaagacgtcccttgagttcttcaagattgacggcgctcgggtctc
 atcagcggcgggagatccgggcaactgggacctcggacgtgttgattgccaacaacaacacctgggttgta
 agatccccgaggatctcgcgccgggcaactacgtgcttcgccacgagatcatcgcccttgacacagcgcgg
 gcaggcggacggcgctcagaactaccctcagtgcttcaacctcgccgtcccaggctccggatctctgcag
 ccgagcggcgctcaagggaaccgcgctctaccactccgatgaacccgggtgtcctcatcaacatctacacca
 gccctcttgctacaccattcctggaccttccgtgggtatcaggcctccccacgagtgtcgcccagggcag
 ctccgcgcgcgacggccaactgccagcgcacactgttcctggcggttagcggaccgggaaacccgaccagtaag
 actacgacgacggcgaggacgacacaggcctcctctagcagggccagctctactcctcctgctactacgt
 cggcacctgggtggaggcccaaccagactttgtacggccagtggtggcagcggctacagtggtcctac
 tegatgcgcgcgcgcggccacttgcctctaccttgaaaccatactacgccacgtgccttaactag

FIG. 40A**SEQ ID NO:52**

Protein sequence of Eg4, an endoglucanase from *Trichoderma reesei*

MIQKLSNLLVTALAVATGVVGHghhindivingvwyqaydpttffpyesnppivvgwtaadldngfvspday
 qnpdiichknaatnakghasvkagdtlffqwpvpwphpgpivdyancngdcetvdkttleffkidgvgl
 lsggdpgtwasdvlisnnntwvkipdnlapgnyvlrheiiialhsaggangagqnyppqcfniavsgsgslq
 psgvlgtldlyhatdpgvliniytsplnyliipgptvvsglptsvaaggssaataatasatvpgggsgptsrtt
 ttarttqassrpsstppattsapaggptqtlyggcggsgysgptrcappatcstlnpyyaqcln

FIG. 40B

44 / 138

SEQ ID NO:53**Nucleic acid sequence of Pa3D, a GH3 family β -glucosidase from *Podospira anserina***

atggctcttcaaaccttcttctctgctggcgccagccatgctggccaacgcagagacaacaggcgaaaagg
tctctcggcaagcaccgtcttggcgtcaagcatgggcccgcgccactcccaggctgccgccactctggc
cagaatgtcacagcaagacaagatcaacatgggtcacgggcatttggtgggacagagggccttgctggga
aacacagctgccatcagctccatcaactatcctcaaatctgtcttcaggatggaccattgggcattogct
tcggcactgggtaccacgccttcacacctggcggtccaagctgcttcgacatgggacgttgatctgatccg
gcagcgcggtgcttacctggcgccgaagccaagggtgctggcgcatcacatcctttggggcccggttggc
ggtgccctgggcaagattcccacggcggtcgcaactgggagggatttggcgccgacccctaccttggcg
gtattgccatgaaggagaccatcgagggtattcagtcagcaggcgtccaggccaacgccaaagcaactacat
tgcaaacgaacaagagctcaaccgcgagaccatgagcagcaatgtggatgaccgcactcagcagagctc
tacctctggccctttgcccagcgccgtgcacgccaaacgtcgccagcgctcatgtgcagttacaacaagctca
atggcacgtgggcttgcgagaatgacaaggctctgaatcagatcttgaagaaggagctcggattccaggg
ctacgttctcagcgactggaatgctcagcagcagcactgctctgtctgctaacagtgggtctggacatgact
atgcccgggtaccgatttcaacggccgcaatgtctactggggccctcaactgaacaacgctgtcaacgccg
gccaggttcagagatccagactagacgacatgtgcaagagaatcttggctggctggtacttgctcgggtca
gaaccagggtatcccgccatcaacatcaggggccaacgttcaggggcaaccataaggagaacgtacgtgct
gttgccagagacggcatcgctcttctgaagaacgatggaattctgccgctttccaagccgagaaagattg
ctgtcgtgggctcccactccgtcaacaatcccaggggaatcaacgcctgtgttgacaagggtgcaatgt
tggcacccttggtcatgggctggggttcaggcagcgtaactacccctatctcgtgtccccgtacgatgt
ctcggactcgtgctcaggccgatggcacacaaatcagcctccacaacactgacagcaccaacgggtgtgt
caaacgttgtgtctgacgtgatgtgtgtgtgtgtcatcactgccgattctggtgaagggtacatcac
tgtcgagggccacgctggcgaccgcagccaccttgaccctggcacaatggcaaccaacttggtcaggct
gcgcggctgccaaacaagaacgtcatcggtgtgtgtgcacagtgttgccagatcacctggagactatcc
tcaacaccaatggagtccgcgcgattgtgtgggctggtcttcggggccaagagaatggcaacgctcttgt
tgatgttctctacggcttggtttcgccatctggaaagcttccctacaccattggcaagagggagtcggac
tatggcacagccgttggttcgtggggatgataacttcagggaaggcccttttgggtgactaccgtcactttg
acaatgccaggatcgagccgcgctatgagtttggtcttgggttaagttccagcggcgaggttgggt
ttgatttcaagctttcctaacctgataaaacagcttacaccaatttcaccttctccgacatcaagattac
ttccaatgtcaagccggggcccgctactggccagaccattcccgccggacctgccgacctgtgggaggac
gttgcgacagtcactgcaaccatcaccaactcgggtgctgtcgagggcgctgaggttgcccagctttaca
tcggcctgcggtcctcggtcctgcctctccccgaagcagctgcgtggattttccaagctgaagctggc
cccgggtgccagcggcactgccacattcaacctcagacgcagagatctcagctattgggatacccgcctc
cagaactgggtcgtgcccagcggcaactttgtcgtcagcgtcggcgccagctcgagagatatccgcttga
cgggcaccatcacggcgtag

FIG. 41A

45 / 138

SEQ ID NO:54

Protein sequence of Pa3D, a GH3 family β -glucosidase from *Podospira anserina*

malqtffllaaamlanaettgekvsrqapsgaqawaaahsqaaatlarmsqqdkinmvtgigwdrGPCVG
ntaaissinypqicldgplgirfgtggttaftpgvqaastwdvdlirrgaylgaeakgcgihillgpva
galgkiphggrnwefgpadpylagiamketiegiqsagvqanakhyanegelnretmssnvddrtqhel
ylwfpadavhanvasvmcsynklngtwacendkalnqilkkelgfggyvlsdwnaqhstalsansgldmt
mpgtdfnggrnvwygpgqlnnavnagqvqrsrlddmckrilagwyllggnqgypainiranvqgnhkenvra
vardgivllkndgilplskprkiavvgshsvnpqginacvdkgcnvgtlgmwgsgsvnpylvspyda
lrtraqadgtqislnhtdstngvsnvvsdadavvvvitadsgegyitveghagdrshldpwhngnqlvqa
aaaanknvivvvhsvgqitletilntngvrailvwaglpqgengnalvdvlyglvspsgklpytigkresd
ygtavvrgddnfreglfvdyrhfdnarietryefgfglsytnftfsdikitsnvkpgpatggtipggpad
lwedvatvtatitnsgavegaevaqllyiglpssapasppkqlrgfsklklapgasgtatfnlrrrdlsyw
dtrlqnwvvpsgnfvsvvgassrdirltgtita

FIG. 41B

46 / 138

SEQ ID NO:55

Nucleotide sequence of Fv3G, a GH3 family β -glucosidase from *Fusarium verticillioides*

atgtttccttcttccatatcttgtttggcgccctgagttctgatgagccagggctctactagctcagagcc
aaccggaatgtcatcacccgatgatacctacttctacgggtcaatcgccaccagtggtatcctacacgtaa
gcactctctctgatttcccaacgaaagcaatactgatctcttgaccagcggaacaggtagacaccggctc
atgggctgcegtgttagccaaagccaagaacttgggtgtcccagttgactcttgaagagaaagtcaacttg
actacaggaggccagacgaccaccggctgctctgggttcateccctggcattcccctgttaggctttccag
gactgtgttttagcagacgctggcaacgggtgtccgcaacacagattatgtgagctcgtttccctccgggat
tcatgtcgggtgcaagctggaatccggagttgacctacagccggagctactacatgggtgctgaggccaaa
gccaaagggcgtaacatccttctcgggtccagttttggacctttgggcccagtagttgaaggtggacgca
actgggaggggttttccaatgatccctacctggcggttaaattagggcattgaagctgtcgccgggtatcca
agaagccggagttgttgcattgcggaacatttcccttgctcaagagcaggagaccatagacttgcggcg
tctgtcactggggctgatgcaatctcatcaaatctcgatgacaagacactccatgaattatatctctgtt
aagcacatcatatcttggctgagtagatgaaccttactaacaccogaactgggcttttccgtgatgcagt
ccacgcccgaacttgccagtgatgtgtgcagctacaacagagcaaacattcacacgcctgccaaaactcg
aagcttctcaatggccttctcaagggcgagttaggattccagggttttgtcgtctcggactggggcgac
agcaatctggtatggcttcagcattggctggcctggatgttgcattgccagctcgatcttgtgggggtgc
caaccttacccttgggtgtgaacaacgggaactattcccaggtcacaggttgacaatatggttacacggtac
gcgaagtctcagccttacttctcaattcttttgaactgacaatcgtgttaggctccttgcaacttgggtatc
agttgaaccaggaccaagacaccgaagccccaggtcacggactcgttgcgaagctttgggagcctcacc
agtagtcgacgctcgcaacgcaagctccaagcctaactatctgggaacgggtgcagtcgagggccatgttctt
gttaagaacaccaacaacgcactgccattcaagcccaacatgaaactcgtttcttgggttcggatactctc
acaaagctcctgataagaacatcccagacccccgcccaaggcatgttctccgcttgggtctatcgggtgcca
atccgccaaacatcactgagctgaacctcggctttctcggaaatttgagtctcacatactccgccatccgcg
cccaacggaaccatcatctcgggtggaggctcgggtgccagcgttggactctgttcagctcacccttcg
atgcattcgtttctcgggcgaagaaagagggtactgcgcttttctgggattttgagagctgggataccta
tgtgaacctacatctgaagcttgcattcgttgcgtggtaatgcattgggttagcgaaggctgggatagacct
gcaacctatgatgcctatactgatgagctcatcaataacgtcgtgacaagtgcgctaacactattgttg
ttcttcacaatgctggaacacgacttgtggatggcttctttgggtcaccccaacgtcaccgctattatcta
cgctcatctcccaggtcaggatagtgagatgctctgggtatctttgctctatggcgatgagaacccatct
ggctgcctcccttacaccggtgcccgcaacgagacggattatgggtcacctgctgaagccagacttgactc
tcgcccccaaccagtagccaacactttcccagtcgcacttctccgaggggtattttcattgactaccgaca
tttcgatgctaagaacatcacgcctcgttccagtttgggttcggcttgagctacacaacctttgagtag
getagtctccagatctcaaagtcccaggcccagacaccggaataaccagctggtgctcttaccgagggag
gcggttcagattttgtgggacgctcgttgcactgtcacagcaagcgtcaggaacactgggtctgtcgacgg
caaggaggttgccacagctatacgttgggtgttccaggtggctctatgagacagctacgtggctttacgaaa
ccagctattaaggctggagagacggctacagtgacctttgagcttactcgcgcgacttgagtgctctggg
atgttaatgcccaggagtggaacttcagcaaggcaactatgctatctacgttggccgaagtagtcgaga
tttgctctgcgaagtagccttgagcatctag

FIG. 42A

47 / 138

SEQ ID NO:56

Protein sequence of Fv3G, a GH3 family β -glucosidase from *Fusarium verticillioides*

mfpssisclaalslmsqgllaqsqpenvitddtyfyggsppvypthtgswwaavakaknlvsqtleekv
nlttggqtttgcsqgfipgiprvvgfpglcladagngvrntdyvssfspgihvgaswnpeltysrsyymgae
akakgvnillgpvfgplgrvveggrnwegfsndpylagklgheavagiqdagvvacgkhflaqeqethrl
aasvtgadaissnlddktlhelylcvmc synrannshacqnsklngllkgelgfgfvvswdgaqqsgm
asalagldvmpssilwganltlgvnngtipesqvdnmvtrllatwyqlngdqdeapghglaaklweph
pvvdarnasskptiwdgaveghvlvkntnnalpfpkpnmkvlslfgyshkapdknipdpaqgmfsawsiga
qsanitelnlgflgnlsltysaiapngtiisgggsgasawtlfsspfdaivsrakkegtalfwdfeswdp
yvnptseacivagnawasegwdrrpatydaytdelinnvadkcantivvlhnagtrldvgffghpnvtaii
yahlpqgdsgdalvslllygdenpsgrlpytvarnetdyghllkpdltlapnqyqhfpqsdfsegifidyr
hfdaknitprfegfglsyttfeyaslqisksqagtpeypagalteggrrsdlwdvvtasvrntgsvd
gkevaqlyvgvpggpmrqlrgftkpaikagetatvtfeltrrdlsvwdvnaqewqlqqgnyaiyvgrssr
dlplqstlsi

FIG. 42B

48 / 138

SEQ ID NO:57**Nucleotide sequence of Fv3D, a GH3 family β -glucosidase from *Fusarium verticillioides***

atggctagcattcgatctgtgttgggtctcgggtcttttggcgcgggtgtcaatgcccagcctacgatg
cgagtgatecgcgctgaagatgctttcagctgggtccagcccaagaacaccactattcttggacagtaagg
ccattcgccctcattaccctgccagtatgttcaccaactacaccaagtgacactgaggctgtactgacatt
ctagacaatgctactggcaagggctgggaagatgccttcgcccaagggtcaaaaactttgtotcccaactaa
ccctcgaggaaaaggccgacatgggtcacaggaactccagggtccttgcgctcggaacatcgtcgccattcc
ccgtctcaacttcaacgggtctctgtcttcacgaaggcccccctcgccatccgagtagcagactacgccagt
gttttccccgctgggtgtatcagccgcttcacgtgggacaaggacctcctctaccagcgcggtctcgcca
tgggtcaagagttcaaggccaaggggtgtcacatcctcctcgcccccgtcgccgggtcctcttggccgctc
ggcatactctgggtcgtaactgggaggggtttctcgccggacccttacctcactgggtattgcatggaggag
actatcatgggacatcaagatgctgggtgttcagggtactgccaagcactttatcggtaatgagcaggagg
tcattgcgaaccctacttttgtcaaggatgggtatatgttggtgaggttgacaaggagggtctttctgtotaa
catggatgatcgaaacctgcacgagctttacctctggccctttgccaatgctgttcatgccaaggcttcc
agcatgatgtgctcgtaaccagcgtctcaacgggtcctacgcctgccagaactcaaagggtcctcaacggaa
ttctgcggtgatgagcttgggtttccaggggtacgtcatgtcagattgggggtgccaccacacgcgggtgtgc
tgccatcaacagcgggtctcgacatggacatgcccgggtgggtatcggtgcctacgggaacatactttaccaag
tccttcttcggcggaaccctcaccgcgcgcgtcaccaacgggcaccctcgacgagaccgcgcgtcaacgaca
tgatcaccgcgatcatgactccctaacttctgggtcggccaggacaaggactatccctccgtcgaccctc
cagcgggtgatctcaacaccttcagccccaagagctcctgggtccgcgagttcaacctcaccggcgagcgc
agccgtgacgtccgcggtaaccacggcgacttgatccgcaagcagggcgccgaggtctaccgtccttctca
agaacgagaagaacgccttccctcaagaagcccaagtcctatcgctgtctttggcaacgatgctgggtga
tatcactgaggggtttctacaaccagaatgactaogaatttggcactcttggttgctgggtgggtctctgga
actgggtcgtttgacataccttgtttcgctctagccgcatcaatgctcgtgctaagcaggacgggtactc
ttgttcagcagtggtgaacaacactcttattgtaccaccaacgtcactgatctctggatccctgctac
tcccgatgtctgctcgttttcttgaagacttgggtctgaggagggtgctgatcgtgagcaccctctccgtt
gactgggacggtaaatgatgttggtgagtcgttgccaagtaactgcaataacactgtcgtcgtcactcact
cttctgggtatcaacactcttcttgggtgaccaccccaacgtcaccgctattctcgtcgtgccacttccc
cgggtcaggaggtctggcaactccctcggtgacctcctctacggcgatgtcaaccctctggtcgtcttccc
tacaccatcgcttcaacgggcaccgaactacaacgctccccccaccactgcggtcaacaccacgggaagg
aggactggcagtccttgggttcgacgagaagctcgagattgactaccgctacttcgacgcgcacacatctc
cgctccgctacgaattcgggttcgggtctctcctactccaccttcgaaatctccgacatctccgctgagcca
ctcgcatccgacattacctcccagcccgaggatctccccgtgcagcccgggcgaaccccgccctctggg
agaccgtctacaacgtgaaccgtctccgtctccaacaacgggcaagggtcgacggcgccactgtccccagct
atacgtgacattccccgacagcgcgcctgcccgtacaccacccaagcagctccgtgggttcgacaagggtc
ttccttgagggtggcgagagcaagagtgctcagctttgagctgatgcgccgtgatctgagctactgggata
tcatttctcagaagtggtcctcctgaggagaggttactattcgtgttggttcagcagtcgggactt
gaaggaggagacaaagggttactgttggttgaggcgtaa

FIG. 43A

49 / 138

SEQ ID NO:58

Protein sequence of Fv3D, a GH3 family β -glucosidase from *Fusarium verticillioides*

masirsvlvsqllaaqvnaqaydasdraedafswvqpknttilgqyghsphypannatgkgwedafaka
qnfvsqtlleekadmvtgtpgpcvgnivaiprlnfnlgclhdgplairvadyasvfpagvsaasswdkd
llyqrglamggefkkakgahillgpvagplgrsaysgrnwegfspdpyltgiameetimghqdagvqata
khfigneqevmrnptfvkdgyigevdkealssnmddrtmhelylwpfanavhakassmmcsyqrlngsy
acqnskvlnqilrdelgfggyvmsdwgathagvaainsgldmdmpggigaygtyftksffggnltravt
ngtldetrvndmitrimtpyfwlgqdkdypsvdpssgdlntfspksswfrelnltgersrdvrgnhgdl
irkhgaestvllkneknalplkkpksiavfgndagditegfynqndyefgtlvagggsgtgrltylvsp
laainarakqdgltvqqwmnntliattnvtdlwipatpdvclvflktwaeeaaadrehlsvdwdgndvve
svakycnntvvvthssgintlpwadhpnvtaillaahfpggesgnsldllygdvnpsgrlpytiafngt
dynappttavnttgkedwqswfdekleidyryfdahnisvryefgfglsystfeisdisaeplasdits
qpedlpvqpggnpalwetvynvtvsvsntgkvdgatvpqlyvtfpdsapagtppkqlrgfdkvfleage
sksvsfelmrrdlsywdiisqkwlipegeftirvgfssrdlkeetkvtvvea

FIG. 43B

50 / 138

SEQ ID NO:59**Nucleotide sequence of Fv3C, a GH3 family β -glucosidase from *Fusarium verticillioides***

atgaagctgaattgggtcgccgcagccctgtctataggtgctgctggcactgacagcgcagttgctcttg
cttctgcagttccagacacttttggctgggtgtaaaggctcagtttttttccaccatttctcgtctaatctc
agccttgttgccatategcccttgttgcgtcggaagccacgcaccagatcgcgatcatttctcctcttgc
agccttgggttctcttacgatcttccctccgcaattatcagcgccttagtctacacaaaaacccccgag
acagtctttcattgagtttgtcgacatcaagttgcttctcaactgtgcatttgctgggtgtctacttct
gctctagacaaccaaatctgggcgcaattgaccgctcaaaccttgttcaaataaccttttttatttcgag
acgcacatttataaatatgcgcttttcaataataccgactttatgcgcggtggtgtgtggcggttgat
cagaaagctgacgctcaaaagggttgctacgagagatacactcgacatactcgccgctcattatccttcac
catggatggaccctaattgctgttggctgggaggaagcttacgccaaagccaagagctttgtgtcccaact
cactctcatggaaaagggtcaacttgaccactgggtgttgggtaagcagctccttgcaaacagggttatctca
atccctcagctaacaaacttctcagatggcaaggcgaacgctgtgtaggaacgtgggatcaattcctcg
tctcggtatgcgaggtctctgtctccaggatgggtcctcttgggaattcgtctgtccgactacaacagcgct
tttcccgctggcaccacagctgggtgcttcttggagcaagctctctctggtatgagagaggtctcctgatgg
gcaactgagttcaaggagaagggtatcgatatcgctcttgggtcctgctaactggacctcttgggtcgcaactgc
tgctgggtggacgaaactgggaaggcttcaccgcttgatccttatatggctggccacgcacatggccgagggc
gtcaagggtattcaagacgcaggtgtcattgcttgtgtcgaagcattacatcgcaaacgagcagggtaagc
cacttggaacgatttgaggaattgacagagaactgacctcttgtagagcacttccgacagagtggcgagg
tccagtcgccgcaagtacaacatctccgagctctctctcctccaacctggatgacaagactatgcacgagct
ctaogcctggcccttcgctgacgcgctccgcgcggcgctcggttccgtcatgtgctcgtacaaccagatc
aacaactcgtacggttggcagaactccaagctcctcaacgggtatcctcaaggacgagatgggcttccagg
gtttcgtcatgagcatttggcgggccacgcataccggtgcgcgttctgcgcgtcgtggtctcgatatgag
catgcctgggtgacactgccttcgacagcggatacagcttctggggcggaacttgactctggctgtcactc
aacggaactgttcccgctggcgagttgatgacatggctctgcgaatcatgtctgccttcttcaagggttg
gaaagacgatagaggatcttcccgacatcaacttctcctcctggaccccgagacaccttcggcttctgtgca
tacatttgtcgaagagaacccgcgagcaggtcaactttggagtcgaacgtccagcagcaccacaagagccac
atccgtgagggcgctgccaagggaagcgctcgtgctcaagaacaccgggtccttccctcaagaacccaa
agttcctcgtgtcattgggtgaggacgcgggtcccaacctgctggacccaatgggtgtggtgacogtgg
ttgcgataatggtaccctggctatggcttggggctcgggaacttcccaattccttacttgatcaccctc
gatcaagggtctctaatcgagctactcaagacggaactcgatatgagagcatcttgaccaacaacgaat
gggttccagtacaagctcttgtcagccagcctaactgacogctatcgtttttcgccaatgccgactctgg
tgaggggatacattgaagtgcagcggaaactttgggtgatcgcaagaacctcaccctctggcagcagggagac
gagctcatcaagaacgtgtcgtccatatgccccaaacaccattgtagttctgcacaccgtcggtcctgtcc
tactcgccgactacgagaagaacccccaaacatcactgccatcgctgtgggtgttcccgcccaagagtc
aggcaatgccatcgctgatctcctctacggcaagggtcagccctggccgatctccttccacttggggccgc
accgcgagagctacggtactgaggttctttatgaggcgaacaacggccgtggcgctcctcaggatgact
tctctgaggggtgtcttcacgactaccgtcacttcgaccgaacgatctccaagcaccgatggaaagagctc
tcccaacaacaccgctgctcctctctacgagttcgggtcacgggtctatcttgggtccacctttgagtaactc
gacctcaacatccagaagaacgtcgagaacccctactctcctcccgctggccagaccatccccgccccaa
ccttttggaacttcagcaagaacctcaacgactaactgttccccaaggcggtccgatacatctacaagtt
catctaccccttctcaacacctcctcactccgcagcagggcatccaacgatggtggccagtttggtaag
actgccgaagagttcctccctcccaacgcctcacaoggtcagcccagcctcgtcttcccgctctgtgtg
ccccaggtggtaacctcaattgtgggacatcttgtacacogtcacagccacaatcaccaacacaggcaa
cgccacctccgacgagattccccagctgtatgtcagcctcgggtggcgagaacgagcccacccgtgttctc
cgcggttttcgacogtatcgagaacattgctcccgggccagagcgccatcttcaacgctcaattgacccgtc
gcgatctgagtaactgggatacaaatgcccagaactgggtcatcactgaccatcccaagactgtctgggt
tggaagcagctctcgcaagctgcctctcagcgccaaagttggagtaagaaagccaaacaagggttggtttt
tggaactgcaattttttgggaggacatagtagccgcgcgccagttacgtc

FIG. 44A

SUBSTITUTE SHEET (RULE 26)

51 / 138

SEQ ID NO:60

Protein sequence of Fv3C, a GH3 family β -glucosidase from *Fusarium verticillioides*

mklnwvaaalsigaagtdsavalasavpdtlagvkkadaqkvvtrdtlayspphyppwmdpnavgw
yakaksfvsqtlmekvnlttgvgwgqgercvgnvgsiprlgmrglclqdgplgirlsdynsafpagttag
aswskslwyergllmgtefkekgidialgpatgplgrtaaggrnwegftvdpymaghamaeavkgiqdag
viacakhyianegehfrqsgevqsrkyniseslssnlddktmhelyawpfadavragvgsvmcsynqinn
sygcqnsklngilkdemgfggfvmstdwaaghtgaasavagldmsmpgdtafdsgysfwggnltlaving
tvpawrvddmalrimsaaffkvgtiedlpdinfsswtrdtfgfvhtfaenreqvnfgvvnqhdhkshir
aaaakgsvglkntgslplknkflavigedagpnpagpncgdrgcdngtlamawgsqtsqfpylitpdq
glsnratqdgtryesiltnnewasvqalvsqpnvtaivfanadsgegyievdgnfgdrknltlwqqgdel
iknvssicpntivvlhtvgpvlladyeknpnitaivwaglpqgesgnaiadllygkvspgrspftwgrtr
esygtevlyeanngrgapqdddfsegvfidyrrhfdrrspstdgksspnntaaplyefghglswstfeysdl
niqknvenpysppagqtipaptfgnfsknldyvfpgkvryiykfiypflntssaseasndggqfgkta
eeflppnalngsaqprlpasgapggnpqlwdilytvtatitntgnatsdeipglyvslggenepirvlrg
fdrieniapggsaifnaqltrrdlsnwdtnaqnwvitdhpktvwvgsssrklplsakle

FIG. 44B

52 / 138

SEQ ID NO:61**Nucleotide sequence of Tr3A, a GH3 family β -glucosidase from *Trichoderma reesei***

atgcgttacccgaacagcagctgcgctggcacttgccactggggccctttgctagggcagacagtcagtata
gctgggtcccatactgggatgtgatatgtatccctggagacaccatgctgactcttgaatcaaggtagotca
acatcgggggctcggctgaggcagttgtacctcctgcagggaactccatggggaaccgcgtacgacaagg
cgaaggccgcattggcgaagctcaatctccaagataaggctcgcatcgtgagcgggtgtcggctggaacgg
cggtccttgcggttggaacacatctccggcctccaagatcagctatccatcgctatgccttcaagacgga
cccctcgggtgttcgatactcgacaggcagcacagcctttacgccgggcgttcaagcgggctcgcagctggg
atgtcaatttgatccgcgaacgtggacagttcatcgggtgaggaggtgaaggcctcggggattcatgtcat
acttggtcctgtggctgggcccgtgggaaagactccgcagggcgggtcgcaactgggagggttcgggtgtc
gatccatatctcacgggcattgccatgggtcaaaccatcaacggcatccagtcggtaggcgtgcaggcga
cagcgaagcactatatcctcaacgagcaggagctcaatcgagaaaccatttcgagcaaccagatgaccg
aactctccatgagctgtatacttggccatttgccgaocgggttcaggccaatgtcgttctgtcatgtgc
tcgtacaacaaggtcaataaccacctgggcctgcgaggatcagtacacgctgcagactgtgctgaaagacc
agctgggggttcccaggctatgtcatgacggactggaacgcacagcacacgactgtccaaagcgcgaattc
tgggcttgacatgtcaatgcctggcacagacttcaacggtaacaatcggctctgggggtccagctctcacc
aatgcggtaaatagcaatcagggtccccacgagcagagtcgacgatatgggtgactcgtatcctcgcgcgat
gggtacttgacaggccaggaccaggcaggctatccgtcgttcaacatcagcagaaatgttcaaggaaacca
caagaccaatgtcagggaattgccagggaaggcatcgttctgtcagaatgacgccaaacatcctgcgcg
ctcaagaagcccgtatgacattgccgtcgttggatctgccgcaatcattggtaaccacgccagaaactcgc
cctcgtgcaacgacaaaaggctgcgacgaacggggccttgggcatgggttgggggttcgggcgcgcgtcaacta
tcggtacttcgtcgcgcctacgatgccatcaataaccagagcgtcttcgcagggcacccagggttaccttg
agcaacaccgacaaacacgtcctcaggcgcacatctgcagcaagaggaaaggacgtcgccatcgtcttcatca
ccgcccactcgggtgaaggctacatcacctggaggggcaacgcgggcgatcgcaacaacctggatccgtg
gcacaacggcaatgccctgggtccaggcgggtggccgggtgccaaacagcaacgtcattgttgttgcactcc
gttggcgccatcattctggagcagattcttgcctcttcgcagggtcaaggccgttgtctgggcgggtcttc
cttctcaggagagcggcaatgcgctcgtcgcagctgctgtggggagatgtcagccctcttggaagctgggt
gtacaccattgcgaagagccccaatgactataacactcgcacatcgtttccggcggcagtgacagcttcagc
gagggaactgttcatcgactataagcaacttcgacgcagccaatatcacgcgcgggtacgagttcggtatg
gactgtgtaagtttgctaacctgaacaatctattagacaggttgactgacggatgactgtggaatgatag
cttacaccaagttcaactactcacgcctctcgtcttgtcgaccgccaagtctggctcctgcgactggggc
cgttgtgcccgggagggcccgagtgatctgttccagaatgtcgcgcagtcaccgttgacatcgcaaaactct
ggccaagtgactgggtgccgaggtagcccagctgtacatcacctacccatcttcagcaccacaggacccctc
cgaagcagctgcgaggctttgccaagctgaacctcacgcctgggtcagagcgggaacagcaacgttcaacat
ccgacgcagagatctcagctactgggacacggcttcgcagaaatgggtgggtgccgtcggggctcgtttggc
atcagcgtgggagcgcagcagccgggatatcaggctgacgagcactctgtcggtagcgtag

FIG. 45A

53 / 138

SEQ ID NO:62

Protein sequence of Tr3A, a GH3 family β -glucosidase from *Trichoderma reesei*

MRYRTAAALALATGPFARADSHSTSGASAEAVVPPAGTPWGTAYDKAKAALAKLNLQDKVGIVSGVGWNG
GPCVGNLTSPASKISYPSTLCLQDGPLGVRYSTGSTAFTPGVQAASTWDVNLIRERGQFIGEEVKASGIHVI
LGPVAGPLGKTPQGGRNWEGFGVDPYLTGIAMGQTINGIQSVGVQATAKHYILNEQELNRETISSNPDDR
TLHELYTWPFADAVQANVASVMCSYNKVNTTWACEDQYTLOTVLKDQLGFPGYVMTDWNQAHTTVQSANS
GLDMSMPGTDENGNNRLWGPALTNAVNSNQVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNISRNVQGNH
KTNVRAIARDGIVLLKNDANILPLKKPASIAVVGSAAIIGNHARNSPSCNDKGCDDGALGMWGSGAVNY
PYFVAPYDAINTRASSQGTQVTLSTNTDNTSSGASAARGKDVAIVFITADSGEGYITVEGNAGDRNNLDPW
HNGNALVQAVAGANSNVIVVHVSVAIILEQILALPQVKAVVWAGLPSQESGNALVDVLWGDVSPSGKLV
YTIKSPNDYNTRIVSGGSDSFSEGLFIDYKHFDANITPRYEFGYGLSYTKFNYSRLSVLSTAKSGPAT
GAVVPGGPSDLFQNVATVTVDIANSQVGTGA EVAQLYITYPSSAPRTPPKQLRGFAKLNLTPGQSGTATF
NIRRRDLSYWD TASQKWVVP SGSEGISVGASSRDIRLTSTLSVA

FIG. 45B

54 / 138

SEQ ID NO:63**Nucleotide sequence of Tr3B, a GH3 family β -glucosidase from *Trichoderma reesei***

atgaagacggttgctcagtggtttgctgcccgccttttggcgccgtagctgaggccaatccctaccgcctc
ctcactccaaccaggcgtaactcgccctcctttctacccttcgccatggatggaccccagtgctccaggctg
ggagcaagcctatgccccagctaaggagttcgctctcggttgactctcttgagagaaggtcaacctcacc
accggtggttgctggatgggtgagaagtgcgttggaacggttggtaccgtgcctcgcttgggcatgcgaa
gtctttgcatgcaggacggccccctgggtctccgattcaacacgtacaacagcgctttcagcggttggtt
gacggccgcgcgcagctggagccgacacctttgggttgaccgcggtaccgctctgggctccgaggcaaa
ggcaagggtgtcgatgttcttctcggaaccgctgggtggccctctcggtcgcaaccccaacggaggccgta
acgtcgagggtttcggtcggatccctatctggcggtttgggtctggccgataccgtgaccggaatcca
gaacgcgggcaccatcgctgtgccaagcacttccctcctcaacgagcaggagcatttccgccaggctcggc
gaagctaaccggttacggatacccatcaccgaggtctgtcttccaacggttgatgacaagacgattcacg
agggtgacggctggcccttccaggatgctgtcaaggctgggtgtcggttccctcatgtgctcgtaacca
ggtcaacaactcgtaacgttgccaaaactccaagctcatcaacggcttgctcaaggaggagtagcggttcc
caaggctttgtcatgagcgactggcaggcccagcacaacgggtgtcgctctgctgttgccggtctcgata
tgaccatgcctgggtgacaccgccttcaacaccggcgcatcctactttggaagcaacctgacgcttgctgt
tctcaacggcaccggtccccgagtgggcgcattgacgacatgggtgatgcgtatcatgggtcccttcttcaag
gtgggcaagacggttgacagcctcattgacaccaactttgattcttggaaccaatggcgagtagcggtacg
ttcaggccgcggtcaatgagaactgggagaaggtcaactacggcgctcgatgtccgcgccaacctatgcgaa
ccacatccgcgaggttggcgccaagggaactgtcatcttcaagaacaacggcatcctgccccttaagaag
cccaagtctcctgaccgtcattgggtgaggatgctggcggaaccctgcggcccccaacgggtgcggtgacc
gggctgtgacgacggcactcttgccatggagtggggatctgggtactaccaacttccccctacctcgctcac
ccccgacgcggccctgcagagccaggctctccaggacggcaccgcgtacgagagcatcctgtccaactac
gccatctcgcagacccaggcgctcgtcagccagcccgatgccattgccattgtctttgccaactcggata
gcggcgagggtacatcaacgtcgatggcaacgagggcgacgcgaagaacctgacgctgtggaagaacgg
cgacgatctgatcaagactgttgctgctgtcaaccccaagacgattgtcgtcatccactcgaccggcccc
gtgattctcaaggactacgccaaccaccccaacatctctgccattctgtgggcccgtgctcctggccagg
agtctggcaactcgctgggtcgacattctgtacggcaagcagagcccgggcccgcactcccttcacctgggg
ccgctcgctggagagctacggagttagtgttatgaccacgcccacaacggcaacggcgctccccaggat
aacttcaacgagggcgcccttcactcgactaccgctactttgacaagggtgggtcccggaagcctcgcagct
cggacaaggctcccacgtacgagtttggttcggactgtcgtggtcgacgttcaagttctccaacctcca
catccagaagaacaatgtcgggccccatgagccccgcccaacggcaagacgattgcggtccctctctgggc
agcttcagcaagaaccttaaggactatggcttcccccaagaacggttcgcgcgcatcaaggagtttatctacc
cctacctgagcaccactacctctggcaaggaggcgctcgggtgacgctcactacggccagactgcgaagga
gttccctccccgcgggtgccttggaacggcagccctcagcctcgctctgcggcctctggcgaaacccggcggc
aacgcgcagctgtacgacattctctacaccgtgacggccaccattaccaaacgggctcggtcatggacg
acgcggttccccagctgtacctgagccacggcggtcccaacgagccgccccaaagggtgctgcgtggcttcga
ccgcatcgagcgcatgtctccggccagagcgctacggttcaaggcagaacctgacgcgcggtgacctgtcc
aactgggacacgaagaagcagcagtggtcattaccgactaccccaagactgtgtacgtgggcagctcct
cgcgcgacctgcccgtgagcgccccgcctgccatga

FIG. 46A

55 / 138

SEQ ID NO:64

Protein sequence of Tr3B, a GH3 family β -glucosidase from *Trichoderma reesei*

mktlsvfaaallaavaeanpyppphsnqaysppfyppspwmdpsapgweqayaqakefvsgltllekvnl
ttgvqwmgekcvgnvgtvprlgmrslcmqdgplglrfntynsafsvgltaaaswsrhllwvdrgtalgse
akgkgvdvllgpvagplgrnpnggrnvegfgsdpylaglaladtvtgignagtiacakhflneqehfr
qvgeangygyppitealssnvddktihevyygwpfqdavgagvgsfmcsynqvnnsyacqnsklngllke
eygfqgfvmsdwqaqhtgvasavagldmtmpgdtafntgasyfgsnltlavlngtvpewriddmvmrim
apffkvgktvdsldtnfdswtngeyygvqaavnenwekvnygvdvranhanhirevgakgtvifknng
ilplkkpkfltvigedaggnpagpnngcgrgddgtlamewsgttnfpylvtpdaalqsqalqdgtry
esilsnysaisqtqalvsqpdaiaivfansdsgegyinvdgnegdrknltlwknngddliktvaavnpti
vvihstgpvilkdyanhpnisailwagapggesgnsldilygkqspgrtpftwgpslesygvsvmttp
nngngapqdnfnegafidyryfdkvapgkprssdkaptyefgfglswstfkfsnlhiqknnvgpmsppn
gktiaapslgsfsknlkdygfpknvrrikefiypylstttsqkeasgdahyggtakeflpagaldgspq
prsaasgepggnrqlydilytvtatitntgsvmdavpqlylshggpnepkvlgfdrieriapggsv
tfkadltrrdlsnwdtkkqqwvitdypktvyvgsssrldplsarl

FIG. 46B

56 / 138

SEQ ID NO:65

Nucleotide sequence of Te3A, a GH3 family β -glucosidase from *Talaromyces emersonii*, optimized for expression in *T.reesei*

atgcgcaacggcctcctcaagggtcgccgcttagccgctgccagcgccgtcaacggcgagaacctcgcc
acagcccccccttctaccccagccccctgggccaacggccagggcgactgggcccaggcctaccagaaggc
cgtccagttcgtcagccagctcaccctcgccgagaagggtcaacctcaccaccggcaccggctgggagcag
gaccgctgcgtcgccaggtcggcagcatcccccgcttaggcttccccggcctctgcatgcaggacagcc
ccctcggcgctcgccgacacgactacaacagcgcccttcctgcccggcgttaacgtcgccgccacctggga
ccgcaacttagcctaccgcagaggcgctcgccatggggcgaggaaacaccgcggcaaggggcgctcgacgtccag
ttagggccccgctcgccggcccccttagggcgctctcctgatgcggcgccgaactgggaggggttcgccccg
accccgctcctcaccggcaacatgatggccagcaccatccagggcacccaggatgctggcgctcattgcctg
cgccaagcaacttcactctctacgagcaggaacacttcggccaggcgcccaggacggctacgacatcagc
gacagcatcagcgccaacgcgcgacgacaagaccatgcacgagttatacctctggcccttcgcccgatgccg
tcgcgcgggtgtcggcagcgctcatgtgcagctacaaccagggtcaacaacagctacgcctgcagcaacag
ctacaccatgaacaagctcctcaagagcgagttaggcttcagggttcgctcatgacgcactggggcggc
caccacagcggcgctcggtctctgcccgcggcctcgacatgagcatgcccggcgacattgccttcgaca
ggggcagctctttctggggcaccaacctcaccggttgccgtcctcaacggctccatccccgagtgggcggt
cgacgacatggccgctcgcctcatgagcgccctactacaagggtcgcccgcgaccgctacagcgctccccatc
aacttcgacagctggaccctcgacacctacggccccgagcaactacgcctcgccaggggccagaccaaga
tcaacgagcacgctcgacgtcgcggcaaccacgcgcgagatcatccaagagatcggcgcgcctccgcggt
cctcctcaagaacaaggggcggtcctccccctcactggcaccgagcgcttcgctcggtgtcttttggaaggat
gctggcagcaacctctggggcggtcaacgggtgcagcgaccgcgggtgcgacaacggcaccctcgccatgg
gctggggcagcggcaccgccaactttccctacctcgtcacccccgagcaggccatccagcgcgagggtcct
cagccgcaacggcaccttcaccggcatcaccgacaacggcgcccttagccgagatggcgctgcccgcctct
caggccgacacctgctcgtcttttgccaacgcgcgactccggcgagggtacatcaccgtcgatggcaacg
agggcgaccgcaagaacctcaccctctggcagggcgccgaccagggtcatccacaacgtcagcgccaactg
caacaacaccgctcgctcgtctttacacaacgctcgggccccgctcctcatcgacgactggtaacgaccacccccaac
gtcaccgcccatacctctggggcggtttaccgggtcaggaaagcgggcaacagcctcgctcgacgtcctctacg
gcccgtcaaaccccggaagaccccccttcacctggggcagagcccgcgaagactatggcgccccctctcat
cgtcaagcctaacaacggcaaggggcgccccccagcaggacttcaccgagggtcatcttcacgactaccgc
cgcttcgacaagtacaacatcaccctcatcagagttcggttcggcctcagctacaccaccttcgagt
tcagccagttaaacgtccagcccataaacgccccctccctacacccccgcagcggttttaagaaggcgcc
ccagagcttcggccagccctccaatgccagcgacaacctctacccctagcgacatcgagcgcgctccccctc
tacatctacccctgggtcaacagcaccgacctcaaggccagcgccaacgacccccgactacggcctcccca
ccgagaagtaagtcccccccaacgcaccacaacggcgacccccagcccattgacctgcccggcggtgcccc
tgggcggaacccccagcctctacgagcccgtcgcccgcgtcaccaccatcatcaccacacccggcaagggtc
accggcgacgaggtccccccagctctatgtcagcttagggcgccctgacgacgcccccaagggtcctccgcg
gcttcgaccgcatcaccctcgccccctggccagcagtaacctctggaccaccacctcactcgccgcgacat
cagcaactgggaaccccgtaaccagaactgggtcgtaaccaactacaccaagaccatctacgtcggcaac
agcagccgcaacctccccctccaggccccccctcaagccctaccccggcacatctgatga

FIG. 47A

57 / 138

SEQ ID NO:66

Protein sequence of Te3A, a GH3 family β -glucosidase from *Talaromyces emersonii*

mrnqlkvaalaaasavngenlaysppfypspwangqgdwaeayqkavqfvsqltlaekvnlttggtgweq
drcvgqvg siprlgfp glcmqdsplgvrdtdynsafpagvnvaatwdrnlayrrgvamgeehrgkgvdvq
lgpvagplgrspdagrnwegfapdpvltgnmmasti qgiqdagviacakhfilyeqehfrqgaqdgydis
dsisanaddktmhelylwpfadavragvgsvmcsynqvnnsyacsnsytmnkllkselgfggfvmtdwgg
hhsgvg salagldmsmpgdiafdsgtsfwgtnltvavlngsipewrvddmavrimsayykvgrdrysvpi
nfdswtldtygpehyavgggqtkinehvdvrgnhaeiiheigaasavllknkgglpltgterfvvgvfgkd
agsnpwgvngcsdrgcdngtlamgwsggtanfpylvtp eqaiqrevlsrngtftgitdngalaemaaaas
qadtclvfanadsgegyitvdgnegdrknltlwggadqvihnvsancnntvvvlhtvgpvliddwydhpn
vtailwaglpqgesgnsldvlygrvnp gktpftwgrarddygaplivkpnnkggapqqdftegifidyr
rfdkynitpiyefgfglsyttfefs qlnvqpinappytpasgftkaagsfgqpsnasdnlypsdiervpl
yiypwlnstdlkasandpdyglptekyvppnatngdpqp idpaggapggnp slyepvarvtiitntgkv
tgdevppqlyvslggpddapkv lrgfdritlapggqylwtttltrrd isnwdpvtqnwvvtnytktiyvgn
ssrnlplqaplkpypgi

FIG. 47B

58 / 138

SEQ ID NO:67**Nucleotide sequence of An3A, a GH3 family β -glucosidase from *Aspergillus niger***

atgcgcttcaccagcagcagggccgctcgccctcaccgcccgcagcctcgccagcgcgcgacgagttagcct
acagccccccctactaccccagccccctgggccaacggccagggcgactgggcccagggcctaccagcgcgc
cgctcgacatcgctcagccagatgacctcgccgagaaggtcaacctcaccaccggccacccggtgggagtta
gagttatgcgctcgccagactggtggcgctcccccgccctcgccatccccggcatgtgcgcccaggacagcc
ccctcggcgctccgagacagcgactacaacagcgccctccctgcccggcgtcaacgtcgccgccacctggga
caagaacctcgccctacctccgcgggccaggccatgggcccaggaattcagcgacaaggggcgccgacatccag
ttaggccccgctgcccggccctttaggccgctctcccgacggcgccagaaaactgggagggccttcagccccg
accccgctctcagcggcgctcctcttcgcccagagactatcaaggccatccaggatgctggcgctcgccac
cgccaagcaactacattgcctacgagcaggaacacttcgcccaggcccccgaggcccagggtacggcttc
aacatcacgagagcgccgagcgccaacctcgacgacaagaccatgcacgagttatacctctggcccttcg
ccgacgcccattagagctggcgctggtgctgtcatgtgcagctacaaccagatcaacaacagctacggctg
ccagaacagctacacctcaacaagctcctcaaggccgagttaggcttcagggttcgtcatgtccgac
tgggcccgcaccacgcggcgctcagcggcgcccttagccggcctcgacatgagcatgcccggcgacgtcg
actacgacagcggccaccagctactggggcaccaacctcaccatcagcgctcctcaacggccaccgtcccca
gtggcgcgctcgacgacatggccgctccgcatcatggccgcctactacaagggtcgcccgcgaccgcctctgg
accccccccaacttcagcagctggaccccgagcaggtacggcttcagtaactactacgtcagcgagggcc
cctatgagaaggtcaaccagttcgtcaacgtccagcgcaaccacagcgagttaatccgcccgcacggcg
cgacagcacccgtcctcctcaagaacgacggcgccctccccctcaccggcaaggaaacgcctcgccctc
atcgcgagaggacgcggcgagcaaccctacggcgccaacggctgcagcgaccggcggtgcgacaacggca
ccctcgccatgggctggggcagcgccaccgccaacttccttacctcgtaacccccgagcaggccatcag
caacgaggtcctcaagaacaagaacggcgctctttaccgccaaccgacaactgggcccacgaccagatcgag
gccttagccaagaccgcctctgtcagcctcgctctttgtcaacggcgacagcgggcgagggctacatcaacg
tcgacggcaacctcgcgacccgcccgaacctcaccctctggcgcaacggcgacaacgtcatcaaggccgc
cgccagcaactgcaacaacaccatcgctcatcaccacagcgctcgcccccgctcctcgtaacgagtggtac
gacaacccccaaacgtcacccgcatcctctggggcggttacccggccaggaaagcggaacagcctcgccg
acgtcctctacggccgctcaaccttgccgccaagagcccccttcacctggggcaagaccgcgaggcccta
tcaggactacctctacaccgagcccaacaacggcaacggcgccccccagggaagatttcgtcgagggcgctc
tttatcgactaccgcggttttgacaagcgcaacgagactcccatctacgagttcggctacggcctcagct
acaccaccttcaactacagcaacctccagggtcgaggtcctcagcgccccctgcctacgagcccgcagcgg
cgagactgagggcgccccccaccttcggcgaggtcggaacggccagcgactacttataccccgacggcctc
cagcgcatcaccaagttcatctacccctgggtcaacagcaaccgacctcgaggccagcagcgggcgacgct
cttacggccaggacgcctccgactacctccccgaggggtgccaccgacggcagcgctcagcccatcttacc
tgccggtggcggtgctggcggaacccccagactctacgacgagctgatccgctcagcgctaccatcaag
aacaccggcaaggtcgctggtgacgaggtccccagctctacgtcagcttaggcggccctaacgagccca
agatcgctcctccgccagttcgagcgcatcacctccagcccagcaaggaaactcagtgagaccaccct
cactcgcccgacctcgccaactggaacgtcgagactcaggactgggagatcaccagctacccaagatg
gtctttgcccggcagcagcagccgcaagctccccctccgcccagcctccccaccgtccactgatga

FIG. 48A

59 / 138

SEQ ID NO:68

Protein sequence of An3A, a GH3 family β -glucosidase from *Aspergillus niger*

mrftsieavaltavslasadelaysppyyppswangggdwaeayqgravdivsqmtlaekvnlttgtgwel
elcvgqgtggvprlgipgmcaqdsplgvrdsdynsafpagvnvaatwdknlaylrgqamgqefsdkgadiq
lgpaagplgrspdggrnwegfspdpalsgvlfaetikgiqdagvvatakhyaieqehfrqapeaaggygf
nitesgsanlddktmhelylwpfadairagagavmcsynqinnsygcqnsytlkllkaelgfggfvm
waahhagvsgalagldmsmpgdvdydsgtsywgtnltisvlngtvpqwrvddmavrimaayykvgrdr
lwtppnfsswtrdeygfkyyyvsegpyekvnqfvnvqrnhselirrigadstvlkndgalpltgkerlval
igedagsnpygangcsdrgcdngtlamgwsgtanfpylvtpaqaisnevlknknvgftatdnwaidqie
alaktasvslvfvnadsgegyinvdgnlgdrrnltlwrngdnvikaaasncnntiviihsvgpvlvnew
ydnpnvtailwgglpqgsgnsladvlygrvnpgakspftwgktreayqdylytepnnngngapqedfvegv
fidyrgfdkrnetpiyefgyglsttfnysnlqvevlsapayepasgeteaaptfgevgnasdylypdgl
gritkfiypwlnstdleassgdasygqdasdylpegatdgsaqpilpagggaggnprlydelirvsvtik
ntgkvagdevpqlyvslggpnepkivlrqferitlqpsketqgsttlttrrdlanwnvetqdweitsypkm
vfagsssrklplraslptvh

FIG. 48B

60 / 138

SEQ ID NO:69**Nucleotide sequence of Fo3A, a GH3 family β -glucosidase from *Fusarium oxysporum***

atgaagctgaactgggtcgccgcagccctctctataggtgctgctggcactgatgggtgcagttgctcttg
cttctgaagttccaggcacttttggtgggtgtaaaggctcggttttttaaccatttccctcacctaattctcag
ccttggttgccatatcgcccttattcgctcggaacgtacgcaccaaatecgcatcatttccctcccttgag
ccttggtttcttttttcgatcttccctccgcaatcgccagcaccccttagcctacacaaaaacccccgaga
cagttctcattgagtttgctgacatcaagttgcttctcaagtgctgcatttgctggtggtgctacttctgcc
tctagaccaccaaactctgggcgcaattgatcgctcaaacccttggttcgaataagcctttttattcgagacgt
ccaatttttacagagaatgtacctttcaataataaccgacgttatgctgctggcggtggctgctgtgatgggt
gttgatcagaataactgacgtcaaaagggttgctacgagagatacactcgcacactcacctcctcactatc
cttcaccatggatggatcctaattgccattggctgggagggaagcttacgccaaagcaaagaactttgtgtc
ccagctcactctcctcgaaaagggtcaacttgaccactgggtgttggttaagtagctccttgcaaacagtg
atctcggtctccttgactaacgactctctcaggtggcaaggcgaaacgtgtgtaggaaacgtgggatcaa
ttcctcgtcttggtatgctgaggtctttgtcttcaggatgggtcctccttggaattcgtctgtccgattacaa
cagtgcttttcccgctggcaccacagctgggtgcttcttgagcaagttctctctggtatgagaggggtctt
ctgatgggaactgagttcaaggggaagggtatcgatatcgctccttgccctgctactggctcctcttgcc
gcaactgctgctggtggacgaaactgggagggtttaccggtgatccttatatggctggccatgccatggc
cgaggccgtcaagggcatccaagacgcaggtgtcattgcttggtgctaagcattacatcgcaaacgagcaa
ggtaagccaattggacgggtttgggaaatcgacagagaactgaccccccttgtagagcacttccgacagagt
ggcgaggtccagtcgccgaagtacaacatctccgagttctctcctccaacctggacgacaagactttgc
acgagctctacgctggccctttgtgatgcccgtccgctggtgctgggttcagtcattgtgtctctttaa
tcagatcaacaactcgtaacggttgccagaactccaagctcctcaacgggtatcctcaaggacgagatgggt
ttccagggtctcgatgagcgattgggcccagcacacccgtgctgcttctgcccgtcgtggtcttg
atatgagcatgctggtgacaccgcttcgacagtggtatagcttctgggggtggaaacctgactcttg
tgtcatcaacggaactgttccgcctggcgagttgatgacatggctctgcgaatcatgtcggccttcttc
aaggttggaagacggtagaggacctccccgacatcaacttctcctcctggacccgcgacaccttcggct
tcgtccaaacatttgctcaagagaaccgcgaacaagtcaactttggagttaacgtccagcacgaccacaa
gaaccacatccgtgagttgcccgaagggaagcgtcatcctcaagaacacccggtcccttccccctcaac
aatcccaagttcctcgctgtcattgggtgaggacgccggtcccaacctgctggacccaatggttgccg
accgtggttgcgacaatggtaccctggctatggcttggggctcgggaaacttctcaattcccttacttgat
cacaccgaccaaggtctccagaaccgagctgcccgaagcggaaactcgatatgagagcatcttgaccaac
aacgaatgggcccagacacaggctcttgctcagccaaacccacgtgacccgtatcggttttgccaacgccc
actctggtgaggggttacattgaagtcgaacggaaaacttcggtgatcgcaagaacctcacctctggaaca
gggagacgagctcatcaagaacgtctcgtccatctgccccaacaccattgtcgttctgcataccgtcggc
cctgtcctgctcgccgactacgagaagaaccccaacatcacccgcatcgtctgggctggtcttcccgcc
aagagtctggcaatgccatcgctgatctcctctacggcaaggtaagccctggccgatctcccttcaactg
gggcccgaaccgtgagagctacggtaccgaggttctttatgaggcgaacaacggccgtggcgctcctcag
gatgacttctcgagggtgtcttctcattgactaccgtcactttgatcgacgatctccagcacccgatggca
agagcgtcccaacaacacogctgctcctctctacgagttcgggtcatggtctgtcttggaactaccttgga
gtattcagacctcaacatccagaagaacgttaactccacctactctcctcctgctgggtcagaccattcct
gccccaacctttggcaacttcagcaagaacctcaacgactacgtgttccctaagggtgtccgatacatct
acaagttcatctaccccttctgaacacttccctcatccgccagcgaggcatctaacgacggcgccaggtt
tggttaagactgccgaagagttcctacctccaaacgccctcaacgggtcagcccagcctcgtcttccctct
tctggtgccccaggcggttaacctcaattgtgggatctcctgtacaccgtcacagccacaatcaccaaca
caggcaacgccacctccgacgagattccccagctgtatgtcagcctcgggtggcgagaacgaacccggttcg
tgtcctccgcggtttcgaccgtatcgagaacattgctcccgccagagcgccatcttcaacgctcaattg
acccgtcgcgatctgagcaactgggtatgtggatgccagaactgggttatcaccgacctccaaagacgg
tgtgggttggaagtagttctcgcgaagctgcctctcagcgccaagttggaataa

FIG. 49A

SUBSTITUTE SHEET (RULE 26)

61 / 138

SEQ ID NO:70

Protein sequence of Fo3A, a GH3 family β -glucosidase from *Fusarium oxysporum*

mklnwvaaalsigaaqtdqavalasevpgtlagvkntdaqkvvtrdtlahspphypspwmdpnaigweea
yakaknfvsqlltlekvnlttgvgwggercvgnvgsiprlgmrglclqdgplgirlsdynsafpagttag
aswskslwyergllmgtefkkgidialgpatgplgrtaaggrnwegftvdpymaghamaeavkgiqdag
viacakhyianegehfrqsggevqsrkyniseslssnlddktlhelayawpfadavragvgsvmcsynqinn
sygcqnsklngilkdemgfqgfvmstdwaaghtgaasavagldmsmpgdtafdsgysfwggnltlaving
tvpawrvddmalrimsaffkvgtvedlpdinfsswtrdtfgfvqtfagenreqvnfgvvnvghdhknhir
esaakgsvilkntgslplnnpkflavigedagpnpagpncgdrgcdngtlamawsgtsqfpylitpdq
glqnraaqdgtryesiltnnewaqtqalvsqpnvtaivfanadsgegyievdgnfgdrknltlwqqgdel
iknvssicpntivlhtvgpvladyeknpnitaivwaglpqgesgnaiadllygkvspgrspftwgrtr
esygtevlyeanngrgapqdddfsegvfidyrhfdrrspstdgksapnntaaplyefghglswttfeysdl
niqknvnstysppagqtipaptfgnfsknldyvfpgkvryiykfiypflntssaseasndggqfgkta
eeflppnalngsaqprlpssgapggnpqlwdilytvtatitntgnatsdeipqlyvslggenepvrvlrg
fdrieniapggsaifnaqltrrdlsnwdvdaqnwvitdhpktvwvgsssrklplsakle

FIG. 49B

62 / 138

SEQ ID NO:71**Nucleotide sequence of Gz3A, a GH3 family β -glucosidase from *Gibberella zeae***

ATGAAGGCCAATTGGCTTGCCGCGGCCGTTTATTTGGCTGCTGGCACCGATGCTGCAGTCCCTGACACTT
TGCGCAGGAGTCAATGTAAGCTACTCTTCAATTTTCATCTCATCTCAACTTTGCCAGGCCACAACAACCTTTT
CTTCACTCACGATCTTTTACCATAAACGCAACAGTTTCACAAAAAATAAAGCCCAAATCATGTCTCTGA
TCGTTGAACTCGCCATCTTCGTTTACATCGCGGTTGTCTTTTTCTTCTTGACTTCTCATTCGTTGTTGT
TCTCTACATTTTTCGACTGGCTGTTTAGCCTTGAGATTCTTCTCACTCCCCGTGATGCCTAGATCACTCTC
TGAGGCGTTTAATCTACTTGTAGAGATGCGCCTCTCATTTGTTGTGTGCTAGTCGCGATAGTTGCTGGA
ATTGCAGTCCTTGATCTTCTACTGACACTCAAAAGCTCGTTGCGCGGGACACACTCGCTCACTCTCCTC
CTCACTATCCCTCGCCATGGATGGACCCTAACGCTGTGCGGCTGGGAGGACGCCCTACGCCAAGGCCAAGGA
CTTTGTCTCCAGATGACTCTCCTAGAAAAGGTCAACTTGACCACTGGTGTGGGTAAGTAACGAGCGAC
AAGACGTCTACAATCCACTAACACGATCTCTAGATGGCAGGGCGAACGTTGTGTTGGAAACGTGGGATCT
ATCCCTCGTCTCGGTATGCGAGGCCTCTGTCTCCAGGATGGTCTCTCGGAATTCGCTTCTCCGACTACA
ACAGCGCTTTCCCTACTGGTGTACCGCTGGTGTCTTGGAGTAAGGCCCTTTGGTACGAGCGAGGACG
ATTGATGGGTACCGAGTTTAAGGAGAAGGGTATCGATATTGCTCTCGGCCCTGCAACTGGTCTCTCGGT
CGCCACGCTGCTGGTGGACGAACTGGGAAGGCTTCACTGTGACCCCTACGCCGCTGGCCATGCTATGG
CTGAGACTGTCAAGGGTATCCAAGATTCTGGAGTCATTGCTTGTGCTAAGCATTACATCGCAAACGAGCA
AGGTATGTACAGGCCCATTC AATGGCTTCAGGAACGAAACTAACTCTTAATAGAACACTTCCGTCAACG
AGGCGATGTATGTCTCAAAAGTTCAACATTTCCGAGTCTCTGTCTTCCAACCTTGACGATAAGACTATG
CACGAGCTCTACAACCTGGCCTTTTCGCCGACGCCGTCCGCGCCGGTGTGGCTCCATTATGTGCTCTTACA
ACCAGGTCAACAACCTCATATGCTTGCCAGA ACTCCAAGCTCCTCAACGGCATCCTCAAGGACGAGATGGG
TTTCCAGGGTTTCGTCATGAGCGATTGGCAGGCTCAGCACACCGGTGCCGCCCTCCGCTGTTGCCGGTCTT
GACATGACCATGCCTGGTGACACCGAGTTCAACACTGGCTTCAGCTTCTGGGGTGGAAACCTGACCCTCG
CTGTTATCAACGGTACTGTTCCCGCCTGGAGAATCGACGACATGGCTACCCGAATTATGGCTGCTTTCTT
CAAGGTTGGCCGATCTGTTGAGGAGGAACCCGACATCAACTTCTCAGCTTGGACTCGTGATGAGTATGGC
TTCGTCCAGACCTACGCCCAAGAGAACCAGAGAAAAGGTCAACTTTGCTGTTAATGTCCAGCACGACCACA
AGCGCCACATTCGCGAGGCTGGCGCAAAGGGATCCGTCGTCCTCAAGAACACTGGCTCACTTCCTCTTAA
GAAGCCCCAGTTCCTCGCTGTCATTGGAGAGGACGCTGGTTCCAACCTGCCGGACCCAACGGTTGCGCT
GACCGTGGATGCGACAACGGTACTCTTGCCATGGCATGGGGTTCCGGAACCTCTCAATTCCCCTACCTTG
TCACCCCCGACCAAGGCATCTCGCTCCAGGCTATTCAGGACGGTACTCGTTATGAGAGCATCCTCAACAA
CAACCAGTGGCCCCAGACACAAGCTCTTGTACGCCAGCCCAACGTCACCGCCATTGTCTTTGCCAATGCC
GATTCTGGTGAGGGCTACATCGAGGTTGACGGCAACTACGGCGACCGCAAGAACCTCACTCTGTGGAAGC
AAGGCGATGAGCTCATCAAGAACGTCTCTGCTATCTGCCCCAACACCATTGTGGTCTTTCACACCGTTGG
CCCCGTCTTCTAACCGAGTGGCACAACAACCCCAACATCACCGCCATTGTTTGGGCTGGTGTGCCTGGA
CAGGAGTCCGGTAACGCCATCGCCGACATCCTCTACGGCAAGACCAGCCCTGGACGTTCTCCCTTCACCT
GGGGTCGCACTTATGACAGCTATGGCACCAAGGTTCTCTACAAGGCCAACAATGGAGAGGGTGCCCTCA
AGAGGACTTTGTGCGAGGGCAACTTCATCGACTACCGCCACTTTGACCGACAATCCCCCAGCACCAACGGA
AAGAGTGCCACCAACGACTCTTCTGCTCCTCTCTACGAGTTCCGGTTTCGGTCTGTCTCTGGACTACCTTTG
AGTACTCTGATCTCAAAGTCGAGTCTGTGAGCAACGCCTCTTACAGCCCCTCTGTGCGAAACACCATTCC
TGCCCCCTACCTACGGCAACTTCAGCAAGAACCTGGACGATTACACATTCCCCTCAGGTGTCCGATACCTC
TACAAGTTCATCTACCCCTACCTCAACACCTCTTCCCTCCGCTGAGAAGGCTTCCGGCGATGTCAAGGGCA
GATTTGGTGAGACCGGCGACGAGTTCCCTCCCTCCCAACGCTCTCAACGGTTCATCGCAGCCTCGTCTCC
TTCCAGTGGTGCTCCCGGCGGTAACCCCTCAGCTCTGGGACATTATGTACACCGTCACTGCCACCATCACC
AACACTGGTGACGCTACCTCGGATGAGGTTCCOCAGCTGTACGTCAGCCTCGGTGGTGAGGGCGAGCCTG
TCCGTGTCTCTCCGTGGCTTCGAGCGTCTTGAAAACATTGCTCCTGGTGAGAGTGCCACATTCACCGCTCA
GCTTACTCGCCGTGACCTGAGCAACTGGGACGTCAACGTCCAGAACTGGGTCACTACCGATCACGCCAAG
AAGATCTGGGTGCGCAGCAGCTCTCGCAATCTGCCCTCAGCGCCGACCTGTAG

FIG. 50A

SUBSTITUTE SHEET (RULE 26)

63 / 138

SEQ ID NO:72

Protein sequence of Gz3A, a GH3 family β -glucosidase from *Gibberella zeae*

mkanwlaaavyllaagtdaavpdtlagvnlvardtlaahspphyppspwmdpnavgwedayakakdfvsqmtl
lekvnlttgvgwqgercvgnvgsiprlgmrglclqdgplgirfsdynsafptgvtagaswskalwyergr
lmgtefkekgidialgpatgplgrhaaggrnwegftvdpaaaghamaetvkgiqdsgviacakhyianeq
ehfrqrgdvmsqkfniseslssnlddktmhelynwpfadavragvgsimcsynqvnnsyacqnsklngi
lkdemgfqgfvmstdwqaqhtgaasavagldmtmpgdtefntgfsfwggnltlavingtvpawriddmatr
imaaffkvgrsveeepdinfawtrdeygvqtyagenrekvnfavnvqhdhkrhireagakgsvvlknt
gslplkkpqflavigedagsnpagpngcadrgcdngtlamawgsqtsqfpylvtpdqqislqaiqdgtry
esilnnnqwpqtqalvsqpnvtaivfanadsgegyievdgnygdrknltlwkggdeliknvsaicpntiv
vlhtvgpvlltewhnnpnitaiwagvpgqesgnaiadilygktspgrspftwgrtydsygtkvlykann
gegapqedfvegnfidyrhfdrrqspstngksatndssaplyefgfglswttefsdlkvesvsnasysps
vgntipptygnfsknlddytfpsgvrylykfliypylntsssaekasgdvkggrfgetgdeflppnalngs
sqprlpssgapggnpqlwdimytvtatitntgdatsdevpqlyvslggegepvrvlrgferleniapges
atftaqltrrdlsnwdvvnvqnwvitdhakkiwvgsssrnlplsadl

FIG. 50B

64 / 138

SEQ ID NO:73**Nucleotide sequence of Nh3A, a GH3 family β -glucosidase from *Nectria haematococca***

atgCGgttcacCGtccttctcgcggcatttttcggggccttgTccccatgggttggttcgcaagctgaccaga
aaccactacagctcgggtgtgaacaataacactctggcgcattcacctcctcactatecttcgccatggat
ggatcctgctgctcctggctgggaggaagcctatctcaaggcgaaagattttgtttcacagcttaccctt
cttgaaaagggtcaacttgaccactgggtgtgggtgagtcacttgTTTTctctctcctgacgtgacactt
tgctttggcctgcttccatatatcgtctactagcatttgctaacactcgaggcagatggatgggCGaacgtt
gCGtcggcaacgtgggttcactccctcgtttttggaatgcgtgggtctctgcatgcaggatggccccctcgg
catccgcttgctgactataactctgcctttcctactgggtattacagctgggtgctccttgagcgctgcc
ctttgggtaccaacgtggcctcctgatgggcaccgagcatcgtgaaaaaggcatcgacgttgcacttgggc
ctgctactgggtcctcttggtcgtactcctactggcgccgcaactgggaggggtttctcgggtgatcccta
cgttgctggcggttgccatggccgagactgttagcggcattcaagatgggtggtactatcgctgtgctaag
cactacatcggaacgaacaagggtatgcctcttcaacttctcctcgtgataaatctgctcacaacaacct
agagcaccatcgccaagccccgaatccattggccgCGgtacaacatcaccgagtcctgtcgtcgaac
gttgatgacaagacctccacgagctctatctctggccgttcgcagatgcCGtcaaggctgggtgttggtg
ctatcatgtgttctaccagcagctgaacaactcttacgggttgccaaaactctaagcttctcaacggaat
tctcaaggacgagctaggattccagggtctcgtcatgagtgaactggcaagcccaacatgctggagctgct
accgctgttgCaggccttgacatgaccatgccccgtgacactttgttcaacaccggatacagcttctggg
gtggtaacctgacctcgtgtgtagtcaatggcactgttcccgactggcgattgacgacatggctatgag
aatcatggcagctttcttcaagggttggaagactgttgaggaccttctgacatcaacttttcttcttg
tctcgagacacttttggtacgttcaagccgctgccccagagaactgggaacagatcaacttcggagttg
atgttcgtcacgaccacagcgaacacattcgactctcggccgccaaggcgaccgtcctccttaagaactc
tggtcattgcctctgaagaagcccaagttccttgCGtCGttggcgaggacgCGggcccgaaacctgct
ggccccaaCGgtgtaaCGaccgCGgatgtaaacCGgcactctggccatgtcctggggctcaggaacag
cccagttcccttacctcgttactccgactcagcgctacagaaccaggctgtcctcgacggcactcgtc
cgagagtgtcttgCGgaacaaccagtgggaacagacacgcagctctcattagccaacctaacgtgacggct
attgtgtttgccaatgccaattccggagagggatatactgatgttgacggcaacgaaggcgatcgggaaga
atgtgacctgtggaacgaggggtgatgacctaatgaagacgtctcctcaatctgccccaacaccattgt
tgttctgcacactgttgGCCctgtcatcctgacggaatgggtatgacaaccgaacattaccGCCatagt
tgggctgggtgtacctggacaggagtcCGgcaatgctcttggtgacatcctttatggcaaaacaagcctg
gtcgtctccttccatggggtcgcacccgaaagagttacggcactgatgtcctatacgagccccaa
tggtcaggggtgctcctcaagatgatttcaCGgagggagtctttatcgactatcgtcattttgaccaggtt
tctcctagcaccgacggcagcaagtctaattgatgagtcagtcctcatcgagtttgGCCatggctgt
cctggaccaCGtttgagtactctgaactcaacattcaagctcacaacaagattcccttcgatcctcctat
tgCGgagacgattgCGctcCGgtccttggaactacagtacCGaccttgCGgattacCGtccCGgat
ggaattcgtctacatctaccagttcatctatccctgggtgaatacttcttcttcCGgaagagaggcttctg
gCGatccCGactacggaagacgCGcgaagagttcctgccccCGgagctctcgaCGgggtcagctcagcc
gCGacctccatcctctgggtgctccagggtggaaccctcatcttgggatgtgtgtgtacactgttagtgct
atcatccaacactggcaacCGccacctCGgaCGagatccCGcagctctacgttagtctcgggtggCGaga
acgagccccgtcCGcgtccttcCGgggttcgacCGaattgagaacattgCGcctggccagagtgtcagatt
cacaactgacatcactcCGcgcagctgagcaactgggacgtcgtctctcagaactgggtcattacagac
tacgagaagaccgtatatgtcgggagcagctccCGcaacctgcctctcaaggcaacctgaagtaa

FIG. 51A

65 / 138

SEQ ID NO:74

Protein sequence of Nh3A, a GH3 family β -glucosidase from *Nectria haematococca*

mrftvllaafsqlypvmvgsqadqkplqlgvnnntlahspphypspwmdpaapgweeaylkakdfvsqtl
lekvnlttgvgwmgercvgnvgslprfmgmrglcmqdgplgirlsdynsafptgitagaswsralwyqrgl
lmgtehrekgidvalgpatgplgrtptggrnwegfsvdpyvagvamaetvsgiqdggtiacakhyigneq
ehhrqapesigrgyniteslssnvddkthelylwpfadavkagvgaimcsyqqlnnsygcqnsklngi
lkdelgfggfvmstdwqaghagaatavagldmtmpgdtlfntgysfwggnltlavvngtvpdwriddmamr
imaaffkvgktvedlpdinfswsrdtfgyvqaaaagenweginfgvdvrhdhsehirlsaakgtvllkns
gslplkkpkflavgedagpnpagpnngcndrgcnngtlamswsgsgtaqfpylvtpdsalqnqavldgtry
esvlrnnqwegtrslisqpnvtaivfanansgegyidvdgnegdrknltlwnegddliknvssicpntiv
vlhtvgpviltewydnpnitaivwagvpqgesgnalvdilygktsppgrspftwgrtrksygtadvlyepnn
gggapqddftegvfidyrhfdqvspstdgsksndesspiyefghglswttfeyselniqahnkipfdppi
getiaapvlgnystdladytfdpgiriyiyqfiypwlnntsssgreasgdpdygktaeeflppgaldgsagp
rppssgapggnphlwdvlytvsaiitntgnatsdeipglyvslggenepvrvlrgfdrieniapggsvrf
tteditrrdlsnwdvvsqnwvitdyektvyvgsssrnlplk

FIG. 51B

66 / 138

SEQ ID NO:75

Nucleotide sequence of Vd3A, a GH3 family β -glucosidase from *Verticillium dahliae*

ATGAAGCTGACCCCTCGCTACTGCCTTACTGGCAGCCAGCGGGTGTGTCTCTGCGGGACAACCCAAGCTCA
AGGTACGTACTTGCCTCTTTTTTCAACAAGGAAACCAAACCCGCACCATAATGGTGATTGAGCAGTCGTGCT
TTCCTCAACCCGAATCAAACCCATGCCGTGTTTCGCGCATGCCCTTTCGATCGTCTGTTGTGTGTGAACCC
ACGCTCTTCAAGCATCGCACATAGCACCCTCCATCTTCATTTTCGAGCAATTTTCGGGCCGAGAGAGCG
GTCTTTCACCTTACCACAATCGTTCATGCCTCGTGCCCCACTGCCATGTTTTCTTCCCAGTATTCTACTTC
TGAGAGCCCTTGACCACCGTTGTGCGACATCTCGTCGCCAAGGCTCGTTGACACGGACTCTGTTTTCCCTTGG
AATTAATATTCGAAACAATGCTGACCAGCATCCTCAGCGCCAGACTAACAGCTCTAGCGAGCTCGCCTTT
TCCCCCTCCGCACTACCCTTCTCCATGGATGAACCCCCAAGCGACTGGGTGGGAGGACGCCCTACGCCCGTG
CCAGAGAGGTGGTAGAGCAGATGACTCTGCTCGAAAAGGTCAACCTGACGACAGGTGTGCGGGTAAGCTTC
ACAGACCCCGTCTTGCCATCCAAAGTCATCTGACAGAATCCTAGCTGGAGCGGTGATCTCTGCGTCGGAA
ACGTCGGCTCGATCCCCCGAATCGGCTGGAGGGGGCTTTGTTTGCAGGATGGCCACAGGGTATCCGTTT
CGCGGACTACGTCTCGTACTTCACTTCGAGCCAGACAGCCGGCGCTACCTGGGACCGAGGGCTTCTGTAC
CAGCGCGCTCACGCCATTGGCGCCGAAGGAGTAGCCAAGGGCGTCGACGTCTCCTCGGGCCCGCCATTG
GCCCTCTAGGTGCGCTTCCCGCCGGAGGTGCTAACTGGGAGGGTTTCGCCGTGGACCCTTACCTCAGTGG
CGTTGCTGTGCGCGAATCCGTCAGGGGCATCCAGGATGCTGGTGCTATTGCCAACGTCAAGCACTACATC
GTCAATGAGCAGGAACATTTCCGCCAGGCTGGCGAGGCTCAAGGTTACGGCTACGATGTGACGAGGCAT
TATCGTCGAACGTTGACGACAAGACCATGCATGAGCTTTACCTTTGGCCATTTGCAGACGCTGTCCGTGC
TGGAGCCGGCAGTGTCTGTGTTCTTATCAACAGGTGGGGGCAATACCATTCTCTCCTCTTTCCCTTGCAG
ACAGTGCCTGACCGACCTTTTTTGGCCAAAGATCAACAACAGTTACGGCTGTCAAACTCACATCTTCTG
AATGGGCTCCTCAAGGACGAACCTCGGCTTTTCAAGGGGTTTCGTCCTCAGCGATTGGCAAGCGCAGCATGCTG
GTGCTGCCACTGCCGTTGCTGGACTTGACATGGCCATGCCCGGTGACACTCGCTTCAACACCGGAGTCGC
CTTCTGGGGCGCTAACCTTACCAATGCCATTTTGAACGGCACCGTTCCCGAATATCGGCTCGATGACATG
GCCATGCGTATTATGGCGGCCCTTTTTCAAAGTTGGAAAGACCCTGGACGATGTTCTCTGACATCAACTCT
CGTCTTGGACAAAAGACACCATCGGCCCGCTGCACTGGGCGGGCCAGGACAATGTGCAGGTCATCAACCA
ACACGTTGATGTCCGTCAAGACCACGGCGCCCTCATTCGCACCATCGCTGCCCGCGGTACTGTCTTACTA
AAAAATGAGGGATCACTGCCTCTGAACAAGCCGAAATTTGTTGCTGTCATTGGTGAAGATGCTGGCCCTC
GTCTGTGTTGGTCCCAATGGCTGCCCTGATCAGGGTTGCAATAACGGCACTCTGGCTGCTGGATGGGGATC
TGGCACCGCCAGTTTCCCTTATCTCATCACTCCTGATAGTGCTCTTCAGTTTCAAGCCGTTTCGGATGGC
TCGCGATACGAAAGCATCCTCAGCAACTGGGATTATGAGCGCACAGAGGCCCTTGGTTTCCCAGGCGGATG
CTACTGCTCTGGTTTTTCGTCAATGCAAACTCTGGCGAAGGATATATCAGCGTTGATGGAAACGAAGGTGA
TCGCAAGAACCTCACTCTCTGGAATGGAGGAGACGAGCTTATTCAACGAGTCGCTGCGGCCAACAACAAC
ACCATCGTCATCATCCATTGCGTTGGTCCCCTTCTAGTCACTGACTGGTACGAGAATCCCAATATCACGG
CTATCATCTGGGCCGGCTTACCCGGACAGGAGTCTGGCAACTCTATCGCCGATATTCTTTACGGCCGCGT
GAACCCTGGTGGCAAGACACCTTTACCTGGGGTCCAACCTGTTGAGAGCTACGGCGTTGACGTCTCTGAGA
GAGCCCAACAATGGCAATGGTGCTCCCCAGAGCGATTTCGACGAGGGAGTCTTCATCGATTACCGTTGGT
TTGACCGGCAGTCGGGTGTTGATAACAATGCATCAGCGCCGAGGAACAGCAGCAGCAGCCACGCCCAAT
CTTCGAGTTTGGCTATGGCCTTTTCGTACACAACCTTTGAATTCTCCAATCTTCAGATTGAGAGGCATGAC
GTTACGATTACGTCCCTACCACTGGGCAGACGAGCCCTGCGCCGAGATTTGGTGCTAACTACAGTACGA
ACTACGACGACTACGTCTTTCCCGAGGGCGAAATCCGTTACATCTATCAACACATCTACCCATACCTCAA
TTCTCAGACCCAAAGGAGGCATTGGCTGATCCTAAATACGGCCAAACTGCAGAAGAGTCTCTCCAGAG
GGCGCTCTTGATGCCCTACCGCAGCCTAGGCTCCCAGCTTCTGGAGGGCCCCGAGGGCAACCCAATGCTTT
GGGACGTCATATTCACGGTCACCGCGACCGTGACCAACACGGGTAAGGTTGCTGGGGACGAAGTGGCACA
GCTTTACGTTTCTCTTGGTGGACCTGACGATCCGATTCGAGTCTCCGTGGGTTTCGACCGCATTACATC
GCGCCTGGAGCCTCGCAAACCTTCCGTGCGGAACTCACGCGCCGGGACCTCAGCAACTGGGATGTTGTCA
CGCAAAATTGGTTCATCAGCCAGTACGAAAAGACGGTCTTTGTGCGGGAGCTCATCCCGAAACCTCCCTCT
CAGCACTCGCCTCGAATAG

FIG. 52A

SUBSTITUTE SHEET (RULE 26)

67 / 138

SEQ ID NO:76

Protein sequence of Vd3A, a GH3 family β -glucosidase from *Verticillium dahliae*

mkltlatalaasgcvsagqpkklkhpqrqtnssselafspphypspwmnpqatgwedayararevveqmt
llekvnlttgvgwsgdlcvgnvgsiprigwrglclqdgpggirfadyvsyftssqtagatwdrgllyqra
haigaegvakgvdvvlgpaigplgrlpaggrnwegfavdpylsgvavaesvrgiqdagaiankhyivne
gehfrqageaggygydvdealssnvddkthelylwpfadavragagsvmcsyqqinnnsygcqnshllng
llkdelgfggfvlswqaghagaatavagldmampgdtrfntgvafwganltnailngtvpeyrlddmam
rimaaffkvgktlddvpdinfsswtkdtigplhwaagdnvqvinqhvdvrgdhgalirtiaargtvllkn
egslplnkpkfvavigedagprpvgpnngcpdggcnngtlaagwsgtasfpylitpdsalqfqavsdgsr
yesilsnwdyertealvsqadatalvfvnansgegyisvdgnegdrknltlwnnggdeliqrvaannnti
viihsvgpvlvtdwyenpnitaiiwaglpqgesgnsladilygrvnpggktpftwgptvesygvdlrep
nngngapqsdfdegvfidyrfdrqsgvdnnasaprnssssshapifefgyglsttfeftnlqierhdvh
dyvpttgqtspaprfganystnyddyvfpegeiriyqhiypylnssdpkealadpkygqtaeeflpega
ldaspqprlpasggpggnpmlwdviftvtatvtntgkvagdevaqlvslggpddpirvlrgfdrihiap
gasqtfraeltrrdlsnwdvvtqnwfwisqyektvfvvgsssrnlplstrle

FIG. 52B

68 / 138

SEQ ID NO:77**Nucleotide sequence of Pa3G, a GH3 family β -glucosidase from *Podospira anserina***

ATGAAACTCAATAAGCCATTCCTGGCCATTTATTTGGCTTTCAACTTGGCCGAGGCTTCGAAAACCTCCGG
ATTGCATCAGTGGTCCGCTGGCAAAGACCTTGGCATGTGATACAACGGCGTCACCTCCTGCGCGAGCAGC
TGCTCTTGTGCAGGCTTTAAATATCACGGAAAAGCTTGTGAATCTAGTGGAGTATGTCAAGTCAAGAGAA
GCTCCTTTAGGGATTTCAATTCAGCTAATCACTCCTCATAGCATGAGCCTCGGTGCAGAAAGGATCGGCC
TTCCAGCTTATGCTTGGTGGAAACGAAGCTCTTCATGGTGTGCGCGTTCGCTGGGGTCTCCTTCAATCA
GGCCGGACAAGAATTCTCACACGCTACTTCATTTGCGAATACTATTACGCTAGCAGCCGCCTTTGACAAT
GACCTGGTTTACGAGGTGGCGGATACCATCAGCACTGAAGCGCGAGCGTTTACAGCAATGCCGAGCTCGCTG
GACTGGATTACTGGACGCCTAACATCAACCCGTACAAAGATCCGAGATGGGGGAGGGGCCATGAGGTTTG
TTACCTTAGCCTTCTTTTCCGTGCCGTGCAGTTGTGTGAGAACTCAAAGACACCCGGAGAAGATCCGGTA
CACATCAAAGGCTACGTCCAAGCACTTCTCGAGGGTCTAGAAGGGAGAGACAAGATCAGAAAGGTGATTG
CCACTTGTAACACTTTGCAGCCTATGATTTGGAGAGATGGCAAGGGGCTCTTAGATACAGGTTCAATGC
TGTTGTGACCTCGCAGGATCTTTCGGAGTACTACCTCCAACCGTTTCAACAATGCGCTCGAGACAGCAAG
GTCGGGTCTTTCATGTGCTCATATAATGCGCTCAACGGAACACCGGCATGTGCAAGCACGTATTTGATGG
ACGACATCCTTCGAAAACACTGGAATTGGACCGAGCACAACAATAATATAACGAGCGACTGTAATGCTAT
TCAGGACTTCCTCCCCAACTTTCACAACCTCAGCCAACTCCAGCTCAAGCCGCCGCTGATGCTTATAAC
GCCGGTACAGACACCGTCTGTGAGGTGCCTGGATACCCCCACTCACAGATGTAATCGGAGCATACAATC
AGTCTCTGCTGTGAGAGAAATTATCGACCGAGCACTTCGCAGATTATACGAAGGCCTCATCCGAGCTGG
CTATCTCGACTCAGCCTCCCCACATCCATACACCAAAATCTCATGGTCCCAAGTAAACACCCCCAAAGCC
CAAGCCCTGGCTCTCCAGTCCGCCACCGACGGGATAGTCCTTCTCAAAAACAACGGCCTCCTTCCCTAG
ACCTCACCAACAAAACCATAGCCCTCATAGGCCACTGGGCCAATGCAACCCGCCAAATGCTAGGCGGCTA
CAGCGGTATCCCCCTTACTACGCCAACCCAATCTATGCAGCCACCCAGCTCAACGTCACTTTTCATCAC
GCCCCAGGACCGGTGAACCAGTCATCTCCCTCCACAAATGACACCTGGACCTCCCCCGCCCTCTCCGCGG
CTTCCAAATCGGATATCATCCTCTACCTCGGCGGCACCGACCTCTCCATCGCAGCCGAAGACCGAGACAG
AGACTCCATCGCCTGGCCATCCGCTCAACTTTCTTGTAACTCCCTCGCCAGATGGGAAAACCCACA
ATCGTAGCAAGACTAGGCGACCAAGTAGACGACACCCCCCTGCTCTCCAACCCAAACATCTCCTCCATCC
TATGGGTAGGCTACCCAGGCCAATCAGGCGGAACAGCCCTCTTGAACATCATCACCGGAGTCAGCTCCCC
CGCCGCTCGACTGCCCCGTACAGTCTACCCAGAACTTACACCTCCCTCATCCCCCTGACAGCCATGTCC
CTCCGCCCAACCTCCGCCCGCCAGGCCGGAATTACAGGTGGTACCCCTCCCCCGTGTCTCCCTTCGGCC
ACGGCCTCCACTACACAACCTTTACCGCCAAATTGCGCGTCTTTGAGTCCCTCACCATCAACATTGCCGA
ACTCGTTTCCAACGTAAACGAACGATACCTCGACCTCTGCCGGTTCCCGCAGGTGTCCGTCTGGGTGTGCG
AATACGGGAGAACTCAAATCTGACTATGTGCGCCCTTGTTTTTGTGAGGGGTGAGTACGGACCGGAGCCGT
ACCCGATCAAGACGCTGGTGGGGTACAAGCGGATAAGGGATATCGAGCCGGGGACTACGGGGGCGGGCC
GGTGGGGGTGGTGGTGGGGGATTTGGCTAGGGTGGATTTGGGGGGGAATAGGGTTTTGTTTCCGGGGAAAG
TATGAGTTTCTGCTGGATGTGGAGGGGGGAGGGATAGGGTTGTGATCGAGTTGGTTGGGGAGGAGGTGG
TGTTGGAGAAGTTCCTTCAGCCGCCTGCGGCGGGTTGA

FIG. 53A

69 / 138

SEQ ID NO:78

Protein sequence of Pa3G, a GH3 family β -glucosidase from *Podospira anserina*

mklmkpflaiylafnlaseasktpdcisgplaktlacdttaspparaaalvqalniteklnlveyvksre
 aplgisiglitphmslgaeriglpayawwnealhgvaaaspvsnqaggefshatsfantitlaaafdn
 dlvyevadtistearafsnalagldywtpninpykdprwgrghevscylsllfravqlrlrtqktpgedpv
 hikgyvqalleglegrdkirkvialtckhfaaydlerwqgalryrfnavvtsqdlseyylqpfqqcardsk
 vgsfmcsynalngtpacastylmddilrkhwnwtehnnyitsdcnaigdfpfnfnfsqtpaqaaadayn
 agtdtvcvpgpypltdvigaynqslseeiidralrrlyeggliragylsasphpytkiswsqvntpk
 qalalgsatdgivllknngllpldltnktialighwanatrqmlggysgipyyanpiyaatqlnvtfhh
 apgpvnqsspstndtwtpalsaaasksdilylggtdlsiaaedrdrdsiawpsaqlslltslaqmgkpt
 ivarlgdqvddtpllslpnissilwvgypgsggtallniitgvsspaarlptvypetytsliptams
 lrptsarpgrrtyrwypspvlpfghglhyttftakfgvfesltiniaelvsncneryldlcrfpqvsvwvs
 ntgelksdyvalvfvrgeygpeypiktlvgkyrirdiepgttgaapvgvvvgdlarvdlggnrvlfpqk
 yeflldveggrdrvvielvgeevvlekfpqppaag

FIG. 53B

SEQ ID NO:79

Protein sequence of Tn3B, a GH3 family β -glucosidase from *Thermotoga neapolitana*

MEKVNEILSQLTLEEKVKLVVGVGLPGLFGNPHSRVAGAAGETHPVPRVGLPAFVLADGPAGLRINPTRE
 NDENTYYTTAFPEIMLASTWNRELLEEVGKAMGEEVREYGVVDVLLAPAMNIHRNPLCGRNFEYYSEDPV
 LSGEMASSFVKGVQSQGVGACIKHFVANNQETNRMVVDITVSEALREIYLRGFEIYVKKSKPWSVMSAY
 NKLNGKYCSQNEWLLKKVLREEWGFEFVMSDWDYAGDNPVEQLKAGNDLIMPGKAYQVNTERRDEIEEIM
 EALKEGKLSEEVLDCEVRNLIKVLVNAPSFKNYRYSNKPDLKHKVAYEAGAEGVLLRNEEALPLSEN
 SKIALFGTGQIETIKGGTGSGDTHPRYASILEGIKERGLNFDEELAKTYEDYIKKMRTEEYKPRRDSW
 GTIIKPKLPENFLSEKEIHKLAKKNDVAVIVISRISGEGYDRKPKVKGDFYLSDDTDLIKTVSREFHEQG
 KKVIVLLNIGSPVEVVSWRDLVDGILLVWQAGQETGRIVADVLTGRINPSGKLPTTFPRDYSVPSWTFP
 GEPKDNPPQKVVEEDYVGYRYDTFGVEPAYEFGYGLSYTTFEYSDLNVSFDGETLRVQYRIENTGGRA
 GKEVSQVYIKAPKGKIDKPFQELKAFHKTRLLNPGEESEVVLEIPVRDLASFNGEEVVEAGEYEVVRGA
 SSRNIKLGTFVSGEERRFKP

FIG. 54

70 / 138

Alignment of amino acid sequences of Fv3C homologs. Alignment was made in Muscle (in accordance with Edgar R.C., BMC Bioinformatics, 2004, 5: 113) using default parameters.

		*	20	*	40	*					
Tn3B	:	-----	-----	-----	-----	-----	-				
Fv3G	:	---MFPSSISCLA	---ALSLMSQGLLAQSQ	PENV	-----	ITDDTY	: 34				
Fv3D	:	MASIRSVLVSGLL	---AAGVNAQAYDASDRAEDAF	SWVQP	---KNTTILGQ		: 45				
Tr3A	:	---MRYRTAAALAL	---ATGPFARADSHS	-----	-----	TSGASA	: 29				
Pa3D	:	---MALQTFLL	---AAAMLANAE	-----	-----	TTGEKV	: 24				
Te3A	:	---MRNGLLKVAAL	---AAASA	-----	-----	VNGENL	: 22				
An3A	:	---MRFTSIEAVAL	---TAVSL	-----	-----	ASADEL	: 22				
Tr3B	:	---MKTLVFAAALL	---AAVAEANPYPPP	-----	-----	HSNQ	: 28				
Nh3A	:	---MRFTVLLAAAFSGLVPMVGSQADQKPLQLG	-----	-----	-----	VNNNTL	: 35				
Gz3A	:	---MKANWLAAAVYL	---AAGTDA	-----	---VPDTLAGV	---NLVARDTL	: 35				
Fv3C	:	---MKLNWVAAAALSIGAAGTDSAVALASAVPDTLAGVKKADAQKVVT	RD	TL			: 48				
Fo3A	:	---MKLNWVAAAALSIGAAGTDGAVALASEVPDTLAGVKN	TD	QAQKVVT	RD	TL	: 48				
Pa3G	:	---MKFSVVVAAAL	---ASGALATPQYPPK	-----	-----	LIKRD	: 30				
Vd3A	:	---MKLTLATALLA	---ASGCVSAGQPKLKH	PQRQT	-----	NSSSEL	: 36				
		60	*	80	*	100					
Tn3B	:	-----	-----	MEKVNEILS	QLTLEE	KVKLVVGVGLP	: 26				
Fv3G	:	FYQGSPPVYP	-----	THTGSWAAAVAKAKNLVS	QLTLEE	KVNLTG	-GQT : 78				
Fv3D	:	YG	---HSPHY	PAN	---NATGKGWEDAF	AKAQN	FVS	QLTLEE	KADMV	TGT	--- : 88
Tr3A	:	EA	---VVP	-----	PAGTPWGTAYD	KAKAALAK	NLQDKV	GIVSGV	GWN		: 69
Pa3D	:	SR	---QAP	-----	SGAQAWAAAH	SQAAAT	LARMS	QODKIN	MVTG	I	GWD : 64
Te3A	:	AY	---SPPFY	PSPWANGQGD	---WAEAYQKAVQ	FVS	QLT	LAEKVN	LT	TGTGWE	: 69
An3A	:	AY	---SPPYY	PSPWANGQGD	---WAEAYQRAVD	IVSQMT	LAEKVN	LT	TGTGWE	: 69	
Tr3B	:	AY	---SPPFY	PSPWMDPSAPGWEQ	AYAQAKE	FVS	GLT	LLEKVN	LT	TG	VGWM : 76
Nh3A	:	AH	---SPPHY	PSPWMDPAAPGWEE	AYLKAKDF	VVS	QLT	LLEKVN	LT	TG	VGWM : 83
Gz3A	:	AH	---SPPHY	PSPWMDPNAVGWED	AYAKAKDF	VVS	QMT	LLEKVN	LT	TG	VGWQ : 83
Fv3C	:	AY	---SPPHY	PSPWMDPNAVGWEE	AYAKAKS	FVS	QLT	LMEKVN	LT	TG	VGWQ : 96
Fo3A	:	AH	---SPPHY	PSPWMDPNAIGWEE	AYAKAKN	FVS	QLT	LLEKVN	LT	TG	VGWQ : 96
Pa3G	:	AY	---SPPVY	PSPWMNPEADGWAE	AYVKARE	FVS	QMT	LLEKVN	LT	TGTGWA	: 78
Vd3A	:	AF	---SPPHY	PSPWMNPQATGWED	AYARAREV	VEQMT	LLEKVN	LT	TG	VGWS	: 84
		*	120	*	140	*					
Tn3B	:	GLFGNPHSRVAGAAGETH	HPVPRVGLPAFVLADGPAGLRINPTRENDENTY				: 76				
Fv3G	:	T	-----	TGCSGFIPGIPRVGFFGLCLADAGNGVRNTD	-----		: 110				
Fv3D	:	P	-----	GPCVGNIVAIPRLNFENGLCLHDGPLAIRVAD	-----		: 120				
Tr3A	:	G	-----	GPCVGNTSPASKISYPSLCLQDGPLGVRYST	-----		: 101				
Pa3D	:	R	-----	GPCVGNTAAISSINYPQICLQDGPLGIRFGT	-----		: 96				
Te3A	:	Q	-----	DRCVGQVGSIPRLGFFGLCMQDSPLGVRD	TD	-----	: 101				
An3A	:	L	-----	ELCVGQTGGVPRLLGIPGMCAQDSPLGVRDSD	-----		: 101				
Tr3B	:	G	-----	EKCVGNVGTVPRLGMRSLCMQDGPLGLRFNT	-----		: 108				
Nh3A	:	G	-----	ERCVGNVGSLPRFGMRGLCMQDGPLGIRLSD	-----		: 115				
Gz3A	:	G	-----	ERCVGNVGSIPLRLGMRGLCLQDGPLGIRFSD	-----		: 115				
Fv3C	:	G	-----	ERCVGNVGSIPLRLGMRGLCLQDGPLGIRLSD	-----		: 128				
Fo3A	:	G	-----	ERCVGNVGSIPLRLGMRGLCLQDGPLGIRLSD	-----		: 128				
Pa3G	:	S	-----	EQCVGQVGAIPLRLGLRSLCMHDAPLGIRGTD	-----		: 110				
Vd3A	:	G	-----	DLCVGNVGSIPLRIGWRGLCLQDGPQGIRFAD	-----		: 116				

FIG. 55A-1

SUBSTITUTE SHEET (RULE 26)

71 / 138

	160	*	180	*	200	
Tn3B	: YTTAFPVEIMLASTWNRELLEEVGKAMGEEVREYGV	DVLLAPAMN--IHRN	: 125			
Fv3G	: YVSSFPSGIHVGASWNPELTYSRSYYMGAEAKAKGVNILLGPVFGPLGRV	: 160				
Fv3D	: YASVFPAGVSAASSWDKDLLYQRLAMGQEFKAKGAHILLGPVAGPLGRS	: 170				
Tr3A	: GSTAFTPGVQAASTWDVNLIRERGQFIGEEVKASGIHVILGPVAGPLGKT	: 151				
Pa3D	: GTTAFTPGVQAASTWDVDLIRQRGAYLGAEAKGCGIHILLGPVAGALGKI	: 146				
Te3A	: YNSAFPAGVNVAATWDRNLAYRRGVAMGEEHRGKGV	DVQLGPVAGPLGRS : 151				
An3A	: YNSAFPAGVNVAATWDKNLAYLRGQAMGQEFSDKGADIQLGPAAGPLGRS	: 151				
Tr3B	: YNSAFSVGLTAAASWSRHLWVDRGTALGSEAKGKGV	DVLLGPVAGPLGRN : 158				
Nh3A	: YNSAFPTGITAGASWSRALWYQRLMGTEHREKGI	DVALGPATGPLGRT : 165				
Gz3A	: YNSAFPTGVTAGASWSKALWYERGRLLMGTEFKEKGI	DIALGPATGPLGRH : 165				
Fv3C	: YNSAFPAGTTAGASWSKSLWYERGLLMGTEFKEKGI	DIALGPATGPLGRT : 178				
Fo3A	: YNSAFPAGTTAGASWSKSLWYERGLLMGTEFKGKGI	DIALGPATGPLGRT : 178				
Pa3G	: YNSAFPSGQTAAATWDRQLMYRRGYAIGKEAKGKGI	NVILGPVAGPLGRM : 160				
Vd3A	: YVSYFTSSQTAGATWDRGLLYQRAHAIGAEGVAKGV	DVVLLGPAIGPLGRL : 166				
	*	220	*	240	*	
Tn3B	: PLCGRNFEYYSEDPVLSGEMASSFVKGVQSQGVGACIKHFVANNQETNRM	: 175				
Fv3G	: VEGGRNWEGFSNDPYLAGKLGHEAVAGIQDAGVVACGKHFLAQEQETHRL	: 210				
Fv3D	: AYSGRNWEGFSPDPYLTGIAEMEETIMGHQDAGVQATAKHFIGNEQEVMRN	: 220				
Tr3A	: PQGGRNWEGFGVDPYLTGIAMGQTINGIQSVGVQATAKHYYILNEQELNR-	: 200				
Pa3D	: PHGGRNWEGFGADPYLAGIAMKETIEGIQSAGVQANAKHYIANEQELNR-	: 195				
Te3A	: PDAGRNWEGFAPDPVLTGNMMASTIQGIQDAGVIACAKHFILYEQEHFRQ	: 201				
An3A	: PDGGRNWEGFSPDPALSGVLFAETIKGIQDAGVVATAKHYYIAYEQEHFRQ	: 201				
Tr3B	: PNGGRNVEGFGSDPYLAGLALADTVTGIQNAGTIACAKHFLLNEQEHFRQ	: 208				
Nh3A	: PTGGRNWEGFSVDPYVAGVAMAETVSGIQDGGTIACAKHYIGNEQEHHFRQ	: 215				
Gz3A	: AAGGRNWEGFTVDPYAAGHAMAETVKGIQDSGVIACAKHYIANEQEHFRQ	: 215				
Fv3C	: AAGGRNWEGFTVDPYMAGHAMA EAVKGIQDAGVIACAKHYIANEQEHFRQ	: 228				
Fo3A	: AAGGRNWEGFTVDPYMAGHAMA EAVKGIQDAGVIACAKHYIANEQEHFRQ	: 228				
Pa3G	: PAAGRNWEGFSPDPVLTGVGMAETVKGHQDAGVIACAKHFFIGNEQEHHFRQ	: 210				
Vd3A	: PAGGRNWEGFAVDPYLSGVAVAESVRGIQDAGAIANVKHYIVNEQEHHFRQ	: 216				
	260	*	280	*	300	
Tn3B	: V-----VDTIVSERALREIYLRGFEIAVKKSKPWSVMSAY	: 210				
Fv3G	: AASVTG-----ADAISSNLDDKTLHELYL-----CVMCSY	: 240				
Fv3D	: PTFVKDGYIGEVDKEALSSNMDDRTMHELYLWPFANAVHAKAS--SMMCSY	: 269				
Tr3A	: -----ETISSNPDDRTLHELYTWPFADAVQANVA--SVMCSY	: 235				
Pa3D	: -----ETMSSNVDDRTQHELYLWPFADAVHANVA--SVMCSY	: 230				
Te3A	: GAQD-----GYDISDSISANADDKTMHELYLWPFADAVRAGVG--SVMCSY	: 245				
An3A	: APEAQG---YGFNITESGSANLDDKTMHELYLWPFADAIRAGAG--AVMCSY	: 248				
Tr3B	: VGEANG---YGYPI TEALSSNVDDKTIHEVYGWPFQDAVKAGVG--SFMCSY	: 255				
Nh3A	: APESIG---RGYNITESLSSNVDDKTLHELYLWPFADAVKAGVG--AIMCSY	: 262				
Gz3A	: RGDVMS---QKFNISESLSSNLDDKTMHELYNWPFADAVRAGVG--SIMCSY	: 262				
Fv3C	: SGEVQS---RKYNISESLSSNLDDKTMHELYAWPFADAVRAGVG--SVMCSY	: 275				
Fo3A	: SGEVQS---RKYNISESLSSNLDDKTLHELYAWPFADAVRAGVG--SVMCSY	: 275				
Pa3G	: VGEARG---YGFNISETLSSNIDDKTMHELYLWPFADAVRAGAG--SFMCSY	: 257				
Vd3A	: AGEAOG---YGYDVDEALSSNVDDKTMHELYLWPFADAVRAGAG--SVMCSY	: 263				

FIG. 55A-2

SUBSTITUTE SHEET (RULE 26)

72 / 138

	*	320	*	340	*	
Tn3B	:	NKLNGKYCSQNEWLLKKVLREEWGFEGFVMSDWYAGDNPVEQLKAGNDLI	:	260		
Fv3G	:	NRANNSHACQNSKLLNGLLKGELGFQGFVSDWGAQQSGMASALAGLDVV	:	290		
Fv3D	:	QRLNGSYACQNSKVLNGILRDELGFQGYVMSDWGATHAGVAAINSGLDMD	:	319		
Tr3A	:	NKVNTTWACEDQYTLQTVLKDQLGFPGYVMTDWAQHTTVQSANSGLDMS	:	285		
Pa3D	:	NKLNGTWACENDKALNQILKKELGFQGYVLSDWNAQHSTALSANSGLDMT	:	280		
Te3A	:	NQVNNSYACSNSYTMNKLKSELGFQGFVMTDWGGHHSVGSALAGLDMS	:	295		
An3A	:	NQINNSYGCQNSYTLNKLKLAELGFQGFVMSDWAAHHAGVSGALAGLDMS	:	298		
Tr3B	:	NQVNNSYACQNSKLLINGLLKEEYGFQGFVMSDWQAQHTGVASAVAGLDMT	:	305		
Nh3A	:	QQLNNSYGCQNSKLLNGILKDELGFQGFVMSDWQAQHAGAATAVAGLDMT	:	312		
Gz3A	:	NQVNNSYACQNSKLLNGILKDEMGFQGFVMSDWQAQHTGAASAVAGLDMT	:	312		
Fv3C	:	NQINNSYGCQNSKLLNGILKDEMGFQGFVMSDWAAQHTGAASAVAGLDMS	:	325		
Fo3A	:	NQINNSYGCQNSKLLNGILKDEMGFQGFVMSDWAAQHTGAASAVAGLDMS	:	325		
Pa3G	:	QQVNNSYGCQNSKLMNGLLKDELGFQGFVLSDWQAQHTGAAAAAAGLDMS	:	307		
Vd3A	:	QQINNSYGCQNSHLLNGLLKDELGFQGFVLSDWQAQHAGAATAVAGLDMA	:	313		
		360	*	380	*	400
Tn3B	:	MPGKAYQVNTERRDEI--EEIMEALKEGKLSEEVLDCEVRNLIKVL----	:	304		
Fv3G	:	MPSS-----ILWGANLTLGVNNGTIPESQVDNMVTRLLATWYQLN	:	330		
Fv3D	:	MPGGIGAYGTYFTKSFFGGNLTRA VTNGTLDETR VNDMITRIMTPYFWLG	:	369		
Tr3A	:	MPG----TDFNGNRLWGPALTN AVNSNQVPTSRVDDMVTRILAAWYLTG	:	331		
Pa3D	:	MPG----TDFNGRNVYWGPQLNNAVNAGQVQSRRLDDMCKRILAGWYLLG	:	326		
Te3A	:	MPGD---IAFDSGTSFWGTNLTVAVLNGSIPEWRVDDMAVRIMSAYYKVG	:	342		
An3A	:	MPGD---VDYDSGTSYWGNTLTISVLNGTVPQWRVDDMAVRIMAAYYKVG	:	345		
Tr3B	:	MPGD---TAFNTGASYFGSNLTLAVLNGTVPPEWRIDDMVMRIMAPFFKVG	:	352		
Nh3A	:	MPGD---TLENTGYSEFWGGNLTLAVVNGTVPDWRIDDMAMRIMAAFFKVG	:	359		
Gz3A	:	MPGD---TEFNTGFSFWGGNLTLAVINGTVPAPWRIDDMATRIMAAFFKVG	:	359		
Fv3C	:	MPGD---TAFDSGYSEFWGGNLTLAVINGTVPAPWRVDDMALRIMSAAFFKVG	:	372		
Fo3A	:	MPGD---TAFDSGYSEFWGGNLTLAVINGTVPAPWRVDDMALRIMSAAFFKVG	:	372		
Pa3G	:	MPGD---TEFNTGVSEFWGTNLTVAVLNGTVPAYRIDDMAMRIMAAFFKVE	:	354		
Vd3A	:	MPGD---TRFNTGVAFWGANLTNAI LNGTVPPEYRLDDMAMRIMAAFFKVG	:	360		
		*	420	*	440	*
Tn3B	:	-----VNAPSEFKNYRYSNKPDL-----EKHAKVAYE	:	330		
Fv3G	:	QDQ-----DTEAPGHGLAAKL-WEPHPVVDARNASSKPTIWD	:	366		
Fv3D	:	QDK-DYPSVDPSSGDLNTFSPKSSWFREF-NLTGERSRDVRGNHGD LIRK	:	417		
Tr3A	:	QDQAGYPSFNISR-----NVQGNHKTNVRA	:	356		
Pa3D	:	QNQ-GYPAINIRA-----NVQGNHKENVRA	:	350		
Te3A	:	RDRYSVP-INFDSWTLDTYGPEHYAVGQG-QTKINEHVDVRGNHAEI IHE	:	390		
An3A	:	RDRLWTP-PNFSSWTRDEYGFKYYYVSEGPYEKVNQFVNVQRNHSELIRR	:	394		
Tr3B	:	KTVDSLIDTNFDSWTNGEYGYVQAAVNEN-WEKVNYGVDVRANHANHIRE	:	401		
Nh3A	:	KTVEDLPDINFSSWSRDTFGYVQAAAQEN-WEQINFGVDVRHDHSEHIRL	:	408		
Gz3A	:	RSVEEEP DINFSAWTRDEYGFVQTYAQEN-REKVNFAVNVQHDH KRHIRE	:	408		
Fv3C	:	KTIEDLPDINFSSWTRDTFGFVHTFAQEN-REQVNFGVNVQHDH KSHIRE	:	421		
Fo3A	:	KTVEDLPDINFSSWTRDTFGFVQTF AQEN-REQVNFGVNVQHDHKNHIRE	:	421		
Pa3G	:	KSIELDP-INFSEWSLDTYGPIHWAAGEG-HQQINYHVDVRADHANLIRE	:	402		
Vd3A	:	KTLDDVPDINFSSWTKDTIGPLHWA AQDN-VQVINQHVDVRQDHGALIRT	:	409		

FIG. 55A-3

SUBSTITUTE SHEET (RULE 26)

73 / 138

	460	*	480	*	500	
Tn3B	: AGAEGVLLRNEE-ALPLS-ENSKIALFG-----TGQ-----					: 360
Fv3G	: GAVEGHVLVKNTNNALPFKPNMKLVSLFGYSHKAPDKNIPDPAQGMFSAW					: 416
Fv3D	: HGAESTVLLKNEKNALPLK-KPKSIAVFG-----NDA--GDITEGFYNQN					: 459
Tr3A	: IARDGIVLLKNDANILPLK-KPASIAVVG-----SAAIIGNHARNSPSCN					: 400
Pa3D	: VARDGIVLLKNDG-ILPLS-KPRKIAVVG-----SHS--VNNPQGINACV					: 391
Te3A	: IGAASAVLLKNKG-GLPLTGTERFVGVFG-----KDA--GSNPWGVNGCS					: 432
An3A	: IGADSTVLLKNDG-ALPLTGKERLVALIG-----EDA--GSNPFYGANGCS					: 436
Tr3B	: VGAKGTVIFKNNG-ILPLK-KPKFLTIVIG-----EDA--GGNPAGPNGCG					: 442
Nh3A	: SAAKGTVLLKNSG-SLPLK-KPKFLAVVG-----EDA--GPNPAGPNGCN					: 449
Gz3A	: AGAKGSVVLKNTG-SLPLK-KPQFLAVIG-----EDA--GSNPAGPNGCA					: 449
Fv3C	: AAAGSVVLKNTG-SLPLK-NPKFLAVIG-----EDA--GPNPAGPNGCG					: 462
Fo3A	: SAAKGSVILKNTG-SLPLN-NPKFLAVIG-----EDA--GPNPAGPNGCG					: 462
Pa3G	: IAAKGTVLLKNTG-SLPLN-KPKFVAVIG-----EDA--GPNPENGNSCA					: 443
Vd3A	: IAARGTVLLKNEG-SLPLN-KPKFVAVIG-----EDA--GPRPVGPNGCP					: 450
		*	520	*	540	*
Tn3B	: -----IETIKGGTGS GDTHPRYAISI					: 381
Fv3G	: SIGAQSANITELNLGFLGNLSLTYSIAIPNGTIIISGGGSGASAWTLFSSP					: 466
Fv3D	: D--YE-----FGTLVAGGGSGTGRLTYLVSP					: 483
Tr3A	: DKGCD-----DGALGMGWGSGAVNYPYFVAP					: 426
Pa3D	: DKGCN-----VGTLGMGWGSGSVNYPYLVSP					: 417
Te3A	: DRGCD-----NGTLAMGWGSGTANFPYLVTP					: 458
An3A	: DRGCD-----NGTLAMGWGSGTANFPYLVTP					: 462
Tr3B	: DRGCD-----DGTLAMEWGSGTTNFPYLVTP					: 468
Nh3A	: DRGCN-----NGTLAMSWGSGTAQFPYLVTP					: 475
Gz3A	: DRGCD-----NGTLAMAWGSGTSQFPYLVTP					: 475
Fv3C	: DRGCD-----NGTLAMAWGSGTSQFPYLITP					: 488
Fo3A	: DRGCD-----NGTLAMAWGSGTSQFPYLITP					: 488
Pa3G	: DRGCN-----NGTLAMGWGSGTANFPYLITP					: 469
Vd3A	: DQGCN-----NGTLAAGWGSGTASFYPYLITP					: 476
		*	560	*	580	*
Tn3B	: LEGIKERGLNFDEELAKTYEDYIKKMRETEEYKPRRDSWGTIIKPKLPEN					: 431
Fv3G	: FDAFVSRAC-----KEGTALF-----W					: 483
Fv3D	: LAAINARAC-----QDGTILVQQWMNNT					: 505
Tr3A	: YDAINTRAS-----SQGTQVT--LSNT					: 446
Pa3D	: YDALRTRAQ-----ADGTQIS--LHNT					: 437
Te3A	: EQAIQREVL-----SRNGTFTGITDNG					: 480
An3A	: EQAISNEVL-----KNKNGVFTATDNW					: 484
Tr3B	: DAALQSQAL-----QDGTRYESILSNY					: 490
Nh3A	: DSALQNQAV-----LDGTRYESVLRNN					: 497
Gz3A	: DQGISLQAI-----QDGTRYESILNNN					: 497
Fv3C	: DQGLSNRAT-----QDGTRYESILTNN					: 510
Fo3A	: DQGLQNRAA-----QDGTRYESILTNN					: 510
Pa3G	: DAALQAQAI-----KDGSRYESILTNY					: 491
Vd3A	: DSALQFQAV-----SDGSRYESILSNW					: 498

FIG. 55A-4

SUBSTITUTE SHEET (RULE 26)

74 / 138

	*	620	*	640	*	
Tn3B	:	FLSEKEIHK---	LAKKNDVAVIVISRISGEGY-----	DRKPVKGDFFYLSDDDE	:	475
Fv3G	:	DFESWDPY----	VNPTSEACIVAGNAWASEGW-----	DRPATY--DAYT	:	521
Fv3D	:	LIATTNVTDLWIPATPDVCLVFLKTWAAE-----	AADREHLSVDWDG	:	547	
Tr3A	:	DNTSSGAS----	AARGKDVAIVFITADSGEGYITVEGNAGDRNNLDPWHNG	:	493	
Pa3D	:	DSTNGVSN----	VVSDADAVVVITADSGEGYITVEGHAGDRSHLDPWHNG	:	484	
Te3A	:	ALAEMAA-----	AASQADTCLVFANADSGEGYITVDGNEGDRKNLTLWQGA	:	526	
An3A	:	AIDQIEA-----	LAKTASVSLVFNADSGEGYINVDGNLGDRRNLTLWRNG	:	530	
Tr3B	:	AISQTQAL----	VSQPDIAIAIVFANSDSGEGYINVDGNEGDRKNLTLWKNG	:	537	
Nh3A	:	QWEQTRSL----	ISQPNVTAIVFANANSGEGYIDVDGNEGDRKNLTLWNEG	:	544	
Gz3A	:	QWPQTQAL----	VSQPNVTAIVFANADSGEGYIEVDGNYGDRKNLTLWKQG	:	544	
Fv3C	:	EWASVQAL----	VSQPNVTAIVFANADSGEGYIEVDGNFGRKNLTLWQQG	:	557	
Fo3A	:	EWAQTQAL----	VSQPNVTAIVFANADSGEGYIEVDGNFGRKNLTLWQQG	:	557	
Pa3G	:	AASQTRAL----	VSQDNVTAIVFVNADSGEGYINFEGNMGDRNNLTLWRGG	:	538	
Vd3A	:	DYERTEAL----	VSQADATALVFNANSGEGYISVDGNEGDRKNLTLWNGG	:	545	

	660	*	680	*	700	
Tn3B	:	TDLIKTVSREFHEQGKKVIVLLNIGSPVEVVSWRDLVDGILLVWQA--	GQ	:	523	
Fv3G	:	DELINNVA-----	DKCANTIVVLHNAGTRLVDGFFGHPNVTAIYYAHLPGQ	:	567	
Fv3D	:	NDVVESVA-----	KYCNNTVVVTHSSGINTLP--WADHPNVTAILAAHFPGQ	:	592	
Tr3A	:	NALVQAVA-----	GANSNVIVVHVSVAIILEQILALPQVKAVVWAGLPSQ	:	539	
Pa3D	:	NQLVQAAA-----	AANKNVIVVHVSVAITLETILNTNGVRAIVWAGLPGQ	:	530	
Te3A	:	DQVIHNVS-----	ANCNNTVVVLHTVGPVLIDDWYDHPNVTAILWAGLPGQ	:	572	
An3A	:	DNVIKAAA-----	SNCNNTIVIIHVSVPVLVNEWYDNPNTAILWGGLPGQ	:	576	
Tr3B	:	DDLIKTVA-----	AVNPKTIVVIHSTGPVILKDYANHPNISAILWAGAPGQ	:	583	
Nh3A	:	DDLIKNVS-----	SICPNTIVVLHTVGPVILTEWYDNPNTAILWAGVPGQ	:	590	
Gz3A	:	DELIKNVS-----	AICPNTIVVLHTVGPVLLTEWHNNPNNTAILWAGVPGQ	:	590	
Fv3C	:	DELIKNVS-----	SICPNTIVVLHTVGPVLLADYEKNPNNTAILWAGLPGQ	:	603	
Fo3A	:	DELIKNVS-----	SICPNTIVVLHTVGPVLLADYEKNPNNTAILWAGLPGQ	:	603	
Pa3G	:	DDLKVNVS-----	SWCSNTIVVIHSTGPVLISEWYDSPNTAILWAGLPGQ	:	584	
Vd3A	:	DELIQRVA-----	AANNNTIVIIHVSVPVLVTDWYENPNNTAILWAGLPGQ	:	591	

	*	720	*	740	*	
Tn3B	:	ETGRIVADVLTGRINPSGKLPTTFPRDYS-----	VPSWTFPG--EPKDN	:	566	
Fv3G	:	DSGDALVSLLYGDENPSGRLPYTVARNETDYGHLLKPDLTAPN--	QYQHF	:	616	
Fv3D	:	ESGNSLVDLLYGDVNPSGRLPYTIAFNGTDYN-----	APPTTAVNTTGKED	:	638	
Tr3A	:	ESGNALVDVLWGDVSPSGKLVYTIKSPNDYN-----	TRIVS-----GGSD--	:	580	
Pa3D	:	ENGNALVDVLYGLVSPSGKLPYTIGKRESYG-----	TAVV-----RGDD--	:	570	
Te3A	:	ESGNSLVDVLYGRVNP--GKTPFTWGRARDDYG-----	APLIVKPN--NGKGA	:	616	
An3A	:	ESGNSLADVLYGRVNP--GAKSPFTWGKTREAYQ-----	DYLYTEPN--NGNGA	:	621	
Tr3B	:	ESGNSLVDILYGKQSP--GRTPFTWGPSLESYG-----	VSVMTTPN--NGNGA	:	627	
Nh3A	:	ESGNALVDILYGKTSP--GRSPFTWGRTRKSYG-----	TDVLYEPN--NGQGA	:	634	
Gz3A	:	ESGNALVDILYGKTSP--GRSPFTWGRTYDSYG-----	TKVLYKAN--NGEGA	:	634	
Fv3C	:	ESGNALVDILYGKVSP--GRSPFTWGRTRRESYG-----	TEVLYEAN--NGRGA	:	647	
Fo3A	:	ESGNALVDILYGKVSP--GRSPFTWGRTRRESYG-----	TEVLYEAN--NGRGA	:	647	
Pa3G	:	ESGNSITDVLYGKVNP--GKSPFTWGATREGYG-----	ADVLYTPN--NGEGA	:	629	
Vd3A	:	ESGNSIADVLYGKVNP--GKTPFTWGPTVESYG-----	VDVLEPN--NGNGA	:	636	

FIG. 55A-5

SUBSTITUTE SHEET (RULE 26)

75 / 138

	760	*	780	*	800	
Tn3B	: P Q K V V Y E E D I Y V G Y R Y Y D T	-----	-----	-----	FGVEPAYEFGYGLSYT	: 601
Fv3G	: P Q S	-----	-----	-----	KNITPRFEFGFGLSYT	: 650
Fv3D	: W Q S	-----	-----	-----	HNISVRYEFGFGLSYS	: 672
Tr3A	: -----	SFSEGLFIDYKHFDD	-----	-----	ANITPRYEFgyGLSYT	: 611
Pa3D	: -----	NFREGLFVDYRHFDN	-----	-----	ARIEPRYEFgyGLSYT	: 601
Te3A	: P Q Q	-----	-----	-----	YNITPIYEFgyGLSYT	: 650
An3A	: P Q E	-----	-----	-----	RNETPIYEFgyGLSYT	: 655
Tr3B	: P Q D	-----	-----	-----	SDKAPTYEFgyGLSWS	: 669
Nh3A	: P Q D	-----	-----	-----	DESSPIYEFghGLSWT	: 679
Gz3A	: P Q E	-----	-----	-----	DSSAPLYEFgyGLSWT	: 680
Fv3C	: P Q D	-----	-----	-----	NTAAPLYEFghGLSWS	: 693
Fo3A	: P Q D	-----	-----	-----	NTAAPLYEFghGLSWT	: 693
Pa3G	: P Q Q	-----	-----	-----	ANTSVIYEFghGLSYT	: 663
Vd3A	: P Q S	-----	-----	-----	DFDEGVFIDYRWFDNRQSGVDNNASAPRNSSSSHAPIFEFGYGLSYT	: 685
		*	820	*	840	*
Tn3B	: T F E Y S D L N V S F	-----	-----	-----	-----	: 612
Fv3G	: T F E Y A S L Q I S K	-----	-----	-----	-----	: 661
Fv3D	: T F E I S D I S A E P	-----	-----	-----	-----	: 683
Tr3A	: K F N Y S R L S V L S T A K S	-----	-----	-----	G-----	: 627
Pa3D	: N F T F S D I K I T S N V K P	-----	-----	-----	G-----	: 617
Te3A	: T F E F S Q L N V Q P I N A P P Y T P A S G F T K A A Q S F G	-----	-----	-----	QPSNASDNLYPSPD-IER	: 697
An3A	: T F N Y S N L Q V E V L S A P A Y E P A S G E T E A A P T F G	-----	-----	-----	EVGNASDYLYPDGL-QR	: 702
Tr3B	: T F K F S N L H I Q K N N V G P M S P P N G K T I A A P S L G	-----	-----	-----	SFSKNLKDYGFPKN-VRR	: 717
Nh3A	: T F E Y S E L N I Q A H N K I P F D P P I G E T I A A P V L G	-----	-----	-----	NYSTDLDADYTFPDG-IRY	: 727
Gz3A	: T F E Y S D L K V E S V S N A S Y S P S V G N T I P A P T Y G	-----	-----	-----	NFSKNLDDYTFPSG-VRY	: 728
Fv3C	: T F E Y S D L N I Q K N V E N P Y S P P A G Q T I P A P T F G	-----	-----	-----	NFSKNLNDYVFPKG-VRY	: 741
Fo3A	: T F E Y S D L N I Q K N V N S T Y S P P A G Q T I P A P T F G	-----	-----	-----	NFSKNLNDYVFPKG-VRY	: 741
Pa3G	: T F E Y S N I Q V T K K N A G P Y K P T T G Q T A P A P T F G	-----	-----	-----	NFSTDLSDYLFDEEFPY	: 712
Vd3A	: T F E F S N L Q I E R H D V H D Y V P T T G Q T S P A P R F G A N Y S T N Y D D Y V F P E G E I R Y	-----	-----	-----	-----	: 735
		860	*	880	*	900
Tn3B	: -----	-----	-----	-----	DGET-----	: 616
Fv3G	: -----	-----	-----	-----	SQAQTPEYPAG	: 672
Fv3D	: -----	-----	-----	-----	LASDITSQPED	: 694
Tr3A	: -----	-----	-----	-----	PAT	: 630
Pa3D	: -----	-----	-----	-----	PAT	: 620
Te3A	: V P L Y I Y P W L N S T D L	-----	-----	-----	KASAND--PDYGLPTEKYVPPNATNGDPQPIDPAG	: 744
An3A	: I T K F I Y P W L N S T D L	-----	-----	-----	EASSGD--ASYGQDASDYLPEGATDGSAQPILPAG	: 749
Tr3B	: I K E F I Y P Y L S T T T S G K E A S G D	-----	-----	-----	AHYGQTAKFLPAGALDGSPQPRSAAS	: 765
Nh3A	: I Y Q F I Y P W L N T S S S G R E A S G D	-----	-----	-----	PDYGKTAEEFLPPGALDGSAQPRPPSS	: 775
Gz3A	: L Y K F I Y P Y L N T S S S A E K A S G D V K G R F G E T G D E F L P P N A L N G S S Q P R L P S S	-----	-----	-----	-----	: 778
Fv3C	: I Y K F I Y P F L N T S S S A S E A S N D	-----	-----	-----	GGQFGKTAEEFLPPNALNGSAQPRLPAS	: 790
Fo3A	: I Y K F I Y P F L N T S S S A S E A S N D	-----	-----	-----	GGQFGKTAEEFLPPNALNGSAQPRLPSS	: 790
Pa3G	: V Y Q Y I Y P Y L N T T D P	-----	-----	-----	RNASGD--PHFGQTAEFFMPPHAIDDSPQPLLPSS	: 759
Vd3A	: I Y Q H I Y P Y L N S S D P	-----	-----	-----	KEALAD--PKYGQTAEFFLPEGALDASPQPRLPAS	: 782

FIG. 55A-6

SUBSTITUTE SHEET (RULE 26)

76 / 138

	*	920	*	940	*
Tn3B	:	-----LRVQYRIENTGGRAGKEVSQVYIKAPKGKI-DKPF	:	650	
Fv3G	:	A-LTEGGRSDLWDVVATVTASVRNTGSVDGKEVAQLYV----	:	GVP-GGPM	: 716
Fv3D	:	LPVQPGGNPALWETVYNVTVSVSNTGKVDGATVPQLYVTFPDSAPAGTPP	:	744	
Tr3A	:	GAVVPGGPNPDLFQNVATVTVDIANSQVGTGAEVAQLYITYPSSAP-RTPP	:	679	
Pa3D	:	GQTIPGGPADLWEDVATVTATITNSGAVEGAEVAQLYIGLPSSAP-ASPP	:	669	
Te3A	:	G--APGGNPSLYEPVARVTTIITNTGKVTGDEVPQLYVSL--GGP-DDAP	:	789	
An3A	:	G--GAGGNPRLYDELIRVSVTIKNTGKVAGDEVPQLYVSL--GGP-NEPK	:	794	
Tr3B	:	G--EPGGNRQLYDILYTVTATITNTGSVMDDAVPQLYLSH--GGP-NEPP	:	810	
Nh3A	:	G--APGGNPHLWDVLYTVSAIITNTGNATSDEIPQLYVSL--GGE-NEPV	:	820	
Gz3A	:	G--APGGNPQLWDIMYTVTATITNTGDATSDEVPQLYVSL--GGE-GEPV	:	823	
Fv3C	:	G--APGGNPQLWDILYTVTATITNTGNATSDEIPQLYVSL--GGE-NEPI	:	835	
Fo3A	:	G--APGGNPQLWDILYTVTATITNTGNATSDEIPQLYVSL--GGE-NEPV	:	835	
Pa3G	:	GKNSPGGNRALYDILYEVTADITNTGEIVGDEVVQLYVSL--GGP-DDPK	:	806	
Vd3A	:	G--GPGGNPMLWDVIFTVTATVTNTGKVAGDEVAQLYVSL--GGP-DDPI	:	827	
		960	*	980	*
				1000	
Tn3B	:	QELKAFHKTRLLNPGESEEVVLEIPVRDLASFN--GEEWVVEAGEYEVRV	:	698	
Fv3G	:	RQLRGFTKP-AIKAGETATVTFELTRRDLNVWDVNAQEWQLQQGNYAIYV	:	765	
Fv3D	:	KQLRGFDKV-FLEAGESKSVSFELMRRDLSYWDIISQKWLIPEGEFTIRV	:	793	
Tr3A	:	KQLRGFAKL-NLTPGQSGTATFNIRRRDLSYWDTASQKWVVPSSGSGFISV	:	728	
Pa3D	:	KQLRGFSKL-KLAPGASGTATFNLRRRDLSYWDTRLQNWVVPSSGNFVSV	:	718	
Te3A	:	KVLRGFDRI-TLAPGQQYLWTTTLTRRDLSNWDVPTQNWVVTNYTKTIYV	:	838	
An3A	:	IVLRQFERI-TLQPSKETQWSTTLTRRDLANWNVETQDWEITSYPKMFVA	:	843	
Tr3B	:	KVLRGFDRIERIAAPGQSVTFKADLTRRDLSNWDTKKQQWVITDYPKTVYV	:	860	
Nh3A	:	RVLRGFDRIENIAPGQSVRFTTDITRRDLSNWDVVSQNWVITDYEKTVYV	:	870	
Gz3A	:	RVLRGFERLENIAPGESATFTAQLTRRDLSNWDVNVQNWVITDHAKKIWV	:	873	
Fv3C	:	RVLRGFDRIENIAPGQSAIFNAQLTRRDLSNWDVNAQNWVITDHPKTVWV	:	885	
Fo3A	:	RVLRGFDRIENIAPGQSAIFNAQLTRRDLSNWDVDAQNWVITDHPKTVWV	:	885	
Pa3G	:	VVLRDFGKL-RIEPGQTAKFRGLLTRRDLSNWDVVSQDWVISEHTKTVFV	:	855	
Vd3A	:	RVLRGFDRI-HIAPGASQTFRAELTRRDLSNWDVVTQNWVIFISQYEKTVFV	:	876	
	*	1020			
Tn3B	:	GASSRNIKLKGTFSVGEERRFKP	:	721	
Fv3G	:	GRSSRDLPLQSTLSI-----	:	780	
Fv3D	:	GFSSRDLKEETKVTVEA-----	:	811	
Tr3A	:	GASSRDIRLTSTLSVA-----	:	744	
Pa3D	:	GASSRDIRLTGTITA-----	:	733	
Te3A	:	GNSSRNPLPLQAPLKPYPGI-----	:	857	
An3A	:	GSSSRKLPLRASLPVH-----	:	860	
Tr3B	:	GSSSRDLPLSARLP-----	:	874	
Nh3A	:	GSSSRNLPLKATLK-----	:	884	
Gz3A	:	GSSSRNLPLSADL-----	:	886	
Fv3C	:	GSSSRKLPLSAKLE-----	:	899	
Fo3A	:	GSSSRKLPLSAKLE-----	:	899	
Pa3G	:	GKSSRDLGLSAVLE-----	:	869	
Vd3A	:	GSSSRNLPLSTRLE-----	:	890	

FIG. 55A-7

77 / 138

Alignment of *T. reesei* Eg4 with TrEGb (or TrEG7, or *T. reesei* Eg7) (SEQ ID NO:80) and TtEG from *Thielavia terrestris* (SEQ ID NO:81). Alignment was made in Muscle (Edgar R.C. BMC Bioinformatics, 2004, 5: 113) using default parameters.

```

          *          20          *          40          *
TtEG   : ---MLANGAIVFLAAALG-VSGHYTWPRVNDGADWQQVRKADNWQ----- : 41
TrEg4  : MIQKLSNLLVTALAVATG-VVGHGHINDIVINGVWYQAYDPTTFPYESNP : 49
TrEGb  : ---MKSCAILAALGCLAGSVLGHGQVQNFETINGQYNQGFILDYYYQKQNT : 47

          60          *          80          *          100
TtEG   : -----DNGYVG---DVTSPQIRCFQATPSPAPSVLNTTAGST : 75
TrEg4  : ---PIVVGWTAADLDNGFVSPDAYQNPDIICHK--NATNAKGHASVKAGDT : 95
TrEGb  : GHFPNVAGWYAEDLDLGFISPDQYTTTPDIVCHK--NAAPGAISATAAAGSN : 96

          *          120          *          140          *
TtEG   : VTY-WANPDVY-HP-GPVQFYMARVPDGEDINSWNGDGAVWFKVYEDHPT : 122
TrEg4  : ILFQWV-PVPWPHP-GPIVDYLANC---NGDCETVDKTTLEFFKIDGVGLL : 141
TrEGb  : IVFQWG-PGVWPHPYGPIVTYVVEC---SGSCTTVNKNLNRWVKIQEAGIN : 143

          160          *          180          *          200
TtEG   : FGAQL-TWPS-----TGKSSFAVPIPPCIKSGYLLRAEQIGLHVAQSVGG : 167
TrEg4  : SGGDPGTWASDVLISNNNTWVVKIPDNLAPGNYVLRHEIIALHSAGQANG : 191
TrEGb  : YNTQV---WAQQDLINQGNKWTVKIPSSLRPGNYVFRHELLAAHGASSANG : 191

          *          220          *          240          *
TtEG   : AQFYISCAQLSVTGGGSTEPPNKVAFPGAYSATDPGILINIYYPVPTS YQ : 217
TrEg4  : AQNYPQCFNIAVSGSGSLQ-PSGVLGTDLYHATDPGVLINI-YTSPLNYI : 239
TrEGb  : MQNYPQCVNIAVTGSGTKALPAGTPATQLYKPTDPGILFNP-YTTITSYT : 240

          260          *          280          *          300
TtEG   : NPGPAVFSC----- : 226
TrEg4  : IPGPTVVSGLPTSVAQGSSAATATASATVPGGGSGPTSRTTTTARTTQAS : 289
TrEGb  : IPGPALWQG----- : 249

          *          320          *          340          *
TtEG   : ----- : --
TrEg4  : SRPSSTPPATTSAFAGGPTQTLYGQCGSGSGPTRCAPPATCSTLNPYY : 339
TrEGb  : ----- : --

TtEG   : ----- : --
TrEg4  : AQCLN : 344
TrEGb  : ----- : --

```

FIG. 56A

78 / 138

SEQ ID NO:80**Protein sequence of TrEGb, or Eg7, from *T. reesei***

MKSCAILAALGCLAGSVLGHGQVQNFTINGQYNQGFILDYYYQKQNTGHFPPNVAGWYAED
LDLGFISPDQYTTDPDIVCHKNAAPGAISATAAAGSNIVFQWGPVWPHPYGPIVTYVVEC
SGSCTTVNKNLNRWVKIQEAGINYNTQVWAQQDLINQGNKWTVKIPSSLRPGNYVFRHEL
LAAHGASSANGMQNYPQCVNIAVTGSGTKALPAGTPATQLYKPTDPGILFNPYTTITSYT
IPGPALWQG

SEQ ID NO:81**Protein sequence of TtEG, from *Thielavia terrestris***

MLANGAIVFLAAALGVSGHYTWPRVNDGADWQQVRKADNWQDNGYVGDTVSPQIRCFQATPSPAPSVLNT
TAGSTVTYWANPDVYHPGPVQFYMARVPDGEDINSWNGDGAVWFKVYEDHPTFQAQLTWPSTGKSSFAVP
IPPCIKSGYYLLRAEQIGLHVAQSVGGAQFYISCAQLSVTGGGSTEPPNKVAFPGAYSATDPGILINIYY
PVPTSYQNPGPAVFSC

FIG. 56B

79 / 138

Partial amino acid alignment of the CBM1 domains of Eg4 with Tr6A from *T. reesei* (SEQ ID NO:82): and Tr7A from *T. reesei* (SEQ ID NO:83). Partial amino acid alignment was made in Muscle (Edgar R.C. (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5: 113) using default parameters.

```

                *           20           *
Tr6A-CBM1   : QACSSVWGQCGGQNSGPTCCASGSTCVYSNDYYSQCLP : 63
TrEg4-CBM1  : -PTQTLYGQCGSGYSGPTRCAPPATCSTLNPYYAQCLN : 343
Tr7A-CBM1   : -PTQSHYGQCGGIGYSGPTVCASGTTTCQVLNPPYYSQCL- : 513

```

SEQ ID NO:82**Protein sequence of Tr6A from *T. reesei***

MIVGILTTLATLATLAASVPLEERQACSSVWGQCGGQNSGPTCCASGSTCVYSNDYYSQ
CLPGAASSSSSTRAASTTSRVSPPTTSRSSSATPPPGSTTTTRVPPVSGGTATYSGNPFVGV
TPWANAYYASEVSSLAIPSLTGAMATAAAVAKVPSFMWLDTLDKTPLMEQTLADIRTAN
KNGGNYAGQFVVYDLPRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLL
VIEPDSLNLVLTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ
DPAAQLFANVYKNASSPRALRGLATNVANYNGWNITSPPSYTQGNVYNEKLYIHAIGPL
LANHGWSNAFFITDQGRSGKQPTGQQQWGDWCNVIGTGFGIRPSANTGDSLLDSFVWVKP
GGECDGTSDSSAPRFDSDHLCALPDALQAPAQAGAWFQAYFVQLLTNANPSFL

SEQ ID NO:83**Protein sequence of Tr7A from *T. reesei***

MYRKLAVISAFATARAQSACTLQSETHPPLTWQKCSSGGTCTQQTGSVVIDANWRWTHA
TNSSTNCYDGNWTSSSTLCPDNETCAKNCCLDGAAYASTYGVTTSGNSLSIGFVTQSAQKN
VGARLYLMASDTTYQEFTLLGNEFSFDVDVSQLPCGLNGALYFVSMDADGGVSKYPTNTA
GAKYGTGYCDSQCPRDLKFINGQANVEGWEPSSNNANTGIGGHGSCCSEMDIWEANSISE
ALTPHPCTTVGQEICEGDGCGGTYSNRYGGTCDPDGCDWNPYRLGNTSFYGPSSFTLD
TTKKLTVVTQFETSGAINRYVQNGVTFQQPNAELGSYSGNELNDDYCTAEAEFGSSSF
SDKGGLTQFKKATSGGMVLVMSLWDDYANMLWLDSTYPTNETSSTPGAVRGSCSTSSGV
PAQVESQSPNAKVTFSTNIKFGPIGSTGNPSGGNPPGGNRGTTTTTRRPATTTGSSPGPTQS
HYGQCGGIGYSGPTVCASGTTTCQVLNPPYYSQCL

SEQ ID NO:206**Protein sequence of a GH61/endoglucanase from *T. aurantiacus***

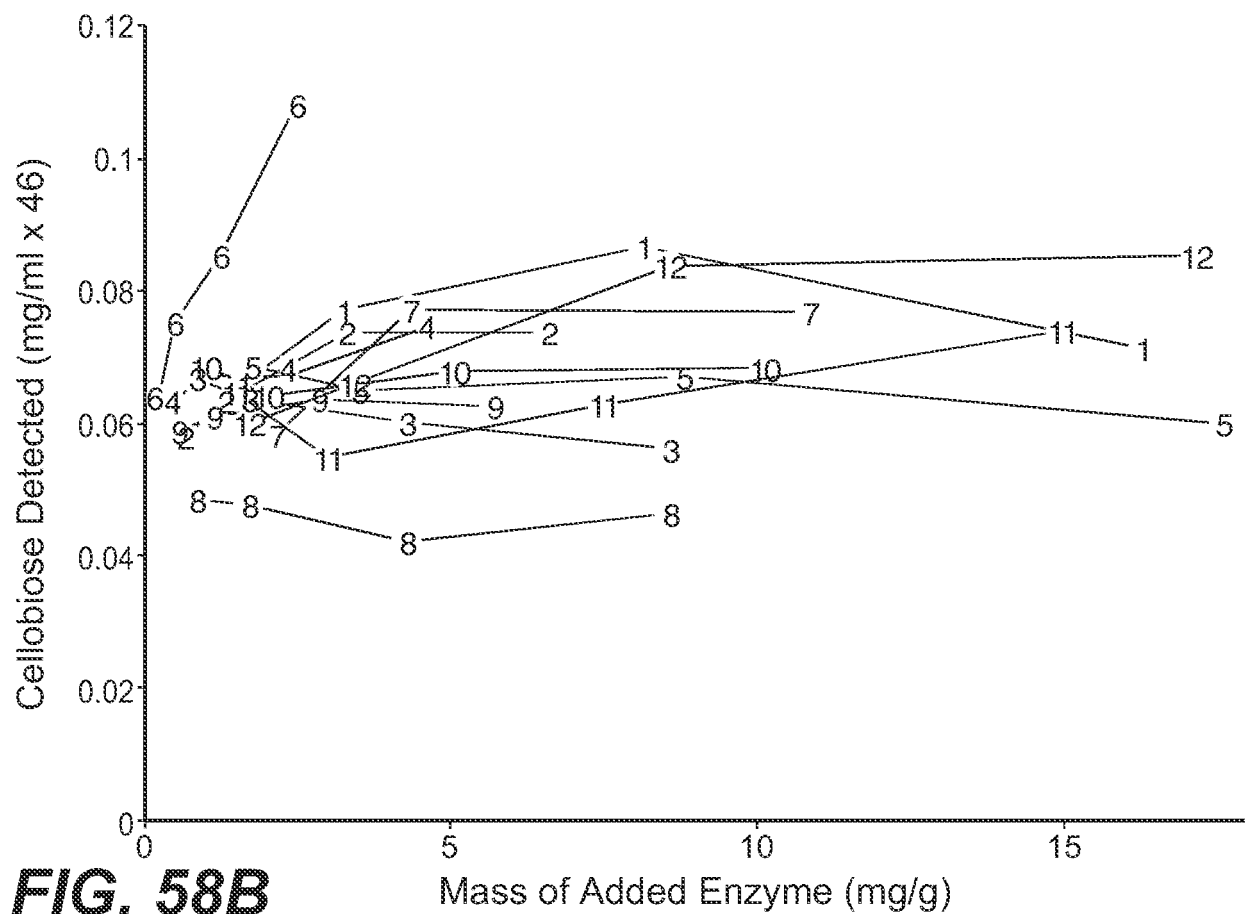
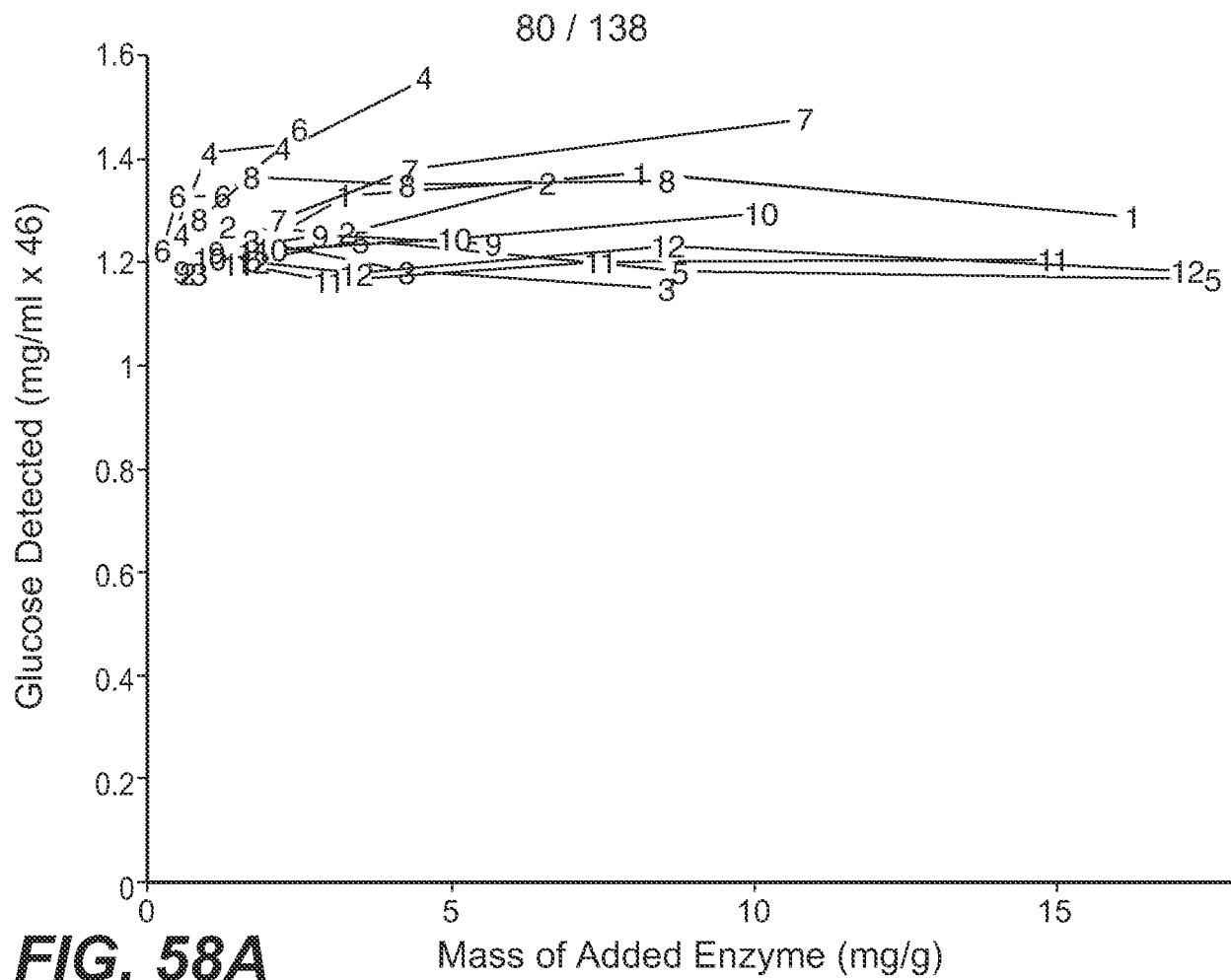
MSFSKIIATAGVLASASLVAGHGFVNIVIDGKKYYGGYLVNQYPYMSNPPEVIAWSTTATDLGFVDGTG
YQTPDIICHRGAKPGALTAPVSPGGTVELQWTPWPDSHHGPVINYLPAPCNGDCSTVDKTQLEFFKIAESG
LINDDNPPGIWASDNLIANNNSWTVTIPTTIAPGNYVLRHEIIALHSAQNQDGAQNYPPQCINLQVTGGGS
DNPAGTLGTALYHDTDPGILINIYQKLSSYIIPGPPLYTG

SEQ ID NO:207**Protein sequence of a GH61/endoglucanase from *T. aurantiacus***

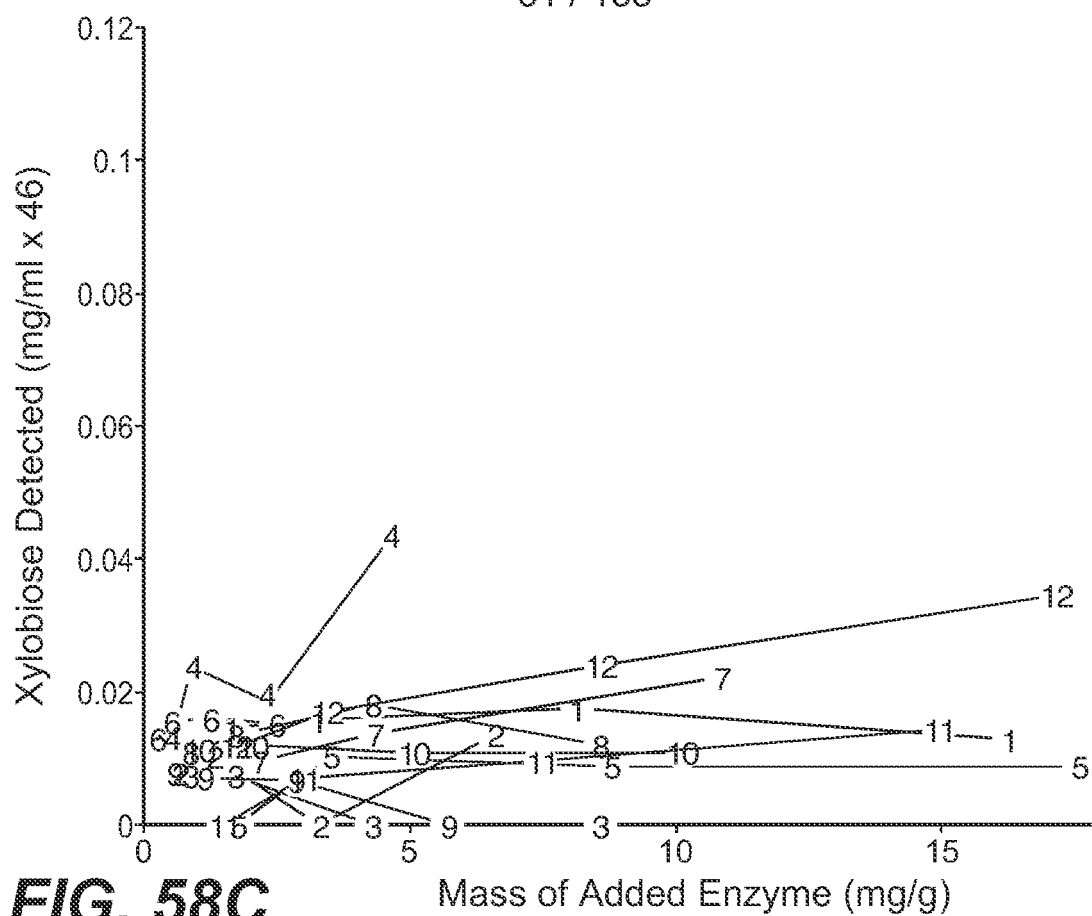
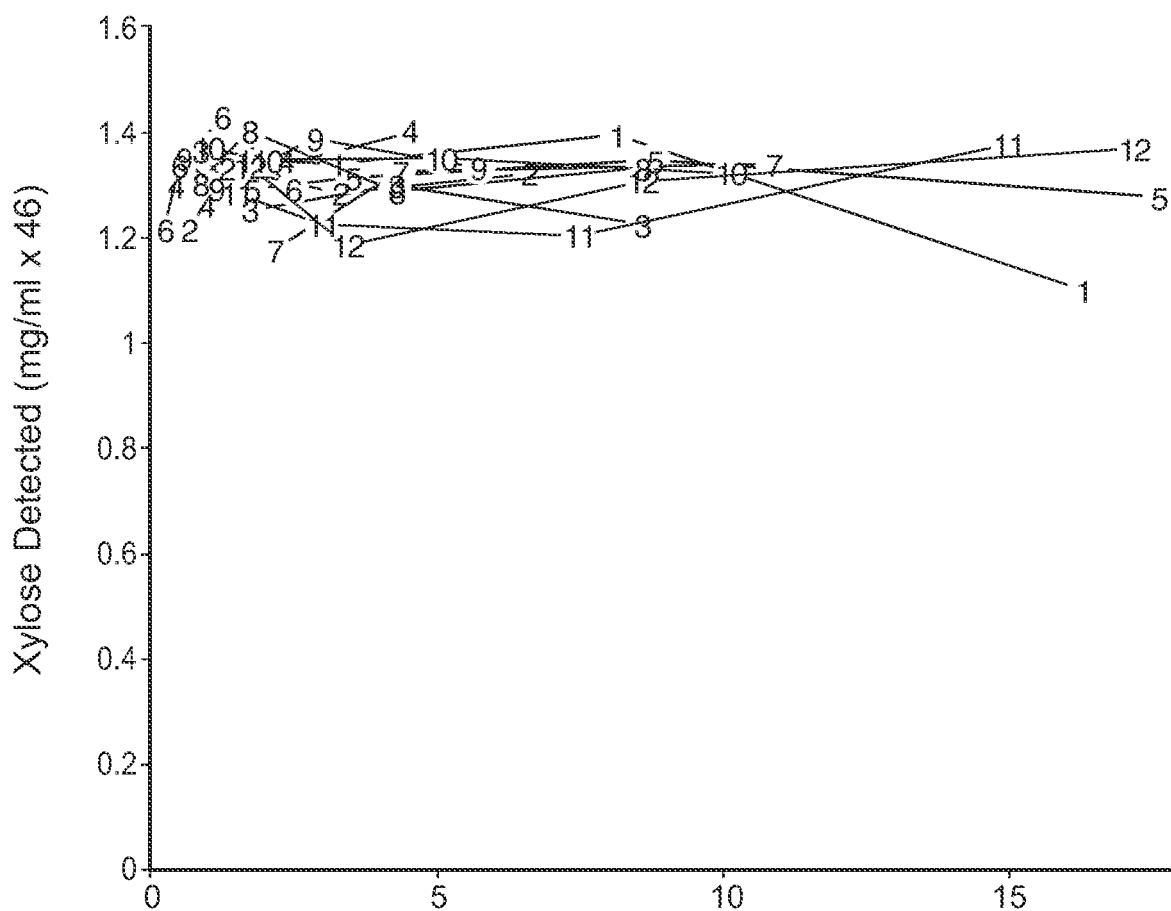
MSFSKIAAITGAITYASLAAAHGYVTGIVADGTYYGGYIVTQYPYMSTPPDVIAWSTKATDLGFVDPSS
YASSDIICHKGAEPGALSAKVAAGGTVELQWTDWPESHKGPVIDYLAACNGDCSTVDKTKLEFFKIDES
GLIDGSSAPGTWASDNLIANNNSWTVTIPTTIAPGNYVLRHEIIALHSAQNTNGAQNYPPQCINLEVTGS
GTDTPAGTLGTELYKATDPGILVNIYQTLTSDIPGPALYTGGSSGSSGSSNTAKATTSTASSSIVTPT
PVNNPTVTQTAVVDVTQTVSQNAAVATTTTAVATAVPTGTTFSFDSMTSDEFVSLMRATVNWLLSN
KKHARDLSY

FIG. 57

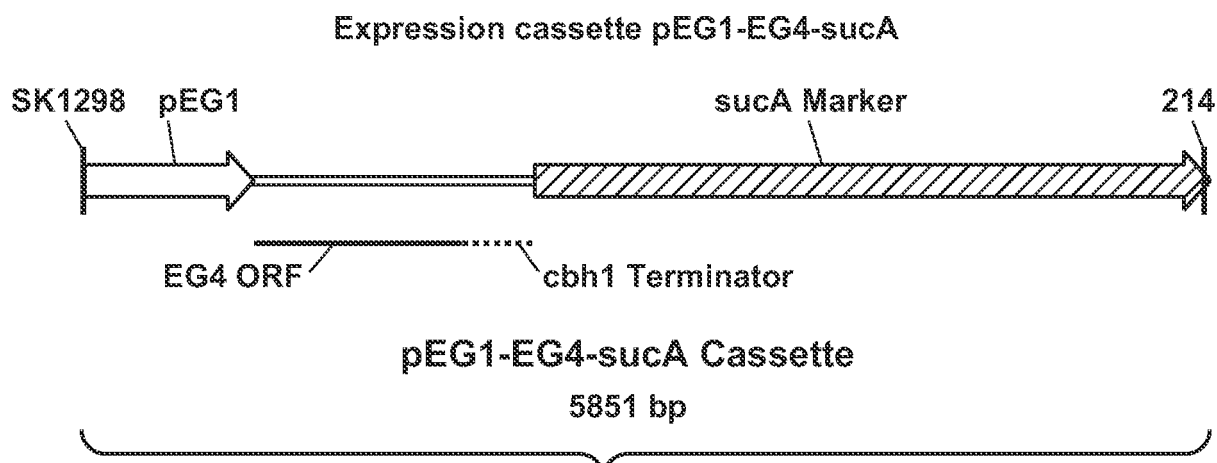
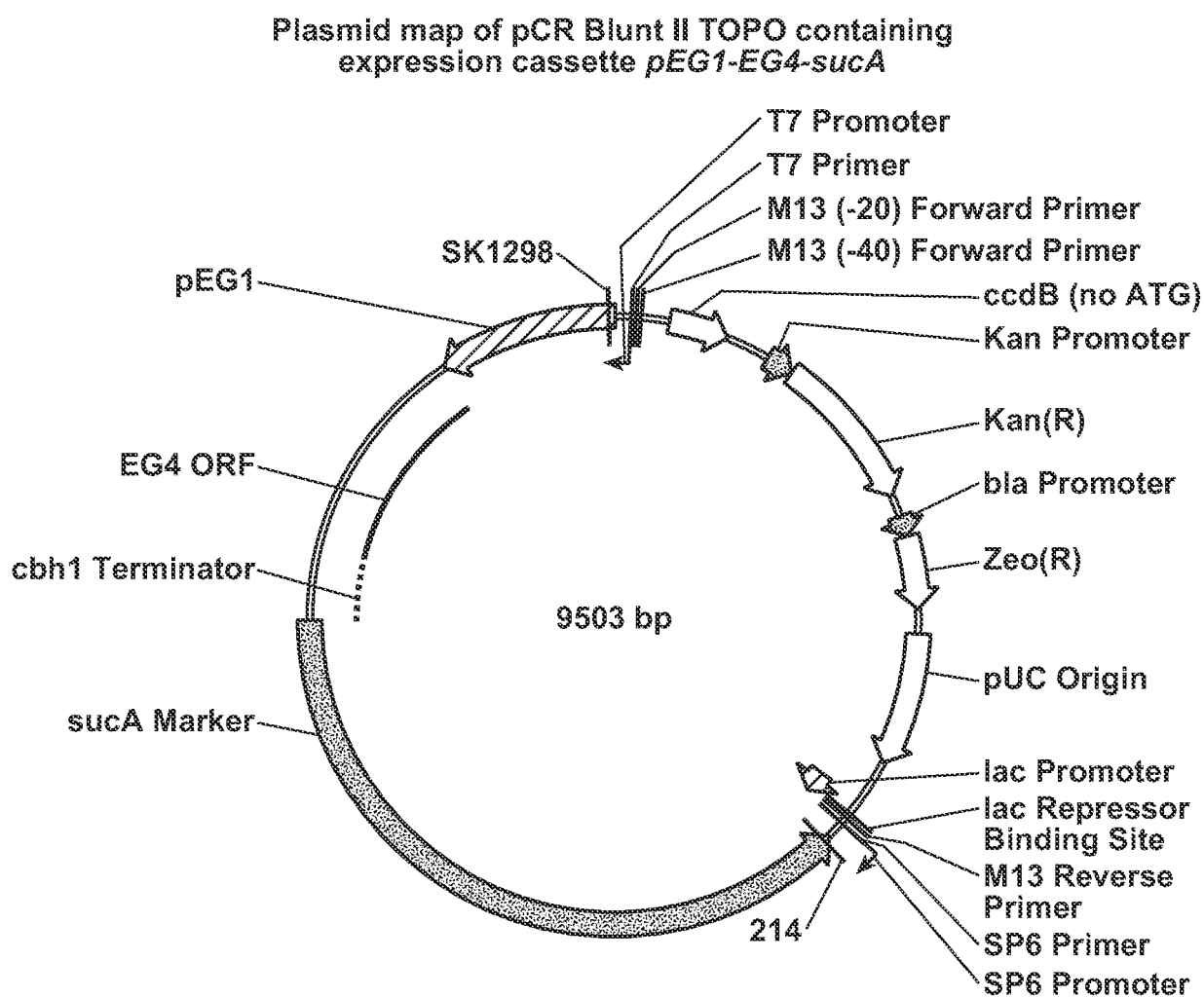
SUBSTITUTE SHEET (RULE 26)



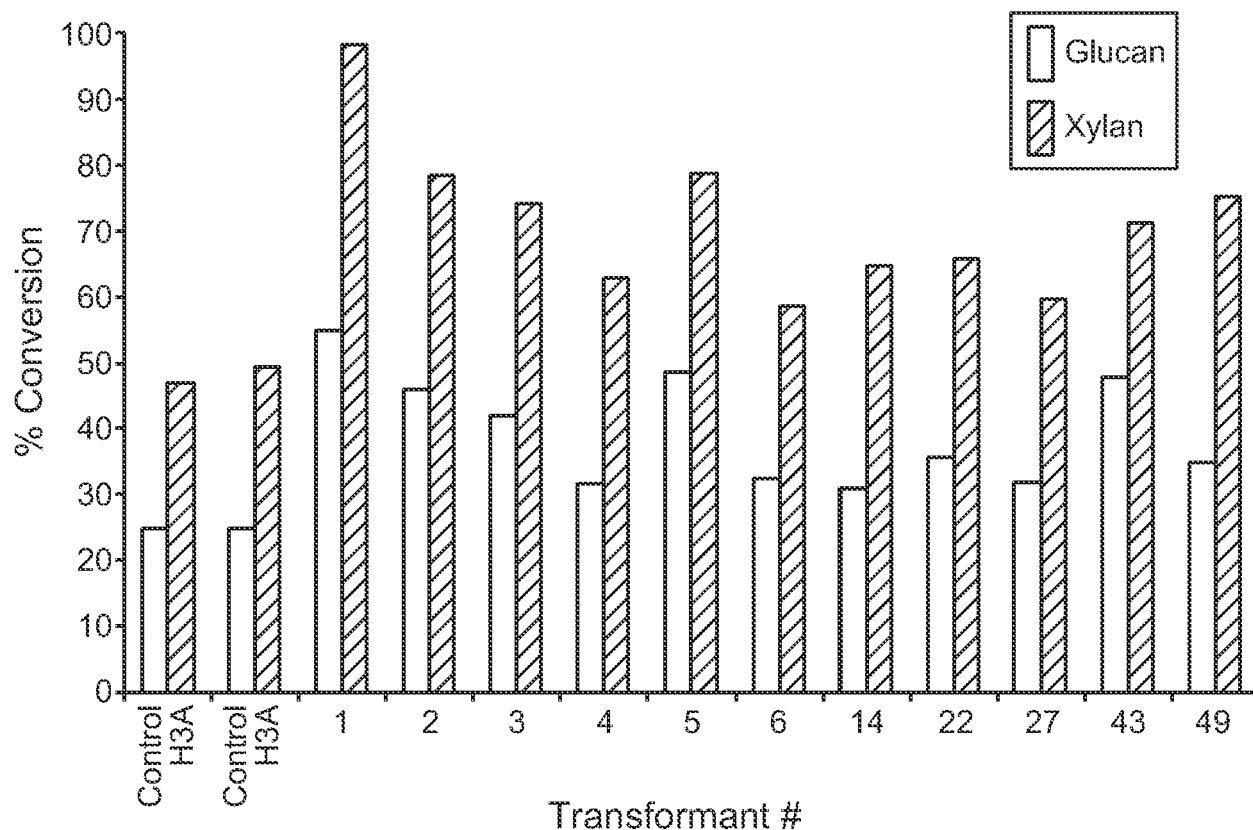
81 / 138

**FIG. 58C****FIG. 58D**

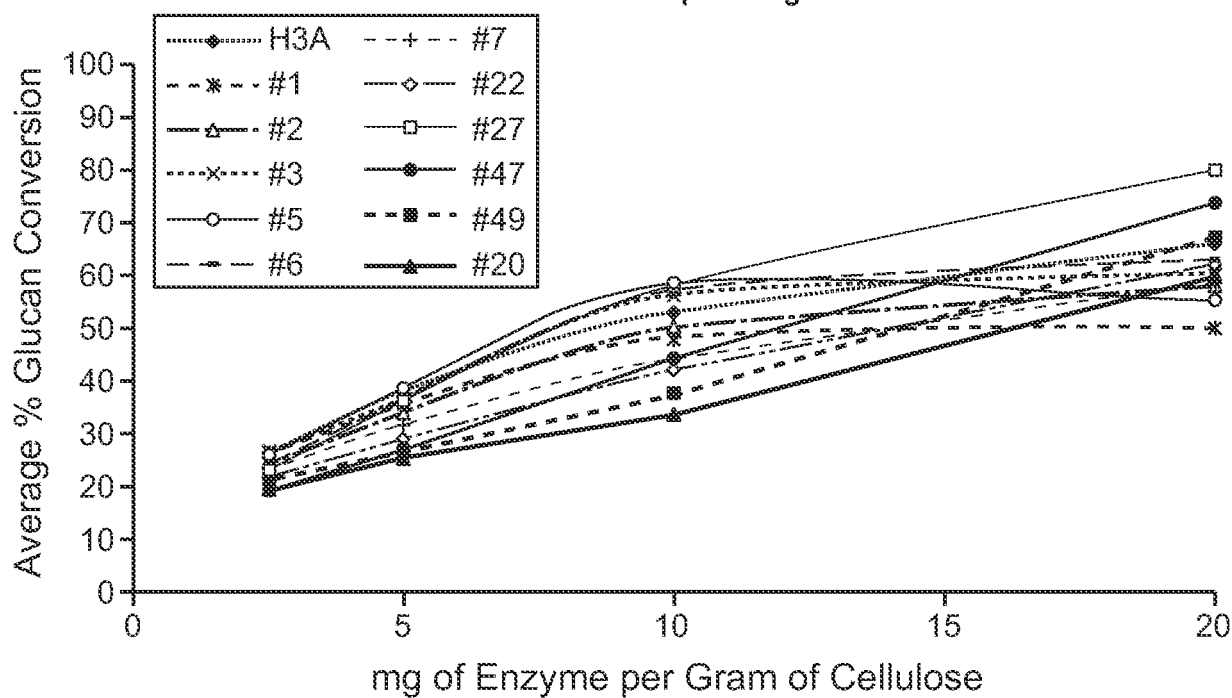
82 / 138

**FIG. 59A****FIG. 59B**

83 / 138

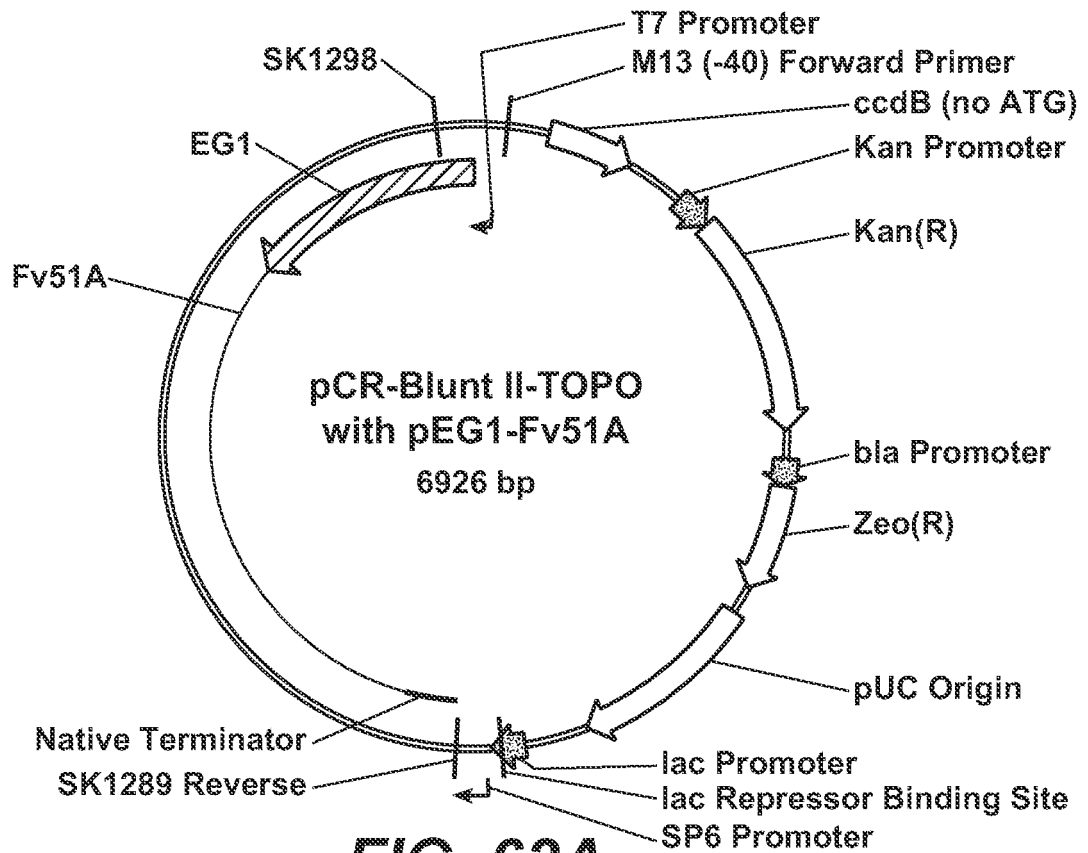
**FIG. 60**

% Glucan conversion with increasing protein dosage of H3A transformants expressing EG4

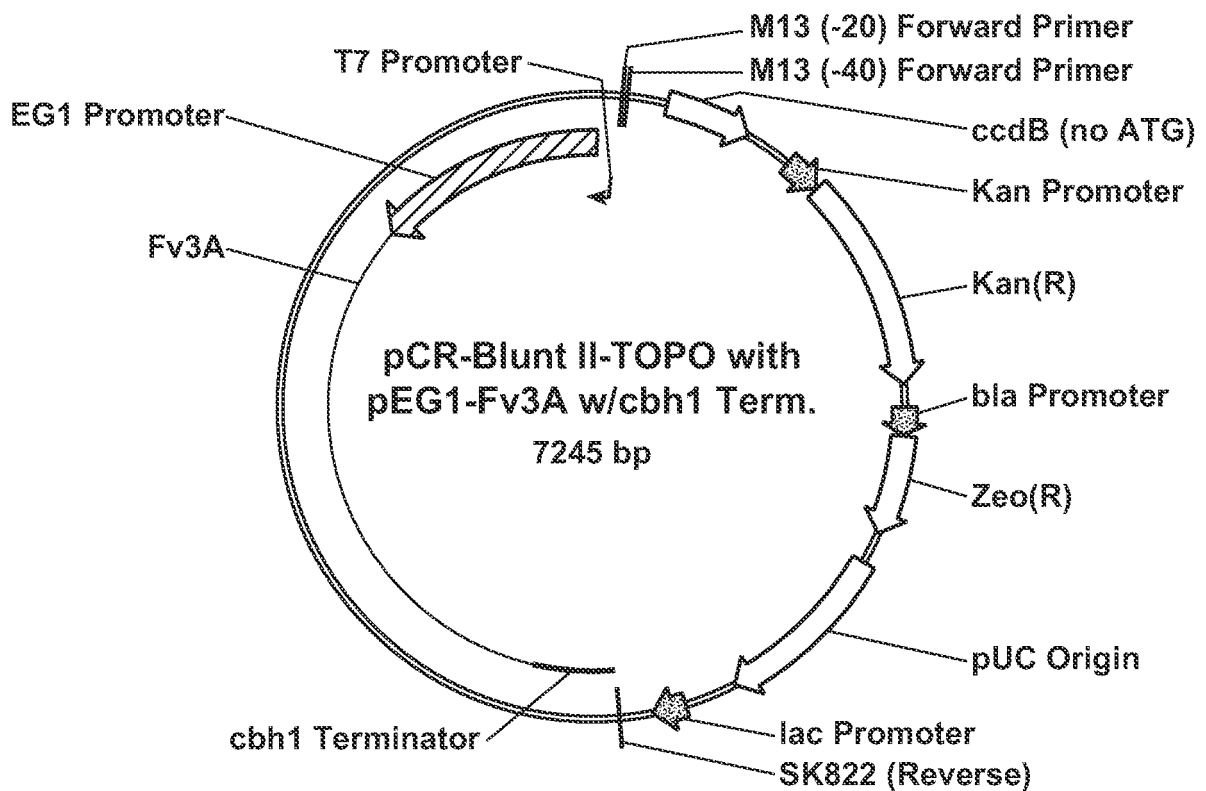
**FIG. 61**

84 / 138

Plasmid map of pCR-Blunt II TOPO plasmid with pEG1-Fv51A expression cassette

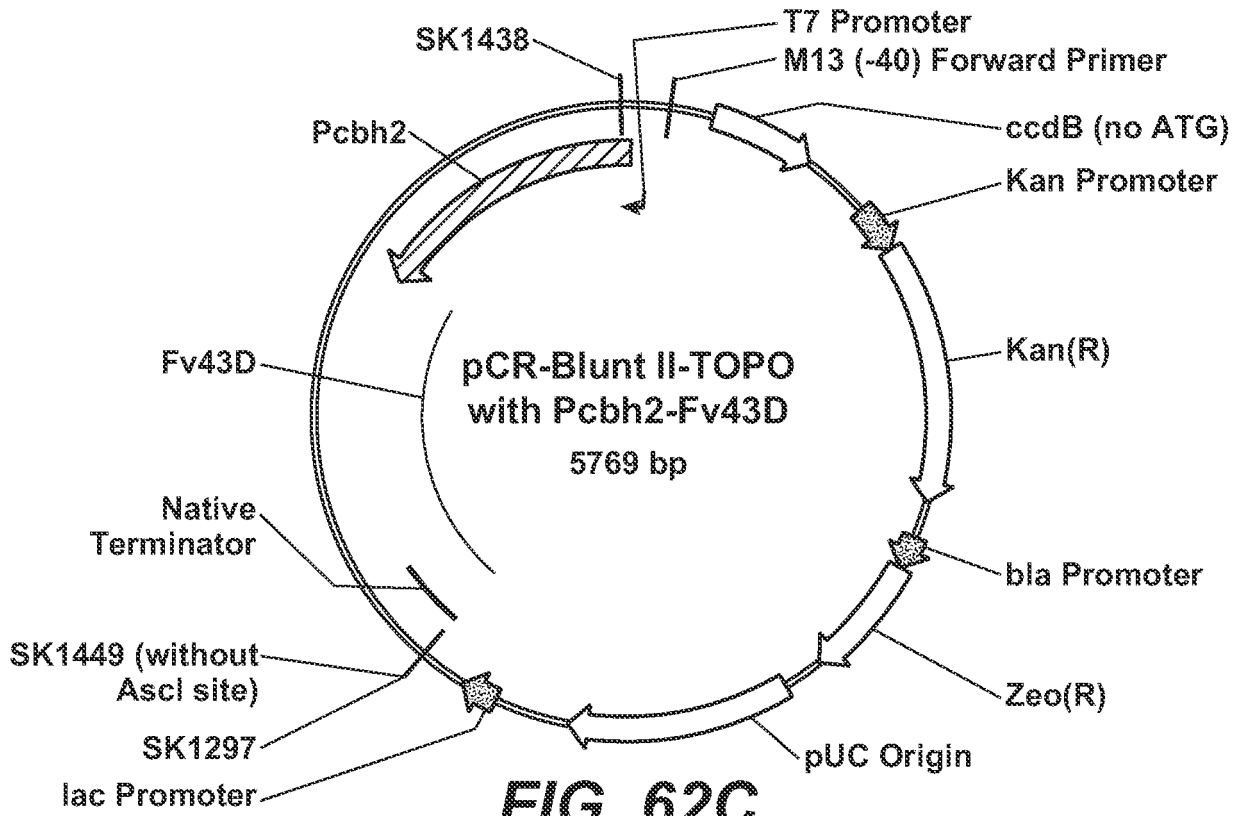
**FIG. 62A**

Plasmid map of pCR-Blunt II TOPO plasmid with pEG1-Fv3A w/ cbh1 term

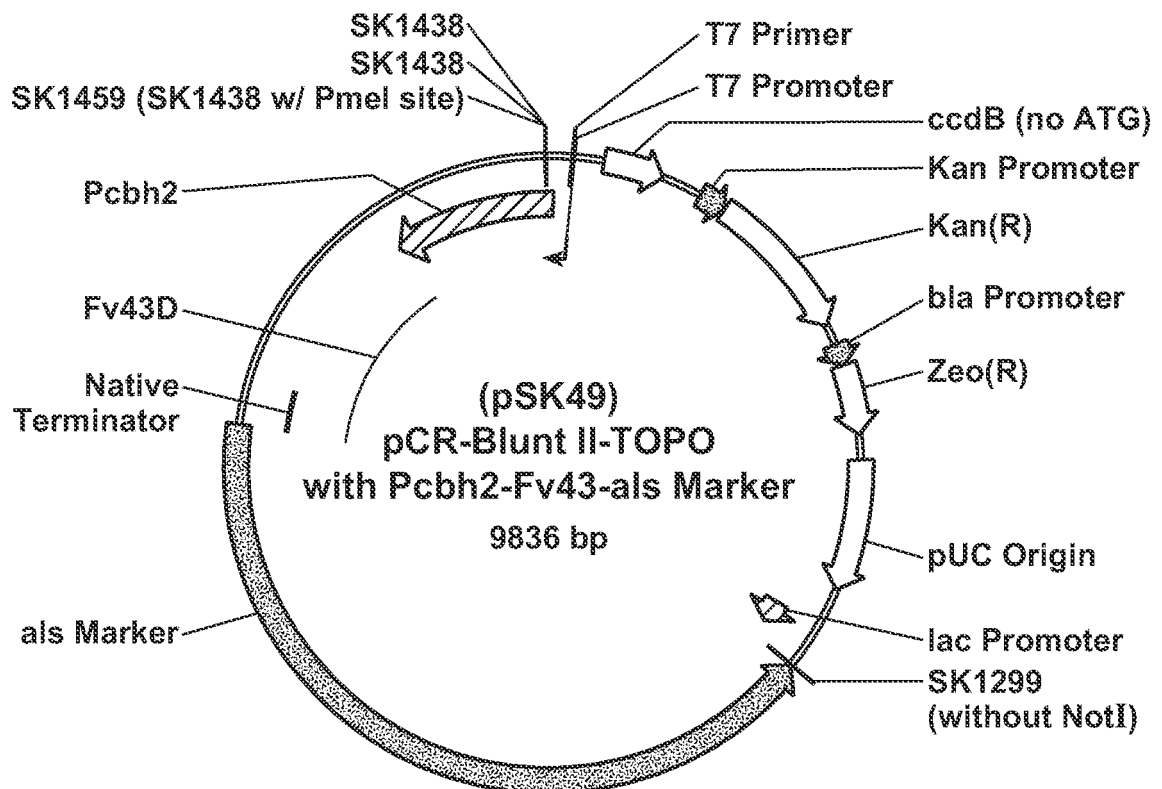
**FIG. 62B**

85 / 138

Plasmid map of pCR-Blunt II TOPO plasmid with Pcbh2-Fv43D



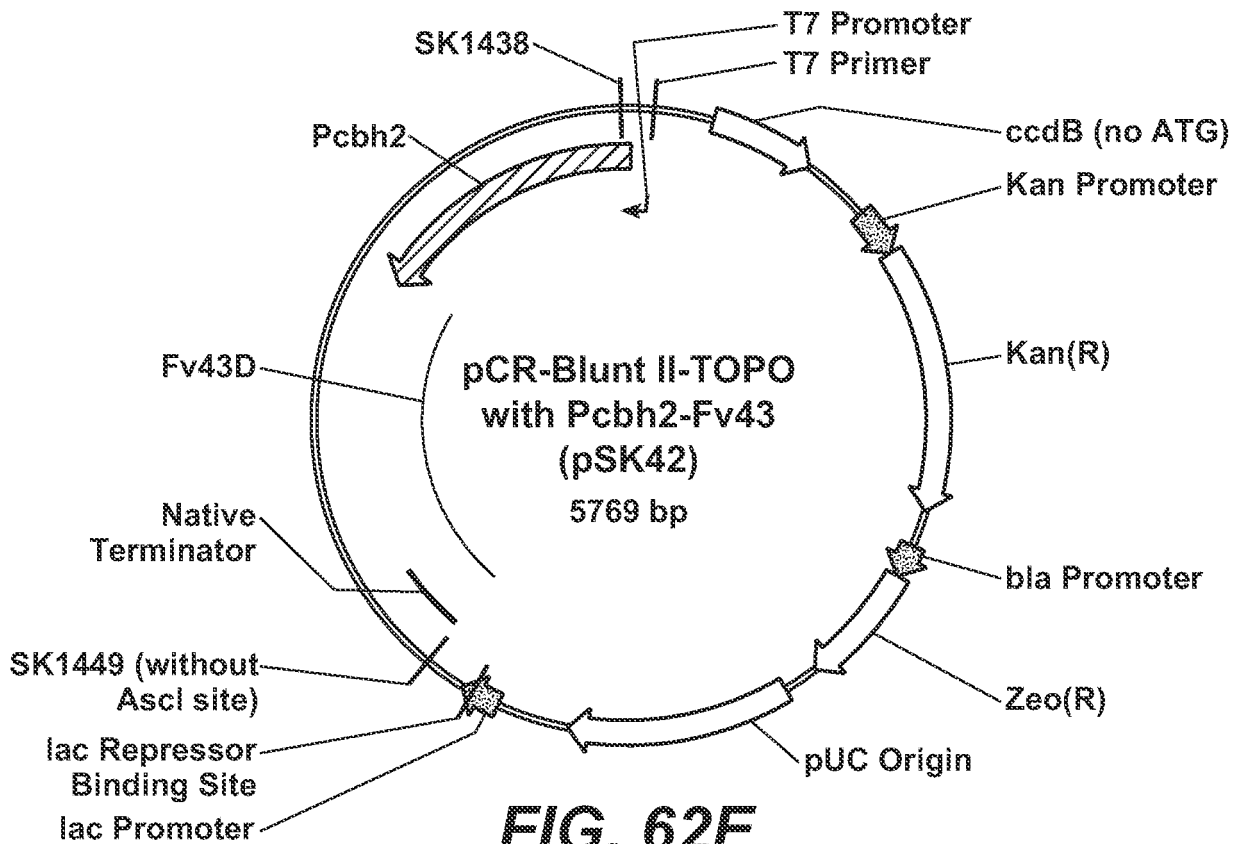
Plasmid map of pCR Blunt-II-TOPO plasmid with Pcbh2-Fv43D-als marker (pSK49)

**FIG. 62D**

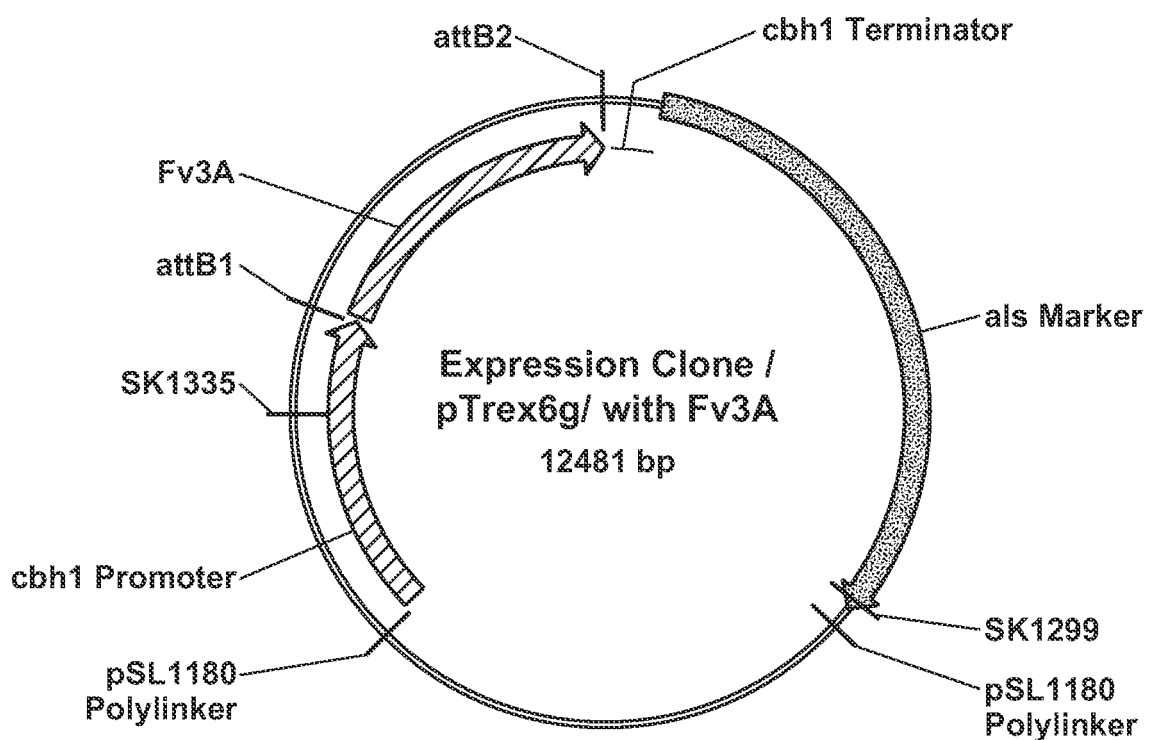
SUBSTITUTE SHEET (RULE 26)

86 / 138

Plasmid map of pCR-Blunt II-TOPO with Pcbh2-Fv43D (pSK42)

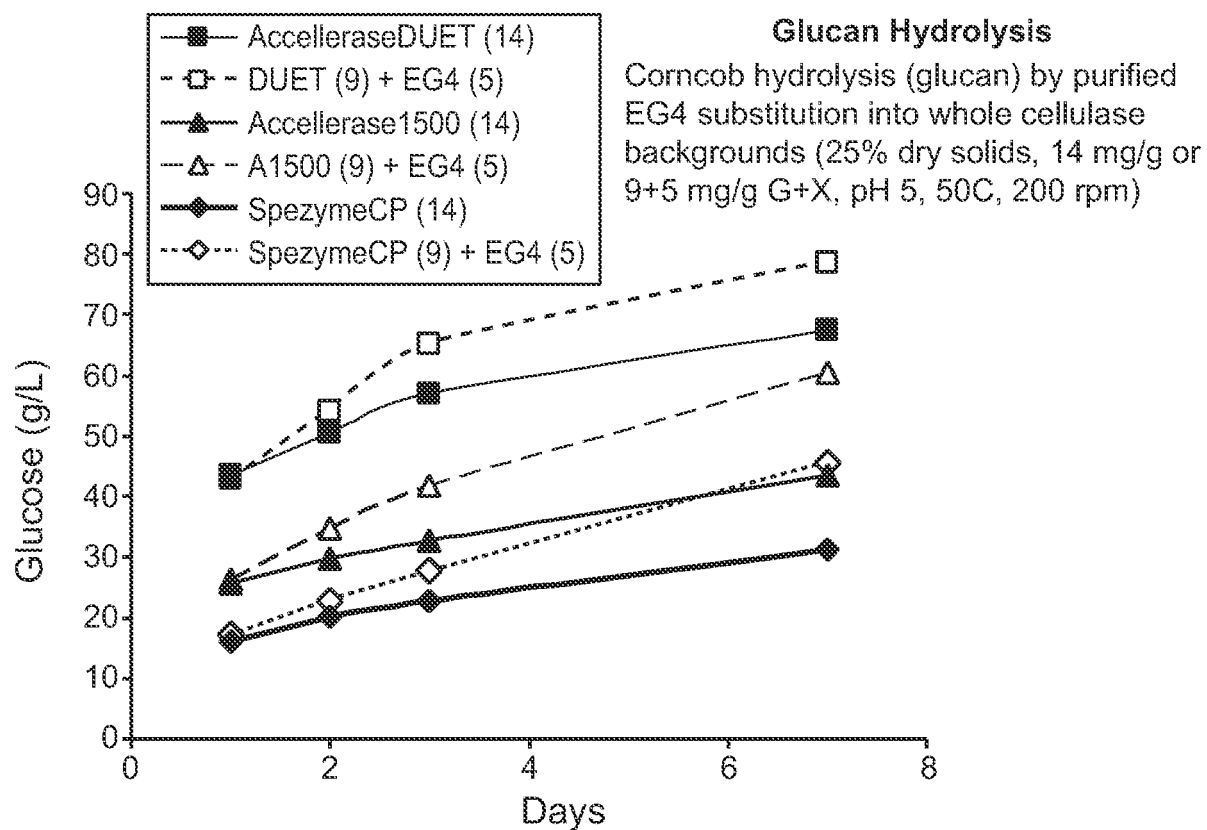
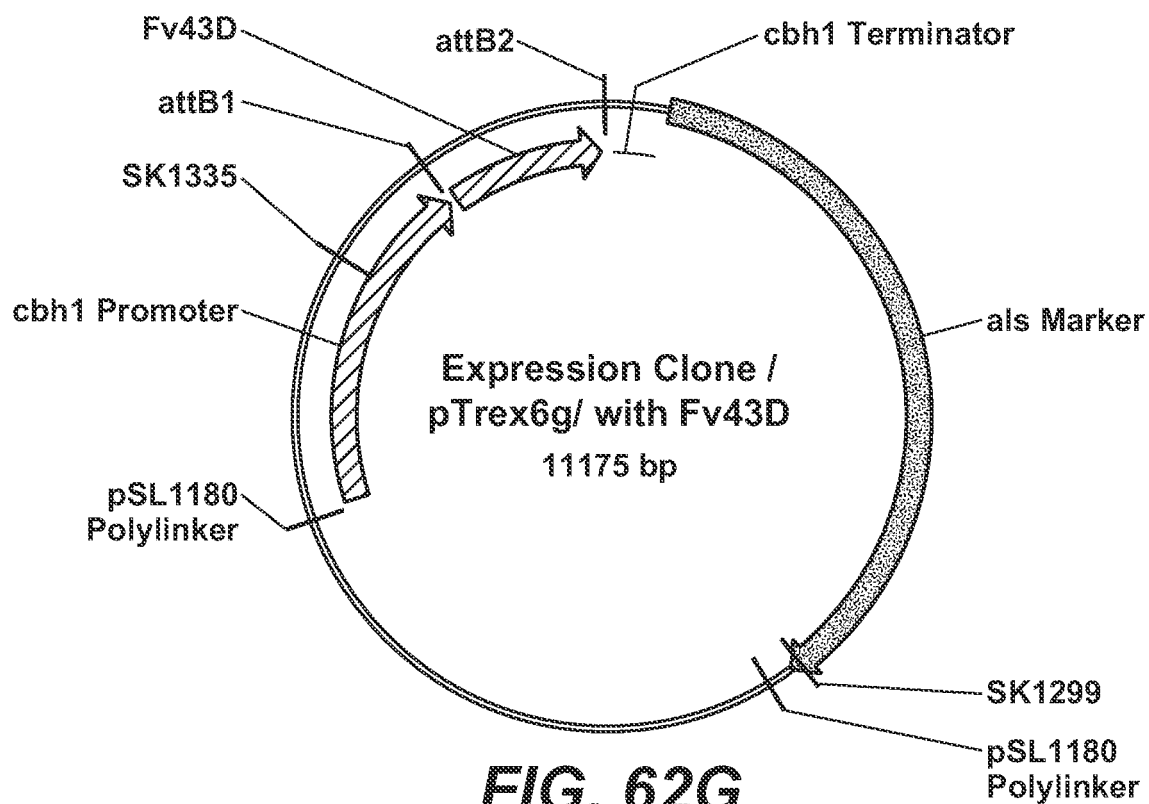


Plasmid map of pTrex6g with Fv3A

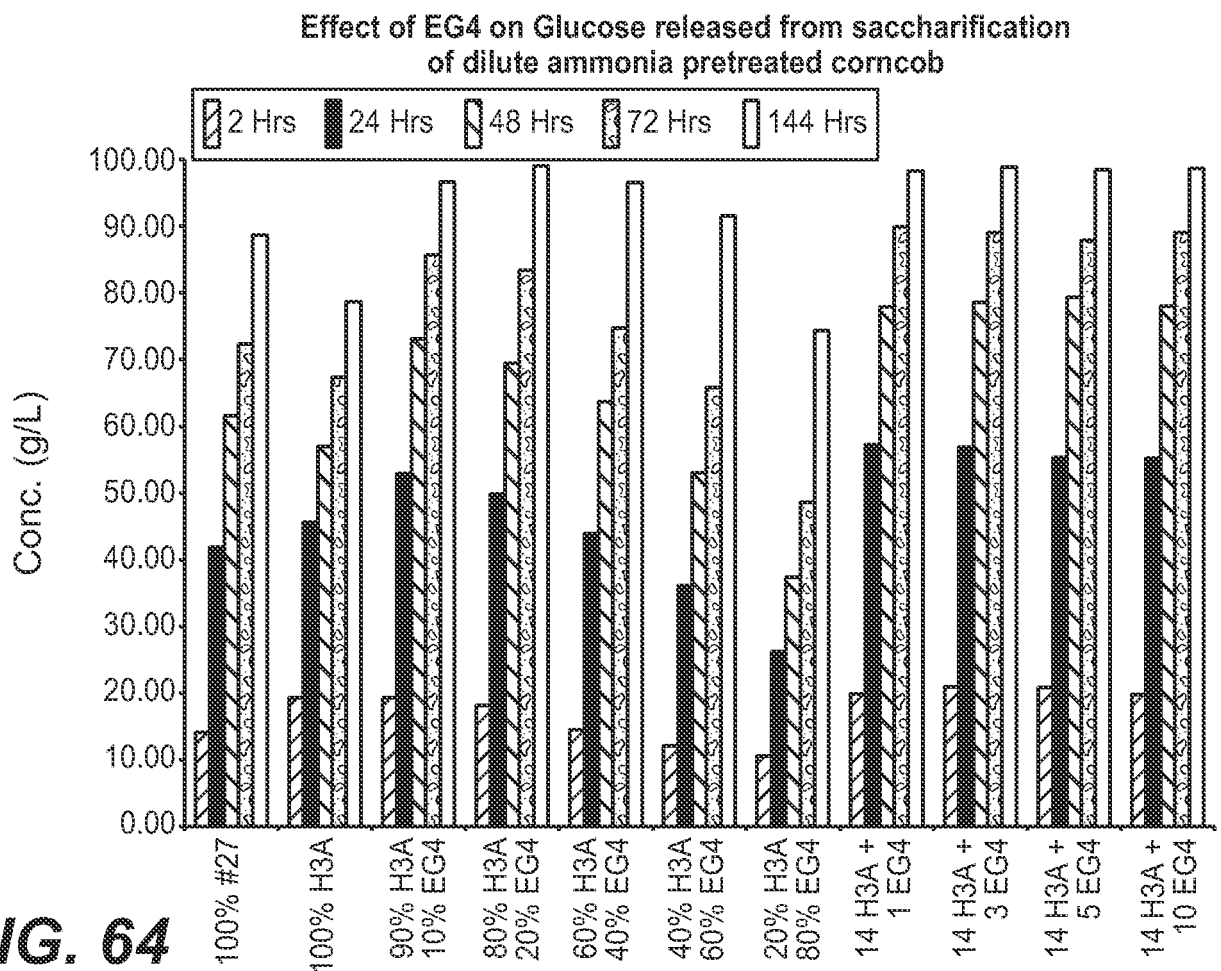
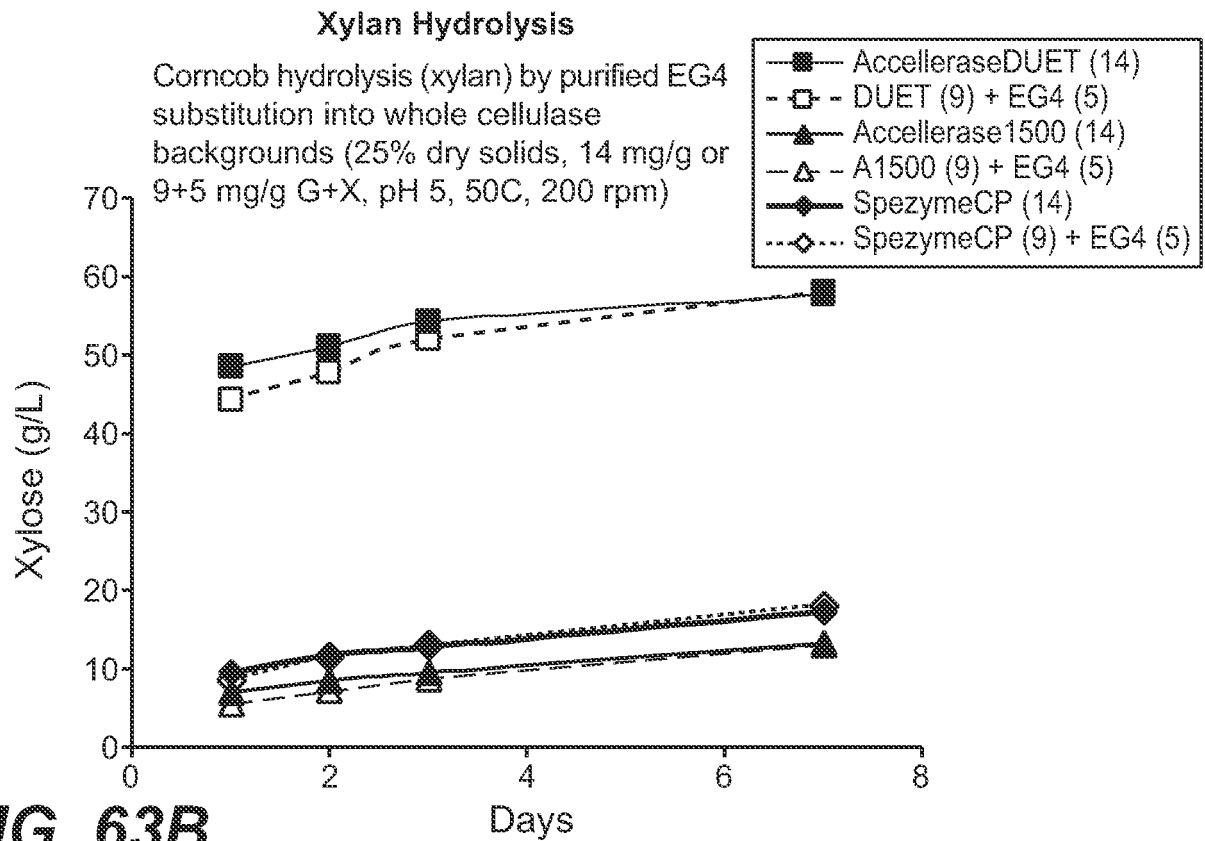


87 / 138

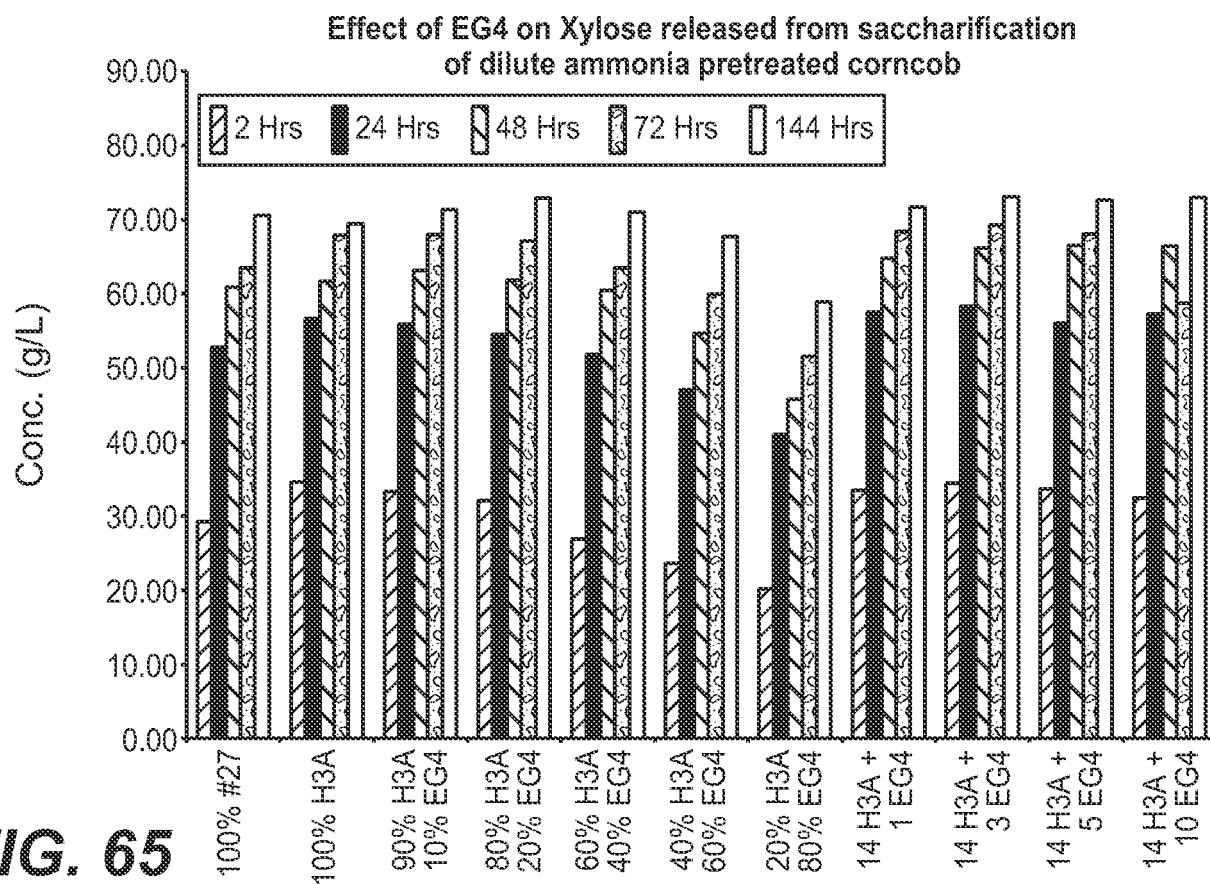
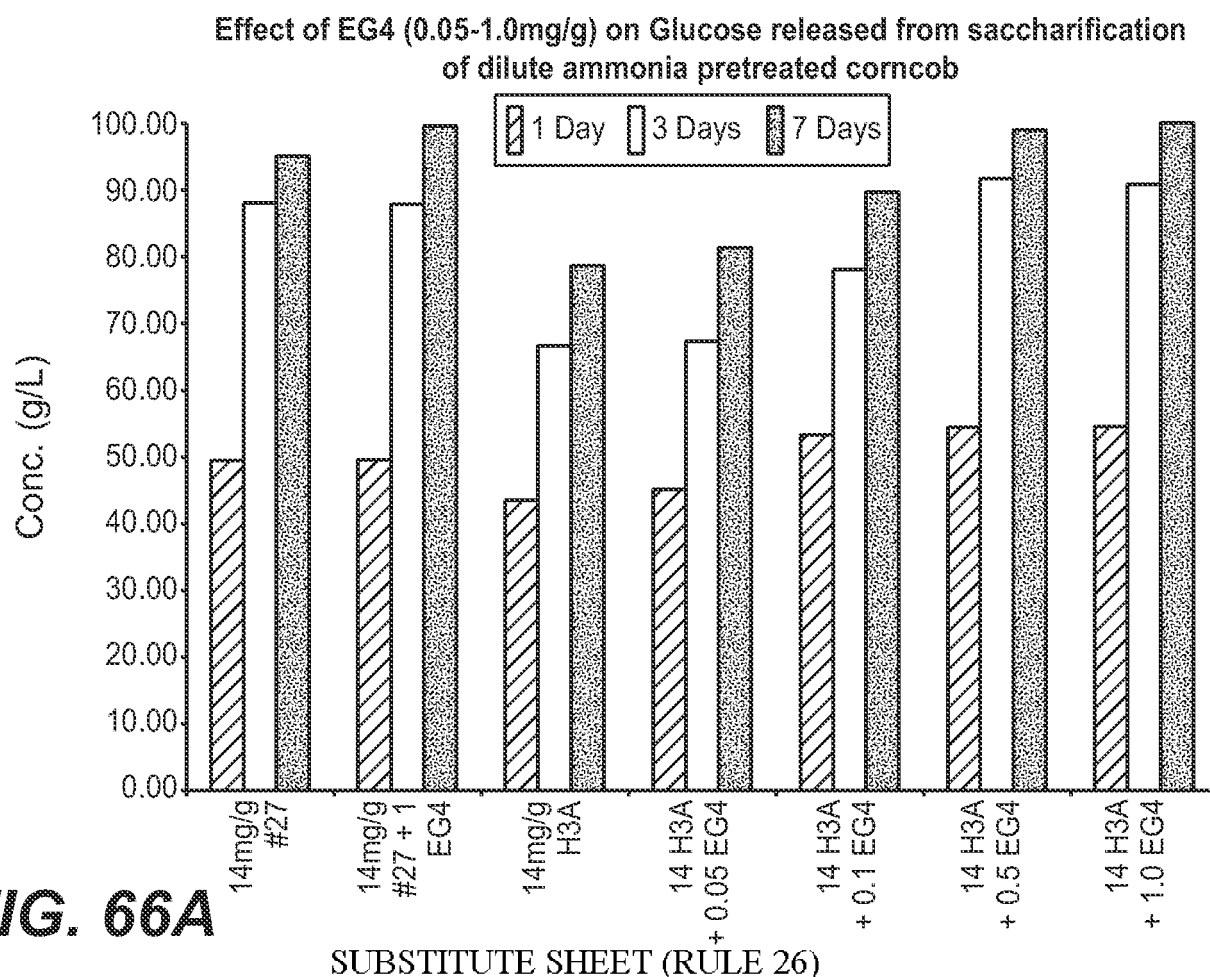
Plasmid map of pTrex6g with Fv43D



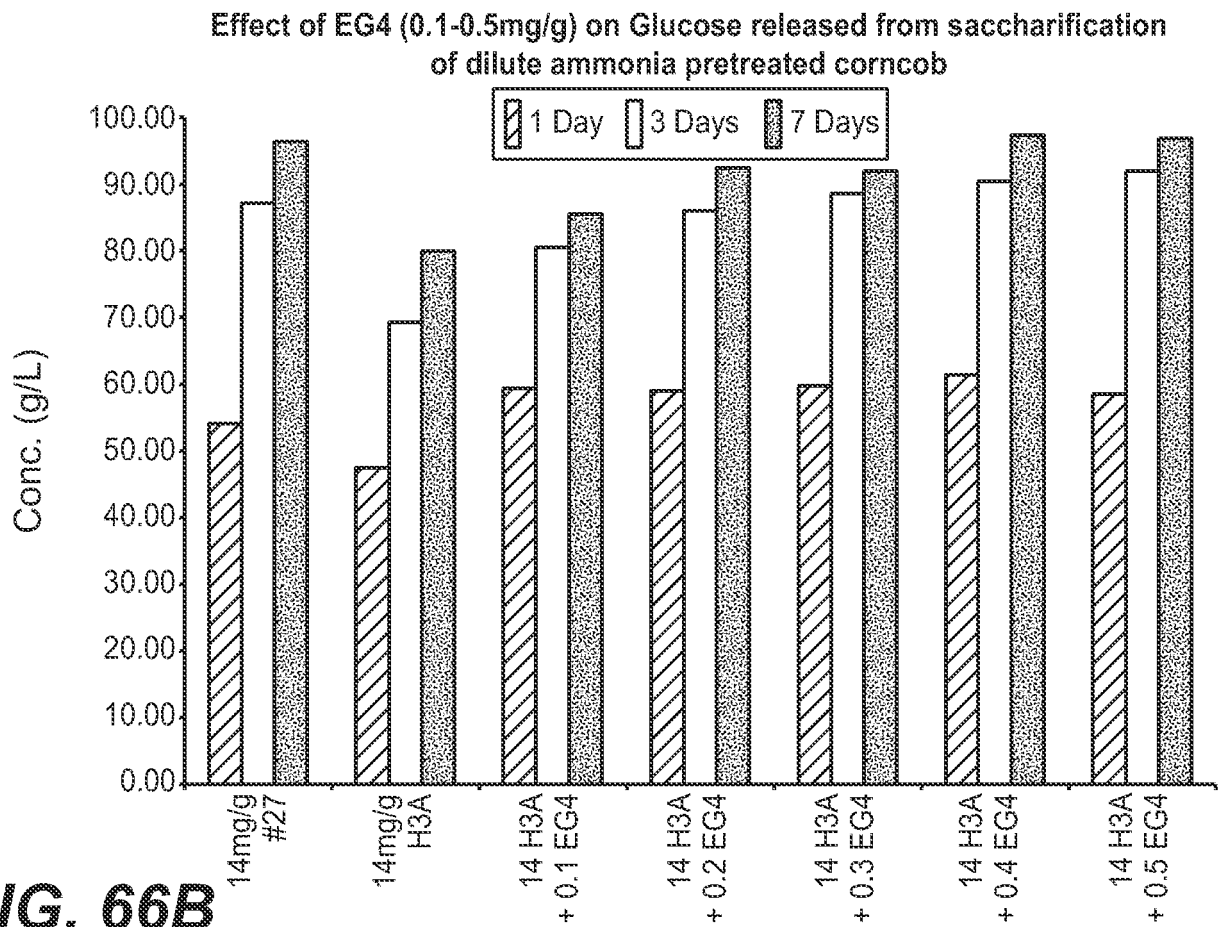
88 / 138



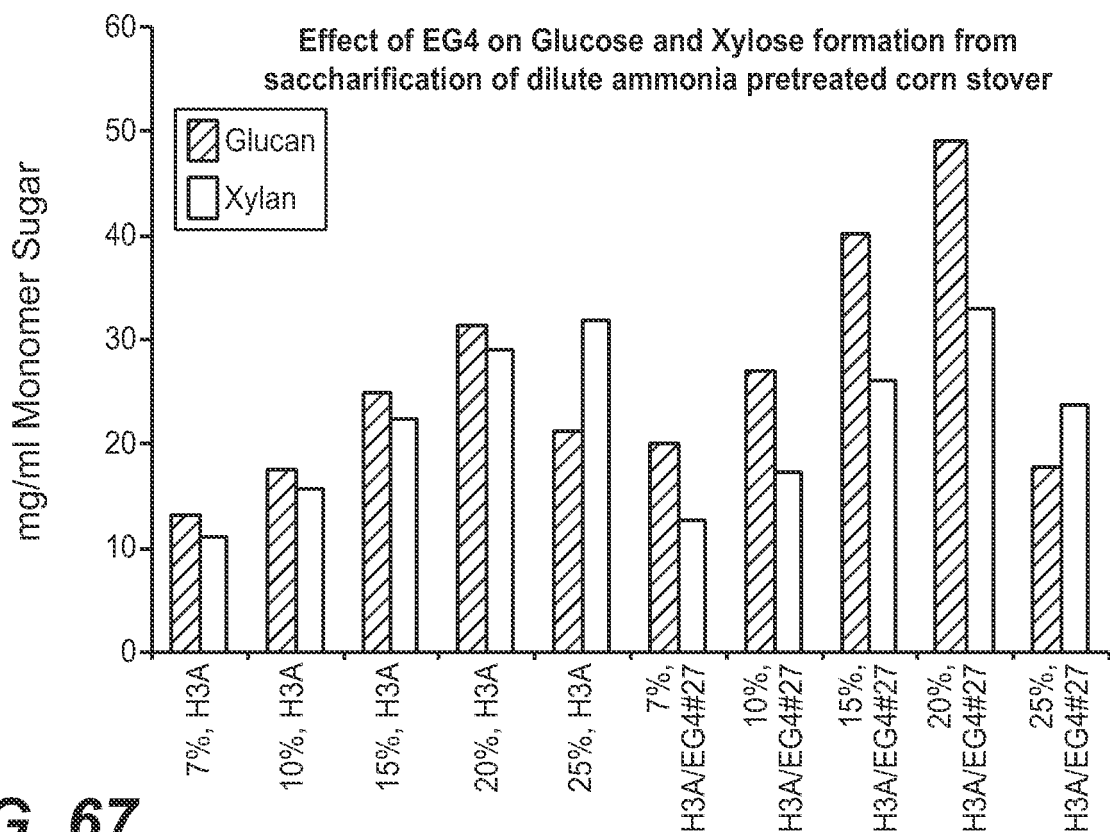
89 / 138

**FIG. 65****FIG. 66A**

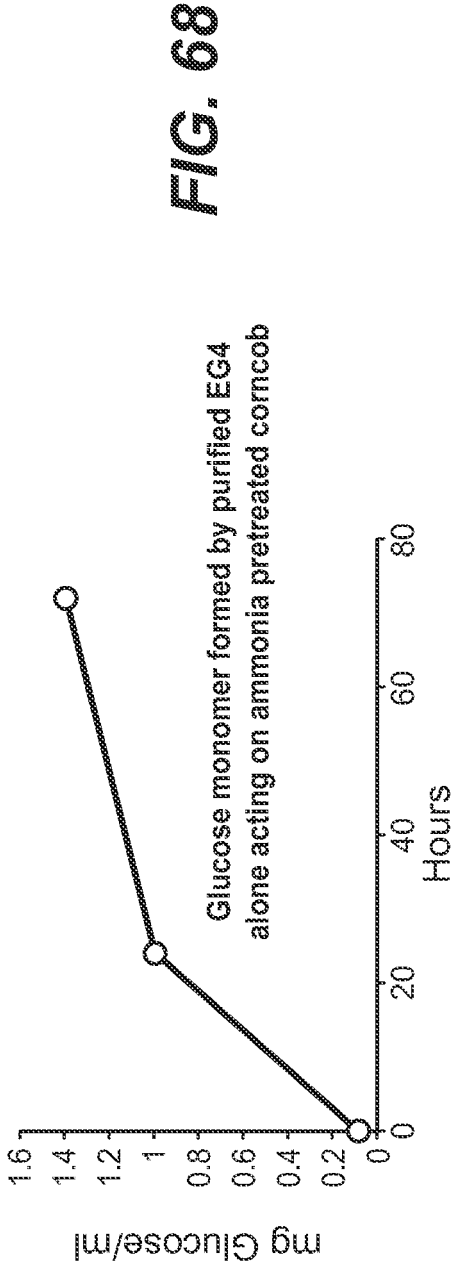
90 / 138



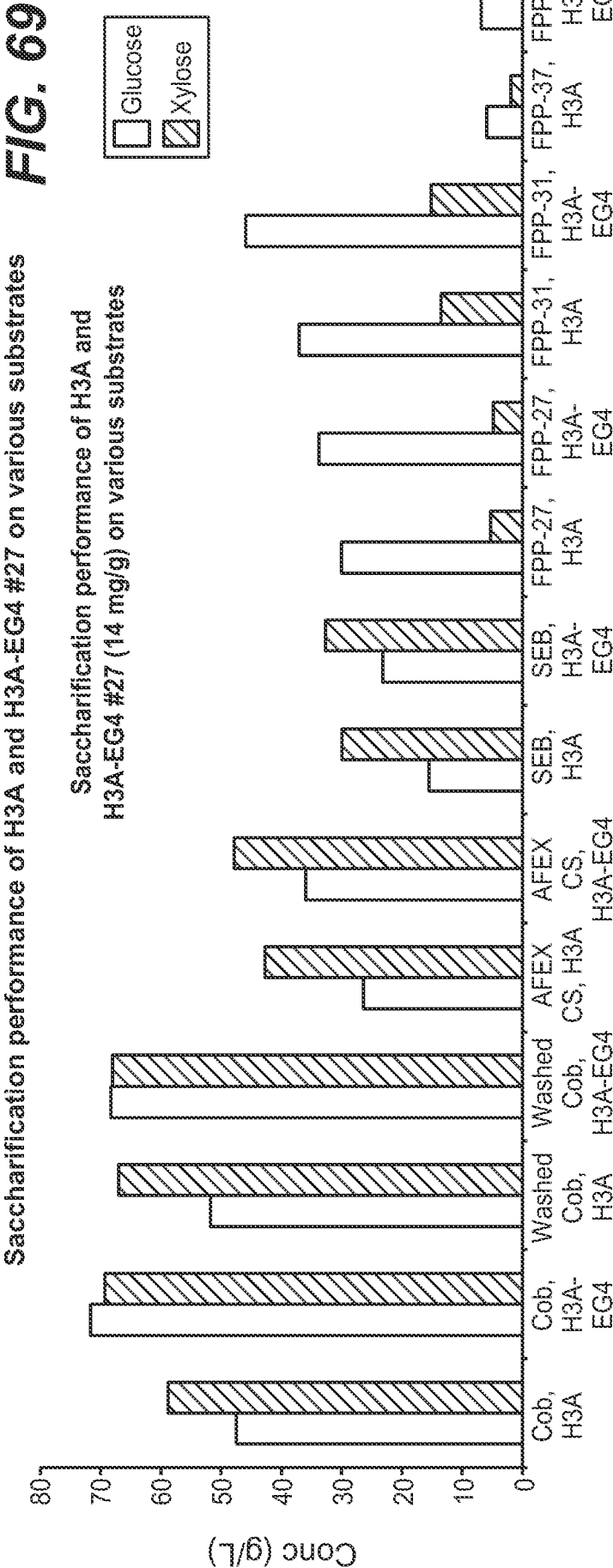
Effect of EG4 on saccharification of dilute ammonia pretreated corn stover



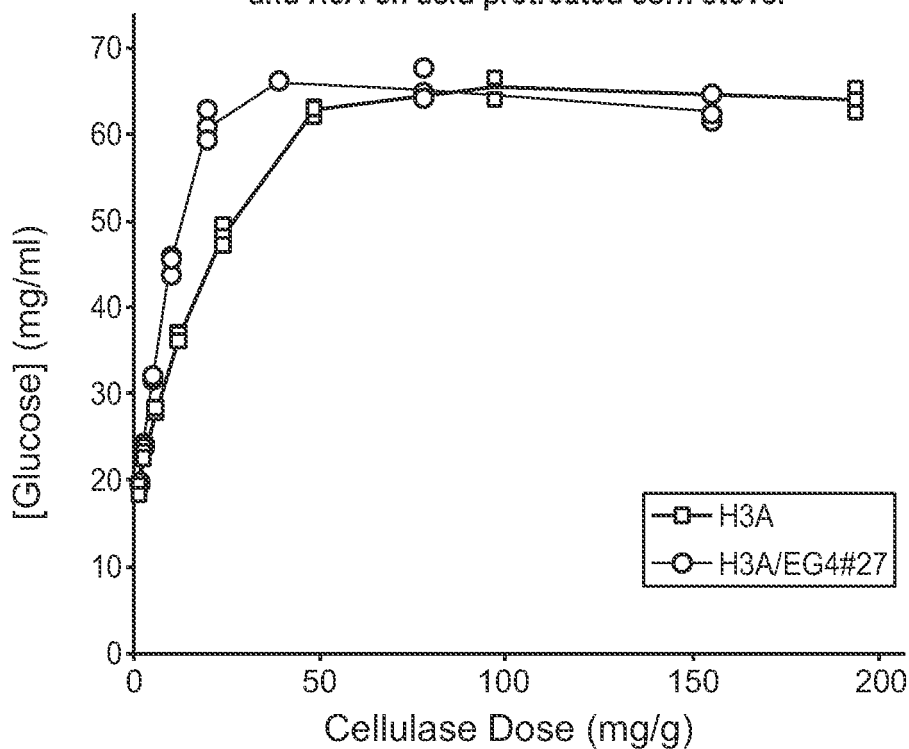
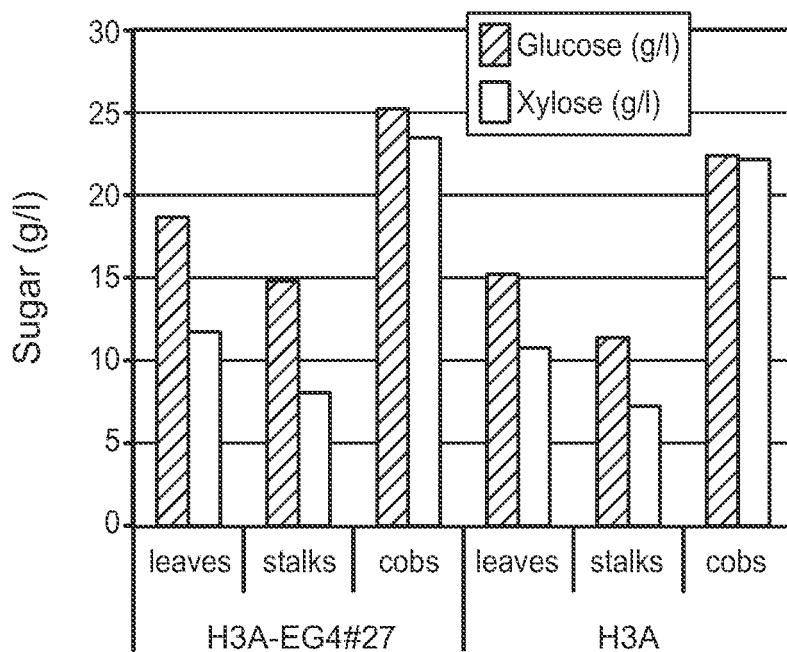
Glucose monomer formed by purified EG4 alone acting on ammonia pretreated corncob.



Saccharification performance of H3A and H3A-EG4 #27 on various substrates



92 / 138

Saccharification performance of H3A/EG4 #27
and H3A on acid pretreated corn stover**FIG. 70**Saccharification performance of *T. reesei*
integrated strains H3A/EG4#27 or H3A on
dilute ammonia pretreated corn leaves, stalks, or cobs**FIG. 71**

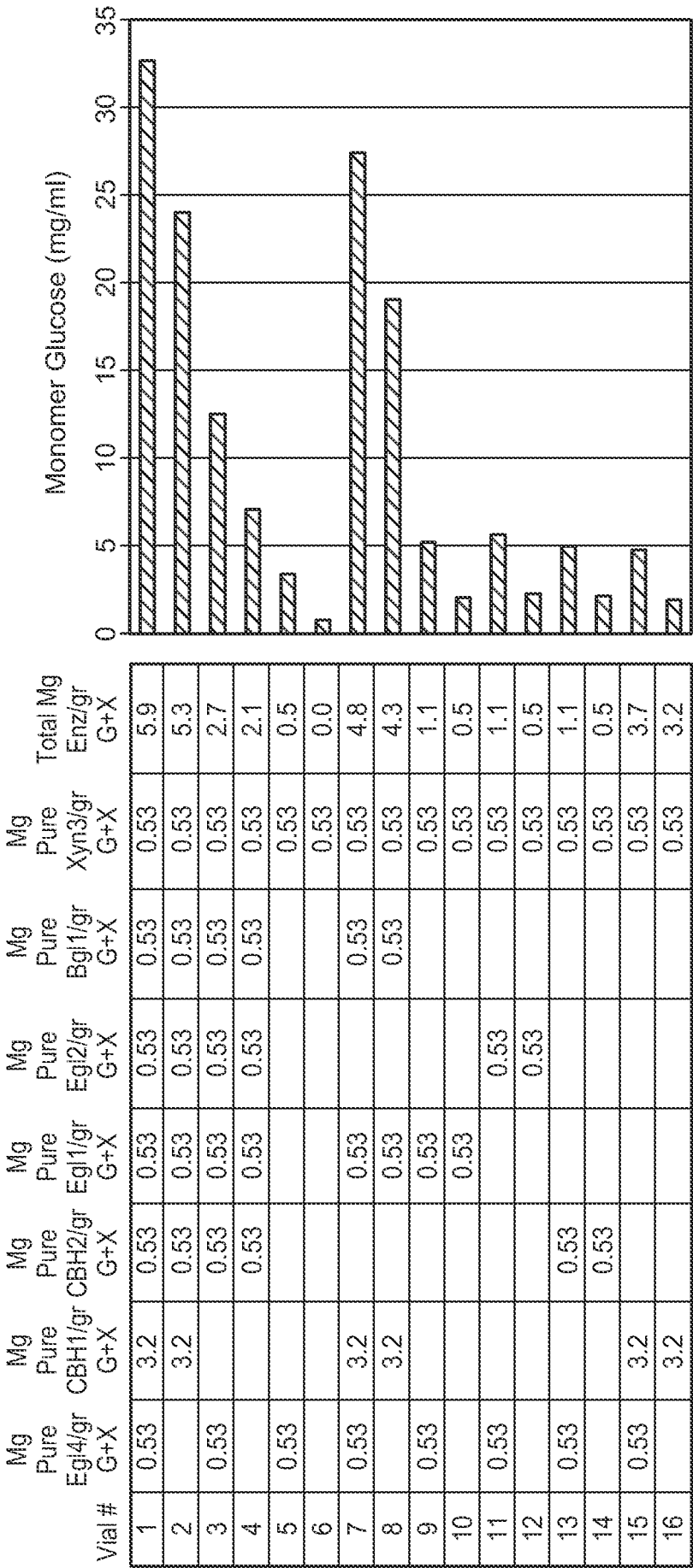


FIG. 72B-1

FIG. 72A-1

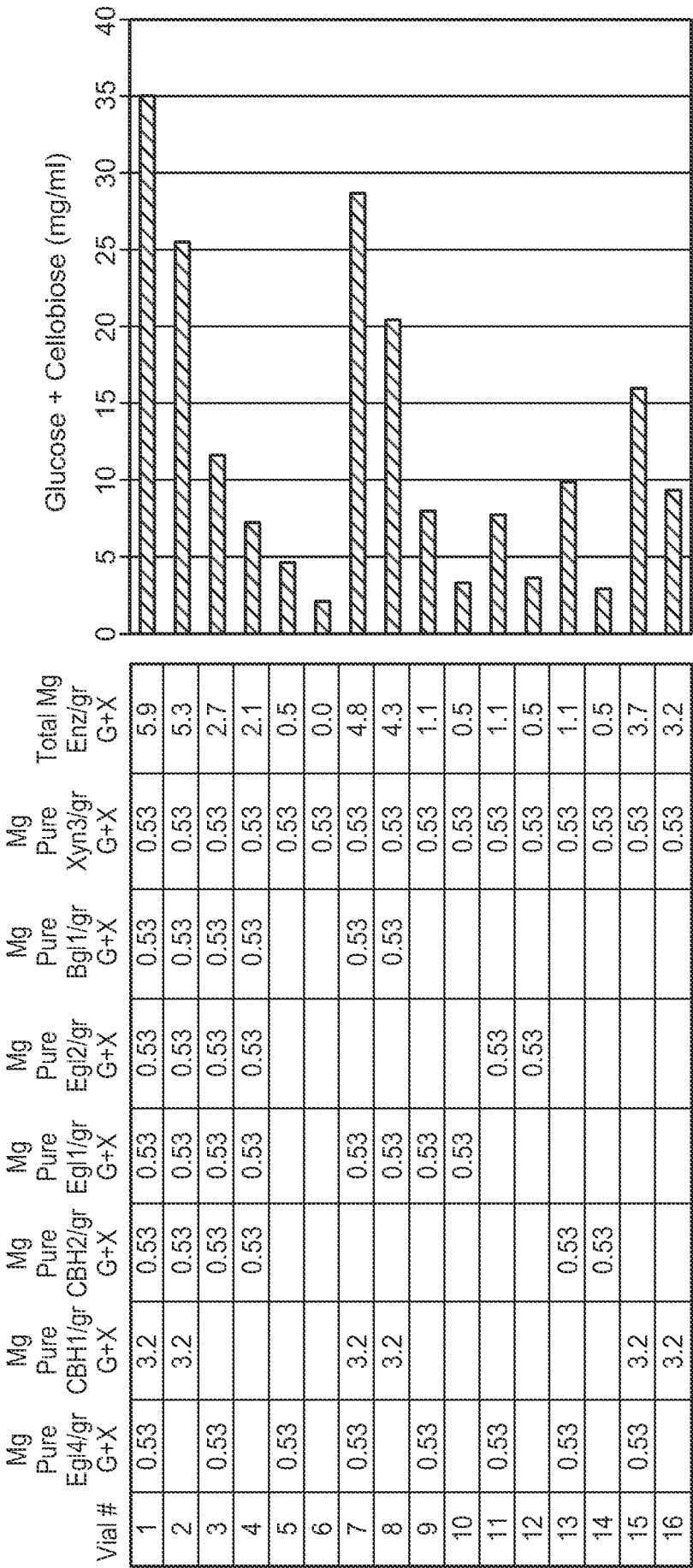


FIG. 72B-2

FIG. 72A-2

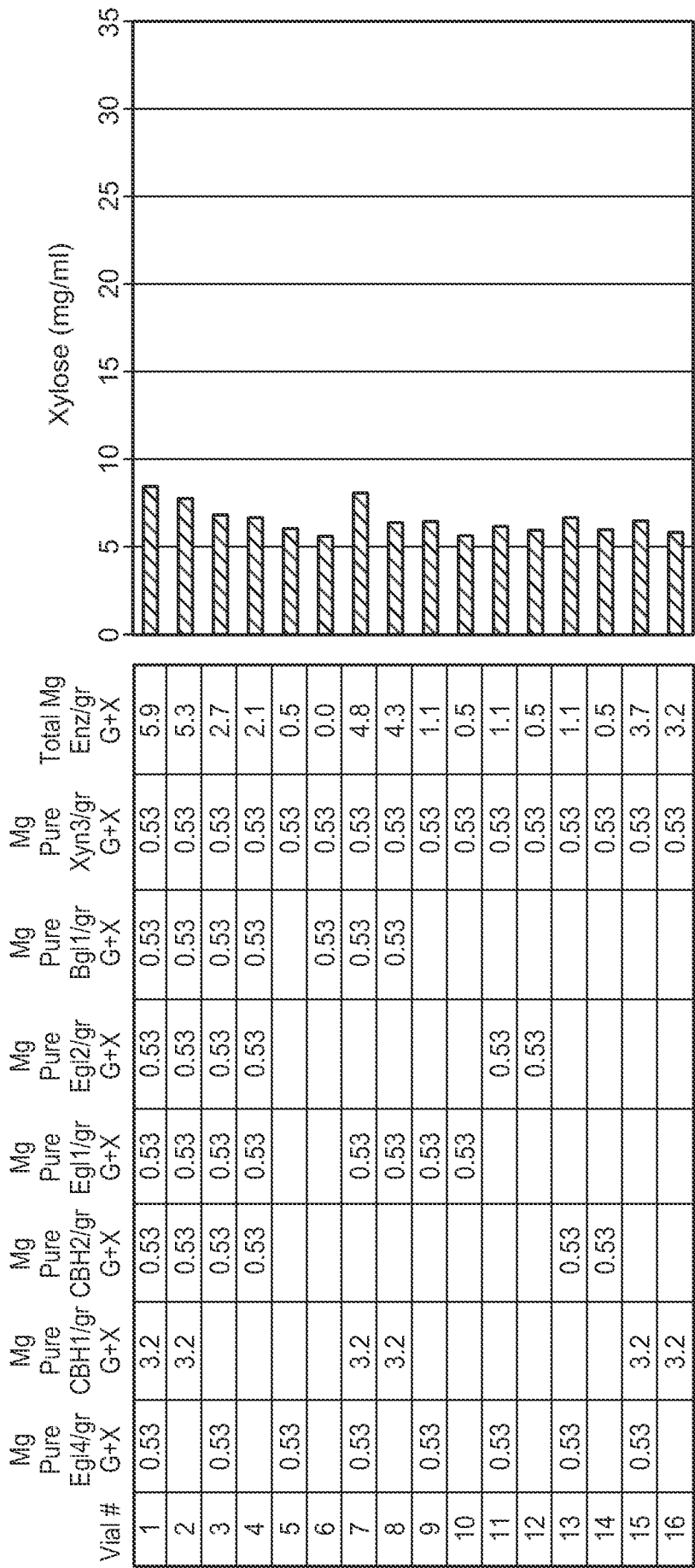
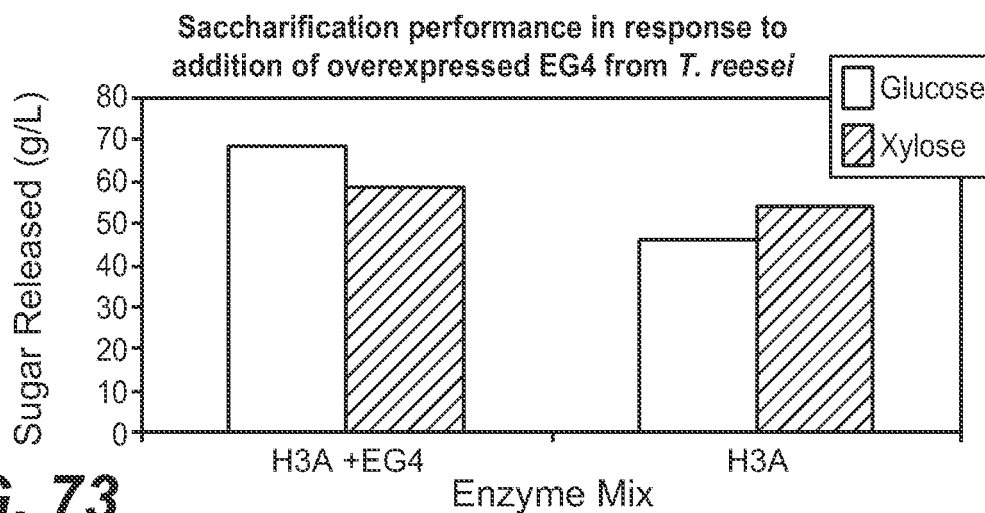


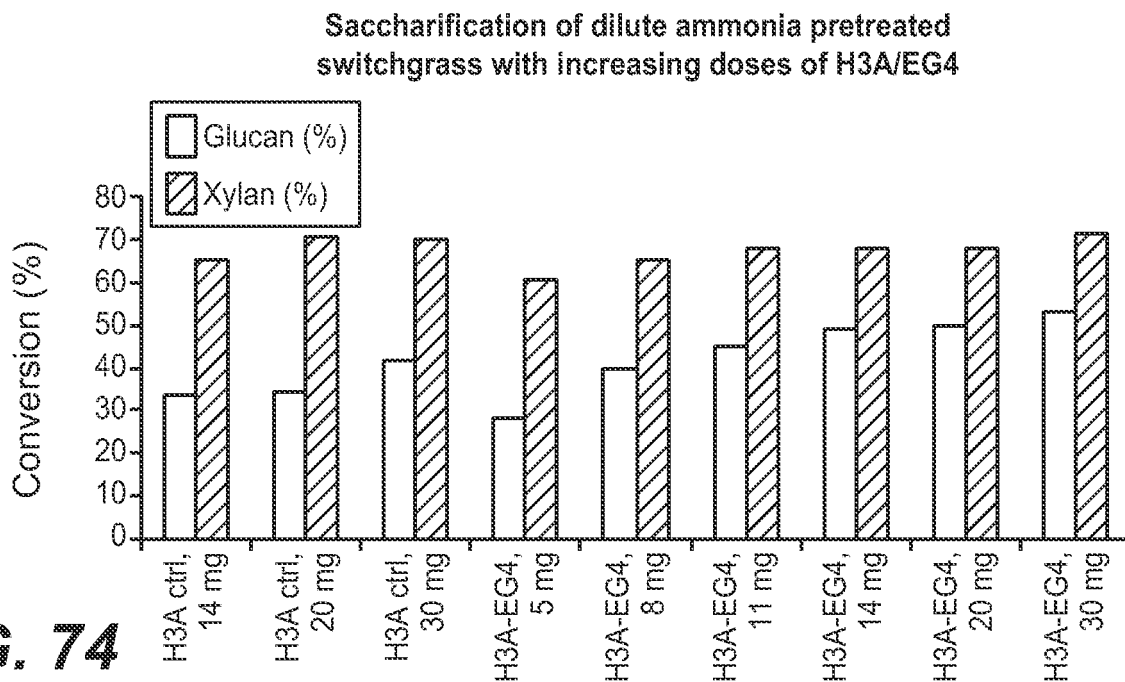
FIG. 72B-3

FIG. 72A-3

96 / 138

**FIG. 73**

Percent glucan and xylan conversion by increasing doses of H3A/EG4#27.

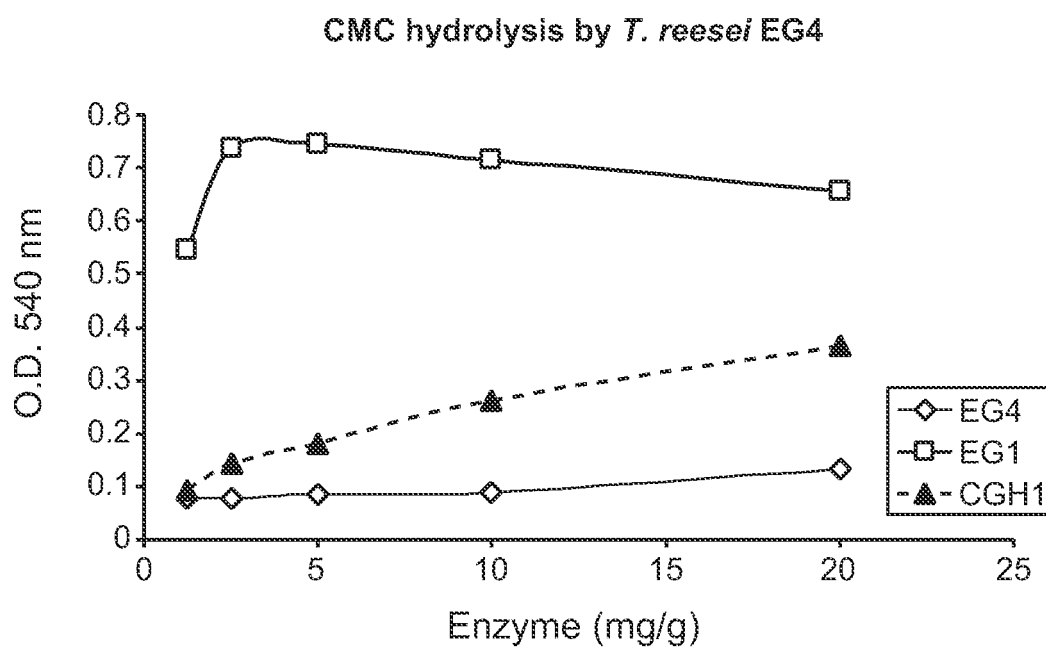
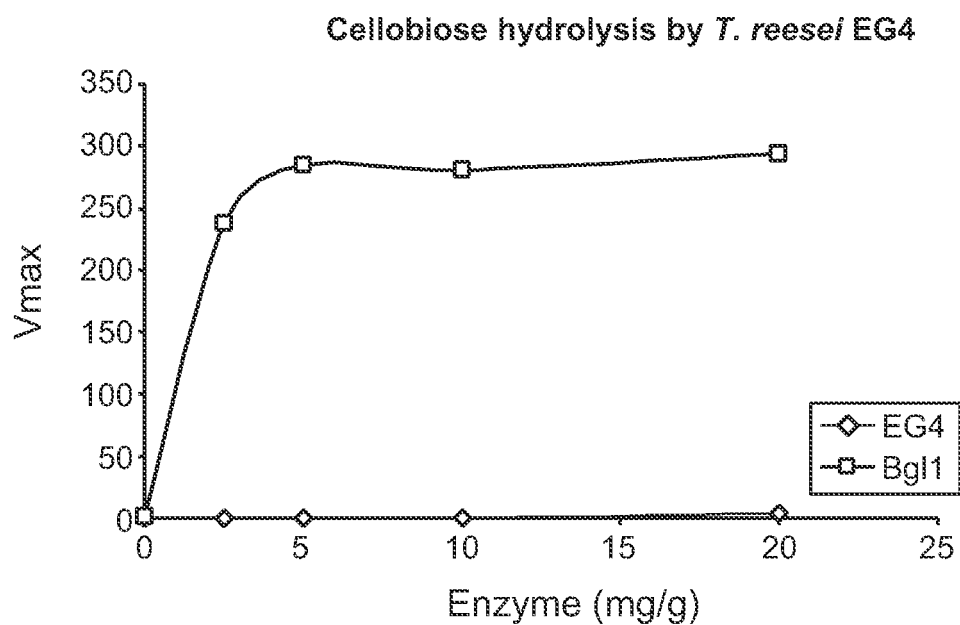
**FIG. 74**

Effect of *T. reesei* EG4 additions on corncob saccharification

Protein Added (mg/g)	CBH1 background (5 mg/g)		Without CBH1 background
	CBH1	EG4	EG4
% glucan conversion			
0	2.7	2.8	2.7
1	3.1	6.6	5.0
2	3.5	7.8	6.9
3	3.4	8.2	7.3
4	3.4	8.8	8.2
5	3.5	7.8	8.8

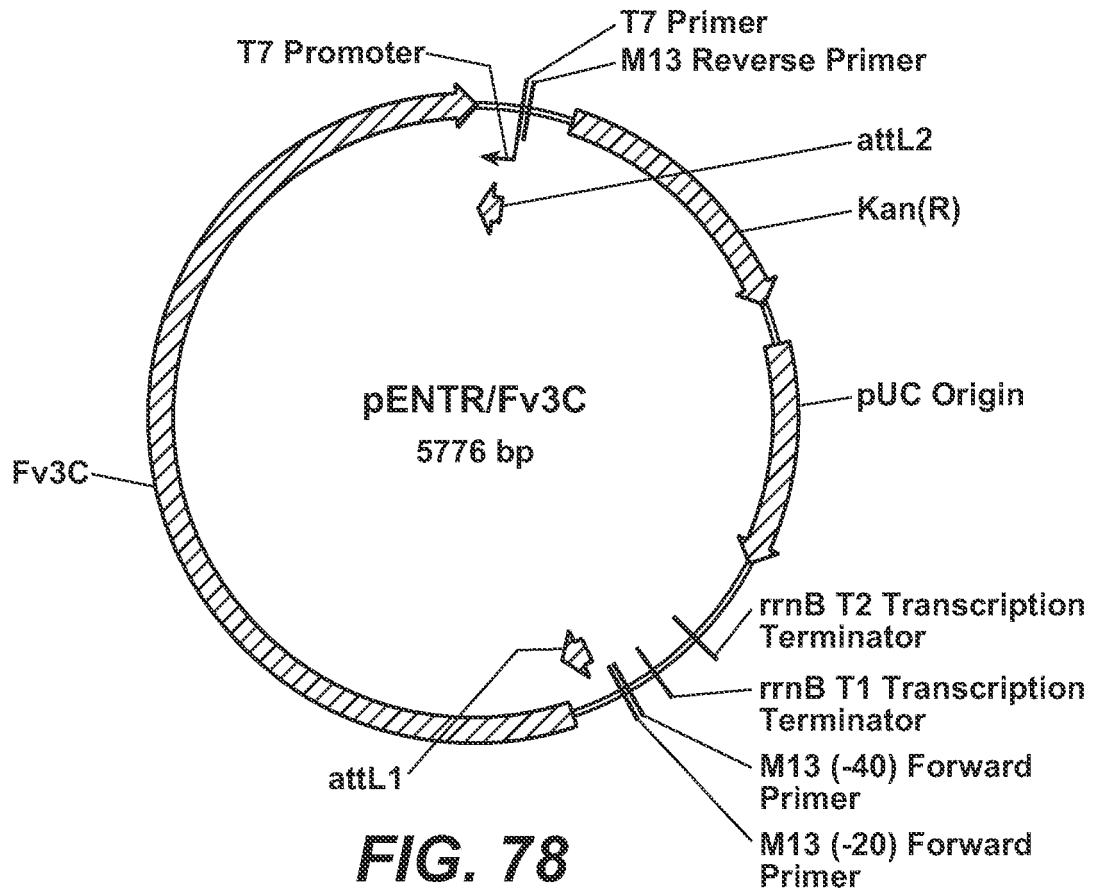
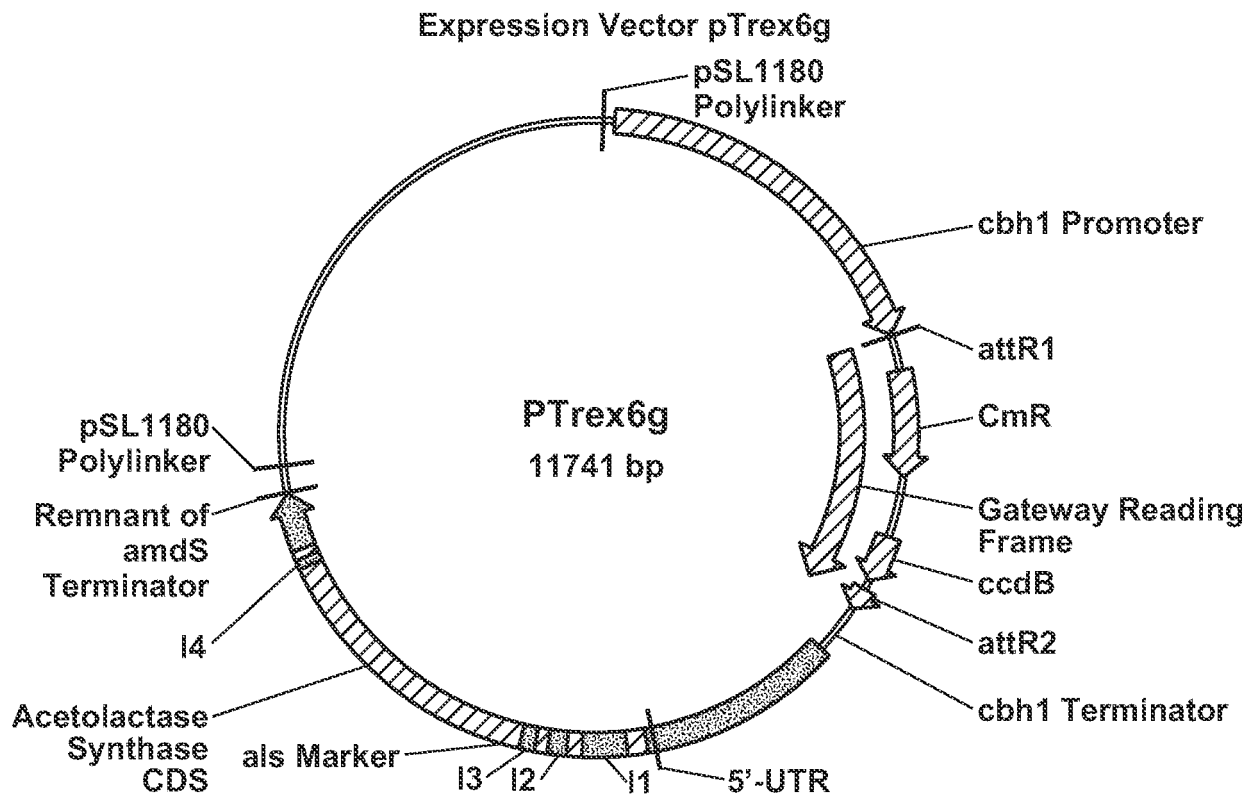
FIG. 75

97 / 138

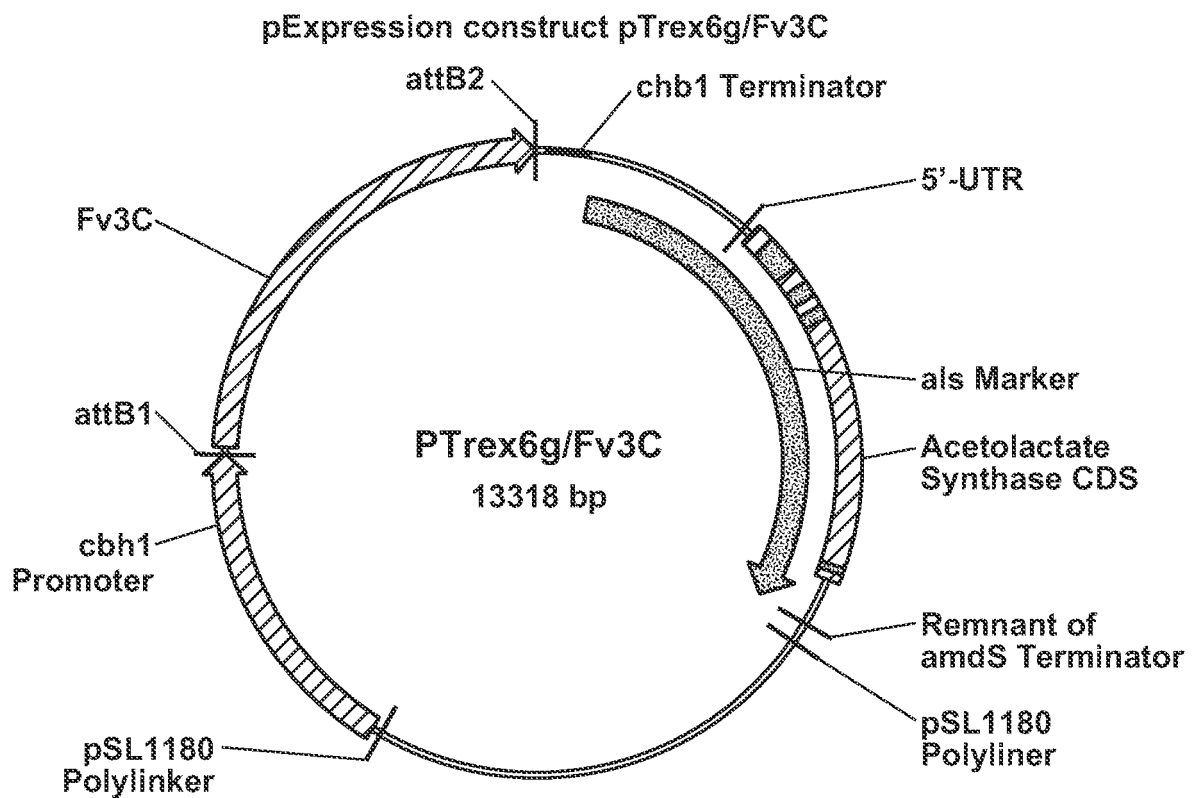
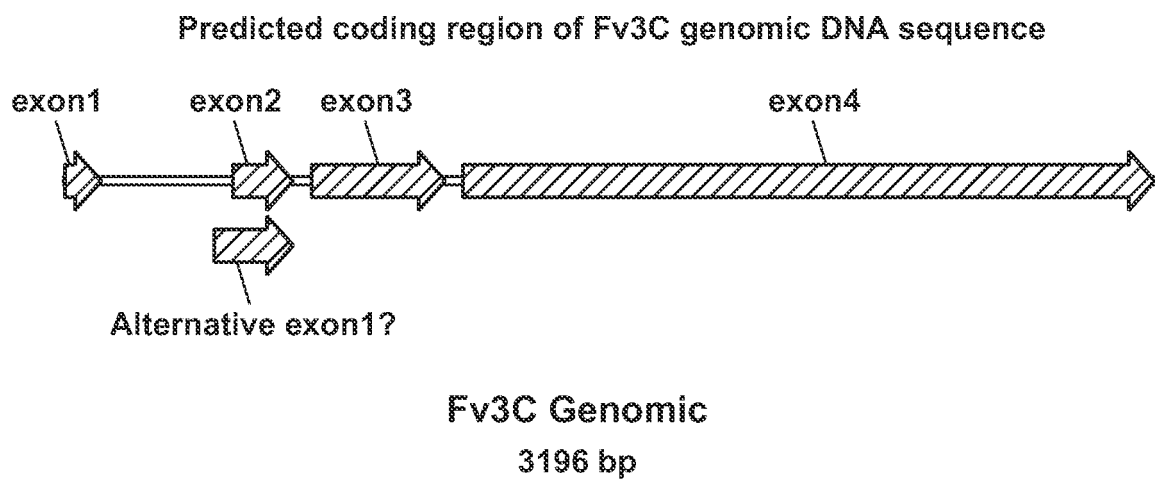
**FIG. 76****FIG. 77**

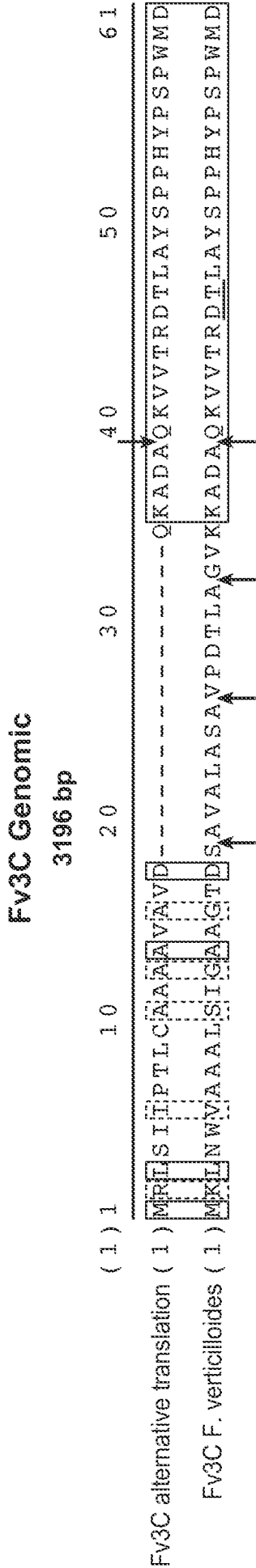
98 / 138

A pENTR/D-TOPO vector with the Fv3C open reading frame

**FIG. 78****FIG. 79A**

99 / 138

**FIG. 79B****FIG. 80**



100 / 138

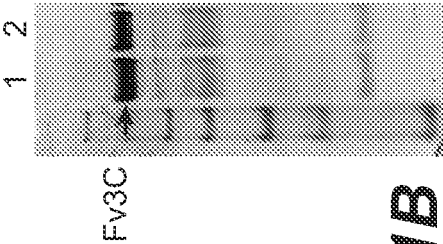
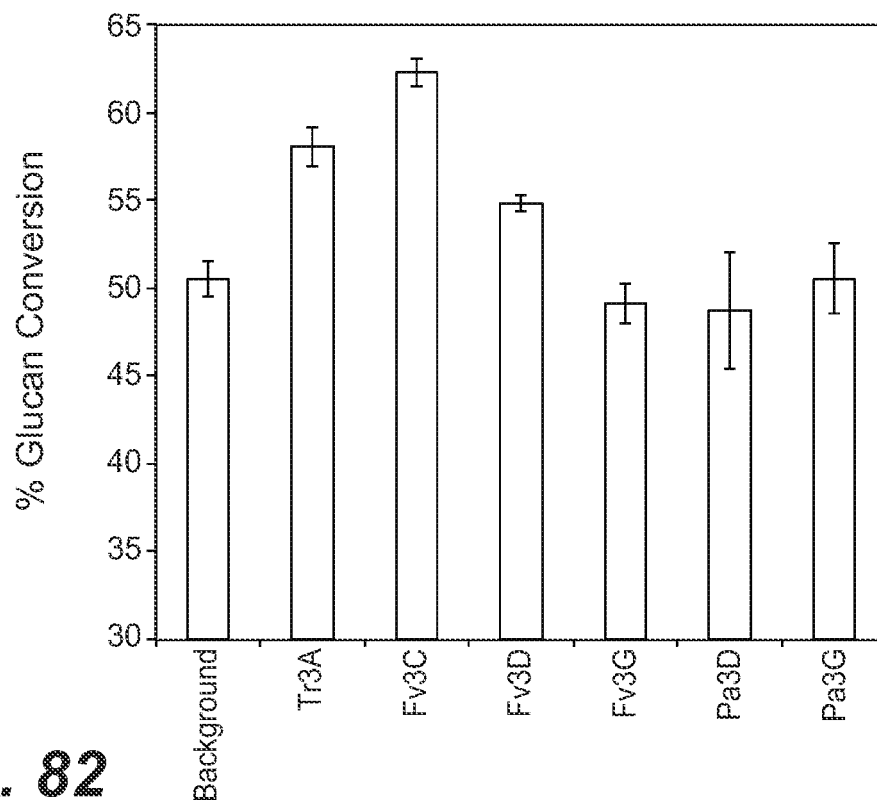
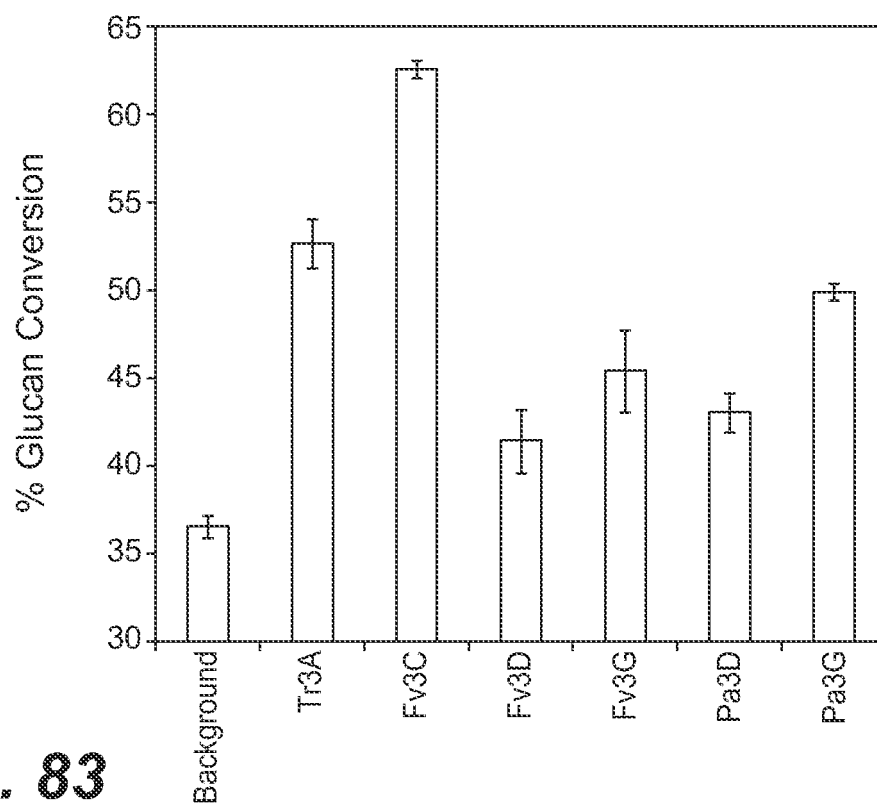
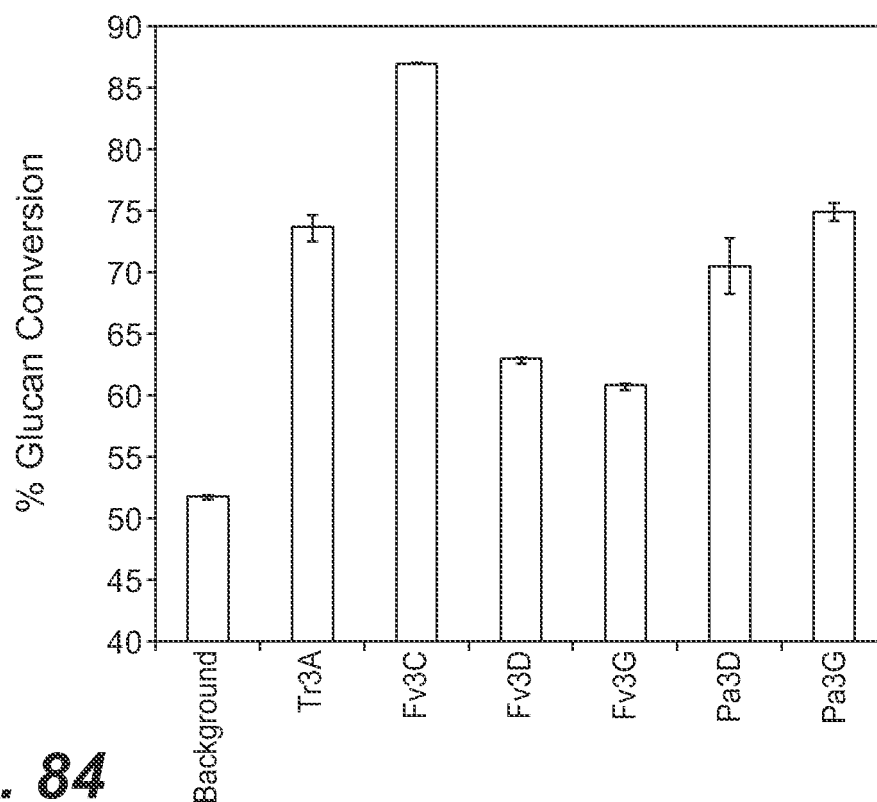
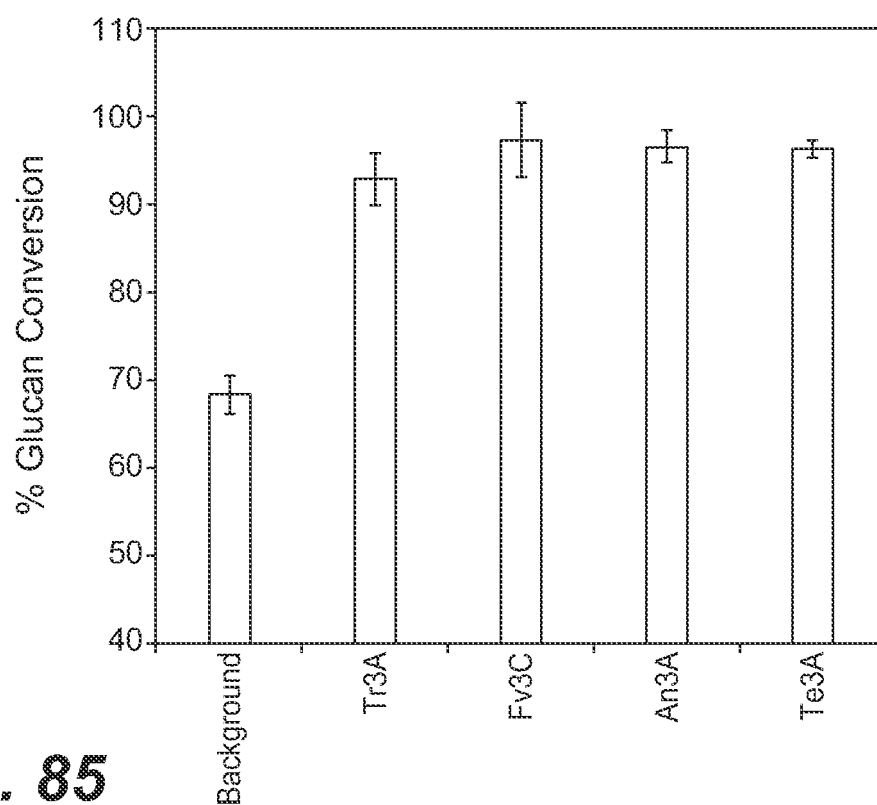


FIG. 81B

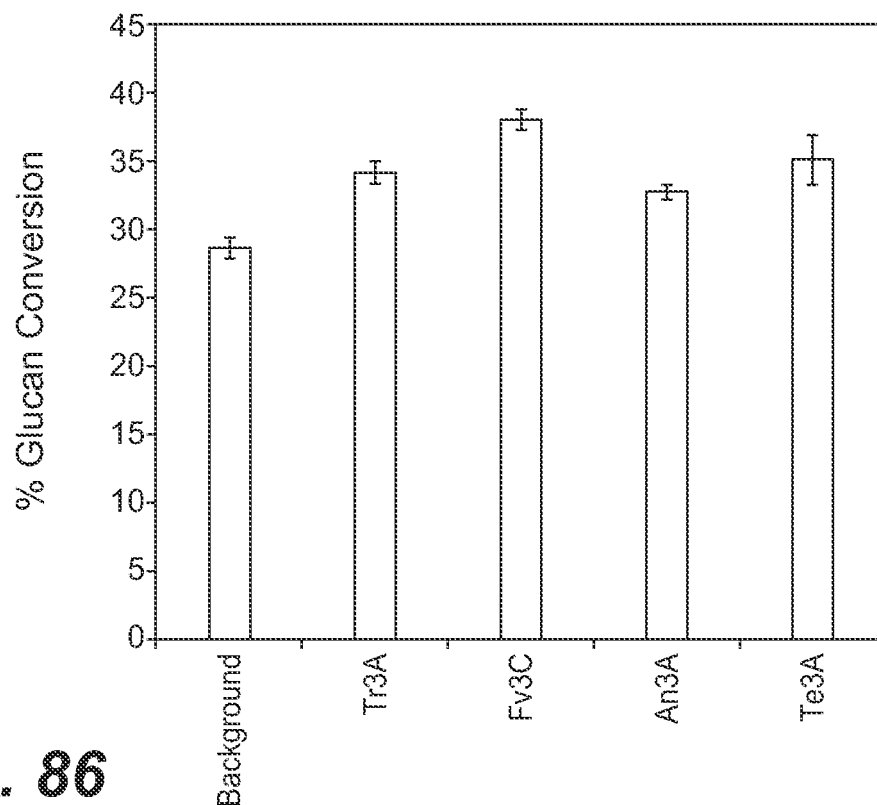
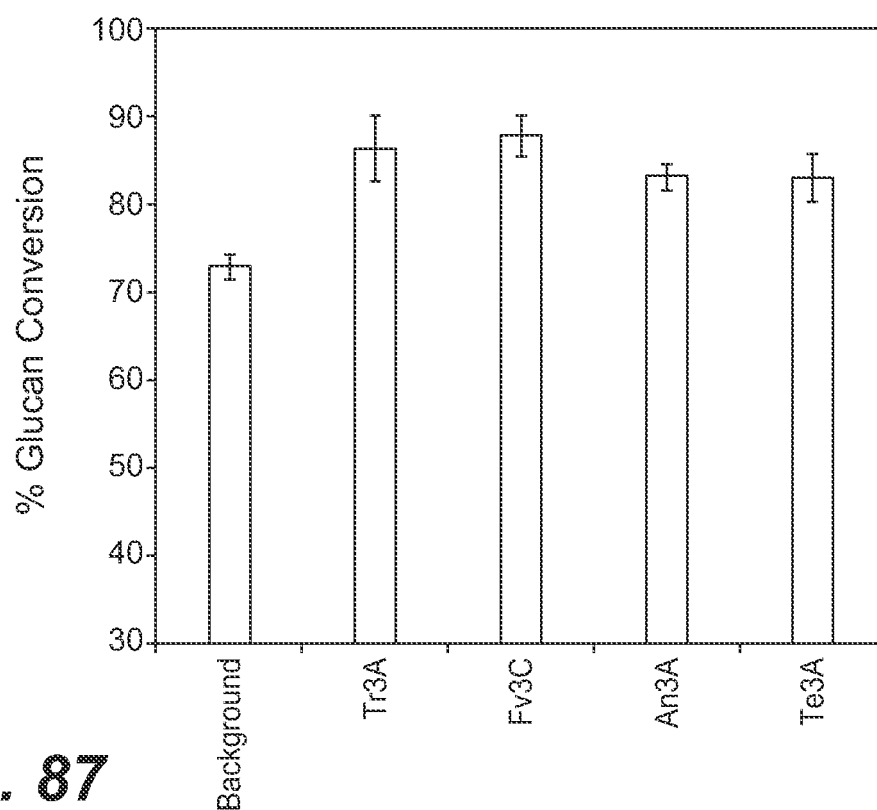
101 / 138

**FIG. 82****FIG. 83**

102 / 138

**FIG. 84****FIG. 85**

103 / 138

**FIG. 86****FIG. 87**

104 / 138

FIG. 88A

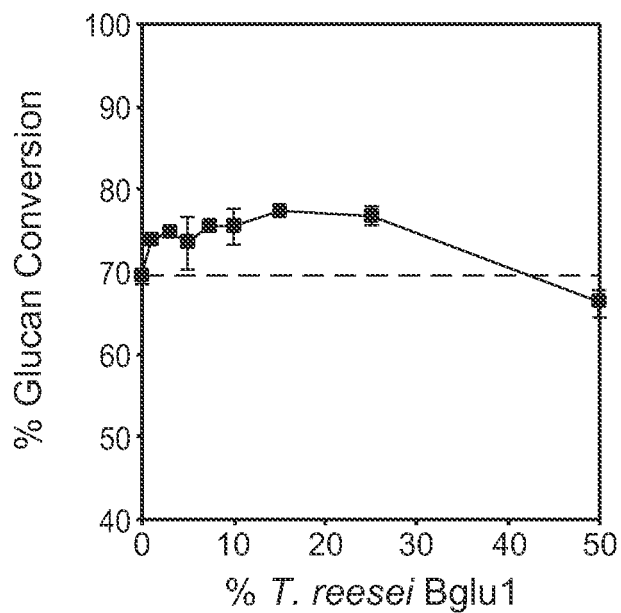


FIG. 88B

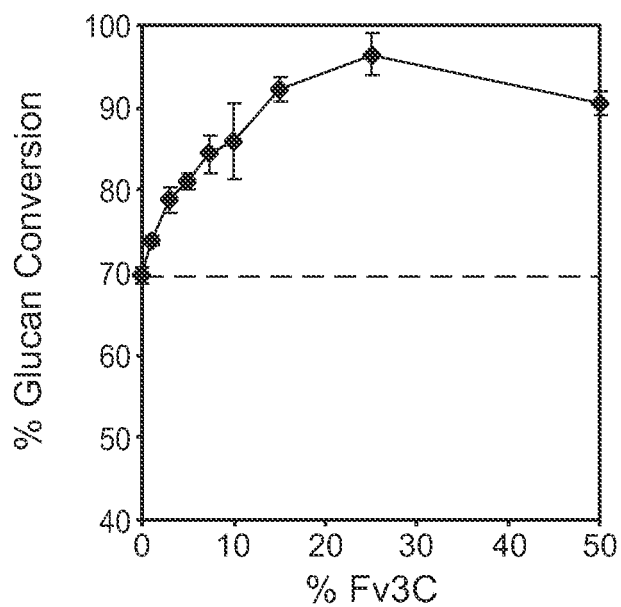
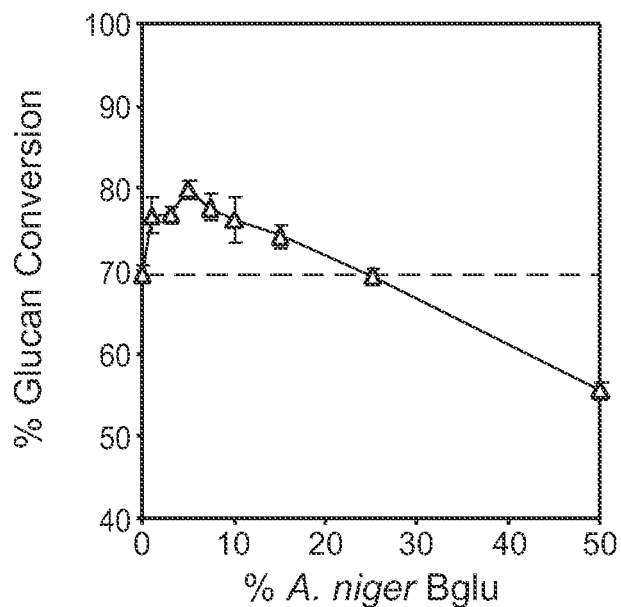
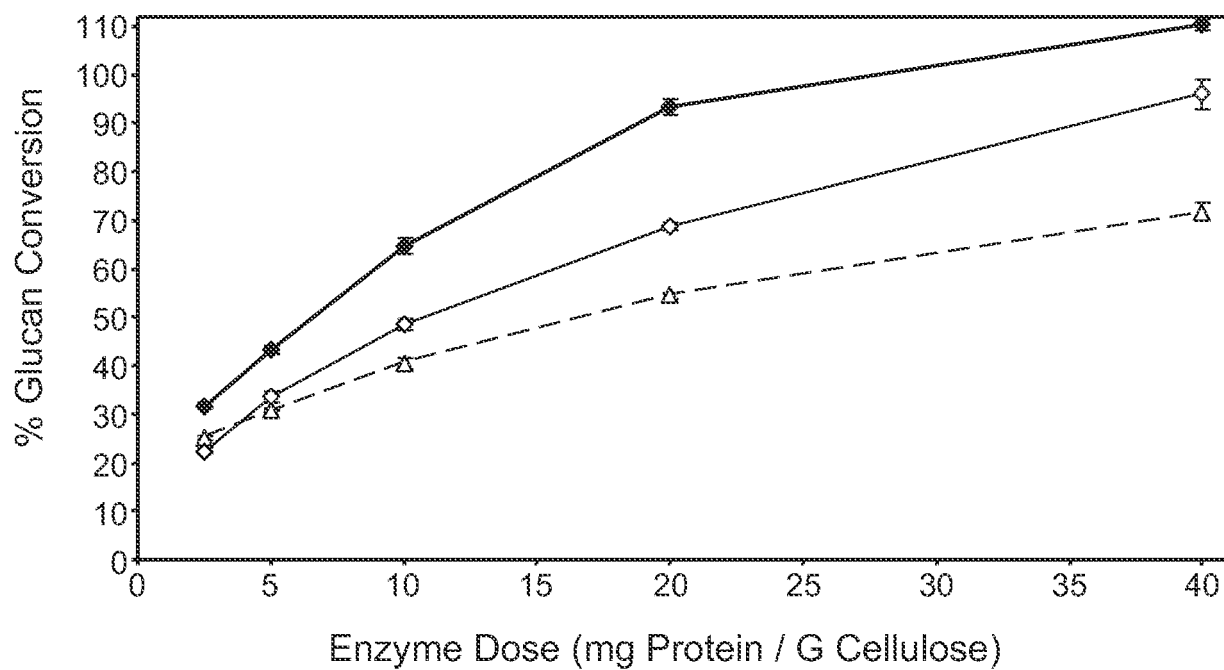


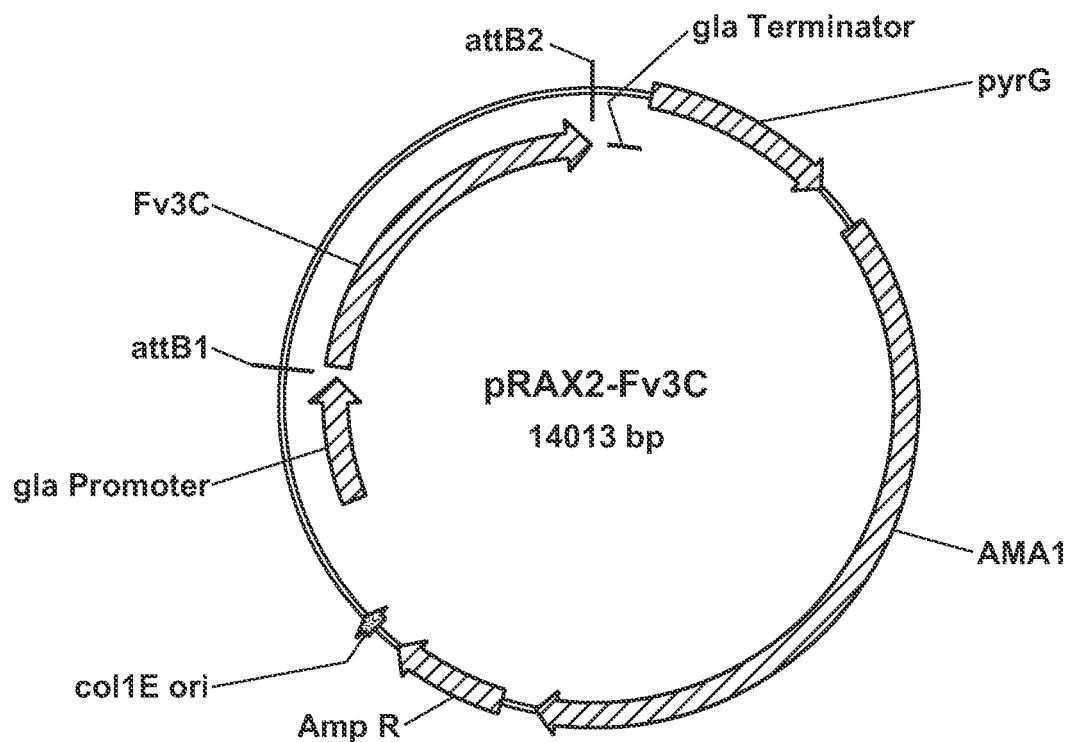
FIG. 88C



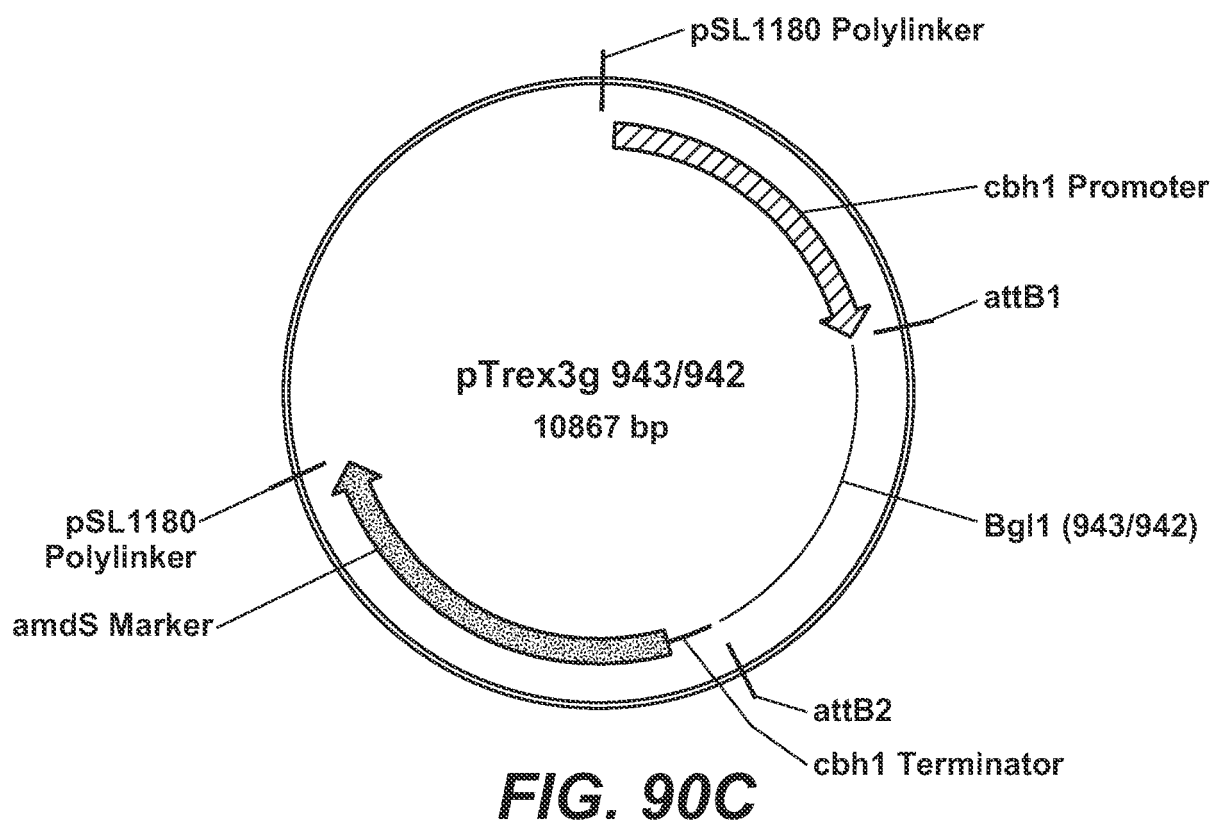
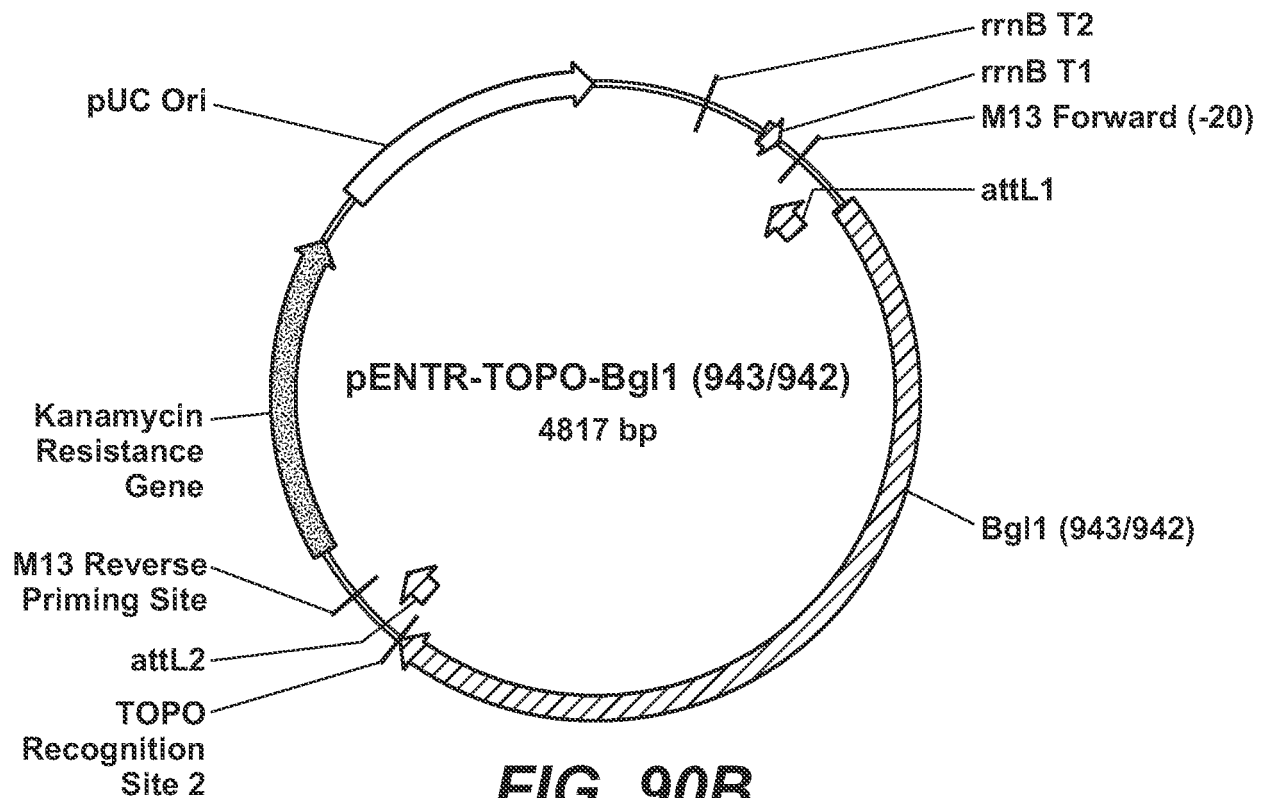
105 / 138

**FIG. 89**

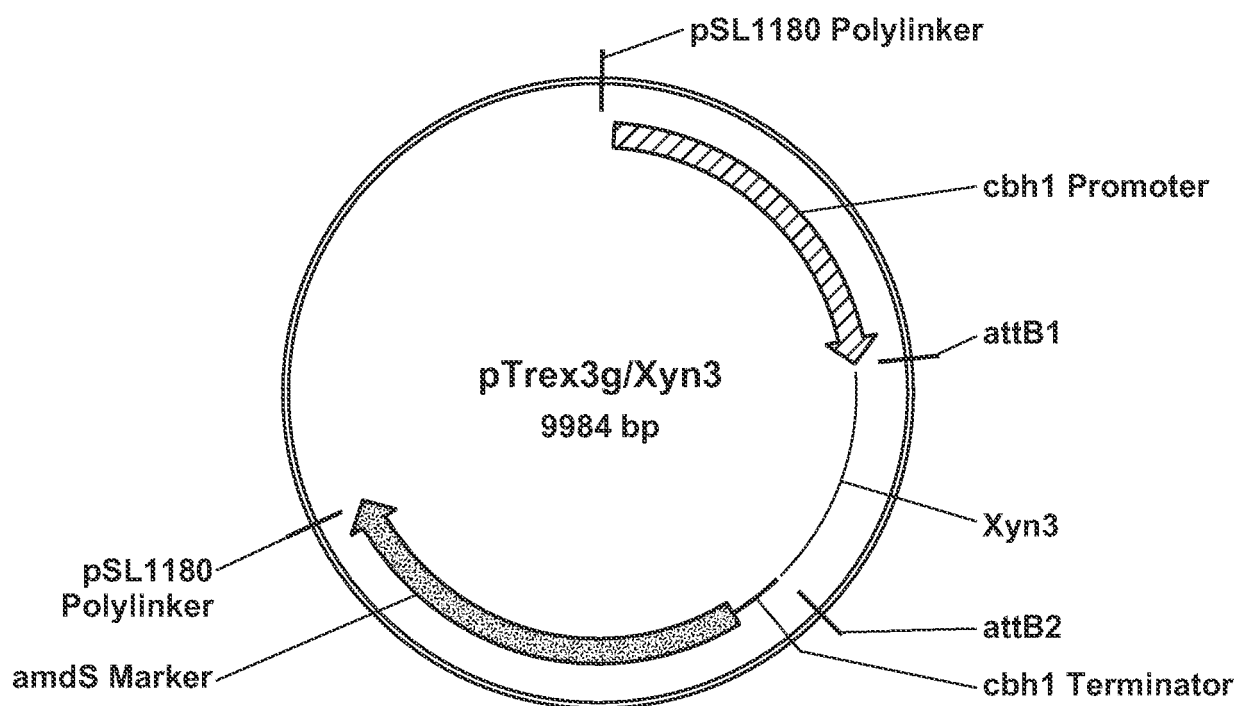
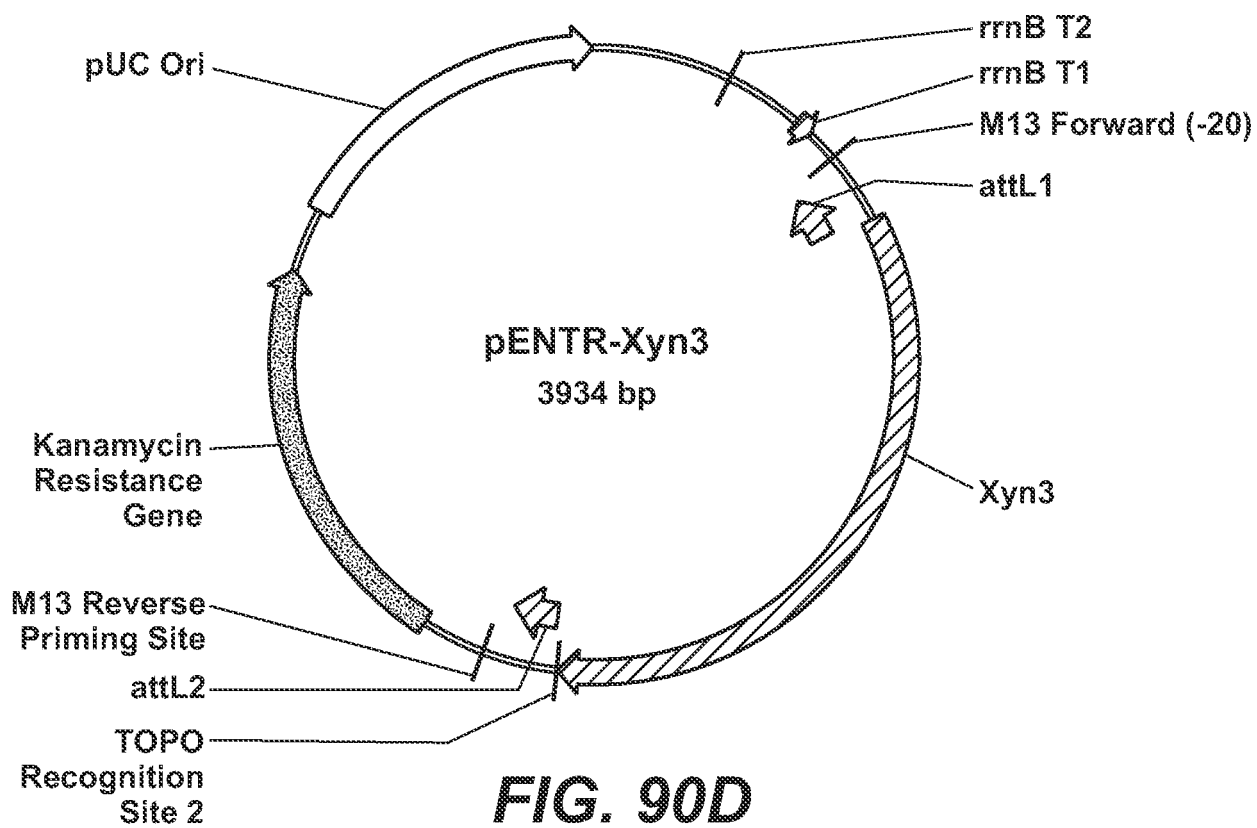
A schematic map of the pRAX2-Fv3C expression plasmid used for expression in *A. niger* var *awamori*.

**FIG. 90A**

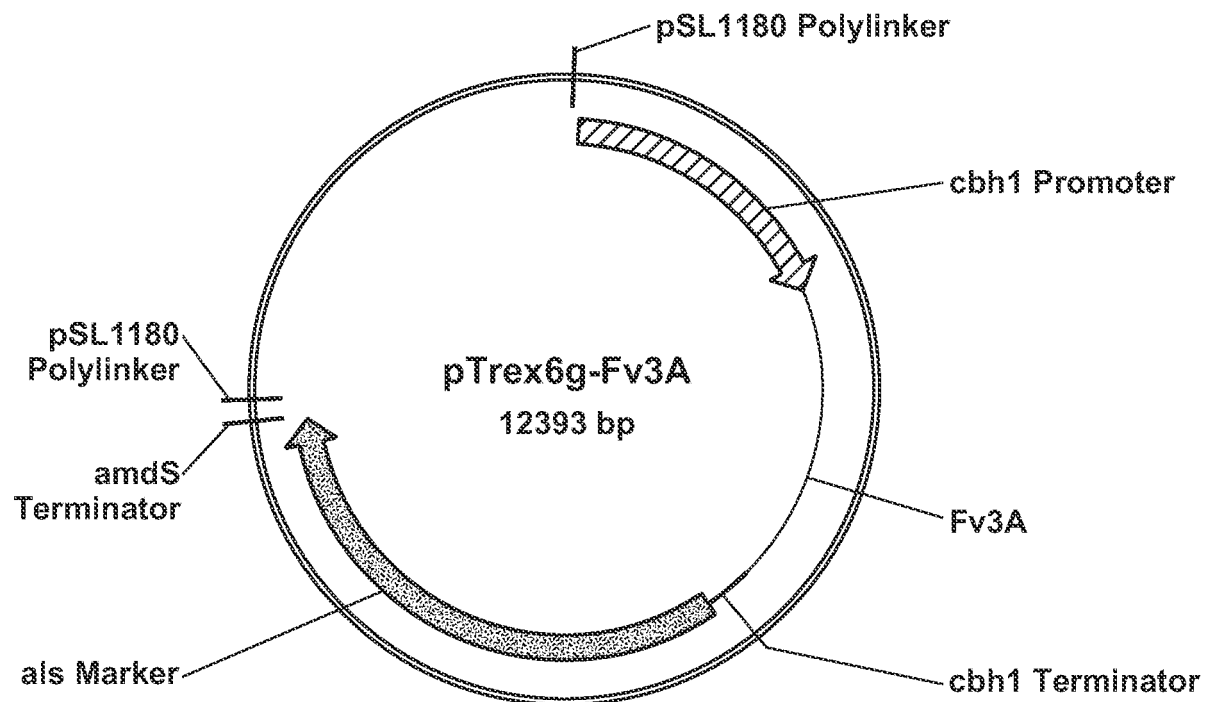
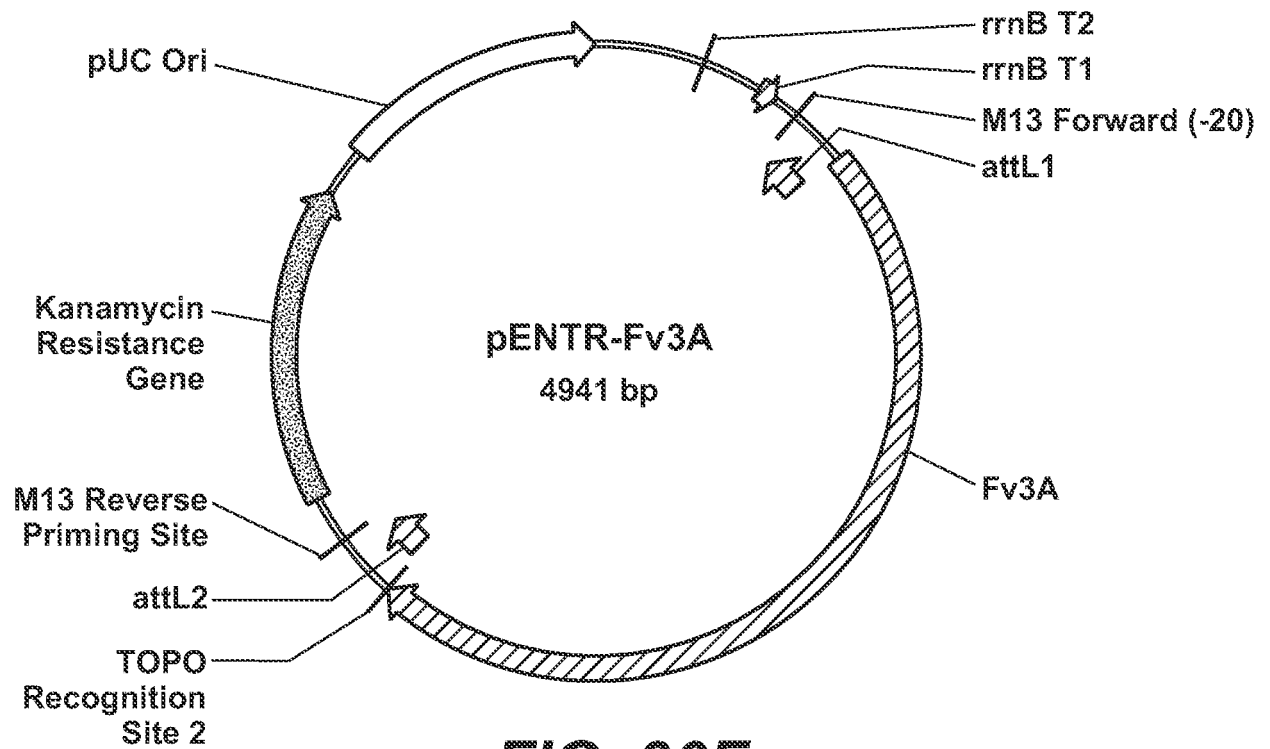
106 / 138



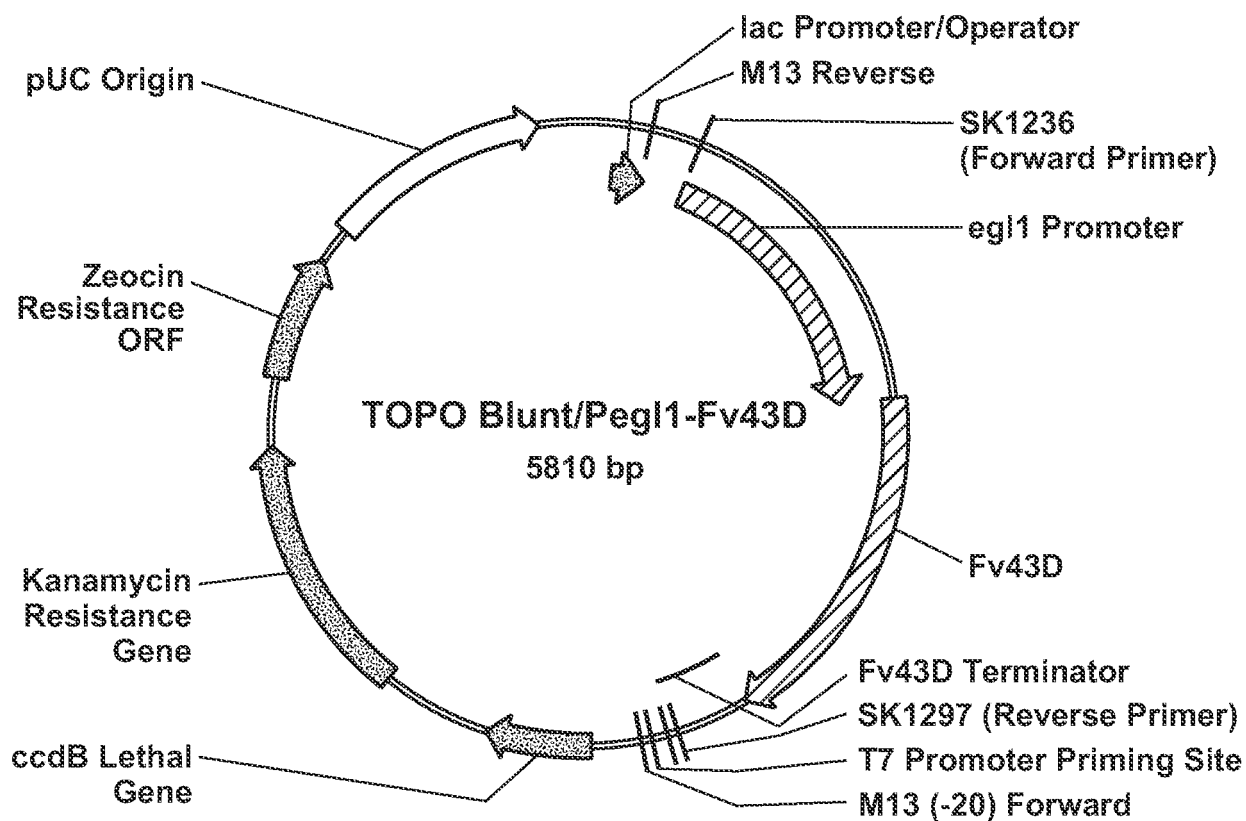
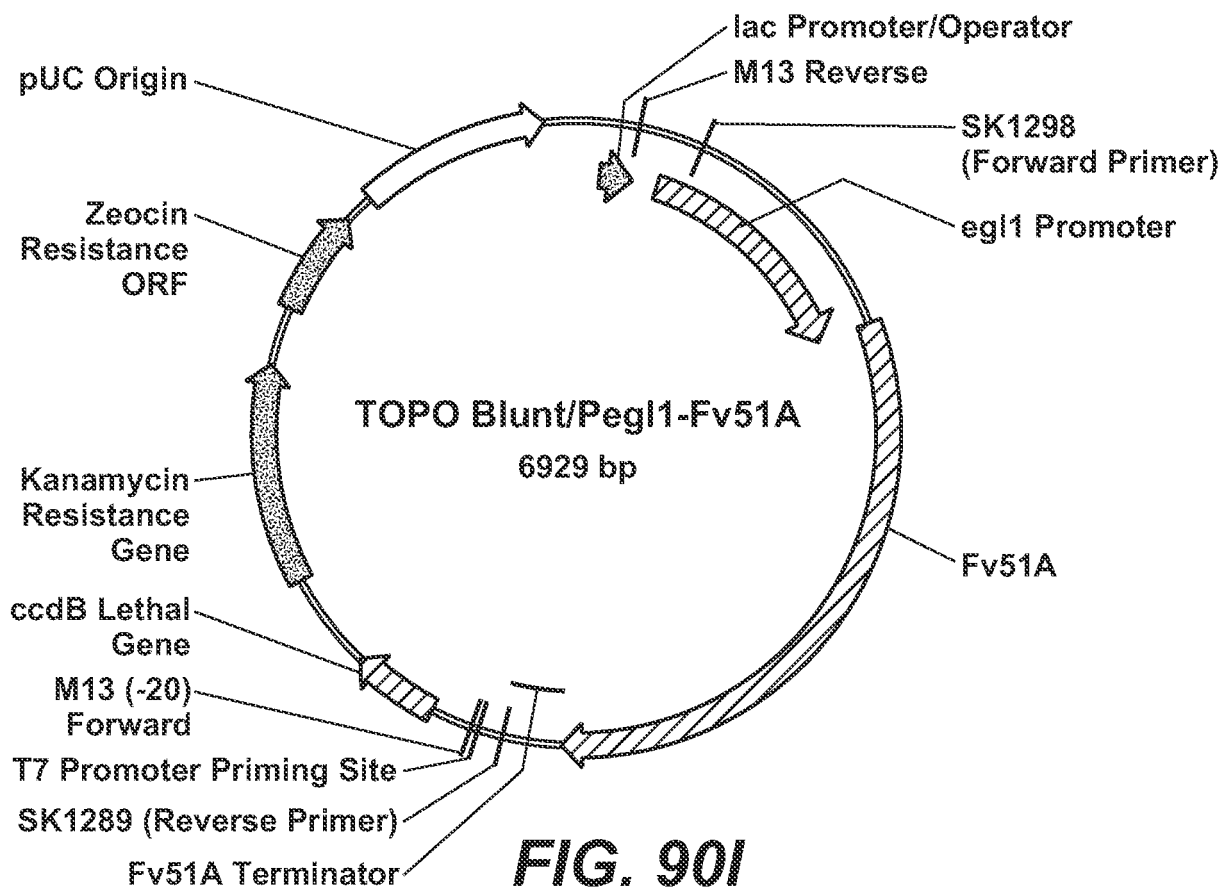
107 / 138



108 / 138



109 / 138

**FIG. 90H****FIG. 90I**

110 / 138

```

gi|reesei|Bxl1      -MVNNAALLAALSALLPTALAQNNQTYANYSAQQQPDLYPETLATLTLSFPDCEHGPLKN
gi|fvert| Fv3A      MLLNLQVAASALSLSLLGGLAEAAATPYT-----LPDCTKGPLSK
                   ::* . :*** * .** :.* :*** :***.:
gi|reesei|          NLVCDSSAGYVERAQUALISLFTLEELILNTQNSGPGVPRGLPNYQVWNEALHGLDRAN-
gi|fvert|          NGICDTSLSPAKRAALVAALTPEEKVGNLVSNATGAPRIGLPRYNWWNEALHGLAGSPG
                   * :*: * . .: ** :*: : * * : * . . . . * .*:***.* : ***** :
gi|reesei|          --FATKGGQFEWATSFPMPILTAAALNRTLHQIADIISTQARAFSNSGRYGLDVYAPNV
gi|fvert|          GRFADTP--PYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEARAFNTGGWRGVDFWTPNV
                   ** . : : *****:* :*: : ***:* . . . . * :***:* * . : : ***
gi|reesei|          NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
gi|fvert|          NPFKDPRWGRGSETPGEDALHVS--RYARYIVRGLEG--DKEQRRIVATCKHYAGNDFEDW
                   * * . * **** .*****:.* * : : * : * * : : .** **:* * :*: *
gi|reesei|          NNQSRLGFDAIITQQDLSEYYTPQFLAAARYAKSRSLMCAYNVNGVPSCANSFFLQTL
gi|fvert|          GGFTRHDFDAKITPQDLAEYYVRPFQECTRDAKVGSIKAYNAVNGIPACANSYLQETIL
                   .. : * .*** ** ***:***. * .: * * * :*****:***:***:***: :*: *
gi|reesei|          RESWGFP-EWGYVSSDCDAVINVENPHDYASNQSSAAASSLRAGTDIDCGQTYPWHLNE
gi|fvert|          RGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAEGAQVAFENGMDSSCEYTTTSDVSDS
                   * * . . : : : :***.* : : : * .* . . . . * : : . * * . * . . . : *
gi|reesei|          FVAGEVSRGEIERSVTRLIANLVLGYFD-KKNQYRSLGWKDVKTDAWNISYEAAVEGI
gi|fvert|          YKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAQAQWNSLSFADVNTKEAQDLALRSAVEGA
                   : * : : . :*: :*: .*: :*** * * .*: : * . : * : : .:****
gi|reesei|          VLLKNDGTLPLSKKVR-SIALIGPWANATTQMGGNYYPAPYLISPLEAAKKAGYHVNF
gi|fvert|          VLLKNDGTLPLKLLKKDSVAMIGFWANDTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVA
                   ***** . * : * :*: * * * :*: * * * : * * * * : * * . . .
gi|reesei|          LGTEIAGNSTTG--FAKAIAAAKKSDAILYLGIDNTIEQEGADRTDIAWPGNQDLIKQ
gi|fvert|          WGPTLQNSSSHDNWTTNAVAAAKKSDYILYFGGLDASAAGEDRDRENLDWPESQLTLLQK
                   * . : .*: . :*:***** * :*: * : * . ** : : ** .** * : :
gi|reesei|          LSEVGKPLVVLQMGQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRL
gi|fvert|          LSSLGKPLVVIQLG-DQVDDTALLKNNKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRL
                   **.:*****:* * .***.:* .***:***:* .****.*.*.:*:*:*:*:*****
gi|reesei|          VTTQYPAEYVHQFPQNDMNLRPDGKSNPGQTYIWYTGKPVYEFSGSLFYTTFKETLASHP
gi|fvert|          PVTQYPSKYTEQIGMTDMDLRPT-KSLPGRTRYRWYS-TPVLPYGFGLHYTKFQAKFKSN-
                   .****:*. .*: .*:*** ** **:* * : .** : * **.*.*.: : * :
gi|reesei|          KSLKENTSSILSAPHPGYTYSEQIPVFTFEANIKNSGKTESPYTAMLFVRTSNAGPAPYP
gi|fvert|          -KLTFDIQKLLKG--CSAQYSDTCALPPIQVSVKNTGRITSDFVSLVFIKS-EVGPKPYP
                   .*.*: .*: * . ** : : . . . .*: * : * :*: * : : .** ***
gi|reesei|          NKWLVGFDRLADIKPGHSSKLSIPIPVSALARVDSHGNNRIVYPGKYELALNTDESVKLEF
gi|fvert|          LKTLAAYGRLHDVAPSSTKDISLEWTLDNIAARRGENGDLVVYPGTYTLLLEPTQAKIQV
                   * * . . . * * : * . . . : * . . : * * . : * : * * * : * . : .
gi|reesei|          ELVGEEVTIENWPLEEQQIKDATPDA
gi|fvert|          TLTGKKAILDKWPQDPKSA-----

```

FIG. 91

111 / 138

```

Fv39A      : MHYATLTTLVLALTNNVAAQQGTATVDLSKNHGPAKALGSGFIYGWPDNGTSVDTSI PDF : 60
Q9ZFM2_XYN : -----MKVVNVPSNGREKFKKNW--KFCVGTGRL-GLALQKEYLDH--LKL : 41
P36906_XYN : -----MIKVRVPDSDKKFSDRW--RYCVGTGRL-GLALQKEYIET--LKY : 41

Fv39A      : LVTDIKENSNRGGGAQIPSLGWARGGYEG-----YLGR-FNSTLSNYRTRKYNA : 109
Q9ZFM2_XYN : VQEKIGFRIIRGHGLLSDDVGIYREVEIDGEMKPFYNFTYIDRIVDSYLA--LNIRPF-I : 98
P36906_XYN : VKENIDFKYIRGHGLLSDDVGIYREDVVGDEVKPFYNFTYIDRIFDSFLE--IGIRPF-V : 98

Fv39A      : DFILLPHDLWGADGG---QGSNSPFFGDNGNWTEM-ELFWNQLVSDLKAHNMLEGLVID : 164
Q9ZFM2_XYN : EFGEMPKALASGDQTVEYWKGNVTP-PKDYNKWRDLIVAVVSHFIERYGIEEVRT-WLFE : 156
P36906_XYN : EIGFMPKKLASGTQTVEYWEKNVTP-PKDYEKWSDLVKAVLHHFISRYGIEEVVK-WPFE : 156

Fv39A      : A * 200 * 220 * 240
Q9ZFM2_XYN : VWNEPDIDIEF--DRPWSQFLEYYNRATKLLRKTLPKTL LSGPAMAHSPILSDDKWHTWLQ : 223
Q9ZFM2_XYN : VWNEPNLVNFWKDKANKQEFYFKLYEVTARAVKSVDPHLQVGGAICGGS---DEWITDFL : 212
P36906_XYN : IWNEPNLKEFEWKDADEKEYFKLYKVTAKAKEVNENLKVGGPAICGGA----DYWIEDFL : 212

Fv39A      : SVAGNKTVP--DIYSWHQIGAW-----EREP-DSTIPDFTTLRA---QYGVF : 264
Q9ZFM2_XYN : HFCAERRVPVDFVSRHAYTSKAPHKKTFEYYYQELELEPPEDMLEQFKTVRALIRQSPFP : 272
P36906_XYN : NFCYEENVPVDFVSRHATTSK-QGEYTPHLIYQ--EIMPSEYMLNEFKTVREIKNSHFP : 269

Fv39A      : N * 320 * 340 * 360
Q9ZFM2_XYN : EKPIDVNEYAARDEQNPAHSVYLLSQLERHNLRLGLRANWGSGLHNMGNLIYSTTGTS : 324
Q9ZFM2_XYN : HLPFHITEYNT--SYSPINPVHDTALNAAAYIARILSEGDDYVDSFSYWTFSDFVEEMDVP : 330
P36906_XYN : NLPFHITEYNT--SYSPQNVPVHDTPFNAAYIARILSEGDDYVDSFSYWTFSDFVEERDVP : 327

Fv39A      : * 380 * 400 * 420
Q9ZFM2_XYN : EGTYY-----PNGEWQAYKYAAMAGQRL-----VT : 350
P36906_XYN : KALFHGGFGLVALHSIPKPTFHAFTHFTFALGDELLYRDGEMIVTRRKDGSIAAVLWNLVM : 390

Fv39A      : KASSDLKFDVFATKQGRKIKI IAGTRTVQAKYNIKISGLEVAGLPMKMGTVKVRTYRFDWA : 410
Q9ZFM2_XYN : EKGEGLTKEVQLVIPVSFSAVFIKRIQIVNEQYGNWVWKQMCPRFPRPSQAVETLPS-A : 449
P36906_XYN : DKTENPDEDEYEVEIPVFRFRDVFIKRQLIDEEHGNPWGTWIHMGRPRYPSKEQVNTLREVA : 447

Fv39A      : * 500 * 520 *
Q9ZFM2_XYN : GP-----NGKVDGPPVDLGEKKYT-YSANTVSSPST----- : 439
P36906_XYN : QPHVMTQORRATDGVIHLSIVLSKNEVTLIEIEQVRDETSTYVGLDDGEITSYSS : 504
P36906_XYN : KPEIMTSQPVANDGYLNLKFKLGKNAVVLVELTERIDESSTYIGLDDSKINGY--- : 500

```

FIG. 92

112 / 138

Fv43D ---MQLKFLSSALLSLTGNCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Fo43A ---MQLKFLSSALLSLTSCCAAQDTNDIPPLITDLWSADPSAHVFEGKLWVYPSHDIEA
Gz43A ---MKSLLFP---LLSFVG---QSLATNDDCPLITSRWTADPSAHVFENDTLWLVP SHDIDA
Pf43A --MLQRFAYILPLALLSVG--VKADN----PFVQSIYTADPAPMVYNDRVYVFMHDHNTG
Fv43A --MWLTSPLLFASTLLGLTGVALADN-----PIVQDIYTADPAPMVYNGRVYLFTGHDNDG
Fv43B --MRFSWLLCPLLAMGSALPETKTDVSTYTNFVLPGWHSDPSC-IQKDGLFLCVTSTFTIS
Af43A -----MAAPSLSYPTGIQSYTNPLFPGWHSDPSCAYVAEQDTFFCVTSTFTI
Pf43B ----MSRSILPYASVFALLGGATAEP----FLVLNSDFPDPSLIETSSGYAFGTTGNGV
Fv43E MKVYWLVAWATSLTPALAGLIGHRRATTENNPIIYSDFPDNDVFLGPDNYYYFSASNHFH

Fv43D NVVNGTGGAQYAMRDYHTYSMKTIYKDPVIDHGVALSVDVDPWAKQQMWAPDAAYK--N
Fo43A NVVNGTGGAQYAMRDYHTYSMKSIYKDPVVDHGVALSVDVDPWAKQQMWAPDAAHK--N
Gz43A GFENDPDGGQYAMRDYHVYSIDKIYGLP-VDHGTALSVEVDPWASRQMWAPDAAHK--N
Pf43A -----ATYYNMTDWHLFSSADMANWQD---HGIPMSLANFTWANANAWAPQVPR--N
Fv43A -----STDENMTDWRLFSSADMVNWQH---HGVMSLKTFSWANSRAWAGQVVAR--N
Fv43B FP----GLPVYASRDLVNWRLISHVWNRE---KQLPGISWKTAGQQQCMYAPTIRYH--K
Af43A AFP---GLPLYASRDLQNWKLASNIFNRP---SQIPDLR-VTDGQQSGIYAPTLRYH--E
Pf43B N-----AQVASSPDFNTWTLLSGT-----DALPGPFPSWVASSPQIWAPDVLVKA-D
Fv43E SP----GAPVLKSKDLLNWDLIGHSSIPRLNFGDGYDLPPGSRYYRG-GTWASSLRYRKS

Fv43D GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTNGE
Fo43A GKYYLYFPAK-DK-DEIFRIGVAVSNKPSGPFK---ADK-SWIPGTYSIDPASYVDTDNE
Gz43A GKYYLYFPAK-DK-DDIFRIGVAVSPTPGGPFV---PDK-SWIPHTFSIDPASFVDDDDR
Pf43A GQFYFYAPVR-HN-DGSMAIGVGVSTITGPYH---DAIGKPLVENNEIDPTVFIDDDGQ
Fv43A GKFFYFVPVRNAK-TGGMAIGVGVSTNILGPYT---DALGKPLVENNEIDPTVYIDTDGQ
Fv43B GTYYVICEYLGVG-DIIGVIFKTTNPWDESSWS---DPV---TFKPNHIDPDLEWDDDGK
Af43A GQFYLVSYLGP--QTKGLLFTSSDPYDDAAWS---DPL---EFAVHGIDPDIFWDHDTGT
Pf43B GTYVMYFASASAASDSGKHCVGAATATSPEGPYTPVDSAVACPLDQGGAIDANGFIDTDGT
Fv43E GQWYWIGCIN-----FWQTWVYTASSPEGPWY---NKGNFGDNNCYYDNGILIDDDDT

Fv43D AYLIWGGI-WGGQLQAWQDHKTFNESWLGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Fo43A AYLIWGGI-WGGQLQAWQDKKNFNESWIGDKAAPNGTNALSPQIAKLSKDMHKITETPRD
Gz43A AYLAWGGI-MGGQLQRWQDKNKYNES--GTEPG-NGTAALSPQIAKLSKDMHTLAEKPRD
Pf43A AYLYWG-----NPDWLWYVKLNQDMISYSGSPTQ
Fv43A AYLYWG-----NPGLYYVKLNQDMLSYSGSINK
Fv43B VYCATHG----ITLQEIDLETGELSPELNIWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Af43A VYVTSAD-QMIKQYTLDLKTGAIGPVDYLWNGTGGVWPEGPHIYKRDGYYYLMIAEGGT
Pf43B IYVVYKID-----GNSLDGDGTTHPTPIMLQQMEADGT
Fv43E MYVVYGSGEVKVSQLSQDGFSSQVKSQVVFKNNTDIGVQDLEGNRMYPKING-----LYYI

Fv43D LVILAPETGKPLQAEDNKRRFFEGP-----WVHKRGKLYYLMYSTG-----
Fo43A LVILAPETGKPLQAEDNKRRFFEGP-----WIHKRGKLYYLMYSTG-----
Gz43A MLILDPKTGKPLLEDEDRRFFEGP-----WIHKRNKIYYLTYSTG-----
Pf43A IPLTTAGFGTRTGNAQRPTTFEEAP-----WVYKRNGIYYIAYAAD-----
Fv43A VSLTTAGFGSRPNNAQRPTTFEEGP-----WLYKRGNLYYMIYAAN-----
Fv43B ----AEDHAITIAARKITGPYEAYNNNPILTNRGTSEYFQTVGHGDLFQDTKGNWWGLC
Af43A ----ELGHSETMARSRTGTGPWEPYPHNPLLSNKGTSEYFQTVGHADLFQDGNGNWWAVA
Pf43B --TPTGSPILQIDRSDLDGPLIEAP-----SLLSNGIYYLSFSSN-----
Fv43E LNDSPSGSQTWIWKS KSPWGPYESKVLADKVTPPISGGNSPHQGS LIKTPNGGWY-----

Fv43D -DTHFLVYATSKN---IYGPYT-----YQGKILDPVDG-----WTTHG
Fo43A -DTHFLVYATSKN---IYGPYT-----YRGKILDPVDG-----WTTHG
Gz43A -TTHYL VYATSKT---PYGPYT-----YQGRILEPVDG-----WTT HS
Pf43A CCSEDIRYSTGTS---ATGPWT-----YRGVIMPTQGSS-----FTNHE
Fv43A CCSEDIRYSTGPS---ATGPWT-----YRGVVMNKAGRS-----FTNHP
Fv43B LATRITAQGVSPMGREAVLFNGTWNKGEWPKLPVRGRMPGNLLPKPTRN-----VPGD
Af43A LSTRSGPAWKNPMPGRETVLAPAAWEKGEWPV IQPVRGQM QG-PFPPPNKR----VPRGE
Pf43B YYNTNYYDTSYAYASSITGPWT-----KQSAPYAPLLVTGT-----ETSND
Fv43E FMSFTWAYPAGRLPVLAPITWG-----SDGFPILVKGANGGWGSSYPTLPGT

FIG. 93A

SUBSTITUTE SHEET (RULE 26)

113 / 138

Fv43D SIVEYKGQWWLFFAD-AHTSGKDYLRQVKARKIWDKDG-----KILLTRPKI-----
Fo43A SIVEYKGQWWLFFAD-AHTSGKDYLRQVKARKIWDKNG-----KILLHRP-----
Gz43A SIVKYQGQWWLFYHD-AKTSKGDYLRQVKAKKIWDYDSKG-----KILTKKP-----
Pf43A GIIDFQNNNSYFFYHNGALPGGGGYQRSVCVEQFKYNADG-----TIPTIEMTTAG-----
Fv43A GIIDFENNNSYFFYHNGALDGGSGYTRSVAVESFKYGS DG-----LIPEIKMTTQG-----
Fv43B GPFNADPDNYNLKKTCKIPPHFVHHRVPRDGAFLSSKG----LHIVPSRNNVTGSVLPG
Af43A GGWIKQPDKVDFRPGSKIPAHFQYWRYPKTEDEFTVSPRGHPNTLRLTPSFYNLTG-----
Pf43B GALSAPGGADFSVDGTMKLFHANLNGQDISGGRALFAAS-----ITEASDVVTLQ----
Fv43E DGVTKNWTRTDTFRGTS LAPSW EWNHNP DVNSFTV NNGLT LRTASITKDIYQARN-----

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A -----PAQIGTLNPPYVRQEAETA AWSSGITTEVCSEGGIDVGFINNG
Fv43A -----PAQLKSLNPPYVKQEAETIAWSEGIETEVCSEGG LNVA FIDNG
Fv43B DEIELSGQRGLAFIGRRQTHTLFKYSVDIDFKPKSDDQEAGITVFRTQFDHIDL GIVRLP
Af43A -TADFKPDDGLSLVMRKQTDTLFTYTVDVSFDPKVADEEAGVTVFLTQQQHIDL GIVLLQ
Pf43B -----
Fv43E -----T LSHRTHGDHPTGIVKIDFSPMKDGDRA GLS AFRDQSAYIGIHRDNGK

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A DYIK-----VKGVAFGS--GAHSFSARVASANS SGGTIAIHLGSTTGTLVGTCTV
Fv43A DYIK-----VKGVD FGSTGAKTFSARVASNSSGGKIELRLGSKTGKLVGTCTV
Fv43B TNQGSNKKSKLA FRFRATGAQNVPAPK---VVPVPD GWEKGVISLHIEAANATHYNLGAS
Af43A TTEG----LSLSFRFRVEGRGN YEGPLPEATVPVPKEWCGQTIRLEIQAVSDTEYVFAAA
Pf43B -----
Fv43E FTIAT---KHGMNMDEWNGTTTDLGQIKATANVPSGR TKIWLRLQLD TNPA GTGNTIFS

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A PSTGGWQTWTTVTCSVSGASGTQ-----DVYFVFGSGTG YLFN-----FDYWQFA
Fv43A TTTGNWQTYKTVD CPVSGATGTS-----DLFFVFTGSGSGSLFN-----FNWWQFS
Fv43B --SHRGKTLDIATASASLVSGGTGSFVGSLLGPYATCNGKSGSVECPKGGDVYVTQW TYK
Af43A PARHPAQRQIISRANSLIVSGDTGRFTGSLVG VYATSNG-GAGSTP-----AYISR WRYE
Pf43B -----
Fv43E YSWDGVKYETLGP NFKLYNG-----WAFFIAYRFGIFNFAETALGGS IKVESFT

Fv43D -----
Fo43A -----
Gz43A -----
Pf43A -----
Fv43A -----
Fv43B PVAQEIDHGVFVKSEL
Af43A GRGQMIDFGRVVP SY--
Pf43B -----
Fv43E AA-----

FIG. 93B

114 / 138

Pa51A MIHLKPALAALLALSTQCVAIDLFEVKSSGGNKTDDIMYGLMHEDINNSGDGGIYAELISN
Fv51A MVRFSSILAAAACF-VAVESVNIKVDKSGGNATSGHQYGFLEHEDINNSGDGGIYAELIRN
Pf51A MGKMWHSILVVLGLLSVGHAITINVSQSGGNKTSPLQYGLMFEDINHGGDGGLYAELVRN

Pa51A RAFQGSEKFPSNLDNWSPVGGATLTLQKLAKPLSSALPYSVNVANPKEGKGKGDTKGKK
Fv51A RAFQYSKKYPVSLSGWRPINDAKLSLNRLDTPLSDALPVSMNVK----PGKGK-----AKE
Pf51A RAFQGSTVYPANLDGYDSVNGAILALQNLTNPLSPSPMPSSLNVA-----KGS-----NNGS

Pa51A VGLANAGFWGMDVVKRQKYTGSEFHVTEGYKGDFFVSLRSALTGETFGKKVVKGGSKKGKWT
Fv51A IGFLNEGYWGM DVKKQKYTGSEFWVKGAYKGHFTASLRNLTDDVFGSVKVKSKANKKQWV
Pf51A IGFANEGWWGIEVKPQRYAGSFYVQGDYQGDFFDISLQSKLTQEVFATAKVRSSGKHEDWV

Pa51A EKEFELVPFKDAPNSNNTFVVQWDAEAGAKDGSLDLNLISLFPPTFKGRKNGLRIDLAQTM
Fv51A EHEFVLTPKNAPNSNNTFAITYDPKGA-DGALDFNLISLFPPTYKGRKNGLRVDLAEAL
Pf51A QYKYELVPKKAASNTNNTLTITFDSKGLKDGSLNFNLISLFPPTYNNRPNGLRIDLVEAM

Pa51A VELKPTFLRFPGGNMLEGNTLDTWWKWEYETIGPLKDRPGMAGVWEYQQTLGLGLVEYMEW
Fv51A EGLHPSLLRFPGGNMLEGNTNKTWWDWKDTLGPLRNRPGFEGVWNYQQTHGLGILEYLOW
Pf51A AELEGKFLRFPGGSDVEGVQAPYWKWNETVGDLDKDRYSRPSAWTYEESNGIGLIEYMNW

Pa51A ADDMNLEPIVGVFAGLALDGSFVPESEMGWVIOQALDEIEFLTGDAKTTKWGAVRAKLGH
Fv51A AEDMNLEIIVGVYAGLSLDGSVTPKDQLQPLIDDALDEIEFIRG-PVTSKWGKKRAELGH
Pf51A CDDMGLEPILAVWDGHYLSNEVISENDLQPYIDDTLNQLEFLMG-APDTPYGSWRASLGY

Pa51A PKPWVKVWVEIGNEDWLAGRPAGFESYINYRFPMMKAFNEKYPDIKIIASPSIFD-----
Fv51A PKPFRLSYVEVGNEDWLAGYPTGWNSYKEYRFPMFLEAIKKAHPDLTVISSGASIDPVGK
Pf51A PKPWTINIVEIGNEDNLYG---GLETYIAYRFQAYYDAITAKYPHMTVMESLTEMPG---

Pa51A ---NMTIPAGAAGDHHPYLTPDEFVERFAKFDNL SKDNVTFLIGEAASTHPNG---GIAWE
Fv51A KDAGFDIPAPGIGDYHPYREPDLVVEEFNLF DN NKYG--HIIGEVASTHPNG---GTGWS
Pf51A -----PAAAASDYHQYSTPDGFVSQFNYFDQMPVTNRTLNGEIA TVYPNNPSNSVAVG

Pa51A GDLMPPLPWWGGSVAEAI FLISTERNGDKIIGATYAPGLRSLDRWQWSMTWVQHAADPALT
Fv51A GNLMPYPWWISGVGEAVALCGYERNADRI PGTFYAPILKNENRWQWAITMIQFAADSAMT
Pf51A SPFPLYPWWIGSVAEAVFLIGEERN SPKIIGASYAPMFRNINNWQWSPTLIAFDADSSRT

Pa51A TRSTSWYVWRILAHHII RETLPVDAPAGKPNFDPLFYVAGKSES-GTGIFKAAVYNSTES
Fv51A TRSTSWYVWSLFAGHPMHTLPTTA-----DFDPLYVAGKNEDKGT LIWKGAAYNTTKG
Pf51A SRSTSWHVIKLLSTNKITQNLPTTWSGG--DIGPLYWVAGRNDNTGSNIFKAAVYNSTSD

Pa51A --IPVSLKFDGLNEGAVANLTVLTGPE-DPYGYNDPFTGINVVEKTTFIKAGKGGKFTF
Fv51A ADVPVSLSFKGVKPGAQAE LTLTNKEKDPFAFNDPHKGN NVDTKKT VLVKADGKGAFNF
Pf51A ---VPVTVQFAGCN-AKSANLTLSSDD--PNASNYPG-GPEVVKTEIQSVTANAHGAFEF

Pa51A TLPGLSVAVLETADAVKGGKGKGGKGGKGN
Fv51A KLPNLSVAVLETLK-----KGKPYSS
Pf51A SLPNLSVAVLKTE-----

FIG. 94

115 / 138

	*	20	*	40	*	
xyn3	:	-MKANVILC--LLAPLVAALPTETIHLDP	ELAALRANL	TERTADLWDRQA	:	47
P56588	:	-----	-----	-----QA	:	2
P23360	:	MVRPTILLTSLLLAPFAAASPI-----	-----	LEERQA	:	28
		60	*	80	*	100
xyn3	:	SQSIDQLIKRKGLYFGTATDRG	LLQRE-KNA	AI IQADLGQVTP	ENSMKW	: 96
P56588	:	SVSIDAKFKAHGKKYLGTIGDQY	TLTKNTKNPA	IIKADFGQLTP	ENSMKW	: 52
P23360	:	AQSVDDQLIKARGKVYFGVATDQ	NRLTTG-KNA	AI IQADFGQVTP	ENSMKW	: 77
	*	120	*	140	*	
xyn3	:	QSLENNQGQLNWDADYLVNFAQ	QNGKSIRGH	TLIWHSQLPAW	VNNINNA	: 146
P56588	:	DATEPNRGQFTFSGSDYLVNFA	QSNGLIRGH	TLVWHSQLPGW	VSSITDK	: 102
P23360	:	DATEPSQGNFNFAGADYLVNWA	QNGKLIRGH	TLVWHSQLPSW	VSSITDK	: 127
		160	*	180	*	200
xyn3	:	DTLRQVIRTHVSTVVGRYK	GKIRAWDVV	NEIFNEDGTL	RSSVFSRLLGEE	: 196
P56588	:	NTLISVLKNHITVMTRYK	GKIYAWDVL	NEIFNEDGSL	RNSVFYNVIGED	: 152
P23360	:	NLTNTVMKNHITTLMTRYK	GKIRAWDVV	NEAFNEDGSL	RQTVFLNVIGED	: 177
	*	220	*	240	*	
xyn3	:	FVSI AFRAARDADPSARLY	INDYNLDRAN	YGKVNGLKTY	VSKWISQGVPI	: 246
P56588	:	YVRI AFETARSVDPNAKLY	INDYNLDSAG	YSKVNGMVSH	VKKWLAAGIPI	: 202
P23360	:	YIPI AFQTARAADPNKLY	INDYNLDSAS	YPKTQAI VNR	VKQWRAAGVPI	: 227
		260	*	280	N	* 300
xyn3	:	DGIGSQSHLSGGGSGTLGAL	QQLATVPVTE	LAIT <u>EL</u> DIQGAPT	TDYTQV	: 296
P56588	:	DGIGSQTHLGAGAGSAVAGA	LNALASAGT	KEIAIT <u>EL</u> DIAGAS	STDYVNV	: 252
P23360	:	DGIGSQTHLSAGQGAGVLQ	ALPLLASAGT	PEVAIT <u>EL</u> DVAGAS	PTDYVNV	: 277
	*	320	*	340	*	
xyn3	:	VQACLSVSKCVGITVWG	ISDKDSWRAS	TNPLLF	DANFNPKPAYNSIVGIL	: 346
P56588	:	VNACLNQAKCVGITVWG	VADPDSWRSS	SSPLLF	DGNYNPKAAYNAIANAL	: 302
P23360	:	VNACLNVQSCVGITVWG	VADPDSWRAS	TTPLLF	DGNFNPKPAYNAIVQDL	: 327
xyn3	:	Q-	:	347		
P56588	:	--	:	-		
P23360	:	QQ	:	329		

FIG. 95A

116 / 138

1	M	V	S	F	S	S	L	L	A	C	S	A	I	G	-	A	L	A	P	V	E	P	E	T	S	F	N	AfuXyn2				
1	M	I	S	S	F	S	L	L	A	C	S	A	I	A	G	A	Y	A	L	P	S	D	-	K	S	V	S	L	A	AfuXyn5		
1	M	V	S	F	T	S	L	L	A	S	P	P	S	R	A	S	C	R	P	A	A	E	V	E	S	V	A	E	Xyn2			
30	E	T	A	L	H	E	F	A	E	R	A	G	T	P	S	S	T	G	N	N	G	Y	Y	S	F	W	T	D	AfuXyn2			
30	E	R	-	-	-	-	-	-	Q	T	I	T	I	T	S	Q	T	G	T	N	N	G	Y	Y	S	F	W	T	N	AfuXyn5		
31	K	R	-	-	-	-	-	-	Q	T	I	T	I	Q	P	G	-	T	G	Y	N	N	G	Y	F	Y	S	W	N	D	Xyn2	
60	G	G	D	V	T	Y	T	N	G	A	G	G	S	Y	S	V	N	W	R	N	-	-	V	G	N	F	V	G	AfuXyn2			
53	G	A	G	S	V	Q	Y	T	N	G	A	G	G	E	Y	S	V	T	W	A	N	Q	N	G	D	F	T	C	G	AfuXyn5		
53	G	H	G	G	V	T	Y	T	N	G	P	G	G	Q	F	S	V	N	W	S	N	-	-	S	G	N	F	V	G	Xyn2		
88	K	G	W	N	P	G	S	A	-	R	T	I	N	Y	G	S	F	N	P	S	G	N	G	Y	L	A	V	Y	G	AfuXyn2		
83	K	G	W	N	P	G	S	D	-	H	D	I	T	F	S	G	S	F	N	P	S	G	N	A	Y	L	S	V	Y	G	AfuXyn5	
81	K	G	W	Q	P	G	T	K	N	K	V	I	N	F	S	G	S	Y	N	P	S	G	N	S	Y	L	S	V	Y	G	Xyn2	
117	W	T	N	P	L	I	E	Y	Y	Y	Y	V	V	S	Y	G	T	Y	N	P	S	G	N	G	Y	L	A	V	Y	G	AfuXyn2	
112	W	T	N	P	L	V	E	Y	Y	Y	Y	V	V	S	Y	G	T	Y	N	P	S	G	N	G	Y	L	S	V	Y	G	AfuXyn5	
111	W	S	R	N	P	L	I	E	Y	Y	Y	V	V	S	Y	G	T	Y	N	P	S	G	N	G	Y	L	S	V	Y	G	Xyn2	
147	N	T	D	G	G	T	Y	N	I	Y	E	R	T	A	V	R	Y	N	A	P	S	I	E	G	T	K	T	Q	Y	Y	W	AfuXyn2
142	T	S	D	G	S	T	Y	D	I	Y	E	R	T	A	V	Q	Q	N	Q	P	S	I	V	G	T	A	T	Q	Y	Y	W	AfuXyn5
141	T	S	D	G	S	V	Y	D	I	Y	E	R	T	A	V	Q	Q	N	Q	P	S	I	I	G	T	A	T	Q	Y	Y	W	Xyn2
177	S	V	R	T	S	K	R	R	T	G	T	V	T	M	A	N	H	F	N	A	W	S	R	L	G	M	N	L	G	T	AfuXyn2	
172	S	I	R	Q	N	K	R	S	S	G	T	V	T	T	A	N	H	F	K	A	W	A	S	L	G	M	N	L	G	T	AfuXyn5	
171	S	V	R	R	N	H	R	S	S	G	S	V	V	T	A	N	H	F	N	A	W	A	Q	Q	G	L	T	L	G	T	Xyn2	
207	H	N	Y	Q	I	V	A	T	E	G	Y	Q	S	S	G	S	A	S	I	T	V	Y	S	G	S	S	S	G	AfuXyn2			
202	H	N	Y	Q	I	V	S	T	E	G	Y	E	S	S	G	S	G	T	S	I	T	V	S	S	G	S	S	G	AfuXyn5			
201	M	D	Y	Q	I	V	A	V	E	G	Y	F	S	S	G	S	A	S	I	T	V	S	S	S	S	S	S	G	Xyn2			
228	S	G	G	S	S	T	S	S	G	S	S	P	T	G	G	S	G	S	C	S	A	L	W	G	Q	C	G	AfuXyn2				
232	S	G	G	S	S	T	S	S	G	S	S	P	T	G	G	S	G	S	C	S	A	L	W	G	Q	C	G	AfuXyn5				
222	S	G	G	S	S	T	S	S	G	S	S	P	T	G	G	S	G	S	C	S	A	L	W	G	Q	C	G	Xyn2				
228	I	G	-	W	S	G	P	T	C	C	S	S	G	T	C	Q	V	S	N	S	Y	S	Q	C	L				AfuXyn2			
262	I	G	-	W	S	G	P	T	C	C	S	S	G	T	C	Q	V	S	N	S	Y	S	Q	C	L				AfuXyn5			
222	I	G	-	W	S	G	P	T	C	C	S	S	G	T	C	Q	V	S	N	S	Y	S	Q	C	L				Xyn2			

FIG. 95B

117 / 138

Fv3A MLLNLQVAASALSLSLLGGLAEATPYT-----LPDCTKGPLSK
Bx11 -MVNNAALLAALSALLPTALAQNNQTYANYSAQGQPDLYPETLATLTLSEFPDCEHGPLKN
Bgl1 MRY-RTAAALALATG-----PFARADSHSTSG---ASAEAVVP-----

Fv3A NGICDTSLSPAKRAAALVAALTPEEKVGNLVSNATGAPRIGLPRYNWWNEALHGLAGSPG
Bx11 NLVCDSSAGYVERAQALISLFTLEELILNTQNSGPGVPRGLGPNYQVWNEALHGLDRAN-
Bgl1 -PAGTPWGTAYDKAKAALAKLNLQDKVGIVSGVGWNGGPCVGNTSPASKISYPSLCLQDG

Fv3A GRFADTP-PYDAATSFPMPLLMAAAFDDDLIHDIGNVVGTEARAFNTGGWRGVDFWTPNV
Bx11 --FATKGGQFEWATSFPMPILTTAALNRTLHQIADIISTQARAFSNSGRYGLDVYAPNV
Bgl1 ---PLGVRYSTGSTAFTPGVQAASTWDVNLIRERGQFIGEEVKASGIHVILGPVAGPLGK

Fv3A NPEKDPRWGRGSETPGEDALHVS-RYARYIVRGLEG---DKEQRRIVATCKHYAGNDFEDW
Bx11 NGFRSPLWGRGQETPGEDAFFLSSAYTYEYITGIQGGVDPEHLKVAATVKHFAGYDLENW
Bgl1 TPQGGRNWEGF---GVDP--YLTGIAMGQTINGIQS-----VGQATAKHYILNEQELN

Fv3A GGFTTR-----HDFDAKITPDLAEEYVRPFQECTRDAKVGSIKAYNAVNGIPACA
Bx11 NNQSR-----LGFDAIITQDDLSEYYTPQFLAAARYAKSRSLMCAYNSVNGVPSCA
Bgl1 R-----ETISSNPDDRTLHELYTWPFADAVQAN-VASVMCSYNKVNTTWACE

Fv3A NSYLQETILRGHWNWTRDNNWITSDCGAMQDIWQNHKYVKTNAGAQAQVAFENGMDSSCEY
Bx11 NSFFLQTLRESWGFP-EWGYVSSDCDAVYNVFNPHDYASNQSSAAASSLRAGTDIDCGQ
Bgl1 DQYTLQTVLKDQLGFP---GYVMTDWNQAQHTTVQSANSGLDMSMPG--TDFNGNRLWGPA

Fv3A TTTSDVSDSYKQGLLTEKLMDRSLKRLFEGLVHTGFFDGAKAQWN-----
Bx11 TYPWHLNESFVAGEVSRGEIERSVTRLYANLVRLGYFD-KKNQYR-----
Bgl1 LTNAVNSNQVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNIS-----

Fv3A -SLSFADVNTKEAQDLALRSAGEGAVLLLKNDG-TLPLKLKKKDSVAMIGFWAN-----
Bx11 -SLGWKDVVKTDAWNISYEAAGEGIVLLLKNDG-TLPLSKKVR-SIALIGPWAN-----
Bgl1 -----RNVQGNHKTNVRAIARDGIVLLLKNDANILPLKKPASIAVVGSAIIIGNHARNSP

Fv3A -----DTSKLQGGYSGRAPFLHSPLYAAEKLGLDTNVAWGPTLQNSSSHDNWTTN
Bx11 -----ATTQMONGNYGPAPYLISPLEAAKKAGYHVNFEGLGTEIAGNSTTG--FAK
Bgl1 SCNDKGCDDGALGMGWGSGAVNPYFVAPYDAINTRASSQGT--QVTLSTNDNTS---SG

Fv3A AVAAAKKSDYILYFGGLDAS-----AAGEDRDRENLDWPESQLTLQKLSSLGKPLVVIQL
Bx11 AIAAAKKSDAIYLGIDNT-----IEQEGADRTDIAWPGNQLDLIKQLSEVGKPLVLQOM
Bgl1 ASAARGKDVAIVFITADSGEGYITVEGNAGDRNNLDPWHNGNALVQAVAGANSNVIVVH

Fv3A G-DQVDDTALLKNKKINSILWVNYPGQDGGTAVMDLLTGRKSPAGRLPVTQYP-----
Bx11 GGGQVDSSSLKSNKKVNSLVWGGYPGQSGGVALFDILSGKRAPAGRLVTTQYP-----
Bgl1 SVGAIIILEQILALPQVKAVVWAGLPSQESGNALVDVLWGDVSPSGKLVTYIAKSPNDYNT

Fv3A -----SKYTEQIGMTDMDLRPT-KSLPGRTYRWYS-TPVLPYGFGLHYTKF
Bx11 -----AEYVHQFPQNDMNLRPDGKSNPGQTYIWTGKPVYEFSGSLFYTTF
Bgl1 RIVSGGS-----DSFSEGLFIDYKHFDDAN-----ITPRYEEGYGLSYTKF

FIG. 96A

118 / 138

Fv3A	QAKFKSN--KLTFDIQKLLKG--CSAQYS-----
Bx11	KETLASHPKSLKFNTSSILSAPHPGYTYS-----
Bg11	NYSRLSVLSTAKS--GPATGAVVP-----
Fv3A	-----DTCALPPIQVSVKNTGR
Bx11	-----EQIPVFTFEANIKNSGK
Bg11	-----GGPSDLFQNVATVTVDIANSQ
Fv3A	ITSDFVSLVFIKS-EVGPKPYPLKTLAAYGRLHDVAPSSTKDISLEWTLDNIAARRGENGD
Bx11	TESPYTAMLFVRTSNAGPAPYPNKWLVGFDRDLADIKPGHSSKLSIPIPVSALARVD SHGN
Bg11	VTGAEVAQLYIITYPSSAPR-TPPKQLRGFAKLN-LTPGQSGTATFNIRRRDLSYWD TASQ
Fv3A	LVVYPGTYTLLLDEPTQAKIQVTLTGKKAILDKWPQDPKSA-----
Bx11	RIVYPGKYELALNTDES VKLEFELVGEEVTIENWPLEEQQIKDATPDA
Bg11	KWVVPSSGSPGISVGASSRDIRLTSTLSVA-----

FIG. 96B

119 / 138

```

Fv30B.pro MNPLSLGLAALLGYVGVNFVAAFPTDSNSGSEVLISVNGHVKHQELDGFASQAFQRAEDILGKDGLSKEGTQHVLDL 80
Fv30A.pro ML-FSLVLPFLA-----FQASLALGDT-----VTVDTSQKLQVIDGFGVSEAYGHAKQF---QNLGPGPQKEGLDL 63

Fv30B.pro LFSKDIGAGESILRNGIGSSNSDDKNFMNSIEFFSPGSPGAKPHYVWDGYDSGQLTVAQEAFFKRGKFLYGDAWSAPGYM 160
Fv30A.pro LFNTTGTAGLSIIRNKIGCDAS-----NSITSTNTDNPDKQAVYHFDGDDDGQ----- 111

Fv30B.pro KTNHDENNGYLGGVTGAACASGDWKQAYADYLLQWVEFYRKSGVKVTNLGFLNEPQFAAPYAGMLNSGTQAADFIRVLG 240
Fv30A.pro -----SAQSMGRLCGTPGVSCSSGDRHRHYVEMIAEYLSYKQAGIPVSHVGFNLEGD-GSDF--MLSTAEQAADVIFLLH 184

Fv30B.pro KTIRKRGIHDLTIAACDGEWDLQ---EDMMAGLTAGPDPAINYLSVVTGHGYVSPPNHPLSTTKKTWLTEWADLTGQFT 317
Fv30A.pro SALQSKGLGDIKMTCCDNIGWKSQMDYTAKLAELEVE-----KYSVITSHEYSSSPNQPMNTTLPWTWMSEGAANDQAF 259

Fv30B.pro PYTFYNNSCQGEGMTWAGRIQTALVDANVSGFLYWIGAENSTNS--ALINMIGDKVIPSKRFFWAFASFSRFARPGARRI 395
Fv30A.pro T-AWYVNGGSNEGFTWAVKIAQGI VNAADLSAYIYWEGVETNNKGSLSHVIDTDGTKFTISSILWAIAHWSRHIRPGAHRL 338

Fv30B.pro EATSSVPLVTVSSFLNTDGTVAQTQVLNNDTVAHSVQLVVSQTG----- 438
Fv30A.pro STSGVVQDTIVGAFENVDSVVMVLTNSGTAQTVDLGVSGSSFTAQFTSDAEAQMVDTKVTLSDGRVKVTVPVHGVV 418

Fv30B.pro -----RNP HSLKPFLTDNSNDLTALKHLKATCKGSFQTTIPPRSLVSFVT----- 483
Fv30A.pro TVKLTTAKSSKPFVSTAVSAQSAPTPTSVKHTLTHQKTSSTLSTAKAPTSTQTTTSVVESAKAVKYPVPFVASKGSSKSAP 498

Fv30B.pro -----DF 485
Fv30A.pro KKGTKKTTTKGSHQSHKAHSATHRRRCRHGYSYRRGHCTN 537

```

FIG. 97

120 / 138

GH61 endoglucanase motifs of the disclosure:

Motif 1 of GH61 Family Endoglucanases:
 SEQ ID NO:84: (I/L/M/V)-P-a-a-a-G-a-Y-(I/L/M/V)-a-R-a-(E/Q)-a-a-a-a-(H/N/Q)

Motif 2 of GH61 Family Endoglucanases:
 SEQ ID NO:85: (I/L/M/V)-p-a-a-a-a-G-a-Y-(I/L/M/V)-a-R-a-(E/Q)-a-a-a-a-(H/N/Q)

Motif 3 of GH61 Family Endoglucanases:
 SEQ ID NO :86: (I/L/M/V)-p-a-a-a-a-G-a-Y-(I/L/M/V)-a-R-a-(E/Q)-a-a-a-A-(H/N/Q)

Motif 4 of GH61 Family Endoglucanases:
 SEQ ID NO :87: (I/L/M/V)-p-a-a-a-a-G-a-Y-(I/L/M/V)-a-R-a-(E/Q)-a-a-a-A-(H/N/Q)

Motif 5 of GH61 Family Endoglucanases:
 SEQ ID NO:88: (F/W)-(T/F)-K-(A/I/V)

Motif 6 of GH61 Family Endoglucanases:
 SEQ ID NO :89: H-a-a-G-P-a-a-a-(Y/W)-(A/I/L/M/V)

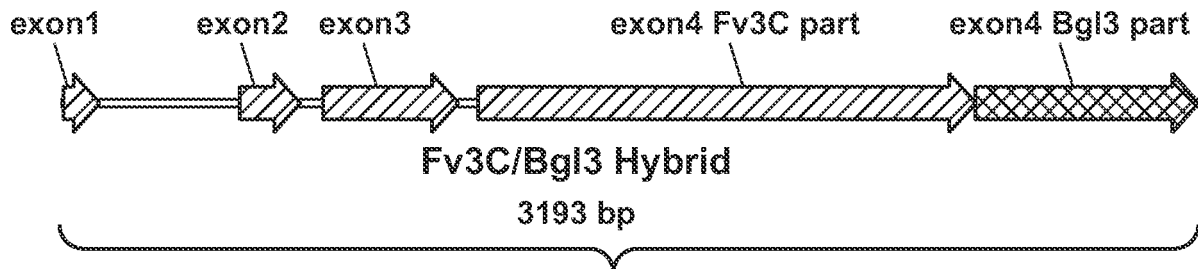
Motif 7 of GH61 Family Endoglucanases:
 SEQ ID NO :90: H-a-G-P-a-a-a-(Y/W)-(A/I/L/M/V)

Motif 8 of GH61 Family Endoglucanases:
 SEQ ID NO :91: (E/Q)-a-Y-a-a-C-a-(E/H/Q/N)-(F/I/L/V)-a-(I/L/V)

FIG. 98

121 / 138

A schematic representation of the Fv3C/Bgl3 hybrid/chimera gene.

**FIG. 99A**

SEQ ID NO: 92

The nucleotide sequence encoding Fv3C/Bgl3

```

atgaagctgaattgggtcgccgcagccctgtctataggtgctgctggcactgacagcgca
gttgcctcttgcttctgcagttccagacacttttggtggtgtaaaggctcagtttttttca
ccatttccctcgtctaatctcagcccttggtgccatatcgcccttggttcgctcggaacgccac
gcaccagatcgcgatcatttccctcccttgccagcccttggttccctcttacgatcttccctcc
gcaattatcagcgcccttagtctacacaaaaacccccgagacagtccttccattgagtttg
tcgacatcaagttgcttctcaactgtgcatttgcggtggctgtctacttctgcctctagac
aaccaaatctgggcgcaattgaccgctcaaaccttggttcaaataaaccttttttattcgag
acgcacatttataaatatgcgcctttcaataataccgactttatgcgcggcggtgctgtg
ggcggttgatcagaaagctgacgctcaaaagggttgacagagagatacactcgcatactc
gccgcctcattatccttcaccatggatggaccctaattgctgttggtggtggagggaagctta
cgccaaagccaagagctttgtgtcccaactcactctcatggaaaagggtcaacttgaccac
tggtgttggttaagcagctccttgcaaacagggtatctcaatccctcagctaacaactt
ctcagatggcaaggcgaacgctgtgttaggaaacgtgggatcaattcctcgtctcggtatg
cgaggtctctgtctccaggatggctcctcttggaattcgtctgtccgactacaacagcgct
tttcccgctggcaccacagctggtgcttcttgagcaagtctctctggtatgagagaggt
ctcctgatgggcactgagttcaaggagaagggtatcgatatcgctcttggtcctgctact
ggacctcttggtcgcactgctgctggtggaacgaaactgggaaggcttcaccgttgatcct
tatatggctggccacgccaatggccgaggccgctcaagggtattcaagacgcaggtgtcatt
gcttggtgctaagcattacatcgcaaacgagcagggttaagccacttggaacgatttgaggaa
ttgacagagaactgaccctcttgtagagcacttccgacagagtggcgagggtccagtcctcg
caagtacaacatctccgagtcctctctcctccaacctggatgacaagactatgcaacgagct
ctacgcctggcccttcgctgacgcgctccgcgcgcggcgctcggttccgctcatgtgctcgta
caaccagatcaacaactcgtaacggttgccagaactccaagctcctcaacgggtatcctcaa
ggacgagatgggcttccagggtttcgctcatgagcgattgggcgggcccagcataccgggtgc
cgcttctgccgctcgctggtctcgatatgagcatgcctggtgacactgccttcgacagcgg
atacagcttctggggcggaacttgactctggctgtcatcaacggaactgttcccgccctg
gcgagttgatgacatggctctgcgaatcatgtctgccttcttcaagggttggaagacgat
agaggatcttcccgacatcaacttctcctcctggaccgcgacaccttcggcttcgtgca
tacatttgctcaagagaacccgcgagcaggtcaacttttgaggtcaacgctccagcaacgacca
caagagccacatccgtgaggccgctgccaaagggaagcgtcgtgctcaagaacaccgggtc
ccttccctcaagaacccaaaagttcctcgtgtcattgggtgaggacgccgggtcccaaccc
tgctggaccctaattggttggtgacccgtgggttgcgataatggtaccctggctatggcttg

```

FIG. 99B-1

SUBSTITUTE SHEET (RULE 26)

122 / 138

gggctcgggaaacttcccaattcccttacttgatcaccctcgatcaagggtctctataatcg
 agctaactcaagacggaactcgatatgagagcatcttgaccaacaacgaatgggcttcagt
 acaagctcttgtcagccagcctaactgacccgctatcgttttcgccaatgccgactctgg
 tgagggatacattgaagtcgacggaacttttggtgatcgcaagaacctcaccctctggca
 gcagggagacgagctcatcaagaacgtgtcggtccatatgccccaacaccattgtagttct
 gcacaccgtcgggccctgtcctactcgccgactacgagaagaacccccaacatcaactgccat
 cgtctgggctggtcttcccggccaagagtcaggcaatgccatcgctgatctcctctacgg
 caaggtcagccctggccgatctcccttcaacttggggccgcaccgcgagagctacgggtac
 tgaggttctttatgagggcaacaacggccgtggcgctcctcaggatgacttctctgaggg
 tgtcttcacgactaccgtcacttcgaccgacgatctccaagcaccgatggaaagagctc
 tcccaacaacaccgctgctcctctctacgagttcggtcacgggtctatcttggtcgacggt
 caagttctccaacctccacatccagaagaacaatgtcggccccatgagcccgcccaacgg
 caagacgattgcggtcctctctggtcagcttcagcaagaaccttaaggactatggctt
 ccccaagaacggttcgcccgcacatcaaggagtttatctacccctacctgagcaccactacctc
 tggcaaggaggcgctcggtgacgctcactacggccagactgcgaaggagttcctccccgc
 cgggtgccctggacggcagccctcagcctcgctctgcggcctctggcgaacccggcgga
 ccgccagctgtacgacattctctacaccgtgacggccaccattaccaacacgggctcggt
 catggacgacgcccgttcccagctgtacctgagccacggcggtcccaacgagccgccccaa
 ggtgctgcgtggcttcgaccgcacatcgagcgcattgctcccgccagagcgctcacgttcaa
 ggcagacctgacgcgcgctgacctgtccaactgggacacgaagaagcagcagtggtcat
 taccgactaccccaagactgtgtacgtgggcagctcctcgcgcgacctgcccgtgagcgc
 ccgctgccatga

FIG. 99B-2**SEQ ID NO:93**

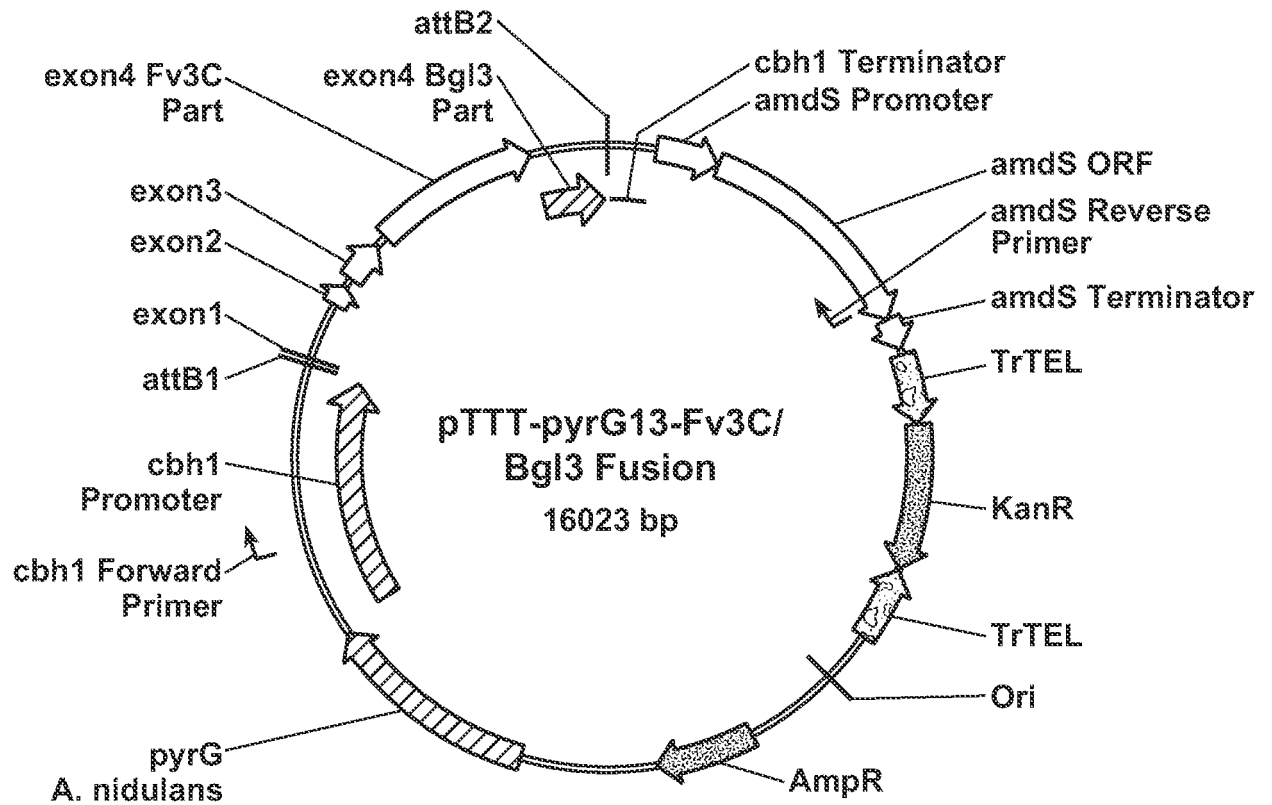
The Fv3C/Bgl3 chimeric polypeptide sequence (the Bgl3 chimeric part is in bold and upper case)

mklmwaaalsigaagtdsavalasavpdtlagvkkadaqkvvtrdtlayspphypspwmdpna
 vgweeayakaksfvsqtlmekvnlttgvgwqgercvgnvgsiprlgmrglclqdgplgirlsd
 ynsafpagttagaswskslwyergllmgtefkekgidialgpatgplgrtaaggrnwegftvdp
 ymaghamaeavkgiqdagviacakhyaneeqehfrqsgevqsrkyniseslssnlddktmhely
 awpfadavragvgsvmcsynqinnsygcqnsklngilkdemgfqgfvmssdwaaqhtgaasava
 gldmsmpgdtafdsygysfwggnltlavingtvpawrvddmalrimsaffkvgtiedlpldinf
 swtrdtfgfvhtfaenreqvnfgvnnvqhdkshireaaakgsvvlkntgslplknkflavig
 edagpnpagpncgdrgcdngtlamawgsqtsqfpylitpdqglsnratqdgtryesiltnew
 asvqalvsgpnvtaivfanadsgegyievdgnfgdrknltlwqqgdeliknvssicpntivvlh
 tvgpvlladyeknpnitaivwaglpqgesgnaiadllygkvspgrspftwgrtresygtevlye
 annrggapqddfsegvfidyrrhfdrrspstdgksspnntaaplyefghgls**WSTFKFSNLHIQK**
NNVGPMSPNGKTIAAPSLGSFSKNLKDYGFPKNVRRRIKEFIYPYLSTTTSGKEASGDAHYGQT
AKEFLPAGALDGSPQPRSAASGEPGGNRQLYDILYTVTATITNTGSMDDAVPQLYLSHGPPNE
PPKVLRGFDRIERIAPGQSVTFKADLTRDLSNWDTKKQQWVITDYPKTVYVGSSSRDLPLSARLP

FIG. 99C

123 / 138

A map of the expression plasmid pTTT-pyr13-Fv3C/Bgl3 fusion

**FIG. 100**

124 / 138

SEQ ID NO :94

Nucleic acid sequence encoding the Fv3C/Te3A/*T. reesei* Bgl3 (FAB) chimera:

atgaagctgaattgggtcgccgcagccctgtctataggtgctgctggcaactgacagcgcagttg
ctcttgcctctgcagttccagacactttggctgggtgtaaaggctcagttttttttcaccatttcc
tcgtctaattctcagccttgttgccatatcgcccttgttcgctcggaacgccacgcaccagatcgc
gatcatttccctcccttgcagccttgggttccctcttacgatcttccctccgcaattatcagcgcgc
ttagtctacacaaaaaaccccgagacagctctttcattgagtttgctgacatcaagttgcttctc
aactgtgcatttgcgtggctgtctacttctgcctctagacaaccaaactctggggcgcaattgacc
gctcaaacccttggtcaaataaccttttttatttcgagacgcacattttataaatatgcgcctttca
ataataaccgactttatgcgcggcggtgtgtgtggcggttgatcagaaagctgacgctcaaaagg
ttgtcacgagagatacactcgcatactcgcgcctcattatccttcaccatggatggaccctaa
tgctgttggctgggaggaagcttacgccaaagccaagagctttgtgtcccaactcactctcatg
gaaaagggtcaacttgaccactgggtgttgggttaagcagctccttgcaaacagggtatctcaatcc
cctcagctaacaacttctcagatggcaaggcgaaacgctgtgttaggaaacgtgggatcaattcct
cgtctcgggtatgcgaggtctctgtctccaggatgggtcctcttggaattcgtctgtccgactaca
acagcgcctttcccgctggcaccacagctgggtgcttcttggaagcaagctctctctggtatgagag
aggctctcctgatgggcactgagttcaaggagaagggtatcgatatcgctcttgggtcctgctact
ggacctcttgggtcgcaactgctgctgggtggacgaaactgggaaggcttcaccggtgatccttata
tggctggccacgcctatggccgagggcgtcaagggtattcaagacgcaggtgtcattgcttgtgc
taagcattacatcgcaaacgagcagggtaagccacttggacgatttgaggaattgacagagaac
tgacctcttgttagagcacttccgacagagtggtcgaggtccagtcgccgaagtacaacatctcc
gagctctctctcctccaacctggatgacaagactatgcacgagctctacgcctggcccttcgctg
acgcgctccgcgcgcggcgtcggttcggtcatgtgtcgtacaaaccagatcaacaactcgtacgg
ttgccagaactccaagctcctcaacggtatcctcaaggacgagatgggttccagggtttcgtc
atgagcgaattgggcggcccagcataccgggtgcgccttctgccgtcgtggtctcgatatgagca
tgccctgggtgacactgccttcgacagcggatacagcttctggggcggaacttgactctggctgt
catcaacggaactgttcccgctggcgagttgatgacatggctctgcgaatcatgtctgccttc
ttcaagggttggaagacgatagaggatcttcccgacatcaacttctcctcctggaccgcgcaca
ccttcgggttcgtgcatacatttgcctcaagagaacgcgcgagcaggtcaactttggagtcacagc
ccagcacgaccacaagagccacatccgtgagggccgctgccaaagggaagcgtcgtgctcaagaac
accgggtcccttccctcctcaagaacccaaagtctcctcgtgctcattgggtgaggacgcgcgggtccca
accctgctggacccaatggttgtggtgaccggtggttgcgataatggtaccctggctatggcttg
gggtcgcgggaacttcccaattcccttacttgatcaaccccgatcaagggtctctctaategagct
actcaagacggaactcgatatgagagcatcttgaccaacaacgaatgggtctcagtacaagctc
ttgtcagccagcctaacgtgaccgctatcgttttcgccaatgccgactctggtgagggatacat
tgaagtgcagcgaaactttggtgatcgcaagaacctcaccctctggcagcagggagacgagctc
atcaagaacgtgtcgtccatatgccccaaacaccattgttagttctgcacacgctcggccctgtcc
tactcgcgcgactacgagaagaaccccaacatcactgccatcgtctgggtggtcttcccggccca
agagtcaggcaatgccatcgtgatctcctctacggcaagggtcagccctggccgatctcccttc
acttggggccgcacccgcgagagctacgggtactgaggttctttatgaggcgaaacaacggccgtg
gcgtcctcaggatgacttctctgaggggtgtcttcacgcactaccgtcacttcgacaagtacaa
catcacgcctatctacgagttcgggtcacgggtctatcttgggtcgacgttcaagttctccaacctc
cacatccagaagaacaatgtcggccccatgagcccgcccaacggcaagacgattgcgggtccct
ctctgggcaacttcagcaagaaccttaaggactatggcttccccaagaacgttcgccgcacatcaa
ggagtttatctacccctacctgaacaccactacctctggcaaggaggcgtcgggtgacgctcac
tacggccagactgcgaaggaggttctccccgcgggtgccctggacggcagccctcagcctcgtc
ctgcggtcctctggcgaaacccggcggaacccgcagctgtacgacattctctacaccgtgacggc
caccattaccaacacgggtcgtggtcatggacgacgcgcgttccccagctgtacctgagccacggc
ggtcccaacgagccgcaccaagggtgctgcgtggcttcgacccgcatcgagcgcattgtctcccgcc
agagcgtcacgttcaaggcagacctgacgcgcgtgacctgtccaactgggacacgaagaagca
gcagtggtgctattaccgactaccccaagactgtgtacgtgggcagctcctcgcgcgacctgccg
ctgagcgcgccgctgccatga

FIG. 101A

SUBSTITUTE SHEET (RULE 26)

125 / 138

SEQ ID NO:95

Amino acid sequence of the Fv3C/Te3A/Bgl3 (FAB) chimera

mklnwvaaalsigaagtdsavalasavpdtlagvkkadaqkvvtrdtlayspphyppspwmdpna
vgweeayakaksfvsqtlmekvnlttgvgwqgercvgnvgsiprlgmrglclqdgplgirlsd
ynsafpagttagaswskslwyergllmgtefkekgidialgpatgplgrtaaggrnwegftvdp
ymaghamaeavkgiqdagviacakhyanegehfrqsgevqsrkyniseslssnlddktmhely
awpfadavragvgsvmcsynqinnsygcqnsklngilkdemgfqgfvmstdwaaqhtgaasava
gldmsmpgdtafdsgysfwggnltlavingtvpawrvddmalrimsaffkvgktiedlpdinfs
swtrdtfgfvhtfaenreqvnfgvvnvghdhkshireaaaakgsvglkntgslplknpkflavig
edagpnpagpngcgdrgcdngtlamawgsqtsqfpylittpdqglsnratqdgtryesiltnew
asvqalvsqpnvtaiivfanadsgegyievdgnfgdrknltlwqggdeliknvssicpntivvlh
tvgpvlladyeknpnitaivwaglpqgesgnaiadllygkvspgrspftwgrtresygtevlye
annrggapqddfsegvfidyrrhfd**KYNITPI**yefghglswSTFKFSNLHIQKNNVGPMSPPNGK
TIAAPSLGNFSKNLKDYGFPKNVRRIKEFIYPYLNTTSGKEASGDAHYGQTAKEFLPAGALDG
SPOPRSAASGEPPGNRQLYDILYTVTATITNTGSVMDDAVPOLYLSHGGPNEPPKVLRGFDRIE
RIAPGQSVTFKADLTRRDLNWDTKKQQWVITDYPKTVYVGSSSRDLPLSARLP

FIG. 101B

126 / 138

Amino acid sequence motifs of a suitable β -glucosidase polypeptide (including a variant, a mutant, or a hybrid/chimeric polypeptide thereof) of the invention

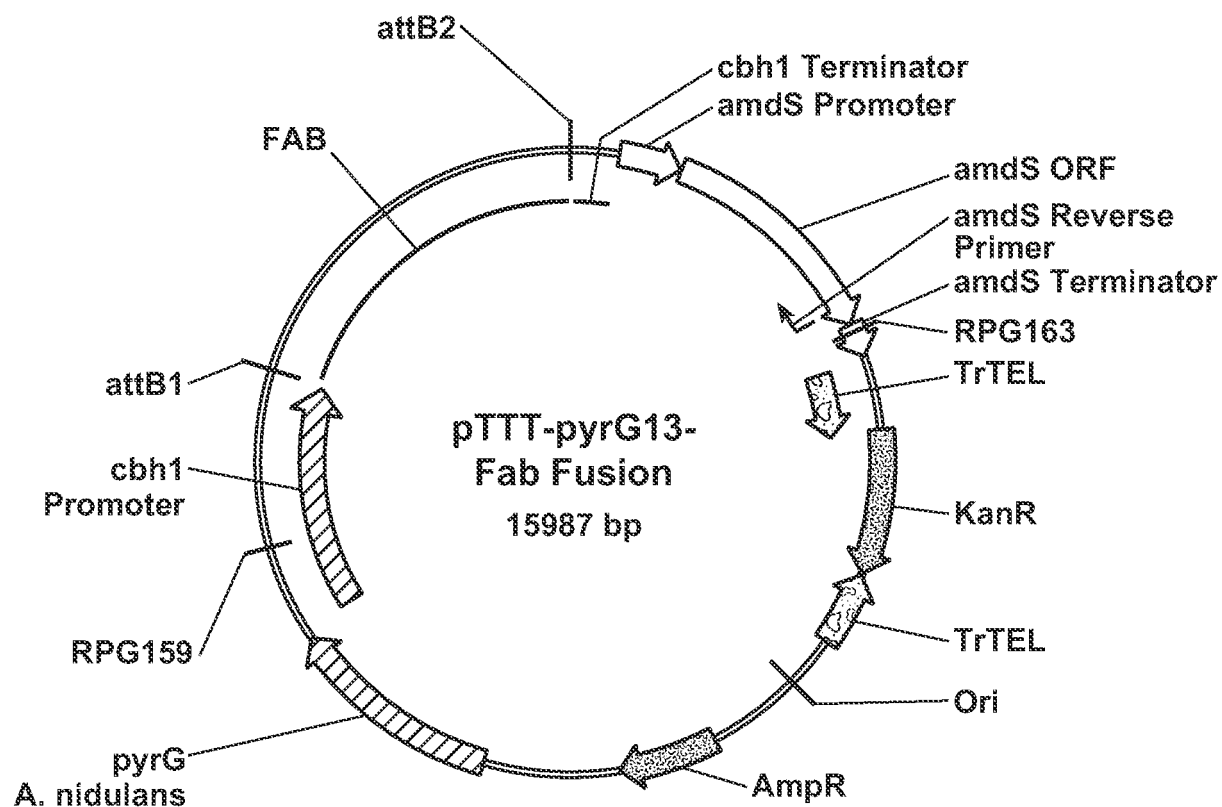
SEQ ID NO:	TERMINAL	SEQUENCE MOTIFS
96	N	A-x-S-P-P-x-Y-P-S-P-W-M-D-P-x-A-x-G-W-E-x-A-Y
97	N	A-K-x-F-V-S-x-x-T-L-x-E-K-V-N-L-T-T-G-V-G-W-x-G-E-x-C-V-G-N-V-G
98	N	P-R-x-G-M-R-x-L-C-x-Q-D-G-P-L-G-x-R
99	N	Y-N-S-A-F-x-x-G-x-T-A-x-A-S-W-S
100	N	G-x-I-A-C-A-K-H-x-x-x-N-E-Q-E-H-x-R-Q
101	N	L-S-S-N-x-D-D-K-T-x-H-E-x-Y-x-W-P-F-x-D-A-V-x-A-G-V-G
102	N	M-C-S-Y-x-Q-x-N-N-S-Y-x-C-Q-N-S-K-L-x-N-G
103	N	G-F-Q-G-F-V-M-S-D-W-x-A-Q-H-x-G-x-A-x-A-V-A-G-L-D-M-x-M-P-G-D-T
104	N	N-L-T-L-A-V-x-N-G-T-V-P-x-W-R-x-D-D-M
105	N	P-x-F-L-x-V-x-G-E-D-A-G-x-N-P-A-G-P-N-G-C-x-D-R-G-C
106	N	G-T-L-A-M-x-W-G-S-G-T-x-F-P-Y-L
107	N	A-I-V-F-A-N-x-x-S-G-E-G-Y-I-x-V-D-G-N-x-G-D-R-K-N-L-T-L-W
108	N	D-x-L-Y-G-K-x-S-P-G-R-x-P-F-T-W-G
109	C	P-x-Y-E-F-G-x-G-L-S-W-x-T-F-x-x-S-x-L
110	C	L-x-D-Y-x-F-P
111	C	E-F-L-P-x-x-A-L-x-G-S-x-Q-P-R
112	C	S-G-x-P-G-G-N-x-x-L-x-D
113	C	Y-T-V-x-A-x-I-T-N-T-G
114	C	V-L-R-G-F-x-R-x-E-x-I-A-P-G-x-S
115	C	T-R-R-D-L-S-N-W-D-x-x-x-Q-x-W-V-I-T-D
116	C	V-G-S-S-R-x-L-P-L-x-A-x-L

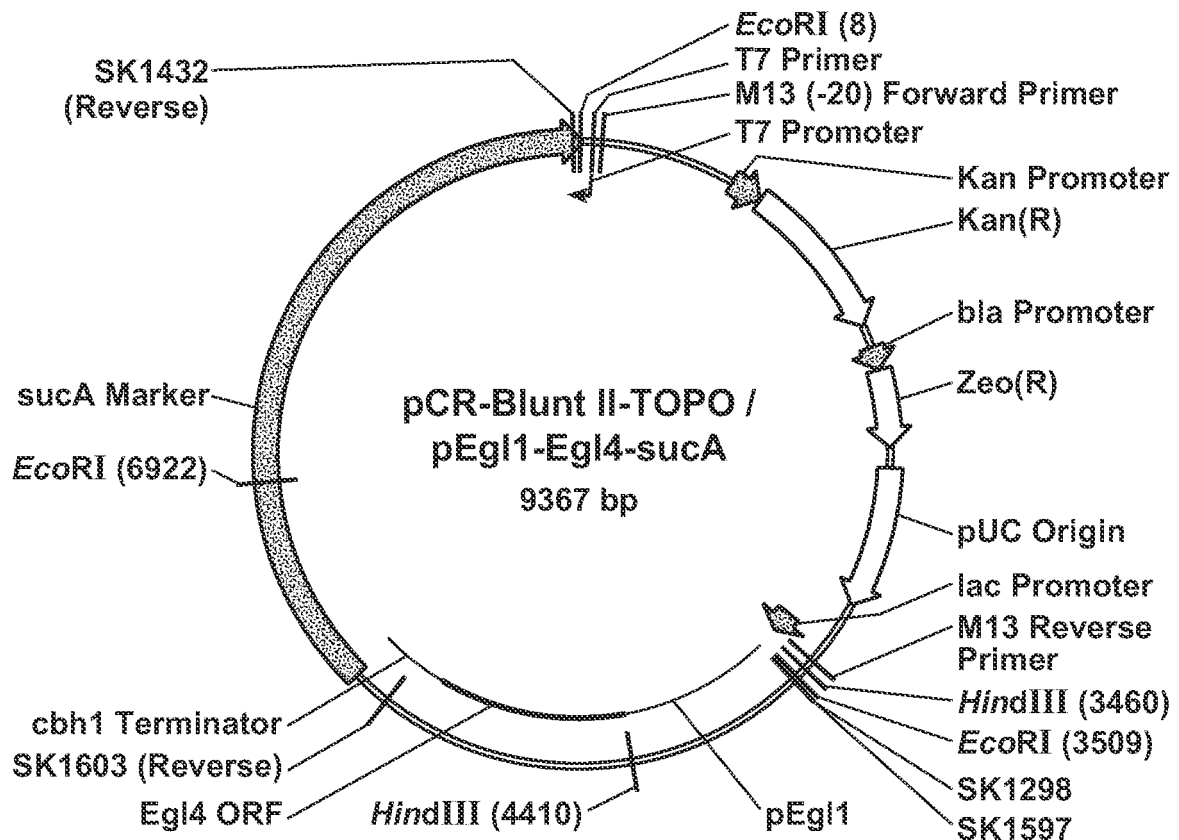
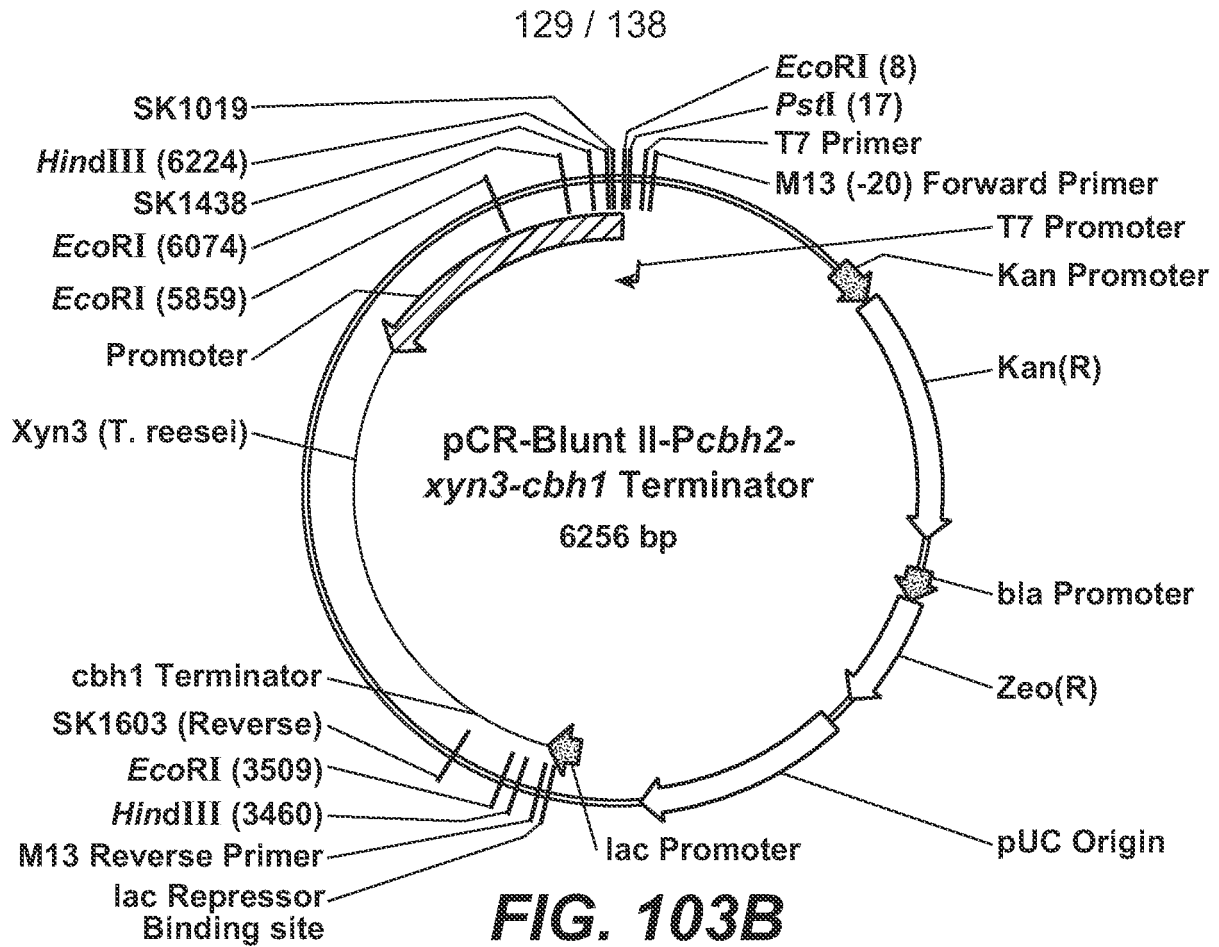
FIG. 102A

FIG. 102B Amino acid sequence motifs used to design a suitable β -glucosidase polypeptide hybrid/chimera of the invention

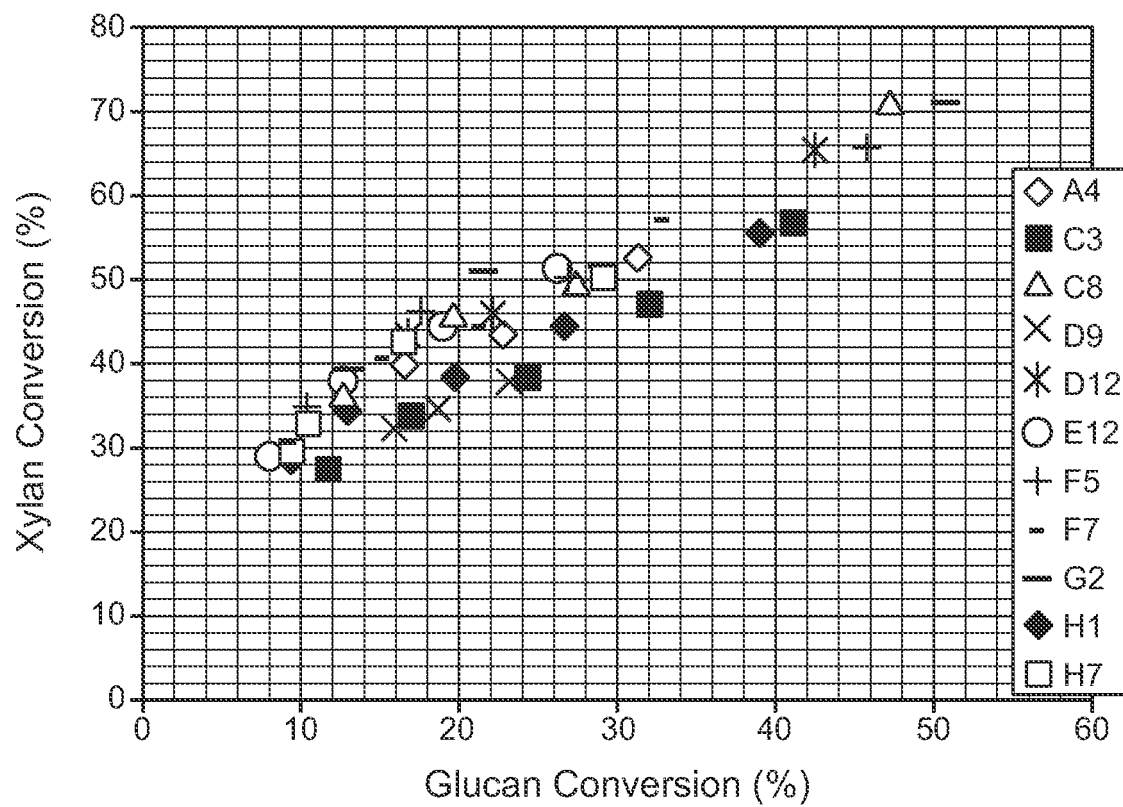
SEQ ID NO:	TERMINAL	SEQUENCE MOTIFS
197	N	Y-P-S-P-W-M-D-P
198	N	E-K-V-N-L-T-T-G-V-G-W
199	N	K-G-(I/V)-D-(V/I)
200	N	C-Q-N-S-K-L-x-N-G
201	N	N-L-T-L-A-V-(L/I/V)-N-G-(S/T)-(V/I)-P-x-W
202	N	S-W-(T/S)-x-D-T-(Y/F)-G
203	C	E-F-L-P-x-x-A-L-x-G-S-x-Q-P-R

128 / 138

**FIG. 103A**



130 / 138

**FIG. 104**

131 / 138

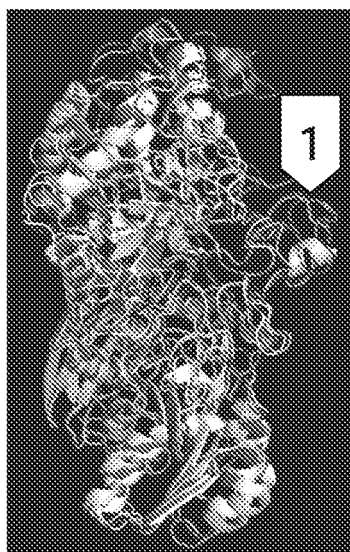


FIG. 105A

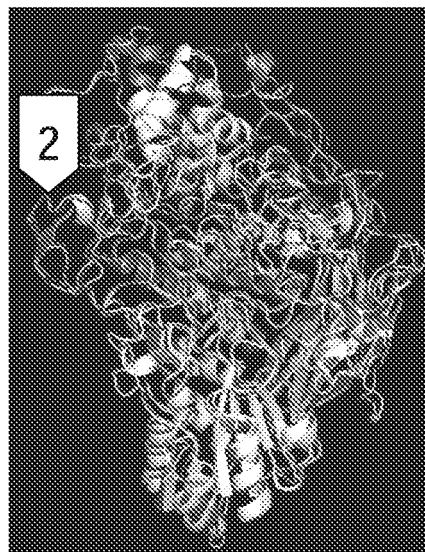


FIG. 105B

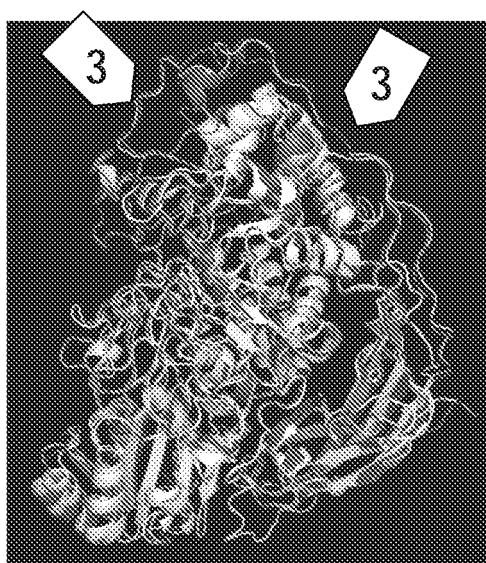


FIG. 105C

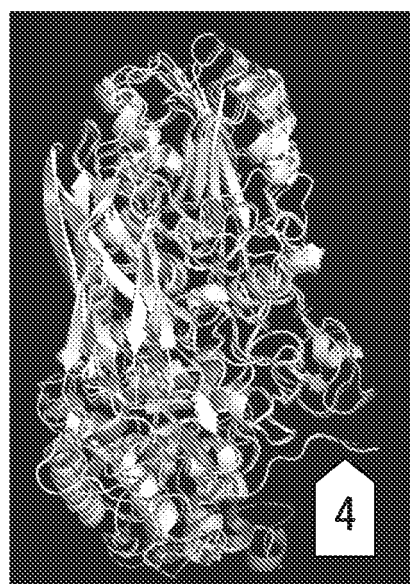


FIG. 105D

132 / 138

Alignment of *T. reesei* BGL1, Te3A, and Fv3C. Mature *T. reesei* BGL1 sequence is shown, mature start of Fv3C and ABC2 indicated by 'Mat'

Q12715 TRI	:	-----	*	20	*	40	*	60	*	120	*	140	*	VVPPAGTPWGTA	:	12
ABC2 T_eme	:	-----	*	20	*	40	*	60	*	120	*	140	*	WVPPAGTPWGTA	:	43
Fv3C	:	-----	*	20	*	40	*	60	*	120	*	140	*	WVPPAGTPWGTA	:	70
Q12715 TRI	:	-----	*	80	*	100	*	120	*	140	*	160	*	YDKAKAALAKINLQDKVIVSGVWNGGPGCVGNTSPASKISYPSLCLQDGLGVRYSTGSTAFTPGVQAA	:	82
ABC2 T_eme	:	-----	*	80	*	100	*	120	*	140	*	160	*	YDKAKAALAKINLQDKVIVSGVWNGGPGCVGNTSPASKISYPSLCLQDGLGVRYSTGSTAFTPGVQAA	:	113
Fv3C	:	-----	*	80	*	100	*	120	*	140	*	160	*	YDKAKAALAKINLQDKVIVSGVWNGGPGCVGNTSPASKISYPSLCLQDGLGVRYSTGSTAFTPGVQAA	:	140
Q12715 TRI	:	-----	*	160	*	180	*	200	*	220	*	240	*	STWDVNLIRERQGFIEEVKASGIHVILGPVAGPLGKTPQGGRNWEGFVDPYLTGTIAMQQTINGIQSVG	:	152
ABC2 T_eme	:	-----	*	160	*	180	*	200	*	220	*	240	*	STWDVNLIRERQGFIEEVKASGIHVILGPVAGPLGKTPQGGRNWEGFVDPYLTGTIAMQQTINGIQSVG	:	183
Fv3C	:	-----	*	160	*	180	*	200	*	220	*	240	*	STWDVNLIRERQGFIEEVKASGIHVILGPVAGPLGKTPQGGRNWEGFVDPYLTGTIAMQQTINGIQSVG	:	210
Q12715 TRI	:	-----	*	220	*	240	*	260	*	280	*	300	*	VQATAKHVILNEQELNR-----ETISSNPDRTLHLYTWPFADAVQANVASVMCSYNKVNVT	:	209
ABC2 T_eme	:	-----	*	220	*	240	*	260	*	280	*	300	*	VQATAKHVILNEQELNR-----ETISSNPDRTLHLYTWPFADAVQANVASVMCSYNKVNVT	:	250
Fv3C	:	-----	*	220	*	240	*	260	*	280	*	300	*	VQATAKHVILNEQELNR-----ETISSNPDRTLHLYTWPFADAVQANVASVMCSYNKVNVT	:	280
Q12715 TRI	:	-----	*	300	*	320	*	340	*	360	*	380	*	TVACEDQYTLQTVLKDQLGFPYVMTDWNAAQHTTVQSANSGLDMSMPG-TDFNGNRLWGPALTNVNSN	:	278
ABC2 T_eme	:	-----	*	300	*	320	*	340	*	360	*	380	*	TVACEDQYTLQTVLKDQLGFPYVMTDWNAAQHTTVQSANSGLDMSMPG-TDFNGNRLWGPALTNVNSN	:	320
Fv3C	:	-----	*	300	*	320	*	340	*	360	*	380	*	TVACEDQYTLQTVLKDQLGFPYVMTDWNAAQHTTVQSANSGLDMSMPG-TDFNGNRLWGPALTNVNSN	:	350
Q12715 TRI	:	-----	*	360	*	380	*	400	*	420	*	440	*	QVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNISR-----NVQGNHKTNR	:	324
ABC2 T_eme	:	-----	*	360	*	380	*	400	*	420	*	440	*	QVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNISR-----NVQGNHKTNR	:	389
Fv3C	:	-----	*	360	*	380	*	400	*	420	*	440	*	QVPTSRVDDMVTRILAAWYLTGQDQAGYPSFNISR-----NVQGNHKTNR	:	420
Q12715 TRI	:	-----	*	440	*	460	*	480	*	500	*	520	*	AIARDGIVLLKNDANILPLK-KPASIAVVGSAIIGNHARNSPSCNDKGCDDGALGMWGSQAVNYPYFV	:	393
ABC2 T_eme	:	-----	*	440	*	460	*	480	*	500	*	520	*	AIARDGIVLLKNDANILPLK-KPASIAVVGSAIIGNHARNSPSCNDKGCDDGALGMWGSQAVNYPYFV	:	456
Fv3C	:	-----	*	440	*	460	*	480	*	500	*	520	*	AIARDGIVLLKNDANILPLK-KPASIAVVGSAIIGNHARNSPSCNDKGCDDGALGMWGSQAVNYPYFV	:	486

FIG. 105E-1

133 / 138

Alignment of <i>T. reesei</i> BGL1, Te3A, and Fv3C. Mature <i>T. reesei</i> BGL1 sequence is shown, mature start of Fv3C and ABG2 indicated by 'Mat'									
	500	*	520	*	540	*	560		
Q12715 TRI	: APYDAINTRASSQGTQVT--LSNTDNTSSGASAAARGKDVAIVF								461
ABG2 T_eme	: TPEQAIQREVLSRNGTFTG-ITDNGALAEMAAASQADTCLVFANAD								525
Fv3C	: TPDQGLSNRATQDGTRYESILTNNEWASVQALVSQPNVTAIVFANAD								556
	*	580	*	600	*	620	*		
Q12715 TRI	: GNALVQAVAGANSNVIVVHVSVAIILEQILALPQKAVVWAGLPSQES								531
ABG2 T_eme	: ADQVIHNVSANCNTIVVLTHTVGPVLI DDWDHPNVTAIILWAGLPGQES								594
Fv3C	: GDELIKNVSSICPNTIVLHTVGPVLLADYEKNPNITAIIVWAGLPGQES								625
	640	*	660	*	680	*	700		
Q12715 TRI	: IAKSPNDYNTRIVSGGSDS-----FSEGLFIDYKHFDDAN-----								582
ABG2 T_eme	: WGRARDDYGAPLIVKPNNGKGAPOQDFTEGIFIDYRRFDKYN-----								652
Fv3C	: WGRTRSYGTEVLYEANNRGAPQDDFSEGVFIDYRHFDRRSPSTDGKSSP								695
	*	720	*	740	*	760	*		
Q12715 TRI	: NYSRLSV-----								598
ABG2 T_eme	: EFSQLNVQPINAPPYTPASGFTKAAQSEFGQ-PSNASDNLVPSDIERV								719
Fv3C	: EYSDLNIQKNVENPSPAGQTIAPPTFGNFESKNLNDYVFPKGVRYIYK								765
	Insertion 3								
	780	*	800	*	820	*	840		
Q12715 TRI	: TGAVV-----								646
ABG2 T_eme	: YGLPTEKYVPPNATNGDPQPIDPPAGGAPGPNPSLYEPVARVTTIIT								787
Fv3C	: FGKTAEEFLPPNALNGSAQPRLLPASGAPGPNQPLWDILYTVTATIT								833
	Loop 3								
	*	860	*	880	*	900	*		
Q12715 TRI	: PPKQLRGFAKL-NLTPGQSGTATFNIRRRDLSYWDTSQKWWVPSGSG								713
ABG2 T_eme	: APKVLRGFDRI-TLAPGQQYLWTTTLTRRDISNWDPTQNWVVTNYTK								856
Fv3C	: PIRVLRGFDRIENIAPGQSAIFNAQLTRRDLNWDTNACQNWVITDHP								899
Q12715 TRI	: - : -								
ABG2 T_eme	: I : 857								
Fv3C	: - : -								

FIG. 105E-2

134 / 138

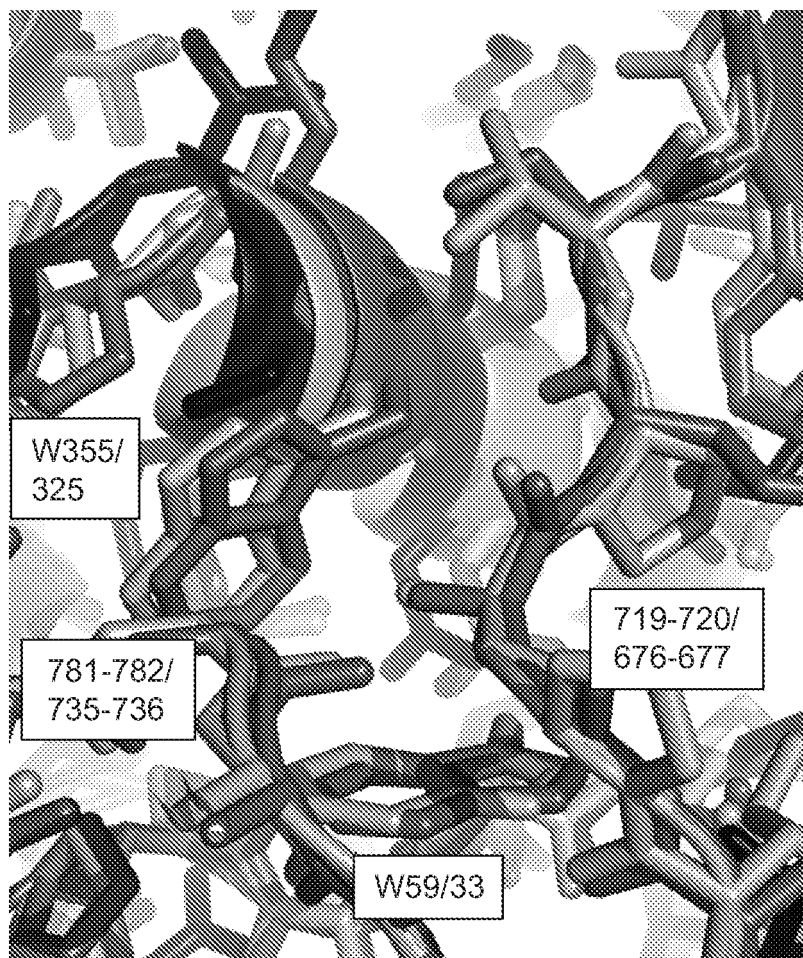


FIG. 105F

135 / 138

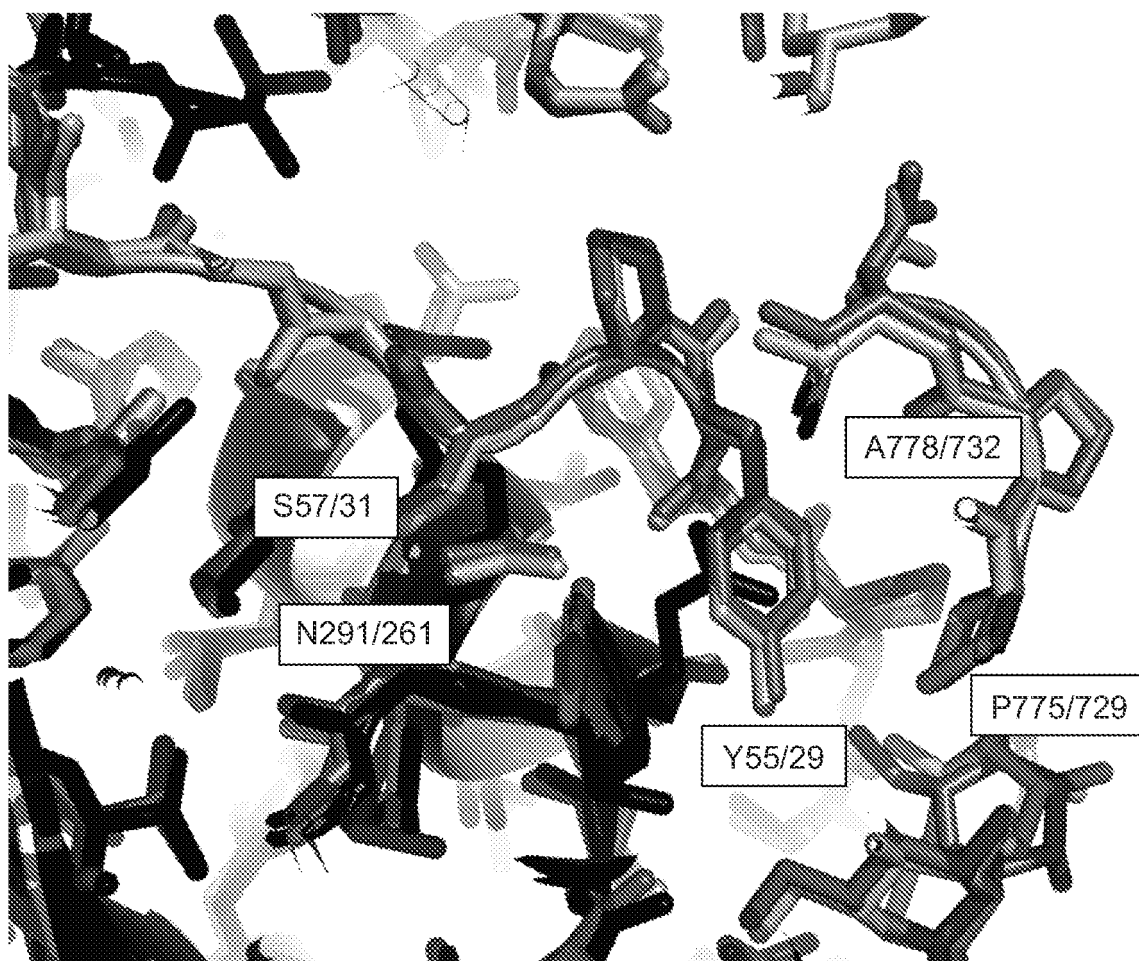


FIG. 105G

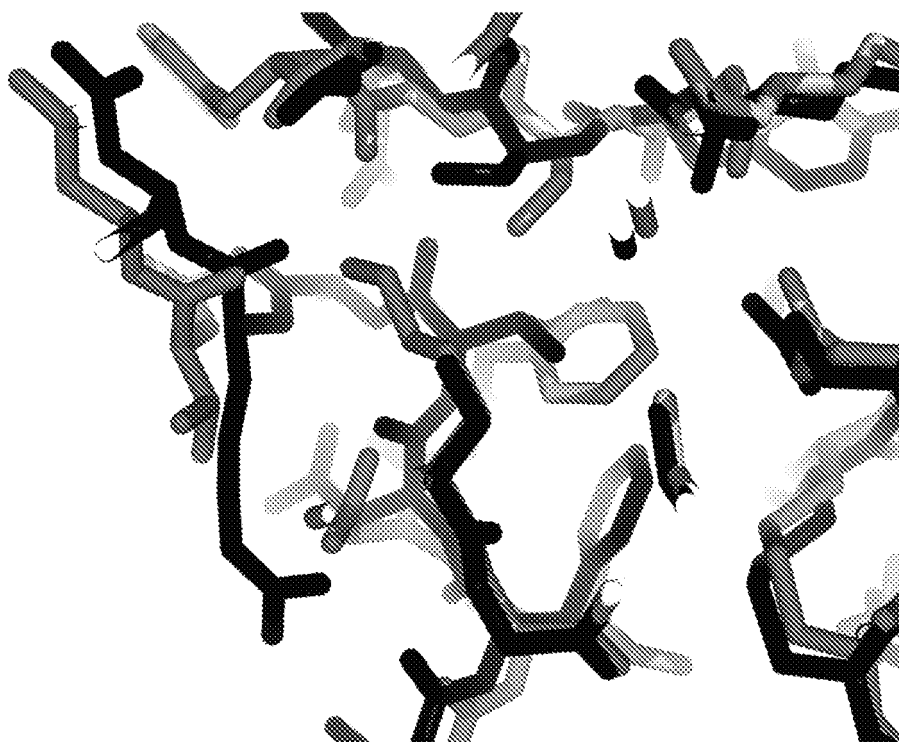


FIG. 105H

136 / 138

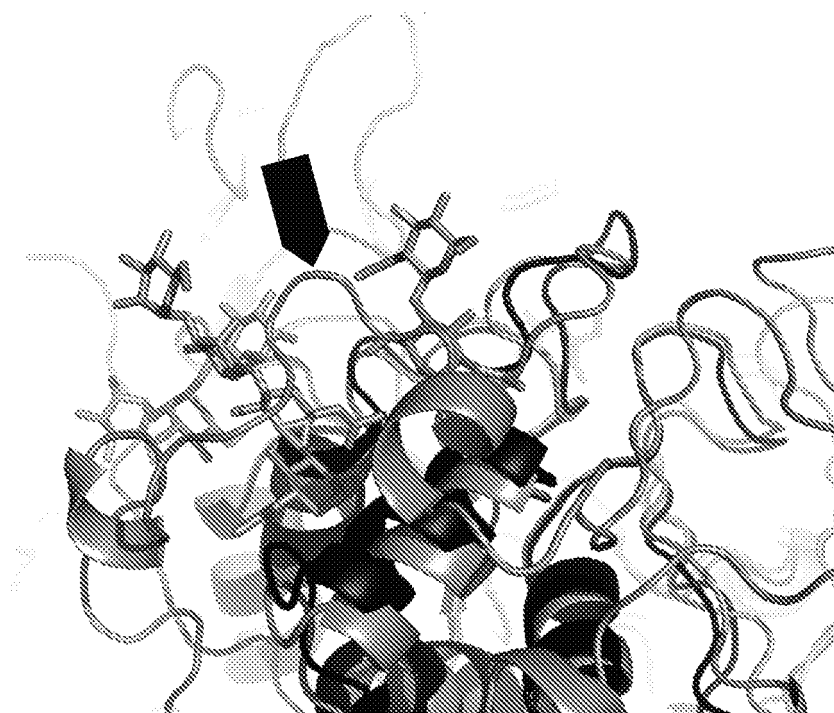


FIG. 105I (a)

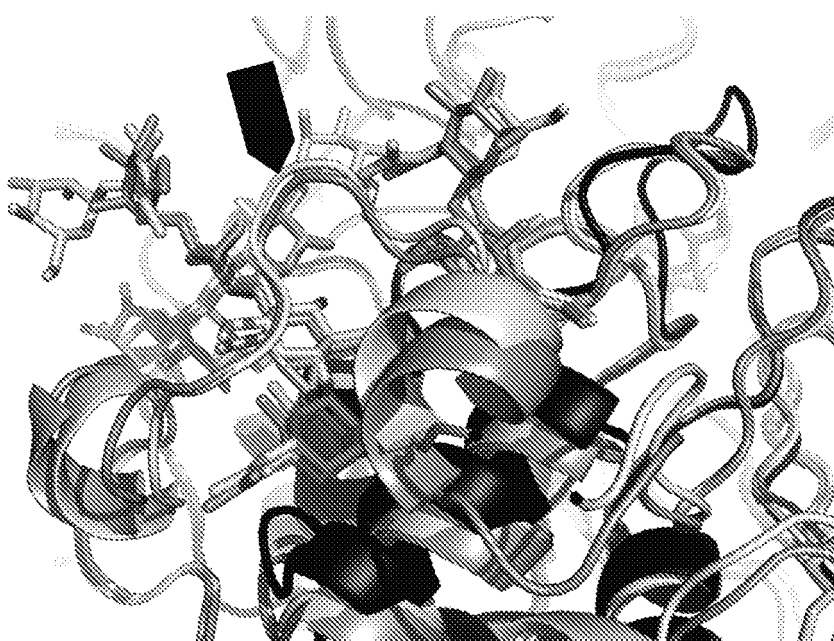


FIG. 105I (b)

137 / 138

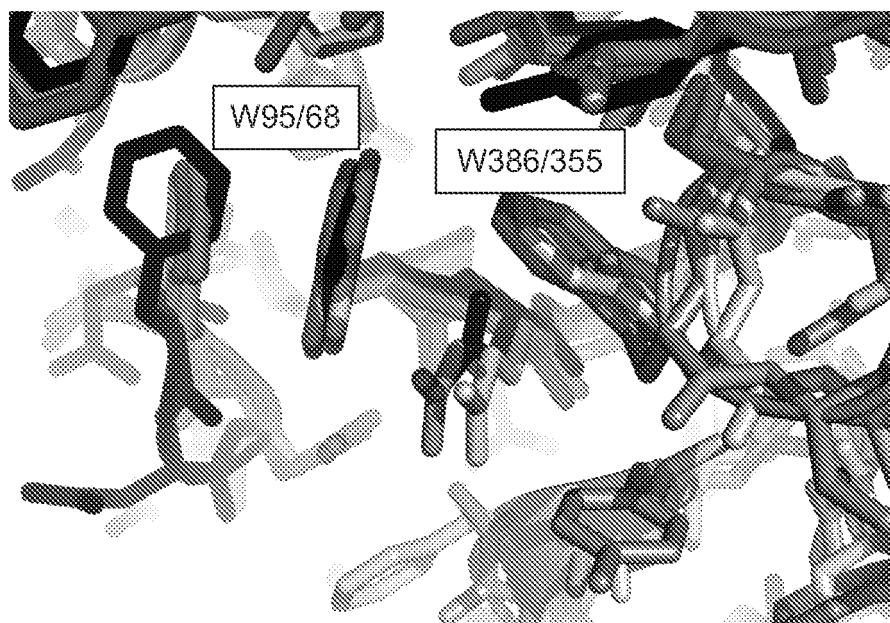


FIG. 105J

An example of a UPLC chromatogram of an enzyme blend:

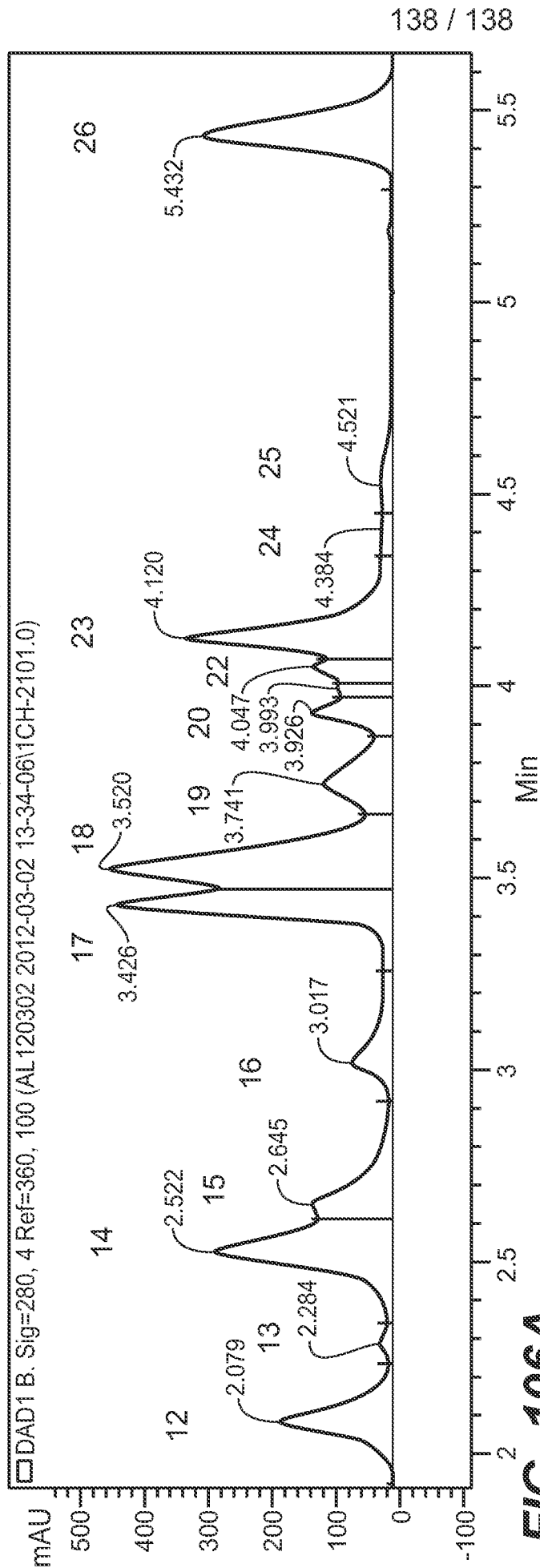


FIG. 106A

A list of major protein components of the enzyme blend used to generate the above UPLC graph:

Protein	Wt. %
CBH1	4 wt. %
CBH2	15 wt. %
GH61/endoglucanase	10 wt. %
Beta-glucosidases	22 wt. %
xylanases	15 wt. %
Beta-xylosidases	12 wt. %
L-a-arabinofuranosidase	1 wt. %
Others	20 wt. %

FIG. 106B