nature plants

Article

https://doi.org/10.1038/s41477-023-01358-4

Structural and biochemical insight into a modular β-1,4-galactan synthase in plants

In the format provided by the authors and unedited

Table of contents

Supplementary Table S1. Primer sequences for generation of GalS1 expression constructs. Primer sequences for construct generation. Fluorescence units and T_m of the resultant soluble secreted sfGFP-fusion proteins. See Extended data Figure 2 for associated schematics

Construct	Primers (5' – 3') Sense (top), Antisense (bottom)	GFP FU*	T_m (°C)[ǂ]
pGEN2-DEST-GalS1 WT	AACTTGTACTTTCAAGGCTCGTTCACCAGTACCTCAT ACAAGAAAGCTGGGTCCTATGAGTAAGCCTGTACATT	2144	59.1 (0.998)
pGEN2-DEST-GalS1 ΔSTEM	GCCGCCGCTCTCTTTGTT GCCTTGAAAGTACAAGTTTTCAGAG	774	57.4 (0.998)
pGEN2-DEST- CBM 95	GTATCTTTACtaaGGGTCTTCTTTGTAC TCGTACTGATAAGGTGGG	1062	57.7 (1.000)
pGEN2-DEST-GalS1 NΔCBM	TATCAGTACGAGTATCTTTAC GCCTTGAAAGTACAAGTTTTC	992	N.D.

*Expression and secretion of the recombinant GFP-fusion proteins and protein variants in transiently transfected HEK293F cells. Expression levels are determined by measuring the relative GFP fluorescence (FU) of the secreted proteins in clarified media.

[ǂ]Protein thermal stability (T_m) determined using SYPRO Orange Fluorescent Dye with R-squared in brackets.

N.D., not determined.

Supplementary Table S2. Primer sequences for generation of GalS1 variants in the GT92 domain. Primers sequences for site-directed mutagenesis and fluorescence units and T_m of the resultant soluble secreted sfGFP-fusion proteins. All site-directed mutagenesis was carried out using the pGEN2-DEST-GalS1WT plasmid as a template.

Variant	Primers (5' – 3') Sense (top), Antisense (bottom)	GFP FU*	T_m (°C)[†]
[€]G242A	TTCTTTGTACgccAATTTGAGTGCG GACCCGCAGTAAAGATAC	1961	51.0 (0.990)
[€]H437A	TTACCACTACgccAATTCCATACAAGTAC TATCGAATCTTTGTTTCAGTC	2383	61.1 (0.998)
[€]R396A	AACTGGGATAgccAGGGATCGCAAG CTTGATTCTCTGAATAAAAGTTTTTC	1546	55.7 (0.990)
[€]D398A	GATACGAAGGgccCGCAAGTACG CCAGTTCTTGATTCTCTGAATAAAAG	1637	60 (1.000)
[€]W166A	GCTTCCTGACgccGGCTATGGTC ATCTTATAAGCCTTGGCC	1437	49.0 (0.990)
[€]D333A	CTTTGATGTGgccGAGTATATCTATTTGC TAAAATGTCCAATTAGCAGAG	1549	51.4 (0.998)
[€]E334A	TGATGTGGATgccTATATCTATTTGCCCG AAGTAAAATGTCCAATTAGC	1989	45.7 (0.998)
[€]K400A	AAGGGATCGCgccTACGCAATACAGG CGTATCCCAGTTCTTGATTC	2129	57.1 (0.998)
[€]Q309A	TTACTATAACgccTTTCTGGTGGTGAATGACTG TACCCATCAAACTCAGCTTG	1929	58.2 (0.998)
[€]H414A	TACAGGAGTGgccATGTCGGAGAATG GCATATGCATTCTTAGCC	296	N.D.
[€]D331A	ATTTTACTTTgcgGTGGATGAGTATATCTATTTG GTCCAATTAGCAGAGTATC	1840	55.3 (0.997)
[€]H435A	TCGATATTACgccTACCACAATTCCATAC ATCTTTGTTTCAGTCTGG	779	63.6 (0.998)
[€]C316S & C236S	GGTGAATGACagcTTGCATCGCT ACCAGAAATTGGTTATAGTAATACC GTATCTTTACagcGGGTCTTCTTTG TCGTACTGATAAGGTGGG	125	N.D.
[€]C236S	GTATCTTTACagcGGGTCTTCTTTG TCGTACTGATAAGGTGGG	390	N.D.

[€]Lowercase nucleotides in the primer sequences represent the bases changed to generate mutant variants.
*Expression and secretion of the recombinant GFP-fusion proteins and protein variants in transiently transfected HEK293F cells. Expression levels are determined by measuring the relative GFP fluorescence (FU) of the secreted GFP fusion proteins in clarified media.
[†]Protein thermal stability (T_m) determined using SYPRO Orange Fluorescent Dye with R-squared in brackets.
N.D., not determined.

Supplementary Table S3. Primer sequences for site-directed mutagenesis for generation of CBM95 mutant variants. Primers used to generate the CBM95 variants and GFP fluorescence units (FU) of soluble sfGFP-CBM95 variants secreted into the media.

CBM95 Variants	Primers (5' – 3') Sense (top), Antisense (bottom)	GFP FU*
€K133A	CCTGGCTTCGgccCCCATTCACG CCAACCACAGCAAAGGTG	1588
€W142A	CAGGCTTCCCgccTACAAATGCGAGTGG AACACGTGAATGGGCTTC	1263
€K144A	TCCCTGGTACgccTGCGAGTGGATC AGCCTGAACACGTGAATG	1286
€K161A	CAAGGCTTATgccATGCTTCCTGACTG GCCCTTATAGAAGAGCCATTG	Not Expressed
€W166A	GCTTCCTGACgccGGCTATGGTC ATCTTATAAGCCTTGGCC	1481
€Y199A	GCTTAATGCCgccTATGACGAGTCTCAGAGGAAGTAC ATGAGCCTTCCACCCGCA	1015
€K206A	GTCTCAGAGGgccTACGAGAAATTCACGG TCGTCATAATAGGCATTAAG	1184
€Y207A	TCAGAGGAAGgccGAGAAATTCACG GACTCGTCATAATAGGCATTAAG	1425
€K209A	GAAGTACGAGgccTTCACGGCATTG CTCTGAGACTCGTCATAATAG	1201

*Expression and secretion of the recombinant GFP-fusion proteins and protein variants in transiently transfected HEK293F cells. Expression levels are determined by measuring the relative GFP fluorescence (FU) of the secreted GFP fusion proteins in clarified media.

€Lowercase nucleotides in the primer sequences represent the bases changed to generate mutant variants.

Supplementary Table S4. **Summary of crystal parameters, data collection, and refinement statistics.** Values in parentheses are for the highest resolution shell.

	GalS1	**GalS1 - Mn^{2+}**	**GalS1 - SeMet**
Wavelength (Å)	1.000	1.000	0.97949
Resolution range (Å)	67.83 – 2.37 (2.45 – 2.37)	49.07 - 2.56 (2.652 - 2.56)	68.09 - 3.61 (3.739 - 3.61)
Space group	P1	P3121	P1
Unit-cell parameters (Å)	a = 70.39, b = 83.42, c = 94.68, α = 76.16, β = 69.66, γ = 88.56	a = 169.98, b = 169.98, c = 73.96, α = β = 90, γ = 120	a = 70.40, b = 83.34, c = 95.25, α = 75.96, β = 69.81, γ = 88.53
Total reflections	202299 (19899)	79592 (7908)	184695 (18365)
Unique reflections	78131 (7586)	39800 (3954)	22479 (2252)
Multiplicity	2.6 (2.6)	2.0 (2.0)	8.2 (8.2)
Completeness (%)	98.2 (94.9)	99.96 (99.97)	99.97 (100.00)
Mean I / sigma (I)	7.18 (1.07)	20.31 (2.35)	15.04 (9.54)
Wilson B-factor	36.67	68.99	38.43
R-merge*	0.103 (1.036)	0.018 (0.350)	0.1228 (0.2211)
CC1/2	0.989 (0.506)	1 (0.755)	0.992 (0.984)
Reflections used in refinement	78079 (7585)	39793 (3953)	
Reflections used for R-free	4010 (387)	2005 (201)	
R-work §	0.197 (0.292)	0.235 (0.318)	
R-free ¶	0.247 (0.348)	0.267 (0.391)	
Number of non-hydrogen atoms	14272	6930	
Macromolecules	12933	6556	
Ligands	699	335	
Solvent	640	39	
Protein residues	1573	798	
RMSD bonds (Å)	0.005	0.009	
RMSD angles (°)	0.980	1.24	
Ramachandran favored (%)	96.29	92.44	
Ramachandran allowed (%)	3.71	6.55	
Ramachandran outliers (%)	0	1.01	
Rotamer outliers (%)	0.07	0.44	
Clashscore	6.17	17.43	
Average B-factor	47.6	108.20	
Macromolecules	46.2	107.15	
Ligands	79.9	133.24	
Solvent	40.3	70.35	
PDB ID	8D3T	8D3Z	

* $R_{merge} = \sum_h \sum_i | I_i(h) - \langle I(h) \rangle | / \sum_h \sum_i I_i(h)$, where $I_i(h)$ is the intensity of an individual measurement of the reflection and $\langle I(h) \rangle$ is the mean intensity of the reflection.

§ $R_{work} = \sum_h ||F_{obs}| - |F_{calc}|| / \sum_h |F_{obs}|$, where F_{obs} and F_{calc} are the observed and calculated structure-factor amplitudes, respectively.

¶ R_{free} was calculated as R_{cryst} using 5.0 % of randomly selected unique reflections that were omitted from the structure refinement.