Supporting Information

Vander Kooi et al. 10.1073/pnas.1009386107

SI Text

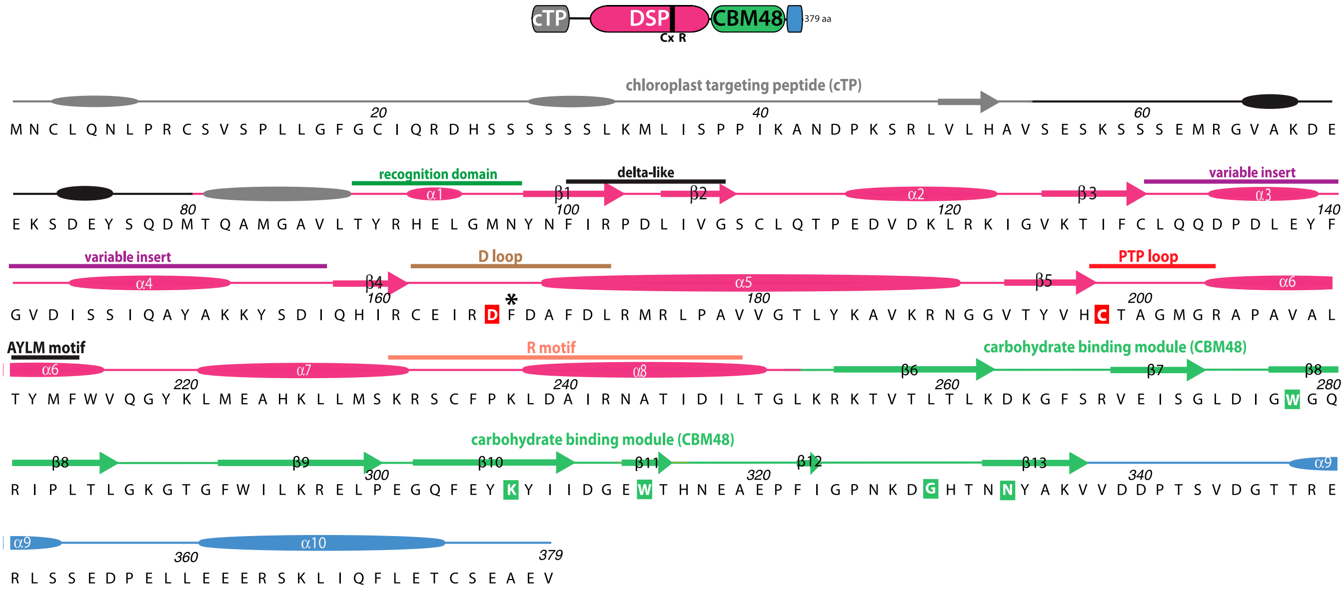


Fig. S1. Schematic of *A. thaliana* SEX4 depicting the domain topography and secondary structure of chain A in the asymmetric unit. The chloroplast targeting peptide (cTP) is gray, the dual specificity phosphatase (DSP) domain is pink, the carbohydrate binding module family 48 (CBM48) is green, the C-terminal domain is blue, and linker regions are black. The different subdomains within the DSP domain are labeled per standard nomenclature (33). α-helices and β-sheets within the DSP domain are numbered 1–8 and 1–5, respectively, the β-sheets within the CBM are numbered 6–13, and the α-helices in the C-terminal domain 9–10. The active site cysteine (C198) and aspartic acid (D166) are highlighted by a red box. F167 is highlighted by an asterisk. Residues known to disrupt binding to glucans are highlighted with a green box.

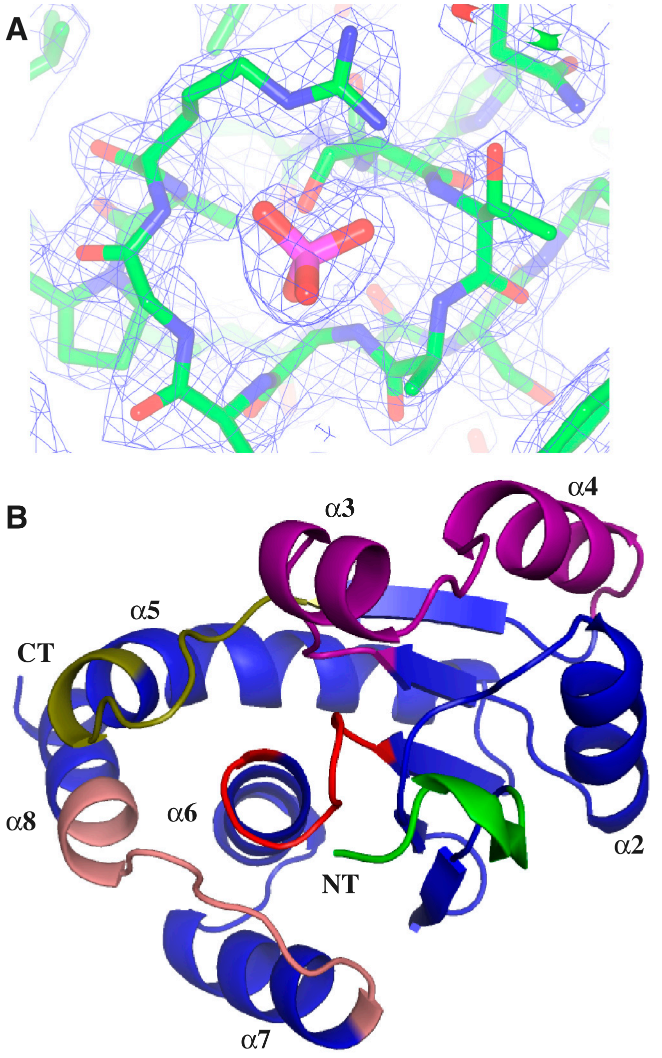
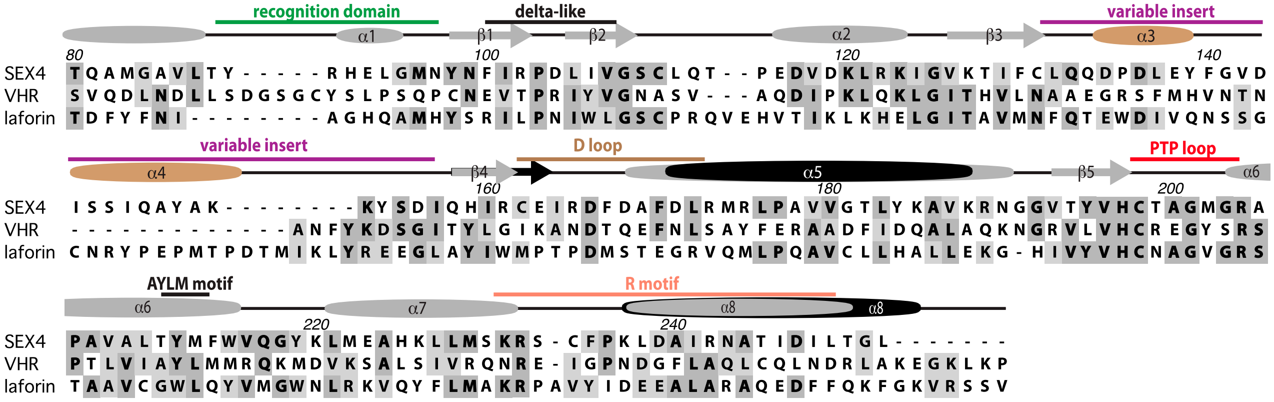


Fig. S2. DSP domain structure. (*A*) The phosphatase active site P-loop of SEX4 including bound phosphate ion (pink and red) shown with 2Fo-Fc electron density map contoured at 1.6σ. (*B*) The DSP domain of SEX4 displayed in the classical orientation and with each subdomain colored. Green, recognition domain; purple, variable insert; beige, D-loop; red, active site; salmon, R-motif.



```
SEX4     TQAMGAVLTY-----RHELGMNYNFIRPDLIVGSCLQT--PEDVDKLRKIGVKTIFCLQQDPDLEYFGVD
VHR      SVQDLNDLLSDGSGCYSLPSQPCNEVTPRIYVGNASV---AQDIPKLQKLGITHVLNAAEGRSFMHVNTN
laforin  TDFYFNI--------AGHQAMHYSRILPNIWLGSCPRQVEHVTIKLKHELGITAVMNFQTEWDIVQNSSG

SEX4     ISSIQAYAK--------KYSDIQHIRCEIRDFDAFDLRMRLPAVVGTLYKAVKRNGGVTYVHCTAGMGRA
VHR      -------------ANFYKDSGITYLGIKANDTQEFNLSAYFERAADFIDQALAQKNGRVLVHCREGYSRS
laforin  CNRYPEPMTPDTMIKLYREEGLAYIWMPTPDMSTEGRVQMLPQAVCLLHALLEKG-HIVYVHCNAGVGRS

SEX4     PAVALTYMFWVQGYKLMEAHKLLMSKRS-CFPKLDAIRNATIDILTGL-------
VHR      PTLVIAYLMMRQKMDVKSALSIVRQNRE-IGPNDGFLAQLCQLNDRLAKEGKLKP
laforin  TAAVCGWLQYVMGWNLRKVQYFLMAKRPAVYIDEEALARAQEDFFQKFGKVRSSV
```

Fig. S3. Multiple sequence alignment of the DSP domain of *A. thaliana* SEX4, *H. sapiens* VHR, and *H. sapiens* laforin. Numbering refers to the amino acids of SEX4. The different subdomains within the DSP domain are labeled per standard nomenclature (33). SEX4 secondary structure is displayed in gray. α-helices 3 and 4 are in tan because they are shared between SEX4 and the predicated secondary structure of laforin. The black α-helices and β-sheet are predicted secondary elements of laforin and VHR.

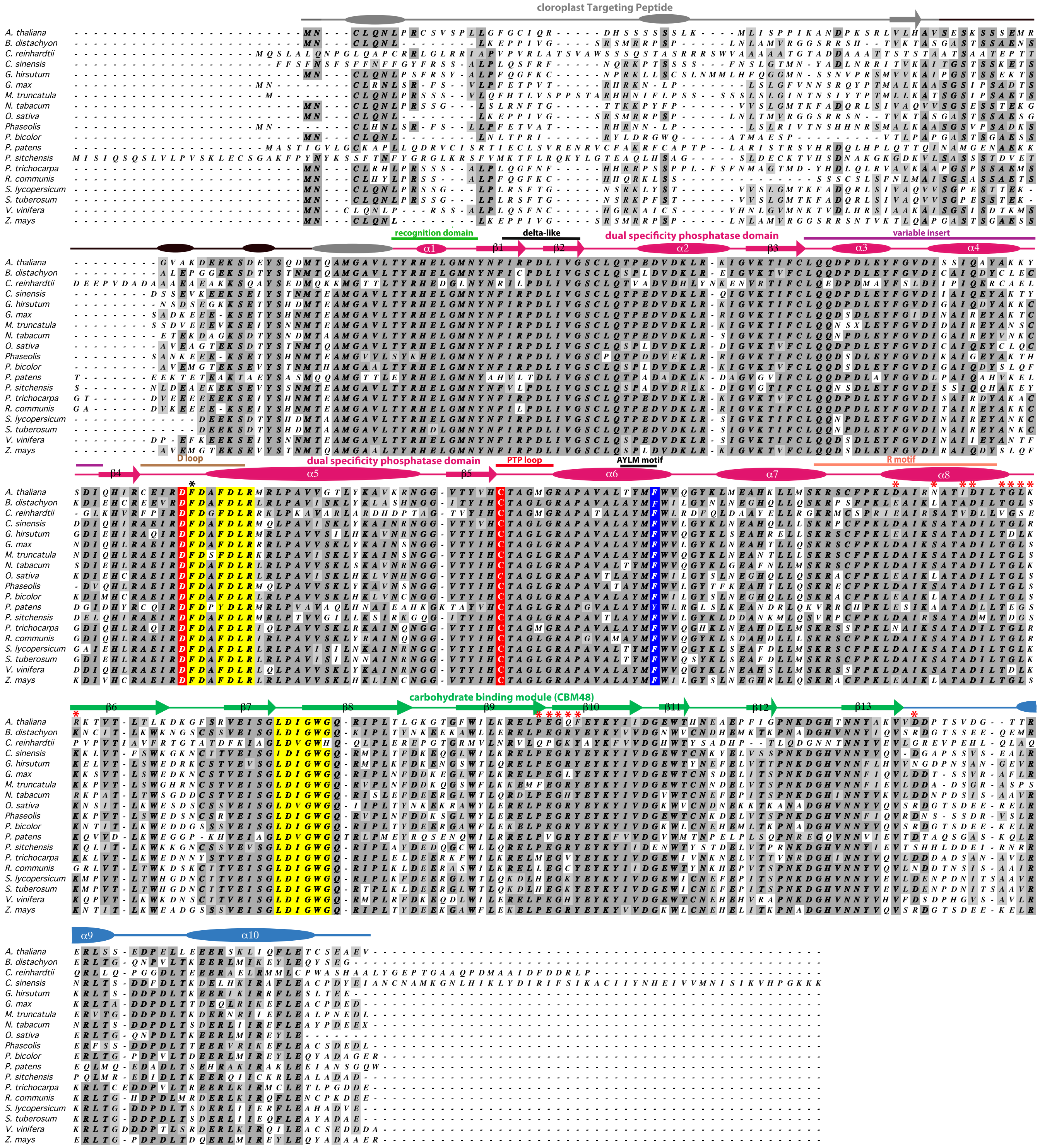


Fig. S4. Multiple sequence alignment of SEX4 orthologs. Secondary structure and elements are labeled as in Fig. S1. The catalytic residues D166 and C198 are boxed in red. F167 is marked with a black asterisk. F214 within the AYLM motif is boxed in blue. The α5 region of the DSP and the β7/8 loop of the CBM that pack together are both boxed in yellow. Residues within α8 of the DSP domain that pack against β10 and the end of the CBM are all marked with red asterisks.

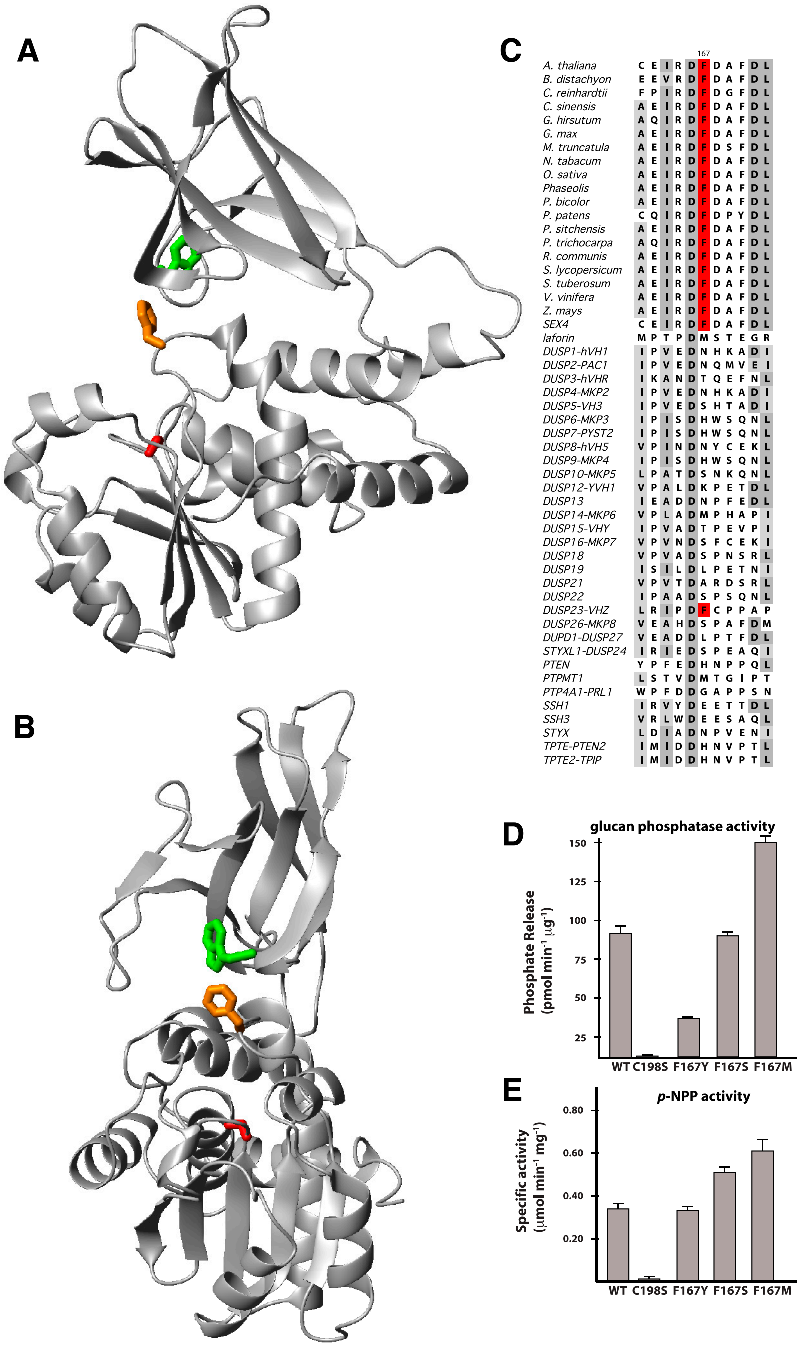


Fig. S5. F167 is critical for SEX4 function as a glucan phosphatase. (*A*) Ribbon diagram of SEX4 (residues 90–379) with the active site S198 labeled with a red stick, F167 in yellow, and W278 in green. This is the same view as in Fig. 1*C*. (*B*) Ribbon diagram as in *B* but rotated 90°. This is the same orientation as the zoom in Fig. 4*A*. (*C*) Multiple sequence alignment of the D-loop of SEX4 orthologs and multiple DSPs. (*D*) Release of phosphate by Δ89-SEX4 (WT) and SEX4 mutants utilizing the specific glucan substrate amylopectin. Error bars indicate means ± the standard deviation. (*E*) Specific activity of Δ89-SEX4 (WT) and SEX4 mutants against the generic substrate *p*-NPP.

Table S1. Percent similarity and identity of SEX4 ortholgs to *A. thaliana* SEX4 full-length protein (At-SEX4), the DSP domain (At-SEX4 DSP), and the CBM (At-SEX4 CBM48) (similarity/identity)

	At-SEX4	At-SEX4 DSP	At-SEX4 CBM48
A. thaliana	100/100	100/100	100/100
B. distachyan	68/57	86/75	71/57
C. reinhardtii	49/34	71/56	52/34
C. sinensis	64/54	93/86	76/64
G. hirsutum	73/62	93/86	76/62
G. max	72/60	90/83	74/65
M. truncatula	71/58	89/81	71/57
N. tabacum	73/64	90/85	68/62
O. sativa	67/57	89/77	66/55
Phaseolis spp.	71/59	86/77	74/62
P. bicolor	68/58	87/77	70/62
P. patens	56/42	72/60	63/43
P. sitchensis	61/48	88/75	67/57
P. trichocarpa	73/61	92/83	70/62
R. communis	73/63	93/82	72/68
S. lycopersicum	73/62	89/83	71/60
S. tuberosum	73/62	89/82	71/60
V. vinifera	71/60	91/83	72/65
Z. mays	68/57	87/76	68/60