Supplementary material associated to the manuscript " Is there a role for tau glutathione transferases in tetrapyrrole metabolism and retrograde signaling in plants? " by Sylvestre-Gonon et al.

"Retrograde signalling" theme issue of Philosophical Transactions of the Royal Society B

Methods

Cloning, protein expression and purification

The sequences coding for GSTU8 (At3g09270), GSTU16 (At1g59700), GSTU19 (At1g78380) and GSTU22 (At1g78340) from *Arabidopsis thaliana* and GSTU1 from *Zea mays* were cloned by PCR into pET-26b plasmids for the production of C-terminal his-tagged proteins. The cloning of Arabidopsis GAPC1 (At3g04120) in pET28a was described previously [1]. Expression was performed using the *Escherichia coli* Rosetta2 (DE3) pLysS strain (Novagen). The cells were classically grown at 37°C in LB medium supplemented with kanamycin (50 $\mu g.ml^{-1}$) and chloramphenicol (34 $\mu g.ml^{-1}$). When the cell culture reached an OD_{600nm} of 0.8, protein expression was induced with 0.1 mM IPTG and cells were grown for a further 4 h at this temperature. Alternatively, to boost bacterial heme synthesis, 0.25 mM $FeCl_3$ and 1.5 mM 5-aminolevulinic acid were added in the medium 2 h before the induction of protein expression with 0.1 mM IPTG performed when the cell culture reached an OD_{600nm} of 0.8. In this case, the cells were additionally grown for 18 h at 30°C. In all cases, cells were harvested by centrifugation, resuspended in a 30 mM Tris-HCl pH 8.0 and 200 mM NaCl lysis buffer and stored at -20°C. Cell lysis was completed by sonication (two times for 1 min with an interval of 1 min). The cell extract was then centrifuged at 42,000 g for 30 min at 4°C to remove cellular debris and aggregated proteins. The C-terminal His_6-tagged proteins were purified as follows: 10 mM imidazole was added in the fraction containing soluble proteins before loading onto a Ni-NTA (Ni^{2+}-nitrilotriacetate)–agarose resin (Qiagen). After a washing step with lysis buffer supplemented with 20 mM imidazole, the proteins were eluted using lysis buffer containing 250 mM imidazole. The fractions of interest were pooled, concentrated by ultrafiltration (Vivaspin turbo 15, Sartorius) then injected onto a gel filtration HiLoad 16/600 Superdex 200 prep grade (GE Healthcare) column connected to an ÄKTA purifier FPLC system (GE Healthcare) and eluted with 30 mM Tris-HCl pH 8.0 and 200 mM NaCl buffer. Again, fractions of interest were pooled, concentrated by ultrafiltration (Vivaspin, Sartorius) and stored in 30 mM Tris-HCl pH 8.0 and 200 mM NaCl buffer at -20°C until use. The homogeneity of the

purified proteins was checked by SDS/PAGE. The concentrations of apo-proteins were determined by measuring the absorbance at 280 nm and using theoretical molar absorption coefficients of 46410, 36340, 44920 $M^{-1}.cm^{-1}$ for AtGSTU8, AtGSTU16, AtGSTU19 respectively, and of 40910 $M^{-1}.cm^{-1}$ for AtGSTU22, ZmGSTU1 and AtGAPC1 as deduced from the amino acid sequences. The concentrations of holo-proteins were determined using a bicinchoninic acid assay (Interchim, France). For electrospray ionization mass spectrometry analyses, AtGSTU8 was desalted onto a PD MidiTrap G-25 column (GE Healthcare) against 30 mM Tris-HCl pH 8.0 and analyzed as described [2]. All proteins were also analyzed after a liquid chromatography separation using the same mass spectrometry equipment.

Analytical gel filtration

100 µg of AtGSTU8 (final concentration of 3 µM) were incubated or not with 100 µM of PPIX during 1 h in 250 µL of 30 mM Tris-HCl pH 8.0, 200 mM NaCl buffer. Samples were centrifuged at 13,400 rpm (minispin, Eppendorf) during 5 min at room temperature and then injected onto a gel filtration 10/300 GL Superdex 200 (GE Healthcare) column connected to an ÄKTA purifier FPLC system (GE Healthcare) and eluted with 30 mM Tris-HCl pH 8.0 and 200 mM NaCl buffer. Data processing was carried out with Unicorn 5.20 software.

Intrinsic fluorescence of tryptophan

Fluorescence measurements were performed on a Cary Eclipse (Varian) fluorescence spectrophotometer. The binding affinity of enzymes for porphyrins was determined by tryptophan fluorescence quenching titrations. The excitation wavelength was set at 290 nm, and the emission recorded between 300 and 400 nm. Titration was carried out in 30 mM of Tris-HCl pH 8.0, 1 mM EDTA buffer with 5 µM of protein and increasing concentrations of PPIX from 0 to 150 µM and of hemin and hematin from 0 to 100 µM. To determine the dissociation constant (K_d) of GSTs for porphyrins, the following equation (Equation 1) was applied using GraphPad Prism 6.0 software:

$$\Delta F / F_{max} = \frac{B_{max} \times [S]}{K_d + [S]}$$

where $\Delta F/F_{max}$ ratio represents the changes in fluorescence intensity relative to the initial value (*i.e.* without ligand) after addition of PPIX at a concentration [S], B_{max} is the maximum number of specific binding sites and K_d corresponds to the dissociation constant.

Molecular docking study

In order to dock PPIX onto AtGSTU23, the coordinates of the ligand (ZINC4208846) from the ZINC15 database [3], together with the coordinates of AtGSTU23 crystal structure in complex with GSH (pdb code 6EP7) were prepared for docking with AutoDock Tools [4]. A search box of 40 $Å^3$ was defined at the center of the active site. Docking calculation was performed with AutoDock Vina [5] using default parameters and side chains lining the active site were set as flexible. The predicted PPIX binding poses returned by the software were visualized with Pymol (The PyMOL Molecular Graphics System, Schrödinger, LLC). The most probable binding pose of PPIX has a configuration such that the negatively charged propionic groups are oriented toward the solvent whereas the tetrapyrrole ring sits in a hydrophobic pocket. The figures 3 and 4 were prepared with Pymol.

Figure S1. Unrooted phylogenic tree showing that the selected AtGSTUs belong to the three major clades.

The sequences used for this analysis are the GSTUs identified in *Arabidopsis thaliana* as well as *Zea mays* GSTU1. Sequences were aligned and poorly aligned positions and divergent regions removed using seaview software [6]. The phylogenetic tree was then constructed with BioNJ [7] and edited with Figtree software (http://tree.bio.ed.ac.uk/software/figtree/). The robustness of the branches was assessed by the bootstrap method with 1000 replications. The scale marker represents 0.1 substitutions per residue.

Figure S2. UV-visible absorption spectra of recombinant AtGSTU8, AtGSTU16, AtGSTU19, AtGSTU22 and ZmGSTU1.

All these proteins were expressed in *E. coli* using a medium supplemented with $FeCl_3$ and aminolevulinic acid, purified in a single IMAC step and dialyzed before recording a UV-Visible absorption spectrum between 250 and 650 nm. AtGSTU8 is in blue, AtGSTU16 in red, AtGSTU19 in green, AtGSTU22 in violet and ZmGSTU1 in orange. In the inset, focus on the 350-450 nm region showing differences in the absorbance band maxima (λ_{max}).

Figure S3. Electrospray ionization (ESI) mass-to-charge ratio (m/z) spectra of AtGSTU8, AtGSTU16, AtGSTU19 and ZmGSTU1 focusing on their co-purified metabolite obtained using liquid chromatography coupled to a Bruker microTOF-Q spectrometer.

A molecule with a molecular mass of 616.1604 or 616.1861 Da corresponding to a heme b was detected for AtGSTU8 and AtGSTU16 respectively. A molecule with a molecular mass of

609.2541 or 609.2816 Da corresponding to harderoporphyrin was detected for AtGSTU19 and ZmGSTU1 respectively. Intensity of the signal is represented as arbitrary units (AU).

Figure S4. Multiple sequence alignment of GSTUs from *Arabidopsis thaliana*.

This multiple alignment was performed with ClustalW implemented in BioEdit. Proteins that present biochemical evidence for PPIX-binding properties are highlighted in magenta including *Zm*GSTU1. The numbering is based on AtGSTU1 sequence. The structure of AtGSTU23 used for docking study is underlined. Putative catalytic signature is coloured in red. Conserved residues are highlighted in yellow and their positions are indicated below the alignment by the symbols G, H and * for residues of the glutathione binding site, of the hydrophobic site and putative PPIX-binding residues identified through docking studies, respectively (this work and [8]).

Class & symbol		**Occurrence**	**Oligomeric state**	**Catalytic residue**	**Catalytic and non-catalytic activities**	**Interaction with porphyrin moieties**
Alpha	A	Mammals, birds, fishes	Dimer	Tyr	GSH-conjugation, peroxide reduction, isomerase	Yes, ligandin/catalytic sites
Beta	B	Bacteria	Dimer	Cys*****	GSH-conjugation, thiol transferase	
Chi	C	Bacteria	Dimer	?	GSH-conjugation, thiol transferase	
Delta	D	Insects	Dimer	Ser	GSH-conjugation	
Epsilon	E	Insects	Dimer	Ser	Peroxide reduction	
Phi	*F*	*Plants, fungi, bacteria, protists*	*Dimer*	*Ser*	*GSH-conjugation, peroxide reduction, isomerase, ligandin, deglutathionylation******	*Yes, ligandin/catalytic sites*
Eta	H	Bacteria	Dimer	Arg	GSH-conjugation, peroxide reduction	
Hemerythrin	*H*	*Plants*	*?*	*Cys*	*?*	
Iota	*I*	*Plants*	*?*	*Cys*	*?*	
Kappa	K	Mammals, bacteria	Dimer	Ser	GSH-conjugation, peroxide reduction, ligandin	
Lambda	*L*	*Plants*	*Monomer*	*Cys*	*Thiol transferase, deglutathionylation, ligandin*	
Mu	M	Mammals, arthropodes, protozoa, cestoda	Dimer	Tyr	GSH-conjugation, peroxide reduction, isomerase, ligandin	Yes, ligandin site
Nu/Ure2p	*N*	*Plants, fungi, nematodes, bacteria*	*Dimer*	*Tyr/Asn*	*GSH-conjugation, thiol-transferase, peroxide reduction, ligandin*	
Omega	O	Mammals, fungi, insects	Dimer	Cys/Ser	GSH-conjugation, thiol transferase, dehydroascorbate (DHA) reductase, ion channel modulation	
Pi	P	Mammals, nematodes	Dimer	Tyr	GSH-conjugation, peroxide reduction, isomerase, NO transport/scavenging	Yes, catalytic site
Rho	R	Fishes, mollusca	Dimer	Ser	GSH-conjugation	
Sigma (including PGDS, PfGST)	S	Animals, insects, algae, trematodes, apicomplexa	Dimer	Tyr	GSH-conjugation, peroxide reduction, isomerase, prostaglandin synthesis	Yes, ligandin/catalytic sites
Theta	*T*	*Animals, plants, insects, algae, bacteria*	*Dimer*	*Ser*	*GSH-conjugation, peroxide reduction*	

Tau	*U*	*Plants*	*Dimer*	*Ser*	*GSH-conjugation, peroxide reduction, ligandin*	*Yes, ligandin/catalytic sites*
GHR (Xi)	-	*Plants, fungi, bacteria, archaea*	*Dimer*	*Cys*	*DHA reductase, thiol transferase, deglutathionylation*	
Zeta	*Z*	*Animals, plants, insects, fungi, bacteria*	*Dimer*	*Ser*	*GSH-conjugation, peroxide reduction, isomerase*	
CLIC	-	Invertebrates, mammals, fishes	Dimer/Monomer	Cys/Asp	Ligandin, ion channel formation and modulation	
DHAR	-	*Plants*	*Monomer*	*Cys*	*DHA reductase, thiol transferase, deglutathionylation*	
EF1B	-	*Animals, plants, fungi*	*Monomer/Dimer*	*?*	*GSH-conjugation, protein translation*	
Grx2	-	Bacteria	Monomer	Cys	GSH-conjugation, esterase, thiol-transferase	
GSTFuA (GTE)		Fungi	Dimer	Ser/Gly	GSH-conjugation, transferase, ligandin, esterase	
GTT	-	Fungi	Dimer	?	GSH-conjugation, peroxide reduction	
MAPEG (PfEXP1)		Animals, plants, fungi, bacteria, apicomplexa	Trimer	?	GSH-conjugation, peroxide reduction	Yes, *catalytic site*
mPGES2	-	*Mammals, plants, birds, fishes*	*Dimer*	*Cys*	*Isomerase*	*Yes, catalytic site*
Metaxin		*Animals, Fungi, Plants*	*?*	*Ser/Cys*	*Membrane protein import in mitochondria*	
SSPA	-	Bacteria (*Francisella*)	Dimer	Asp ?	Ligandin ?	
TCHQD	-	*Plants, bacteria*	*Dimer*	*Ser*	*Dehalogenase*	

Table S1. Occurrence of GST classes and associated properties.

The classes found in plants are in italics. Note that Prostaglandin-D synthases (PGDS) and *Plasmodium falciparum* GST (PfGST) have been grouped with sigma class GSTs based on their sequence similarity. Noteworthy, the Iota and Xi nomenclature was also used for defining some atypical GSTs in *Aedes aegypti* [9] but they are different from those indicated in this table. Although structurally divergent, MAPEG (membrane-associated proteins in eicosanoid and glutathione metabolism) proteins have been listed here because they exhibit GST-type activities and a

MAPEG-like from *P. falciparum* (EXP1) was shown to bind heme. The conclusions shown in the column “interaction with porphyrin moieties” are essentially derived from the enzymatic studies listed in Table 1.

*Both serine and cysteine residues are present in the active site signature of some members which explain the existence of both GSH-conjugation and thiol transferase/deglutathionylation activities. This table was updated from a previous version [10].

References

1. Bedhomme M, Adamo M, Marchand CH, Couturier J, Rouhier N, Lemaire SD, Zaffagnini M, Trost P. 2012 Glutathionylation of cytosolic glyceraldehyde-3-phosphate dehydrogenase from the model plant *Arabidopsis thaliana* is reversed by both glutaredoxins and thioredoxins in vitro. *Biochem. J.* **445**, 337–347. (doi:10.1042/BJ20120505)

2. Koh CS *et al.* 2008 An atypical catalytic mechanism involving three cysteines of thioredoxin. *J. Biol. Chem.* **283**, 23062–23072. (doi:10.1074/jbc.M802093200)

3. Sterling T, Irwin JJ. 2015 ZINC 15--Ligand Discovery for Everyone. *J Chem Inf Model* **55**, 2324–2337. (doi:10.1021/acs.jcim.5b00559)

4. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. 2009 AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **30**, 2785–2791. (doi:10.1002/jcc.21256)

5. Trott O, Olson AJ. 2010 AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**, 455–461. (doi:10.1002/jcc.21334)

6. Gouy M, Guindon S, Gascuel O. 2010 SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224. (doi:10.1093/molbev/msp259)

7. Gascuel O. 1997 BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. *Mol. Biol. Evol.* **14**, 685–695. (doi:10.1093/oxfordjournals.molbev.a025808)

8. Dixon DP, Lapthorn A, Madesis P, Mudd EA, Day A, Edwards R. 2008 Binding and Glutathione Conjugation of Porphyrinogens by Plant Glutathione Transferases. *Journal of Biological Chemistry* **283**, 20268–20276. (doi:10.1074/jbc.M802026200)

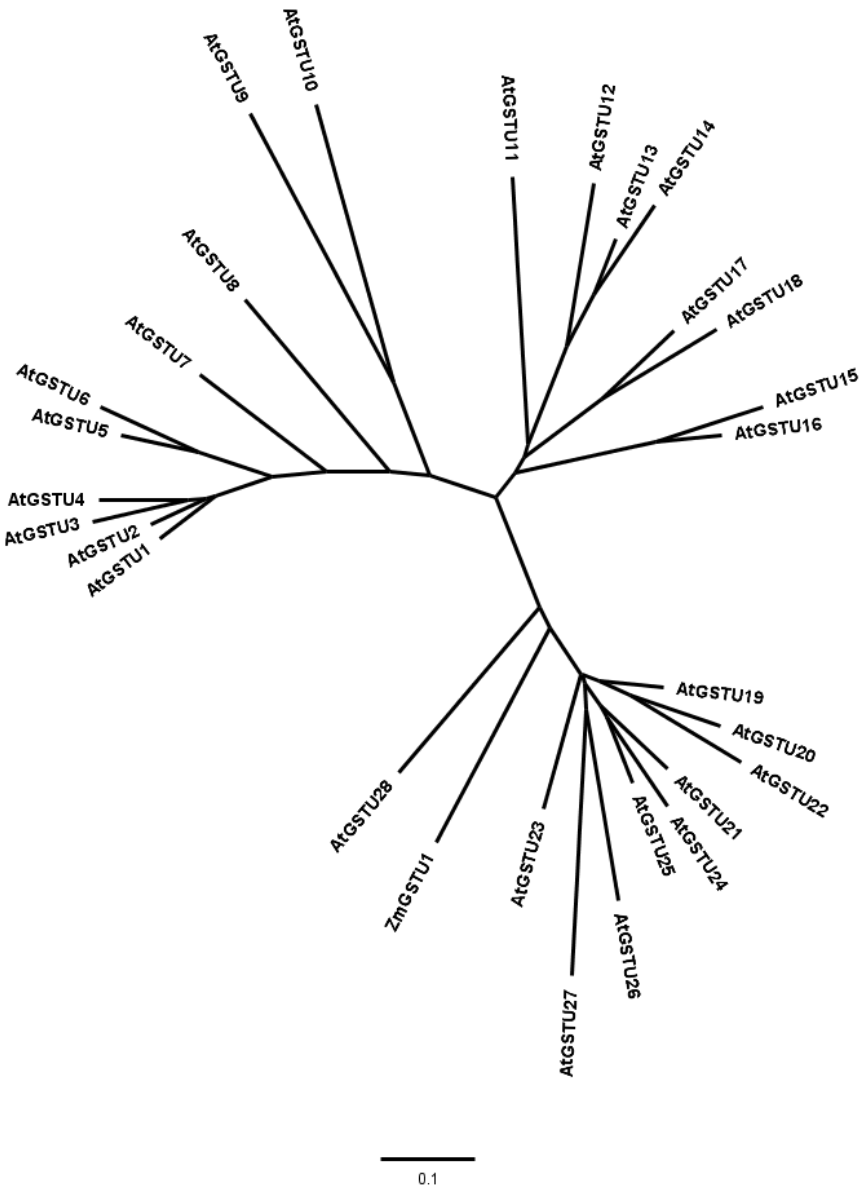

AtGSTU9
AtGSTU10
AtGSTU11
AtGSTU12
AtGSTU13
AtGSTU14
AtGSTU8
AtGSTU17
AtGSTU18
AtGSTU7
AtGSTU6
AtGSTU15
AtGSTU5
AtGSTU16
AtGSTU4
AtGSTU3
AtGSTU2
AtGSTU1
AtGSTU19
AtGSTU20
AtGSTU22
AtGSTU28
AtGSTU21
AtGSTU25
AtGSTU24
ZmGSTU1
AtGSTU23
AtGSTU26
AtGSTU27
0.1

Figure S1

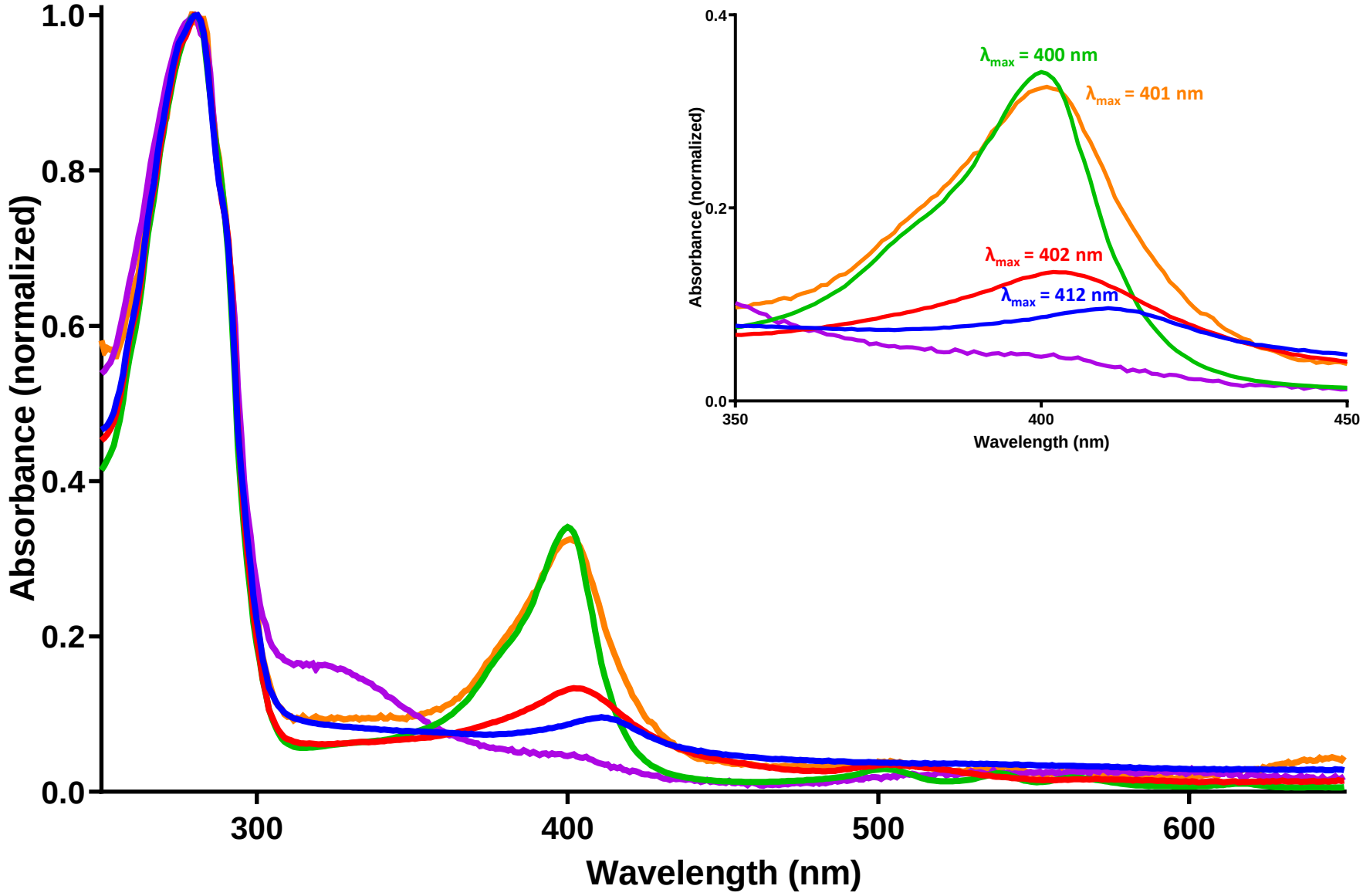
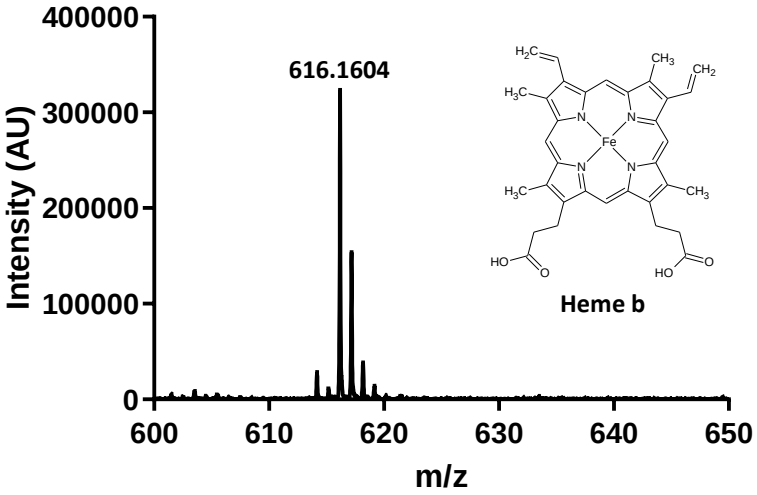
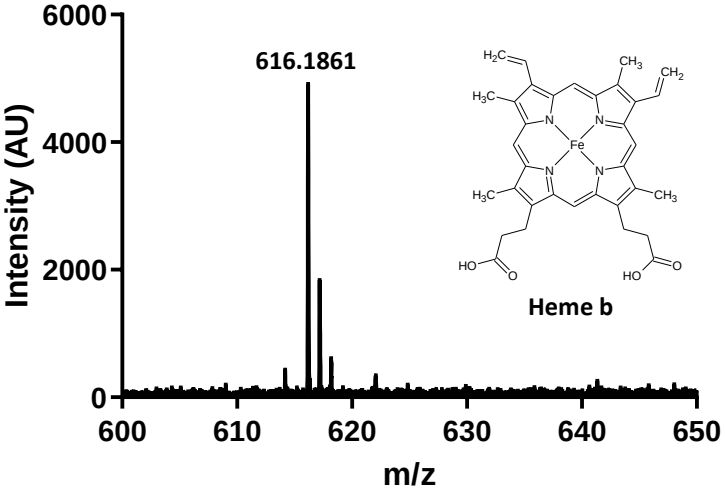

Absorbance (normalized)
1.0
0.8
0.6
0.4
0.2
0.0
300
400
500
600
Wavelength (nm)
Absorbance (normalized)
0.4
0.2
0.0
350
400
450
Wavelength (nm)
λmax = 400 nm
λmax = 401 nm
λmax = 402 nm
λmax = 412 nm

Figure S2

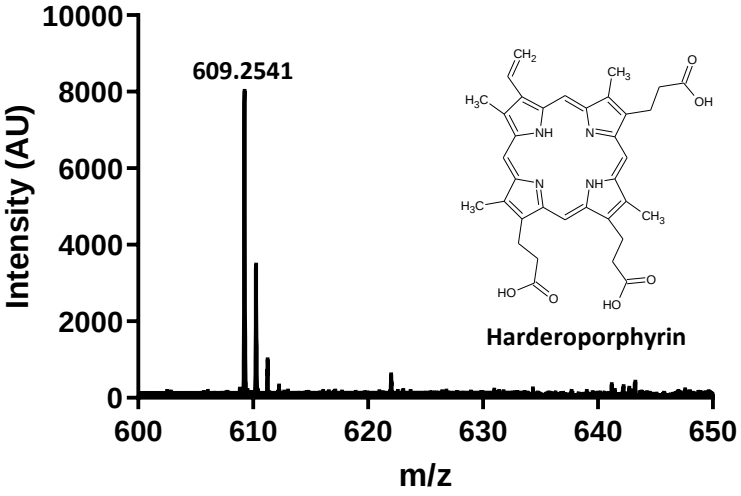
AtGSTU8



AtGSTU16



AtGSTU19



ZmGSTU1

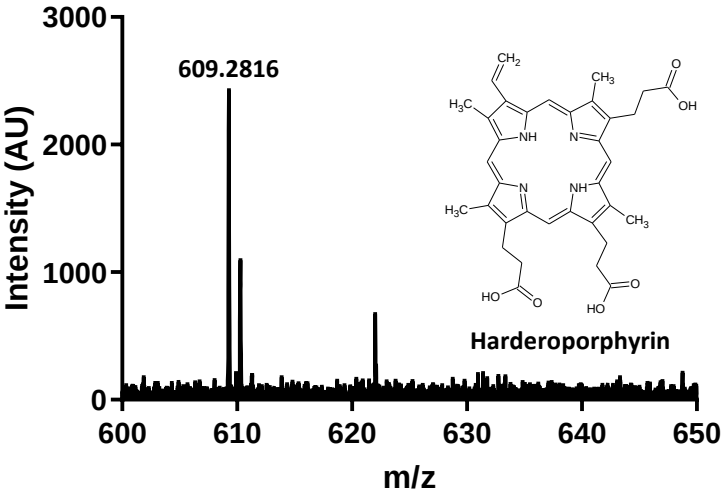


Figure S3

```
                                              β1          α1          β2         α2         β3
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                                                 8          18         28         38         45         55
AtGSTU1     ---------- ---------- -----MAE-- KEESVKLLGF WASPFSRRVE MALKLKGVPY EYLEED---L PNKTPLLLEL NPLHKKVPVL
AtGSTU2     ---------- ---------- -----MAK-- KEESVKLLGF WISPFSRRVE MALKLKGVPY EYLEED---L PKKSTLLLEL NPVHKKVPVL
AtGSTU3     ---------- ---------- -----MAE-- KEEGVKLIGS WASPFSRRVE MALKLKGVPY DYLDEDY--L VVKSPLLLQL NPVYKKVPVL
AtGSTU4     ---------- ---------- -----MAE-- KEEDVKLLGF WASPFTRRVE MAFKLKGVPY EYLEQD---I VNKSPLLLQI NPVYKKVPVL
AtGSTU5     ---------- ---------- -----MAE-- KEE-VKLLGI WASPFSRRVE MALKLKGIPY EYVEEI---L ENKSPLLLAL NPIHKKVPVL
AtGSTU6     ---------- ---------- -----MGK-- NEE-VKLLGI WASPFSRRIE MALKLKGVPY EYLEED---L ENKSSLLLAL SPIHKKIPVL
AtGSTU7     ---------- ---------- -----MAERS NSEEVKLLGM WASPFSRRIE IALTLKGVSY EFLEQD---I TNKSSLLLQL NPVHKMIPVL
AtGSTU8     ---------- ---------- -----MNQ-- -EEHVKLLGL WGSPFSKRVE MVLKLKGIPY EYIEEDV--Y GNRSPMLLKY NPIHKKVPVL
AtGSTU9     ---------- ---------- ----MDEE-- VENKVILHGS FASPYSKRIE LALRLKSIPY QFVQED---L QNKSQTLLRY NPVHKKIPVL
AtGSTU10    ---------- ---------- -----MEE-- KKSKVILHGT WISTYSKRVE IALKLKGVLY EYLEED---L QNKSESLIQL NPVHKKIPVL
AtGSTU11    ---------- ---------- --MGLMNRSK NDEYVKLLGA WPSPFVLRTR IALNLKNVAY EYLEEED-TL SSES--VLNY NPVHKQIPIL
AtGSTU12    MLKNKKSDNS LSRDTLQIKK RKKTTMAQNG SNTTVKLIGT WASPFAIRAQ VALHLKSVEH EYVEETD-VL KGKSDLLIKS NPIHKKVPVL
AtGSTU13    ---------- ---------- -----MAQN- --DTVKLIGS WSSPYSLRAR VALHLKSVKY EYLDEPD-VL KEKSELLLKS NPIHKKVPVL
AtGSTU14    ---------- ---------- -----MAQN- --DTVKLIGC SDDPFSIRPR VALHLKSIKY EYLEEPDDDL GEKSQLLLKS NPIHKKTPVL
AtGSTU15    ---------- ---------- -----MGER- --EQVKLLGT WYSPVVIRAK IDLRLKSVDY DYVEENL--F GSKSELLLKS NPVYKKVPVL
AtGSTU16    ---------- ---------- -----MGEK- --EEVKLLGV WYSPYAIRPK IALRLKSVDY DYVEENL--F GSKSELLLKS NPVHKKVPVL
AtGSTU17    ---------- ---------- -----MAS-- --SDVKLIGA WASPFVMRPR IALNLKSVPY EFLQET---F GSKSELLLKS NPVHKKIPVL
AtGSTU18    ---------- ---------- -----MAT-- --EDVKLIGS WASVYVMRAR IALHLKSISY EFLQET---Y GSKSELLLKS NPVHKKMPVL
AtGSTU19    ---------- ---------- -----MA--- --NEVILLDF WPSMFGMRTR IALREKGVEF EYREEDL--R N-KSPLLLQM NPIHKKIPVL
AtGSTU20    ---------- ---------- -----MA--- --NLPILLDY WPSMFGMRAR VALREKGVEF EYREEDF--S N-KSPLLLQS NPIHKKIPVL
AtGSTU21    ---------- ---------- -----MA--- --AEVILLGF WPSMFGMRTM IALEEKGVKY EYREEDV--I NNKSPLLLEM NPIHKTIPVL
AtGSTU22    ---------- ---------- -----MA--- --DEVILLDF WPSPFGVRAR IALREKGVEF EYREENL--R D-KSPLLLQM NPVHKKIPVL
AtGSTU23    ---------- ---------- -----ME--- --EEIILLDY WASMYGMRTR IALEEKKVKY EYREEDL--S N-KSPLLLQM NPIHKKIPVL
AtGSTU24    ---------- ---------- -----MA--- --DEVILLDF WASMFGMRTR IALAEKRVKY DHREEDL--W N-KSSLLLEM NPVHKKIPVL
AtGSTU25    ---------- ---------- -----MA--- --DEVILLDF WPSMFGMRTR IALEEKNVKF DYREQDL--W N-KSPILLEM NPVHKKIPVL
AtGSTU26    ---------- ---------- -----MAN-- --DQVILLDY WPSMFGMRTK MALAEKGVKY EYKETDP--W V-KTPLLIEM NPIHKKIPVL
AtGSTU27    ---------- ---------- -----MSE-- --EEVVVLNF WPSMFGARVI MALEEKEIKF EYKEEDV--F GQKTDLLLQS NPVNKKIPVL
AtGSTU28    ---------- ---------- -----MGKE- N-SKVVVLDF WASPYAMRTK VALREKGVEF EVQEEDL--W N-KSELLLKS NPVHKKVPVL
ZmGSTU1     ---------- ---------- -----MAEE- KKQGLQLLDF WVSPFGQRCR IALDEKGLAY EYLEQDL--R N-KSELLLRA NPVHKKIPVL
                                                      H                                    G             GGG
                                                      *                                    *              *

            β3 β4      α3                                   α4                                    α5
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                65         75         84         92         102        110        119        129        139
AtGSTU1     VHNDKILLES HLILEYIDQT WKN-SP--IL PQDPYEKAMA RFWAKFIDDQ --ILTLGFRS LVKA-EKGRE VAIEETRELL MFLEKEVTGK
AtGSTU2     VHNDKLLSES HVILEYIDQT WNN-NP--IL PHDPYEKAMV RFWAKFVDEQ --ILPVGFMP LVKA-EKGID VAIEEIREML MFLEKEVTGK
AtGSTU3     VHNGKILPES QLILEYIDQT WTN-NP--IL PQSPYDKAMA RFWAKFVDEQ --VTMIGLRS LVKS-EKRID VAIEEVQELI MLLENQITGK
AtGSTU4     VYKGKILSES HVILEYIDQI WKN-NP--IL PQDPYEKAMA LFWAKFVDEQ --VGPVAFMS VAKA-EKGVE VAIKEAQELF MFLEKEVTGK
AtGSTU5     VHNGKTILES HVILEYIDET WPQ-NP--IL PQDPYERSKA RFFAKLVDEQ --IMNVGFIS MARADEKGRE VLAEQVRELI MYLEKELVGK
AtGSTU6     VHNGKTIIES HVILEYIDET WKH-NP--IL PQDPFQRSKA RVLAKLVDEK --IVNVGFAS LAKT-EKGRE VLIEQTRELI MCLEKELAGK
AtGSTU7     VHNGKPISES LVILEYIDET WRD-NP--IL PQDPYERTMA RFWSKFVDEQ --IYVTAMKV VGKT-GKERD AVVEATRDLL MFLEKELVGK
AtGSTU8     IHNGRSIAES LVIVEYIEDT WKTTHT--IL PQDPYERAMA RFWAKYVDEK --VMLAVKKA CWGP-ESERE KEVKEAYEGL KCLEKELGDK
AtGSTU9     VHNGKPISES LFIIEYIDET WSNGPH--IL PEDPYRRSKV RFWANYIQLH --LYDLVIKV VKSE-GEEQK KALTEVKEKL SVIEKEGLKE
AtGSTU10    VHDGKPVAES LVILEYIDET WTNSPR--FF PEDPYERAQV RFWVSYINQQ --VFEVMGQV MSQE-GEAQA KSVEEARKRF KVLD-EGLKK
AtGSTU11    IHGNKPIRES LNIVMYVDET WLS-GP-PIL PSDPFDRAVA RFWDVYIDEH --CFTSINGV AVAKGEENIN AAIAKLEQCM ALLEETFQEC
AtGSTU12    IHGDVSICES LNIVQYVDES WPS-DL-SIL PTLPSERAFA RFWAHFVDGK --LFESIDAV AGAKDDAARM TLAGNLMENL AALEEAFQKS
AtGSTU13    LHGDLSISES LNVVQYVDEA WPS-VP-SIL PSDAYDRASA RFWAQYIDDK --CFAAVDAV VGAKDDEGKM AAVGKLMECL AILEETFQKS
AtGSTU14    IHGDLAICES LNIVQYLDEA WPS-DP-SIL PSNAYDRASA RFWAQYIDDK --CFEAANAL TGANNDEERI AATGKLTECL AILEETFQKS
AtGSTU15    IHNTKPVCVS LNIVEYIDET WNSSGS-SIL PSHPYDRALA RFWSVFVDDK --WLPTLMAA VVAKSEEAKA KGMEEVEEGL LQLDAAFIAL
AtGSTU16    LHNNKPIVES LNIVEYIDET WNSSAP-SIL PSHPYDRALA RFWSDFVDNK --WFPALRMA AITKSEDAKA KAMEEVEEGL LQLEDAFVSI
AtGSTU17    LHADKPVSES NIIVEYIDDT WSSSGP-SIL PSDPYDRAMA RFWAAYIDEK --WFVALRGF LKAGGEEEKK AVIAQLEEGN AFLEKAFIDC
AtGSTU18    IHADKPVCES NIIVHYIDEA WNSSGP-SIL PSHPYDRAIA RFWAAYIDDQ --WFISVRSI LTAQGDEEKK AAIAQVEERT KLLEKAFNDC
AtGSTU19    IHNGKPVNES IIQVQYIDEV WSHKN--PIL PSDPYLRAQA RFWADFIDKK --LYDAQRKV WA-TKGEEQE AGKKDFIEIL KTLESELGDK
AtGSTU20    VHNGKPVCES LNVVQYVDEA WPEKN--PFF PSDPYGRAQA RFWADFVDKK --FTDAQFKV WG-KKGEEQE AGKKEFIEAV KILESELGDK
AtGSTU21    IHNGKPVLES LIQIQYIDEV WSDNN--SFL PSDPYHRAQA LFWADFIDKK EQLYVCGRKT WA-TKGEELE AANKEFIEIL KTLQCELGEK
AtGSTU22    IHNGKPVCES MNVVQYIDEV WSDKN--PIL PSDPYQRAQA RFWVDFVDTK --LFEPADKI WQ-TKGEEQE TAKKEYIEAL KILETELGDK
AtGSTU23    IHEGKPICES IIQVQYIDEL WPDTN--PIL PSDPYQRAQA RFWADYIDKK --TYVPCKAL WS-ESGEKQE AAKIEFIEVL KTLDSELGDK
AtGSTU24    IHNGKPVCES LIQIEYIDET WPDNN--PLL PSDPYKRAHA KFWADFIDKK --VNVTARRI WA-VKGEEQE AAK-ELIEIL KTLESELGDK
AtGSTU25    IHNGNPVCES LIQIEYIDEV WPSKT--PLL PSDPYQRAQA KFWGDFIDKK --VYASARLI WG-AKGEEHE AGKKEFIEIL KTLESELGDK
AtGSTU26    IHNGKPICES LIQLEYIDEV WSDAS--PIL PSDPYQKSRA RFWAEFIDKK --FYDPSWKV WA-TMGEEHA AVKKELLEHF KTLETELGDK
AtGSTU27    IHNGKPVCES NIIVEYIDEV WKDDKTLRLL PSDPYQKSQC RFWADLIDKK --VFDAGRRT WT-KRGKEQE EAKQEFIEIL KVLERELGDK
AtGSTU28    IHNNTPISES LIQVQYIDET WTDAA--SFL PSDPQSRATA RFWADYADKT -ISFEGGRKI WGNKKGEEQE KGKKEFLESL KVLEAELGDK
ZmGSTU1     LHDGRPVCES LVIVQYLDEA FPEAAP-ALL PADPYARAQA RFWADYVDKK --LYDCGTRL WK-LKGDGQA QARAEMVEIL RTLEGALGDG
                    GG                                                H   H   H
                                                                      *   *   *

                                      α6                                 α7       α8          α9
            ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                           150        160        170        180        190        200        210        220
AtGSTU1     DF-------- FGG-KTIGFL DMIAGSMIPF CLARLWKGIG IDMIPEEKFP ELNRWIKNLE EVEAVRGCIP PREKQIERMT KIAET-----
AtGSTU2     DF-------- FGG-KTIGFL DMVAGSMIPF CLARAWECLG IDMTPEDTFP ELNRWIKNLN EVEIVRECIP PKEKHIERMK KIIER-----
AtGSTU3     KL-------- FGG-ETIGFL DMVVGSMIPF CLARAWEGMG IDMIPEEKFP ELNRWIKNLK EIEIVRECIP DREKHIEHMM KIVGR-----
AtGSTU4     DF-------- FGG-KTIGFL DLVAGSMIPF CLARGWEDMG IDMIPEEKFP ELNRWIKNLK EIEIVRECIP PREEQIEHMK KVVER-----
AtGSTU5     DY-------- FGG-KTVGFL DFVAGSLIPF CLERGWEGIG LEVITEEKFP EFKRWVRNLE KVEIVKDCVP PREEHVEHMN YMAER-----
AtGSTU6     DY-------- FGG-KTVGFL DFVAGSMIPF CLERAWEGMG VEMITEKKFP EYNKWVKKLK EVEIVVDCIP LREKHIEHMN NMAEK-----
AtGSTU7     DF-------- LGG-KSLGFV DIVA-TLVAF WLMRTEEIVG VKVVPVEKFP EIHRWVKNLL GNDVIKKCIP PEDEHLKYIR ARMEKL----
AtGSTU8     LF-------- FGG-ETIGFV DIAA-DFIGY WLGIFQEASG VTIMTAEEFP KLQRWSEDFV GNNFIKEVLP PKEKLVAVLK AMFGS-----
AtGSTU9     IFSDTDGEPT VTN-ETMSLV DIVMCTLLSP YKA-HEEVLG LKIIDPEIVP GVYGWINAIN ETSVVKDLSP PYEQILEILR AFRQMS----
AtGSTU10    HFPNKN---I RRN-DDVGLL EITIIATLGG YKA-HREAIG VDIIGPVNTP TLYNWIERLQ DLSVIKEVEV PHDTLVTFIQ KYRQKC----
AtGSTU11    SKGRG----F FGG-ENIGFI DIGFGSMLGP LTV-LEKFTG VKFIHPENTP GLFHWADRFY AHEAVKPVMP DIEKLVQFAR L---------
AtGSTU12    SKGGD----F FGG-GNIGFV DITVGAIVGP ISV-IEAFSG VKFLRPDTTP GLIQWAEKFR AHEAVKPYMP TVAESIEFAK K---------
AtGSTU13    SKGLG----F FGG-ETIGYL DIACSALLGP ISV-IEAFSG VKFLRQETTP GLIKWAERFR AHEAVKPYMP TVEEVVAFAK Q---------
AtGSTU14    SKGLG----F FGG-ETIGYL DIACAALLGP ISV-IEMFSA DKFVREETTP GLIQWAVRFR AHEAVRPYMP TVEEVTELVK QRIEEGFKRN
AtGSTU15    SKGKS----F FGG-ETIGFI DICLGSFLVL LKA-REKLKN EKILDELKTP SLYRWANQFL SNEMVKNVVP DIDKVAKFIE EFED------
AtGSTU16    SKGKP----F FGG-EAIGFM DICFGSFVVL LKA-REKFKA EKLLVESKTP SLCKWADRFL SDETVKNVAP EIEKVAEFLQ ELEV------
AtGSTU17    SKGKP----F FNG-DNIGYL DIALGCFLAW LRV-TELAVS YKILDEAKTP SLSKWAENFC NDPAVKPVMP ETAKLAEFAK KIFP------
AtGSTU18    SQGKP----F FNG-DHIGYL DIALGSFLGW WRV-VELDAN HKFLDETKTP SLVKWAERFC DDPAVKPIMP EITKLAEFAR KLFP------
AtGSTU19    PY-------- FSG-DDFGYV DIALIGFYTW FPA--YEKFA NFSIESE-VP KLIAWVKKCL QRESVAKSLP DPEKVTEFVS ELRKK-----
AtGSTU20    PY-------- FGG-DSFGYV DISLITFSSW FQA--YEKFG NFSIESE-SP KLIAWAKRCM EKESVSKSLP DSEKIVAYAA EYRKN-----
AtGSTU21    PY-------- FGG-DKFGFV DIVLIGFYSW FPA--YQKFG NFSIEPE-CL KLIAWGKRCM QRESVAKALP DSEKVVGYVL QLKKL-----
AtGSTU22    PY-------- FGG-DTFGFV DIAMTGYYSW FEA--SXKLA NFSIEPX-CP TLMASAKRCL QRESVVQSLH DSEKILAFAY KIRKI-----
AtGSTU23    YY-------- FGG-NEFGLV DIAFIGFYSW FRT--YEEVA NLSIVLE-FP KLMAWAQRCL KRESVAKALP DSDKVLKSVS DHRKII----
AtGSTU24    KY-------- FGD-ETFGYV DIALIGFHSW FAV--YEKFG NVSIESE-CS KLVAWAKRCL ERESVAKALP ESEKVITFIS ERRKK-----
AtGSTU25    TY-------- FGG-ETFGYV DIALIGFYSW FEA--YEKFG SFSIEAE-CP KLIAWGKRCV ERESVAKSLP DSEKIIKFVP ELRKK-----
AtGSTU26    PY-------- YGG-EVFGYL DIALMGYYSW FKA--MEKFG EFSIETE-FP ILTTWTKRCL ERESVVKALA DSDRIIEYVY VLRKK-----
AtGSTU27    VY-------- FGGNDNVSMV DLVLISYYPW FHT--WETIG GFSVEDH-TP KLMDWIRKCL TRPAISKSLP DPLKIFDRVT QIIKVH----
AtGSTU28    SY-------- FGG-ETFGYV DITLVPFYSW FYA--LEKCG DFSVEAE-CP KIVAWGKRCV ERNSVAATLP ESEKVYQQVL KLRQI-----
ZmGSTU1     PF-------- FGG-DALGFV DVALVPFTSW FLA--YDRFG GVSVEKE-CP RLAAWAKRCA ERPSVAKNLY PPEKVYDFVC GMKKR-----
                                           H                                                    H    H
                                           *                                                    *    *
```