MATERIALS AND METHODS

Phylogenetic and sequence analyses. Homologs of AtCHI were gathered by psi-blast from the NCBI NR database, followed by iterative profile HMM building and searches against public protein, expressed sequence tag (EST) and genome sequence databases. Gene models were built with Genewise[1] and refined manually. Protein sequences in each class were aligned with MUSCLE[2] and manually edited, and classes were aligned by ClustalW[3] and further edited on the basis of the structural superposition of the available structures. An alignment of all available protist and plant proteins and selected fungal and bacterial proteins is attached (Supplementary File 1). A consensus phylogenetic tree was built from 30,000 trees sampled from a 4 million-generation run of MrBayes[4] (Supplementary Fig. 2, Supplementary File 2). A maximum likelihood tree build using PhyML[5] produced a highly similar branching pattern supporting the overall family classification, though several terminal branches were not conserved (not shown).

Expression and purification of *A. thaliana* CHI-fold proteins. The coding sequence for each *A. thaliana* protein's CHI-fold domain was inserted between the NcoI and BamHI sites of the expression vector pHIS8, which under the control of a T7 promoter, yields the target protein fused to a thrombin-cleavable amino-terminal His_8-tag[6]. Proteins were heterologously expressed in *E. coli* BL21 (DE3) (Novagen, USA) expression host. *E. coli* cultures were grown at 37 ºC in Terrific Broth (TB) to an optical density (600 nm) of 1.5, induced with 1 mM isopropyl-β-D-thiogalactoside (IPTG), and then grown for an additional 6 h at 22 ºC. Bacterial cells were harvested by centrifugation, resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0; 0.5 M NaCl; 20 mM imidazole; 1% Tween20; 10% glycerol; and 20 mM 2-mercaptoethanol), and lysed by sonication. Proteins were isolated from *E. coli* lysates by affinity chromatography with Ni-affinity (Ni^{2+}-NTA) agarose, and eluted with lysis buffer containing 0.25 M imidazole. Partially purified proteins were treated with thrombin for cleavage of the His_8-tag, and then further purified by gel-exclusion chromatography with a Superdex 200 HR26/60 column (Pharmacia Biosystems, Germany). Selenomethionine-substituted proteins for AtCHIL and AtFAP3 were generated from *E. coli* cultures grown in M9 minimal medium supplemented with an amino acid cocktail including L-selenomethionine[7] (Sigma-Aldrich), and were otherwise handled as described above for unlabeled protein.

Analysis of FA binding by *A. thaliana* CHI-fold proteins. Bound ligands were extracted from protein samples of AtFAP1, AtFAP2, and AtFAP3 (1-2 mg in 20 mM ammonium bicarbonate) by addition of ethanol to a final concentration of 80%. After incubation at –20°C for 3 d, the treated samples were centrifuged (16000 x *g* at 4°C) to remove denatured protein and other particulates, and the supernatant was evaporated under vacuum. Residual material was re-dissolved in 200 µl 3-propanol, which was then passed through a 0.45 µm nylon filter in preparation for further characterization by HPLC-MS. The analysis of binding of exogenously provided alpha-linolenic acid by AtFAP1 utilized the aforementioned amino-terminal His_8-tagged form of the protein coupled to Ni-affinity resin. Immobilized protein (2 mg) was incubated in buffer (12.5 mM Tris-HCl, pH 8.0, 50 mM NaCl) with 0.5 mM alpha-linolenic acid, and subsequently washed thoroughly with buffer including 5% ethanol. The AtFAP1 protein was released with 0.25 M acetic acid, and then processed for extraction of bound ligand as described above. Control samples were treated identically, except that the AtFAP1 protein was omitted.

Extracts were analyzed for FA content by HPLC-MS on an Agilent 1100 Series LC-MSD

instrument with electrospray-ionization eluent introduction into an XCT ion trap mass spectrometer (Agilent). Chromatographic separations employed an Agilent Zorbax Eclipse XDB-C18 (4.6 x 150 mm, 5 mm particle size) reversed-phase column run at a flow rate of 0.5 ml min^{-1}, and a linear gradient with initial and final mobile phases consisting of 95% water:5% acetonitrile:0.1% formic acid and 5% water:95% acetonitrile:0.1% formic acid, respectively. The identities of FA components were established by mass determination and comparison of chromatographic retention times with authentic FA standards (Sigma-Aldrich).

Thermal-shift binding assay. The shift in melting temperature for AtFAP proteins observed under increasing concentrations of fatty acids (FAs) was measured using a Thermofluor-type assay similar to one previously published[8]. Protein melting temperatures were assayed using the LightCycler480 System II (Roche), using the following program: 30 seconds at 20°C, ramp up to 85°C at 0.06°C/second, 30 seconds at 20°C, excitation wavelength 483 nm, emission wavelength 568 nm. Using SYPRO Orange (Sigma), an environmentally-sensitive dye that interacts with hydrophobic amino acid residues, the melting of a protein can be observed as a fluorescence increase when the hydrophobic core residues of a protein fold are exposed to the environment in a denatured protein. SYPRO Orange dye can interact with the unfolded protein and cause an increase in fluorescence. The maximum or minimum of a first derivative curve of the fluorescence profile will indicate the melting temperature (Tm). The melting temperature of a protein in the absence of any additional compound provides a baseline melting temperature for the protein (Tm_o). Compounds can be screened for binding interactions with the protein of interest. When a ligand binds a protein, a change in Gibbs free energy occurs and may cause a change in observed Tm for that protein. ΔTm can be measured for a collection of putative ligands and the strength of binding interactions can be ranked and compared for compounds with similar physicochemical properties, providing a relative measure of binding affinities.

A quantitative analysis of these relative binding affinities for AtFAPs versus the collection of FAs shown in Supplementary Table 1 (12 different FAs) can be obtained by a simple dose-response analysis of ΔTm with respect to FA concentration. 5 uM of each protein and 10X SYPRO Orange (from 5000x stock solution) were mixed with from 10^{-1} to 10^{5} nM of the twelve different FAs, incubated at room temperature for 20 minutes, then assayed for protein melting temperature. Mean ΔTm (Tm_o - $Tm_{measured}$) was plotted against FA concentration for each FA/FAP combination using data from four replicate experiments. A 'dose-response: stimulation (three parameters)' standard (Hill) slope non-linear regression was fit to the data, using this model: $Y=Tm_{max} + (Tm_{max}-Tm_{min})/(1+10\text{^}(LogEC50-X))$. Calculated EC50s ± SE were reported as relative measures of binding constants, while maximum ΔTm was reported as a qualitative measure of free energy change in the protein + ligand system.

Crystallization and structure elucidation of *A. thaliana* CHI-fold proteins. Crystals of AtCHI-fold proteins were grown by vapor diffusion at 4ºC, from 1:1 mixtures of protein solution (10-15 mg ml^{-1} in 12.5 mM Tris-HCl, pH 7.5, 50 mM NaCl) and reservoir solution. The reservoir solution contained 28 % polyethylene glycol (PEG) 8000, 0.3 M magnesium nitrate, 2 mM dithiothreitol (DTT), and 100 mM HEPES-Na^+ (pH 7.0) for AtCHI; 28 % PEG 8000, 0.2 M calcium acetate, 2 mM DTT, and 100 mM TAPS-Na^+ (pH 8.5) for AtCHIL; 19 % PEG 3350, 0.3 M potassium chloride, 2 mM DTT, and 100 mM TAPS-Na^+ (pH 8.5) for AtFAP1; and 7-9 % PEG 8000, 0.2 M calcium acetate, 2 mM DTT, and 100 mM PIPES-Na^+ (pH 6.5) for AtFAP3. Crystal growth typically occurred over a period of 2-10 days, and was sometimes expedited

through seeding with finely crushed microcrystals. For heavy-atom complexes, crystals of AtFAP1 were soaked overnight in mother liquor supplemented with 1 mM K_2PtCl_4.

Crystal samples were flash frozen by immersion in liquid nitrogen after a brief incubation in a cryoprotectant solution (consisting of reservoir solution supplemented with 17-20% ethylene glycol). X-ray diffraction data were collected from frozen crystals at the FIP beamline of the European Synchrotron Radiation Facility (ESRF), beamlines 8.2.1 and 8.2.2 of the Advanced Light Source (ALS), Lawrence Berkeley National Laboratory, or beamlines 1-5 and 9-1 of the Stanford Synchrotron Radiation Laboratory (SSRL). Diffraction intensities were measured on ADSC Quantum or MarResearch charge-coupled device (CCD) detectors, and were indexed, integrated, and scaled with MOSFLM[9] and SCALA[10], HKL2000[11], or XDS and XSCALE[12] programs.

Crystallographic structure solutions for the CHI-fold proteins were obtained through either single/multiple-wavelength anomalous dispersion (SAD/MAD) or molecular-replacement (MR) analyses, as detailed in Supplementary Table 8. For SAD/MAD analyses, the location of anomalous scatterers and initial phase estimates were determined with the program SOLVE[13], and preliminary structural models were automatically built with the program RESOLVE[14]. For MR analyses, Molrep[15] was used, and where necessary, search coordinate-sets were constructed through homology modeling with Modeller[16]. ARP/wARP[17] was used for automated rebuilding of initial structure models for AtCHIL and AtFAP3. Subsequent structural refinements utilized CNS[18] or REFMAC[19]. Xfit[20] and Coot[21] were used for graphical map inspection and manual rebuilding of atomic models. Programs from the CCP4 suite[22] were employed for all other crystallographic calculations. Structural superpositions were calculated with SSM[23].

The identity of two ordered small-molecules bound in the active site of AtCHI was inferred to be nitrate because of the high concentration (0.3 M) of magnesium nitrate in the crystallization medium. From the X-ray crystallographic analysis, the shape and level of the electron density associated with these small molecules are also consistent with nitrate.

Plant growth. This study uses wild type *A. thaliana* Columbia (Col-0), and mutant and transgenic lines derived from Col-0. Seeds (approximately 5/pot) were sown on soil, in flats containing either 21 or 32 pots; genotypes were distributed in a completely randomized design. After incubation for 3 d at 4ºC to break dormancy, flats were moved to a growth-chamber under long day conditions (16:8-h light/dark cycle) at 22ºC and 75% relative humidity, and after a week pots were thinned such that each pot contained 2 plants. For additional plant morphological analyses and for seed collection, plants were grown in the greenhouse at 20-23ºC under continuous illumination at (170 µmol m^{-2} s^{-1}) in soil treated with granular Marathon (Olympic Horticultural Products, Bradenton, FL) to protect against insect damage.

In the specified experiments, after 3 weeks of growth at 22°C, plants were further grown at 15°C, 22°C, and 26°C until senescence (control plants were grown at 22°C for this period). Genotypes were distributed in a randomized design. FA determinations were made from expanded rosette leaves harvested after 10 days of growth at 15°C, 22°C, and 26°C. Growth and morphological phenotype of plant lines were observed every 2 days throughout development, and any differences between wild type and mutants were recorded.

Molecular constructs, transgenic *A. thaliana* lines and microscopy. To evaluate sub-cellular location of the proteins, we generated promoter + target-GUS/GFP fusion constructs using promoters of two different lengths. The primers employed for amplification were: *FAP1* prom1 +

T FW (5'-GGGGA CAAGTTT GTACAAAAA
AGCAGGCTACTGGTAACTTCATTTAAACTT CCT-3'); *FAP1* prom2 + T FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTGTCTTC AGACTGAATTATCTTCAAC-3');
FAP1 prom + T Rev 5'-GGGGACCACTTTGTA
CAAGAAAGCTGGGTTACGATCACACGGGAAATTTCG-3'); *FAP2* prom + T FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAGAAAGTGAAACACAACACAAAA-
3'); *FAP2* prom + T Rev (5'-
GGGGACCACTTTGTACAAGAAAGCTGGGTTTATAAGCTTTTGAATATAACTCCA-3');
FAP3 prom1 + T FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTTCATTGAAACCAATAACTAGCAA-
3'); *FAP3* prom2 FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAAACGACAAGCAGTTGGTTATT-3");
FAP3 prom + T Rev (5'
GGGGACCACTTTGTACAAGAAAGCTGGGTTGTTCAAAACTCCGGCGATAA
-3'). These constructs harbored GFP fused to the N-terminal region of *FAP1/2/3* gene, under the control of the specified promoter. The constructs were cloned into the pDONR221 entry vector, integrated into the binary vector pBWSF7[24], and transformed into *A. tumefaciens* strain GV3101 via electroporation, which we used to transform *A. thaliana* (Col-0) by "floral-dip" method[25]. After identification of putative transgenic T0 lines by bar resistance[26], plants from 10 independent lines for each construct were analyzed by PCR using the attB-adapted primers (5'-GGGGACAAGTTTGTACAAAA AAGCAGGCTTC-forward primer, 5'-GGGGACCACTTTGTACAAGAAAGCTGGG TC-reverse primer). Five independent transgenic lines in the T2 generation (5 individual plants from each transgenic line at each stage of development) were evaluated for GFP location by confocal microscopy[27].

To evaluate the spatial distribution of FAP expression *in planta*, we generated promoter-GUS/GFP fusion constructs using promoters of two different lengths. The primers used for amplification were: *FAP1* prom1 FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTACTGGTAACTTCATTTAAACT TCCT-3');
FAP1 prom2 FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTGTCTTCAGACTGAATTATCTTCAAC-3');
FAP1 prom REV (5'-GGGGACCACTTTG
TACAAGAAAGCTGGGTCGTCGTCTTTTTGGTGTGAG-3'); *FAP2* prom FW (5'-
GGGGACCACTTTGTACAAGAAAGCTGGGTTTATAAGCTTTTGAATATAACTCCA-3');
FAP3 prom Rev (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTTAGAAAGTGAAACACAACACAAAA-3');
FAP3 prom1 FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCTTCATTGAAACCAATAACTAGCAA-
3'); *FAP3* prom2 FW (5'-
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCAAAACGACAAGCAGTTGGTTATT-3');
FAP3 prom Rev (5'-
GGGGACCACTTTGTACAAGAAAGCTGGGTGTTCAAAACTCCGGCGATAA-3').
Cloning, transformation, and selection methods were as previously described in this section. 10 independent lines were obtained for each construct and analyzed by PCR. Representative transgenic lines in the second generation (T2) were subjected to histological tissue analysis.

In an additional set of constructs, the full open reading frame (ORF) of FAP1, 2, and 3 was introduced into pEarleyGate 103 [38], to generate an in frame fusion with GFP under the control of

the cauliflower mosaic virus 35S promoter. The cloning procedure was similar to the one described above. The primers used for amplification of the ORFs were *FAP1*p103 FW (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGTTTCGTTTCGCTTCC-3'); *FAP1*p103 Rev (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTGTTGTCGAAGGCTAGAGAAGC-3'); *FAP3*p103 FW (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATGGGATTCTTGCAG-3'); *FAP3*p103 Rev (5'-GGGGACCACTTTGTACAAGAAAGCTGGGTTCACTAGGGTCATGGCTAATTG-3'). After insertion of FAPs ORF in pDONR221 which contains kanamycin resistance (bacterial selection marker also found in the binary vector pEarleygate 103), we cleaved the transformed pDONR221 using restriction enzymes that do not cut the inserted sequence and purified the fragment of interest.

Near fully expanded leaves from 2-week old T2 mutant plants (5 independent lines) were harvested for confocal microscopy observation using a Leica TCS NT laser scanning microscope system (Confocal Microscopy Facility, Iowa State University, Ames IA, USA). For each line, 5 plants were analyzed. GFP signal and auto fluorescence were detected at wavelengths of 488 nm FITC and 568 nm TRITC, respectively (Supplementary Fig. 4). We imported the images into Adobe Photoshop 6.0 (Adobe, San Jose, CA, USA).

For histochemical studies, transgenic plants and control plants (siblings not containing the promoter-GUS constructs) were harvested at different stages of development. Plant material was incubated at room temperature in the dark with a GUS[28] substrate solution containing 10 mg ml^{-1} X-Gluc in dimethyl sulphoxide, 0.1 M potassium phosphate buffer (pH 7.0), Triton/ethanol (Triton X-100:ethanol:water; 1:4:5), 0.1 M $K_3[Fe(CN)_6]$ (pH 7.0), 0.1 M $K_4[Fe(CN)_6]$ (pH 7.0), washed with 70% ethanol, observed and photographed with the Olympus stereomicroscope[27] (Bessey Microscopy Facility at Iowa State University, Ames IA, USA).

T-DNA insertion lines and mutant isolation. Seeds of *AtFAP1-1 (Salk_130560)* and *AtFAP1-3 (Salk_039829)* were obtained from the Arabidopsis Biological Resource Center (ABRC); *AtFAP1-2 (Flag_389_G05)* seeds were obtained from the Institut National de la Recherche Agronomique (INRA). PCR reactions with a combination of specific primers LP + RP and LB + RP were used to verify the site of T-DNA insertion for each line (Supplementary Table 9). The T-DNA insertion in *AtFAP1-1* is located in the first intron; in *AtFAP1-2* and *AtFAP1-3*, it is in the second exon (Supplementary Fig. 14). Knock-out mutants were identified as homozygous via PCR and by back-crossing to WT and mutant parents (once each; both crosses resulted in 1:1 segregation pattern).

We previously determined that the putative AtFAP3 mutant lines available from SALK and SAIL did not have the T-DNA insertion in AtFAP3. Specifically, we came to this conclusion after ordering and testing multiple batches of seeds of the SALK_107742 and SAIL_598_D12 putative AtFAP3 mutant lines. The plants from these lines were screened by PCR using the primer combinations as suggested in TAIR ((SALK_107742-LP + SALK_107742-RP, lbba1/lbba1-3 + SALK107742-RP); (SAIL_598-LP + SAIL_598-RP, LB2/LB3 + SAIL_598-RP)), in parallel with controls (SALK/SAIL lines representing the other FAP mutants). We were unable to identify either homozygote mutant or heterozygote mutant plants in the SALK_107742 or SAIL_598 lines. Unlike the multiple independent AtFAP3-RNAi lines we generated, the SALK_107742 or SAIL-598 lines were visibly indistinguishable from those of the corresponding

WT.

DNA and RNA isolation from plant tissue and RT-PCR. Genomic DNA was isolated from leaves using a cetyl trimethylammonium bromide (CTAB) extraction protocol[29]. Tissue was frozen and ground to a powder in liquid nitrogen. Buffer composed of 1% CTAB, 50 mM Tris-HCl, pH 8.0, 0.7 mM NaCl, 10 mM EDTA, 0.5% polyvinylpyrrolidone (PVP), was added, homogenized and incubated at 65ºC for 1 h. The solution was extracted with chloroform, followed by isopropanol precipitation of DNA. After incubation for 10-30 min at -20ºC, the pellet was treated with RNAse before precipitation with sodium acetate, centrifuged, and washed with 70% ethanol. After centrifugation, the pellet was dried, resuspended in water, and DNA was amplified by PCR as previously described.

Total RNA was isolated from leaves from two week-old wild type and mutant plants with TRIzol reagent (Invitrogen) following the manufacturer's instructions. Isolated RNA was treated with amplification-grade DNaseI (Invitrogen) to remove genomic DNA in the samples. The purified RNA was used for RT-PCR amplification with the SuperScript III One-Step Kit (Invitrogen) according to the manufacturer's instructions. The forward and reverse primers used were: 5'-AT GGTTTCGTTTCGCTTCC -3' and 5'-CTAGTTGTCGAAGG CTAGAGAAGCT -3' respectively; tubulin primers (FW 5'-CGTGGATCACAGCAATACAGAGCC-3' and REV 5'-CCTCCTGCACTTCCACTTCGTCTTC-3') were used as a control for RNA content. The thermal profile used was: 1 cycle: 45-60 ºC for 15-30 min (cDNA synthesis); 1 cycle: 94 ºC for 2 min (denaturation); 28 cycles: 94 ºC for 15 s (denature), 56 ºC for 30 s (anneal), 68 ºC for 1 min kb^{-1} (extend). The final extension comprised 1 cycle: 68 ºC for 5 min.

Western blot analysis. Total protein was extracted from 3-week-old plant leaves and 7-8 DAF siliques; protein concentration was determined using the method of Bradford[39]. Extract containing 100 ug of total protein were loaded in each well for SDS-PAGE; for a positive control, 0.1 ug of purified protein was used. Proteins were transferred to nitrocellulose membrane using Bio-Rad apparatus. FAP3-specific antiserum was generated in rabbits at the Protein Facility of the Iowa State University by using FAP3 protein expressed and purified from *E.coli.*

FA analysis from plant tissue. FAs were extracted using a barium hydroxide hydrolysis protocol[30-32]. Recently expanded rosette leaves were collected and placed in liquid nitrogen. Each replicate comprises leaves from a single plant. Approximately 0.1 g fresh weight of leaves or 5 mg of seeds was used for each extraction. Tissues were frozen in liquid nitrogen upon harvest, and placed in a pre-cooled tissue homogenizer. 20 µl of internal standard (nonadecanoic acid, C19:0, 2 mg ml^{-1} dissolved in chloroform) and 1 ml of barium hydroxide were added, and the mixture was further homogenized according to the procedure in plant metabolomics.org. The homogenate was transferred to a glass tube and 550 µL of 1,4-dioxane (Aldrich cat # 123-91-1) added. The tube was tightly capped and incubated for 24 h at 110ºC. The solution was then acidified with 6 drops of 6 M HCl, hexane (2 x 3 ml) was added, and the mixture was centrifuged. The hexane layer was transferred to a new tube, dried under nitrogen gas, methylated with 2 ml of HCl : methanol for 1 h at 80ºC, followed by two extractions in hexane (2 ml each). Samples were dried under nitrogen gas, acetylated with 1 ml of acetonitrile and 70 µl of bis-(trimethylsilyl) trifluoroacetamide for 20 min at 60ºC. The solvent was evaporated under nitrogen gas and 200 µL (for leaf samples) or 1.5 mL (for seed samples) of chloroform was

added.

Chloroform-dissolved material was subjected to FA methyl ester (FAME) analysis using a Model 6890 series gas chromatograph (Agilent, Palo Alto, CA, USA) equipped with a Mass detector 5973 (Agilent) and an HP-1 silica capillary column (30 m x 30.32 mm, inner diameter), using helium as a carrier gas[33,34]. The injector and detector temperature was 300ºC. The initial oven temperature was 100ºC for 2 min, increased to 240ºC at a rate of 5ºC min^{-1}, and maintained at 240ºC for 5 min. The total running time was 40 min. AMDIS software version 2.65 was used for peak integration[35]; peaks were identified through the Agilent NIST05 mass spectra libraries (http://www.nist.gov/srd/nist1a.htm). FAMEs were quantified in each sample in accordance with the amount of nonadecanoic acid internal standard added.

Statistical and bioinformatics analysis. All experiments were conducted a minimum of three times. For each experiment, plants were harvested and analyzed in random block design. Data is presented as mean ± standard deviations. We compared two sets of independent samples using student's *t* test (two-tailed) with assumption of equal variances and $P < 0.05$ was considered significant.

Pearson correlation coefficients of *A. thaliana* transcript accumulation across changes in plant genotype, environment, and development, were calculated using MetaOmGraph (MOG, http://www.metnetdb.org)[36] and a dataset consisting of 951 chips from 72 public microarray experiments[27,37].

SUPPORTING REFERENCES

1. Birney, E., Clamp, M. & Durbin, R. GeneWise and Genomewise. *Genome Res.* **14**, 988-995 (2004).
2. Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792-1797 (2004).
3. Larkin, M. A., *et al*. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
4. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17**, 754-755 (2001).
5. Guindon, S., Lethiec, F., Duroux, P.& Gascuel, O. PHYML Online – a web server for fast maximum likelihood-based phylogenetic inference. *Nucleic Acids Res.* **33**, W557-W559 (2005).
6. Jez, J. M., Ferrer, J. L., Bowman, M. E., Dixon, R. A. & Noel, J. P. Dissection of malonyl-coenzyme A decarboxylation from polyketide formation in the reaction mechanism of a plant polyketide synthase. *Biochem.* **39**, 890-902 (2000).
7. Doublié, S. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276**, 523–530 (1997).
8. Niesen, F. H., Berglund, H. & Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protocols* **9**, 2212-2221 (2007).
9. Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography, No. 26 (1992).

10. Evans, P. R. Scaling and assessment of data quality. *Acta Crystallog.* **D62**, 72-82 (2005).
11. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307-326 (1997).
12. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallog.* **26**, 795-800 (1993).
13. Terwilliger, T. C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Cryst.* **D55**, 849-861 (1999).
14. Terwilliger, T. C. Automated main-chain model building by template matching and iterative fragment extension. *Acta Cryst.* **D59**, 38-44 (2003).
15. Vagin, A. & Teplyakov, A. MOLREP: an automated program for molecular replacement. *J. Appl. Crystallog.* **30**, 1022-1025 (1997).
16. Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**, 779–815 (1993).
17. Perrakis, A., Morris, R. & Lamzin, V. S. Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.* **6**, 458-463 (1999).
18. Brunger, A. T. & Warren, G. L. Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Cryst.* **D54**, 905-921 (1998).
19. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst.* **D53**, 240-255 (1997).
20. McRee, D. E. XtalView/Xfit: a versatile program for manipulating atomic coordinates and electron density. *J Struct. Biol.* **125**, 156-165 (1999).
21. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst*. **D60**, 2126-2132 (2004).
22. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. *Acta Cryst.* **D50**, 760-763 (1994).
23. Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Cryst.* **D60**, 2256-2268 (2004).
24. Invitrogen. Gateway technology: A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. Catalog nos. 12535-019 & 12535-027 (2003).
25. Clough, S. J. & Bent, A. F. Floral dip: A simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J.* **16**, 735-743 (1998).
26. Lohar, D. P., Shuller, K., Buzas, D. M., Gresshoff, P. M. & Stiller, J. Transformation of *Lotus japonicum* using the herbicide resistance *bar* gene as a selectable marker. *J. Exp. Bot.* **52**, 1697-1702 (2001).
27. Li, L., Ilarslan, H., James, M. G., Myers A. M. & Wurtele E. S. Genome wide co-expression among the starch debranching enzyme genes AtISA1, AtISA2, and AtISA3 in *Arabidopsis thaliana*. *J. Exp. Bot.* **58**, 3323-3342 (2007).
28. Jefferson, R. A. Assaying chimeric genes in plants: the GUS gene fusion system. *Plant Mol. Biol. Rep.* **5**, 387-405 (1987).
29. Doyle, J. J. & Doyle, J. L. Isolation of plant DNA from fresh tissue. *Focus* **12**, 13-15 (1990).
30. Bonaventure, G., Salas, J. J., Pollard, M. R., & Ohlrogge J. B. Disruption of the *FATB* gene in *Arabidopsis* demonstrates an essential role of saturated fatty acids in plant growth. *Plant Cell* **15**, 1020-1033 (2003).
31. Markham, J. E., Li J., Cahoon, E. B. & Jaworski J. G. Separation and identification of

major plant sphingolipid classes from leaves. *J. Biol. Chem.* **281**, 22684-22694 (2006).
32. Moon, S. & Nikolau, B. J. Plantmetabolomics Platform2: fatty acids. plantmetabolomics.org (2009).
33. Perera, M. A. D. N., Dietrich, C. R., Meeley, R., Schnable, P. S. & Nikolau, B. J. Dissecting the maize epicuticular wax biosynthetic pathway via the characterization of an extensive collection of *glossy* mutants. *In* Advanced Research on Plant Lipids (eds. M Murata, M Yamada, I Nishida, H Okuyama, J Sekiya, W Hajima, H Wada). Kluwer Academic Press, Boston, 225-228 (2003).
34. Fatland, B. L., Nikolau, B. J. & Wurtele, E. S. Reverse genetic characterization of cytosolic acetyl-CoA generation by ATP-citrate lyase in *Arabidopsis*. *Plant Cell* **17**, 182–203 (2005).
35. Davies, N. A. The new Automated Mass Spectrometry Deconvolution and Identification System (AMDIS). ISAS Dortmund (1998).
36. Mentzen, W. I. & Wurtele, E. S. Regulon organization of Arabidopsis. *BMC Plant Biol.* **8**, 99 (2008).
37. Mentzen, W. I., Peng, J., Ransom, N., Nikolau, B. J. & Wurtele, E. S. Articulation of three core metabolic processes in Arabidopsis: fatty acid biosynthesis, leucine catabolism and starch metabolism. *BMC Plant Biol.* **8**, 76 (2008).
38. Earley, K.W., Haag, J.R., Pontes, O., Opper, K., Juene, T., Song, K., & Pikaard, C.S. Gateway- compatible vectors for plant functional genomics and proteomics. *Plant J.* **45**, 616-629 (2006).
39. Bradford, M.M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Annu Rev Biochem*. **72**, 248-254 (1976).