

Amino Acid Sequence of Anionic Peroxidase from the Windmill Palm Tree *Trachycarpus fortunei*

Margaret R. Baker,[†] Hongwei Zhao,^{†,‡} Ivan Yu. Sakharov,[§] and Qing X. Li*,[†]

[†]Department of Molecular Biosciences and Bioengineering, University of Hawaii at Manoa, Honolulu, Hawaii 96822, United States

[‡]School of Environmental Science and Engineering, Tianjin University, Tianjin 300072, China

[§]Department of Chemistry, Lomonosov Moscow State University, Moscow 119991, Russia

ABSTRACT: Palm peroxidases are extremely stable and have uncommon substrate specificity. This study was designed to fill in the knowledge gap about the structures of a peroxidase from the windmill palm tree *Trachycarpus fortunei*. The complete amino acid sequence and partial glycosylation were determined by MALDI-top-down sequencing of native windmill palm tree peroxidase (WPTP), MALDI-TOF/TOF MS/MS of WPTP tryptic peptides, and cDNA sequencing. The propeptide of WPTP contained N- and C-terminal signal sequences which contained 21 and 17 amino acid residues, respectively. Mature WPTP was 306 amino acids in length, and its carbohydrate content ranged from 21% to 29%. Comparison to closely related royal palm tree peroxidase revealed structural features that may explain differences in their substrate specificity. The results can be used to guide engineering of WPTP and its novel applications.

KEYWORDS: amino acid sequencing, glycosylation, palm, peroxidase, MALDI-TOF, top-down mass spectrometry

INTRODUCTION

Plant secretory peroxidases (class III peroxidases; EC 1.11.1.7) are ubiquitous in nature. They are extracellular or vacuolar glycoproteins and catalyze redox reactions that facilitate a myriad of biological processes including cell wall synthesis and response to abiotic and biotic stresses.¹ The peroxidase-catalyzed reaction occurs in three steps according to the “ping-pong” mechanism. First, the resting state peroxidase (E) is oxidized by hydrogen peroxide (H_2O_2). Next, two back-to-back single-electron transfers to an aromatic reducing substrate (AH_2) proceed. The reaction ends with the return of peroxidase to the resting state and generation of water and radical products (AH^{\bullet}).² The reaction can be summarized as $H_2O_2 + 2 AH_2 \xrightarrow{E} 2 H_2O + 2 AH^{\bullet}$.

Peroxidases are an important component of biosensors and immunochemical kits due to their sensitive and quantitative detection of H_2O_2 ^{3,4} and persistent organic compounds such as pesticides.⁵ Peroxidase-based biosensors have been commonly used for detection of antioxidants, such as polyphenols, flavonoids, and carotenoids, which gives an indication of the nutritional quality of foodstuffs.⁶ Some enzyme immunoassays have been developed using horseradish peroxidase (HRP) to detect food allergens⁷ and metal ions that may contaminate food and agricultural soil.⁸

Anionic peroxidases purified from palm tree leaves possess extremely high stability.⁹ Moreover, these enzymes showed distinct substrate specificity compared with other plant peroxidases. Their unique properties have allowed for development of novel and improved applications. Palm peroxidase-based biosensors were more stable and could tolerate higher concentrations of H_2O_2 .¹⁰ The exceptional stability of palm peroxidases at acidic pH met the requirement for synthesis of conductive and chiral polymers under environmentally safe conditions.^{11–13}

Some plant peroxidases, particularly HRP, are well studied; however, relatively little is known about the structure of palm peroxidases. The X-ray crystal structure of native, highly glycosylated royal palm tree peroxidase (RPTP) was recently solved.¹⁴ However, it is a unique work, and structural data for other palm tree peroxidases are practically absent in the literature, except for the sequence of 20 amino acids of the N-terminus of windmill palm tree peroxidase (WPTP, *Trachycarpus fortunei*).¹⁵ Also, it is known that WPTP uniquely contains 6.5 calcium cations per molecule, whereas most plant peroxidases, including HRP and RPTP, contain only 2 Ca²⁺.^{14,16}

To fill in the knowledge gap about the structures of palm peroxidases, herein we describe the complete amino acid sequence of WPTP and provide evidence of its modification with N-linked glycans. The complete amino acid sequence and glycosylation were determined by MALDI-top-down sequencing of native WPTP, MALDI-TOF/TOF MS/MS of WPTP tryptic peptides, and cDNA sequencing. The propeptide of WPTP contained N- and C-terminal signal sequences. Mature WPTP was 306 amino acids in length and its carbohydrate content was in the range of 21–29%. The results can be used to guide engineering of WPTP and its novel applications. A more robust peroxidase such as WPTP would enable broader use in agricultural and food applications.

MATERIALS AND METHODS

Materials. WPTP was isolated and purified from *T. fortunei* leaves as previously described.¹⁵ Trypsin (MS grade) was purchased from Promega (Madison, WI). Bovine serum albumin (BSA; 98% purity), used as a protein standard, was from Sigma (St. Louis, MO). The

Received: September 17, 2014

Revised: November 3, 2014

Accepted: November 10, 2014

Published: November 10, 2014

MALDI matrix 2,5-dihydroxybenzoic acid (DHB), peptide calibration standard II, and protein calibration standard II were from Bruker (Billerica, MA). C₄ ziptips were from EMD Millipore (Darmstadt, Germany). Other reagents, including HPLC grade acetonitrile (ACN), trifluoroacetic acid (TFA), and ammonium bicarbonate (AMBIC) were from Fisher Scientific (Waltham, MA). Water was purified on a Milli-Q Advantage A10 system (EMD Millipore). Dithiothreitol was from Acros Organics (New Jersey) and iodoacetamide was from BioRad (Hercules, CA). For gene cloning, an RNeasy Plant Mini Kit (Qiagen, Venlo, Limburg), a SMARTer RACE cDNA Amplification kit, and an Advantage 2 PCR kit from Clontech (Mountain View, CA), a pGEM-T vector (Promega), and *Escherichia coli* strain JM109 were used.

Molecular Mass Measurement. Purified WPTP was dissolved in 0.1% TFA, mixed 1:1 with a MALDI matrix solution (20 g/L DHB in 50% ACN, 0.1% TFA), spotted onto a polished steel target plate (Bruker), and allowed to air-dry. The molecular mass was measured with an Ultraflex III MALDI-TOF/TOF mass spectrometer (Bruker) in linear positive mode. In 200 shot increments, 1500 laser shots were accumulated. The matrix suppression cutoff, using gating, was at 9500 *m/z*. Pulsed ion extraction (PIE) delay was set to 150 ns. Instrument voltages were at 25 kV (ion source 1 [IS1]), 23.1 kV (IS2), and 6.5 kV (lens). The WPTP molecular mass value was calibrated using the protein mix II calibration standard (Bruker). Spectrum processing, consisting of peak detection (centroid algorithm, peak width 1000 *m/z*) and smoothing (Savitzky–Golay algorithm, 1 cycle at 20 *m/z*), was performed in Flex Analysis 3.4 (Bruker).

Top-Down Sequencing. Top-down sequencing (TDS) of native WPTP was performed after solid-phase reduction of its disulfide bonds. BSA was used as the control and the molecular mass calibrator. For solid-phase disulfide bond reduction, a C₄ ziptip was wetted with 50% ACN in 250 mM AMBIC solution several times. The protein sample (25 pmol in 15 μ L of 100 mM AMBIC) was loaded on the tip equilibrated with 100 mM AMBIC by slowly repipetting at least 10 times. Then, 50 mM dithiothreitol (200 μ L) was introduced to the tip by aspirating and dispensing the solution for 2 min at 60 °C. The ziptip was then incubated in that solution for 30 min at 60 °C. The tip was washed three times with 0.1% TFA. The protein was eluted with 5 μ L of a MALDI matrix solution (40 g/L DHB in 70% ACN, 0.1% TFA), and 0.25 μ L was spotted directly on the target plate and allowed to air-dry.

TDS spectra were acquired with an Ultraflex III MALDI-TOF/TOF mass spectrometer in reflector positive mode. In 200 shot increments, 19 000 laser shots were accumulated. The matrix suppression cutoff, using deflection, was at 900 *m/z*. PIE delay was set to 40 ns. Instrument voltages were at 25 kV (IS1), 21.4 kV (IS2), 9.7 kV (lens), 26.3 kV (reflector 1), and 13.8 kV (reflector 2). Spectrum processing, consisting of peak detection (SNAP algorithm, peak width 0.75 *m/z*), baseline subtraction (TopHat algorithm), and smoothing (Savitzky–Golay algorithm, 5 cycles at 1.5 *m/z*), as well as peak annotation was carried out in Flex Analysis 3.4. Peaks were assigned with a mass tolerance of 300 ppm. Interpretation of the ions comprising the near C-terminus was aided by the characteristic *m/z* difference of 15 ($\Delta m/z$) between γ - and α -ion series. Additional annotation was performed with BioTools (Bruker) and Sequence Editor (Bruker).

Bottom-Up Sequencing. WPTP (10 μ g, 1 μ g/ μ L in 25 mM AMBIC) was reduced with 50 mM dithiothreitol (20 μ L) for 30 min at 60 °C, alkylated with 100 mM iodoacetamide (25 μ L) for 45 min at room temperature in the dark. The alkylation reaction was quenched by addition of 1 μ L of 143 mM dithiothreitol. Reduced and alkylated WPTP was diluted with water and then digested in solution with 0.5 μ g of trypsin (50 μ L) at 37 °C for 12 h. Peptides were dried via SpeedVac and resuspended in 5% ACN in 0.1% TFA. The peptide solution was mixed 1:1 with a MALDI matrix solution (20 g/L DHB in 50% ACN, 0.1% TFA) and then spotted on a target plate.

Mass spectra of the peptides were acquired with an Ultraflex III MALDI-TOF/TOF mass spectrometer in reflector positive mode. In 200 shot increments, 600 shots were accumulated. The matrix suppression cutoff, using deflection, was at 850 *m/z*. PIE delay was set to 0 ns. Instrument voltages were at 25 kV (IS1), 21.45 kV (IS2), 9.5 kV (lens), 26.3 kV (reflector 1), and 13.8 kV (reflector 2). External calibration was performed using peptide mix II calibration standard

(Bruker). Spectrum processing, consisting of peak detection (SNAP algorithm, peak width 0.75 *m/z*) and smoothing (Savitzky–Golay algorithm, 1 cycle at 0.2 *m/z*), was performed in Flex Analysis 3.4.

Several peaks in the mass spectrum were selected for fragmentation in LIFT mode.¹⁷ Several thousand laser shots were accumulated in 200 shot increments. PIE delay was set to 0 ns. Instrument voltages were at 8 kV (IS1), 7.2 kV (IS2), 3.6 kV (lens), 29.5 kV (reflector 1), and 13.85 kV (reflector 2). Spectrum processing, consisting of peak detection (SNAP algorithm, peak width 0.75 *m/z*) and smoothing (Savitzky–Golay algorithm, 4 cycles at 0.15 *m/z*) as well as peak annotation, was carried out in Flex Analysis 3.4. Additional annotation was performed with BioTools and Sequence Editor with a mass tolerance of 0.3 *m/z* for peptides and 1 *m/z* for glycopeptides. Fragmentation spectra were interpreted and used for primer design in the gene cloning experiment.

Cloning and Sequencing of the cDNA. Total RNA was isolated from 0.1 g of *T. fortunei* leaves using an RNeasy Plant Mini Kit (Qiagen). The first-strand cDNA was synthesized from the isolated RNA using a SMARTer RACE cDNA Amplification kit (Clontech). The WPTP gene was cloned in two steps. In the first step, the forward degenerate primers (5'-GAYCTNCARATHGGNTT-3' and 5'-GAYTRCARATHGGNTT-3') targeted the previously determined amino acid sequence of the N-terminus of WPTP (DLQGIFY).¹⁵ The reverse primer, UPM, was provided with the same kit. PCR amplification was performed using an Advantage 2 PCR kit (Clontech). The PCR protocol consisted of an initial denaturation step at 94 °C (2 min), which was followed by 5 cycles of denaturation at 94 °C (30 s), annealing at 60 °C (30 s), and extension at 68 °C (1 min); then 5 cycles of denaturation at 94 °C (30 s), annealing at 55 °C (30 s), and extension at 68 °C (1 min); then followed by 25 cycles of denaturation at 94 °C (30 s), annealing at 50 °C (30 s), and extension at 68 °C (1 min); and then a final extension for 3 min at 68 °C. The resulting 1200 bp product was used as a template in the second step. The forward primer (5'-ATGCAYTTYCAYGAYTGYTT-3') was designed from the amino acid sequence MHFHDCE, obtained by tandem mass spectrometry analysis of WPTP tryptic peptides in this study. The PCR protocol was the same as above except that the annealing temperatures were 65 °C, 62 °C, and 60 °C, respectively. The PCR product was cloned into a pGEM-T vector and transformed into *E. coli* strain JM109. Positive transformants were screened by PCR. Plasmids were extracted from positive clones and confirmed by sequencing. For 5'-RACE, the reverse primer, 5'-GGACCTGGTAGGTGATGTTGCCG-3', was designed from the DNA sequence obtained in the previous step. The forward primer, UPM, was from the same kit. The PCR protocol was the same as above except that the annealing temperatures were 65 °C, 62 °C, and 58 °C, respectively. The PCR product was cloned, transformed, screened, and sequenced as before. The complete cDNA sequence of WPTP was obtained by combining the 5'- and 3'-cDNA sequences which were overlapping with each other. Analysis of the data was conducted with the software Mega5.¹⁸ These data can be accessed at NCBI accession: KMS04969.

RESULTS AND DISCUSSION

Determination of the Molecular Mass of WPTP. The MALDI-TOF mass spectrum of WPTP indicated the presence of a single glycoprotein with considerable heterogeneity (Figure 1A). An expanded view of the +2 charged peak revealed that it was composed of many partially resolved peaks, probably corresponding to WPTP glycoforms, i.e., differentially glycosylated WPTP (Figure 1B). A similar view of the +1 charged peak also showed multiple, partially resolved peaks.

The obtained data allowed for calculation of the MW of WPTP which ranged from 41 to 45 kDa (Figure 1C). These values are a little lower than the value previously reported for WPTP (50 kDa), which was measured using SDS-PAGE.¹⁵ Due to decreased binding of SDS to glycans relative to the polypeptide backbone, the migration of a glycoprotein during SDS-PAGE is altered relative to what is expected for a similarly sized nonglycosylated protein.¹⁹ Therefore, the previous measurement

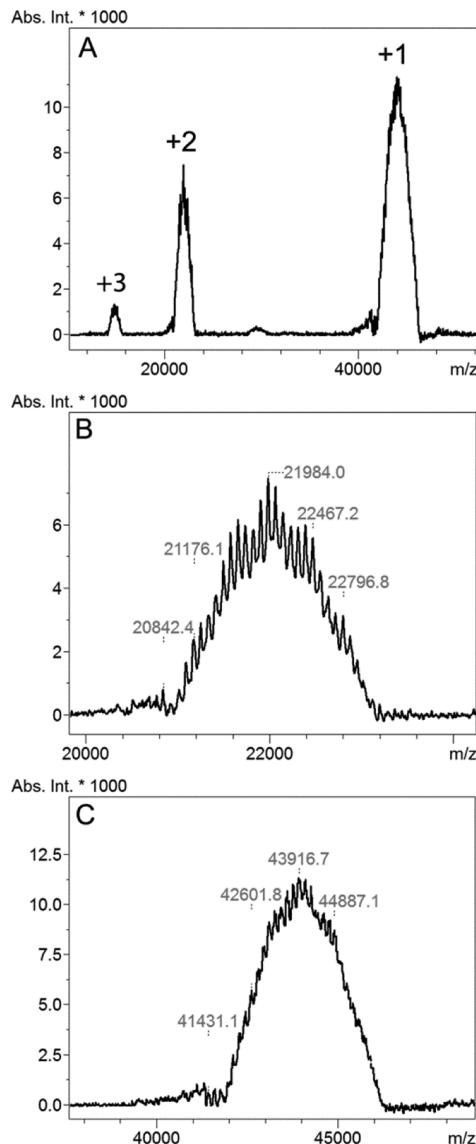


Figure 1. MALDI-TOF MS of WPTP to determine the molecular mass: (A) +1, +2, and +3 charged peaks and an expanded view of the (B) +2 charged and (C) +1 charged peak.

was likely an overestimation. Similarly, the first report of RPTP's MW, as measured by SDS-PAGE, was 51 kDa²⁰ and the MW as measured by mass spectrometry was 44 596 Da.¹⁴

Top-Down Sequencing. It is well-known that drawbacks of MALDI-TDS are the requirement of a high purity of analyte, low sensitivity, and lack of fragmentation in the presence of disulfide bonds.²¹ The isolated WPTP used in this study was high purity.¹⁵ WPTP was immobilized onto a C4 ziptip by hydrophobic interactions, and its disulfide bonds were then reduced. This provided a quick and effective way to reduce WPTP's disulfide bonds, remove contaminating salts, and concentrate the reduced protein in one simple step.

MALDI-TDS results in ladder fragmentation of the intact protein, leaving glycans attached to the modified Asn residues.²² Partial amino acid sequence near the N-terminus of WPTP was gained through interpretation of the c-ion series in the MALDI-TDS spectrum (Figure 2A, MHFHD). The y- and z+2-ions were interpreted as a continuous 30-amino acid residue sequence (Figure 2A, DN*LTAWVAKFAQQAIVKMGQIQVLGTQ-

GEI). A protein BLAST search of the NCBI database (blast.ncbi.nlm.nih.gov) predicted that this sequence originated from a protein in the peroxidase superfamily. The top hit was RPTP (PDB: 3HDL; E value: 6×10^{-16}). The y ions at m/z 5290.1 and 6575.7 (N^*) (and the corresponding $z+2$ -ions) had a mass difference corresponding to a glycosylated Asn, Asn-GlcNAc₂Man₃FucXyl (exact $\Delta m/z$ 1284.5 and accurate $\Delta m/z$ 1285.6). This glycan is the predominant glycan found on HRP,²³ soybean peroxidase (SBP),²⁴ RPTP,¹⁴ and other plant peroxidases. A scheme of the glycan is depicted in Figure 2B according to the conventions of the Consortium for Functional Glycomics. All assigned peaks had a mass error within 200 ppm and the root-mean-square error (RMSError) was 65.32 ppm.

Bottom-Up Sequencing. Additional amino acid sequence information was gained through a more traditional "bottom-up" approach. For this, WPTP was digested with trypsin, and the resulting tryptic peptides were subjected to MALDI-TOF MS (Figure 3A). Tandem mass spectrometry was performed to obtain the amino acid sequence of the tryptic peptides of WPTP (Figure 3B–F). Among these, fragments of the parent ion at m/z 1264.6 could be unambiguously interpreted as $M^{[ox]}HFHDCFVR$ (Figure 3D). Furthermore, this sequence overlapped with an amino acid sequence ladder from the MALDI-TDS experiment. Thus, this amino acid sequence was used for primer design for cDNA cloning and sequencing of WPTP.

Fragments of the parent ion at m/z 2890.3 displayed a characteristic spectrum of a glycopeptide (Figure 3F). Man and GlcNAc oxonium ions were observed at m/z 162.7 and 203.7, respectively. We also observed cross-ring fragmentation of the innermost GlcNAc (^{1,2}X). Additionally, the glycan moiety fragmented into a series of y-ions corresponding to mass shifts indicative of the glycan GlcNAc₂Man₃XylFuc (Figure 2C).

cDNA Sequencing. The gene for WPTP was cloned and sequenced from mRNA isolated from the leaves of *T. fortunei*. The gene consisted of 5'- and 3'- untranslated regions, which included a poly-A tail, and 1032 bp encoding a polypeptide of 344 amino acids (Figure 4). There was an N-terminal signal peptide upstream of the previously sequenced N-terminus of mature WPTP, which targets class III peroxidases to the secretory pathway.²⁵ Some peroxidases additionally bear a C-terminal signal peptide for vacuolar targeting.²⁵ Insights into in vivo processing of the C-terminus of WPTP were obtained by combining information from the MALDI-TDS spectrum of intact WPTP and its cDNA sequence. In a MALDI-TDS spectrum, the ion with the lowest m/z value in each ion series corresponds to the mass of the N-terminus (for c-ions) or the C-terminus (for y- and z-type ions). Therefore, the y 11-ion corresponded to the mass of RTN*CSVVNSAS (Figure 2A; exact m/z 2307.9 and accurate m/z 2307.8), where the mass shift for Asn²⁹⁸ (N^*) in the WPTP sequence corresponded to GlcNAc₂Man₃FucXyl. Thus, the vacuolar targeting signal for WPTP was LGDIVMASGHLTEVATS (Figure 4, gray italics).

Using the same logic for the N-terminus, the c 38-ion at m/z 6098.4 (Figure 2A) corresponded to the sequence DLQIG-FYN[#]QSCPASAEVLVQQAVAAFAN^{*}NSGIAPGLIR (Figure 4), where the mass difference of 2178.3 relative to the expected peptide mass (3920.1 Da) could be accounted for by adding a glycan to Asn⁸ (N^*) and one to Asn²⁸ with the potential sequences GlcNAc₂Man₂FucXyl ($\Delta m/z$ 1008.4) and GlcNAc₂Man₃FucXyl ($\Delta m/z$ 1170.4), giving an overall mass error of 65.6 ppm (exact m/z 6098.7 and accurate m/z 6098.4).

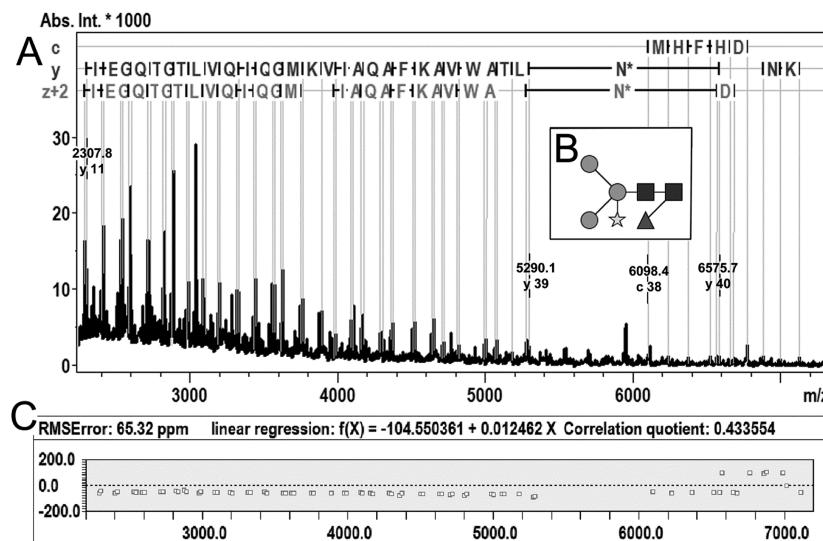


Figure 2. MALDI-top-down sequencing of WPTP to determine the amino acid sequence near the protein termini. The amino acid sequence was interpreted from c-, y-, and z+2-ion series. (A) $\Delta m/z$ between y 39 at m/z 5290.1 and y 40 at 6575.7 corresponded to Asn modified with the glycan GlcNAc₂Man₃XylFuc (N*). (B) A cartoon structure of the glycan GlcNAc₂Man₃XylFuc. (C) A plot showing the mass error (ppm) for each matched ion. (■) N-Acetylglucosamine (GlcNAc); (●) mannose (Man); (▲) fucose (Fuc); (★) xylose (Xyl).

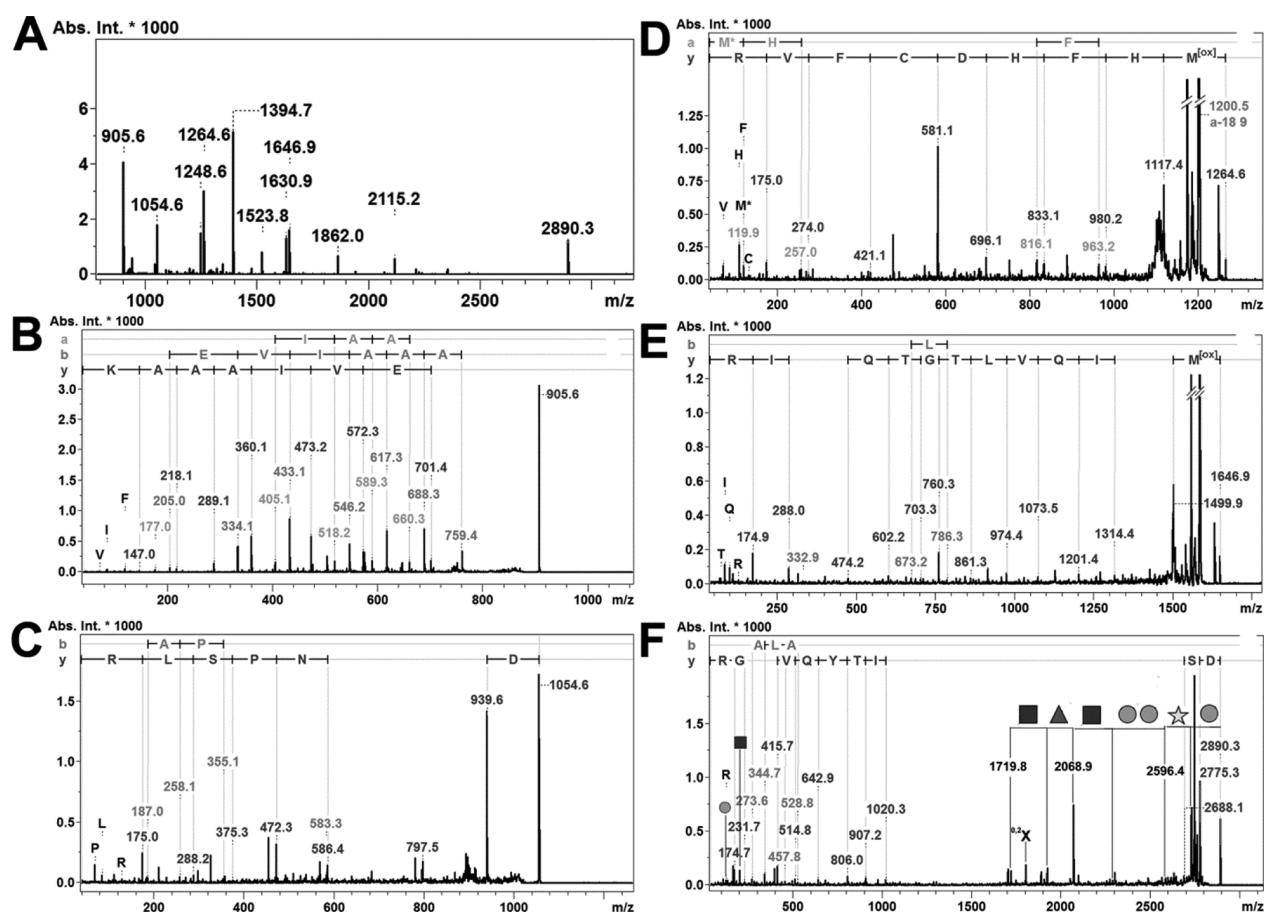


Figure 3. MALDI-TOF/TOF MS/MS of WPTP tryptic peptides for primer design: (A) mass spectrum WPTP tryptic peptides and fragmentation spectra of the parent ion at m/z (B) 905.6, (C) 1054.6, (D) 1264.6, (E) 1646.9, and (F) 2980.3. ($M^{[ox]}$) oxidized methionine; (■) N-acetylglucosamine (GlcNAc); (●) mannose (Man); (▲) fucose (Fuc); (★) xylose (Xyl).

Structure Overview. Mature WPTP was composed of a single polypeptide, 306 amino acid residues in length, with an average mass of 32 172 Da. The cDNA sequence showed that there were 13 potential glycosylation sites, identified by the

sequence Asn-Xxx-Ser/Thr, where Xxx is any amino acid except Pro. The MW, as measured by MALDI-TOF MS (Figure 1), ranged from 41–45 kDa. Considering this and the polypeptide mass, WPTP has 21–29% glycans. This level of complexity made

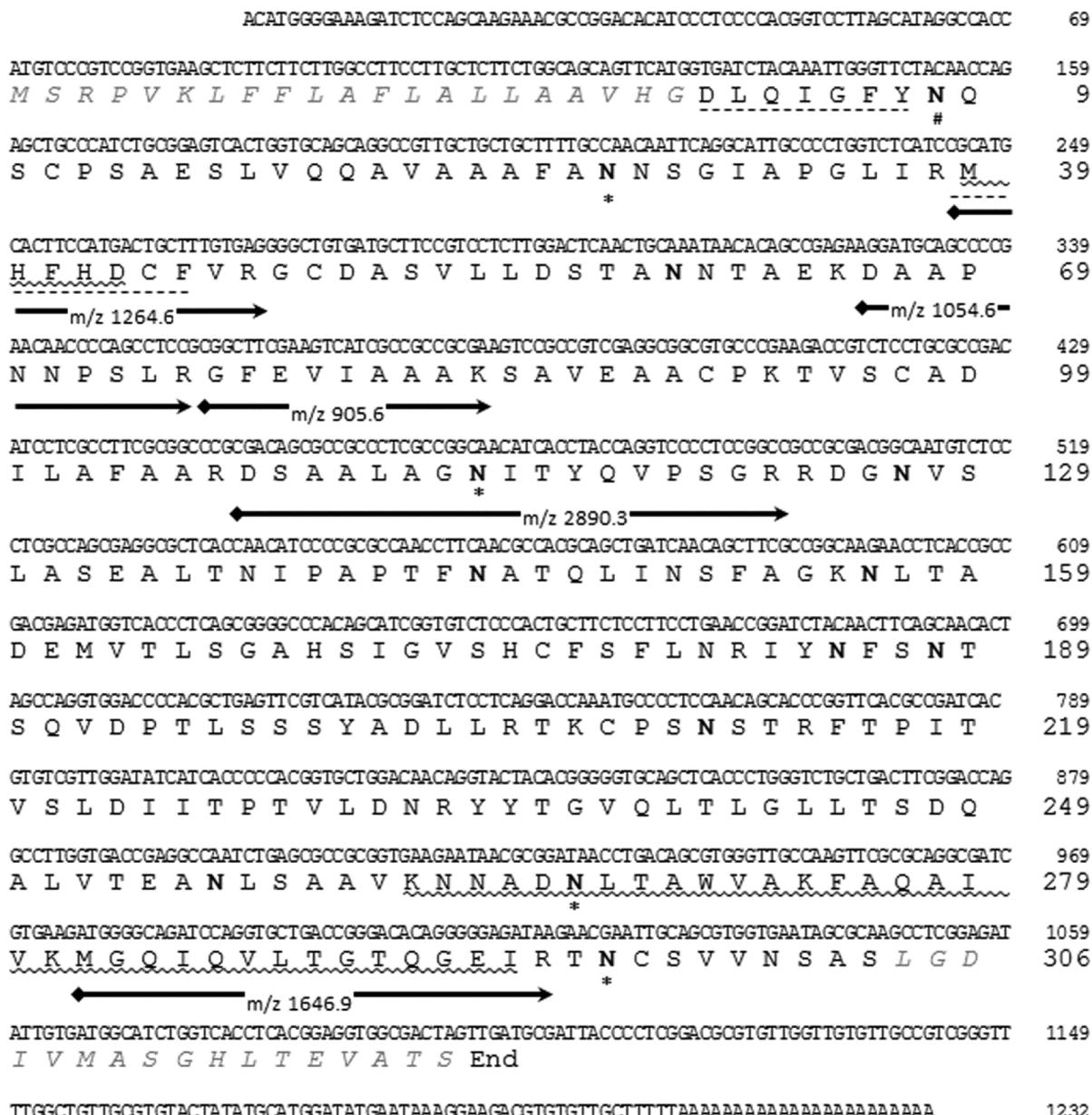


Figure 4. cDNA sequencing of WPTP to determine the complete amino acid sequence. Nucleotide and amino acid residue numbers are on the right. (gray italicized) N- and C-terminal signal peptides; amino acid sequence (----) used for primer design, (zig-zag) covered by MALDI-TDS, and (→) determined by MALDI-TOF/TOF MS/MS; (N) potential glycosylation site; (#) GlcNAc₂Man₂FucXyl; (*) GlcNAc₂Man₃FucXyl.

it difficult to assign a specific glycoform structure to each peak in that spectrum. Positions 8, 28, 114, 267, and 298 were glycosylated. A more detailed study of WPTP's glycosylation profile is underway.

Class III peroxidases share a common 3D structure despite a low sequence identity.^{2,16,26} For example, RPTP and HRP were only 36% identical, but their Cα backbone structures were closely related, having an overall root mean squared deviation of 1.07 Å.¹⁴ A structurally guided alignment revealed that WPTP and RPTP were 88% identical in primary structure and likely were very similar in secondary and tertiary structures (Figure 5).

Structural integrity of plant peroxidases is maintained largely by an extensive hydrogen bonding network that extends above and below the heme-containing active site to distal and proximal Ca²⁺.³² Additionally, four invariant disulfide bridges tie the protein together. These eight cysteine residues were present in WPTP (Figure 5, c). Importantly, the disulfide bridge between

Cys⁴⁴ and Cys⁴⁹ stabilizes the BC loop, which comprises the majority of the distal Ca²⁺ binding residues (Figure 5, +). Distal Ca²⁺ was critical for maintaining the proper tertiary structure of the active site.³³ The disulfide bridge between Cys¹⁷⁶ and Cys²⁰⁸ stabilizes the region between helices F and H, which is important for substrate binding and contains residues necessary for binding proximal Ca²⁺. The additional 4.5 Ca²⁺ found in the WPTP structure were difficult to account for because only one other peroxidase, anionic peanut peroxidase,³⁴ had elevated Ca²⁺ content and its 3D structure was still not available.

Features of Potential Importance to Substrate Specificity. The new WPTP structure information presented in this study allowed for exploration of features of potential importance in palm tree peroxidases. One striking functional difference is substrate reactivity differences between WPTP, RPTP, and HRP (Table 1). Ferulic acid was the better substrate for RPTP and HRP.²⁰ Moreover, the difference in reactivity

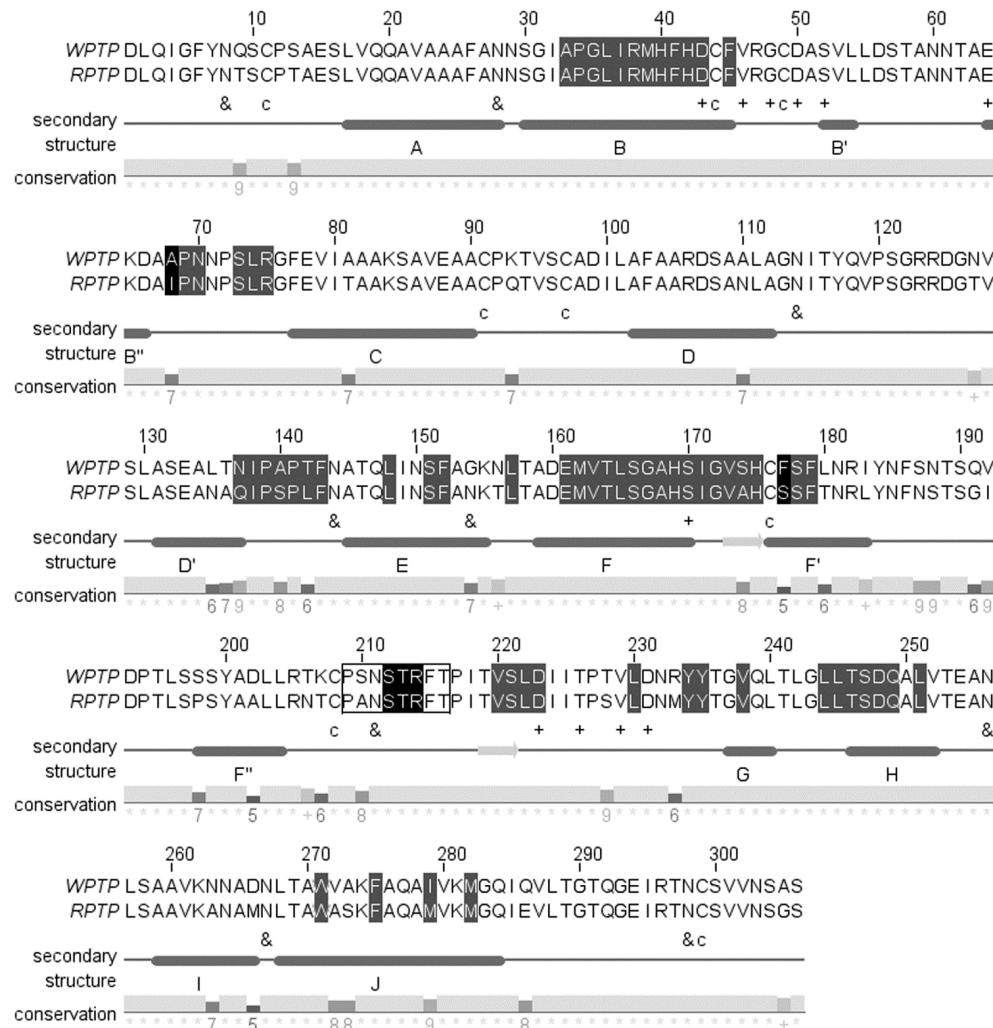


Figure 5. Alignment of WPTP and RPTP (PDB: 3HDL) according to the method of Armougom et al.²⁷ to determine features related to substrate specificity. Determination of (highlighted in gray) residues less than 12 Å from heme iron and (highlighted in black) less than 4.6 Å from MES with Swiss-PdbViewer 4.1.0²⁸ and (box) residues with significant structural deviation between RPTP and HRP (PDB: 1ATJ) with a protein structure alignment tool.²⁹ (&) Occupied glycosylation site;¹⁴ (c) cysteine; (+) Ca²⁺ binding. Secondary structure: (dark gray) helices; (light gray) sheets. Helices labeled according to Watanabe et al.¹⁴ and conservation according to Livingston and Barton.³⁰ The figure was drawn in JalView.³¹

Table 1. Substrate Reactivity As Measured by k_{app} ($M^{-1} s^{-1}$)

	WPTP ^a	RPTP ^b	HRP ^b
ferulic acid	3.0×10^6	6.3×10^7	1.3×10^7
ABTS ^c	6.0×10^7	5.0×10^7	4.0×10^6

^aThe data for WPTP are from ref 15. ^bThe data for RPTP and HRP are from ref 20. ^c2,2'-Azinobis(3-ethylbenzothiazoline-6-sulfonic acid).

between ferulic acid and ABTS was not too different for those two peroxidases, i.e., there was a 1.26-fold difference for RPTP and 3.25-fold difference for HRP.²⁰ The better substrate for WPTP was ABTS, and ferulic acid had 20-fold lower reactivity.¹⁵ Investigation of the 3D structure of RPTP, using Swiss-PdbViewer 4.1.0,²⁸ revealed amino acid residues and features of known and putative importance to substrate binding (Table 2).

Amino acids involved in substrate binding should be within 12 Å of the heme iron, as indicated by NMR studies of HRP–substrate complexes.³⁸ There were 64 residues within this distance in RPTP (Figure 5, highlighted in gray) and 5 of those differed between WPTP and RPTP. Of those 5, the lowest conservation was at position 142. Interactions between residue

142 and conserved Phe¹⁴³ resulted in topological features thought to affect interactions with substrates.^{2,39} Residue 142 was hydrophobic in HRP, RPTP, and WPTP, except that it was large in HRP (Phe) and RPTP (Leu) but small in WPTP (Thr). This difference could be a major factor causing the large deficit in WPTP's reactivity toward ferulic acid. The absence of a 3D structure complex between peroxidase and ABTS prevented speculation about features responsible WPTP's superior catalysis of that large substrate.

The RPTP crystal structure contained a potential inhibitor, an MES molecule.¹⁴ Residues within van der Waals radius (4.6 Å) of MES may affect substrate reactivity. Nine residues (Ile⁶⁸, Ser¹⁴⁰, Pro¹⁴¹, Leu¹⁴², Ser¹⁷⁷, Ser¹⁷⁸, Ser²¹², Thr²¹³, and Arg²¹⁴) were within that region. Five of the nine residues did not overlap with the residues that were 12 Å from the heme iron (Figure 5, highlighted in black). The position with the lowest conservation between RPTP and WPTP was at position 177 and WPTP was at position 177. In WPTP, it was large, hydrophobic, and aromatic Phe, whereas in RPTP, it was small and polar Ser. Moreover, this MES binding site was in the most structurally divergent region between HRP and RPTP (Figure 5, box). In

Table 2. Residues of Known and Potential Importance for Protein–Substrate Interactions

position ^a	WPTP	RPTP	predicted effect	evidence
Distance within 12 Å of Heme Iron				
38	R	R	critical	two H-bonds to ferulic acid (FA) (HRP); ³⁵ Arg ³⁸ to Leu 10–100 times slower substrate oxidation (HRP) ³⁶
69	P	P	minor	putative hydrophobic interaction between FA and anionic <i>Arabidopsis thaliana</i> peroxidase (ATP A2) ³⁷
138	I	I	minor	putative hydrophobic interaction with FA (ATP A2) ³⁷
139	P	P	major	H-bond to active site H ₂ O which is H-bonded to FA (HRP) ³⁵
140	A	S	minor	Ala: tiny ^b and hydrophobic vs Ser: small ^c and polar; hydrophobic interaction with FA (HRP: Ala ¹⁴⁰); ³⁵ H-bond to MES ^d (RPTP) ¹⁴
142	T	L	major	Thr: small ^c and hydrophobic vs Leu: large, hydrophobic, and aliphatic; hydrophobic interaction with FA (HRP) ³⁵
Distance within 4.6 Å of MES ^d				
68	A	I	minor	Ile: large, hydrophobic, and aliphatic vs Ala: tiny ^b and hydrophobic; hydrophobic interaction with FA (HRP: Gly ⁶⁸) ³⁵
177	F	S	major	Phe: large, hydrophobic, and aromatic vs Ser: small ^c and polar
212–214	STR	STR	major	structurally distinct region (HRP vs RPTP) and H-bond with MES ^d (RPTP: Arg ²¹⁴) ¹⁴

^aPosition in the amino acid sequence (Figure 5). ^bLess than 35 Å³. ^cLess than 60 Å³. ^d2-(N-Morpholino)ethanesulfonic acid.

HRP, this loop region was pointing away from the active site, whereas in RPTP, it was pointing toward it. This difference in topology brought Arg²¹⁴ into hydrogen bonding distance with MES.¹⁴ Further investigation into this distinct region is warranted.

AUTHOR INFORMATION

Corresponding Author

*Telephone: 808-956-2011; fax 808-956-3542; e-mail: qingl@hawaii.edu.

Funding

M.R.B. was a NIFA Pre-Doctoral Fellowship recipient. This project was supported in part by the Agriculture and Food Research Initiative Competitive Grant 2012-67011-19671 from the USDA National Institute of Food and Agriculture, NIH RCMI Grant 8G12MD007601, and the Korean RDA Grant 2008131.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors thank Mr. Sam Fu for consultation about the method for solid-phase reduction of disulfide bonds.

ABBREVIATIONS USED

ABTS, 2,2'-azinobis(3-ethylbenzothiazoline-6-sulfonic acid); ACN, acetonitrile; AMBIC, ammonium bicarbonate; ATP A2, anionic *Arabidopsis thaliana* peroxidase 2; DHB, 2,5-dihydroxybenzoic acid; FA, ferulic acid; HRP, horseradish peroxidase; IS, ion source; MES, 2-(N-morpholino)ethanesulfonic acid; PDB, protein databank; PIE, pulsed ion extraction; PNP, peanut peroxidase; TDS, top-down sequencing; TFA, trifluoroacetic acid; WPTP, windmill palm tree peroxidase

REFERENCES

- Zipor, G.; Oren-Shamir, M. Do vacuolar peroxidases act as plant caretakers? *Plant Sci.* **2013**, *199*, 41–47.
- Veitch, N. C. Structural determinants of plant peroxidase function. *Phytochem. Rev.* **2004**, *3*, 3–18.
- Gaspar, S.; Popescu, I. C.; Gazaryan, I. G.; Gerardo Bautista, A.; Sakharov, I. Yu.; Mattiasson, B.; Csöregi, E. Biosensors based on novel plant peroxidases: a comparative study. *Electrochim. Acta* **2000**, *46*, 255–264.
- Chen, S.; Yuan, R.; Chai, Y.; Hu, F. Electrochemical sensing of hydrogen peroxide using metal nanoparticles: A review. *Microchim. Acta* **2013**, *180*, 15–32.
- Farré, M.; Kantiani, L.; Barceló, D. Advances in immunochemical technologies for analysis of organic pollutants in the environment. *TrAC, Trends Anal. Chem.* **2007**, *26*, 1100–1112.
- Litescu, S. C.; Ereminia, S.; Radu, G. L. Biosensors for the determination of phenolic metabolites. In *Bio-Farms for Nutraceuticals*; Giardi, M. T.; Rea, G.; Berra, B., Eds.; Springer: New York, 2010; pp 234–240.
- Zhang, H.; Lu, Y.; Ushio, H.; Shiomi, K. Development of sandwich ELISA for detection and quantification of invertebrate major allergen tropomyosin by a monoclonal antibody. *Food Chem.* **2014**, *150*, 151–157.
- Zhao, H.; Nan, T.; Tan, G.; Gao, W.; Cao, Z.; Sun, S.; Li, Z.; Li, Q. X.; Wang, B. Development of two highly sensitive immunoassays for detection of copper ions and a suite of relevant immunochemicals. *Anal. Chim. Acta* **2011**, *702*, 102–108.
- Sakharov, I. Yu. Palm tree peroxidases. *Biochemistry (Moscow)* **2004**, *69*, 823–829.
- Alpeeva, I. S.; Niculescu-Nistor, M.; Leon, J. C.; Csöregi, E.; Sakharov, I. Yu. Palm tree peroxidase-based biosensor with unique characteristics for hydrogen peroxide monitoring. *Biosens. Bioelectron.* **2005**, *21*, 742–748.
- Sakharov, I. Yu.; Vorobiev, A. C.; Leon, J. J. C. Synthesis of polyelectrolyte complexes of polyaniline and sulfonated polystyrene by palm tree peroxidase. *Enzyme Microb. Technol.* **2003**, *33*, 661–667.
- Caramyshev, A. V.; Evtushenko, E. G.; Ivanov, V. F.; Barceló, A. R.; Roig, M. G.; Shnyrov, V. L.; van Huystee, R. B.; Kurochkin, I. N.; Vorobiev, A. K.; Sakharov, I. Yu. Synthesis of conducting polyelectrolyte complexes of polyaniline and poly(2-acrylamido-3-methyl-1-propanesulfonic acid) catalyzed by pH-stable palm tree peroxidase. *Biomacromolecules* **2005**, *6*, 1360–1366.
- Caramyshev, A. V.; Lobachov, V. M.; Selivanov, D. V.; Sheval, E. V.; Vorobiev, A. K.; Katasova, O. N.; Polyakov, V. Y.; Makarov, A. A.; Sakharov, I. Yu. Micellar peroxidase-catalyzed synthesis of chiral polyaniline. *Biomacromolecules* **2007**, *8*, 2549–2555.
- Watanabe, L.; de Moura, P. R.; Bleicher, L.; Nascimento, A. S.; Zamorano, L. S.; Calvete, J. J.; Sanz, L.; Pérez, A.; Bursakov, S.; Roig, M. G. Crystal structure and statistical coupling analysis of highly glycosylated peroxidase from royal palm tree (*Roystonea regia*). *J. Struct. Biol.* **2010**, *169*, 226–242.
- Caramyshev, A. V.; Firsova, Y. N.; Slasty, E. A.; Tagaev, A. A.; Potapenko, N. V.; Lobakova, E. S.; Pletjushkina, O. Y.; Sakharov, I. Yu. Purification and characterization of windmill palm tree (*Trachycarpus fortunei*) peroxidase. *J. Agric. Food Chem.* **2006**, *54*, 9888–9894.
- Gajhede, M.; Schuller, D. J.; Henriksen, A.; Smith, A. T.; Poulos, T. L. Crystal structure of horseradish peroxidase C at 2.15 Å resolution. *Nat. Struct. Biol.* **1997**, *4*, 1032–1048.
- Suckau, D.; Resemann, A.; Schuerenberg, M.; Hufnagel, P.; Franzen, J.; Holle, A. A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics. *Anal. Bioanal. Chem.* **2003**, *376*, 952–965.

- (18) Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGAS: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **2011**, *28*, 2731–2739.
- (19) Segrest, J. P.; Jackson, R. L.; Andrews, E. P.; Marchesi, V. T. Human erythrocyte membrane glycoprotein: A re-evaluation of the molecular weight as determined by SDS polyacrylamide gel electrophoresis. *Biochem. Biophys. Res. Commun.* **1971**, *44*, 390–395.
- (20) Sakharov, I. Yu.; Vesga B, M. K.; Galaev, I. Yu.; Sakharova, I. V.; Pletjushkina, O. Y. Peroxidase from leaves of royal palm tree *Roystonea regia*: Purification and some properties. *Plant Sci.* **2001**, *161*, 853–860.
- (21) Resemann, A.; Wunderlich, D.; Rothbauer, U.; Warscheid, B.; Leonhardt, H.; Fuchser, J.; Kuhlmann, K.; Suckau, D. Top-down *de novo* protein sequencing of a 13.6 kDa camelid single heavy chain antibody by matrix-assisted laser desorption ionization-time-of-flight/time-of-flight mass spectrometry. *Anal. Chem.* **2010**, *82*, 3283–3292.
- (22) Hanisch, F.-G. Top-down sequencing of O-glycoproteins by in-source decay matrix-assisted laser desorption ionization mass spectrometry for glycosylation site analysis. *Anal. Chem.* **2011**, *83*, 4829–4837.
- (23) Yang, B. Y.; Gray, J. S.; Montgomery, R. The glycans of horseradish peroxidase. *Carbohydr. Res.* **1996**, *287*, 203–212.
- (24) Gray, J. S. S.; Montgomery, R. Asymmetric glycosylation of soybean seed coat peroxidase. *Carbohydr. Res.* **2006**, *341*, 198–209.
- (25) Welinder, K. G.; Justesen, A. F.; Kjaersgård, I. V. H.; Jensen, R. B.; Rasmussen, S. K.; Jespersen, H. M.; Duroux, L. Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*: 73 peroxidases from *Arabidopsis*. *Eur. J. Biochem.* **2002**, *269*, 6063–6081.
- (26) Schuller, D. J.; Ban, N.; van Huystee, R. B.; McPherson, A.; Poulos, T. L. The crystal structure of peanut peroxidase. *Structure* **1996**, *4*, 311–321.
- (27) Armougom, F.; Moretti, S.; Poirot, O.; Audic, S.; Dumas, P.; Schaeli, B.; Kedua, V.; Notredame, C. Expresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Res.* **2006**, *34*, W604–W608.
- (28) Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. The SWISS-MODEL workspace: A web-based environment for protein structure homology modelling. *Bioinformatics* **2006**, *22*, 195–201.
- (29) Shindyalov, I. N.; Bourne, P. E. Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng.* **1998**, *11*, 739–747.
- (30) Livingstone, C. D.; Barton, G. J. Protein sequence alignments: a strategy for the hierarchical analysis of residue conservation. CABIOS, *Comput. Appl. Biosci.* **1993**, *9*, 745–756.
- (31) Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **2009**, *25*, 1189–1191.
- (32) Smulevich, G.; Feis, A.; Howes, B. D. Fifteen Years of Raman Spectroscopy of Engineered Heme Containing Peroxidases: What Have We Learned? *Acc. Chem. Res.* **2005**, *38*, 433–440.
- (33) Szigeti, K.; Smeller, L.; Osváth, S.; Majer, Z.; Fidy, J. The structure of horseradish peroxidase C characterized as a molten globule state after Ca²⁺ depletion. *Biochim. Biophys. Acta, Proteins Proteomics* **2008**, *1784*, 1965–1974.
- (34) Hu, C.; Lee, D.; Chibbar, R. N.; Huystee, R. B. Ca²⁺ and peroxidase derived from cultured peanut cells. *Physiol. Plant.* **1987**, *70*, 99–102.
- (35) Henriksen, A.; Smith, A. T.; Gajhede, M. The structures of the horseradish peroxidase c-ferulic acid complex and the ternary complex with cyanide suggest how peroxidases oxidize small phenolic substrates. *J. Biol. Chem.* **1999**, *274*, 35005–35011.
- (36) Rodriguez-Lopez, J. N.; Smith, A. T.; Thorneley, R. N. F. Role of arginine 38 in horseradish peroxidase: A critical residue for substrate binding and catalysis. *J. Biol. Chem.* **1996**, *271*, 4023–4030.
- (37) Nielsen, K. L.; Indiani, C.; Henriksen, A.; Feis, A.; Becucci, M.; Gajhede, M.; Smulevich, G.; Welinder, K. G. Differential activity and structure of highly similar peroxidases. Spectroscopic, crystallographic, and enzymatic analyses of lignifying *Arabidopsis thaliana* peroxidase A2 and horseradish peroxidase A2. *Biochemistry* **2001**, *40*, 11013–11021.
- (38) Veitch, N. C. Aromatic donor molecule binding sites of haem peroxidases. *Biochem. Soc. Trans.* **1995**, *23*, 232–240.
- (39) Henriksen, A.; Mirza, O.; Indiani, C.; Teilum, K.; Smulevich, G.; Welinder, K. G.; Gajhede, M. Structure of soybean seed coat peroxidase: A plant peroxidase with unusual stability and haem-apoprotein interactions. *Protein Sci.* **2001**, *10*, 108–115.