Purification, biochemical characterization and cloning of a new cationic peroxidase isoenzyme from artichoke

Angela Cardinali a,*, Nunzia Tursi a, Angela Ligorio b, Maria Gabriella Giuffrida c, Lorenzo Napolitano c, Rocco Caliandro d, Lucrezia Sergio a, Donato Di Venere a, Vincenzo Lattanzio e, Gabriella Sonnante f

a Institute of Sciences of Food Production-CNR, Via Amendola 122/O, 70126 Bari, Italy
b Department of Plant Protection and Applied Microbiology, University of Bari, Via Amendola 165/A, 70126 Bari, Italy
c Institute of Sciences of Food Production-CNR, Sezione di Torino c/o BioIndustry Park del Canavese, Via Ribes, 5 – 10010 Colleretto Giacosa (TO), Italy
d Institute of Cristallography-CNR, Via Amendola 122/O, 70126 Bari, Italy
e Department of Agro-Environmental Sciences, Chemistry and Crop Protection, University of Foggia, Via Napoli 25, 71100 Foggia, Italy
f Institute of Plant Genetics-CNR, Via Amendola, 165/A, 70126 Bari, Italy

Abstract

A cationic soluble peroxidase isoenzyme (CysPrx) has been purified and characterized from artichoke (Cynara cardunculus subsp. scolymus (L.) Hegi) leaves by combination of aqueous two phase extraction, ion exchange chromatography, and gel filtration. The purification fold was 149 and the activity recovery 5.5%. CysPrx was stable from 5 to 45 °C with a pH optimum around 5.5; the pI was 8.3 and the MW of 37.7 ± 1.5 kDa. MALDI-TOF MS analysis provided partial peptide sequences and resolved CysPrx isoenzyme into two putative isoforms. The presence of these isoforms was confirmed by the isolation of full-length cDNA encoding CysPrx that generate two slightly different sequences coding for two putative CysPrx: CysPrx1 and CysPrx2. The obtained MS peptides showed a 35% coverage with 100% identity with the two CysPrx deduced protein sequences. A molecular modeling analysis was carried out to predict in silico the protein structure and compare it with other plant Prx structures. Considering that CysPrx is quite stable, the study carried out in this paper will offer new insights for the production of the recombinant protein for utilization of CysPrx as an alternative Prx for food technology, biomedical analysis and bioremediation.

© 2011 Elsevier Masson SAS. All rights reserved.

1. Introduction

Peroxidases [donor: hydrogen peroxide oxidoreductase, EC 1.11.1.7] (Prx), are ubiquitous enzymes found in bacteria, fungi, plants and animals. They all accept hydrogen peroxide or hydroperoxide analogs as oxidant and form water as by-product. Prxs occur in plants as several isoenzymes or glycoforms, with different cellular localization. They catalyze the reduction of H2O2 by taking electrons from various donor molecules. The Prx super-family has essentially been classified into three classes: class I intracellular Prx of prokaryotic origin, class II fungal Prxs and class III the higher plant Prxs targeted for the secretory pathway [48].

In the case of class III plant Prxs, the donor molecules can be phenolics, lignin precursors, or secondary metabolites [19]. Moreover, these peroxidases are secreted glycoproteins that contain a FeIII-protoporphyrin-IX as the prosthetic group linked to a proximal histidine residue. The steric arrangement of the residues next to heme is crucial in determining the enzyme activity and stability [13,18]. Prxs have been found soluble in the apoplastic fluid or linked by ionic or covalent bond to cell wall components. Iso-peroxidases are widely distributed within both the intra and extracellular environment and are usually classified into 3 subgroups: anionic, neutral, and cationic Prxs based on their isoelectric points. However, their actual physiological role is often still unclear [11]. The numerous Prx enzymes can arise from the transcription of different genes or from posttranslational modification, in fact plant Prxs are coded by a large multigene family [32], and participate in a broad range of physiological processes including the removal of hydrogen peroxide from chloroplasts and cytosol, the oxidation of toxic compounds, the biosynthesis of cell

Abbreviations: Prx, peroxidase; CysPrx, Prx extracted from artichoke (Cynara cardunculus subsp. scolymus (L.) Hegi) leaves; HRP, horseradish peroxidase; BP, barley grain peroxidase; rpm, rotation per minute; MALDI-TOF, matrix assisted laser desorption ionization – time-of-flight; TFA, trifluoroacetic acid; EST, expressed sequence tag; PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends.

* Corresponding author. Tel.: +39 080 5929303; fax: +39 080 5929374.
E-mail address: angela.cardinali@ispa.cnr.it (A. Cardinali).
walls (lignin and suberin), indole-3-acetic acid regulation, ethylene biosynthesis, defense responses towards wounding, and other stresses. Prxs can also catalyze the formation of diferuloyl and isodityrosine linkages in both primary and secondary cell walls. The specific functions of individual Prxs are often difficult to understand because of their low substrate specificity and the existence of many isoenzymes [141].

The most thoroughly studied example of a higher plant Prx is horseradish (HPR), which catalyzes the oxidative coupling of phenolic compounds using H2O2 as the oxidizing agent. The reaction is a three-step cyclic reaction in which the enzyme is first oxidized by H2O2 and then reduced in two one-electron transfer steps by reducing substrates, typically small molecule phenol derivatives. The oxidized phenolic radicals can polymerize, with the final product depending on the chemical character of the radical, the environment and the Prx isoenzyme used. HPR is also the most commercially available Prx, even if one of the major challenges associated with this enzyme is its susceptibility to inactivation during thermal treatments [44]. Soybean seed coat was also identified as a rich source of Prx. This enzyme was found very stable at elevated temperature and was less susceptible to HRP to permanent inactivation by hydrogen peroxide, even if catalytically slower against aqueous phenols [49].

Relatively little is known about the structure and properties of artichoke (Cynara cardunculus subsp. scolymus (L.) Hegi) peroxidases [9,27,37,38]. In our previous study, a soluble cationic Prx (ALSP) was purified from artichoke leaves and partially characterized. The enzyme is a glycoprotein with a molecular weight of 51 kDa, an isoelectric point of 9 and is very stable at different temperatures and pH values. Its substrate specificity is characteristic of class III Prxs (guaiacol-type) [9]. Another Prx isoenzyme is present also in artichoke flower heads, where it displayed at lower activity [27].

Considering that Prx is a large gene family with several isoenzymes, the idea to learn more about the others isoenzymatic forms present in artichoke has become crucial for several reasons. First of all, the characterization of new enzymes could find interesting uses in food biotechnology. Cationic peroxidases may be related to the modification of functional properties of food proteins and carbohydrates [5]. Artichoke, in fact, can be affected by enzymatic browning during handling and storage. Besides to chemical interactions, Prx is involved in this mechanism together with polyphenoloxidase (PPO), another enzyme very abundant in artichoke [26]. In addition, cationic Prxs are more active with phenolic compounds compared to anionic peroxidases and laccases [45]. Thus, cationic peroxidases may be of interest for biocatalytic application such as analysis of glucose or O2, even in non aqueous medium, removal of dyes, chlorophenols and carcinogenic amines from residual water, and treatment of polluted water containing phenols [34].

With the aim of finding alternative Prxs for a wide range of applications, for a deeper knowledge about artichoke Prxs, the present paper deals with the purification, biochemical characterization and cloning of a new soluble cationic Prx isoenzyme from artichoke leaves. Mass spectrometry analysis and full-length cDNA cloning were performed and modeling analysis was carried out.

2. Materials and methods

2.1. Materials

Leaves of artichoke, cv. Violette di Provenza, were provided by the experimental field Azienda Pantanelli (Policoro, Matera, Italy). Leaves were washed, deprived of midribs before analysis (about 80% of total weight) and stored at −20°C until use. Analytical-grade H2O2 (30% v/v) was purchased from Carlo Erba Reagents (Rodano, Milano, Italy). Amicon YM 10 membrane from Millipore Corporation (290 Concord Road, Billerica, MA), Phenyl Sepharose CL-4B, CM Sepharose Fast Flow and Sephacryl S-300 High Resolution were purchased from GE HealthCare (Uppsala, Sweden). All the other analytical-grade chemicals were supplied by Sigma or Aldrich (St. Louis, MO, USA).

2.2. Isolation and purification of Prx

All purification steps were carried out at 4°C.

Fifty g of leaves were homogenized for 3 min in a Waring blender (15000 rpm) with 200 mL of 10 mM phosphate buffer, pH 7.0 and then incubated with constant agitation for 1 h at room temperature. After 1 h incubation, the tissue debris was removed by filtration and centrifugation at 7000 g, 4°C for 15 min. Solid polyethylene glycol (PEG 6000) and solid (NH4)2SO4 were dissolved in the supernatant up to 14% (w/v) and 20% (w/v), respectively. Three phases (the top polymer phase containing pigments, the bottom aqueous phase containing Prx and the solid phase which was accumulated at the interface) were formed after the addition of (NH4)2SO4 to the supernatant containing PEG. The suspension was centrifuged at 6000 g for 10 min and Prx activity was measured.

For all chromatographic steps, the AKTA Prime Plus (GE HealthCare Uppsala, Sweden) system was used. The lower clear phase containing Prx activity was applied directly to a Phenyl-Sepharose CL-4B 16/40 column (Volume 80 ml) equilibrated with 100 mM phosphate buffer, pH 7.0, containing 1.7 M (NH4)2SO4. After sample application, the column was eluted with a step gradient of 1.7, 1.0 and 0.0 M ammonium sulphate in phosphate buffer, pH 7.0 at a flow rate of 4 mL min−1. Fractions were collected and assayed for Prx activity using guaiacol as reducing substrate. Fractions showing Prx activity were pooled and concentrated using an Amicon YM-10 (Amicon, cut-off Mr 10 000) membrane to reduce the volume to 4 mL.

The active concentrated fractions, collected from hydrophobic chromatography, were desalted using Sephadex G-25 gel filtration column, pre-equilibrated with 0.02 M sodium acetate buffer, pH 6.0 and applied on CM Sepharose Fast 0.8/10 column (Volume 5 ml) equilibrated with the same buffer and eluted with a salt gradient of 0–0.5 M NaCl in the same buffer at a flow rate of 5 mL min−1. Two separate protein fractions showing Prx activity were collected from CM Sepharose, one in the void volume and another eluted with the salt gradient. The proteins recovered in the void volume were collected and stored for further purification. The active fractions from CM Sepharose were pooled and concentrated using an Amicon YM-10 membrane to reduce the volume to 3.8 mL.

Concentrated enzyme was loaded onto a gel-filtration column (Sephacryl S-300 High Resolution 16/70 column), equilibrated with 5 mM Tris—HCl, pH 8.3 and the elution was carried out in the same buffer at a flow rate of 1 mL min−1. The fractions having Prx activity were collected and named from now CysPrx.

2.3. Electrophoresis

The purity and molecular weight of Prx was determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). Before run, the protein was reduced with 0.01 M dithiothreitol (DTT) (100°C, 3–5 min) and alkylated by addition of 1 µL of 4-vinyl pyridine (room temperature, in the dark, 20 min). The result of this procedure made the protein ready for MS analysis. SDS-PAGE was performed using a modified Laemmli system [24]. Samples containing about 20 µg of protein were applied on 10% polyacrylamide gel; the electrophoretic conditions were 25 mA per gel for 1 h and 30 min. After run, the proteins were stained with
Colloidal Coomassie Brilliant Blue G-250 and for glycan analysis proteins were stained using the PAS method [30].

2.4. Two-dimensional gel electrophoresis (2DE)

Twenty μg of protein extract were solubilized in 150 μl of rehydration buffer solution (7 M urea, 2 M thiourea, 4% w/v 3-[(3-Cholamidopropyl) dimethylammonio]-1-propanesulfonate, 20 mM Tris, 0.5% Triton X-100, 1% DTT w/v and 0.2% BioLyte Ampholyte (Bio-Rad) and 125 μg of the protein extract were loaded onto a 7 cm Strip from BioRad, pH 3–10. IEF was carried out on an Protean IEF Cell (Bio-Rad) at 20 °C constant temperature and 5000 V for a 20,000 total Vh. After IEF, strips were incubated at room temperature in 4 mL of Lithium Dodecil Sulfate (LDS) sample buffer (Invitrogen) enriched with both 2% w/v DTT for 15 min and 4.5% (w/v) iodoacetamide for the same period of time. The second dimension of 2DE was carried out on NuPAGE gel 12% T (Invitrogen) using 3-(N-morpholino) propane sulfonic acid) MOPS as running buffer. Running conditions were 200 V constant voltages, 125 mA, 15 W/gel, for about 70 min. Gels were automatically stained using the Processor Plus (Amersham Biosciences, Uppsala, Sweden) with freshly prepared Blue Coomassie Colloidal stain.

2.5. Isoelectrofocusing (IEF)

IEF was performed using the Amersham Ampholine PAGplate system on precast polyacrylamide gels (pH range 3.5–9.5). The gel was calibrated using the calibration kit with the following protein standards: trypsinogen (pI 9.3); lentil lectin-basic band (pI 8.65), middle band (pI 8.45) and acidic band (pI 8.15); myoglobin-basic band (pI 7.35), acidic band (pI 6.85); human carbonic anhydrase (pI 6.55); bovine trypsin inhibitor (pI 5.20); soybean trypsin inhibitor (pI 4.55); B-lactoglobulin (pI 3.50); IEF gel, containing CysPrx, after run was stained with Colloidal Coomassie Brilliant Blue G-250.

2.6. Protein determination

Protein concentration was determined by the Bradford method using bovine serum albumin as standard [7].

2.7. Enzyme assay

Prx activity was determined with guaiacol as reducing substrate in a reaction mixture (1 ml) containing 0.1 M K phosphate buffer pH = 6.0, 0.02 M guaiacol and 0.03 M H2O2. The oxidation of guaiacol was followed by observing the increase in absorbance (A) at 470 nm (ε470nm = 26.6 mM−1 cm−1) and 25 °C. One unit of the enzyme is defined as the amount of enzyme that oxidizes 1 nmol guaiacol/min under the above assay conditions [42]. For the determination of the CysPrx classification, the activity of the enzyme was monitored by spectrophotometric measurement using ascorbic acid, guaiacol and veratryl alcohol. The rate of oxidation of guaiacol was followed at 470 nm (ε470nm = 26.6 mM−1 cm−1), of ascorbic acid at 290 nm (ε290nm = 2.8 mM−1 cm−1), and of veratryl alcohol at 310 nm (ε310nm = 9.3 mM−1 cm−1).

2.8. pH and thermal inactivation

The effect of pH on the substrate depletion activity of CysPrx was determined using guaiacol following tetraguaiacol formation at 470 nm. For this purpose four different buffers were used: 30 mM sodium acetate (pH 3.5–5.5), 30 mM sodium phosphate (pH 5.5–7.5), 30 mM Tris–HCl (pH 7.5–8.0) and 30 mM glycine–NaOH (pH 8.0–9.0). The thermostability of purified CysPrx was measured by incubating the enzyme at different temperatures (5–85 °C) for 10 min in a water bath and immediately cooled on ice [6]. Activity was assayed after 5 min using guaiacol as substrate.

2.9. Molecular weight determination by gel filtration

A Sephacryl S-300 High Resolution 16/70 column was equilibrated at room temperature with 50 mM phosphate buffer, pH 7.2, containing 0.15 M NaCl. Purified protein and calibration proteins [0.50 mg of ferritin (440 kDa), 3.9 mg of ovalbumin (45 kDa) and 3 mg of carbonic anhydrase (29 kDa) supplied by Sigma] were applied on the column and eluted at a flow rate of 1 ml/min. The molecular weight (MW) value of the protein was estimated from a semi-logarithmic plot of the MW values for the calibration proteins against elution volume.

2.10. N-Terminal sequencing

The bands present in SDS-PAGE gel corresponding to proteins with peroxidase activity, were excised from the gel, passively eluted and analyzed by microsequencing as already described [9].

2.11. Mass spectrometry analysis

The bands of interest in the SDS-PAGE gel were cut and “in gel” digested with trypsin (Promega, Madison, WI, USA) according to Hellmann and co-workers procedure [17]. The tryptic peptide mixtures were also analyzed by LC-MS/MS using an Agilent (Palo Alto, CA, USA) 1100 series nano HPLC coupled to an Agilent XCT Plus ion trap fitted with a nano electrospray nebulizer. The chromatographic separations were run on a C18 nanocolumn Zorbax 300SB 150 × 0.075 mm (Agilent). Injection volume was 1 μl and flow rate 300 nl/min. Gradient mobile phase composition was adopted: 95:5 for the first 10 min to 45/55 in 35 min and then to 70/30 in 5 min. Eluent A: water/0.1% formic acid/5% acetonitrile, eluent B: acetonitrile/0.1% formic acid. Capillary voltage was 1600 V and fragmentation voltage was 1.3 V. For elaboration of the LC and MS/MS data the Data Analysis software (Agilent) was used. Data interpretation were achieved by Mascot software (Matrix Science Ltd, London UK). Data were searched against the NCBI database subset “other green plants” (20091202) using the following parameters: three missed cleavages were allowed for trypsin, cysteins were considered completely modified by 4-vinyl pyridine and methionine partially oxidized. For further Prx confirmation, Peptide Mass Fingerprinting (PMF) interpretation was achieved by MS-Fit software (http://prospector.ucsf.edu/prospector/cgi-bin/msform) using the amino acid sequences deduced from cDNA with the same parameters used for LC-MS/MS data interpretation.

2.12. Full-length cDNA isolation

Total RNA was isolated from artichoke leaves using RNeasy Plant mini kit (Qiagen) according to the manufacturer’s instructions. Total RNA (1 μg) was used for the synthesis of single-strand cDNA with the AccuScript High Fidelity (Stratagene) following the provided instructions. Primers forward (NewPOD_F: GTCTCTTG TGCTGATATTACTGC) and reverse (NewPOD_R2: ACGGATTTCTCC TGTTGTTGCC) were designed on the artichoke ESTs GE580188 and GE592500. cDNA was amplified with an initial denaturation at 95 °C for 1 min, followed by 40 cycles of 95 °C for 30 s, 55 °C for 30 s, 68 °C for 3 min, with a final extension at 72 °C for 10 min. The PCR product was purified, cloned in a pGEM T-Easy vector (Invitrogen, Leek, The Netherlands), following the instruction manual, and sequenced. New oligonucleotide primers were synthesized for the completion of Prx transcript sequence, using the 5′ or 3′ RACE.
System for Rapid Amplification of cDNA Ends (Invitrogen). To identify the 5′-end, 5RACE_POD-R1 (AATCCAAATAGTTCG) primer and 1.5 μg total RNA were used for cDNA synthesis. Then, a first round of PCR was performed with the antisense primer 5RACE_POD-R2 (TGGCTAGCGTTTGTGTTG) along with the anchored RACE-specific AAP primer. Amplification reaction contained 1x buffer, 1.5 mM MgCl2, 0.2 mM dNTP mix, 400 mM each primer, 2.5 U recombinant Taq (Invitrogen), and 5 μl cDNA, in a final volume of 50 μL. Amplification cycles were: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 56 °C for 30 s, 72 °C for 90 s, and a final extension of 7 min at 72 °C. A second round PCR was carried out using 1 μL of a 1:100 dilution of the first PCR product as a template, and the nested antisense primer 5RACE_POD-R3 (CGTCAGACGGAGACGAACTAAT) with the universal RACE-specific AAUAP primer. The reaction mixture and PCR amplification protocol were the same as above, but with an annealing temperature (AT) of 59 °C. The amplified cDNA band was purified, cloned, and sequenced. For the completion of the 3′-end, 1 μg total RNA and the AAUAP primer were used for cDNA synthesis, and then PCR was performed with primers 3RACE_POD-F1 (CAATCAAGTGTGT TTTCGA) and AAUAP. Amplification reaction contained 1x buffer, 1.5 mM MgCl2, 0.2 mM dNTP mix, 200 mM each primer, 2.5 U recombinant Taq (Invitrogen), and 2 μL cDNA, in a final volume of 50 μL. Amplification cycles were: 94 °C for 3 min, 35 cycles of 94 °C for 30 s, 52 °C for 30 s, 68 °C for 1 min, and additional 10 min at 72 °C. For the nested PCR, 1 μL of a 1:500 dilution of the first PCR product was used as a template, together with the primers AAUAP and 3RACE - POD-F2 (ATGCCAGCATTGGCAGTTGG), in the same conditions as above, but with an AT of 54 °C, and an extension step at 72 °C for 45 s. Full-length cDNA clones were amplified using end-specific primers designed on the 5′ and 3′ RACE fragments (POD-FL-F: CTATCGCCTAGTGCCTTGT, and POD-FL-R: GGAACCAACAAACCAT ATATCCA). The PCR fragments were cloned and sequenced.

2.13. Sequence data analysis

Artichoke Prx nucleotide sequences were translated in silico (http://www.ebi.ac.uk/Tools/emboss/translate/index.html) and aligned with sequences from other plants, retrieved from the Prx database (http://peroxibase.toulouse.inra.fr/). Protein sequence alignment was subjected to phylogenetic analysis using the program MEGA 4 [39]. A phylogenetic tree was constructed by means of the neighbour-joining method of Saitou and Nei [35] with experimental evidence showed that after 10 min incubation at 3.5 using guaiacol as substrate 9 using guaiacol as substrate CysPrx had demonstrate to have high affinity for the phenolic compounds such as: chlorogenic and caffeic acids, present in artichoke leaves (ALSPr), which was found to have a MW of 51 kDa and the same pl [9]. The presence of more than one Prx enzyme in artichoke supports the hypothesis that a gene family gives rise to multiple isoenzymes belonging to class III Prx as in several other plant species [6,41]. In the majority of plant species, specific Prx isoenzymes can be constitutive or induced in response to external factors such as wounding, abiotic stress and attack by pathogens [44].

2.14. Structure prediction

The structure of CysPrx was predicted by comparative modeling, starting from its deduced sequences. The public servers I-TASSER (zhang.bioinformatics.ku.edu/I-TASSER) [52] and Phyre (www.sbg.bio.ic.ac.uk) [22] were used, as they were top performers in the recent CASP7 experiment for critical assessment of techniques for protein structure prediction [10]. They are based on the iterative threading algorithm applied on a representative sample of homologous structures [4,53]. The best models were generated by analyzed by means of the graphic interfaces Coot [18] and PyMOL [12] and subjected to molecular dynamics simulations by using the program NAMD [33].

3. Results and discussion

3.1. Purification of Cynara scolymus Peroxidase (CysPrx)

The cationic Prx artichoke isoenzyme (CysPrx) was purified through a combination of two phase systems extraction method, hydrophobic interaction, ion exchange and size exclusion chromatography. The two phase extraction [36] that utilizes the polyethylene glycol/(NH4)2SO4 system seems to be more convenient and promising compared to the polyvinylpyrrolidone (PVP) method utilized previously [9]. This method permits to almost completely eliminate the enzyme PPO, particularly abundant in artichoke leaves [26], from the aqueous (NH4)2SO4 phase. The lower clear salt phase containing Prx activity was submitted to hydrophobic chromatography column. After this purification step, the specific activity increased 7 times with 6.6% of quantitative recovery of the enzyme activity.

These active concentrated fractions were applied on CM Sepharose Fast flow ion exchange chromatography column (Supplementary Fig. S1A). After this purification step two proteins with Prx activity were recovered in the starting buffer. The presence of these two active proteins (anionic and cationic) was then confirmed by means of 2D electrophoresis analysis (Supplementary Fig. S2A) showing two spots with different pl but similar molecular weight (39 kDa by gel filtration analysis). After ion exchange chromatography the pooled fractions containing the cationic peroxidase were concentrated and applied on Sephacryl S-300 chromatography as last step of purification and CysPrx began to elute in Tris–HCl buffer at around 130 min. The results are shown in Supplementary Fig. S1B. Detailed results on purification procedure are summarized in Supplementary Table S1. The enzyme was purified 149 times, with a yield of 5.5%; the final specific activity of Prx enzyme for the H2O2-dependent oxidation of guaiacol was 2645 U mg⁻¹ of protein. Fraction obtained from the peak of Sephacryl S-300 generates a band in SDS-PAGE with an apparent molecular weight of about 39 kDa (Supplementary Fig. S2B). Gel filtration on Sephacryl S-300 was then performed to determine the native MW of the enzyme that was calculated to be 37.7 ± 1.5 kDa and was consistent with the SDS-PAGE results. The IEF of purified CysPrx resulted in a single band with a pl value of 8.3.

In a previous investigation we purified another Prx isoenzyme from artichoke leaves (ALSPr), which was found to have a MW of 51 kDa and the same pl [9]. The presence of more than one Prx enzyme in artichoke supports the hypothesis that a gene family gives rise to multiple isoenzymes belonging to class III Prx as in several other plant species [6,41]. In the majority of plant species, specific Prx isoenzymes can be constitutive or induced in response to external factors such as wounding, abiotic stress and attack by pathogens [44].

3.2. Substrate specificity, pH dependence of enzymatic activity, and thermostability

CysPrx had demonstrate to have high affinity for the phenolic compounds such as: chlorogenic and caffeic acids, present in artichoke [25], as was already reported by other authors [27].

CysPrx oxidized guaiacol very quickly but showed no detectable activity with ascorbic acid or veratryl alcohol. This result allows to attribute CysPrx to class III peroxidases. In addition CysPrx activity was tested in the pH range 3.5–9 using guaiacol as substrate (Supplementary Fig. S3A). The data show that the enzyme exhibited more than 50% activity in the pH range between 3.5 and 7 with its pH optimum at 5.5. This peroxidase could be considered a potential candidate for application in industry requiring a broader pH activity. Periodate/Schiff analysis showed that CysPrx is a glycoprotein but with a very low degree of glycosylation confirmed by the light staining of the gel (data not shown). The low degree of glycosylation of CysPrx could also explain the low stability of the purified enzyme at high temperature treatments compared to HRP. In fact, experimental evidence showed that after 10 min incubation at
different temperatures, the enzyme has high stability between 5 °C and 45 °C; instead, from 45 °C to 65 °C, there is an abrupt reduction of activity up to 5% of the original (Fig. S3B see supplementary material). Conversely, artichoke ALSP [9] and HRP (Supplementary Fig. S3B) are stable even above 65 °C, and this could be probably explained with the degree of glycosylation. It has been demonstrated that the carbohydrate portion plays a crucial role in determining the thermal stability of the protein [40]. Artichoke ALSP possesses a higher MW compared to CysPrx, possibly due to the presence of carbohydrates [9]. Plant peroxidases have eight potential N-glycosylation sites (Asn-X-Ser/Thr, where X could be any amino acid apart from Pro [15]), and the extent of N-glycosylation is highly variable, ranging from one for turnip Prx (TP7) and barley grain Prx (BP) [29,20] to the unusual eight sites for HRP [47].

3.3. N-Terminal sequencing and peptide mass spectrometry

Edman degradation of CysPrx did not provide any sequence suggesting that, like other classical plant peroxidases, the N-terminal is blocked by pyroglutamate [46]. In order to identify the sequence of the purified CysPrx LC-MS/MS analysis was carried out. After MS analysis, the identification of CysPrx as peroxidase was obtained by matching the peptides in the NCBInr database. The identified peptides matching peroxidases from different species were shown in Table 1. When the two amino acid sequences were deduced from cDNA sequences a PMF analysis using MALDI-TOF MS was performed to confirm the identified isoforms (see the following paragraph). As support to this hypothesis, the 2D electrophoresis showed a hazy spot in the zone corresponding to CysPrx (Supplementary Fig. S2A). Among the peptide sequences covered by PMF only one difference was found, the amino acid Arg instead of His. This situation is compatible with the presence of more than one protein in the same place with a very high degree of similarity.

3.4. Isolation of full-length cDNA encoding CysPrx and phylogenetic analysis

The artichoke peroxidase peptide sequences obtained from MS analysis (Table 2) were all used to blast (http://blast.ncbi.nlm.nih.gov/Blast.cgi) artichoke EST sequences available in the NCBI database. Numbers on the nodes indicate bootstrap values.

### Table 1

<table>
<thead>
<tr>
<th>Peptides</th>
<th>MW (Da) experimental masses MH⁺¹</th>
<th>Organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPNVFDNR</td>
<td>474.64</td>
<td>Catharanthus roseus</td>
</tr>
<tr>
<td>SAIPNLTR</td>
<td>984.36</td>
<td>X</td>
</tr>
<tr>
<td>YVVDLMNR</td>
<td>1073.50</td>
<td>X</td>
</tr>
<tr>
<td>YVVDLMNR (Met OX)</td>
<td>1089.485</td>
<td>X</td>
</tr>
<tr>
<td>AQAVIENLR</td>
<td>1273.68</td>
<td>X</td>
</tr>
<tr>
<td>AQAVIENLHR</td>
<td>1410.74</td>
<td>X</td>
</tr>
<tr>
<td>AQAVIENLR</td>
<td>1429.77</td>
<td>X</td>
</tr>
<tr>
<td>DPTMDQTFNRL</td>
<td>1598.71</td>
<td>X</td>
</tr>
<tr>
<td>DPTMDQTFNRL (Met OX)</td>
<td>1614.66</td>
<td>X</td>
</tr>
<tr>
<td>MSQSLVTGQGERI</td>
<td>1619.791</td>
<td>X</td>
</tr>
<tr>
<td>MSQSLVTGQGERI (Met OX)</td>
<td>1635.788</td>
<td>X</td>
</tr>
<tr>
<td>DAVVLGGPNYINPLGR</td>
<td>1741.91</td>
<td>X</td>
</tr>
<tr>
<td>QLEIFSDQQLYIDSR</td>
<td>1745.762</td>
<td>X</td>
</tr>
<tr>
<td>QLEIFSDQQLYIDSRQAGGLR</td>
<td>2059.112</td>
<td>X</td>
</tr>
<tr>
<td>LFPSRDPMTDQTFNNLR</td>
<td>2199.24</td>
<td>X</td>
</tr>
<tr>
<td>LFPSRDPMTDQTFNNLR (Met OX)</td>
<td>2215.12</td>
<td>X</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>Peptides</th>
<th>MW (Da) experimental masses MH⁺¹</th>
<th>CysPrx1</th>
<th>CysPrx2</th>
<th>MS/MS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPNVFDNR</td>
<td>948.44</td>
<td>X</td>
<td>X</td>
<td>+</td>
</tr>
<tr>
<td>SAIPNLTR</td>
<td>984.36</td>
<td>X</td>
<td>X</td>
<td>–</td>
</tr>
<tr>
<td>YVVDLMNR</td>
<td>1073.50</td>
<td>X</td>
<td>X</td>
<td>+</td>
</tr>
<tr>
<td>YVVDLMNR (Met OX)</td>
<td>1089.485</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>AQAVIENLR</td>
<td>1273.68</td>
<td>X</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>AQAVIENLHR</td>
<td>1410.74</td>
<td>X</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>AQAVIENLR</td>
<td>1429.77</td>
<td>X</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>DPTMDQTFNRL</td>
<td>1598.71</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>DPTMDQTFNRL (Met OX)</td>
<td>1614.66</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>MSQSLVTGQGERI</td>
<td>1619.791</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>MSQSLVTGQGERI (Met OX)</td>
<td>1635.788</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>DAVVLGGPNYINPLGR</td>
<td>1741.91</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>QLEIFSDQQLYIDSR</td>
<td>1745.762</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>QLEIFSDQQLYIDSRQAGGLR</td>
<td>2059.112</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>LFPSRDPMTDQTFNNLR</td>
<td>2199.24</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>LFPSRDPMTDQTFNNLR (Met OX)</td>
<td>2215.12</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

**Fig. 1.** Phylogenetic tree of type III peroxidase sequences. Protein sequences were retrieved from the Peroxidase Database and follow its nomenclature. Organism abbreviations are as follows: Pop: Populus; Vit: Vitis; Hel: Helianthus; Lac: Lactuca; Lin: Linum; Sol: Solanum; Nic: Nicotiana; Cap: Capsicum; Lyc: Lycopersicon; Cat: Catharanthus; Vig: Vigna; Pha: Phaseolus; Gly: Glycine; Lot: Lotus; Cyn: Cynara; Tar: Taraxacum; Spi: Spinacia; Ara: Arabidopsis; Bra: Brassica; Sor: Sorghum; Pic: Picea; Pin: Pinus.
The best matches were found with ESTs GE580188 and GE592500. These two ESTs were translated in silico into proteins, and blasted against the NCBI protein sequence database which provided high levels of identity to other plant peroxidases. On the basis of the two detected artichoke ESTs, two primers were designed and used to amplify artichoke leaf cDNA. A fragment of 585 bp was obtained, which was sequenced and revealed a high level of similarity to other plant Prx sequences. In order to obtain the full-length CysPrx cDNA, this partial sequence was extended towards the 5' and 3' ends by RACE-PCR, and new sets of primers were subsequently designed around the start and downstream the stop codons. After cloning and sequencing, two slightly different kind of sequences were obtained from a total of seven clones: a group of six clones were identical, while one possessed three substitutions causing the changing of three amino acids, in position 120 (Arg instead of His), in position 242 (Ala instead of Val), and in position 348 (Glu instead of Gly). Both sequences were 1179 bp and were composed of a 1059 bp open reading frame, encoding a 353-residue protein, with an estimated molecular weight of 38.46 and 38.53 kDa respectively. The identity between the two translated sequences was 99.2%. The sequences were submitted to the NCBI database under the accession numbers GU145300 (CysPrx1) and GU145301 (CysPrx2).

A Blast analysis against the Prx database revealed that CysPrx1 and CysPrx2 have 86% identity with lettuce LsPrx04, 82% with sunflower HaPrx07 [23], 71% with Catharanthus roseus CroPrx03 [3] and tobacco NtPrx15 [14], 66% with flax LuPrx03 [31]. Conversely, a very low degree of similarity was observed with Class I ascorbate Prx Cycas CruAPx02 [8] and Class II lignin Prx Basidiomycota TvLiP07 [21] (below 20% each), providing evidence that our enzyme is related to the Class III Prxs.

A phylogenetic analysis was performed with the two artichoke sequences together with a number of type III plant Prxs, showing at least 60% identity, and including sequences from species belonging to the Asteraceae family. As out-groups, Prx from monocotyledonous (maize, sorghum) and gymnosperm (pine, spruce) species were considered.

The phylogenetic tree (Fig. 1) showed that artichoke Prx sequences cluster with other sequences from species belonging to the Asteraceae family (Lactuca spp., Helianthus annuus, Taraxacum officinale). In general, Prx sequences from species of the same family are grouped together. It is interesting to notice that for the Asteraceae family, for instance, three separate clusters are formed, and for Solanaceae (Solanum, Lycopersicum, Nicotiana, Capsicum), two separate groups can be observed. Sequences from different species clustering together might represent orthologs, i.e., homologous genes in different species that diverged from a single ancestral gene after a speciation event. Orthologs normally share functional characteristics. On the other hand, within the same species, Prx proteins may be encoded by paralogous genes, arisen from duplication events, which might be devoted to different functions [28].

As expected, Prx from maize and sorghum on one side, and pine and spruce on the other, were the most distantly related sequences. The sequence alignment of CysPrx1 and CysPrx2 with the same type III peroxidase sequences used in the phylogenetic tree is presented in Supplementary Fig. S4. The active site is highlighted in gray.

To confirm the hypothesis that the two deduced protein sequences, CysPrx1 and CysPrx2, are the same proteins identified by MALDI-TOF analysis, PMF was carried out against both sequences (see previous paragraph). The results obtained showed that coverage of 35% was achieved for both isoforms (Fig. 2). In addition, the identified peptides of the purified enzyme matched perfectly with CysPrx1 and CysPrx2 (Table 2), and the only difference found in MS analysis was the amino acid Arg instead of His in one peptide (Table 2 in bold), which could also be found in the proteins deduced from cDNA sequences. No further data of
coverage by PMF method are available. All the results lead to the conclusion that the purified protein overlap perfectly with the deduced sequences from cDNA. CysPrx was different even from ALSP, in fact beyond the molecular weight and the substrate specificity, in the sequenced peptides of ALSP there was another difference with CysPrx: the presence of the aminoacid Val instead of Thr[9].

The availability of the whole CysPrx1 and CysPrx2 coding sequences will allow investigating the precise physiological role of these genes. Some plant peroxidase genes have been expressed transgenically and the phenotypes of these transgenic plants yield information about their functions[50]. It has been reported that over expression of the tomato TPX2 gene or the sweet potato swpa1 gene conferred increased salt-tolerance or oxidative-stress tolerance, respectively[2,51].

Moreover, a better understanding of Prx role and function might help develop new methods to increase the efficiency of plant peroxidase in several biotechnological processes.

3.5. Structure prediction

To improve the knowledge of CysPrx structure, also with the aim of possibly increasing the stability and catalytic efficiency of CysPrx, a structure prediction and comparative model were carried out. Negligible differences were found between the predicted structures of the two deduced proteins, therefore only the CysPrx1 structure was considered for further analysis. Both I-TASSER and Phyre servers chose the same set of representative crystal structure, reported in Table 3 with their sequence identities with CysPrx1. The best models they generated, denoted CP1-ModTas (red in Fig. 3A and B) and CP1-ModPhy (blue in Fig. 3A and B) respectively, are compared with BP (PDB code 1BGP, cyan in Fig. 3A and B) and HRP (PDB code 1ATJ, magenta in Fig. 3A and B) crystal structures. Fig. 3A shows the two models superimposed to 1BGP and 1ATJ structures where only the common domain, consisting of residues 37-334 of the CysPrx1 sequence is reported. The overall structure appears well conserved, with the central heme group (from 1BGP and 1ATJ) sandwiched between the distal N-terminal domain and the proximal C-terminal domain. The backbone of CP1-ModPhy strictly follows that of 1BGP (the cyan structure is almost completely hidden by the blue one), apart from the region highlighted by the blue arrow in Fig. 3A, comprising the residues 126–129. The backbone of CP1-ModTas, instead, seems to follow more strictly that of 1BGP in the proximal domain and that of 1ATJ in the distal

<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>Id (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1BGP</td>
<td>Barley Grain Peroxidase</td>
<td>53</td>
</tr>
<tr>
<td>1PA2</td>
<td>Arabidopsis Thaliana Peroxidase</td>
<td>42</td>
</tr>
<tr>
<td>1ATJ</td>
<td>Horseradish Peroxidase C1A</td>
<td>41</td>
</tr>
<tr>
<td>1FHF</td>
<td>Soybean Peroxidase</td>
<td>43</td>
</tr>
<tr>
<td>1SCH</td>
<td>Peanut Peroxidase</td>
<td>42</td>
</tr>
</tbody>
</table>
domain. The largest deviations from 1BGP occur for the loop residues 93–113, highlighted by the red arrow in Fig. 3A. The comparison among the heme pocket residues is shown in Fig. 3B, where the same colour code of Fig. 3A has been used and the labels referring to the CysPrx1 sequence. It can be noted that the side chain of CP1-ModPhy overlaps with that of 1BGP, apart for Met252, the aligned residue of which is Leu224 for 1BGP and Phe221 for 1ATJ. The side chain of CP1-ModPha reveals peculiar features which deserve further discussion. Its proximal histidine (His207) and catalytic residues Arg74 and Phe77 are differently orientated with respect to the corresponding residues of 1BGP and 1ATJ, despite the orientation of the hydrogen-bonded invariant proximal aspartate (Asp278) is conserved in all the considered structures and the three residues are incompatible with the position of heme in 1BGP and 1ATJ. Moreover, the orientation of its distal histidine His78 is closer to that of 1ATJ than to that of 1BGP. In fact, the distal histidine N2-Fe distance is 6.1 Å in CP1-ModTas and 5.9 Å in 1ATJ, while it is 8.2 Å in 1BGP and CP1-ModPhy. This is extremely important, since BP, contrary to the other plant peroxidases listed in Table 3, is known to be inactive for pH values greater than 5. The 1BGP structure refers to the inactive form of BP and this behaviour has been ascribed to the altered orientation of the distal histidine, which results to be located too far from the heme iron atom to be able to catalyze the formation of compound I [43]. In addition, CP1-ModTas reproduces 1ATJ also in the hydrogen bond between O31 of Asn106 and N11H of His78, while CP1-ModPhy follows 1BGP in disrupting this hydrogen bond. In 1BGP the closest residue to the distal His is a Thr, which lies on the Ser96 of CP1-ModPhy. This different behaviour of CP1-ModTas with respect to CP1-ModPhy is related to the above mentioned main chain deviations of residues 93–113.

Short (50 ps) molecular dynamics simulations have been performed on CP1-ModPhy and CP1-ModTas to equilibrate their structure in presence of the heme group (the initial coordinates of which have been taken from 1BGP). The analysis of the heme pocket residues reveals interesting features: the orientation of His78 remains stable for CP1-ModPha (the root mean square deviation of their atoms between its initial and final configuration is RMSD = 0.9 Å), but not for CP1-ModPhy (RMSD = 2.3 Å) and this latter converges towards the configuration exhibited by CP1-ModTas and 1ATJ. The contrary occurs for Arg74 and His207, for which the CP1-ModPhy configurations are confirmed (RMSD = 1.3 Å and 1.0 Å respectively), while the CP1-ModTas configurations, which were incompatible with the heme orientation, drift towards those exhibited by CP1-ModPhy and 1BGP (RMSD = 2.0 Å and 1.3 Å respectively).

In summary, the CysPrx1 structure is predicted to follow the BP backbone in the proximal domain and to assume the configuration of the reaction pocket typical of HRP and other active plant peroxidases, where the major role is played by the orientation of the distal His residue. In this respect, the predictions diverge from the inactive crystal structure of BP, even if BP has a higher sequence homology with CysPrx1. Peculiar properties could arise for CysPrx1 due to the presence of Met252 in its catalytic site: the presence of a sulphur atom close to the heme pocket might affect the stability of the heme group by stabilisation/delocalisation of the porphyrin-cationic proximal aspartate interplay [18] and could differentiate the catalytic properties of CysPrx from those of BP and HRP.

4. Conclusion

CysPrx was purified by using a combined purification protocol involving two phase system, ion exchange and gel filtration chromatography. Systematic biochemical characterization of the enzyme (optimum of pH and temperature, substrate specificity) demonstrated that CysPrx possesses interesting properties, such as stability at a wide pH range, suggesting that CysPrx could be a potential candidate for application in industry requiring large pH stability.

MALDI-TOF MS analyses resolved the CysPrx isoenzyme into two cationic isoforms differing for one aminoadic. The presence of these isoforms was confirmed by the isolation of full-length cDNAs encoding CysPrx that generate two slightly different sequences coding for two putative CysPrx: CysPrx1 and CysPrx2 sharing a very high degree of similarity with sunflower and lettuce Prx. Finally a three-dimensional model was predicted from CysPrx1 by homology modeling, by using different computational tools. The study of three-dimensional model, the roles for particular amino acid residues and structural motifs or regions of these isoenzymes will offer new insights into the relationship between plant peroxidases and their physiological substrates. In addition, the obtained information could open new methods for improving the stability of the enzyme that might be involved in several processes of biotechnological interest for peroxidase application.

Acknowledgements

The authors wish to thank Dr Graziano Lamba for his useful suggestions during the preparation of the manuscript and Anita Morgese and Giuseppe Sonnante for their excellent technical assistance.

Appendix. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.plaphy.2011.01.028.

References
