Predicting the functionally distinct residues in the heme, cation, and substrate-binding sites of peroxidase from stress-tolerant mangrove specie, *Avicennia marina*

Uzma Jabeen · Atiya Abbasi · Asmat Salim

Received: 9 January 2011 / Revised: 25 April 2011 / Accepted: 10 May 2011 / Published online: 10 June 2011
© Cell Stress Society International 2011

Abstract Recent work was conducted to predict the structure of functionally distinct regions of *Avicennia marina* peroxidase (AP) by using the structural coordinates of barley grains peroxidase as the template. This enzyme is utilized by all living organisms in many biosynthetic or degradable processes and in defense against oxidative stress. The homology model showed some distinct structural changes in the heme, calcium, and substrate-binding regions. Val53 was found to be an important coordinating residue between distal calcium ion and the distal heme site while Ser176 is coordinated to the proximal histidine through Ala174 and Leu172. Different ionic and hydrogen-bonded interactions were also observed in AP. Analyses of various substrate–enzyme interactions revealed that the substrate-binding pocket is provided by the residues, His41, Phe70, Gly71, Asp138, His139, and Lys176; the later three residues are not conserved in the peroxidase family. We have also performed structural comparison of the *A. marina* peroxidase with that of two class III salt-sensitive species, peanut and soybean. Four loop regions were found to have largest structural deviation. The overall protein sequence was also analyzed for the presence of probable post-translational modification sites and the functional significance of these sites were outlined.

Keywords Homology modeling · Structure prediction · Peroxidase · *Avicennia marina* · Substrate–enzyme interactions

Abbreviations

AP	<i>Avicennia marina</i> peroxidase
BGP	Barley grain peroxidase
PNP	Peanut peroxidase
SBP	Soybean peroxidase
HRP	Horseradish peroxidase
BHA	Benzhydroxamate
FCN	Ferulic acid-cyanide
FA	Formic acid
CO	Carbon monoxide
TRIS	Tris (hydroxy methyl) aminomethane

Introduction

Plant peroxidase superfamily consists of a characteristic central heme group which is sandwiched between a proximal and a distal domain (Henriksen et al. 1998a). Heme peroxidases catalyze the one-electron oxidation of a variety of organic substrates. Most heme peroxidases undergo the following multi-step reaction scheme (Sundaramoorthy et al. 1994).

[P=porphyrin and S=substrate]

U. Jabeen · A. Abbasi
H.E.J. Research Institute of Chemistry, University of Karachi,
Karachi 75270, Pakistan

A. Salim (✉)
Dr. Panjwani Center for Molecular Medicine and Drug Research,
International Center for Chemical and Biological Sciences,
University of Karachi,
Karachi 75270, Pakistan
e-mail: asmat.salim@iccs.edu

Plant peroxidases are grouped into three classes. Class I peroxidases are of prokaryotic origin which include yeast cytochrome c peroxidase, a soluble protein found in the mitochondrial electron transport chain, where it protects against toxic peroxides; ascorbate peroxidase, the main enzyme responsible for hydrogen peroxide removal in the chloroplasts and cytosol of higher plants (Dalton 1991); and prokaryotic catalase peroxidases that exhibit both peroxidase and catalase activities and provide protection to cells under oxidative stress (Welinder 1991). Class II peroxidases consists of secretory fungal peroxidases, e.g., lignin peroxidases, which are monomeric glycoproteins involved in the degradation of lignin and manganese-dependent peroxidases in which Mn^{2+} serves as the reducing substrate (Reddy and D'Souza 1994). Class II proteins contain four conserved disulfide bridges and two conserved calcium-binding sites. Class III peroxidases consist of the secretory plant peroxidases and are targeted via the endoplasmic reticulum to the outside of the plant cell or to the vacuole. They play a vital role in a wide range of physiological processes, for example in the metabolism of reactive oxygen and nitrogen species, cross-linking of cell wall components, lignin and suberin formation, and synthesis of phytoalexins. Peroxidases perform their defensive role against several biotic and abiotic stresses by trapping of heavy metals and degradation of toxic molecules (Agostini et al. 2003; Lavid et al. 2001). During plant defense, peroxidases directly or indirectly participate in the signaling pathways, which regulate peroxidase gene expression (Almagro et al. 2009). The most well known members of this class of peroxidases are horseradish peroxidase (HRP) (Gajhede et al. 1997), peanut peroxidase (PNP) (Schullar et al. 1996), soybean peroxidase (SBP) (Henriksen et al. 2001), barely grain peroxidase (Henriksen et al. 1998a) and *Avicennia marina* peroxidase.

Due to their versatile nature of peroxidases, the presence of a large number of isoenzymes have been reported in a single plant such as *Arabidopsis*, 73; rice, 138; and tobacco, 12 (Passardi et al. 2004; Welinder et al. 2002). The peroxidase database, PeroxiBase (<http://peroxibase.toulouse.inra.fr/>), has revealed 143 sequences of distinct class III peroxidases in maize (*Zea mays* L.). Recently, phylogenetic analysis of two new class III peroxidase genes, CrPrx3 and CrPrx4, from *Catharanthus roseus* has been performed. The analysis revealed that both genes have diverse expression patterns in a variety of plant tissues. Quantitative real-time PCR confirmed their maximum expression in stem tissues followed by flower tissue. The expression levels of CrPrx3 and CrPrx4 were down-regulated under salt and dehydration stress, respectively (Kumar et al. 2011).

A number of plant peroxidases have been either structurally elucidated or their 3D structures were predicted (Koua et al. 2009; Passardi et al. 2007). They consist of highly conserved domains as well as variable regions. Multigene families of peroxidases have catalyzed the diversity of processes indicating that each isoform may have potential in functional specialization (Cosio and Dunand 2009). Invariant amino acid residues necessary for catalytic activity and proper folding are present in all plant peroxidases (Welinder et al. 2002). Heme in these peroxidases is protoporphyrin IX. The proximal histidine is the fifth ligand of the heme iron whereas the other histidine is known as the distal histidine. This class of peroxidase also consists of two calcium-binding sites, proximal and distal to the heme. Heme is not covalently linked to the protein matrix and is thus possibly more sensitive to react to altered nonbonded interactions in its vicinity (Laberge et al. 2003). Various substrates, including carbon monoxide (CO), cyanide, azide, and fluoride, as well as formate and acetate coordinate with the heme (Fe) at the distal catalytic site and form six coordinated peroxidase complexes (Carlsson et al. 2005; Veitch 2004). These ligands present in their protonated forms are stabilized through hydrogen-bonded interactions with the distal catalytic residues. Binding of such ligands is mutually exclusive and fully competitive toward peroxide binding (Carlsson et al. 2005). Class III peroxidases recognize a number of vacuolar metabolites, such as phenols, flavonoids, and alkaloids as substrates, and assume specific functions in the metabolism of these compounds in vivo (Costa et al. 2008; Passardi et al. 2004; Welinder et al. 2002).

In the present study, we have predicted the 3D homology model of peroxidase from salt-tolerant specie, *A. marina* (AP). The model was built with heme and calcium ions and analyzed as regards the active site residues, heme, and cation-binding sites, which play important roles in catalysis. The amino acid sequences of salt-tolerant and salt-sensitive plant species were also aligned and conserved and nonconserved amino acids were analyzed. The 3D model of AP was compared with peroxidase structures of salt-sensitive and salt-tolerant plant species and regions having structural deviations were analyzed. The 3D models of AP were also constructed with various substrates, i.e., benzhydroxamate (BHA), ferulic acid-cyanide (FCN), formic acid (FA), carbon monoxide (CO), and tris (hydroxy methyl) amino-methane (TRIS). The structural information obtained from the models was compared with those of the previously determined crystal structures. The 3D models were studied with respect to the residues which play important role in catalysis.

Methods

Sequence analysis

Pairwise sequence alignment

Primary sequence of peroxidase from *A. marina* (Q9FXL6_AVIMR) was retrieved from Uniprot database (Bairoch et al. 2005). The sequence was submitted to BLAST (Altschul et al. 1997) search against protein databank, PDB (Berman et al. 2002) in order to identify the template for molecular modeling. The search gave highest homology (36%) with barley grain peroxidase, pdb id: 1bgp (Henriksen et al. 1998a). Residues 31–328 of the target sequence were aligned with residues 8–305 of 1bgp by the BLAST program. Some manual adjustments were done where needed.

Multiple sequence alignment and phylogenetic analysis

Amino acid sequences of peroxidase belonging to salt-tolerant and salt-sensitive plants were retrieved from Uniprot database (Bairoch et al. 2005). Homology sequence searches were performed using BLAST (Altschul et al. 1997) against non-redundant (nr) database and class III PeroxiBase (Koua et al. 2009). Multiple sequence alignment of plant peroxidases was performed using Clustal X program (Thompson et al. 1997). Phylogenetic trees were constructed by using neighbor-joining method using Drawtree and Drawgram programs of PHYLIP tree package version 367 (Felsenstein 1989; Howe et al. 2002).

Prediction of post-translational modification sites

Primary sequence of *A. marina* was searched at PROSITE database using ScanProsite (Obenauer et al. 2003) for the possible post-translational modification sites.

Model building

The three-dimensional homology models of *A. marina* peroxidase with heme, and two calcium ions were constructed separately with and without different substrates. The substrates used were ferulic acid-cyanide, benzhydroxamic acid, formic acid, carbon monoxide, and tris (hydroxymethyl) aminomethane. The models were built using the structural coordinates of 1bgp, 7atj, 2atj, 1w4w, 1w4y, and 1fhf, respectively, by the protein structure-modeling program MODELLER 9v3 (Sali and Blundell 1993). For the model without the substrate, 1bgp was used as the template. The program was executed using the default parameters. Protein structures were visualized and analyzed using Swiss PDB

Viewer (SPDV) 3.7 (Guex and Peitsch 1997) and DS viewer (<http://accelrys.com/products/discovery-studio/>).

Model assessment

Evaluation of the predicted homology models was based on the analyses of stereochemistry, geometry, and energy distributions. The consistency of the models was assessed by using the ENERGY command of the MODELLER. Assessment was also performed by the programs, PROCHECK (Laskowski et al. 1993) and protein structure analysis (ProSa) (Sippl 1993; Wiederstein and Sippl 2007). The variability among the models was compared by superposition of the C α traces of the model and the template and the root mean square deviation (RMSD) between the equivalent atoms was determined.

Structural comparison of *A. marina* peroxidase with that of salt-sensitive and salt-tolerant plants

The three-dimensional homology model of *A. marina* peroxidase was compared with the 3D structures of salt-sensitive (peanut and soybean) and salt-tolerant (barley grain) plant species by means of structural superpositions using the program, DS viewer (<http://accelrys.com/products/discovery-studio/>). The regions of highest structural deviations were then analyzed.

Results and discussion

Multiple sequence alignment

Table 1 shows the multiple sequence alignment of *A. marina* peroxidase with the members of peroxidase family. Sequences identified by the BLAST server against non-redundant (nr) database were used for multiple alignment using Clustal X program (Thompson et al. 1997). Multiple sequence alignment has shown that the distal (Arg45, His49) and proximal catalytic (His175) residues of heme and co-essential residue Phe48 are highly conserved in the family of class III heme peroxidase. Arg45 is a transition state stabilizer and provides access of the substrate to the catalytic site while His49 is a proton acceptor. Asn46 is present in 12 members of the peroxidase. It is replaced by Leu in more than 70% of the sequences. The conserved fold, ¹³F/YYXXT/SCP¹⁹, is present at the distal site with yet unknown function. Another conserved motif, ⁵⁰DCFVQ/NGCDGSVLI/L⁶², is present in most of the family of peroxidases. These residues provide distal site coordination with Ca in peroxidases of *A. marina*, *Nicotiana tabacum*, *Vinca rosea*, *Arabidopsis thaliana*,

Table 1 CLUSTAL X multiple sequence alignment showing the sequence of members of the peroxidase family

Q9FXL6_AVIMR	-----FAEN-----EGHGLAMNYRDS	19
B2G335_CATRO	-MVFMSFSSSSSSTLLFLISSLLISTHFN-VHIEAQTTPPIVSGLSFTFYDSSCP	56
B2G334_CATRO	--MALIAFSSSTS---LLLICSLLLVSTHFN-PHIEAQTTPPIVSGLSFAFYNSTCP	51
Q0VYC8_CATRO	-----MAFSSSTS---LLLLLISLLISAHFNHVAQTTRPPTVSGLSYTFHNSRCP	51
Q6PQF2_EUPCH	-----MAS---KLVLVSCLLVAFWFC--AIEAQTTPPIVNGLSWTFYKSSCP	42
P93545_SPIOL	-----MMVSFKIIPMLLVVAISCLVGIS---KSQST-VPVNVGLSYSYYSRSCP	45
P93547_SPIOL	-----MAATKGFPILILVLS-SLIIGLS---QGQST-IPVVPGLSYTFYSSSCP	43
Q5W514_PICAB	-----MEKLMILCLFCTLIWACASREN---VLTLSNDPPLVNGLSWTFYKSSCP	46
Q5W512_PICAB	-----MAKHITPLASVILCMFVIYGRA---VHSLPTPVAGLSWTFYSTSCP	43
Q40069_HORVU	-----MARVPLLAALVVAMAVLVASSLG---PRASAAEPPVAPGLSFDIFYRTCP	47
Q49866_HORVU	-----MASRAAAAIAVLALVCAAVHSS---EGQLSPNFHAATCP	36
Q58A85_ZINEL	-----MAIIS-ISPFSAFILILTIVIPFS---EAGLSPHYHQTCP	36
Q9ZP15_TRIRP	-----MGRFNNVILVWSLLMLCFIPYTT---FAQLSPNHYANTCP	38
PERN1_TOBAC	-----MEYYHHSINKMAMFMVILVLADVT---MVLGQGTTRVGFYSSTCP	42
Q4ADU9_POPAL	-----MSQKVLMFLLVAMAGTA---TVQGGTTRVGFYATTCR	35
Q4A3Z3_SENSQ	-----MEVYSHIRTPIIILFVVIFAAL---TSLALGCKVGFYQATCP	38
Q4A3Z1_SENSQ	-----MEVYSHLRTPIIILFVVVFAAL---TSLALGCKVGFYQATCP	38
Q4A3Z2_SENSQ	-----MEVYSHLRTPIIILFVVVFAAL---TSLALGCKVGFYQATCP	38
Q4A3Y8_SENSQ	-----MEVYSHLRTPIIILFVVVFAAL---TSLALGCKVGFYQATCP	38
P93551_SPIOL	-----LANSAK---SOLS---IAYYASSCP	19
PER13_ARATH	-----MITIALFLVLLYFHDQL---GYSAAQLQFGFYSETCP	34
PER39_ARATH	-----MTRFGLALLMLVIQGLV---TFSEAQLKMGFYDQTC	35
Q5I3F0_TRIMO	-----MASSGGLAGLAVLLTAVLCLQLP---VHSRAQLRVGFYNTTCP	40
A1YBM0_HELAN	-----	
P93546_SPIOL	-----MRISLFSMMFLGLILVTLV---GHCYGQLKVGFFYEGKCGG	38
PER2_ARMRU	-----MHSSSS---LIKGLFLLLLNVS---LSHAQLSPSFYDKTCP	36
PER1B_ARMRU	-----MHSPSSTSFTWILITLGLAFYAS---LSDAQLTPTFYDTSCP	40
PER3_ARMRU	-----MGFSPLISCSAMGALILSCLLLQAS---NSNAQLRPFDFYRFTCP	41
PERE5_ARMRU	-----QLRDPFYRSTCP	12
Q43099_POPTR	-----MHLSKAIVAAFFVVLGG---TLAHGQLTPTFYDQTC	36
Q43101_POPTR	-----MHLSKAIVAAFFVVLGG---TLAHGQLTPTFYDRTCP	36
Q43050_POPKI	-----MKLSKLMVVALFYAFLVGG---PLAYGQLTPTFYDET	36
Q43102_POPTR	-----MIDKALHPLVASLFFVIFWGGSL---PYAYAQLTPTFYDGTCP	40
Q42905_LINUS	-----MASLFLNKFFVSSVLAMLLFAG---ASYAQLTPTFYDSTCP	39
PERA2_ARMRU	-----QLNATFYSGTCP	12
PERX_TOBAC	-----MSFLRFVGAIFLVAIFGASN---AQLSATFYDTTCP	34
Q40555_TOBAC	-----MIPALMLQVLYVVLWIKGN---VLMLELVKLFLV	32
O65773_CUCPE	-----MALLFFLALLFGAS---YAQLTETFYDQTC	28
PER2_ZINEL	-----MSYHKSSGTTLMVPLFMLLISVNYF---MSCNAQLSTTFYDTTCP	42
PER1_ZINEL	-----MSYHKSSGTTLMVPLFMLLISVNYF---MSCNAQLSTTFYDTTCP	42
PERX_NICSY	-----MNTPTQSFRAKAAIFSLLLSC---MQCHAQLSATFYDNTCP	39
PER1_HORVU	-----MASSSYTSLVLVLAIVTAAS---AQLSPTFYDTSCP	33
Q5I3F6_TRIMO	-----MAASASCSISLVLVALATAAS---GQLSPTFYDTSCP	34
Q5I3F7_TRIMO	-----MAASASCSISLVLVALATVAS---AQLSPTFYDTSCP	34
Q5I3F5_TRIMO	-----MAMGSASRISLVLVALATAAS---GQLSSTFYDTSCP	35
Q5I3F4_TRIMO	-----MASSS-LSAVLLCLAVAAS---AQLSPTFYDTTCP	32
Q9XIV9_TOBAC	-----MASL-KINAIVLFI-LVSLIGS---SSA-QLS-TGFYKSSCP	36
PER4_VITVI	-----MASS-SFSIVVALGVLFALFAGS---SSA-QLS-TNFYSKTCP	37
Q4W2V5_PICAB	-----MKT-LACIVLCVLCISSIN-NA---AHG-QLT-STFYNKLC	35
Q4W2V2_PICAB	-----MMMRT-LLCIGVMMAVFCSSININA---VSG-QLS-STFYDKSCP	37
Q5I3F2_TRIMO	-----MASS-SGCRRAWHCLLALFLLSSS---AYG-QLS-PSFYAKSCP	38
P93548_SPIOL	-----MASN-LVVGFIAIFSIILLAGT---SDA-WLRKPHFYASSCP	38
Q40372_MEDTR	-----MASS-SPCQIFLVFVMVTLVTSL---IPSNALLTPHFYDNVCP	39
A0SWU6_SESRO	-----MASSGYLSVFFHAFVFASLATG---FS---QLSPNYDYTCP	37
P93549_SPIOL	-----GKLLFFFQILFGLSL---TVNGQLS-PNFYSSTCP	32
Q9MAX9_ASPOF	-----MALSMN-LLFLVLIISLSLAHLC---FADG-SLTPQFYDHS	38
Q8H285_ANACO	-----MALSMS-GIVAVLMV-LSLAPLS---LGGG-YLCPYDYDKSCP	37
Q5W513_PICAB	MARIKHSPLTLQFQSVLITAVALLMLWIQTL	49
PERN_ARMRU	-----MKTQTKVMGGHVLLTVFTLCMLC---SAVRAQLSPDIYAKSCP	40
Q5I3F1_TRIMO	-----MAPPRVHLLALLALCLAFARAD---AAVR-ELRVGYAATCP	38
Q9XFL2_TOBAC	-----MGFKVFFFAI---LFFSAVS---AFAE---DNSGLVMDYYKDS	37
A3KLN6_CATRO	-----MASKTLFLFLVI---LSFSALS---TFAENEA-EADPGLVMNYYKDS	41
Q08671_GOSHI	-----MRKIMGAKVLFFAL---LSFSAVS---AFAQD---EEDQGLVMNYYKDS	43
B9RXL3_RICCO	-----MAAKALFFFAI---LFFSAVSFRPAFAEEN---EEDPGLVMNYYKDT	43
PER42_ARATH	-----MGKGVMVAI---LCLWALS---ATSEAVT-EAEPGLMMNYYKDT	41
A9XN55_THEHA	-----MGKGAMMVAI---LCLWALS---ATSEAVTEEPGLMMNYYKDT	42
B7FHN5_MEDTR	-----MSPNKAFLFLAL---LSFSPQLFFIVSSAA---EDNGLVMNYYKEA	42
Q9FT05_CICAR	-----MAPNHAFILFVL---LSFSPQLFFTLSSAQ---QDNGLLMNYYKES	42
A9NNP1_PICSI	-----MGRWVLLALALSQALLQSQAYNVGSSNTNGSYGENGLVMNYYGDS	37
Q9FXL6_AVIMR	-QAEIIKEQVQLLYKRHKNTAFSWLRNIFHDCFVESCDA	77
B2G335_CATRO	-DLESIIIRNRLRVFRNDIQAAAGLLRLHFHDCFVQCGDGSVLLVGSASGPGEQDAPPNL	115
B2G334_CATRO	-DLESIIIRNRLREVFNQNDIEQAAAGLLRLHFHDCFVQCGDGSVLLVGSASGPGEQDAPPNL	110
Q0VYC8_CATRO	-DLKSIVRNRLREVFNQNDIEQAAAGLLRLHFHDCFVQCGDGSVLLVGSASGPGEQDAPPNL	110

Table 1 (continued)

Q6PQF2_EUPCH	-KVESIIQKELKKLFKKDVEQAAGLLRLHFHDCFVLGCDGSVLLNGSAGGPSEQSELPNL	101
P93545_SPIOL	-DLDFIIRDLHFDVFERDITQAAGLLRLHFHDCFVKGCDGSVFLVGSSSTPSEKDAPPNL	104
P93547_SPIOL	-GLDFIIRGHLWQIFQSDLTQAAGLLRLHFHDCFVQGGCDGSVLLDGSAGGPSEKAPPNL	102
Q5W514_PICAB	-KLESIVKQRIDFYLYKQDITQAAGLLRLHFHDCFVQGGCDGSVLLAGSTSGPSEQGAPPNL	105
Q5W512_PICAB	-SLESIVRQRMGAYLSADITQAAGLLRLHFHDCFVQGGCDGSVLLN-STSG--EQTTPPNL	99
Q40069_HORVU	-RAESIVREFVQEAVRKDIGLAAGLLRLHFHDCFVQGGCDASVLLDGSATGPGEQQAPPNL	106
Q49866_HORVU	-DLERIVEFHVAETFRRDVGVPALIRILFHDCFPQGGCDASVLLKGAG---SELNEIPNQ	92
Q58A85_ZINEL	-QAEDIIIFQTVRNASIYDPKVPARLLRMFHDHCFIRGCDASLLLDSTPANKAEKDGPPNI	95
Q9ZP15_TRIRP	-NVQSIVRSAVQKKFQQTFTVTPATLRLFFHDCFVQGGCDASVLVASSGGNQAEKDNPDNL	97
PERN1_TOBAC	-RAESIVQSTVRAHFQSDPTVAPGILRMHFHDCFVLGCDGSILIE---GSDAERTAI PN	97
Q4ADU9_POPAL	-RAESIVRATVQSHFTSDSSIAPGLLRMHFHDHCFVNGCDASILLID---GANTEKTAGPN	90
Q4A323_SENSQ	-RAESIVQSVVKS AIRSNPTYAPGILRLFFHDCFVNGCDASVLLD---GSTSEQTASTN	93
Q4A321_SENSQ	-RAETIVQSVVKS AIRSNPTYAPGILRLFFHDCFVNGCDASVLLD---GSTSEQTASTN	93
Q4A322_SENSQ	-KAETIVQSVVKS AIRTNPTYAPGILRLFFHDCFVNGCDASVLLD---GSTSEQTASTN	93
Q4A3Y8_SENSQ	-KAETIVQSVVKS AIRTNPTYAPGILRLFFHDCFVNGCDASVLLD---GSTSEQTASTN	93
P93551_SPIOL	-QAEIVRSTVQSHFNSDPTIAPGLLRHLFHDCFVQGGCDASILIS---GTSSERTAF TN	74
PER13_ ARATH	-SAESIVRDVVQAVTNDPGKAAVLLRLQFHDHCFVEGCDGSILIKHG---GNDDERFAAGN	91
PER39_ ARATH	-YAEKIVQDVVNQHINNAPS LAGSLIRMHFHDHCFVRCGDSILINATSSNGPGEQAPPNL	94
Q5I3F0_TRIMO	-NAEAIVRQAVTAATNPVGAAGLIRLHFHDCFVEGCDSSVLLSVNPGGGTTEREAPN	99
A1YBM0_HELAN	-----RLHFHDCFVNGCDASILLD---GPNSEKTAPPN	30
P93546_SPIOL	NNVENVIYNVVKQKIKADPDTVSDLVVSVFHDHCFVRCGDSIFLD---GANAEKSAVPN	94
PER2_ARMRU	-QVFDIATNTIKTALRSDPRIAASILRLHFHDCFVNGCDASILLDNTTSFRTEKDAFGN-	94
PER1B_ARMRU	-NVSNIVRDIIINELRSDPRITASILRLHFHDCFVNGCDASILLDNTTSF LTEKDALGN-	98
PER3_ARMRU	-SVFNIIIGDIIIVDELRTDPRIAASLIRLHFHDCFVRCGCDASILLDNTSFRTEKDAAPN-	99
PERE5_ARMRU	-SVFNIIKKNVIVDELQTDPRIAASILRLHFHDCFVRCGCDASILLDTSKSFRTKDAAPN-	70
Q43099_POPTR	-NVSSIIIRNVITETLVCDRRIGGSLIRLHFHDCFVNGCDGSLLLDNTDPTIESEKBAAGN-	94
Q43101_POPTR	-NVSSIIISNVITETLVCDRRIGGSLIRLHFHDCFVNGCDGSLLLDNTDPTIESEKBAAGN-	94
Q43050_POPKI	-YVISIIRGVIAETLIFDPRIGASLIRLHFHDCFVNGCDGSILLDKTATIDTEKEALAN-	94
Q43102_POPTR	-NVSTIIRGVLAQALQTDPRIGASLIRLHFHDCFVNGCDGSILLDNTDPTIESEKBAAPN-	98
Q42905_LINUS	-NVIGIVRTVLQNAAMADPRIGASLIRLHFHDCFVNGCDGSLLLDNSATILSEKQALGN-	97
PERA2_ARMRU	-NASAIVRSTIQAFQSDTRIGASLIRLHFHDCFVNGCDASILLDDSGSIQSEKNAGPN-	70
PERX_TOBAC	-NVTISIVRGVMDQRQRTDARAGAKIRLHFHDCFVNGCDGSILLD-TDGTQTEKDAPAN-	91
Q40555_TOBAC	-FISMIV-----LLIGCDGSILLD-TDGTQTEKDAPAN-	63
O65773_CUCPE	-RLPNIVRQEVKRAIETDIRAGAKLIRFHFHDCFVQGGCDGSVLLDAPGIDSELNGLGN-	86
PER2_ZINEL	-TALSTIRTSIRSSVSSNRNRNALVIRLLFHDCFVQGGCDASLLLSGAG---SERASPN-	97
PER1_ZINEL	-TALSTIRTSIRSSVSSNRNRNALVIRLLFHDCFVQGGCDASLLLSGAG---SERASPN-	97
PERX_NICSY	-NALNTIRTSVRQAISERRMAASLIRLHFHDCFVQGGCDASILLDETPTIESEKTALPN-	97
PER1_HORVU	-RALATIKSGVMAAVTSDPRMGASLLRLHFHDCFVQGGCDASVLLSGME-----QNAIPN-	86
Q5I3F6_TRIMO	-RALATIKSGVMAAVTSDPRMGASLLRLHFHDCFVQGGCDASVLLSGME-----QNAIPN-	87
Q5I3F7_TRIMO	-RALATIKSGVMAAVTSDPRMGASLLRLHFHDCFVQGGCDASVLLSGME-----QNALPN-	87
Q5I3F5_TRIMO	-RALATIKSGVMAAVTSDPRMGASLLRLHFHDCFVQGGCDASVLLSGME-----QNAGPN-	88
Q5I3F4_TRIMO	-NALATIKSAVTAAVNENRNGASLIRLHFHDCFVQGGCDASVLLSGME-----QNAIPN-	85
Q9XIV9_TOBAC	-KLYQTVKSAVQSAINKETRMGASLLRLFHDCFVNGCDGSLLLDSTSSFTGEKRAAPN-	94
PER4_VITVI	-KVFDTVKSGVQSAVSKERRMGASLLRLFHDCFVNGCDASVLLDSTSSFTGEQTAVPN-	95
Q4W2V5_PICAB	-TALSIVKAAVNKAVERNMGASLIRLHFHDCFVNGCDGSILLDDNSTPTGEKTALPN-	93
Q4W2V2_PICAB	-RAQSIKRVVVKQAVAKEKRMGASLRLHFHDCFVNGCDGSILLDDNATFTGEKTAGPN-	96
Q5I3F2_TRIMO	-TLQLIVRATMIKALLAERRMGASLLRLHFHDCFVQGGCDGSILLDDVGSFVGEKTAFPN-	95
P93548_SPIOL	-NVEQIVFNTMKQAVSKEPRMGASLIRLFFHDCFVNGCDGSVLLDDTPTSGEKTALPN-	97
Q40372_MEDTR	-QALPTIKSVVLAHAILREKRIGASLRLHFHDCFVNGCDGSVLLDDTPTNFTGEKTALPN-	97
A0SWU6_SESRO	-NALSTIKSVVEGAVVKERRMGASLRLHFHDCFVNGCDGSILLDPTSSIDSEKNAGPN-	95
P93549_SPIOL	-NALRIVKQGIKAKRIKKEARVGASLIRLHFHDCFVNGCDGSILLDDTSTFRGEKTAIPN-	90
Q9MAX9_ASPOF	-RAQQIVKGVVEKAVAKDRMAASLRLHFHDCFVKGCDGSVLLDSSGTIVSEKRSNPR-	96
Q8H2B5_ANACO	-QVKEIVWSMVAKAVAKEPRMAASLRLHFHDCFVKGCDASILLDSSGSIVSEKSGKPN-	95
Q5W513_PICAB	-KAQAIKISMVEDAVKKEARIAASLRLHFHDCFVKGCDASILLDDNASTGFKTAIPN-	107
PERN_ARMRU	-NLLQIVRDQVKIALKAEIRMAASLIRLHFHDCFVNGCDASVLLDGT---NSEKLAIPN-	95
Q5I3F1_TRIMO	-RAEEIVRRVMARALAREARSVASVRLQFHDHCFVNGCDGSVLLMDATPTMAGEKEALS-	96
Q9XFL2_TOBAC	-QAEDIIREQVKLLYKRHKNTAFSWLRNIFHDHCFVESCDSALLDSTRRLSEKETDRS-	95
A3KLN6_CATRO	-QAEDIIREQVKLLYKRHKNTAFSWLRNIFHDHCFVESCDSALLDSTRRLSEKETDRS-	99
Q08671_GOSHI	-QAEDIIKEQVKLLYKRHKNTAFSWLRNIFHDCAVQSCDASLLLDSTRRLSEKETDRS-	101
B9RXL3_RICCO	-QAEDVIKEQVRLLYKRHKNTAFSWLRNIFHDCAVE-----STRRLSEKETDRS-	78
PER42_ ARATH	-QAEDIVREQVKLLYKRHKNTAFSWLRNIFHDCAVEESCDSALLDSTRRELGEKEHDS-	99
A9XN55_THEHA	-QAEDIIKEQVKLLYKRHKNTAFSWLRNIFHDCAVEESCDSALLDSTRRELGEKEHDS-	100
B7FHN5_MEDTR	-QAEIIKEQVRLLYKRHKNTAFSWLRNIFHDCAVQSCDASLLLTSTRRLSEKEHDS-	100
Q9FT05_CICAR	-QAEIIKEQVRLLYKRHKNTAFSWLRNIFHDCAVQSCDASLLLTSTRRLSEKEHDS-	100
A9NNP1_PICSI	-QAEIIIAEQVRLLYKRHKNTAFSWLRNIFHDCAVEESCDSALLDSTRKSISEKETDRS-	106
Q9FXL6_AVIMR	GR-KSRAEILEQHLDPDHNESLSVLDRLFANMG-IDAPGVVAL-LGAHSGVRTHCVKLVHR	188
B2G335_CATRO	GLNFATQNETLNNLPPPSNTSELTSLATKN-FNATDVVAL-SGGHTTIGVGHCVSFEER	228
B2G334_CATRO	GLNFATQNETLNNLPPPSNTSELTSLATKN-FNATDVVAL-SGGHTTIGVGHCVSFEER	223
Q0VYC8_CATRO	GLNFATRADTIANLPPPTSNTSALLTSLATKN-FNATDVVAL-SGGHTTIGIGHCVSFEER	223
Q6PQF2_EUPCH	GVKFAEVNATFEHLVGPTANVTTLIAKLARKG-LDTTDAVSL-SGGHTTIGIGHCVSFEER	214
P93545_SPIOL	GLSFATQSETLANLPPPFNTTQLLNFAFATKN-LNATDLVAL-SGGHTTIGISHCVSFTNR	217
P93547_SPIOL	GLNFATLSDTLANLPPPSFNTGQLLDSLANKK-LNATDLVAL-SGGHTTIGISHCVSFTNR	215
Q5W514_PICAB	SLKFATQNTVLANLPAPSSKVTTLIKAFATKN-LNVTDLVAL-SGGHTTIGIGHCVSFTNR	218
Q5W512_PICAB	SLTFANLSTTLANLPSPTSNTVTTLISVLGPKG-LTFTDLVAL-SGGHTTIGISHCVSFTNR	212
Q40069_HORVU	SRSFASTQDVLSDLPGPSSNVQSLALLLGRGLG-LDATDLVTI-SGGHTTIGLAHCSSFEEDR	220
Q49866_HORVU	GLAPAS-SALVGLLPVPFDPVPTLISFFANRS-LDVADLVSL-SGAHTTIGVGHCVSFEEDR	204
Q58A85_ZINEL	GR-VSKANETIN-LPSPFNATTLIQSFARKG-LDVKDLVTL-SGGHTTIGLSHCNRSFSNR	204
Q9ZP15_TRIRP	GL-VSRSSDVNGRLPQPGFNLNQLNSLFASNG-LTQTDMLAL-SGAHTTIGLSHCNRSFSNR	211
PERN1_TOBAC	GR-VSRADAGD-LPAFFDSVDIQKRKFLTKG-LNTQDLVAL-TGAHTTIGTAGCAVIRDR	207

Table 1 (continued)

Q4ADU9_POPAL	GR-VSLASDTSN-LPGFTDSVDVQKQKFAAFG-LNAQDLVTL-VGGHTTIGTTACQFFRYR	200
Q4A3Z3_SENSQ	GL-VSRAEDALK-LPGSRDSAEOVIEKFAAKG-LNIEELVTL-VGGHTTIGTSACARFVHR	203
Q4A3Z1_SENSQ	GL-VSRAEDALK-LPGSRDSAEOVIEKFAAKG-LNIEELVTL-VGGHTTIGTSACARFVHR	203
Q4A3Z2_SENSQ	GL-VSRAEDALK-LPGSRDSAEOVIEKFAAKG-LNIEELVTL-VGGHTTIGTSACARFVHR	203
Q4A3Y8_SENSQ	GL-VSRAEDALK-LPGSRDSAEOVIEKFAAKG-LNIEELVTL-VGGHTTIGTSACARFVHR	203
P93551_SPIOL	GK-RSSASDAVN-LPSPLESIAVHRQKQKFAADKG-LNDHDLVTL-VGAHTTIGTQDCRFQYR	184
PER13_ARATH	GL-IANVDHAKN-LPDVQDSINTLKSFKREKG-LSDQDLVLLSAGAHHTIGTTACFFVPIR	202
PER39_ARATH	GR-INSFAEAMNNIPPPFNTFTLITLFGNQG-LDVKDLVLL-SGAHTTIGVSHCSSFSNR	205
Q5I3F0_TRIMO	GN-LSTDGTGAFT-LPGPNLTADGLVRGFGADR-LTAEDMVVL-SGSHHTLGRSHCNFSFIVR	211
A1YBM0_HELAN	GR-VSLAQNTQT-LPPLSASVSSAIRAFARKG-LNPDTMLYL-LGGHTTIGIAHCSLFQDR	140
P93546_SPIOL	GF-VSRKNEAQASIPPTMPVPQAIQLFASKG-LNKDDFVVL-LGGHTTIGTGAHCSFRER	205
PER2_ARMRU	SL-RGFMDLANDNLPGPSTLQVLKDKFRNVGLDRPSDLVAL-SGGHTTIGKNCQCFIMDR	207
PER1B_ARMRU	SL-QAFLDLANANLPAPFFTLPLQKDAFAKVGGLDRPSDLVAL-SGGHTTIGKNCQCFIMDR	211
PER3_ARMRU	SV-EAFFDLANTALPSPFFTLAQLKKAFAADVGLNRPDLVAL-SGGHTTIGRAQCQFVTPR	212
PERE5_ARMRU	SV-EAFFDLANTALPSPFFTLAQLKKAFAADVGLNRPDLVAL-SGGHTTIGRAQCQFVTPR	183
Q43099_POPTR	ST-TASRAAANASLPAPFLLDQLRESFTNVGLNNNTDLVAL-SGAHTTIGRAKCSFNFNR	207
Q43101_POPTR	ST-TASRAAANASLPAPFLLDQLRESFTNVGLNNNTDLVAL-SGAHTTIGRAKCSFNFNR	207
Q43050_POPKI	SL-TANRTAANAFIPGPQDTLRLRSRFTVVGNNNTDLVAL-SGAHTTIGRAKCSFNFNR	207
Q43102_POPTR	SL-TANRSGANSSIPAPSESLAVLKSFAAVGLNTSSDLVAL-SGAHTTIGRAQCCLNFISR	211
Q42905_LINUS	SL-TANRSLADDQLPPFFTVDELKANFATVGLNTTEDLVAL-SGAHTTIGRARCQGVFVNR	210
PERA2_ARMRU	SL-TANLAGANSAIPSPFETLRLRSRFTVVGNNNTDLVAL-SGAHTTIGRARCQGVFVNR	182
PERX_TOBAC	SL-TANRSGANSDIPSPFETLAVMIPOFTNKGMDLT-DLVAL-SGAHTTIGRARCCTFEQR	202
Q40555_TOBAC	SL-TANRSGANSDIPSPFETLAVMIPOFTNKGMDLT-DLVAL-SGAHTTIGRARCCTFEQR	174
O65773_CUCPE	SR-TANRTGAD-ELPSPFETLRLRSRFTVVGNNNTDLVAL-SGAHTTIGRARCCTFEQR	196
PER2_ZINEL	ST-TSNAQAATDLPRGNMVLSQLISNFANKG-LNTREMVVAL-SGSHHTLQARCFRFRGR	208
PER1_ZINEL	ST-TSNAQAATDLPRGNMVLSQLISNFANKG-LNTREMVVAL-SGSHHTLQARCFRFRGR	208
PERX_NICSY	ST-TASKTLAETDLPGPFDPDLNRLISSFSKQ-LSTRDMVAL-SGAHTTIGQAQCQCFVFR	209
PER1_HORVU	SI-DANEANEANTDLPGFNSSRAEELAAFLKKGGLNTVDMVAL-SGAHTTIGQAQCSTFRAR	199
Q5I3F6_TRIMO	SI-DANEANEANTDLPGFNSSRAEELAAFLKKGGLNTVDMVAL-SGAHTTIGQAQCSTFRAR	200
Q5I3F7_TRIMO	SI-DANEANEANTDLPGFNSSRAEELAAFLKKGGLNTVDMVAL-SGAHTTIGQAQCSTFRAR	199
Q5I3F5_TRIMO	ST-TASASLANSDLPGPSSSRSLQEAFLKKN-LNTVDMVAL-SGAHTTIGQAQCSTFRAR	200
Q5I3F4_TRIMO	ST-TASASLANSDLPGPSSSRSLQEAFLKKN-LNTVDMVAL-SGAHTTIGQAQCSTFRAR	197
Q9XIV9_TOBAC	SR-TASQSAANSIPIPATSNLNLRISSFSKQ-LSTKDMVAL-SGAHTTIGQAQCSTFRAR	206
PER4_VITVI	SK-TASLSGANNIPPTSSSLNLISKFQAG-LSTRDMVAL-SGAHTTIGQAQCSTFRAR	207
Q4W2V5_PICAB	ST-TASLSGANNIPPTSSSLNLISKFQAG-LSTRDMVAL-SGAHTTIGQAQCSTFRAR	205
Q4W2V2_PICAB	ST-TASLSGANNIPPTSSSLNLISKFQAG-LSTRDMVAL-SGAHTTIGQAQCSTFRAR	208
Q5I3F2_TRIMO	ST-TASLTEANSNDLPAPSLSLGLIKAFDKKQ-LSPQDLTAL-SGAHTTIGFSQCLNFRDH	207
P93548_SPIOL	AR-TANMTLANLNLPPGNAPLANLTELFAFQON-LNIREMTAL-SGSHHTIGFARCTNFRDH	209
Q40372_MEDTR	AR-TASKAANANLPSPTFNFSQLISNFKQ-LNVKDLVAL-SGSHHTIGFARCTNFRDH	212
A0SWU6_SESRO	ST-TASRAAANKDIPAPFFSLSQLIENFKNG-LDEKDLVVL-SGSHHTIGYARCATFRDH	208
P93549_SPIOL	SL-TANRSAANAFIPAPSNLNLRISSFTTVG-LSFKDMVVL-SGAHTTIGFARCTNFRDH	202
Q9MAX9_ASPOF	SL-GASLSGSNNYIPAPNNTLPTIITKFKLKG-LDIVDLVTL-LGSHHTIGFARCTNFRDH	208
Q8H285_ANACO	SL-GASLSGSNNYIPAPNNTLPTIITKFKLKG-LDIVDLVTL-LGSHHTIGFARCTNFRDH	207
Q5W513_PICAB	SR-SASKSGANEDLPAPNSTHTLTELKFKLQG-LNVVDLVAL-SGAHTTIGLARCSAFKQR	219
PERN_ARMRU	GL-VANQ-SSANNLPSPFELDLAIKFAAVG-LNVTDVVAL-SGAHTTIGLARCSAFKQR	206
Q5I3F1_TRIMO	SL-TASQEDSDNIMPSPRANASALIRLFAGYK-LTVTDLVAL-SGSHSIGEARCFSIVFR	208
Q9XFL2_TOBAC	GR-KSRADILEQHLDPDHNESMSVLERFANVG-INAPGVVAL-LGAHSGVGRTHCVKLVHR	206
A3KL6_CATRO	GR-RSRAEILEQHLDPDHNESMSVLERFANVG-INAPGVVAL-LGAHSGVGRTHCVKLVHR	210
Q08671_GOSHI	GR-RSRADVVEEYLPDHNETISGLVDRFAAMG-IDTPGVVAL-LGAHSGVGRTHCVKLVHR	212
B9RXL3_RICCO	GR-KSRADVLEQYLPDHNESISVLERFANVG-IDTPGVVAL-LGAHSGVGRTHCVKLVHR	149
PERA2_ARATH	GL-KSRTDMLESYLPDHNESISVLEKFKSIG-IDTPGVVAL-LGSHSGVGRTHCVKLVHR	210
A9XN55_THEHA	GR-KSRTDMLESYLPDHNESISVLEKFKSIG-IDTPGVVAL-LGSHSGVGRTHCVKLVHR	211
B7FHN5_MEDTR	GR-KSRVDLLEAYLPDHNESISAVLDKFGAMG-IDTPGVVAL-LGAHSGVGRTHCVKLVHR	211
Q9FT05_CICAR	GR-KSRVDLLEAYLPDHNESISAVLDKFGAMG-IDTPGVVAL-LGAHSGVGRTHCVKLVHR	211
A9NNP1_PICSI	GR-KSRADVVENYLPDHNESISTVLSRFKAMG-IDTRGVVAL-LGAHSGVGRTHCVKLVHR	217
Q9FXL6_AVIMR	LYPEVDPA-----LNPSHVEHMLHKCPDPIPDPAVQYVRNDRGTPMKLDNNYYRNILD	242
B2G335_CATRO	LYPT-----QDPTMDQTFARNLRL-TCPALNTTN---TVFLDIRSPNRFNDRYYVDLMN	278
B2G334_CATRO	LYPT-----QDPTMDQTFARNLRL-TCPALNTTN---TVFLDIRSPNRFNDRYYVDLMN	273
Q0VYC8_CATRO	IYPN-----IDPTMDQTFARNLRL-TCPPTDSNN---RFFLDIRSPNRFNDRYYVDLMN	273
Q6PQF2_EUPCH	LYPS-----QDPTLDKTFANNLKR-TCPNVNTEN---STFLDLRTPNEFDNRYVDLMN	264
P93545_SPIOL	LYPT-----QDPSMDQTLANNLKL-TCPATATNS---TNFLDLRTPNVFDNKYYVDLMN	267
P93547_SPIOL	LYPT-----QDPTMAQTFANNLKV-TCPATATNA---TNFLDLRTPNVFDNKYYVDLMN	265
Q5W5I4_PICAB	LYPK-----QDPTLNKSFAQRLYT-ACPPKTSSN---TVFLDIRTPNVFDNKYYVDLMN	268
Q5W5I2_PICAB	LYNSTTGISMQDSTLDQNFANLKL-TCPNTSVN---TNFLDLRTPNVFDNKYYVDLMN	268
Q40069_HORVU	LFPR-----PDPTISPTFLSRLKR-TCPVKGTD---RFLVDLRTPNVFDNKYYVDLMN	270
O49866_HORVU	FKPVFD---TNPAIDGKFATALRN-KCAGDNPAGTL-TQNLVDLRTPNVFDNKYYVDLIA	258
Q58A85_ZINEL	IHN-----SIDPTINSEFAMSLKK-KCPLKNKDRN-AGEFLDSTSSRFNDRYYKRTIM	255
Q9ZP15_TRIRP	IFNFNNQ-SPVDPTLNKQYAAQLQ-MCPNRVDPRI--AINMDPTTPROVDNAYYQNLQ	267
PERN1_TOBAC	LFNFNSTG-GPDPSIDATFLPQLRA-LCPQNGDASR--RVGLDTCGSVNNFDTSYFNLNR	263
Q4ADU9_POPAL	LYNFTTTGNGADPSINPSFVSQLQT-LCPQNGDASR--RIALDTCGSQNSFDSSFFANLRS	257
Q4A3Z3_SENSQ	LYNYSNTN-APDPHIDQAFPLHLQT-LCPEHGDRTI--RVDLDTGGSVNNFDTSYYENLRK	259
Q4A3Z1_SENSQ	LYNYSNTN-APDPHIDQAFPLHLQT-LCPEHGDRTI--RVDLDTGGSVNNFDTSYYENLRK	259
Q4A3Z2_SENSQ	LYNYSNTN-APDPHIDQAFPLHLQT-LCPEHGDRTI--RVDLDTGGSVNNFDTSYYENLRK	259
Q4A3Y8_SENSQ	LYNYSNTN-APDPHIDQAFPLHLQT-LCPEHGDRTI--RVDLDTGGSVNNFDTSYYENLRK	259
P93551_SPIOL	LYNFTPTG-NADPSINQPNIAQLQT-LCPKNGNGLT--KVALDRDSRTKFDVNFKNIRD	240
PER13_ARATH	LD-----AQDPTINPEFFQILRS-KCPQGGDVNV--RIPLDWDSQFVFNQIFQNLK	252
PER39_ARATH	LFNFTGVG-QDPSLDSEYADNLKSRRLSIADNTT--KVEMDPSGRNTFDLSYRLVLK	262
Q5I3F0_TRIMO	NRRERLASG-----TISPAYQALLEA-LCPANTSQFTNVTTEIDLSPTPVLDNNYYKLVL	265
A1YBM0_HELAN	LYNFQNSG-KPDPTMDATLIASLRN-RCAQNATVDRTVNLQDNPLSSKVIDKSFYXQIML	198

Table 1 (continued)

P93546_SPIOL	LYNFRNTK-KPDSTISPTLLQLLQK-TCPRNSQTDNETFLDQTPNSHF	KIDNGYYKQILA	263
PER2_ARMRU	LYNFSNSG-KPDPTLDKSYLSTLRK-QCPRNGN--LSVLVDFDLRTPT	IFDNKYYVNLKE	263
PER1B_ARMRU	LYNFSNTG-LPDPPTLNTTYLQTLRQ-QCPLNGN--QSVLVDFDLRTPT	VFDNKYYVNLKE	267
PER3_ARMRU	LYNFNGTN-RPDPTLDPTYLVQLRA-LCPQNGN--GTVLVNFDVVT	PNPTFDRQYYTNLRN	268
PERE5_ARMRU	LYNFNGTN-RPDPTLNPSYLADLRK-LCPRNGN--GTVLVNFDVMT	PNPTFDNQFYTNLRN	239
Q43099_POPTR	LYDFNGTG-APDPTLDPPFLAALQE-LCPQGGN--DSVLTDLDLT	TPDAFDSNYYSNLQC	263
Q43101_POPTR	LYDFNGTG-APDSTIDPPFLEALQK-LCPENG- GSVITDLDVT	TADAFDSKYYSNLQC	263
Q43050_POPKI	LYNFNNTG-LPDPPTLDTTYLATLQR-LCPQGGN--GTVLADLDPT	TPDGFDDNNYFSNLQA	263
Q43102_POPTR	LYNFSGSG-NPDPTLNTTYLAALQE-LCPQGGN--RSVLTNLDRT	TPDFTDGNYYSNLQ	267
Q42905_LINUS	LYNFNSTG-GPDPTINATFLETLRQ-ICPQNGN--GSVLTNLDRT	TADAFDSNYYFTNLQT	266
PERA2_ARMRU	LFNFSGTN-GPDPTLNSTLLSSLQ-LCPQNGS--ASTITNLDLS	TPDAFDDNNYFANLQS	238
PERX_TOBAC	LFNFNGSG-NPDLTVDATFLQTLQG-ICPQGGN-GNTFTNLDIS	TPNDFDNDYFTNLQS	259
Q40555_TOBAC	LFNFNGSG-NPDLTVDATFLQTLQG-ICPQGGN-GNTFTNLDIS	TPNDFDNDYFTNLQS	231
Q65773_CUCPE	FSNFNGTG-QPDPALDPAYRQELER-ACTDGET----	RVNFDPTTPDFTDKNYYTNLQA	249
PER2_ZINEL	IYNSTLR-----IEPNFNRSLQ-ACPPTGN--DATLRPLDLV	TPNSFDNNYYRNLT	258
PER1_ZINEL	IYNSTLR-----IEPNFNRSLQ-ACPPTGN--DATLRPLDLV	TPNSFDNNYYRNLT	258
PERX_NICSY	IYNGTD-----IDAGFASTRRR-QCPQEGE--NGNLAPLDLV	TPNQFDNNYFKNLIQ	259
PER1_HORVU	IYGGDTN-----INAAAYASLRA-NCPQTVGSGDGS	LANLDTTANTFDNYYTNLMS	251
Q513F6_TRIMO	IYGGDTN-----INAAAYASLRA-NCPQTVGSGDGS	LANLDTTTPNAFDNYYTNLMS	252
Q513F7_TRIMO	IYN-ETN-----IDTFATSLRA-NCPRSGG--DGS	LANLDTTANTFDNYYTNLMS	248
Q513F5_TRIMO	IYGGDTN-----INTAFATSLRA-NCPQSGG--NTNL	ANLDTTTPNAFDNYYTNLMS	250
Q513F4_TRIMO	LYN-ETN-----INSGFATSLRA-NCPQTVGSGDRN	LANLDTTTPYSDNYYTNLKS	248
Q9XIV9_TOBAC	IYNETNN-----LDASFARTQS-NCPRSSGSGDNN	APLDLQTPNKFDDNNYFKNLVD	258
PER4_VITVI	IYNETN-----IDSSFATRQA-SCPSASGSGDNN	APLDLQTPPTFDNYYKNLIN	258
Q4W2V5_PICAB	IYNESN-----IDTSFATSVKS-SCPSAG--GDNTL	SPLDLATPTTFDNYYTDLGN	254
Q4W2V2_PICAB	IYNESN-----INAAFATSVKA-NCPASAG--GDNTL	SPLDLVTSIKFDNYYGNLKI	257
Q513F2_TRIMO	IYNGTN-----IDPAFATLRKA-TCPAQAPNGDKN	APFDVQTLQFDNYYKNLVA	258
P93548_SPIOL	IYNDN-----IDPNFAATRKA-SCPRPTGTGDFN	APMDIQTPNTFDNYYKNLVA	260
Q40372_MEDTR	IYNETN-----IDPIFAASLRK-TCPRNG--GDNNL	TPLDF-TPTVENTYRDLLY	260
A0SWU6_SESO	IYKDT-----INSEFAQQLYK-ICPING--GDSN	SPLDL-TAANFDVAYYSNLLQ	256
P93549_SPIOL	IHNNTN-----INAAFASLQK-KCPQSG--NGKVL	QPLDYQTKFBDKYYQNLV	251
Q9MAX9_ASPOF	LYNQSGNG-LPDATLDKTYAAQLRQ-RCPQSGG-DQN-L	FALDFNTQKFDDNYYKNLVA	264
Q8H285_ANACO	LYNQSGNG-LADSTLDESAMQLRW-GCPRSGS-DDN-L	FPLDVYSPAQFDNYYKNLVA	263
Q5W5I3_PICAB	LYNQGTN--KPDQTLDTTYLKQLRT-VCPQGTG-DNNQ	TRFPDVSPTKFDVNYKNVVA	275
PERN_ARMRU	LFNFTGAG-TPDSTLETLLSLDQ-LVCPIGG--GNKT	APLDRNSTDAFDDNNYFKNLLE	262
Q513F1_TRIMO	LYNQSGSG-RPDPHMDPAYRQALDA-LCLPTGD--QNV	TGGLD-ATPVDFDNDYFTNLVH	263
Q9XFL2_TOBAC	LYPEVDPA-----LNPDPHVMHLKKCPDPI	PDPKAVQYVRNDRGTPMKLDNNYYRNILE	260
A3KLN6_CATRO	LYPEVDPA-----FPESHVQHMLKKCPDPI	PDPKAVQYVRNDRGTPMKLDNNYYRNILD	264
Q08671_GOSHI	LYPEVDPA-----LSPDHVPHMLHKKCPDI	PDPKAVQYVRNDRGTPMVLNDDNNYYRNILD	266
B9RXL3_RICCO	LYPEVDPA-----LNPDPHVMHLHKKCPDPI	PDPKAVQYVRNDRGTPMVLNDDNNYYRNILD	203
PER42_ARATH	LYPEVDPS-----LNPDPHVMHLHKKCPDSI	PDPKAVQYVRNDRGTPMVLNDDNNYYRNILD	264
A9XN55_THEHA	LYPEVDPS-----LNPDPHVMHLHKKCPDSI	PDPKAVQYVRNDRGTPMVLNDDNNYYRNILD	265
B7FHN5_MEDTR	LYPEVDPA-----LNPEHIPHMLKKCPDSI	PDPKAVQYVRNDRGTPMVLNDDNNYYRNILD	264
Q9FT05_CICAR	LYPEVDPA-----LNPEHIPHMLKKCPDSI	PDPKAVQYVRNDRGTPMVLNDDNNYYRNILD	265
A9NNP1_PICSI	LYPEVDPT-----LDPGHVHEMHKKCPDAI	PNPKAVQYVRNDRGTPMKLDNNYYVNLMN	271
Q9FXL6_AVIMR	NKGLLIIVDQHLATDK---RTKPFVKKMAK----	SQDYFFKEFGRAITVLSNNPLTGT	293
B2G335_CATRO	RQGLFTSDQDLYTDR---RTRGIVTDFAV----	NQTLFFEKFVYAMIKMGQLNVLGT	329
B2G334_CATRO	RQGLFTSDQDLYTDR---RTRGIVTDFAV----	NQTLFFEKFVYAMIKMGQLNVLGT	324
Q0VYC8_CATRO	RQGLFTSDQDLYTDR---RTRGIVTDFAI----	NQTLFFEKFVYAMIKMSQLNVLGT	324
Q6P62H_EUPCH	RQGLFTSDQDLYTDK---RTRQIVDFAV----	NQTLFFEKFVYAMIKMGQLEVTGT	315
P93545_SPIOL	HQGLFTSDQDLYTDS---RTKAIVTSFAT----	NQNLFFEKFIDAMVKMSQLSVLTGT	318
P93547_SPIOL	RQGLFTSDQDLYTDS---RTKDIVTSFAL----	NQNLFFQKFIDAMVKMGQLNVLGT	316
Q5W5I4_PICAB	RQGLFTSDQDLYSDS---RTKAIVNDFAL----	DQDLFFEKFVAMVKMGQLNVLGT	319
Q5W5I2_PICAB	EQTLFTSDQSLYTDY---RTRDIVKSFAL----	NQSLFFQFVLSMLKMGQLDVLGT	319
Q40069_HORVU	REGLFVSDQDLFTNA---ITRPIVERFAR----	SQDDFFEQFVSIKMGQMRVTS	321
Q49866_HORVU	RQGLFKSDQGLIDHP---TTKRMATRFSL----	NQGAFFEQFARSMTNMDILTGT	309
Q58A85_ZINEL	GKGVFGSDQALYGDS---RTKGIVDSYAK----	DEKLFKEFAASMVKLGNVGVIED	305
Q9ZP15_TRIRP	GKGLFTSDQILFTDT---RSRATVNSFAS----	SGNVFNANFINAMTKLGRIGVKTAR	318
PERN1_TOBAC	GRGVLESDDKLWTD---STQVFVQRFGLG-IRGLLGLT	FGVEFGSRSMVKMSNIEVKTGT	318
Q4ADU9_POPAL	GQGILESDDKLWTD---TTRTFVQRFGLG-VRGLAGLT	FGVEFGSRSMVKMSNIGVKTGT	312
Q4A3Z3_SENSQ	GRGVLESDDTKLWTH---ITQNLVQQFIS-VGRPNQLT	FSKKFARAMVKLSQVEVKTGN	314
Q4A3Z1_SENSQ	GRGVLESDDTKLWTH---ITQNLVQQFIS-VGRPNQLT	FSKKFARAMVKLSQVEVKTGN	314
Q4A3Z2_SENSQ	GRGVLESDDTKLWTH---ITQNLVQQFIS-VGRPNQLT	FSKKFARAMVKLSQVEVKTGN	314
Q4A3Y8_SENSQ	GRGVLESDDTKLWTH---ITQNLVQQFIS-VGRPNQLT	FSKKFARAMVKLSQVEVKTGN	314
P93551_SPIOL	GNAVLESDDRLWGDD---ATQAIQVQNYAGNLRGL	FGVRFNDFPKAMVKMSGIVKSGS	296
PER13_ARATH	GRGVILSDSVLYQDN---NMKKIIDSYLE-TNQSSKAN	FADFTKAMIKMGAIGVKIGA	307
PER39_ARATH	RRGLFESDAALTMNP---AALAQVKRFAG---GSEQE	FFAEFNSMEKMGRIQVKTGS	314
Q513F0_TRIMO	NLGLHFSDDQLIRNA---TLKAFVDAFAA-----	NETLWKDKFLAAMIKMGNISPKTGT	316
A1YBM0_HELAN	RRGILQIDQQLALDG---LTKSKVAASIAS-----	SSD-FNTKFGQAMVKMGN-----	241
P93546_SPIOL	HNGVMEIDQNLALYP---STRCLVTGLAH-----	NPNQFLERFGPAMVKMARIGVLTGS	314
PER2_ARMRU	NKGLIQSDQELFSSP-DASDTIPLVRAYAD----	GQKGFDAFVEAMIRMGNLSPSTGK	317
PER1B_ARMRU	QKGLIQSDQELFSSP-NATDTIPLVRSFAD----	GTQKFFNAFVEAMNRMGNIPTLTGT	321
PER3_ARMRU	GKGLIQSDQELFSTP--GADTIPLVNLVYSS----	NTFAFFGAFVAMIRMGNLRLTGT	321
PERE5_ARMRU	GKGLIQSDQELFSTP--GADTIPLVNLVYSS----	NTLSFFGAFADAMIRMGNLRLTGT	292
Q43099_POPTR	NRGLLQTDQELFSTP-GADDVIALVNAFSA----	NQTAFVESFVESMIRMGNLSPLTGT	317
Q43101_POPTR	NRGLLQTDQELFSTP-GADDVIALVNAFSA----	NQTAFVESFVESMIRMGNLSPLTGT	317
Q43050_POPKI	SKGLLQSDQELFSTP-EADDIIELVDFST-----	DETAFFESFVESMIRMGNLSPLTGT	317
Q43102_POPTR	NEGLLQSDQELFSTT-GAD-TIAIVNNFSS----	NQTAFVESFVSMIRMGNISPLTGT	320
Q42905_LINUS	REGLLQTDQELISTP-GSD-TIELVNRFAA----	NQTAFQSFVNSMIRMGNIPPPGS	319

Table 1 (continued)

PERA2_ARMRU	NNGLLQSDQELFSTL-GSA-TIAVVTSFAS-----NQTLFFQAFQSMINMGNISPLTGS	291
PERX_TOBAC	NQGLLQTDQELFSTS-GSA-TIAIVNRYAG-----SQTQFFDDFVSSMIKLGNIPLTGT	312
Q40555_TOBAC	NQGLLQTDQELFSTS-GSA-TIAIVNRYAG-----SQTQFFDDFVSSMIKLGNIPLTGT	284
O65773_CUCPE	NRGLLTSDQVLFSTP-GAD-TIEIVNRLGS-----REGTFFRQFRVSMIKMGNIPLTGN	302
PER2_ZINEL	SRGLLISDQVLFN---ADS-TDSIVTEYVN-----NPATFAADFAAAMVKMSEIGVVTGT	309
PER1_ZINEL	SRGLLISDQVLFN---ADS-TDSIVTEYVN-----NPATFAADFAAAMVKMSEIGVVTGT	309
PERX_NICSY	KKGLLQSDQVLFN---GGS-TDNIVSEYSN-----SARAFSSDFAAAMIKMGDISPLSGQ	310
PER1_HORVU	QKGLLHSDQVLFN---NDT-TDNTVRNFAS-----NPAAFSSSFTTAMIKMGNIAPKTGT	302
Q513F6_TRIMO	QRGLLHSDQVLFN---NDT-TDNTVRNFAS-----NPAAFSSAFTTAMIKMGNIAPKTGT	303
Q513F7_TRIMO	QKGLLHSDQVLFN---NDT-TDNTVRNFAS-----NPAAFSSAFTTAMIKMGNIAPKTGT	299
Q513F5_TRIMO	QKGLLHSDQVLFN---NDT-TDNTVRNFAS-----NPAAFSSAFTTAMIKMGNIAPKTGT	301
Q513F4_TRIMO	QKGLLHSDQVLFN---NDT-TDNTVRNFAS-----NPAAFSSAFTTAMIKMGNIAPKTGT	301
Q9XIV9_TOBAC	KKGLLHSDQQLFNG--GS--ADSIVTSYSN-----NPSSFSSDFVTAMIKMGDIRPLTGS	309
PER4_VITVI	QKGLLHSDQVLYNG--GS--TDSIVTKTYVN-----NPKTFTSDVFAGMIKMGDIRPLTGS	309
Q4W2V5_PICAB	RKGLLHSDQQLFSG--GS--TNSQVTTYSA-----NQNTFFTDFAAAMVKMGNIPLTGT	305
Q4W2V2_PICAB	QKGLLHSDQQLFNG--GP--TDSQVTAYST-----NQNSFFTDFAAAMVKMGSNISPLTGT	308
Q513F2_TRIMO	KRGLLNSDQVLFNG--GS--QDALVRQYVA-----NPALFASDFVTAMIKMGNIPLTGT	309
P93548_SPIOL	KRGLLHSDQQLFNG--GS--QDALVRQYVA-----NPALFASDFVTAMIKMGNIPLTGT	311
Q40372_MEDTR	KRGVLHSDQQLFKGQ-GSE-SDKLVQLYSK-----NTFAPASDFKTSIKMGNIKPLTGR	313
A0SWU6_SESRO	TKGLLHSDQELFNG--GS--TDELVKQYSY-----YPEAFFQDFAKSMIKMGNIPLTGD	307
P93549_SPIOL	KKGLLHSDQQLYSGN--NN--ADAYVRKYAS-----KQGEFFQEFNGSMIRMGNIKPLTGT	303
Q9MAX9_ASPOF	SEGLLSSDEILF-TQ-SST-TMALVKKYAE-----DNGAFFEQFAKSMVKGNIPLTGT	316
Q8H285_ANACO	GKGLLNSDQILF-TK-SAT-TRQLVELYAA-----NIGIFYDHFASMIKMGNIPLTGL	315
Q5W5I3_PICAB	GKGLLNSDEILYSTK-GSR-TAGFVKYYTT-----NTHAFFKQFAASMIKMGNIPLTGF	328
PERN_ARMRU	GKGLLSSDQILFSSDLAVNTTKRLVEAYS-----SQYLFRRDFTCSMIRMGSVLV--NGA	315
Q513F1_TRIMO	LRGFLNSDQTLFSDN---EGTRRVVTQFSQ-----NQDAFFRAFIEGMVKLGELQ--NPR	313
Q9XFL2_TOBAC	NKGLMLVDHQLATDK---RTKPYVKKMAK-----SQDYFFKEFARAITILSENPLTGT	311
A3KLN6_CATRO	NKGLLLVDHQLATDK---RTKPYVKKMAK-----SQDYFFKEFARAITILSENPLTGT	315
Q08671_GOSHI	NKGLLIVDHQLAYDK---RTRPYVKKMAK-----SQDYFFKEFSRAITILSENPLTGS	317
B9RXL3_RICCO	NKGLLIVDHQLATDK---RTRPYVKKMAK-----SQDYFFKEFSRAITILSENPLTGS	254
PER42_ARATH	NKGLLLVDHQLAHDK---RTRPIVKKMAK-----DQAYFFKEFTRAIQILSENPLTGS	315
A9XN55_THEHA	NKGLLLVDHQLAHDK---RTRPIVKKMAK-----DQAYFFKEFTRAIQILSENPLTGS	316
B7FHN5_MEDTR	-----	-----
Q9FT05_CICAR	NKGLLSVDHQLAHDK---RTKPYVKKMAK-----SQDYFFKEFSRAITILSENPLTGT	316
A9NNP1_PICSI	NKGLLIVDQQLYADS---RTRPYVKKMAK-----SQDYFFKYFARALITILSENPLTGN	322

Amino acid sequence of *Avicennia marina* peroxidase (Q9FXL6_AVIMR): Distal site: F/YYXST/SCP conserved fold (red), catalytic residues (blue) while another conserved motif DCFVQ/NGCDGSVLI/L (green) are present. These residues provide distal site coordination with Ca. Proximal site: His175 and Ser/Thr176 (in red) are involved in coordination with heme and calcium respectively which are conserved in all plant peroxidases. Asp250 (in red) maintains imidazolate nature of His175 ring by forming two salt bridges. Other proximal residues, Val221- Lys231 (in blue) are conserved in some peroxidase sequences while hydrophobic amino acids are present in place of Lys231 in most of the peroxidases. These are also involved in coordination with heme and calcium

Thellungiella halophila, and *Picea sitchensis*. At position 54 and 55 of the conserved motif, ⁵⁰DCFVQ/NGCDGSVLI/L⁶², Glu and Ser are present in place of Asn/Gln and Gly, respectively. At positions 67, 77, and 79, Arg, Ser, and Gly are present, respectively, in *A. marina* while in most of peroxidase family proteins Ser/Thr, Asn, and Ser are present at these positions; these residues are involved in hydrogen-bonding interactions with distal catalytic residues. These H-bonding interactions are not observed in *A. marina* peroxidase.

His175 at the proximal site of heme and Ser/Thr176 at the proximal site of calcium are conserved in all plant peroxidases. Val222, Arg223, Asn224, Asp225, Arg226, Gly227, Thr228, and Lys231 are present at the proximal site of calcium in *A. marina* peroxidase. Asp225 and Thr/Ser228 are mostly conserved and Leu224 is 70% conserved in the family of class III peroxidase while Asn224 is present in place of Leu224 in *A. marina* and some other peroxidases (Table 1). At position 231 variable amino acid residues are present at the proximal site of calcium in the peroxidase family.

Phylogenetic analysis

Phylogenetic analysis based on the multiple sequence alignment of 54 sequences of salt-tolerant plant peroxidases has shown that peroxidases can be grouped in several distinct branches (Fig. 1). The branches that were recognized as clusters were considered to be a sign of biological similarity. An individual partition contains different number of classes, where each class consists of a cluster of similar sequences originating from a common node within the partition. There is a close phylogenetic relationship between peroxidases belonging to *A. marina*, *N. tabacum*, *V. rosea*, *Gossypium hirsutum*, and *Ricinus communis* indicating that they share a common ancestor.

Analysis of the amino acid sequence deduced for AP using the databank, PeroxiBase (Koua et al. 2009), revealed that *A. marina* has the highest homology with that of *Triphysaria versicolor*, *N. tabacum*, *C. roseus*, and *Capsicum annum*. AP also shares significant identity with barley grain peroxidase (BGP) whose 3D structure has been determined (Henriksen et al. 1998a).

Fig. 1 Phylogenetic analysis of *A. marina* peroxidase showing relatedness with peroxidases belonging to stress-tolerant plants. Evolutionary tree was constructed from distance data by using online

neighbor-joining method and represented using the Drawgram program of the PHYLIP tree package version 3.67

3D model of *A. marina* peroxidase

Theoretical models of *A. marina* peroxidase were constructed by the program MODELLER 9v3 using templates that gave highest sequence homology with the target sequence. Table 2 shows the overall summary of the

templates used for unbound model and models with bound substrates and the stereochemical quality of the models as assessed by PROCHECK (Laskowski et al. 1993). The Ramachandran plots of these models showed that the quality of each model was reasonably good. The ProSA energy plot illustrates the local model quality by plotting

Table 2 Homology models of *Avicennia marina* peroxidases

Models	Templates (pdb id)	Procheck statistics	RMSD (Å)
Unbound model	Barley grain peroxidase (1bgp)	271 residues (allowed) 1 residue (disallowed)	0.4
Ferulic acid-cyanide bound model	Horseradish peroxidase-bound ferulic acid-cyanide (7atj)	265 residues (allowed) 1 residue (disallowed)	0.5
Benzhydroxamic-acid-bound model	Horseradish peroxidase-bound benzhydroxamic acid (2atj)	265 residues (allowed) 1 residue (disallowed)	0.5
Formate-bound model	Horseradish peroxidase complexed with formate (1w4w)	265 residues (allowed) 1 residue (disallowed)	0.5
Carbon monoxide-bound model	Horseradish peroxidase complexed with carbon monoxide (1w4y)	265 residues (allowed) 1 residue (disallowed)	0.5
Tris(hydroxymethyl) aminomethane-bound model	TRIS-bound soy bean peroxidase (1fhf)	265 residues (allowed) 1 residue (disallowed)	0.36

energies as a function of amino acid sequence position. In case of the homology models of *A. marina* peroxidase, ProSA analysis revealed Z scores which lie in the low energy conformation states (Fig. 2). Moreover, the root mean square deviations for the C α coordinates of the models and the respective templates indicate that there is structural conservation between the folds (Table 2). The overall homology model of AP consists of 11 α helices and two β strands while remaining part of the structure comprises loops between the main secondary structure elements. A middle region along with the heme comprises proximal and distal domains and the two-cation-binding sites containing calcium ions (Fig. 3a). These calcium ions are linked to the heme-binding residues by a number of hydrogen bonds. Four disulfide bridges, 18cys–97cys, 51cys–56cys, 103cys–301cys, and 182cys–209cys were predicted by WHATIF (Chinea et al. 1998) (Fig. 3b).

Calcium-binding sites

Proximal site Cation binding is vital in maintaining heme environment and hence structural and functional properties (Morishima et al. 1986). In AP, the proximal calcium site is five coordinated with a mean coordination distance of 2.35 Å (Fig. 4). In case of barley grain (BGP), horseradish (HRP), and peanut peroxidase the proximal calcium site is seven-coordinated with a mean coordination distance of 2.4 Å (Schullar et al. 1996). The additional amino acid residue that provides coordination site is Asp225 while two coordination sites are provided by Thr228. Asp225 in AP is located too far away from the proximal calcium site and is located in the loop region having residues Val222, Arg223, Asn224, Arg226, and Gly227 in the vicinity (Fig. 4). These residues are variable in the plant peroxidase family (Table 1). With the exception of Ser176 and Lys231 which are replaced by Thr and Val, the rest of the amino acids that provide Ca coordination are conserved. These amino acids

superpose well and show same coordination with calcium. However, different hydrogen-bonding interactions were observed between these residues (Table 3a).

Distal site The presence of cation in the distal site is stabilized by the hydrogen-bonding network, thereby reducing the mobility of the catalytic distal histidine. This site is also five coordinated with the donor oxygen ligands provided by Asp50, Asp57, Val53, and Ser55 and an average mean coordination distance of 2.32 Å (Fig. 5a). All sites do not show same coordination strength. On the other hand, BGP is six coordinated with the additional site provided by Ser59 while Gly55 is present in place of Ser55. In AP, Ser59 show weak coordination to distal calcium as it lies at a distance of 3.5 Å (Fig. 5a). Different ionic and hydrogen-bonded interactions were also observed in AP (Table 3b, Fig. 5b).

Substrate binding and the role of active site residues

Unbound model At the distal site of the heme plane His49, Arg45, and Phe48 are conserved in all plant peroxidases. The distance between His49 (ND1) and Heme–Fe is 7.92 Å, which is similar as in case of BGP (8.2 Å). The interactions mainly the hydrogen-bonding pattern involving active site residues are shown in Fig. 6a. The region of greatest variation is the connecting loop comprising residues 65–83 present at the distal site of the enzyme. Arg45 is coordinated with heme forming three hydrogen bonds while it is also hydrogen bonded to His49 through water molecule. Arg45 also hydrogen bonds with Phe48 and prevents the substrate contact to the ferryl oxygen of compound I (Eq. 1), which is generated by the reaction between H₂O₂ and Fe (III). In BGP, Arg45, and His49 also show hydrogen-bonding interactions with other residues, i. e., Thr, Asn, and Thr at positions 67, 77, and 79,

Fig. 2 Protein Structure Analysis (ProSA) plots of homology models of *Avicennia marina* peroxidase; **a** unbound and bound with Cyanide and Ferulic acid **b** benzhydroxamic acid, Formate, and **c** Carbon Monoxide and Tris with z scores of -6.92, -6.35, -6.23, -6.86, -6.65 and -7.42 respectively. ProSA plot shows local model quality by

plotting energies as a function of amino acid sequence position. The program calculates an overall quality score for a specific input structure. The thick line shows the average energy over each 40-residue fragment while the thin line shows a smaller window size of 10 residues

respectively. On the other hand, these residues are replaced by Arg, Ser, and Gly, respectively, in AP. Therefore, no such interactions were observed in AP except that Arg67 hydrogen bonds with His49. However, a new interaction

was also observed in which Asn46 is hydrogen bonded to Asp50 which in turn is hydrogen bonded to Arg67. Glu71 is a conserved residue in the plant peroxidase superfamily; in AP it is involved in hydrogen-bonding interaction with

Fig. 2 (continued)

Arg67, Met68, and Thr74. In other peroxidases, Glu71 did not show such interaction and instead form another interaction in which it is involved along with Asn77 in hydrogen-bonding network connecting distal His to distal calcium whereas this network is not present in AP due to the presence of Ser at position 77.

The proximal site of the heme crevice of the class III peroxidases closely resembles one another. The proximal

His175 is coordinated with heme iron atom. It has been shown that His175 play a role in heme stability. It was demonstrated that His175 to Ala mutation caused heme degradation (Veitch 2004). In AP, the carboxylate side chain of Asp250 maintains imidazolate nature of His175 ring by forming two salt bridges (Fig. 6b). Hydrophobic interactions are also predicted in the vicinity of heme moiety because of the presence of Leu172 and Leu247 (Table 4). A number of salt bridges and hydrogen bonds were also predicted with His181 which is also present at the proximal site of Heme. It is hydrogen bonded to O1A atom of heme and also forms an ionic interaction with Asp215 and hydrogen bond with Val183 and Lys184 (Table 4). A different pattern of interaction is observed in BGP. An extended hydrogen-bonding network connects Asp225-Asp239-Leu224-Asp250 to proximal histidine through water molecule.

Ferulic-acid-cyanide-bound model (APFCN) The distal catalytic residues Arg45, His49, and Phe48 are present near the cyanide site and coordinate with the aromatic ring of the ferulic acid while Arg45, heme, and His49 are coordinated with the cyanide Arg45 forms hydrogen bonds with ferulic acid as well as with Phe41, Phe48, His49, Phe52, and Gly79. The catalytic water molecule is H bonded with ferulic acid, cyanide, His49, and Pro145 (Fig. 7a). It has been demonstrated experimentally that the flexibility of the aromatic binding pocket of peroxidase to accommodate ferulic acid resides in the aromatic nature of the substrate that interacts with the surrounding aromatic amino acids of the enzyme thus assisting the substrate for access in the substrate-binding pocket (Henriksen et al. 1999). In addition, Arg45 is also found to be actively involved in both substrate oxidation and ligand coordination (Henriksen et al. 1999). Coordination between Arg45 and cyanide play an important role in stabilization of peroxidase bound peroxy transition state and O–O bond cleavage. In general, it has been observed that coordination of substrate in the access channel motivates the distal site residues near the substrate (Henriksen et al. 1999).

Benzhydroxamic acid-bound model (APBHA) At the catalytic site of the predicted model of APBHA, Arg45, His49, and Pro145 are involved in H bonding with benzhydroxamic acid either directly or through water molecule (Fig. 7b). The distance between heme (Fe) and catalytic water molecule is 2.65 Å which is the same as in case of HRP and is enough to control the electronic structure of heme system (Howes et al. 1997). Crystallographic study has revealed that BHA was found at the peroxidase catalytic site, whereas hydrophobic residues are present near the aromatic portion of the substrate (Henriksen et al. 1998b). The conserved hydrophobic residues, Phe78, Gly79, and Leu144 surround the aromatic ring of APBHA.

Fig. 3 **a** The 3D homology model of *A. marina* peroxidase showing the secondary structural features. Also shown are the heme prosthetic group and calcium ions; **b** Cysteine residues involved in the four disulfide bridges

Formate-bound model (APFA) Formate coordinates directly in the heme pocket in a similar fashion as in case of HRP. Hydrogen-bonding interactions between formate, heme, and Arg45 are almost similar. Formate also forms hydrogen bonds with His49 while Arg45 in turn hydrogen bonds with Phe41 and His49 (Fig. 7c). It has been demonstrated that the channel providing access to the heme iron, involves site where formate binds (Carlsson et al. 2005). In the proposed mechanism for the formation of compound I (Eq. 1), hydrogen peroxide can readily be contained in that site forming hydrogen bonds with Arg45 and His49. A proton has to be removed from hydrogen peroxide and transferred to His49 which acts as proton acceptor (Berglund et al. 2002). This interaction involving His49 and formate is also predicted in the APFA model.

Carbon monoxide-bound model (APCO) Carbon monoxide coordinates with heme (Fe–C) at a distance of 2.04 Å and it binds perpendicularly to the heme plane with a Fe–C–O angle of 168.6°. It has been demonstrated that the geometry of carbon monoxide coordinating to the heme resembles the coordinating geometries of carbon monoxide in other heme proteins (i.e., Fe–C–O angle=1.58–1.74° and Fe–C distance=1.7–2.1 Å) (Berglund et al. 2002). Two hydrogen

bonds between *A. marina* peroxidase and bound CO were predicted; one with His49 and the other with Arg45. It has been demonstrated that all these contacts result in the slow dissociation of CO from the peroxidase, indicating that the complex is more stable as compared to other hemoprotein–CO complexes (Hayashi et al. 1976; Wittenberg et al. 1967). Furthermore, an extended hydrogen-bonding network was found upon CO binding with AP resulting in the change in conformation of catalytic distal residues (Fig. 7d).

Tris (hydroxymethyl)aminomethane-bound model (APTRIS) TRIS also coordinates with the distal catalytic residues that correspond to the residues involved in the structure of TRIS bound soybean peroxidase, SBP (Henriksen et al. 2001) and ferulic-acid-bound HRP (Henriksen et al. 1999). TRIS acts as the secondary substrate coordinating site in plant peroxidases thus providing a shuttle station for reducing substrates allowing water molecules and hydrogen peroxide to diffuse to the heme center (Henriksen et al. 2001). TRIS hydroxyl group coordinates with His49 through three catalytic water molecules while the second catalytic water molecule (W2) is coordinated to Pro145 and a third water molecule (W3; Fig. 7e). Arg45 has shown weak

Fig. 4 **a** Superpositions of *A. marina* peroxidase (gray) and barley grain peroxidase (red); loop region not superimposed very well is marked with arrow, **b** same loop region superimposed with amino acid

residues shown in different orientations, **c** superimposed region showing coordination with proximal calcium

coordination (3.26 Å) with W2 both in the APTRIS model as well as in the template, SBP. This interaction was found nearest to the TRIS hydroxyl group. Furthermore, no hydrogen-bonding interaction was seen between Arg45, Gly79, and heme in APTRIS while in SBP, Ser79 form hydrogen bonds with Arg45 and heme. It was also observed that in the APTRIS model, His181 (ND1) present within the 3 Å distance of heme is too far away (7.5 Å) from TRIS whereas Arg183 present at the equivalent position in SBP forms a long range interaction (3.8 Å) with TRIS (Henriksen et al. 2001). This is mainly due to the structural difference between the two amino acids; His being aromatic while Arg has an extended structure that is situated closer to the heme propionate. It was also observed that the distance between the catalytic water (W1) and heme (Fe; 3.8 Å) is relatively longer as compared to that found in other unbound peroxidases due to the presence of positively charged TRIS molecule present at the exterior substrate coordinating site. In the APTRIS model, Met80 (equivalent to Ile80 in SBP)

produces significant orientation of distal residues. Met44 is responsible for higher thermal stability in SBP (Henriksen et al. 2001), whereas equivalent Leu44 residue in APTRIS seems to have no distinct role in providing thermal stability to the protein.

Structural comparison of *A. marina* peroxidase with that of salt-sensitive and salt-tolerant plants

Structural superposition studies of the peroxidases from *A. marina* and barley grain (salt-tolerant), peanut and soybean (salt-sensitive) plant species have revealed that there are four regions of structural deviation (Fig. 8; Table 5). Regions having RMSD >1.0 in any of the species were considered as the regions of structural deviation and the corresponding regions in other species were compared. These regions were analyzed with respect to their functions and the role of significant residues was highlighted.

Table 3 Comparison between proximal and distal calcium-coordinating residues and hydrogen-bonding pattern in *Avicennia marina* peroxidase (AP) and barley grain peroxidase (BGP)

AP		BGP	
a Proximal site			
Ca coordination	H bonding	Ca coordination	H bonding
Ser176(O)-Ca	Ser176(N)-(O)Gly173	Thr180(O)-Ca	Thr180(N)-(O)Gly177
Ser176(OG)-Ca	Ser176(OG)-(O)Asp233	Thr180(OG)-Ca	Thr180(OG)-(O)Asp233
	Ser176(OG)-(OD1)Asp233		Thr180(OG)-(OD1)Asp233
No coordination with Asp225	Asp225(OD1)-(N)Arg226	Asp225(OD2)-Ca	Asp225(OD1)-(N)Arg227
	Asp225(OD1)-(N)Asp225		Asp225(OD1)-(N)Thr228
Thr228(O)-Ca	Thr228(N)-(O)Arg223	Thr228(O)-Ca	Thr228(N)-(OD1)Asp225
	Thr228(O)-(N)Lys231	Thr228(OG)-Ca	Thr228(OG)-(OD2)Asp233
Asp233(OD1)-Ca	Asp233(O)-(OG) Ser176	Asp233(OD1)-Ca	Asp233(O)-(OG)Thr180
Lys231(O)-Ca	Lys231(N)-(O)Thr228	Val231(O)-Ca	Val231(N)-(O)Thr228
b Distal site			
Asp50(O)-Ca	Asp50(N)-Asn46(O)	Asp50(O)-Na	Asp50(N)-Leu46(O)
Asp50(OD1)-Ca	Asp50(OD2)-Ala58(N)	Asp50(OD1)-Na	Asp50(OD1)-Ser59(N)
	Asp50(OD1)-Ser59(N)		Asp50(OD1)-Val60(N)
Val53(O)-Ca	Val53(N)-His49(O)	Val53(O)-Na	Val53(N)-His49(O)
Ser55(O)-Ca	Ser55(N)-Cys51(O)	Gly55(O)-Na	Gly55(O)-Asp57(N)
Asp57(OD1)-Ca	Asp57(N)-Ser55(O)	Asp57(OD1)-Na	Asp57(OD1)-Asp57(N)
	Asp57(OD1)-Asp57(N)		Asp57(OD1)-Ser59(N)
	Asp57(O)-Arg298(NH2)		
No coordination with Ser59	Asp57(OD1)-Ser59(N)	Ser59(OG)-Na	Ser59(OG)-Gln54(OE1)
	Ser59(OG)-Glu54(OE1)		Ser59(OG)-Val60(N)
	Ser59(O)-Leu62(O)		
	Glu54(OE2)-His143(NE2)		Gln54(OE1)-Asp143(OD1)

Region 64–85

This region has shown highest deviation, i.e., 2.2 Å in SBP while in case of PNP this region has an RMSD of 1.10 Å. The corresponding region in BGP is only slightly deviated (0.38 Å). This region also includes the connecting loop comprising residues 65–83 and is present at the distal site of the enzyme. In BGP, Thr67, Asn77, and Thr79 show hydrogen-bonding interactions with distal catalytic residues, Arg45 and His49, while no interaction was observed among these residues in AP except that Arg67 hydrogen bonds with His49 also because of sequence variation. Different substrates show variations in the hydrogen-bonding pattern in AP, e.g., in case of bound ferulic acid, Gly79 show hydrogen bond interaction with Arg45 while no hydrogen bonding was observed between these residues in case of bound TRIS. In SBP, Ser79 is present in place of Gly which also form hydrogen bonds with Arg45. Arg81 is present both in BGP and AP but is replaced by Ile in salt-sensitive species, SBP and PNP. The change from hydrophilic to hydrophobic residue may lead to a decrease in ionic interaction as the surface accessibilities of Arg81 in

AP was found to be 24.06 Å² as compared to Ile81 as 3.76 Å² and 8.49 Å² in PNP and SBP, respectively. Another sequence variation observed in this region is at position 85 where Tyr is present in AP. This residue is replaced by similar aromatic amino acid Phe in BGP while acidic amino acids, Glu and Asp are present in PNP and SBP, respectively. Again change in the nature of amino acid may lead to different interactions as the acidic amino acids were found to be located on the surface whereas aromatic amino acids were located in the interior of the protein.

Region 184–194

This region has shown highest deviation, i.e., 2.08 Å in PNP while in case of SBP this region has a low deviation with an RMSD of 0.64 Å. The corresponding region in BGP is only slightly deviated (0.1 Å). In AP, Lys184 is involved in hydrogen-bonding interaction with His181 at the proximal heme site which in turn forms another hydrogen bond with O1A atom of heme (Table 4). This residue is not conserved and is replaced even in BGP by Ser and in SBP and PNP by Thr and Ala, respectively. No

Fig. 5 **a** *A. marina* peroxidase region showing amino acid residues with the distal Ca; **b** Hydrogen bonding network between coordinating catalytic distal histidine and distal Ca

such interaction was therefore observed between residues at position 184 and heme. No other residue in this region seems to have any functional significance. The large structural deviation in this region in case of PNP seems to be because of the presence of Arg at position 186. This position is occupied by similar amino acid residues, Val and Ile in AP and SBP, respectively.

Region 210–216

This region has shown highest deviation, i.e., 1.04 Å in PNP while in case of SBP this region has a low deviation with an RMSD of 0.7 Å. The corresponding region in BGP is only slightly deviated (0.14 Å). No residue in this region seems to have any functional significance.

Region 222–227

This region has shown similar structural deviation of 2.06 Å in both SBP and PNP. The corresponding region in BGP is only slightly deviated (0.34 Å). In most of the members of peroxidase superfamily, Asp225 provides an additional coordination site for proximal calcium. But in case of AP, Asp225 does not coordinate with proximal calcium because it is located away from the latter having

different amino acid residues in the vicinity. This region has shown greater structural variation in case of PNP and SBP as compared to BGP although sequence variation was also observed in case of the latter. The large structural deviation in this region in case of PNP and SBP seems to be because of the presence of Pro at positions 222 and 225 whereas no proline is present in BGP and AP in this region.

Post-translational modification

Glycosylation

One Asn-linked glycosylation site was predicted by the PROSITE database (Obenauer et al. 2003) in the sequence patch, ¹⁴⁸Asn-Glu-Ser-Leu¹⁵¹ (Table 6). We analyzed this site in the homology model of peroxidase from *A. marina* (AP). In the AP model, Asn148 is present on the surface, being part of the loop structure. This analysis was consistent with the previous findings regarding the location of the glycosylated residues in different proteins. These residues should be present on the surface of the protein molecule in order to be glycosylated. They are more likely to be present in the loop structure emphasizing the role of these structures in molecular recognition. Moreover, it has also been documented that the position of proline residue in

Fig. 6 Hydrogen-bonding network between **a** distal and **b** proximal heme with catalytic residues in *Avicennia marina* peroxidase. Also shown are the ionic interactions between Asp250 and His175

the peptide segment is critical regarding the probability of glycosylation site. Pro can prevent glycosylation if it is present at the next position C-terminal to Ser/Thr or between Asn and Ser/Thr (Bause 1983). We compared this sequence patch with the proposed criteria for such a site in glycosylated proteins. Pro145 is present N-terminal to Asn148 but not between the patch, and no proline is found within the next 10 residues beyond Leu151. We have also found one recent experimental study that reported the presence of nine potential glycosylation sites in Royal palm tree peroxidase (Watanabe et al. 2010). Other studies have also demonstrated that glycans are required for the stable conformation of peroxidase and play significant role in the appropriate folding of the protein (Bause 1983). The glycosylation level of a protein can affect its structure, stability, and movement through the secretory pathway

Table 4 Coordinating catalytic residues with heme at the distal and proximal sites in *Avicennia marina* peroxidase (AP)

Heme distal site	Heme proximal site
His49(N)-(O)Arg45	Leu172(O)-(N)Ala174
His49(O)-(N)Val53-Ca	Ala174(O)-(N)Ser176-Ca
His49(ND1)-water-(NH2) Arg45	His175-heme
Heme-Fe-(ND1)His49	His175(ND1)-(OD1)Asp250
Heme-Fe-(NH1)Arg45	His175(N)-(O)Leu172
Arg45(O)-(N)Phe48	His181(N)-Heme
Phe48(O)-(N)Phe52	His181(O)-(N)Lys184 and (N) Val183
	Asp250(OD1)-(O)Tyr236
	Asp250(OD2)-(N)Leu247
	Asp250(N)-(O) Leu247
	Asp250(OD2)-(N)Leu246

(Bao et al. 2001). It has been shown that HRP contains eight asparagine-linked glycans located on the surface of the protein but these glycans hardly affect the thermodynamic stability, yet increased the kinetic stability significantly (Tams and Welinder 1998).

Phosphorylation

Protein kinases have profound effects on a cell and their activity is highly regulated. Kinases are turned on or off by phosphorylation, by binding of activator or inhibitor protein or small molecules, or by controlling their location in the cell relative to their substrates. Protein kinases have been shown to be activated by hyperosmotic stresses in different plant species. Ca-dependent protein kinases are a group of well-characterized protein kinases and have been found in many plants including *Arabidopsis*, soybean, rice, tobacco, and potato (Hrabak 2000). AP is also a salt-tolerant plant, in which several protein kinase C phosphorylation sites were predicted by PROSITE database (Obenauer et al. 2003) (Table 6). These sites were analyzed with respect to their positions in the 3D homology models. Some of the sites were found to have low surface accessibilities and therefore these sites may not be the potential phosphorylation sites. The remaining sites having relatively high surface accessibilities in the homology model were Ser64 in ⁶⁴Ser-Thr-Arg⁶⁶, Ser70 in ⁷⁰Ser-Glu-Lys⁷², Thr88 in ⁸⁸Thr-Ile-Lys⁹⁰, and Thr255 in ²⁵⁵Thr-Asp-Lys²⁵⁷ sequences patches with the surface accessibilities of 18.53 Å², 11.49 Å², 16.08 Å², and 16.37 Å², respectively.

Some potential phosphorylation sites picked up by the PROSITE search also include the phosphorylation sites that are specific for casein kinase II. Two of these sites were also potential targets for protein kinase C. The predicted

Fig. 7 The 3D homology models of *A. marina* peroxidase with various substrates. **a** Ferulic acid-cyanide (FCN), **b** benzhydroxamic acid (BHA), **c** formate (FA), **d** carbon monoxide (CO), and **e** Tris (hydroxymethylaminomethane) (TRIS)

Fig. 8 Superposition of *A. marina* peroxidase with peroxidases belonging to salt-sensitive plants. Regions labeled 1–4 are structurally variable having RMSD between 0.64 and 2.27 Å. Details of these regions are presented in text and Table 5

sites specific for casein kinase II that also showed increased surface accessibilities are Ser70 in ⁷⁰Ser-Glu-Lys-Glu⁷³, Thr88 in ⁸⁸Thr-Ile-Lys-Glu⁹¹, Ser135 in ¹³⁵Ser-Arg-Ala-Glu¹³⁸, Ser200 in ²⁰⁰Ser-His-Val-Glu²⁰³, and Thr293 in ²⁹³Thr-Lys-Gly-Glu²⁹⁶ sequences patches.

Amidation

The PROSITE pattern search results predicted two amidation sites, Thr-Gly¹²⁸-Arg-Arg and Asp-Gly¹³²-Arg-Lys in AP sequence that was consistent with the consensus pattern x-G-RK (Table 6). As the requisite for amidation, C-terminal glycine residue provide the amide group followed by at least two consecutive basic residues, which generally function as an active peptide precursor cleavage site (Bradbury and Smyth 1987). Analysis of the AP model showed that both the probable sites showed are buried (i.e., surface accessibility=0Å²) and therefore are less likely to be the amidation sites.

Conclusion

Plant cells have to survive constantly in the presence of various metabolites that are produced in response to various stress conditions. Mangroves that are salt-tolerant plants are better protected from oxidative damage under salt stress. They have evolved a complex series of enzymatic and non-enzymatic antioxidant protective mechanisms. Various genes

Table 5 Sequence patches that show structural deviation between *Avicennia marina* peroxidase (AP) and that of salt-sensitive (peanut, PNP; soybean, SBP) and salt-tolerant (barley grain, BGP) plant species

BGP	SBP	PNP
Overall RMSD of AP and BGP (0.38 Å)	Overall RMSD of AP and SBP (1.2 Å)	Overall RMSD of AP and PNP (1.2 Å)
64–85	64–85	64–85
AP: STRRMLSEKETDRSFGMRNFRY	AP: STRRMLSEKETDRSFGMRNFRY	AP: STRRMLSEKETDRSFGMRNFRY
BGP: GSATGPGEQQAPPNLTLPSPAF	SBP:NTDTIESEQDALPNINSIRGLD	PNP:DTSNFTGEKTAGPNANSIRGFE
(0.3 Å)	(2.27 Å)	(1.10 Å)
184–194	184–194	184–194
AP: KLVHRLYPEVD	AP: KLVHRLYPEVD	AP: KLVHRLYPEVD
BGP: SFEDRLFPRPD	SBP: TFINRLYNFSN	PNP: AFRTRIYNESN
(0.1 Å)	(0.64 Å)	(2.08 Å)
210–216	210–216	210–216
AP: PDPIPD	AP: PDPIPD	AP: PDPIPD
BGP: PAKGTDR	SBP: PQNATGD	PNP: PSVGGDT
(0.14 Å)	(0.7 Å)	(1.04 Å)
222–227	222–227	222–227
AP: VRNDRG	AP: VRNDRG	AP: VRNDRG
BGP: TVLDVR	SBP: LSTPDQ	PNP: PFDVTT
(0.34 Å)	(2.06 Å)	(2.06 Å)

Aligned regions of maximum deviation are shown; RMSD is shown in brackets

Table 6 Sequence patches that contain the predicted post-translational modification sites in *Avicennia marina* peroxidase as predicted by PROSITE

Glycosylation site	Phosphorylation sites	Amidation site
148–151 *NESL	Protein kinase C site	127–130 T*GRR
	64–66 *STR	131–134 D*GRK
	70–72 *SEK	
	88–90 *TIK	
	255–257 *TDK	
	Casein kinase II site	
	70–73 *STRE	
	88–91 *TIKE	
	135–138 *SRAE	
	200–203 *SHVE	
	293–296 *TKGE	

The predicted residue in the patch is denoted by asterisk (*). In case of glycosylation, phosphorylation, and amidation, Asn (N), Ser (S)/Thr (T), and Gly (G) are modified, respectively

responsible for providing oxidative and salt tolerance and ability to grow at elevated salt concentration have been identified. The present study is related to the sequence and structural analysis of peroxidase, a stress-related protein of *A. marina*. The structural features shows that Arg45 and His49 were catalytically significant residues as they have shown hydrogen-bonding interactions with all the substrates used in this study, i.e., ferulic acid-cyanide, benzhydroxamic acid, formate, carbon monoxide, and Tris while Phe40 also showed interaction in case of ferulic acid-cyanide. The connecting loop comprising residues 64–84 present at the distal site of the enzyme is the noticeable site for structural difference in peroxidases from various species. The structural comparison of the *A. marina* peroxidase (belonging to salt-tolerant plant) with that of two class III salt-sensitive species, peanut and soybean revealed four loop regions having largest structural deviation. The study also outlines some sites for post-translational modification in the peroxidase that play significant roles in the protein function. The structural and functional aspects presented here may further assist in understanding the role of *A. marina* peroxidase in resistance to oxidative and salt stress.

References

- Agostini E, Coniglio MS, Milrad SR, Tigier HA, Giulietti AM (2003) Phytoremediation of 2,4-dichlorophenol by *Brassica napus* hairy root cultures. *Biotechnol Appl Biochem* 37:139–144
- Almagro L, Gómez Ros LV, Belchi-Navarro S, Bru R, Ros Barcel A, Pedreño MA (2009) Class III peroxidases in plant defence reactions. *J Exp Botany* 60:377–390
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402
- Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh LS (2005) The universal protein resource (UniProt). *Nucleic Acids Res* 33:D154–D159
- Bao L, Shengwu M, van Huystee RB (2001) The effects of the site-directed removal of *N*-glycosylation from cationic peanut peroxidase on its function. *Arch Biochem Biophys* 386:17–24
- Bause E (1983) Structural requirements of *N*-glycosylation of proteins studies with proline peptides as conformational probes. *Biochem J* 209:331–336
- Berglund GI, Carlsson GH, Smith AT, Szołke H, Henriksen A, Hajdu J (2002) The catalytic pathway of horseradish peroxidase at high resolution. *Nature* 417:463–468
- Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, Feng Z, Gilliland GL, Iype L, Jain S, Fagan P, Marvin J, Padilla D, Ravichandran V, Schneider B, Thanki N, Weissig H, Westbrook JD, Zardecki C (2002) The protein data bank. *Acta Crystallogr D Biol Crystallogr* 58:899–907
- Bradbury AF, Smyth DG (1987) Biosynthesis of the C-terminal amide in peptide hormones. *Biosci Rep* 7:907–916
- Carlsson GH, Nicholls P, Svistunenko D, Berglund GI, Hajdu J (2005) Complexes of horseradish peroxidase with formate, acetate, and carbon monoxide. *Biochemistry* 44:635–642
- China RG, Lopez N, Pons T, Vriend G (1998) Homology modeling, model and software evaluation: three related resources. *CABIOS* 14:523–528
- Cosio C, Dunand C (2009) Specific functions of individual class III peroxidase genes. *J Exp Bot* 60:391–408
- Costa MMR, Hilliou F, Duarte P, Pereira LG, Almeida I, Leech M, Memelink J, Barcelo AR, Sottomayor M (2008) Molecular cloning and characterization of a vacuolar class III peroxidase involved in the metabolism of anticancer alkaloids in *Catharanthus roseus*. *Plant Physiol* 146:403–417
- Dalton DA (1991) Ascorbate peroxidases. *Peroxidases in chemistry and biology*. CRC, Boca Raton, pp 139–153
- Felsenstein J (1989) PHYLIP—phylogeny inference package (Version 3.67). *Cladistics* 5:164–166
- Gajhede M, Schuller DJ, Henriksen A, Smith AT, Poulos TL (1997) Crystal structure of horseradish peroxidase C at 2.15 Å resolution. *Nat Struct Biol* 4:1032–1038
- Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18:2714–2723
- Hayashi Y, Yamada H, Yamazaki I (1976) Heme-linked proton dissociation of carbon monoxide complexes of myoglobin and peroxidase. *Biochim Biophys Acta* 427:608–616
- Henriksen A, Mirza O, Indiani C, Teilum K, Smulevich G, Welinder KG, Gajhede M (2001) Structure of soybean seed coat peroxidase: a plant peroxidase with unusual stability and haem-apoprotein interactions. *Protein Sci* 10:108–115
- Henriksen A, Schuller DJ, Meno K, Welinder KG, Smith AT, Gajhede M (1998a) Structural interactions between horseradish peroxidase C and the substrate benzhydroxamic acid determined by X-ray crystallography. *Biochemistry* 37:8054–8060
- Henriksen A, Smith AT, Gajhede M (1999) 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* 274:35005–35011
- Henriksen A, Welinder KG, Gajhede M (1998b) Structure of barley grain peroxidase refined at 1.9-Å resolution. *J Biol Chem* 273:2241–2248
- Howe K, Bateman A, Durbin R (2002) QuickTree: building huge neighbour-joining trees of protein sequences. *Bioinformatics* 18:1546–1547

- Howes BD, Rodriguez-Lopez JN, Smith AT, Smulevich G (1997) Mutation of distal residues of horseradish peroxidase: influence on substrate binding and cavity properties. *Biochemistry* 36:1532–1543
- Hrabak EM (2000) Calcium-dependent protein kinase and their relatives. In: Krie M, Walker J (eds) *Advances in botanical research incorporating advances in plant pathology*. Academic, New York, pp 185–223
- Koua D, Cerutti L, Falquet L, Sigrist CJ, Theiler G, Hulo N, Dunand C (2009) PeroxiBase: a database with new tools for peroxidase family classification. *Nucleic Acids Res* 37:D261–D266
- Kumar S, Jaggi M, Taneja J, Sinha AK (2011) Cloning and characterization of two new Class III peroxidase genes from *Catharanthus roseus*. *Plant Physiol Biochem* 49:404–412
- Laberge M, Huang Q, Stenner RS, Fidy J (2003) The endogenous calcium ions of horseradish peroxidase C are required to maintain the functional nonplanarity of the heme. *Biophys J* 84:2542–2552
- Laskowski RA, McAurthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Cryst* 26:283–291
- Lavid N, Schwartz A, Yarden O, Tel-Or E (2001) The involvement of polyphenols and peroxidase activities in heavy-metal accumulation by epidermal glands of the waterlily (Nymphaeaceae). *Planta* 212:323–331
- Morishima I, Kurono M, Shiro Y (1986) Presence of endogenous calcium ion in horseradish peroxidase. *J Biol Chem* 261:9391–9399
- Obenauer JC, Cantley LC, Yaffe MB (2003) Scansite 20: proteome-wide prediction of cell signaling interactions using short sequence motifs. *Nucleic Acids Res* 31:3635–3641
- Passardi F, Longet D, Penel C, Dunand C (2004) The class III peroxidase multigene family in rice and its evolution in land plants. *Phytochemistry* 6:1879–1893
- Passardi F, Theiler G, Zamocky M, Cosio C, Rouhier N, Teixeira F, Margis-Pinheiro M, Ioannidis V, Penel C, Falquet L, Dunand C (2007) PeroxiBase: the peroxidase database. *Phytochemistry* 68:1605–1611
- Reddy CA, D'Souza TM (1994) Physiology and molecular biology of the lignin peroxidases of *Phanerochaete chrysosporium*. *FEMS Microbiol Rev* 13:137–152
- Sali A, Blundell TL (1993) Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol* 234:779–815
- Schullar DJ, Ban N, Huystee RB, Mcpherson A, Poulos TL (1996) The crystal structure of peanut peroxidase. *Structure* 4:311–321
- Sippl MJ (1993) Recognition of errors in three-dimensional structures of proteins. *Proteins* 17:355–362
- Sundaramoorthy M, Kishi K, Gold MH, Poulos TL (1994) The crystal structure of manganese peroxidase from *Phanerochaete chrysosporium* at 206 Å resolution. *J Biol Chem* 269:32759–32767
- Tams JW, Welinder KG (1998) Glycosylation and thermodynamic versus kinetic stability of horseradish peroxidase. *FEBS Lett* 421:234–236
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876–4882
- Veitch NC (2004) Horseradish peroxidase: a modern view of a classic enzyme. *Phytochemistry* 65:249–250
- Watanabe L, de Moura PR, Bleicher L, Nascimento AS, Zamorano LS, Calvete JJ, Sanz L, Pérez A, Bursakov S, Roig MG, Shnyrov VL, Polikarpov I (2010) Crystal structure and statistical coupling analysis of highly glycosylated peroxidase from royal palm tree (*Roystonea regia*). *J Struct Biol* 169:226–242
- Welinder KG (1991) Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. *Biochim Biophys Acta* 1080:215–220
- Welinder KG, Justesen AF, Kjærsgård IVH, Jensen RB, Rasmussen SK, Jespersen HM, Duroux L (2002) Structural diversity and transcription of class III peroxidases from *Arabidopsis thaliana*. *Eur J Biochem* 269:6063–6081
- Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:W407–W410
- Wittenberg BA, Antonini E, Brunori M, Noble RW, Wittenberg JB, Wymann J (1967) Studies on the equilibria and kinetics of the reactions of peroxidases with ligands. The dissociation of carbon monoxide from carbon monoxide ferro-horseradish peroxidase. *Biochemistry* 6:1970–1974