Journal of Molecular Graphics and Modelling 79 (2018) 15–26

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Topical Perspectives

Comparative modeling and molecular docking analysis of white, brown and soft rot fungal laccases using lignin model compounds for understanding the structural and functional properties of laccases

Ayyappa Kumar Sista Kameshwar, Richard Barber, Wensheng Qin*

Department of Biology, Lakehead University, 955 Oliver Road, Thunder Bay, Ontario, P7 B 5E1, Canada

ARTICLE INFO

Article history:
Received 14 July 2017
Received in revised form 25 October 2017
Accepted 25 October 2017

Keywords:
Multicopper oxidase
Laccases
White rot
Brown rot
Soft rot fungi
Homology modeling
Molecular docking

ABSTRACT

Extrinsic catalytic properties of laccase enable it to oxidize a wide range of aromatic (phenolic and non-phenolic) compounds which makes it commercially an important enzyme. In this study, we have extensively compared and analyzed the physico-chemical, structural and functional properties of white, brown and soft rot fungal laccases using standard protein analysis software. We have computationally predicted the three-dimensional comparative models of these laccases and later performed the molecular docking studies using the lignin model compounds. We also report a customizable rapid and reliable protein modelling and docking pipeline for developing structurally and functionally stable protein structures. We have observed that soft rot fungal laccases exhibited comparatively higher structural variation (higher random coil) when compared to brown and white rot fungal laccases. White and brown rot fungal laccase sequences exhibited higher similarity for conserved domains of *Trametes versicolor* laccase, whereas soft rot fungal laccases shared higher similarity towards conserved domains of *Melanocarpus albomyces* laccase. Results obtained from molecular docking studies showed that aminoacids PRO, PHE, LEU, LYS and GLN were commonly found to interact with the ligands. We have also observed that white and brown rot fungal laccases showed similar docking patterns (topologically monomer, dimer and trimer bind at same pocket location and tetramer binds at another pocket location) when compared to soft rot fungal laccases. Finally, the binding efficiencies of white and brown rot fungal laccases with lignin model compounds were higher compared to the soft rot fungi. These findings can be further applied in developing genetically efficient laccases which can be applied in growing biofuel and bioremediation industries.

1. Introduction

Laccase (EC 1.10.3.2) is highly studied commercially important enzyme representing the major subgroup of multicopper oxidase (MCO) family, widely distributed among bacteria (prokaryotes), fungi and plants (eukaryotes) [1]. The function of laccases varies widely based on their host organisms, in plants it is involved in lignin biosynthesis, where as in fungi and bacteria it is involved in lignin degradation [1,2]. It was first discovered in the sap of plants (*Rhus vernicifera*) [3] and later it was demonstrated in fungi [4]. Other enzymes belonging to multicopper oxidases (copper containing enzymes) family are ferroxidase (EC 1.16.3.1), ascorbate oxidase (EC 1.10.3.3), ceruloplasmin monoxygenases, dioxygenases and various manganese oxidases [5]. Multicopper oxidase family enzymes usually found to contain one to six copper atoms per molecule, with the aminoacids ranging between 100–1000 per a single peptide chain [5,6]. Laccases are characterized by the presence of four catalytic copper atoms: the T1 copper site and the T2/T3 trinuclear copper cluster [7]. Substrate oxidation occurs at the T1 copper due to its high redox potential (up to +800 mV). The one electron substrate oxidation is coordinated with the four electron reduction of molecular oxygen at the T2/T3 cluster; oxidation of four substrates is necessary for complete reduction of molecular oxygen to water [7].

Laccases extensively uses the redox ability of copper ions for oxidation of various aromatic substrates concomitantly reducing the molecular oxygen to water [2,8]. Laccases directly oxidize ortho, *para*-diphenols, aminophenols, polyphenols, polyamines, aryl diamines and also some inorganic ions [2]. The use of the laccase mediator system allows for oxidation of non-phenolic compounds and substrates too large to bind to the active site [9–12]. A mediator is a low molecular weight compound (acting as elec-

* Corresponding author.
E-mail address: wqin@lakeheadu.ca (W. Qin).

https://doi.org/10.1016/j.jmgm.2017.10.019
1093-3263/

tron shuttle) with higher redox potential than the T1 copper (>900 mV) [13]. The most common laccase mediators used are 2,2′-azino-bis (3-ethybenzthizoline-6-sulfonic acid) (ABTS) and triazole 1-hydroxybenzotriazole (HBT). The mediator is initially oxidized at the T1 site, generating a strong oxidizing intermediate, which then diffuses out of the active site and oxidizes the substrate [13]. In this way, the laccase mediator acts as an electron transport shuttle. Laccases typically show low substrate-specificity, and the range of substrates oxidized can vary between laccases. Oxidizing ability of laccases also depends on the nature of substrate whether it is monomeric, dimeric, or tetrameric [13]. Possible substrates of laccases include polyphenols, methoxy-substituted phenols, aromatic amines, and ascorbate [14].

The comparative modeling of fungal and bacterial laccases was reported in the past, however studies on fungal laccases typically focused on white rot basidiomycetes due to the extrinsic lignolytic abilities. Rivera-Hoyos et al., reported the three dimensional (3D) homology models of white rot fungal (*Ganoderma lucidum* and *Pleurotus ostreatus*) laccase proteins, which revealed the laccase interactions with ABTS [15]. The 3D homology models of white rot fungi *Pycnoporus cinnabarinus* [16], *Lentinula edodes* [17], were reported earlier. Tamboli et al. has compared physio-chemical properties of bacterial and fungal (*Cryphonectria parasitica, Ganoderma lucidum, Phomopsis liquidambaris, Pycnoporus coccineus and Trametes sanguine*) laccases and generated the 3D comparative models of bacterial and fungal laccase proteins which can be used for molecular docking studies [18]. Molecular docking studies with fungal laccases were performed using various chemical substrates such as ABTS [15] and also with lignin model compounds such as sinapyl alcohol, dimer, trimer and tetramer [19] were reported earlier. However, studies comparing the structural and functional properties of white, brown and soft-rot fungal laccases were not been reported till today. As these fungi exhibit differential wood decaying properties white rot (can efficiently degrade lignin, cellulose and hemicellulose), brown rot fungi (efficient cellulose, hemicellulose degrading with lignin modifying) and soft rot fungi (exhibits partial decaying abilities). It would be interesting to understand the structural and functional differences among the laccases of these fungi.

In our present study, we have reported the three-dimensional homology models of the selected white, brown and soft fungal laccase protein sequences retrieved from public repositories and extensively discussed about their structural and functional properties using standard tools. Using a set of lignin model compounds (monomers, dimer, trimer and tetramer) we have performed the molecular docking experiments. Results obtained in our study demarcates the structural and functional properties of white, brown and soft rot fungi and highlights the significant aminoacids which are involved in its catalysis. These results can be further applied for designing and developing recombinantly efficient laccases having wide range applications in clinical, chemical, environmental and industrial sectors.

2. Materials and methods

2.1. Selection and retrieval of laccase protein sequences

Laccase protein sequences of six different fungi viz., *Phlebia brevispora* HHB-7030 SS6 v1.0 [20], *Dichomitus squalens* CBS463.89 (White rot), *Fomitopsis pinicola* FP-58527 SS1 [21], *Wolfiporia cocos* MD-104 SS10 [21] (Brown rot) and *Chaetomium globosum* v1.0 [22], *Cadophora* sp. DSE1049 (Soft rot), were retrieved from JGI (Joint Genome Institute) MycoCosm database. The *D. squalens* CBS463.89 and *Cadophora* sp. DSE1049 laccase protein sequences (Dsqual_59186 and Cadophora_560981) were produced by the "US Department of Energy Joint Genome Institute http://www.jgi.doe.gov/in collaboration with the user community". We have used CAZy (Carbohydrate active enzymes), KOG (Eukaryotic orthologous groups) and GO (Gene Ontology) tools of JGI MycoCosm database during the retrieval of laccase protein sequences. Initially, we have retrieved a total of 56 laccase protein sequences (*P. brevispora* (5), *D. squalens* (12), *F. pinicola* (6), *W. cocos* (4), *C. globosum* (6) and *Cadophora* sp. (22)) respectively, from JGI MycoCosm database. All the retrieved laccase protein sequences from each organism was queried through BLAST against protein data bank (PDB) database using PSI-BLAST algorithm a variation of BLAST (sensitive to low-similarity, provides biologically relevant sequences and three times faster than regular BLAST) [23]. Laccase protein sequences showing highest sequence similarity and query coverage was designated as the template for the comparative modeling studies.

2.2. Phylogenetic analysis

All the retrieved laccase protein sequences of each organism were aligned using ClustalW algorithm (fast, accurate, and robust method, which uses a residue comparison matrix and position specific gap penalties to align sequences) of MEGA v7 software [24]. The ClustalW aligned sequences were considered for the construction of phylogenetic trees using Neighbour Joining method and Bootstrap resampling of 1000 replicates parameters were used for the estimation of phylogenetic tree topologies [25]. The phylogenetic trees were constructed for both intra and inter organism level to determine the laccase target sequences which are closely related to the template during the evolution.

2.3. Physico-chemical properties of selected laccases

Physico-chemical properties of above selected laccase protein sequences were determined using the ExPASy ProtParam tool [26]. Our analysis included the parameters such as aminoacid composition, number of positively (+R) and negatively (-R) charged aminoacid residues, predicted molecular weight, theoretical isoelectric point (pI), extinction coefficient (EC) [27], instability index (Ii) [28], aliphatic index (Ai) [29] and GRAVY (grand average hydropathicity) [30].

2.4. Structural and functional properties of laccases

The above selected laccase protein sequences were studied for their structural and functional properties for which we have used SOPMA (Self-optimized prediction method with alignment) tool for determining secondary structure elements [31]. We have used Motif Scan web server to identify the well-known motif sequences using the motif sources such as PeroxiBase, HAMAP, PROSITE patterns and profiles, More profiles, Pfam HMM (both local and global) profiles [32]. To understand the cellular localization of selected laccases, the protein sequences were subjected CELLO v2.5 web server [33]. We have used EDBCP (Ensemble-based Disulfide Bonding Connectivity Pattern) for understanding the presence of cysteine residues and to predict the most possible disulfide (S—S) bonds [34]. To predict the location and presence of signal peptide cleavage sites the protein sequences of laccases were analyzed using SignalP v4.1 web server [35]. And to predict the presence of transmembrane helices we have analyzed the selected protein sequences using TMHMM v2.0 web server (http://www.cbs.dtu.dk/services/TMHMM/). Acetylation of the selected fungal laccase proteins are assessed using the NetAcet v1.0 web server [36].

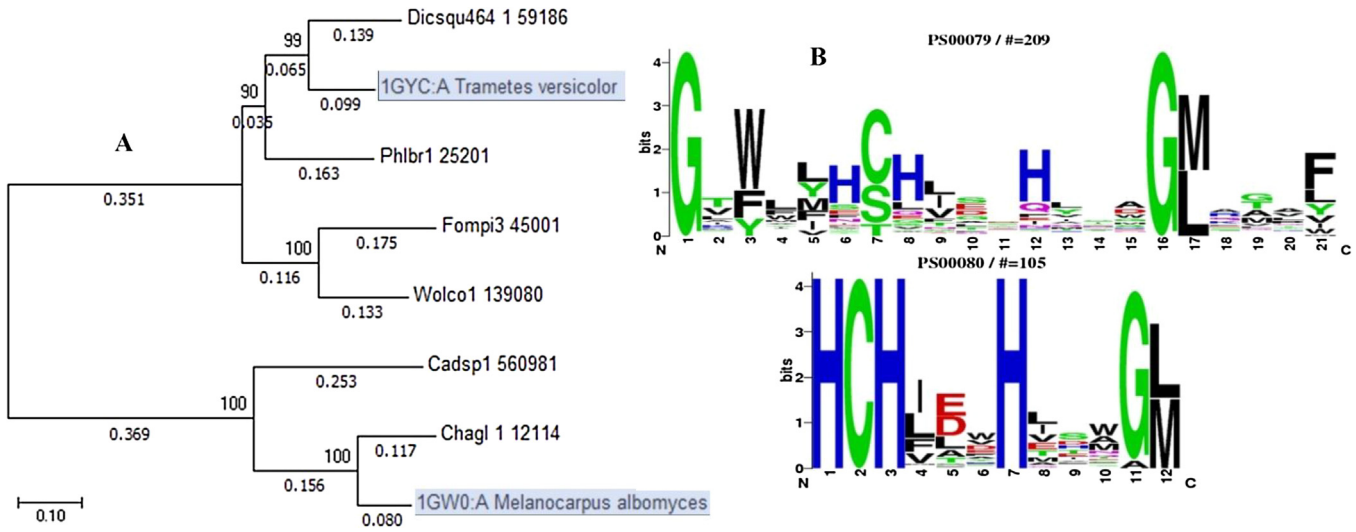


Fig. 1. Shows the analysis of fungal laccase protein sequences A) phylogenetic analysis of laccase protein sequence (*P. brevispora, D. squalens, F. pinicola, W. cocos, C. globosum*) and experimentally determined laccases *T. versicolor* (1GYC) and *M. albomyces* (1GW0), B) sequence logos of MCO signature 1 (PS00079) and 2 (PS00080) patterns.

Table 1
Lists the physico-chemical properties of laccase protein sequences calculated from ExPASy ProtParam and EDBCP tools:.

Organism (protein-ID)	Length	M.W*	pI*	EC	-R	+R	Ii	Ai	GRAVy	#. of Cys	Predicted S—S bonds
Phlebia brevispora (25201)	520	55977.21	5.20	69705	45	23	31.55	90.04	0.082	7	106–509, 138–227
Dichomitus squalens (59186)	520	56477.01	4.63	71195	48	22	28.04	84.79	−0.044	6	106–509, 138–226
Fomitopsis pinicola (45001)	539	58990.52	4.75	67185	55	26	34.35	88.85	−0.055	7	109–513, 141–228, 349–352
Wolfiporia cocos (139080)	479	52277.43	4.46	68550	54	18	40.44	83.47	−0.135	5	64–468, 96–183
Chaetomium globosum (12114)	619	68515.16	6.11	120820	55	45	35.90	74.44	−0.295	8	51–344, 112–378, 161–586
Cadophora sp. DSE1049 (560981)	582	63262.99	6.88	127810	34	33	38.02	75.74	−0.231	8	26–34, 135–563, 327–361

Note: M.W* = Molecular weight in Daltons, pI* = Theoretical pI values, EC = Extinction coefficient (M^{-1} cm^{-1}), −R = number of negatively charged amino acids, +R = number of positively charged amino acids, Ii = Instability index, Ai = Aliphatic index, GRAVy = grand average of hydropathy, #. Of Cys = Number of cysteine residues.

2.5. Initial protein model generation and refinement

Laccase protein sequences exhibiting highest sequence similarity and highest query coverage (obtained from the BLAST analysis), were further considered as input to SWISS-MODEL protein prediction web server (generates tertiary protein structure from a queried protein sequence based on the SWISS-MODEL template library), for the generation of initial unrefined 3D modelled structure of the target protein [37]. The 3D structures of modelled laccase proteins were further refined using GalaxyRefine (method improves both local and global qualities of template-based predicted protein structures) [38] and KoBaMIN (knowledge based minimization method, which minimizes the potential of mean force derived from the experimental structures of PDB) [39] web servers. GalaxyRefine server optimizes the side chain conformations and performs the energy minimization on each conformation and subsequently relaxes the overall protein structure through molecular dynamic simulation methods. Whereas KoBaMIN refined structures are stereochemically optimized with MESHI software, KoBaMIN energy function indirectly includes the effects of both solvent interactions and the crystal environments. The refined laccase protein structures were validated to assess the stereochemical quality, using a set of software such as PROCHECK [40], RAMPAGE [41], ERRAT [42] and PROQ [43].

2.6. Preparation of ligands (Lignin model compounds)

The NMR based structural studies of lignin and plant cell wall compounds explained by Ralph et al., were derived for the present study [44]. The lignin model compounds considered for the present study were monomers (sinapyl alcohol, coniferyl alcohol and *p*-coumaryl alcohol), dimer (guaiacyl 4-O-5 guaiacyl), trimer (syringyl β-O-4 syringyl β-O-4 sinapyl alcohol) and tetramer (guaiacyl β-O-4 syringyl β-β syringyl β-O-4 guaiacyl) were sketched using ChemDraw Ultra v7.0. and later internally transferred to Chem3D Pro v7.0 software and further saved the structure in protein data bank format. The ligands were subjected for energy minimization before using it for the protein docking simulation experiments using AutoDock software [45].

2.7. Protein docking of refined models

The refined and validated 3D modelled laccase proteins were used for further protein docking studies. We have used AutoDock Tools v1.5.6 and AutoDock Vina v1.1.2 [46] for the simulated protein docking experiments. The refined protein model in PDB file format were initially opened in AutoDock Tools and following functions were performed a) add all hydrogens b) merge non-polar hydrogens c) compute Gasteiger charges d) finally, save the protein model in PDBQT format. The ligand (lignin model compounds) are loaded into AutoDock Tools and following functions were performed a) detect root using the option torsion tree b) the number of torsions were set to maximum c) then finally save the ligand in PDBQT format. Later we have prepared the AutoDock Vina configuration files for all the laccase modelled structures and ligands, the above prepared files were used for performing the protein docking analysis. Ligand docking was performed using AutoDock Vina which is installed on an instance of the Galaxy platform based on In-house High-Performance Computing Cluster (LUHPCC). We have performed blind docking with all the lignin ligand models with the exhaustiveness set to 32. The best-fit ligand conformations were selected based on their minimum binding energies.

Aminoacid residues which established a contact with ligand and the residues which are involved in hydrogen bonding with ligands were recorded using AutoDock Tools using the results obtained from AutoDock Vina. The validated 3D laccase modelled structures were compared using the root mean square deviation (RMSD) using SWISS-PDB viewer v4.1 [47]. We have used Edu PyMOL v1.7.4 (https://pymol.org/educational) for visualizing the interactions of the ligand and modelled protein structure and for developing the respective docked images.

3. Results

3.1. Sequence retrieval, analysis and physicochemical properties

We have retrieved genome wide laccase encoding protein sequences from *P. brevispora* and *D. squalens* (white rot fungi), *F. pinicola* and *W. cocos* (brown rot fungi), and *C. globosum* and *Cadophora* sp. (soft rot fungi). Total of 56 laccase protein sequences from *P. brevispora* (5), *D. squalens* (12), *F. pinicola* (6), *W. cocos* (4), *C. globosum* (6) and *Cadophora* sp. (22), were retrieved from JGI MycoCosm database. The laccase protein sequences exhibiting highest sequence similarities with the known laccase structures from PDB database upon BLASTP search, were considered for the present study. The following protein sequences of fungi exhibiting higher sequence identities were used for the modeling studies: Phlbr1-25201, Dsqual-59186, Fompi3-45001, Wcocos-139080, Chagl1-12114 and Cadophora-560981 Table-S2).

The selected laccase protein sequences were found to contain about 520–619 aminoacid residues with an exception of 479 (*W. cocos*), with a theoretical molecular weight and pI ranging between 52277.43–68515.16 Da and 4.46–6.88 respectively (Table 1) (Fig. S1). All the above fungal laccase protein sequences were found to be stable with an instability index value ranging between 28.04–38.02, with an exception of *W. cocos* laccase protein instability index of (40.44) (Table 1). Aliphatic index values of the laccase proteins were under the range of 74.44–90.4. The grand average of hydropathicity (GRAVY) index values of the laccase proteins were mostly negative indicating the hydrophilic nature of these proteins, with an exception of *P. brevispora* laccase protein which gave a positive value indicating its hydrophobic nature and the hydropathicity plots generated using DiscoveryStudio 2016 Client® were also found to be in accordance with the GRAVY results, these plots were listed in (Fig. S4) supplementary information S1 (Table 1). The selected laccase protein sequences exhibited higher evolutionary similarities with the template protein sequences (Fig. S2). All the above considered laccase protein sequences were found to contain sequences coding for cupredoxin superfamily conserved domains specifically trinuclear Cu binding site (CuRO_1_Tv-LCC_like), domain 3 interface (CuRO_2_Tv-LCC_like) and type-1 copper binding site (CuRO_3_Tv-LCC_like) with an exception of *W. cocos* laccase only CuRO_3_Tv-LCC_like domain sequences (Fig. S3). Laccase protein sequences of soft rot fungi contain cupredoxin superfamily conserved domains matching to *Melanocarpus albomyces*, whereas white and brown rot fungi possessed cupredoxin superfamily conserved domains matching to *Trametes versicolor* (Fig. 1) (Fig. S3).

The selected laccase protein sequences were also subjected to series of analysis using SignalP (for the detection of signal peptides cleavage sites), TMHMM (for detecting transmembrane helices in proteins) and NetAcet (for predicting N-acetyltransferase A substrates) webservers. According to Nakai, the signal peptide ranges between the 15–40 aminoacid residues, required for the protein secretion and eventually these sites are cleaved from the mature protein [48]. We have observed that all the laccases except *W. cocos* (13980) possessed a signal peptide, which supports that these laccases are secretory proteins (Fig. S6). The amino acid sequence involved in formation of transmembrane helices were only found to be present in *D. squalens* and these results explains that majority of aminoacid residues occur in the outside region (Fig. S7, Table S3). Results obtained from CELLO v2.5 subcellular location predictor showed that the laccase proteins of *P. brevispora* (3.981), *D. squalens* (4.644), *F. pinicola* (3.688), *W. cocos* (3.871), *C. globosum* (3.995) and *Cadophora* (3.782) where located in the extracellular regions of fungi. All the selected laccase protein sequences showed *N*-Myristoylation, *N*-glycosylation, Amidation (except *Cadophora* sp.), Casein kinase-II phosphorylation site, protein kinase c phosphorylation site and multi copper oxidase type-1&2 (Cu_oxidase_1, 2 and 3) motif sites (Table 2) (Table S1). Analysis from NetAcet v1.0 server has showed that laccase proteins of *P. brevispora* and *Cadophora* sp. possess the substrates for N-acetyltransferases with values of 0.511 and 0.522 respectively. Whereas laccase protein sequences of *D. squalens* (0.491), *F. pinicola* (0.478), *W. cocos* (0.483) and *C. globosum* (no Ala, Gly, Ser or Thr at positions 1–3) do not show any possible substrates for N-acetyltransferases. We have also analyzed the number of cysteine residues present and reported the number of predicted disulfide bridges in the selected laccase protein sequences using EDBCP web server. We have observed that laccase protein sequences of *P. brevispora, D. squalens* and *W. cocos* contain two disulfide bridges, laccase protein sequences of *F. pinicola, C. globosum* and *Cadophora* sp. contain three disulfide bridges. The secondary structure elements of the selected laccase proteins were calculated using the SOPMA web server. These results show that higher percentage of aminoacids were found to be involved in formation of random coils, the alpha helical content of soft rot fungi is less when compared to brown and white rot fungi (Table 3).

Table 2
Computationally predicted motifs in *P. brevispora, D. squalens, F. pinicola, W. cocos, C. globosum* and *Cadophora* laccase protein sequences obtained from the MOTIF SCAN server:.

Organism, Protein-ID	Motif Description/Aminoacid residues (# of sites)
P. brevispora (25201)	Multi copper oxidase 1: 125–145 (1)
	Multi copper oxidase (Cu_oxidase): 163–307 (1)
	Multi copper oxidase (Cu_oxidase_2): 369–494 (1)
	Multi copper oxidase (Cu_oxidase_3): 30–152 (1)
D. squalens (59186)	Multi copper oxidase 1: 125–145 (1)
	Multi copper oxidase (Cu_oxidase): 163–305 (1)
	Multi copper oxidase (Cu_oxidase_2): 365–494 (1)
	Multi copper oxidase (Cu_oxidase_3): 30–152 (1)
F. pinicola (45001)	Multi copper oxidase 1: 128–148 (1)
	Multi copper oxidase (Cu_oxidase): 166–310 (1)
	Multi copper oxidase (Cu_oxidase_2): 372–497 (1)
	Multi copper oxidase (Cu_oxidase_3): 33–155 (1)
W. cocos (139080)	Multi copper oxidase 1: 83–103 (1)
	Multi copper oxidase (Cu_oxidase): 121–267 (1)
	Multi copper oxidase (Cu_oxidase_2): 327–453 (1)
	Multi copper oxidase (Cu_oxidase_3):1–110, 10–110 (1)
C. globosum (12114)	Multi copper oxidase 1, 2: 543–563, 548–559
	Multi copper oxidase (Cu_oxidase): 212–361 (1)
	Multi copper oxidase (Cu_oxidase_2): 427–569 (1)
	Multi copper oxidase (Cu_oxidase_3): 87–206 (1)
Cadophora sp. (560981)	Multi copper oxidase 1, 2: 523–543, 528–539 (2)
	Multi copper oxidase (Cu_oxidase): 186–344 (1)
	Multi copper oxidase (Cu_oxidase_2): 410–549 (1)
	Multi copper oxidase (Cu_oxidase_3): 61–180 (1)

3.2. Homology modeling and model validation

The selected laccase protein sequences were used to develop the corresponding 3D protein structures using SWISS-MODEL server. In SWISS-MODEL server the best suitable template protein structures for the target protein sequence were identified using BLAST and HHblits. The server will also evaluate the global and local qualities of the modelled protein structure using QMEAN (a complex

Table 3
Computationally predicted secondary structure elements of laccase protein sequences calculated using SOPMA web server.

Protein-ID	Alpha helix (%)	Extended Strand (%)	Beta turn (%)	Random coil (%)
Phlbr1_25201	15.96 (83)	30.77 (160)	10.96 (57)	42.31 (220)
Dsqual1_59186	15.38 (80)	31.92 (166)	11.73 (61)	40.96 (213)
Fompi3_45001	16.70 (90)	29.31 (158)	9.09 (49)	44.09 (242)
Wcocos _139080	10.23 (49)	31.94 (153)	10.65 (51)	47.18 (226)
Chagl1_12114	12.12 (75)	33.12 (205)	10.82 (67)	43.94 (272)
Cadoph_560981	9.79 (57)	28.01 (163)	9.28 (54)	52.92 (308)

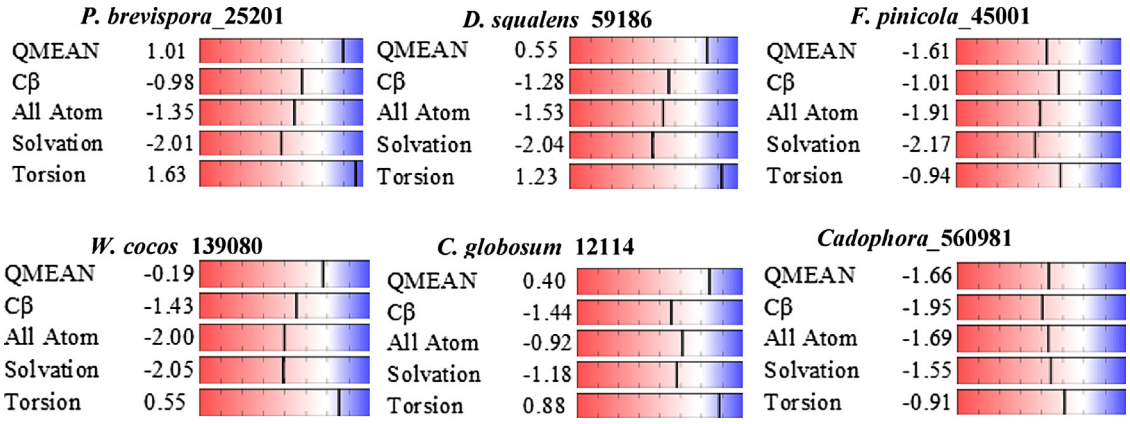


Fig. 2. QMEAN scores for the 3D modelled laccase structures obtained from SWISS-MODEL server for fungal protein sequences *P. brevispora, D. squalens, F. pinicola, W. cocos, C. globosum* and *Cadophora* sp.

Table 4
Comparison of results obtained from the Swiss model server (BLAST and HHblits) and BLASTP (NCBI PSI-BLAST against PDB server).

Organism, Protein-ID	Template Swiss PDB-server	Sequence Identity	GQME	QMEAN	Template BLASTP	Sequence Identity
Phlbr1 (25201)	5a7e.1A	70.24	0.86	1.08	3KW7	71
Dsqual1 (59186)	1kya.1A	80.92%	0.97	0.55	2QT6	82
Fompi3 (45001)	5ehf.1A	60.57	0.78	−1.6	3KW7	63
Wcocos (139080)	5anh.1A	63.83%	0.84	−0.19	3KW7	66
Chagl1 (12114)	1gw0.1A	80.29	0.84	0.41	1GW0	80
Cadophora (560981)	3pps.1A	61.75%	0.79	−1.66	3PPS	59

Table 5
Ramachandran plot scores of laccase modelled protein structures obtained after customized refining pipeline using different model refining softwares and results from RAMPAGE, PROQ, ERRAT and RMSD (initial and final protein structures) webservers.

Method Used	Phlbr1 (25201)	Dsqual (59186)	Fompi3 (45001)	Wcocos (139080)	Chagl1 (12114)	Cadophora (560981)
SWISS model (Initial quality)	87.0	87.1	82.7	82.4	85.9	81.6
Galaxy Refine (after SWISS)	90.5	90.4	88.3	90.3	91.3	**89.4**
KoBaMIN (after GR)	90.7	**92.3**	**89.3**	89.5	90.5	**88.3**
Only KoBaMIN	**91.7**	89.4	87.8	**90.0**	**92.2**	87.9
RAMPAGE (RFR, RAR, ROR)	97.6, 2, 0.4	97.4, 1.8, 0.8	94.8, 2.6, 2.6	93.9, 4, 2.1	98.4, 1.4, 0.2	96.2, 2.3, 1.4
RMSD (Initial vs Final Model)	0.36 A°	0.40 A°	0.34 A°	0.33 A°	0.35 A°	0.40 A°
RMSD (Final vs Template)	0.35 A°	0.52 A°	0.37 A°	0.34 A°	0.41 A°	0.51 A°
PROQ (LG and Max Sub)	4.631, 0.348	5.002, 0.359	5.228, 0.389	4.486, 0.295	4.562, 0.369	5.050, 0.325
ERRAT Server	87.780	86.735	75.102	87.152	89.091	84.432

Note: The pipeline used for refining the protein models was SWISS model → Galaxy Refine → KoBaMIN; only KoBaMIN, RAMPAGE results were RFR = residues in favored regions, RAR = residues in allowed regions and ROR = residues in outlier regions; The highlighted regions represent the final refined laccase modelled structures used for the molecular docking analysis and their respective Ramachandran plot scores.

function used for the estimation of both local and global qualities calculation which includes four parameters all atom, Cβ, solvation and torsion) and GMQE (global model quality estimation, which retrieves information from template and target alignment). All the selected laccase protein sequences were found by BLAST search with >60% of sequence identity. The 3D modelled protein structures of fungal laccase proteins were found to be statistically acceptable with higher QMEAN and GMQE scores, *P. brevispora* (1.08), *D. squalens* (0.55) and *C. globosum* (0.41), *F. pinicola* (−1.6), *W. cocos* (−0.19) and *Cadophora* sp. (−1.66). Both the QMEAN and GMQE scores were expressed as a number in a range between 0–1, where the higher number represents higher quality (Fig. 2) (Table 4). These initial protein models were refined using GalaxyRefine and KoBaMIN web servers which performs the side chain refinement,

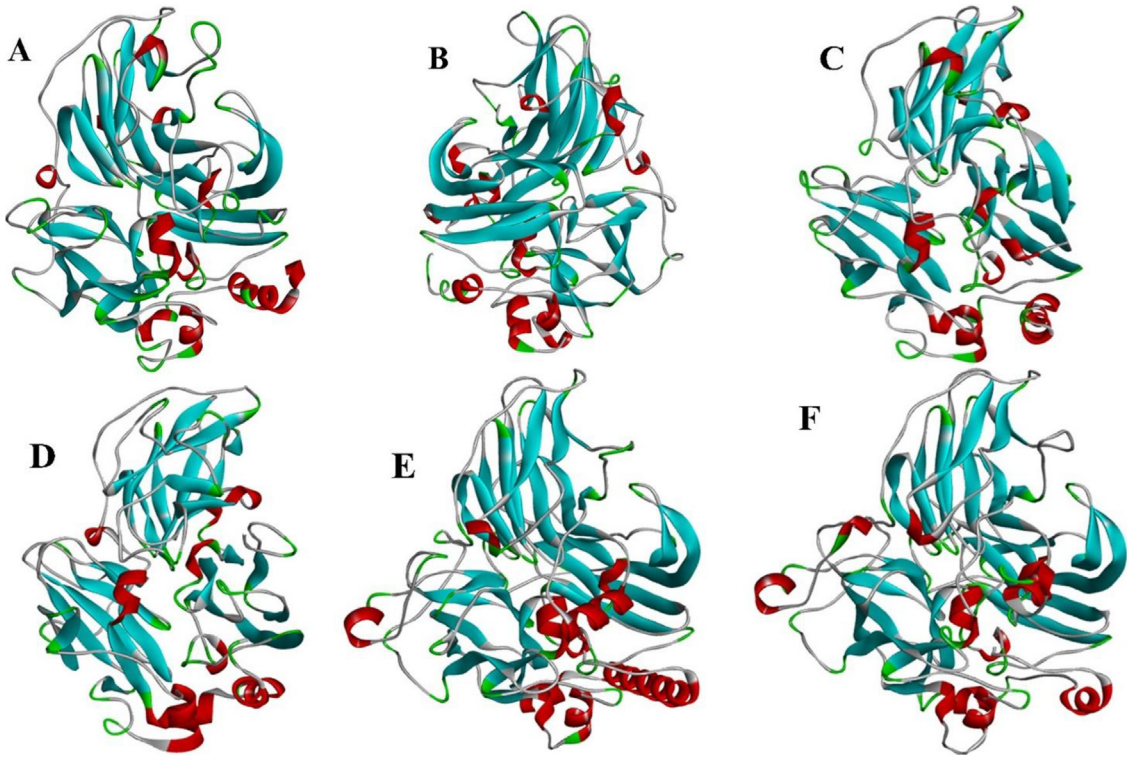


Fig. 3. Homology models of laccase protein sequences A) *Phlebia brevispora* B) *Fomitopsis pinicola* C) *Dichomitus squalens* D) *Wolfiporia cocos* E) *Chaetomium globosum* F) *Cadophora* DSE1049 v1.0.

energy minimization and relaxes the overall modelled structure. The above obtained refined laccase protein structures were validated using PROCHECK software and Ramachandran plots (Fig. S5). The results obtained from Ramachandran plots were found to be statistically acceptable, all the laccase refined models attained >90% residues in most favored regions (except for *F. pinicola*-89.3% and *Cadophora* sp. 89.4%) (Table 4). The refined protein structures were also validated using PROQ, ERRAT and RAMPAGE. PROQ web tool provides mainly LG (a −log P-value, models are good if the score is >3 and very good if the score is >5) and Max sub scores (ranges between 0–1, 1 being highly reliable and 0 is insignificant). All the refined laccase structures achieved significant scores with an LG score of >4 and Max sub score of > 0.3 (except *W. cocos*). Similarly, overall quality factor values obtained from ERRAT server for the refined structures were >84 (except *F. pinicola*). Generally, the predicted protein models attaining overall quality factor >50% infers that the homology models were stable and reliable (Fig. S8). Results from RAMPAGE server showed that all the refined laccase structures exhibited >93% of residues in most favored regions. Results obtained from QMEAN4, Ramachandran plots, PROQ, ERRAT and RAMPAGE web servers convey that predicted fungal laccases were of good quality (Table 5) (Figs. 2 & 3). The structural variation observed in the Ramachandran plots generated from PROCHECK and RAMPAGE web servers can be explained by the advanced refined and more reliable protein structure validation methods implemented by the RAMPAGE web server [41].

3.3. Molecular docking of modelled laccases with lignin model compounds

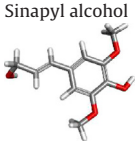
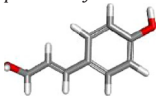
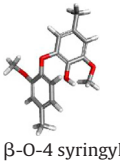
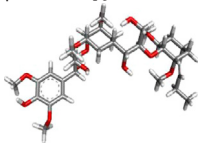
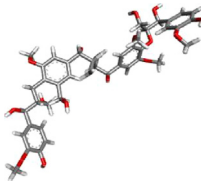
Molecular docking experiments of 3D modelled fungal laccases (white rot, brown rot and soft rot) with lignin model compounds that is monomers (sinapyl, coniferyl and *p*-coumaryl alcohol), dimer (guaiacyl 4-O-5 guaiacyl), trimer (syringyl β-O-4 syringyl β-O-4 sinapyl alcohol) and tetramer (guaiacyl β-O-4 syringyl β-β syringyl β-O-4 guaiacyl) was performed using AutoDock Tools and Vina software. Results obtained from this study were reported in Tables 6–8. Based on the results obtained we have observed a sharp increase in binding efficiencies from monomers to tetramers with increase in size of the ligand with all the fungal (white, brown and soft rot) laccases (Table 6). The binding efficiencies of white rot fungal laccases (*P. brevispora* and *D. squalens*) has exhibited a clear ascending order of minimum binding efficiencies from monomers to tetramers (−6.0 to −8.2). Brown rot fungal laccases (*F. pinicola* and *W. cocos*) has also shown a clear increase in minimum binding efficiencies from monomers to tetramers however, unlike white rot laccases brown rot laccases exhibited selective specificity among the ligands. When compared to white rot and brown rot fungal laccases, soft rot fungal laccases (*C. globosum* and *Cadophora* sp.) exhibited lesser minimum binding efficiencies and were variably specific among the ligands. Based on these results we can conclude that fungal laccases exhibit higher minimum binding efficiencies for lignin model compounds and out of which trimers and tetramers bind more efficiently to the fungal laccases. The higher binding efficiencies exhibited by white rot fungal laccases towards lignin model compounds are evident to their extrinsic lignin degrading abilities.

The lignin model compounds sinapyl (SA), coniferyl (CA), *p*-coumaryl alcohol (CoA), dimer, trimer and tetramer were found to interact with a total of 7, 7, 3, 6, 10,13 aminoacid residues of *P. brevispora* respectively. Out of which following aminoacid ligand interactions were found to be common among SER134, GLU481, PHE471, PHE369 (SA: CA), PRO371, HIS132 (SA: CA: CoA), ALA101 (SA: CoA) and PRO371, PHE102 (dimer and trimer). We have observed that hydrogen bond formations were between SA-ALA101, CA-HIS132 and GLU481, CoA-HIS132 and tetramer GLU325 residues of *P. brevispora* laccase (Fig. S9). Similarly, lignin model compounds were found to interact with 4, 6, 7, 6, 10 and 11 aminoacid residues of *D. squalens* laccase protein. Following aminoacid ligand interactions were observed in common

Table 6
Lists the final predicted minimum binding energy scores (kcal/mol) of predicted laccase models with lignin model compounds obtained from AutoDock Vina software.

Ligand (Modelled Protein-ID)	Phlbr1 (25201)	Dsqualm (59186)	Fompi3 (45001)	Wcocos (139080)	Chagl1 (12114)	Cadophora (560981)
Sinapyl alcohol	−6.0	−5.4	−5.7	−5.8	−6.0	−5.6
Coniferyl alcohol	−6.0	−6.2	−7.0	−6.0	−5.6	−5.8
p-coumaryl alcohol	−6.1	−6.3	−6.7	−5.8	−5.5	−6.8
Guaiacyl 4-O-5 guaiacyl	−6.9	−6.7	−6.2	−7.1	−6.5	−6.8
Syringyl β-O-4 syringyl β-O-4 sinapyl alcohol	−7.8	−7.1	−7.0	−8.2	−6.8	−6.9
Guaiacyl β-O-4 syringyl β-β syringyl β-O-4 guaiacyl	−7.7	−7.4	−7.2	−7.0	−7.2	−6.1

Table 7
Lists the aminoacid residues of modelled fungal laccases in contact with the lignin based model compounds Monomers (Sinapyl alcohol, Coniferyl alcohol and *p*-coumaryl alcohol), Dimer, Trimer and Tetramer.

Ligand	Phlbr1 (25201)	Dsqual (59186)	Fompi3 (45001)	Wcocos (139080)	Chagl1 (12114)	Cadophora (560981)
Sinapyl alcohol 	PRO371 HIS132 **SER134** GLU481 PHE471 GLY483 **PHE369**	PHE102 LEU133 TYR512 **LEU480**	**ARG95** PRO124 ASN125 GLN94	GLN81 PHE85 **SER384** GLY203 ALA82	GLY192 **TYR222** ASP220 VAL196 MET219 TYR221 THR278	ASP579 ALA397 SER536 VAL400 GLN537
Coniferyl alcohol 	SER134 **HIS132** PRO371 **GLU481** ALA101 PHE369 PHE471	**SER134** LEU133 PRO367 ASP481 HIS132 **PHE365**	PHE481, ILE183 ASP229, **ALA187** TYR176, ARG46 CYS141	**ALA82** GLN81 PHE48 GLY203 SER384 PHE85	PHE598 VAL601 THR134 ASN135 TRP430 ASN596 ASP227	LYS401 VAL400 ARG410 GLN537 SER536
p-Coumaryl alcohol 	**HIS132** PRO371 ALA101	PRO367 GLY483 PHE365 ALA101 ALA482 **ASP481** **HIS132**	**ASP229** TRP175 CYS141 CYS228 ASP174 TYR176	ASP202 **ALA82** GLY203 GLU204 GLN81 ASP405 SER384	TYR221 LYS101 TRP553 TYR222 ASP220 LEU229 GLY192 **THR280**	VAL264 LEU582 TRP156 TYR165 VAL535 **ALA532** SER170 PRO262
Guaiacyl 4-O-5  guaiacyl	PRO371 TYR512 LEU133 SER134 PHE102 PHE369	LEU133 PHE365 PHE102 PRO367 SER134 TYR512	LEU59 GLU166 ILE211 **VAL169** ASN195	SER384, LYS19 ALA82, PHE85 ASN205, GLY203 PHE48, **GLN81**	GLY490 VAL520 **ARG513** ARG491 PRO519 PRO382 VAL488	TYR244 PRO237 ASN332 TRP317 ASN348 **ALA336** VAL235
Syringyl β-O-4 syringyl β-O-4 sinapyl alcohol 	GLU481, PRO371 PHE102, LEU480 LYS178, ALA101 PHE471, PRO182 VAL370, LYS361	LYS92 TRP96 ARG64 PRO502 ALA504 GLN119 ASP117 THR94 **LEU115** TRP505	ASP82, GLU523 LEU483, SER181 LEU138, PRO110 **SER374**, ALA484 ILE183, ALA104 PHE105, GLY137	PHE48, **ASN424** GLU204, ASN205 **ALA82**, HIS383 LEU291, ARG403 ASP422, PHE85 SER384, GLU292 PRO111, SER386 PRO297, GLN81	**ARG400** ASN403 GLY402 **ILE451** PHE397 VAL452 **GLN453** TYR462 VAL398	THR150 **THR144** **ARG148** LEU121 LYS143 ASN123 SER146 ASN126 ILE125
Guaiacyl β-O-4 syringyl β-β syringyl β-O-4 guaiacyl 	ASN284, ASN324 TYR261, GLN264 LEU316, THR259 TYR266, THR232 **GLU325**, SER433 PHE293, SER234 LEU323	THR369 PRO367 ALA101 **ASP481** SER134 LEU133 TYR512 ALA517 LEU480 HIS132 PHE102	THR340, PRO343 **ASP152**, GLY338 THR466, GLY341 ARG448, PHE465 PHE430, LYS64 ASN250, SER431	PRO282, GLN216 LEU283, LEU284 GLU285, SER408 PHE218, SER279 **GLN214**, ALA391 GLY392	VAL452 **GLN459** ASP455 GLY402 **TYR462** **THR447** THR461 ASN422 ARG400 **ILE451** GLN453 TRP463	**ASN562** THR144 ARG148 LYS143 PHE147 ASN123 THR560 SER146 ASN126 ILE125

Note: aminoacids represented in bold are involved in hydrogen bonding between protein and ligand.

LEU480 (SA-Tetramer), ALA101 (CoA-Tetramer), PHE102, TYR512 (SA-Dimer-Tetramer), PHE365 (CA-CoA-Dimer), ASP481, HIS132 (CA-CoA-Tetramer), SER134 (CA-Dimer-Tetramer), LEU133 (SA-CA-Dimer-Tetramer) and PRO367 (CA-CoA-Dimer-Tetramer). And hydrogen bond formations were observed between SA-LEU480, CA-SER134 and PHE365, CoA- HIS132 and ASP481, Trimer- LEU115 and Tetramer-ASP481 (Fig. S10).

Table 8
Lists the root mean square deviation (RMSD Å) values obtained from the comparison studies of white, brown and soft rot fungal laccases.

Organism	*P. brevispora*	*D. squalens*	*F. pinicola*	*W. cocos*	*C. globosum*	*Cadophora* sp.
P. brevispora	n/a	0.62	0.83	0.46	1.11	1.06
D. squalens	0.62	n/a	0.84	0.68	1.13	1.08
F. pinicola	0.83	0.84	n/a	0.80	1.23	1.17
W. cocos	0.46	0.68	0.80	n/a	1.07	1.12
C. globosum	1.11	1.13	1.23	1.07	n/a	0.68
Cadophora sp.	1.06	1.08	1.17	1.12	0.68	n/a

Brown rot fungal laccases considered for this study *F. pinicola* and *W. cocos* were found to interact with lignin model compounds (SA, CA, CoA, dimer, trimer and tetramer) through 4, 7, 6, 5, 12, 12 and 5, 6, 7, 8, 16, 11 aminoacids residues respectively. *F. pinicola* laccase protein and ligands were involved in hydrogen bond formations between SA-ARG95, CA-ALA187, CoA-ASP229, dimer-VAL169, trimer-SER374 and tetramer-ASP152, with commonly interacting residues ASP229, TYR176 and CYS141 between CA and CoA, ILE183 between CA-tetramer respectively (Fig. S11). In *W. cocos* laccase protein following aminoacids were found in common GLU204 (CoA-trimer), ASN205(dimer-trimer), GLY203 (SA, CA, CoA, dimer), PHE85 (SA, CA, dimer, trimer) and GLN81, SER384, ALA82 (SA, CA, CoA, dimer and tetramer) with formation of hydrogen bonds between SA-SER384, CA-ALA82, CoA-ALA82, dimer-GLN81, trimer-ALA82, ASN424 and tetramer-GLN214 respectively (Fig. S12) (Table 7).

Soft rot fungal laccases considered for this study *C. globosum* and *Cadophora* sp. were found to interact with lignin ligand model compounds SA, CA, CoA, dimer, trimer and tetramer by 7, 7, 8, 7, 9, 12 and 5,5,8,7,9, 10 aminoacid residues respectively. In *C. globosum* laccase following aminoacid residues were found to interact commonly GLY192, TYR222, ASP220, TYR221 with SA and CoA, ARG400, GLY402, ILE451, VAL452, GLN453, TYR462 with trimer and tetramer respectively. We have observed hydrogen bond formation between SA-TYR222, CoA-THR280, dimer-ARG513, trimer-ARG400, ILE451, GLN453 and tetramer-THRE447, ILE451, GLN459, TYR462 residues respectively (Fig. S13). In *Cadophora* sp. laccase protein following residues SER536, VAL400 and GLN537 found to be interact commonly with SA and CA, whereas THR144, ARG148, LYS143, ASN123, SER146, ASN126, ILE125 with trimer and tetramer respectively. And hydrogen bond formation was observed between CoA- ALA532, dimer-ALA336, trimer-THR144, ARG148 and tetramer-ASN562 residues (Table 7) (Fig. S14).

4. Discussion

In our present study, we have performed a comparative modeling and molecular docking study of fungal laccase protein sequences, to understand and reveal the lignin degrading abilities exhibited by white rot, brown rot and soft rot fungi. We have retrieved all the laccase protein sequences (genome wide) of the selected fungi however, based on the sequence similarity against experimentally validated protein structures (PDB) we have considered Phlbr1-25201, Dsqual-59186, Fompi3-45001, Wcocos-139080, Chagl1-12114 Cadophora-560981. The above considered laccases possess the essential domains belonging to cupredoxin superfamily and specifically contains Cu_oxidase type 1, 2 and 3 of multicopper oxidase domains. We have observed that *P. brevispora*, *D. squalens* (white rot) laccase sequences were closely related to the experimentally determined laccase of *Trametes versicolor* (1GYC) and laccase sequences of *F. pinicola* and *W. cocos* fall under the same branch and share good similarity with *T. versicolor* (1GYC) and white rot fungal laccases. Whereas soft rot fungal laccases *C. globosum* and *Cadophora* sp. were found to be closely related to the experimentally determined laccase of *Melanocarpus albomyces* (1GW0) (Fig. 1A). Multicopper oxidases (MCO) are group of enzymes performing single electron oxidation of various substrate with an associated four electron reduction of molecular oxygen to water molecule [49]. The enzymes belonging to MCO class were further classified into laccase, ferroxidase, ascorbate oxidase and ceruloplasmin. Previous studies have revealed that MCOs consists of two active sites a) substrate oxidizing site (blue type-1 (T1) copper site) and b) oxygen binding site (tri nuclear copper site containing three type-2 (T2) or type-3 (T3) coppers) [49]. The electrons are transferred from T1 copper site to T2/T3 copper site through a set of highly conserved amino acid residues (MCO-specific patterns) [6,50,51] (Fig. 1B). Except ceruloplasmin (six domains) and bacterial laccases (two domains), most of the MCOs contain three cupredoxin domains, based on the presence of these domains the length of MCOs ranges between 300–1000 residues and can contain up to six copper ions [6,49].

All the fungal laccase protein sequences selected for the present study ranges between 479–619 (Table 1) containing four copper ions mostly in contact with histidine residues (Table S1). The aminoacid composition of the selected laccases show the higher content of negatively charged amino acids (aspartic and glutamic acids) can explain about the acidic nature (theoretical pI values obtained) of the laccases. The concentration of tyrosine, tryptophan and cysteine residues reflects the extinction coefficients of fungal laccases [27]. The physico-chemical properties such as theoretical pI and molecular weight, extinction coefficient values for soft rot fungal laccases (*C. globosum* and *Cadophora* sp.) were comparatively higher than brown and white rot laccases. The lower extinction coefficients of white and brown rot laccases might be due to the lower content of phenylalanine, tyrosine, tryptophan and cysteine residues [27]. Studies have reported that proteins exhibiting an instability index lesser than 40 possess an in vivo half-life of 5 h and instability index greater than 40 has an in vivo half-life period of 16 h [52]. The instability index (used for estimation of in vivo half-life of proteins) values report that the selected laccases have a long in vivo half-life period of 16 h except for *W. cocos* laccase protein (in vivo half-life period of <5 h) [28,52]. The aliphatic index (determined using the relative volume occupied by aliphatic side chains of alanine, valine, leucine and isoleucine) is used as a positive factor for increase in thermal stability of globular proteins [29]. The selected fungal laccases exhibited aliphatic index values in the range of 74.44–90.04, which suggests the stability of selected laccases at wider temperature ranges. The hydrophobicity/hydrophilicity of protein can be defined using the GRAVY (grand average of hydropathicity) index values, where the positive and negative values denote for hydrophobic and hydrophilic natures of the protein. All the selected fungal laccase proteins except *P. brevispora* were found to be hydrophilic in nature reporting that these laccases interact better with water. The hydropathicity plots generated using Discovery studio visualizer® for the laccase protein sequences also support the above reported values (Fig. S4).

The secondary structure analysis of the selected laccase protein sequences using SOPMA web server has revealed that the per-

centage of random coiled secondary structures content is higher followed by extended strand percentage (Table 3). Earlier studies have reported that random coiled secondary structures are involved in imparting flexibility, turnover and conformational changes of the enzymes [53]. Upon Motif scan analysis, all the selected laccase proteins were commonly found to contain motif sequences for phosphorylation, glycosylation, myristoylation and multi copper oxidase patterns (Table 2). Phosphorylation is significant process which effects the functional and structural activities of proteins and regulates the cell behavior in eukaryotes, through controlling its intrinsic biological activity, cellular localization and interaction with other proteins [54]. The laccase protein sequences commonly showed casein kinase-II, protein kinase C and cAMP/cGMP dependent protein kinase (*F. pinicola* and *W. cocos*) phosphorylation sites. Similarly, we have observed myristoylation patterns commonly among the laccases, myristoylation (modification of proteins with myristic acid) alters the conformational stability of proteins through interacting with hydrophobic membranes and domains of the protein, it also plays crucial role in cellular signaling and extracellular export of the proteins [18,55–57]. Other commonly observed patterns include glycosylation and amidation these post translational modifications are involved in imparting thermal stability, copper retention and retains its biological activities [58]. According to Marion et al., most of the laccases are extracellular glycoproteins [59], results obtained from CELLO v2.5 subcellular location predictor showed that the above selected laccases were extracellular. We have observed two disulfide bonds in *P. brevispora, D. squalens, W. cocos* and three disulfide bonds in *F. pinicola, C. globosum* and *Cadophora* sp. Previous studies have reported that extracellular proteins contain more cysteine residues and disulfide bonds, compared to intracellular proteins, thus disulfide bonds in extracellular laccase proteins play crucial role during protein folding and stability of the protein [60,61]. Presence of signal peptide cleavage sites (except *W. cocos*) and absence of transmembrane helices (except *D. squalens*), supports the extracellular nature of the selected fungal laccases.

Homology modeling studies are highly significant in determining protein structures for the proteins which lack the experimental structures, in the past few decades a wide range of efficient tools and servers were developed which can perform modeling studies of proteins even with 30% of sequence identity with an accuracy achieved from low resolution X-ray structures [18,62]. Thus performing computational studies predicting the structural and functional properties of commercially important enzymes will significantly support in planning biological experiments based on these enzymes [18]. Homology modeling of the selected fungal laccases was performed using SWISS Model automated server, it selects best template based on the sequence identity results obtained through BLAST and HHblits [37]. SWISS-Model server performs a range of quality checks and refines the side chains and loops of the targets using template structures, it generates quality scores such as QMEAN4 and GMQE scores. According to Benkert et al., QMEAN is a compound scoring function used for the estimation of global and local model qualities [63]. QMEAN uses four structural descriptors a) torsion angle potential b) all atom c)C-beta interactions d) solvation potential, for the estimation of local and global qualities of the modelled structures between 0 (unreliable) to 1(better model) [63]. Global model quality estimation (GMQE) combines properties from the template and target alignment, GMQE is estimated in between 0 (unreliable) to 1 (reliable mode) [37]. We have obtained positive QMEAN scores for *P. brevispora, D. squalens* and *C. globosum* whereas negative QMEAN scores for *F. pinicola, W. cocos* and *Cadophora* sp. laccases. SWISS Model server provides the quality estimations in the form of a chart where 3D modelled structures have a color gradient in between blue to red which defines the resolutions between 1 and 3.5 Å respectively, where the greater blue values represents a reliable structures [37]. Estimated global model qualities (GMQE) for the modelled laccases where found to be in between the range of 0.78–0.97, which indicates a good quality and reliable structures (Fig. 3).

The above obtained 3D modelled structures of fungal laccases were further refined using GalaxyRefine and KoBaMIN web servers. KoBaMIN web server performs a knowledge based potential of mean force correction of the modelled structures, KoBaMIN internally uses knowledge based potential derived from PDB structures which also uses the energy function which implements the effects of solvent and crystal environment and MESHI for optimizing stereochemistry of the models [39]. GalaxyRefine is online web server which primarily rebuilds the side chains and also performs repacking of side chains and later relaxes the overall structure through molecular dynamic simulation methods, It improves both local and global structural qualities of the predicted models [38]. We have applied a combination of above mentioned model refining methods differentially to obtain a best protein model with reliable Ramachandran scores and higher acceptable residues. We have performed validation of the refined models using PROCHECK, RAMPAGE and ERRAT analysis (Table 4). The refined selected laccase models upon Ramachandran plot analysis revealed that 3D-modelled protein structures have >90% of residues in allowed regions. For the predicted protein structures, it is ideal to contain at least 90% of the residues must be in regions of allowed regions, which suggests the predicted laccase structures were of good quality. We have also validated the refined laccase structures using PROQ web server, it generates two quality metrics for the estimation of model quality they are LG score and Max sub-score [43]. Where LG score must be greater than 4 and max sub score must be greater than 0.8 (very good) for the reliable protein structures. All the predicted laccase models have attained LG and max sub scores between the range of (4.48–5.22) and (0.295–0.389) respectively, which suggests that the refined 3D models of fungal laccases were of good quality [43]. Finally, we have used ERRAT web server for the validation of the refined models, ERRAT analyses the non-bonded interactions among different atom types based on characteristic atomic interactions [42]. We have observed that all the predicted model structures have attained an overall quality factor >50 which confirms that the refined laccase models were of good quality [18,42]. From the predicted 3D modelled fungal laccase structures, we can infer that soft rot fungal laccases *C. globosum* and *Cadophora* sp. have shown significant differences in structural and physico-chemical properties when compared to white rot and brown rot fungal laccases. In order, to reveal these differences we have superimposed all the refined laccase models using SWISS PDB viewer. Results obtained from the superimposition studies have revealed that RMSD values obtained for soft rot fungal laccases (*C. globosum* and *Cadophora* sp.) showed slight structural differences when compared to white and brown rot laccases (Table 8).

Several structural studies were conducted in the past which has revealed the structural and functional properties of plant cell wall based compounds and especially lignin based compounds. Ralph et al. has conducted structural studies of lignin and other plant cell wall compounds, and developed a single source database for lignin model compounds [44]. We have considered six lignin model compounds for the present protein docking studies, lignin building monomers (sinapyl, coniferyl and *p*-coumaryl alcohol), dimer (guaiacyl 4-O-5 guaiacyl), trimer (syringyl β-O-4 syringyl β-O-4 sinapyl alcohol) and tetramer (guaiacyl β-O-4 syringyl β-β syringyl β-O-4 guaiacyl) [19]. We have used AutoDock Vina and Tools for achieving the protein docking studies using the above-mentioned lignin model compounds [46]. AutoDock Vina is a fast and accurate method of ligand protein docking tools which will facilitate flexible docking studies [46]. We have clearly observed that in *P. brevispora, W. cocos* laccase protein, monomers, dimer and trimer were

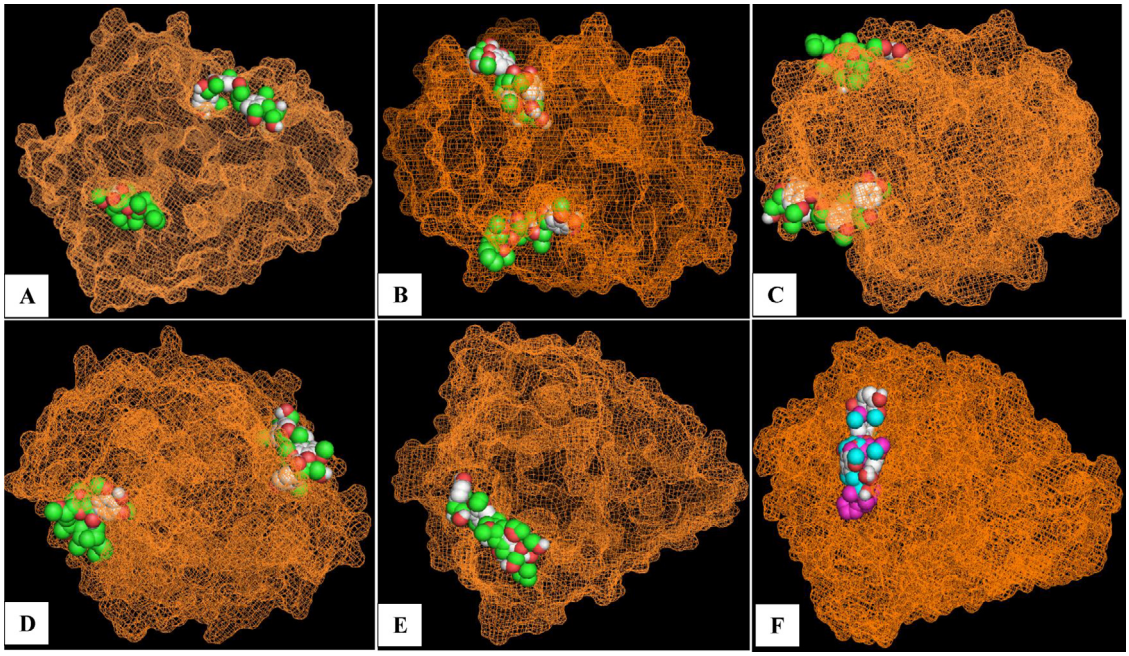


Fig. 4. Protein docking of laccase protein molecular models with syringyl β-O-4 syringyl β-O-4 sinapyl alcohol (Trimer) and guaiacyl β-O-4 syringyl β-β syringyl β-O-4 guaiacyl (Tetramer), A) *Phlebia brevispora,* B) *Wolfiporia cocos,* C) *Dichomitus squalens,* D) *Fomitopsis pinicola,* E) *Chaetomium globosum* and F) *Cadophora* DSE1049 v1.0.

mostly found to bind to the same pocket in different conformations, with tetramer binding at different (large) pocket. Whereas in *D. squalens* laccase protein we have observed that monomers, dimer and tetramer bind at the same pocket and contrastingly trimer binds separately at different pocket (Fig. S15). The lignin model compounds (monomer, dimer, trimer and tetramers) were found bind separately at different pockets in *F. pinicola* laccase protein. Both soft rot fungal (*Cadophora* sp. and *C. globosum*) laccases when docked with lignin model compounds, showed similar binding patterns monomers, dimer bound at different pockets however, trimer and tetramer bound at the same pockets respectively (Fig. S15). The binding patterns of trimer and tetramer by modelled fungal laccases were shown below (Fig. 4).

We have compared the aminoacid residues of fungal laccases which are occurring in close interactions with lignin model compounds. Commonly found aminoacid residues between the white rot fungal laccases were PRO, HIS, SER, PHE, GLY, ALA, TYR, LEU, LYS, GLN, Thr. Similarly, common aminoacid residues between brown rot fungal laccases were ARG, PRO, ASN, GLN, PHE, ASP, ALA, LEU, GLU, SER, GLY and Lys. Finally, common aminoacid residues between soft rot fungal laccases were TYR, ASP, VAL, THR, PHE, ASN, TRP, LYS, LEU, ARG, PRO, ILE and Gln. The commonly occurring aminoacid residues among the selected white, brown and soft rot fungal laccases were found to be PRO, PHE, LEU, LYS and GLN, other commonly found aminoacids were SER, GLY, ALA (white and brown rot), ARG, ASN, ASP (brown and soft rot) and TYR, THR (white and soft rot) respectively.

In support of our present study, protein docking studies using lignin model compounds (sinapyl alcohol, dimer, trimer and tetramer) earlier by Awasthi et al., has revealed the, interactions of 11 aminoacid residues commonly (LEU, ASP, ASN, PHE, SER, PHE, GLY, ALA, PRO, ILE, and HIS) with all the lignin models [19]. Crystallographic studies of *Melanocarpus albomyces* laccase using lignin model compounds by Kallio et al., has revealed the interactions of seven amino acid residues (ALA, PRO, GLU, LEU, PHE, TRP, and HIS) with the lignin model compounds [64]. According to Awasthi et al., fungal laccases exhibiting higher redox potential were found to include phenylalanine in the active binding site of the lignin model compounds, which is in accordance with the present results [19]. Molecular docking studies of *G. lucidum* and *P. ostreatus* 3D-predicted laccases with ABTS, has revealed the close interactions of aminoacid residues (Phe, Asp, Ser, Pro, Gly, Ile, His and Gly, Val, Pro, Asp, Ser, Asn, Phe, Ile, Trp and His) with the ligand [15]. Similarly, studies conducted by Morozova et al. have revealed that laccases exhibiting high potential were usually found to contain phenylalanine residue at its axial ligand of type-T1 copper binding site. [65]. According to Xu, laccases containing PHE residue at type-T1 copper binding site has showed higher redox potential than MET containing laccases [66].

Results obtained in our present study highlights the structural and functional properties exhibited by white, brown and soft rot fungal laccase models. We have clearly observed that white and brown rot fungi exhibited clear and strong binding efficiencies towards lignin model compounds when compared to soft rot fungal laccases. We have also seen that physico-chemical and structural properties of soft rot fungi exhibited little but significant differences upon comparison with white and brown rot fungi. Further molecular dynamic simulation and corresponding wet lab experiments must be performed to understand the catalytic efficiencies of fungal laccases. Increasing genome and transcriptome wide studies were continuously revealing the molecular complexities of several fungi. Efficient methods for fishing high potential laccases will play a crucial role in developing genetically efficient microorganisms for the degradation of lignin, which will significantly help the growing biofuel and bioremediation industries.

Conflict of interest

The authors have declared that no competing interest exists.

Acknowledgements

This work was supported by NSERC-RDF Fund to Wensheng Qin and Ontario Trillium Scholarship (OTS) to Ayyappa Kumar Sista Kameshwar.

References

[1] U.N. Dwivedi, P. Singh, V.P. Pandey, A. Kumar, Structure–function relationship among bacterial, fungal and plant laccases, J. Mol. Catal. B: Enzym. 68 (2011) 117–128.
[2] A.K.S. Kameshwar, W. Qin, Lignin degrading fungal enzymes, in: Production of Biofuels and Chemicals from Lignin, Springer, 2016, pp. 81–130.
[3] H. Yoshida, LXIII.—chemistry of lacquer (Urushi). Part I. communication from the chemical society of Tokio, J. chem. Soc. Trans. 43 (1883) 472–486.
[4] G. Bertrand, Sur la presence simultanee de la laccase et de la tyrosinase dans le suc de quelques champignons, CR Hebd Seances Acad Sci. 123 (1896) 463–465.
[5] D. Singh, K.K. Sharma, M.S. Dhar, J.S. Virdi, Molecular modeling and docking of novel laccase from multiple serotype of Yersinia enterocolitica suggests differential and multiple substrate binding, Biochem. Biophys. Res. Commun. 449 (2014) 157–162.
[6] A. Messerschmidt, R. Huber, The blue oxidases, ascorbate oxidase, laccase and ceruloplasmin Modelling and structural relationships, Eur. J. Biochem. 187 (1990) 341–352.
[7] B. Yao, Y. Ji, Lignin biodegradation with laccase-mediator systems, Front. Energy Res. 2 (2014) 12.
[8] P. Giardina, V. Faraco, C. Pezzella, A. Piscitelli, S. Vanhulle, G. Sannia, Laccases a never-ending story, Cell. Mol. Life Sci. 67 (2010) 369–385.
[9] C. Eggert, U. Temp, K.-E. Eriksson, The ligninolytic system of the white rot fungus Pycnoporus cinnabarinus: purification and characterization of the laccase, Appl. Environ. Microbiol. 62 (1996) 1151–1158.
[10] R. Bourbonnais, M. Paice, B. Freiermuth, E. Bodie, S. Borneman, Reactivities of various mediators and laccases with kraft pulp and lignin model compounds, Appl. Environ. Microbiol. 63 (1997) 4627–4632.
[11] S. Camarero, D. Ibarra, M.J. Martínez, Á.T. Martínez, Lignin-derived compounds as efficient laccase mediators for decolorization of different types of recalcitrant dyes, Appl. Environ. Microbiol. 71 (2005) 1775–1784.
[12] J. Rencoret, A. Pereira, C. José, A.T. Martínez, A. Gutiérrez, Laccase-mediator pretreatment of wheat straw degrades lignin and improves saccharification, Bioenergy Res. 9 (2016) 917–930.
[13] A. Kunamneni, A. Ballesteros, F.J. Plou, M. Alcalde, Fungal laccase—a versatile enzyme for biotechnological applications, Communicating current research and educational topics and trends in applied microbiology 1 (2007) 233–245.
[14] C.F. Thurston, The structure and function of fungal laccases, Microbiology 140 (1994) 19–26.
[15] C.M. Rivera-Hoyos, E.D. Morales-Álvarez, S.A. Poveda-Cuevas, E.A. Reyes-Guzmán, R.A. Poutou-Piñales, E.A. Reyes-Montaño, et al., Computational analysis and low-scale constitutive expression of laccases synthetic genes GILCC1 from Ganoderma lucidum and POXA 1B from Pleurotus ostreatus in Pichia pastoris, PLoS One 10 (2015) e0116524.
[16] R.J. Meshram, A. Gavhane, R. Gaikar, T. Bansode, A. Maskar, A. Gupta, et al., Sequence analysis and homology modeling of laccase from Pycnoporus cinnabarinus, Bioinformation 5 (2010) 150.
[17] K.-S. Wong, M.-K. Cheung, C.-H. Au, H.-S. Kwan, A novel Lentinula edodes laccase and its comparative enzymology suggest guaiacol-based laccase engineering for bioremediation, PLoS One 8 (2013) e66426.
[18] A.S. Tamboli, N.R. Rane, S.M. Patil, S.P. Biradar, P.K. Pawar, S.P. Govindwar, Physicochemical characterization, structural analysis and homology modeling of bacterial and fungal laccases using in silico methods, Network Model. Anal. Health Inform. Bioinform. 4 (2015) 17.
[19] M. Awasthi, N. Jaiswal, S. Singh, V.P. Pandey, U.N. Dwivedi, Molecular docking and dynamics simulation analyses unraveling the differential enzymatic catalysis by plant and fungal laccases with respect to lignin biosynthesis and degradation, J. Biomol. Struct. Dyn. 33 (2015) 1835–1849.
[20] M. Binder, A. Justo, R. Riley, A. Salamov, F. Lopez-Giraldez, E. Sjökvist, et al., Phylogenetic and phylogenomic overview of the Polyporales, Mycologia 105 (2013) 1350–1373.
[21] D. Floudas, M. Binder, R. Riley, K. Barry, R.A. Blanchette, B. Henrissat, et al., The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes, Science 336 (2012) 1715–1719.
[22] R.M. Berka, I.V. Grigoriev, R. Otillar, A. Salamov, J. Grimwood, I. Reid, et al., Comparative genomic analysis of the thermophilic biomass-degrading fungi Myceliophthora thermophila and Thielavia terrestris, Nat. Biotechnol. 29 (2011) 922–927.
[23] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, et al., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.
[24] S. Kumar, G. Stecher, K. Tamura, MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets, Mol. Biol. Evol. (2016) msw054.
[25] M.A. Larkin, G. Blackshields, N. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, et al., Clustal W and clustal X version 2.0, Bioinformatics 23 (2007) 2947–2948.
[26] E. Gasteiger, C. Hoogland, A. Gattiker, M.R. Duvaud S. e. Wilkins, R.D. Appel, et al., Protein Identification and Analysis Tools on the ExPASy Server, Springer, 2005.
[27] S.C. Gill, P.H. Von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, Anal. Biochem. 182 (1989) 319–326.
[28] K. Guruprasad, B.B. Reddy, M.W. Pandit, Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence, Protein Eng. 4 (1990) 155–161.
[29] I. Atsushi, Thermostability and aliphatic index of globular proteins, J. Biochem. 88 (1980) 1895–1898.
[30] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[31] C. Combet, C. Blanchet, C. Geourjon, G. Deleage, NPS@: Network Protein Sequence Analysis, Elsevier Current Trends, 2000.
[32] M. Pagni, V. Ioannidis, L. Cerutti, M. Zahn-Zabal, C.V. Jongeneel, J. Hau, et al., MyHits: improvements to an interactive resource for analyzing protein sequences, Nucleic Acids Res. 35 (2007) W433–W437.
[33] C.S. Yu, Y.C. Chen, C.H. Lu, J.K. Hwang, Prediction of protein subcellular localization, Proteins: Struct. Funct. Bioinform. 64 (2006) 643–651.
[34] H.-H. Lin, J.-C. Hsu, Y.-N. Hsu, R.-H. Pan, Y.-F. Chen, L.-Y. Tseng, Disulfide connectivity prediction based on structural information without a prior knowledge of the bonding state of cysteines, Comput. Biol. Med. 43 (2013) 1941–1948.
[35] T.N. Petersen, S. Brunak, G. von Heijne, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nat. Methods 8 (2011) 785–786.
[36] L. Kiemer, J.D. Bendtsen, N. Blom, NetAcet: prediction of N-terminal acetylation sites, Bioinformatics 21 (2005) 1269–1270.
[37] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modeling server, Nucleic Acids Res. 31 (2003) 3381–3385.
[38] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, Nucleic Acids Res. 41 (2013) W384–W388.
[39] J.P. Rodrigues, M. Levitt, G. Chopra, KoBaMIN: a knowledge-based minimization web server for protein structure refinement, Nucleic Acids Res. (2012) gks376.
[40] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr. 26 (1993) 283–291.
[41] S.C. Lovell, I.W. Davis, W.B. Arendall, P.I. de Bakker, J.M. Word, M.G. Prisant, et al., Structure validation by Cα geometry: φ, ψ and Cβ deviation, Proteins: Struct. Funct. Bioinform. 50 (2003) 437–450.
[42] C. Colovos, T.O. Yeates, Verification of protein structures: patterns of nonbonded atomic interactions, Protein Sci. 2 (1993) 1511–1519.
[43] S. Cristobal, A. Zemla, D. Fischer, L. Rychlewski, A. Elofsson, A study of quality measures for protein threading models, BMC Bioinform. 2 (2001) 5.
[44] S.A. Ralph, J. Ralph, L. Landucci, L. Landucci, NMR Database of Lignin and Cell Wall Model Compounds, US Forest Prod. Lab., Madison, WI, 2004 http://ars.usda.gov/Services/docs.htm.
[45] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, et al., AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791.
[46] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.
[47] N. Guex, M.C. Peitsch, SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling, Electrophoresis 18 (1997) 2714–2723.
[48] K. Nakai, Protein sorting signals and prediction of subcellular localization, Adv. Protein Chem. 54 (2000) 277–344.
[49] D. Sirim, F. Wagner, L. Wang, R.D. Schmid, J. Pleiss, The Laccase Engineering Database: a classification and analysis system for laccases and related multicopper oxidases, Database 2011 (2011) bar006.
[50] C. Ouzounis, C. Sander, A structure-derived sequence pattern for the detection of type I copper binding domains in distantly related proteins, FEBS Lett. 279 (1991) 73–78.
[51] S. Kumar, P.S. Phale, S. Durani, P.P. Wangikar, Combined sequence and structure analysis of the fungal laccase family, Biotechnol. Bioeng. 83 (2003) 386–394.
[52] S. Rogers, R. Wells, M. Rechsteiner, Amino acid sequences common to rapidly degrade proteins: the PEST hypothesis, Science 234 (1986) 364–369.
[53] E. Buxbaum, Fundamentals of Protein Structure and Function, Springer, 2007.
[54] P. Cohen, The regulation of protein function by multisite phosphorylation–a 25 year update, Trends Biochem. Sci. 25 (2000) 596–601.
[55] S. Podell, M. Gribskov, Predicting N-terminal myristoylation sites in plant proteins, BMC Genomics 5 (2004) 37.
[56] J. Zheng, D.R. Knighton, S.S. Taylor, N.H. Xuong, J.M. Sowadski, L.F.T. Eyck, Crystal structures of the myristylated catalytic subunit of cAMP-dependent protein kinase reveal open and closed conformations, Protein Sci. 2 (1993) 1559–1573.
[57] H.B. Olsen, N.C. Kaarsholm, Structural effects of protein lipidation as revealed by LysB29-myristoyl, des (B30) insulin, Biochemistry 39 (2000) 11893–11900.
[58] G. Walsh, Post-translational modifications in the context of therapeutic proteins: an introductory overview, in: Post-translational Modification of Protein Biopharmaceuticals, 2009, pp. 1–14.
[59] M. Heinzkill, L. Bech, T. Halkier, P. Schneider, T. Anke, Characterization of laccases and peroxidases from wood-rotting fungi (family Coprinaceae), Appl. Environ. Microbiol. 64 (1998) 1601–1606.
[60] R.A. Bradshaw, Protein translocation and turnover in eukaryotic cells, Trends Biochem. Sci. 14 (1989) 276–279.
[61] H. Nakashima, K. Nishikawa, Discrimination of intracellular and extracellular proteins using amino acid composition and residue-pair frequencies, J. Mol. Biol. 238 (1994) 54–61.

[62] Z. Xiang, Advances in homology protein structure modeling, Curr. Protein Pept. Sci. 7 (2006) 217–227.
[63] P. Benkert, M. Künzli, T. Schwede, QMEAN server for protein model quality estimation, Nucleic Acids Res. (2009) gkp322.
[64] J. Kallio, S. Auer, J. Jänis, M. Andberg, K. Kruus, J. Rouvinen, et al., Structure–function studies of a Melanocarpus albomyces laccase suggest a pathway for oxidation of phenolic compounds, J. Mol. Biol. 392 (2009) 895–909.
[65] O. Morozova, G. Shumakovich, M. Gorbacheva, S. Shleev, A. Yaropolov, Blue laccases, Biochemistry (Moscow) 72 (2007) 1136–1150.
[66] F. Xu, Oxidation of phenols, anilines, and benzenethiols by fungal laccases: correlation between activity and redox potentials as well as halide inhibition, Biochemistry 35 (1996) 7608–7614.