X-ray structure of a two-domain type laccase: A missing link in the evolution of multi-copper proteins

Hirofumi Komori a,b, Kentaro Miyazaki c,d,*, Yoshiki Higuchi a,b,∗

1Department of Life Science, Graduate School of Life Science, University of Hyogo and Himeji Institute of Technology, 3-2-1 Koto, Kamigori-cho, Ako-gun, Hyogo 678-1297, Japan
2RIKEN SPring-8 Center, 1-1-1 Koto, Mikazuki-cho, Sayo-gun, Hyogo 679-5248, Japan
3Institute for Biological Resources and Functions, National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan
4Department of Medical Genome Sciences, Graduate School of Frontier Sciences, University of Tokyo, Tsukuba Central 6, 1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

A R T I C L E   I N F O

Article history:
Received 19 December 2008
Revised 2 March 2009
Accepted 4 March 2009
Available online 11 March 2009
Edited by Hans Eklund

Keywords:
Crystal structure
Laccase
Metagenome
Molecular evolution
Multi-copper protein
Nitrite reductase

A B S T R A C T

A multi-copper protein with two cupredoxin-like domains was identified from our in-house metagenomic database. The recombinant protein, mgLAC, contained four copper ions/subunits, oxidized various phenolic and non-phenolic substrates, and had spectroscopic properties similar to common laccases. X-ray structure analysis revealed a homotrimeric architecture for this enzyme, which resembles nitrite reductase (NIR). However, a difference in copper coordination was found at the domain interface. mgLAC contains a T2/T3 tri-nuclear copper cluster at this site, whereas a mononuclear T2 copper occupies this position in NIR. The trimer is thus an essential part of the architecture of two-domain multi-copper proteins, and mgLAC may be an evolutionary precursor of NIR.

© 2009 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

Copper is a metal essential for life, and copper-containing proteins play various roles in biosystems, such as in electron transfer (e.g., azurin and plastocyanin), oxygen transfer (e.g., hemocyanin), and catalysis (e.g., laccase, ascorbate oxidase, tyrosinase, and galactose oxidase). These proteins use cupredoxin-like domains as structural units. The number of domains in a subunit, and the pattern of subunit assembly, varies from protein to protein, which, together with point mutations, appear to promote the functional evolution of copper proteins [1]. For example, ascorbate oxidase (AO) is a monomeric protein consisting of three cupredoxin domains [2]. Ceruloplasmin (CP) is also monomeric, but contains six cupredoxin domains [3]. Nitrite reductase (NIR) is composed of three identical subunits, each of which contains two cupredoxin domains [4]. Each cupredoxin domain contains a single copper-binding motif, in which various types of copper ions are included. Copper is classified into types I, II, and III (T1, T2, and T3) based on spectroscopic properties. T1 copper has an absorption peak at ~600 nm and a narrow hyperfine coupling in EPR (electron paramagnetic resonance) spectroscopy. This copper is generally coordinated by three strong ligands (one cysteine and two histidines) and one weaker ligand, typically methionine. In some fungal laccases, the methionine is replaced by non-ligating residues leucine or phenylalanine. T2 copper has a much weaker absorption, broader hyperfine interactions, and is generally coordinated with 4–5 histidines and water–oxygen ligands. T3 is a copper pair, antiferromagnetically coupled, usually coordinated by three histidines per copper and a bridging moiety [5]. High conservation of copper-binding motifs in cupredoxin domains allows for easy identification of copper proteins based solely on the amino acid sequence.

Taking advantage of the rapid accumulation of bacterial genomic sequences, Nakamura et al. [6] identified numerous protein sequences that possess cupredoxin-like domains. Based on the topology of the copper-binding motifs, they classified the proteins into several categories and proposed a possible evolutionary scheme, as shown in Fig. 1. In addition to well-characterized proteins, such as NIR, AO, and CP, there are numerous other proteins, designated as [X], [Y], [A], [B], and [C] in Fig. 1. Most of these
proteins varied between simple cupredoxin to well-established, multi-domain proteins. Although these “missing links” are uncharacterized, some of them have been studied. SLAC [7] and EpoA [8,9] are type [B] proteins composed of two cupredoxin domains. These proteins were identified in Streptomyces and have laccase activity. Recently, the crystal structure of SLAC was determined [10,11]. The enzyme formed trimer and resembled NIR [4] and CP [3]. A type [C] protein has also been reported. More than 20 years ago, a “blue copper oxidase” was purified from Nitrinosomas europaea [12]. Biochemical studies revealed that the enzyme was trimeric and displayed both laccase and NIR activities. Recent genetic studies [13,14] have confirmed that the enzyme is a type [C] protein.

In this paper, we report the crystal structure of a novel type [C] laccase, designated mgLAC, which was identified from our in-house metagenomic sequence database [15–18]. We determined the three-dimensional structure of mgLAC to better understand the molecular basis for the evolution of multi-copper proteins.

2. Materials and methods

2.1. Enzyme activity assays

mgLAC was prepared as previously described [18]. Enzyme activities were measured from the decrease in the concentration of dissolved molecular oxygen at 25 °C in McIlvaine buffer (pH 4.5 for ABTS and 8.5 for others), containing 1 mM substrate on a PreSens oxygen meter (Microx TX3).

2.2. Spectroscopic measurements

The UV–visible absorption spectrum of mgLAC (~13 μM in 20 mM Tris–HCl, pH 8.0) was recorded at 25 °C on a Jasco UV/VIS spectrophotometer (model V-550). The EPR spectrum was recorded on a Bruker ESP350E spectrometer at 77 K using 0.13 mM enzyme in 20 mM Tris–HCl (pH 8.0). The inductively coupled plasma-atomic emission spectrometry (ICP-AES) measurement was carried out using 0.13 mM enzyme in 20 mM Tris–HCl (pH 8.0) on a SII NanoTechnology SPS4000.

2.3. Bioinformatics tools

A BLAST search was carried out in the non-redundant database of GenBank, via the internet at http://www.ncbi.nlm.nih.gov/BLAST/ [19]. Multiple sequence alignment was carried out using a web-based version of ClustalW (http://www.ebi.ac.uk/clustalw/), with default parameter settings [20]. To analyze protein domain structures, the simple modular architecture research tool (SMART) was used (http://smart.embl-heidelberg.de/) [21]. For identification of signal peptides, the SignalP facility in SMART was used [22].

2.4. X-ray crystallographic analysis

Crystallization and preliminary X-ray crystallographic results are described as previously noted [18]. The crystal belongs to space group P212121, with unit-cell parameters of a = 74.67, b = 100.95, and c = 124.11 Å. The structure was solved by the single-wavelength anomalous diffraction (SAD) technique using Cu atoms. The heavy-atom refinement, density modification, and initial structure modeling were performed using autoSHARP [23]. Further model building and structure refinement were carried out using the COOT [24] and reflmac [25] programs. The progress and validity of the refinement process were checked by monitoring the R-free value for 5% of the total reflections [26]. Model geometry was analyzed using the MOLPROBITY program [27]. The data collection and refinement statistics are summarized in Table 1. The root-mean-square deviation was calculated using the LSQMAN program [28]. Domain interfaces were analyzed using the Protein–Protein Interaction Server [29]. The figures were prepared by the PyMOL program (http://pymol.sourceforge.net), using the coordinates from PDB files 2BW4 (NIR from Achromobacter cycloclaster), 1AOZ (AO from Cucurbita pepo var. melopepo), and 2J5W (CP from human).

3. Results and discussion

3.1. Identification of a two-domain laccase in a metagenomic library

In silico screening of our in-house metagenomic sequence database identified an open reading frame containing two copper-binding motifs [18]. The protein, designated mgLAC, is comprised of...
359 residues, which is small compared to three-domain type laccases commonly found in fungi [30] and bacteria [31]. At the N-terminus, a putative signal peptide consisting of 34 amino acid residues was identified.

The amino acid sequence of mgLAC was compared to the all non-redundant GenBank cds database [19]. mgLAC shared the highest overall identities (51–58%) with multi-copper oxidases from *Nitrosomonas eutrophia* (YP_747619), *Nitrosomonas europaea* (NP_841001), *Nitrobacter hamburgensis* (YP_578478), and *Sphingomonas wittichii* (YP_001262290). All of these proteins are type [C], which contain two cupredoxin domains (Fig. 1). The enzyme from *Nitrosomonas europaea* was purified and characterized more than 20 years ago [12]. However, since this first report, no additional studies have been conducted, and thus, the structure–function relationship of the type [C] enzyme remains unclear. Other two-domain type laccases have been also documented: SLAC from *Streptomyces coelicolor* (343 aa, CAB45586, [7]) and EpoA from *Sphingomonas* sp. (NP_841001). These laccases belong to other laccase superfamilies (Fig. 1). The enzyme from *Nitrosomonas europaea* do not react with some phenolic substrates, such as 2,2′-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), 2,2-dimethyl-1,3-cyclobutanediol (DMPO), syringaldazine, and guaiacol. Ascorbic acid, which is usually not a laccase substrate, was a substrate here, demonstrating the broad substrate specificity of mgLAC. EpO does not react with some syringaldazine and guaiacol [8], and thus, substrate specificity appears to vary among two-domain laccases.

### 3.3. X-ray crystallographic analysis of the recombinant mgLAC

#### 3.3.1. Overall structure

The final model was refined to an R-factor of 0.161 at a resolution of 1.7 Å. There were 22 residues at the C-terminus (including His-tag) that remain undefined due to disorder. Most (98.8%) of the residues were found to be in favored regions of the Ramachandran plots. Each asymmetric unit in a crystal cell contained three tightly associated molecules of mgLAC, resulting in a triangular structure with approximate dimensions of 57 × 77 × 78 Å (Fig. 2). Each monomer is comprised of two cupredoxin-like domains (hereafter, Domain-1 denotes the first or N-terminal half domain and Domain-2 denotes the second or C-terminal half domain). The quaternary structure, consisting of six cupredoxin-like domains, is essentially the same as those of NIR, SLAC, and CP (Fig. 2).

#### 3.3.2. Structurally unconserved regions

An individual domain of mgLAC has essentially the same topology as that of other multi-copper proteins. Eight conserved β-strands, constituting the core structure of each cupredoxin domain, are connected by six structurally unconserved regions (SURs) [1] (Fig. 3). These unconserved regions are likely to be involved in the modulation of substrate recognition and molecular stability in multi-copper proteins.

Despite the overall similarity in the quaternary structures, a significant difference is found at the C-terminal region between mgLAC and NIR. In mgLAC, the C-terminal loop region (residues 278–318) fills the cleft between Domains 1 and 2, stabilizing the intra-molecular domain–domain interaction. In contrast, in NIR the corresponding cleft is filled by the N-terminal loop region (residues 7–36). In addition, the C-terminal loop region of NIR extends to interact with Domain-1 of the neighboring monomer of the trimer. This extended C-terminal region of NIR provides additional stabilizing interactions, resulting in much more extensive inter-domain–domain interfaces (~2000 Å²) than in mgLAC (~1500 Å²). In mgLAC, domain interfaces are filled with three coppers, which stabilize the protein. However, two of these three coppers are absent in NIR. Therefore, we believe the extended C-terminal residues in NIR compensate for the loss of stability due to the depletion in copper. This is also consistent with previous reports on CP. In CP, the interactions between Domain-1 and Domain-6 are significantly weaker than those observed in AO. In addition, a pair of T3 coppers between the domains plays the primary role of stabilizing CP [32].

#### 3.3.3. Copper-binding site

The anomalous difference Fourier map showed four clear peaks corresponding to the copper ions per subunit (12 copper ions per trimer; Figs. 2, 4 and Supplementary Fig. 3). The presence of four copper ions per monomer is consistent with the results obtained by ICP-AES analysis. There are two types of copper-binding sites (mono- and tri-nuclear) in the mgLAC monomer. Mononuclear T1...
copper is located in Domain-1, and tri-nuclear copper atoms (T2/T3) are embedded in the inter-molecular interface. The electron density corresponding to an oxygen atom was observed between T3 copper. The T1 copper site is easily accessible to the solvent.
which is presumed to account for the broad substrate specificity (Supplementary Table 1). The electron density of T2 copper is lower than the other coppers; the occupancy of T2 copper is estimated to be ~0.7, which supports the slight leak in copper at this site, being consistent with the results of ICP-AES analysis.

Both Domain-1 and -2 of mgLAC originated from a cupredoxin-like structural unit, but appear to have evolved differently. T1 copper is found only in Domain-1. A potential T1 site in Domain-2 is missing, due to the substitution of a Cys ligand to Asp to produce a “pseudo-” T1 site (Figs. 3 and 4). This asymmetric distribution of T1 sites is conserved in NIR, but differs from SLAC and three-domain type laccases that have a T1 site in Domain-3. This Domain-3 corresponds to the Domain-2 of the neighboring molecule, which shares the tri-nuclear copper site with Domain-1 in the mgLAC trimer (Fig. 2A and E). T1 copper of the SLAC trimer and CP are located near the surface of the central cavity (Fig. 2C and D). In mgLAC, the T1 site of Domain-1 and the pseudo-site of Domain-2 are related by a pseudo twofold symmetry. While the T1 copper site of mgLAC is positioned on the opposite side of the tri-nuclear copper site, in contrast with three-domain type laccases, structural features around the T1 copper site are similar to three-domain type laccases. The T1 copper is ligated by His57, His112, Cys105, and Met118 and is located approximately 12 Å away from the tri-nuclear cluster, which is linked by the electron-transferring Cys-His pathway (Fig. 4).

Supplementary Table 2 summarizes the copper ligands and B-factors of copper atoms. The residues involved in the ligation of T2/T3 copper ions are eight histidines, as in three-domain type laccases. However, the histidine ligands are donated by two subunits (Domain-1 of subunit A and Domain-2 of subunit B, as shown in Supplementary Fig. 4) that are related by a pseudo-twofold axis. One T2 copper is ligated by two histidines (His60A and His204B; “A” and “B” after the residue number represent subunit A and B in the trimeric form, respectively) and two T3 coppers are ligated by three histidines (His62A, His104A, and His254B in mgLAC), ligate to the mononuclear copper ion, and a fourth histidine, His255B (His204B in mgLAC), is located nearby, but is not involved in copper ligation. In addition to the fourth histidine, substitutions from potential copper ligand histidines (His252B, His206B, and His60A in mgLAC) to hydrophobic residues (Val304B and Ile257B) and aspartate (Asp98A) are likely responsible for the nitrite (but not oxygen) reduction (Fig. 4).

3.4. Evolutionary relationships of the multi-copper protein family

As shown in Fig. 2, structural studies confirmed the relationship of mgLAC to NIR and three-domain type laccases. The type [B] laccase, SLAC, also forms trimers [11], indicating that a homotrimer is the essential architecture of two-domain type multi-copper proteins. Variations in the copper-binding motif at domain interfaces...
seem to have created an array of functional diversity. Based on the evolutionary scheme proposed by Nakamura et al. [6], mgLAC-type laccases are likely the precursors of NIRs. Formation of intimate domain–domain interaction appears to be essential for the creation of functional diversity in multi-copper proteins. Gain, or loss, of inter-domain copper sites is coupled to the formation of quaternary structures. Multi-copper proteins favor three repeating units. Despite the difference in the oligomeric states of the polypeptide chains, the triangular formation of cupredoxin domains acts as the functional unit for all existing multi-copper proteins.

Acknowledgments

This work was partially supported by the GCOE Program (Y.H.), the Japanese Aerospace Exploration Agency Project (Y.H.), a Grant-in-Aid for Young Scientists (H.K.), the Hyogo Science and Technology Association (H.K.), and Interface Technology from JST (Y.H.), the Sumitomo Foundation (K.M.), the Japanese Aerospace Exploration Agency Project (Y.H.), a Grant-in-Aid for Scientific Research (18GS0207) (Y.H.), the basic research program CREST type, Development of the Foundation for Nano-Science and Technology (K.M.), and K. M. thanks Shiori Mizuta (H.K.), the Hyogo Science and Technology Association (H.K.), and Interface Technology from JST (Y.H.), the Sumitomo Foundation (K.M.)

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.febslet.2009.03.008.

References


Fig. 4. Stereo-view of the copper-binding sites of mgLAC (A), SLAC (B), NIR (C), and AO (D). Residues from molecule A (Domain-1) are shown in green and those from the adjacent molecule B (Domain-2) are shown in light green. T1 copper atoms are depicted as blue spheres and the other coppers are depicted as cyan spheres. For mgLAC (A), a local pseudo twofold axis is drawn as a dotted line.


