Infection, Genetics and Evolution 13 (2013) 49–55

Contents lists available at SciVerse ScienceDirect

Infection, Genetics and Evolution

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

The evolution and putative function of phosducin-like proteins in the malaria parasite *Plasmodium*

Catherine Putonti [a,b,c,*], Bryan Quach [c], Rachel L. Kooistra [a], Stefan M. Kanzok [a]

[a] Department of Biology, Loyola University Chicago, 1032 W Sheridan Rd, Chicago, IL 60660, USA
[b] Department of Computer Science, Loyola University, 820 N Michigan Ave., Chicago, IL 60611, USA
[c] Bioinformatics Program, Loyola University Chicago, 1032 W Sheridan Rd, Chicago, IL 60660, USA

ARTICLE INFO

Article history:
Received 2 July 2012
Received in revised form 29 August 2012
Accepted 31 August 2012
Available online 17 September 2012

Keywords:
Phosducin-like proteins
Thioredoxin-fold domain
Plasmodium

ABSTRACT

Ubiquitous to the proteomes of all living species is the presence of proteins containing the thioredoxin (Trx)-domain. The best characterized Trx-domain containing proteins include the enzymes involved in cellular redox metabolism facilitated by their cysteine-containing active site. But not all members of the Trx-fold superfamily exhibit this catalytic motif, e.g., the phosducin-like (PhLP) family of proteins. Genome sequencing efforts have uncovered new Trx-domain containing proteins, and their redox activity and cellular functions have yet to be determined. The genome of the malaria parasite *Plasmodium* contains multiple thioredoxins and thioredoxin-like proteins which are of considerable interest given their role in the parasite's antioxidant defense. While adaptations within the Trx-domain have been studied, primarily with respect to redox active structures, PhLP proteins have not been examined. Using the uncharacterized phosducin-like protein from *Plasmodium berghei* PhLP-1, we investigated the evolution of PhLP proteins across all branches of the tree of life. As a result of our analysis, we have discovered the presence of two additional PhLP proteins in *Plasmodium*, PhLP-2 and PhLP-3. Sequence homology with annotated PhLP proteins in other species confirms that the *Plasmodium* PhLP-2 and PhLP-3 belong to the PhLP family of proteins. Furthermore, as a result of our analysis we hypothesize that the PhLP-2 thioredoxin was lost over time given its absence from higher-order eukaryotes. Probing deeper into the putative function of these proteins, inspection of the active sites indicate that PbPhLP-1 and PbPhLP-2 may be redox active while PbPhLP-3 is very likely not. The results of this phylogenetic study provide insight into the emergence of this family of Trx-domain containing proteins.

1. Introduction

The thioredoxin (Trx)-fold is a highly conserved protein structure that occurs in proteins with very diverse functions and biochemical pathways (Martin, 1995), including but not restricted to redox metabolism (Holmgren et al., 2005), protein folding (Ito and Inaba, 2008), cell proliferation (Immenschuh and Baumgart-Vogt, 2005), and signal transduction and transcriptional regulation (Brigelius-Flohé and Flohé, 2011). Thioredoxin-fold containing proteins can be found in organisms throughout all branches of the tree of life, from bacteria to higher-order eukaryotes. The basic structure of this thioredoxin-fold exhibits a beta sheet, consisting of four beta strands, sandwiched by at least two alpha helices (βαββββα) (Martin, 1995; Atkinson and Babbitt, 2009).

Proteins that contain one or more Trx-fold domains are generally consolidated in the Trx superfamily, also called the Trx-fold superfamily. The Pfam database of the Sanger Institute lists a thioredoxin clan (PFam ID: CL0172) containing 43 member families (Punta et al., 2012). The best-characterized proteins of this superfamily are enzymes involved in the cellular redox metabolism, which includes the canonical single domain redox-active thioredoxins. Their principal biochemical reaction mechanism is based on the oxidation and reduction of the thiol-groups of one or more cysteine residues in their active sites (Martin, 1995). Therefore early organization of the Trx-superfamily was based, in addition to the Trx-fold, on the presence of a cysteine or selenocysteine containing active site in a characteristic [Cys/SeCys-X1-X2-Cys] motif. The protein family was then expanded following the characterization of Trx-fold containing proteins with atypical redox active sites, such as glutatredoxins and peroxiredoxins which exhibited [Ser/Thr-X1-X2-Cys] or [Cys-X1-X2-Ser/Thr] motifs (Atkinson and Babbitt, 2009). Furthermore, it was shown that the molecular mechanisms of the dithiol and the monothiol containing proteins are related but distinct from one another (Fernandes and Holmgren, 2004).

* Corresponding author at: Department of Biology, Loyola University Chicago, 1032 W Sheridan Rd, Chicago, IL 60660, USA. Tel.: +1 773 508 3277; fax: +1 773 508 3646.
E-mail address: cputonti@luc.edu (C. Putonti).

1567-1348/$ - see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.meegid.2012.08.023

In addition to the single domain thioredoxins and thioredoxin-like proteins, the recent explosion of genome sequencing data has led to the discovery and characterization of Trx-domain containing proteins that do not contain a discernible cysteine-containing active site. It is also not clear whether they exhibit redox activity or whether the Trx-fold domain plays primarily a structural role. The single Trx-fold domain proteins containing a redox active site are referred to as Class I Trx. Conversely, multi-domain proteins, that contain one or more Trx-fold domains and vary in possessing a redox active site, belong to the Class II Trx or Trx-like proteins (Hirt et al., 2002). The number of Class II Trx-like proteins has dramatically increased in recent years and these proteins form numerous families in the Trx-fold superfamily of proteins.

The phosducin-like (PhLP) family is a Class II thioredoxin and includes the signaling protein phosducin that interacts with trimeric G-proteins upon activation (Bauer et al., 1992). It was first discovered within the retina and pineal glands of mammals (Lee et al., 1990; Reig et al., 1990). Subsequently, phosducin and phosducin-related proteins have been found in vertebrates as well as lower eukaryotes (Blaauw et al., 2003). The PhLP family (Pfam ID: PF02114) contains three subfamilies: *PhLP Phd* proteins, involved in G-protein related signal transduction; *PhLP viral inhibitors of apoptosis* (VIAF) involved in apoptosis signaling; and *PhLP TxnDC9* which have been shown to interact with cytoskeletal elements, such as actin and tubulin (Blaauw et al., 2003; Stirling et al., 2006). Phosducins do not exhibit a classical catalytic motif and it is therefore not clear if they are redox active enzymes (Atkinson and Babbitt, 2009; Lou et al., 2009).

In the malaria parasite *Plasmodium*, thioredoxins and thioredoxin-like proteins are of major interest, as they comprise a large part of its antioxidant defense arsenal (Müller, 2004). Their inhibition would leave the parasite defenseless to oxidative stresses. Consequently, these proteins represent potential drug targets. At present there are over 40 putative thioredoxin, thioredoxin-like and thioredoxin-related genes annotated in the *Plasmodium berghei* genome (PlasmoDB, Aurrecoechea et al., 2009). Included in this list is a gene annotated as a *putative thioredoxin* (PlasmoDB ID: PBANKA_120480) and as a hypothetical protein in GenBank (GenBank ID: XP_674598). PBANKA_120480 is a multi-domain protein consisting of an N-terminal helical domain and a C-terminal Trx-fold domain which does not contain a classical [Cys-X1-X2-Cys] active site motif, thus designating it as a Class II Trx-like protein. Published transcriptomics and proteomics data indicates that PBANKA_120480 has a biological function in the malaria parasite (e.g., Hall et al., 2005; Khan et al., 2005). Furthermore, its primary sequence shows significant homology with the human PhLP3 or Thioredoxin-domain containing protein #9 (TxnDC9) (Marchler-Bauer et al., 2009).

The evolution of the Trx fold has been studied, mostly with respect to redox active dithiol and monothiol containing enzymes (e.g., Copley et al., 2004; Meyer et al., 2006; Atkinson and Babbitt, 2009; Hall et al., 2009; Pedone et al., 2010). In an effort to better understand the emergence and functionality of PhLP proteins and their putative role in malaria, we conducted an extensive search and analysis of homologous sequences across all domains of life. At present, there are no *Plasmodium* genes which have been identified in PlasmoDB as PhLP (Aurrecoechea et al., 2009). This search was initiated using two *P. berghei* genes, PBANKA_120480 and PBANKA_051970 (PlasmoDB, Aurrecoechea et al., 2009). Although neither of these genes has been biologically characterized, the high sequence similarity between PBANKA_120480 and human PhLP3 suggests that this is potentially a phosducin-like protein. Given our prior examination of their amino acid sequences and the presence of the Trx-fold domain, we introduce the gene designations of PbPhLP-1 for PBANKA_120480 and PbPhLP-2 for PBANKA_051970, respectively. Integrating a phylogenetic approach and an analysis of the active sites of these proteins, we can better understand how these proteins, which emerged from prokaryotes, have evolved over millennia.

2. Methods

2.1. Plasmodium genes of interest

While there are several annotated thioredoxin and thioredoxin-like proteins within the *Plasmodium* genome, we selected two for our analysis: the putative thioredoxin PBANKA_120480 and PBANKA_051970, annotated within PlasmoDB as "putative thioredoxin" and "conserved Plasmodium protein, unknown function", respectively (Aurrecoechea et al., 2009). Both of these genes are from the genome of *P. berghei* ANKA, a species of malaria, which infects rodents and is used as a model system in malaria research (Sinden, 1978; Aikawa and Seed, 1980). PlasmoDB lists PBANKA_051970 as a paralog to PBANKA_120480 (Aurrecoechea et al., 2009). We refer to these two sequences as *P. berghei* phosducin-like proteins, PbPhLP-1 and PbPhLP-2, respectively.

2.2. Sequence comparison

The PbPhLP-1 and PbPhLP-2 amino acid sequences, retrieved from PlasmoDB (Aurrecoechea et al., 2009), were aligned using EBI's CLUSTALW2 tool online (Larkin et al., 2007). The conserved thioredoxin domain (phosducin (Phd)-like family, Thioredoxin (TRX) domain containing protein 9 or Phd_like_TxnDC9) was identified within each sequence via NCBI's Conserved Domain Database (Marchler-Bauer et al., 2009). The sequence identity between the two sequences was calculated based upon the CLUSTALW alignment in BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html).

2.3. Identifying paralogs and homologs

The PbPhLP-1 and PbPhLP-2 amino acid sequences were BLASTed against the NCBI nr protein database using blastp, retrieving the 1000 best hits given a word size of two (Altschul et al., 1997). (It is important to note that while the PbPhLP-1 sequence is a record in NCBI's Protein database, GenBank ID: XP_674598, PbPhLP-2 in its entirety does not have a record; GenBank ID: XP_680424 includes 74 of the 213 amino acid PbPhLP-2 sequence). The PbPhLP-1 sequence BLAST results all had an e-value less than 0.002; the PbPhLP-2 sequence BLAST results all had an e-value less than 0.037.

The BLAST results were downloaded and parsed, using in-house parsers developed in C++, producing FASTA format files. Two FASTA format files were created for the PbPhLP-1 and PbPhLP-2 searches – the homologous sequences and the orthologous sequences. The homologous files include all 1000 hits for the individual BLAST. These files include orthologous sequences as well as paralogous sequences as there are many copies of the thioredoxin domain within the genomes of all species, not just *Plasmodium*. While the vast majority of the sequences returned for the BLAST searches encompassed more than just the thioredoxin domain, there were several smaller partial sequences, many of which are annotated as thioredoxins or putative thioredoxins.

The second pair of FASTA files created contained only orthologs of PbPhLP-1 and orthologs of PbPhLP-2. The BLAST of the PbPhLP-2 sequence against NCBI's Protein database identified its paralog, PbPhLP-1, with an e-value of 8×10^{-12}, 68% coverage, and 32% identity. Using this as our threshold, we restricted orthologs to hits with a coverage $\geqslant$68% and a sequence identity >32%. This is, however, a stringent threshold as two orthologous sequences from very divergent species can share a lower identity than two paralogs in