Available online at www.sciencedirect.com
ScienceDirect
Biochimica et Biophysica Acta 1783 (2008) 535–548
Biochimica et Biophysica Acta
BBA
www.elsevier.com/locate/bbamcr

Review

The human PDI family: Versatility packed into a single fold

Christian Appenzeller-Herzog, Lars Ellgaard *

Department of Molecular Biology, Universitetsparken 13, University of Copenhagen, DK – 2100 Copenhagen Ø., Denmark

Received 18 September 2007; received in revised form 16 November 2007; accepted 20 November 2007
Available online 3 December 2007

Abstract

The enzymes of the protein disulfide isomerase (PDI) family are thiol–disulfide oxidoreductases of the endoplasmic reticulum (ER). They contain a CXXC active-site sequence where the two cysteines catalyze the exchange of a disulfide bond with or within substrates. The primary function of the PDIs in promoting oxidative protein folding in the ER has been extended in recent years to include roles in other processes such as ER-associated degradation (ERAD), trafficking, calcium homeostasis, antigen presentation and virus entry. Some of these functions are performed by non-catalytic members of the family that lack the active-site cysteines. Regardless of their function, all human PDIs contain at least one domain of approximately 100 amino acid residues with structural homology to thioredoxin. As we learn more about the individual proteins of the family, a complex picture is emerging that emphasizes as much their differences as their similarities, and underlines the versatility of the thioredoxin fold. Here, we primarily explore the diversity of cellular functions described for the human PDIs.

Keywords: Disulfide-bond formation; Endoplasmic reticulum; Protein disulfide isomerase; PDI; Protein folding

1. Introduction

Disulfides are covalent bonds produced by oxidation of two free thiols — in proteins they form between two cysteine residues. They serve important functions during protein folding and in stabilizing protein structures – both as intramolecular bonds and when formed intermolecularly between two protein chains – but also through their capacity to work as regulatory switches in redox signaling.

Mammalian cells produce a fascinating collection of complicated disulfide-bonded structures. Disulfides are present in a large fraction of the close to 1/3 of all human proteins that traverse the secretory pathway [1]. The synthesis and folding of these proteins take place in the ER where the environment is conducive to disulfide-bond formation — the ER lumen is relatively more oxidizing than the cytosol, and a number of enzymes catalyze the formation of native disulfides. This activity was first shown for PDI [2] that is an abundant protein of the ER and the founding member of the PDI family (the members of which we refer to as PDIs). It is an essential protein in *S. cerevisiae* [3], and the human as well as the yeast protein have been thoroughly investigated in vitro. Based on these and similar studies on other family members, we now have a quite detailed understanding of the structure and enzymatic properties of single catalytically active thioredoxin-like domains, the basic unit of PDIs. As the name implies, these domains are structural homologs of the cytosolic reductase thioredoxin.

Current work is directed at understanding the interplay between domains within the multi-domain structure found in most PDIs. For PDI itself, the domain composition has been known for years [4–6] and recently the crystal structure of the entire protein from *S. cerevisiae*, Pdi1p, was solved [7] (Fig. 1A and B). Another center of attention is the physiological function of individual proteins, including studies of their cellular redox regulation and the identification of endogenous substrates. Together, these topics also constitute the main focus of this review, where we will limit ourselves to the mammalian system and the processes that take place in the ER unless directly otherwise stated (for recent studies on the extracellular function of P5 and ERp57 see [8,9]).

2. The proteins of the human PDI family — an overview

With the increasing number of cDNA sequences deposited in the public domain in recent years, the number of known human

* Corresponding author. Tel.: +45 35 32 17 25; fax: +45 35 32 15 67.
E-mail address: lellgaard@aki.ku.dk (L. Ellgaard).

0167-4889/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbamcr.2007.11.010

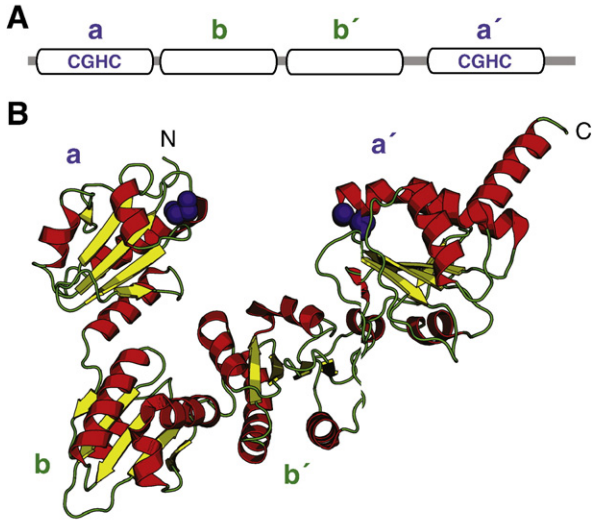

Fig. 1. Domain composition and three-dimensional structure of PDI. A, schematic overview of the PDI domain composition. Thioredoxin-like domains are shown as rounded rectangles with purple lettering denoting catalytic domains (**a** and **a′**) and green lettering non-catalytic domains (**b** and **b′**). The active-site sequences are shown. The C-terminal extension beyond the **a′** domain constitutes the prominent α-helix visible at the C-terminus in the crystal structure of Pdi1p (B). The α-helices are shown in red, β-strands in yellow and loops in green. The active-site cysteines are depicted as purple spheres.

PDIs has also quite rapidly expanded. The family comprises 19 published members that all contain a predicted signal sequence and at least one thioredoxin-like domain (Table 1)[10–12]. Although the family name implies an isomerase function, this activity has not been demonstrated experimentally for all. PDILT has been shown not to have any appreciable oxidoreductase activity [13], which probably also holds true for other PDIs with only one cysteine in the active site, such as the newly discovered Hag2 and Hag3 [14]. Moreover, ERp27 and ERp29 are non-catalytic family members. It is therefore important to stress that the grouping of proteins into the human PDI family is not based on one common function or the exact same enzymatic properties, but rather on sequence and structural similarity as well as ER localization (see below).

Table 1 provides an overview of important traits of the human PDIs. The unifying feature in terms of sequence and structure is the thioredoxin-like domains that can be either catalytic or non-catalytic. By convention these two types of domains are called **a**- and **b**-type domains, respectively. The **a**-type domains usually contain two cysteines in a CXXC active-site motif with an intervening GH sequence being the most common in the PDIs. By sequence similarity rather than catalytic activity some domains lacking one or even two active-site cysteines are also categorized as **a**-type domains (Table 1). The **b**-type domains do not have cysteines in the active site and are therefore not redox active. Generally, the sequence similarity is much higher for **a**-type than **b**-type domains since the latter are missing a number of residues important for catalysis that are conserved among the **a** domains. Only two PDIs contain additional types of domains in the ER — ERdj5 has an N-terminal J-domain that binds and stimulates the ATPase activity of the ER chaperone BiP in vitro [15,16], and ERp29 comprises an α-helical so-called D-

Table 1
Overview of the proteins of the human PDI family listed according to size with the soluble proteins above the thick line and the transmembrane proteins below

Name	Accession	Length	ER-localization motif	Domain composition	Number of **a**-type domains	Active-site sequence	PDB accession numbers[a]
Hag 3	Q8TD06	165	QSEL[b]	**a**	1	CQYS	
ERp18	O95881	172	EDEL	**a**	1	CGHC	1sen
Hag 2	O95994	175	KTEL[b]	**a**	1	CPHS	
ERp28/29[c]	P30040	261	KEEL	**b–D**	0	n.a.	1g7e; 2c0e; 1ovn
ERp27	Q96DN0	273	KVEL	**b–b′**	0	n.a.	
ERp44	Q9BS26	406	RDEL	**a–b–b′**	1	CRFS	
ERp46[d]	Q8NBS9	432	KDEL	**a°–a–a′**	3	CGHC, CGHC, CGHC	**a°**: 2diz
P5	Q15084	440	KDEL	**a°–a–b**	2	CGHC, CGHC	**a**: 2dml; **a′**: 1x5d
ERp57	P30101	505	QEDL	**a–b–b′–a′**	2	CGHC, CGHC	**a**: 2alb; **b–b′**: 2h8l; **a′**: 2dmm
PDI	P07237	508	KDEL	**a–b–b′–a′**	2	CGHC, CGHC	**a**: 1mek; **b**: 2bjx; **a′**: 1x5c; Pdi1p: 2b5e
PDIr	Q14554	519	KEEL	**b–a°–a–a′**	3	CSMC, CGHC, CPHC	
PDIp	Q13087	525	KEEL	**a–b–b′–a′**	2	CGHC, CTHC	
PDILT	Q8N807	584	KEEL	**a–b–b′–a′**	2	SKQS, SKKC	
ERp72	P13677	645	KEEL	**a°–a–b–b′–a′**	3	CGHC, CGHC, CGHC	**a**: 2dj1; **a′**: 2dj2; **a°**: 2dj3
ERdj5	Q8IXB1	793	KDEL	**J–a″–b–a°–a–a′**	4	CSHC, CPPC, CHPC, CGPC	
TMX	Q9H3N1	280	Unknown	**a**	1	CPAC	1x5e
TMX2[c]	Q9Y320	296	KKDK	**a**	1	SNDC	2dj0
TMX4	Q9H1E5	349	RQR	**a**	1	CPSC	
TMX3	FLJ20793	454	KKKD	**a–b–b′**	1	CGHC	

The lengths include the predicted signal sequence and the grey coloring of thioredoxin-like domains denotes those for which a three-dimensional structure has been solved. The assignment of domains is based on previously published data and our own bioinformatics analysis. J: J-domain, D: D-domain.

[a]PDB accession numbers are given only for structures of metazoan proteins with the exception of the Pdi1p protein from *S. cerevisiae*.

[b]The ER localization of Hag2 and Hag3 has been confirmed for myc-tagged variants of both proteins (Ruddock, L., personal communication).

[c]The names ERp28 and ERp29 are often used interchangeably in the literature, although the human protein was originally named ERp28 and the rat protein ERp29.

[d]ERp46 is also known as EndoPDI.

[e]The ER localization for TMX2 has not been experimentally confirmed. Moreover, it is unclear whether its thioredoxin-like domain faces the ER or the cytosol.

n.a. — not applicable.