Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Structural Biology 151 (2005) 130–150
Journal of Structural Biology
www.elsevier.com/locate/yjsbi

PEDF and the serpins: Phylogeny, sequence conservation, and functional domains

Joyce Tombran-Tink [a,b,*], Samuel Aparicio [a], Xuming Xu [b], Amber R. Tink [a], Nuria Lara [a], Supriya Sawant [b], Colin J. Barnstable [b], Samuel Shao-Min Zhang [b]

[a] *Division of Pharmaceutical Sciences, University of Missouri-Kansas City, Kansas City, MO 64110, USA*
[b] *Department of Ophthalmology and Visual Science, Yale University School of Medicine, New Haven, CT 06520, USA*

Received 25 August 2004, and in revised form 28 April 2005
Available online 9 June 2005

Abstract

Pigment epithelium derived factor (PEDF) is non-inhibitory serpin with neurotrophic and antiangiogenic functions. In this study, we have assembled PEDF sequences for 9 additional species by data base mining and performed cross-species alignment for 14 PEDF sequences to identify conserved structural domains. We found evolutionary conservation of a leader sequence, a single C-terminal glycosylation site, collagen-binding residues, and four specific conserved PEDF peptides. The C-terminus, 384–415 and an N-terminal region 78–95, show close homology with many other serpins, and there is strong conservation of 39 of 51 consensus key residues involved in serpin structure and function. Two peptide regions, 40–67 and 277–301, are unique to PEDF but conserved in all species. Conserved residues at the N-terminus, helix d (hD), and helix A (hA) of PEDF form a structure similar to the heparin-binding groove of other serpins. We identified a motif in PEDF that is homologous to the nuclear localization signals of other proteins. A bitopographical localization of PEDF was confirmed by immunocytochemistry and Western blots. Our results suggest that secretion is required for PEDF's activity, that PEDF can migrate to the nucleus, and that PEDF has structural and functional features more common with inhibitory serpins.

Keywords: PEDF; Serpins; Bioinformatics tools; Protein alignment; Reactive center loop; Peptides; Structure–function relationships; Heparin binding; Nuclear localization signal; Glycosylation

1. Introduction

Serpins are a large family of proteins that include the predominant protease inhibitors in mammals. They have an extraordinary ability to bind, with a high degree of selectivity, to a wide range of proteases. This function is tightly controlled by structural constraints of the serpin molecule that allow them to modulate key processes controlling proteolytic cascades such as blood coagulation, fibrinolysis, and complement activation. Although the largest group of serpins is known to be protease inhibitors, a substantial number have not been shown to have such inhibitory activity. It remains an open question as to whether these serpins have inhibitory activity or are functionally different from those that can inhibit specific target proteases.

Pigment epithelium derived factor (PEDF), an inhibitor of angiogenic processes and a neurotrophic protein, is a member of the serpin gene family (Steele et al., 1993). It is an extracellular serpin that shows the typical serpin secondary and tertiary structure (Simonovic et al., 2001) and binds to heparin and collagen type 1 (Alberdi et al., 1998, 2003; Kozaki et al., 1998; Meyer et al., 2002). The PEDF gene is highly conserved in evolution and resides on human chromosome

* Corresponding author. Fax: +1 816 235 5190.
E-mail address: jttink@aol.com (J. Tombran-Tink).

1047-8477/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jsb.2005.05.005

17p13.3 (Tombran-Tink et al., 1994; Xu et al., submitted). PEDF is grouped in serpin clade f subgroup 1 and its closest serpin relatives are α2-antiplasmin (clade f subgroup 2) and C1-inhibitor (clade g) (Xu et al., submitted).

The crystal structures of many serpins show that they are α/β proteins with three β-sheets (A, B, and C) and nine α-helices (Silverman and Lomas, 2004). The Aβ-sheet is the dominant unit of secondary structure and forms the core structural domain of the protein as well as being intimately involved in the dynamic movements that are part of serpin function (Gettins, 2002). Another remarkable feature of serpins is their reactive center loop (RCL), which is the proteinase recognition site and a critical component of the function of serpins (Carrell and Huntington, 2003). The extended loop is complementary to the substrate-binding site of the target proteinases and contains a stretch of approximately 17 amino acids tethered between the A and C β-sheets. The RCL of some serpins such as α1-antitrypsin is exposed and poised for action. Other serpins, such as antithrombin, a plasma serine proteinase inhibitor, normally exist in a state with an RCL that is partially inserted into the center of the A sheet (Whisstock et al., 2000a). Antithrombin becomes a more potent inhibitor only when the RCL is expelled from the protein and assumes a conformation more suitable for interaction with the target proteinase (Quinsey et al., 2004). The complex conformational change activating antithrombin is mediated by its interaction with other ligands such as heparin and related glycosaminoglycans.

The function of the RCL in PEDF is still not known, as is true for most serpins that do not inhibit specific proteinases. Site-directed mutagenesis studies of PEDF have suggested that one function of the highly exposed PEDF RCL is associated with the secretion of the serpin (Shao et al., 2003).

In this study, we used data mining tools to compare the sequence of the human PEDF protein with a wide range of species and to examine its structural homology with inhibitory and non-inhibitory serpins. This strategy has allowed us to identify regions of PEDF that show striking cross-species conservation, highly conserved peptides that are unique to PEDF, a C-terminus that is homologous with many other serpins, a heparin-binding site, and a possible nuclear localization signal. The structural features we have identified in this study may represent key functional domains of the PEDF molecule.

2. Material and methods

2.1. Data source

The protein sequences for PEDF in other species as well as human and mouse serpins were obtained from NCBI GenBank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Protein) and our previous paper (Xu et al., submitted).

The crystal structures of PEDF and other serpins were obtained from NCBI structure database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Structure). RasMol (http://openrasmol.org/OpenRasMol.html) and a RasMol-derivative, Protein Explorer, (http://molvis.sdsc.edu/ protexpl/pe_lit.htm) software packages were used to analyze the macromolecular structure of PEDF and other heparin-binding serpins. The following proteins were examined for structure–function relationships with PEDF.

Protein	PDB	Accession No.	Chain	Primary references
PEDF	1imv	gi:15988024	Chain A	Steele et al. (1993), Simonovic et al. (2001)
Antithrombin	2ant	gi:2392668	Chain I	Whisstock et al. (2000a)
Heparin cofactor 2	1jmj	gi:23200173	Chain B	Baglin et al. (2002)
Plasminogen activator inhibitor I	1c5g	gi:6730157	Chain A	Ehrlich et al. (1992)

The serpin sequence alignments studies were performed using our currently submitted serpin sequences and the dataset by (Irving et al., 2000) (http://www.med.monash.edu.au/biochem/research/projects/serpins/serpins). The amino acid numbering given throughout this manuscript is based on the full-length human PEDF sequence (Accession No. M76979), not the numbering of Fig. 1 or the numbering used by Simonovic et al. (2001).

2.2. Protein sequence alignment

The non-redundant protein database nr (ftp://ftp.ncbi.nih.gov/blast/db/nr.tar.gz) was used to search for homologies between PEDF peptides and other proteins. FASTA biological sequence comparison programs (http://fasta.bioch.virginia.edu/; http://www.ddbj.nig.ac.jp/search/fasta-e.html), BLAST (http://www.ddbj.nig.ac.jp/search/blast-e.html), PSI-BLAST (http://www.ddbj.nig.ac.jp/search/psi_blast-e.html) and CLUSTALW (http://www.ddbj.nig.ac.jp/search/clustalw-e.html) were used to search protein and DNA sequence databases to obtain alignments and homologies between PEDF and other proteins. ClustalW (http://www.ebi.ac.uk/clustalw/) was used to construct the alignment of PEDF in different phyla. The cross-species PEDF RCLs alignments, the alignments of the RCL of several serpins, and the putative PEDF NLS in other proteins were defined using the multi-alignment results of the PEDFs generated by ClustalW.

2.3. Peptide sequence comparison

The amino acid homology of PEDF peptides across species was calculated based on an alignment of