**Pfam**

Pfam is a database of a large collection of protein families, where each family is represented by multiple sequence alignments and Hidden Markov Models. The definition of domain boundaries, family members and alignment is done semi-automatically based on expert knowledge, sequence similarity, other protein family databases and the ability of HMM-profiles to correctly identify and align the members[1]. Pfam-A is manually curated and contains well-characterized protein domain families with high quality assignment. One can search the database for a given protein sequence against the Pfams collection of HMMs and its corresponding domain assignment by visiting http://pfam.xfam.org.

instead you should just write about how a HMM profile is constructed at the domain level and proteins are classified at the domain level in this database.

Pfam is a database of protein domains - so you introduce it as a resource for domain information.

Rewrite.

**References**

[1]: Identifying protein domains with the Pfam database Finn R, Griffiths-Jones S, Bateman A.Curr Protoc Bioinformatics. 2003 May;Chapter 2:Unit 2.5. doi:10.1002/0471250953.bi0205s01.