



ARCPHdb: A comprehensive protein database for SF1 and SF2 helicase from archaea



Mirna Moukhtar^{a,1}, Wafi Char^{b,1}, Ziad Abdel-Razzak^c, Mohamad Khalil^a, Samir Taha^c, Hala Chamieh^{c,d,*}

^a Centre Azm pour la recherche en Biotechnologie et ses applications, Ecole Doctorale Sciences et Technologies, Mitein Street, Tripoli, Lebanon

^b Hloul Business Analytics Solutions, Omar Daouk Street, Beirut, Lebanon

^c Laboratoire de Biotechnologies appliquées, Centre Azm pour la recherche en Biotechnologie et ses applications, Ecole Doctorale Sciences et Technologies, Mitein Street, Tripoli, Lebanon

^d Faculty of Science, Department of Biology, Lebanese University, Tripoli, Lebanon

ARTICLE INFO

Keywords:

Archaea
Superfamily helicase 1
Superfamily helicase 2
Protein database

ABSTRACT

Purpose: Superfamily 1 and Superfamily 2 helicases, two of the largest helicase protein families, play vital roles in many biological processes including replication, transcription and translation. Study of helicase proteins in the model microorganisms of archaea have largely contributed to the understanding of their function, architecture and assembly. Based on a large phylogenomics approach, we have identified and classified all SF1 and SF2 protein families in ninety five sequenced archaea genomes. Here we developed an online webserver linked to a specialized protein database named ARCPHdb to provide access for SF1 and SF2 helicase families from archaea.

Methods: ARCPHdb was implemented using MySQL relational database. Web interfaces were developed using Netbeans. Data were stored according to UniProt accession numbers, NCBI Ref Seq ID, PDB IDs and Entrez Databases.

Results: A user-friendly interactive web interface has been developed to browse, search and download archaeal helicase protein sequences, their available 3D structure models, and related documentation available in the literature provided by ARCPHdb. The database provides direct links to matching external databases.

Conclusions: The ARCPHdb is the first online database to compile all protein information on SF1 and SF2 helicase from archaea in one platform. This database provides essential resource information for all researchers interested in the field.

1. Introduction

Archaea, the third domain of life, have long been recognized as important model microorganisms [1]. Known to be prokaryotes, they share with eukaryotes important cellular processes such as replication, transcription and translation [2]. At the middle of these processes are essential motor protein enzymes that use the energy of ATP to bind and remodel nucleic acids termed helicases [3]. Helicases are one of the largest protein families and have been classified into 6 superfamilies SF (From SF1 to SF6) according to enzyme sequence, function and oligomeric state [4,5]. SF1 and SF2 are the non-hexameric helicases, they are the most widespread and found in all genomes sequenced to date. These families were subject to intensive research over the past few years, since they are linked to numerous diseases including cancer,

neurodegenerative diseases and infectious diseases [6,7]. All helicases share a conserved domain termed as helicase core that consists of nine characteristic sequence motifs named Q, I, Ia, Ib and II to VI [5]. Motifs I and II also known as Walker A and B are highly conserved and serve to bind NTPs. The other motifs are less conserved and vary between SF1 and SF2. SF1 and SF2 can be further subdivided into two groups A and B according to their translocation polarity: SF1/2 A has a 3'–5' translocation polarity and SF1/2B has a 5'–3' translocation polarity [8,9]. Fairman-Williams and co-workers proposed a classification of SF1 and SF2 protein families based on typical sequence, structural, and mechanistic features [10]. SF1 includes three families: (1) the UvrD-like/Rep family, (2) Pif-1-like family, and (3) Upf1-like family [11]. SF2 consists of nine families: (1) Rec-G like family, (2) RecQ-like family, (3) XPD/Rad3/DinG, (4) Ski-2 like family, (5) Type 1

* Corresponding author at: Faculty of Science, Department of Biology, Lebanese University, Tripoli, Lebanon.

E-mail address: hala.chamieh@ul.edu.lb (H. Chamieh).

¹ The two authors equally contributed to the work.

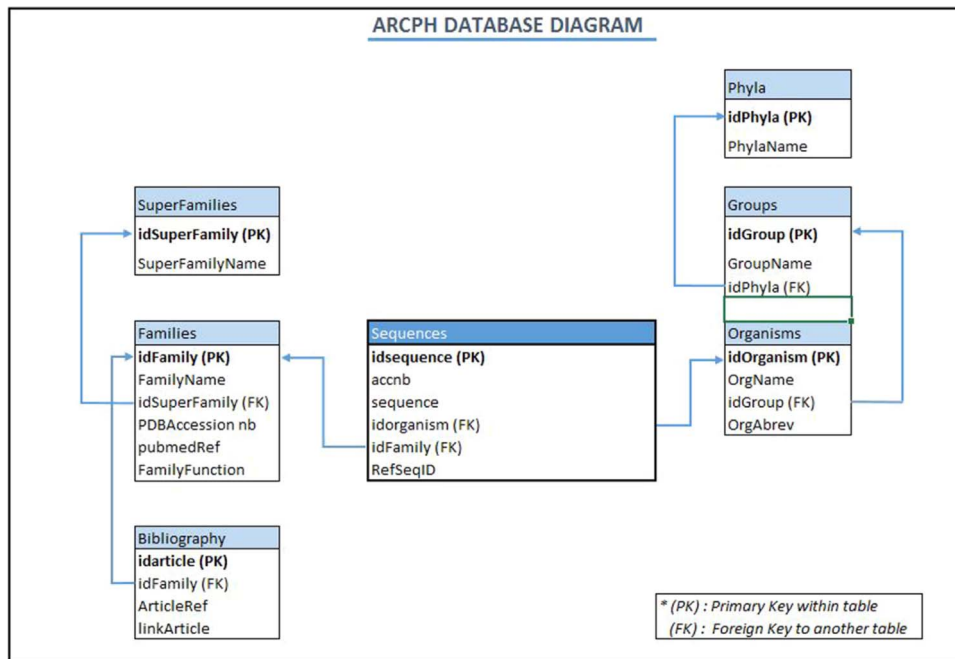


Fig. 1. Dataflow of the ARCPH helicase database.

ARCPHdb
ARChaea Protein Helicase Database

Home About us Research Interests Contact

SF1

- PIF-1 like
- Upf1-like
- UvrD-like

SF2

- ASH
- Cas3
- DEAD-box
- Hjm/Hel308
- Lhr
- Rad3-like
- RecQ
- Reverse Gyrase
- Sfth
- snf2
- T1R
- UvrB
- XPB/ Rad25
- XPF/Hef/ERCC4

Phylogenetic Tree

Phylogenetic Tree

ARCPHdb is a specialized protein database for archaea helicase sequences from the SF1 and SF2 superfamily.

Helicases are proteins that use the energy of nucleotides to unwind DNA/RNA/protein complexes participating in almost every aspect of DNA and RNA metabolism. Helicases have been classified into six superfamilies (SF1-6) according to sequence, enzyme activity and oligomeric state. SF1s and SF2 are the non hexameric helicases and are the most abundant helicases. These helicases have been studied in many organisms and most importantly archaea.

Archaea are important model microorganisms that thrive in extreme environments of salt, temperature, pressure and pH. Archaea are prokaryotic microorganisms regarding their shape and morphology, they have the particularity to share with eukaryotes their cellular processes. Archaeal proteins are easier to manipulate than their eukaryotic counterparts and have been used as model proteins to study the structure and function relationship. Using a Genome-wide approach, our group has curated SF1 and SF2 helicases proteins from 95 archaea sequenced genomes belonging to four taxa Euryarchaeota, Crenarchaeota, Thaumarchaeota, and Korarchaeota (Chamieh et al., Gene, 2016). SF1 and SF2 families were screened for their presence in archaea genomes. Retrieved helicase sequences were reclassified after phylogenetic and domain architecture analysis. Sequences were then attributed to each family and stored into a database accordingly.

ARCPHdb allows scientists to browse helicase sequence database for SF1 and SF2 families. Users can perform a search by Phyla, groups, organisms, and families. The database includes in addition links to Uniprot, Ref Seq ID, Protein Data Bank (PDB) and related publications from PUBMED.

Azm Center for Research in Biotechnology and its applications
Bioinformatics Lab & Applied Biotechnology Lab
http://esbst.ul.edu.lb/

Developed By: Mirna Moukhtar
Release Date: 19/12/2015

Fig. 2. Screenshot of the main website homepage. At the left-side, menu allows to choose for information on SF1 and SF2 families.

Download English Version:

<https://daneshyari.com/en/article/4965022>

Download Persian Version:

<https://daneshyari.com/article/4965022>

[Daneshyari.com](https://daneshyari.com)