EVOLUTIONARY ANALYSIS OF ARCHITECTURE AND EXTRACELLULAR DOMAIN SEQUENCE OF SID-1 FAMILY PROTEINS

SAPOZHNIKOVA V.O., BARANNIK T.V.

V.N. Karazin Kharkiv National University, Kharkiv, Ukraine
e-mail: valeria93@ukr.net

Introduction

Systemic RNA interference-defective-1 (SID-1) family includes transmembrane dsRNA channels, that passively transport dsRNA and do not require ATP. Proteins play role in systemic RNA interference and are widely spread among animals [1]. In human two paralogues are present: SIDT1 playing role in systemic RNA interference and SIDT2 transporting dsRNA into lysosomes.

Bioinformatics and biochemical data predict SIDT1 to contain a large (300 amino acids) extracellular domain (ECD) and 11 transmembrane domains (by UniProt), but the exact mechanisms of RNA transport are unknown [2]. ECD could play role in substrate recognition and transport regulation through Cys-Pro motifs having been predicted as potential heme-binding sites in human SIDT1 (our previous investigation). So the aim of this research was to analyze domain structure and evolutionary conservation of ECD in SID-1 family in *Metazoa*.

Materials and methods

The search of homologous protein sequences was performed using BLASTP algorithm against non-redundant database with SIDT1 protein of *Homo sapiens* as query (NP_060169). Only sequences with a threshold value lower than 10^{-10} were chosen for the analysis. Domain architecture was analysed using Pfam server (http://pfam.sanger.ac.uk). Multiple alignment was constructed in MEGA 6.0 tool using ClustalW algorithm. Phylogenetic tree was built in MEGA 6.0 tool using Maximum Likelihood method with 1000 bootstrap replicates.

Results and discussion

Sequence analysis showed SIDT1 homologues to exist in most *Metazoa, Choanoflagellata* and *Amoebozoa* taxa sequenced for today (Fig.1).

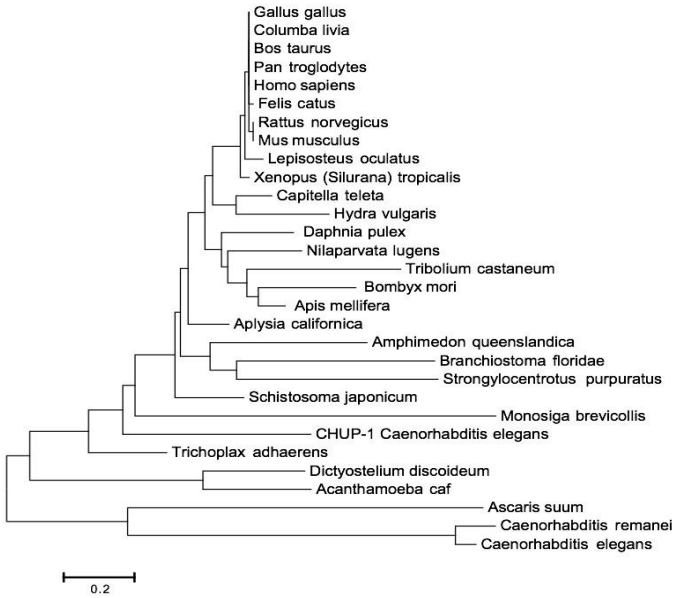

Fig.1. Phylogenetic tree of SID-homologues present in animals.

SID-1 gene possibly emerged in common ancestor of these groups might be lost later in fungi and some invertebrates (Table 1).

Group	Number of SID domains	PC/CPV motifs
Choanoflagellata	1-2	-/-
Amoebozoa	1	-/-
Placozoa	1	-/-
Spongia	1-5	-/-
Cnidaria	1	C/-
Platyhelminthes	1-3	-/-, C/CPV
Nematoda	1	-/-,C/CPV
Crustacea	1-2	-/-, C/-
Insecta	1-2	-/CPV, PC/CPV
Mollusca	1-2	-/-, C/CPV
Echinodermata	1-2	-/-
Cephalochordata	1-2	C/CPV
Vertebrata	1	C/CPV

Table 1. Variants of structure of proteins and motifs in different groups. (C- Cys, PC – ProCys, CPV – CysProVal)

For example, SID-1 homologue is absent in *Drosophila melanogaster*. The analysis of phylogenetic trees also revealed the presence of one to six SID-1 paralogues with different domain architecture in one specie. Thus the evolution of this gene was accompanied by several gene duplications and losses.

The sequence analysis showed domain architecture of this protein to be very diverse in invertebrates with 1 to 5 dsRNA-gated channel SID-1 domains. Some of homologues contain other functional domains involved in reception or in signal transduction, such as scavenger receptor cysteine rich domain in *Strongylocentrotus purpuratus* (*Echinodermata*), protein kinase domain in *Danaus plexippus* (*Insecta*) and WD40 domain in *Crassostrea gigas* (*Mollusca,* Fig.2, by Pfam). Several SID-1 homologues with different domain architecture may be present in one organism as it is shown in *Amphimedon queenslandica* (*Spongia*) having four different SID-1 homologues with 1, 2, 3 and 5 SID-1 domains.

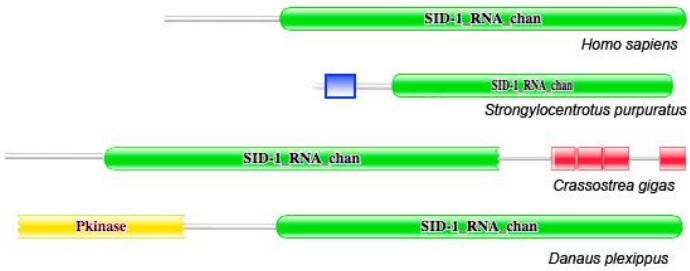

Fig.2. Domain architecture of SID-1 homologues in animals.

Comparing with SID-1 domain, ECD is poorly conserved in invertebrates and varies greatly by length from absence (*Amoebozoa*, *Placozoa*) to more than 400 amino acids (*Mollusca*). In vertebrates these proteins share one type of domain architecture with one SID-1 domain and ECD of about 300 amino acid residues (Fig.2, upper model).

Cys-Pro-Val (CPV) motif is considered to be used for reversible heme binding and regulation of protein activity [3]. Such motif has evolved in *C.elegans* in CHUP-1 protein but is absent in SID-1 protein involved in capture of environmental RNA. Starting from nematode this motif is present in SID-1 proteins of all invertebrates taxa. Vertebrates have ECD with absolutely conservative CPV motifs and new vertebrate-specific PC motif.

CHUP-1 *Caenorhabditis elegans*	--VAVHLDSNSTICMTVSVQKIGCPVFDLP
SID-1 *Caenorhabditis elegans*	SQFHVTLYSEDDICANLITVPANESIYDRS
Aplysia californica	ETVLLKVTSPTETCMTVSVQTVKCPVFDLD
Bombyx mori	ENVIFMIESDDELCAVVSIQNFSCPVFDNE
Branchiostoma floridae	TVVVRATSQNNEKCSVLSLQRAKCPVYDLD
Homo sapiens	SVIIKVVSEMAYPCSVVSVQNIMCPVYDLD
Rattus norvegicus	SVIIKVVSEKAYPCSVVSVQNIMCPVYDLD
Bos taurus	SVIIKVVSTLAYPCSVVSVQNIMCPVYDLD
Gallus gallus	SVIIKVVSDAVYPCSVVSVQDIVCPVYDLD
Xenopus (Silurana) tropicalis	SVIIKVKSPENYPCSVVSVQDISCPVYDLD

Cysteines in both motifs are predicted not to form disulfide bridges. The ability of human SIDT1 CP/PC motifs to bind heme molecule was proved by bioinformatics tools in our previous study. Taking into account the wide involvement of cysteine residues in redox regulation and metal or heme binding we can suggest the regulatory role of ECD with conservative CP/PC motifs in the functioning of vertebrate SID homologues.

Conclusion

So in spite of various types of SID-1 architecture in invertebrates that have SID homologues, the single type of SID structure with ECD was inherited in common ancestor of vertebrates. The conservation of SID structural features including cysteine-containing patterns PC/CP that are not involved in disulfide bonds formation may indicate their importance for redox regulation of systemic RNA interference and dsRNA transport in vertebrates especially under oxidative agents action and hemolytic conditions.

References

1. Feinberg E.H. et al./ Science.– 2003.– V. 301.– P.1545-1547.
2. Elhassan M. O. et al./ J.Biol.Chem.– 2012.– V. 287.– P. 5267–5277.
3. Zhang L, Guarente L./ EMBO J.– 1995.–V.14.– P.313-320.