UDC 577.112.7:593.4
ISSN 1330-9862 *scientific note*

(FTB-1152)

Signal Recognition Particle 54 kD Protein (SRP54) from the Marine Sponge *Geodia cydonium*

Sonja Durajlija-Žinić[1], *Helena Ćetković*[1], *Werner E. G. Müller*[2] *and Vera Gamulin*[1]*

[1]Department of Molecular Genetics, Ruđer Bošković Institute, Bijenička 54, HR-1000 Zagreb, Croatia

[2]Institut für Physiologische Chemie, Abteilung Angewandte Molekularbiologie, Johannes Gutenberg-Universität Mainz, Duesbergweg 6, D-55099 Mainz, Germany

Received: May 10, 2002
Accepted: June 18, 2002

Summary

In the systematic search for phylogenetically conserved proteins in the simplest and most ancient extant metazoan phylum – Porifera, we have identified and analyzed a cDNA encoding the signal recognition particle 54 kD protein (SRP54) from the marine sponge *Geodia cydonium* (Demospongiae). The signal recognition particle (SRP) is a universally conserved ribonucleoprotein complex of a very ancient origin, comprising SRP RNA and several proteins (six in mammals). The nucleotide sequence of the sponge cDNA predicts a protein of 499 amino acid residues with a calculated M_r of 55175. *G. cydonium* SRP54 displays unusually high overall similarity (90 %) with human/mammalian SRP54 proteins, higher than with *Drosophila melanogaster* (88 %), or *Caenorhabditis elegans* (82 %). The same was found for the majority of known and phylogenetically conserved proteins from sponges, indicating that the molecular evolutionary rates in protein coding genes in Porifera as well as in highly developed mammals (vertebrates) are slower, when compared with the rates in homologous genes from invertebrates (insects, nematodes). Therefore, genes/proteins from sponges might be the best candidates for the reconstruction of ancient structures of proteins and genome/proteome complexity in the ancestral organism, common to all multicellular animals.

Key words: Porifera, Metazoa, molecular evolution, common ancestor, signal recognition particle, SRP54

Introduction

Signal recognition particle (SRP) is a ubiquitous ribonucleoprotein complex that assists in the co-translational translocation of specific proteins to cellular membranes *(1)*. When ribosomes translate proteins destined for either secretion or integration into the plasma membrane, they are directed by SRP to translocational pores in the endoplasmatic reticulum in eukaryotes, or to the plasma membrane in prokaryotes. The SRP-dependent mechanism is highly conserved from prokaryotes to mammalian cells, SRP-related particles have been identified in all three domains of living organisms (Eukarya, Archea and Bacteria) and SRPs may function in protein translocation in every living organism *(2)*. The most studied mammalian SRP is an 11S cytoplasmic ribonucleoprotein particle that consists of a 300-nucleotide SRP RNA (7SL RNA) complexed with six proteins: SRPs 72, 68, 54, 19, 14 and 9 (*3*). The prominent role in protein translocation is designated to a 54 kD protein, SRP54 *(4)*,

* Corresponding author; Phone: ++385-1-4561-115; Fax: ++385-1-4561-177; E-mail: gamulin@rudjer.irb.hr
** This paper was presented at the 1st Croatian Congress on Molecular Life Sciences (CCOMLIS) in Opatija, 9–13 June, 2002

which binds the signal sequence of the nascent polypeptide as it emerges from a ribosome. Signal sequences of the secretory and membrane proteins possess an amino acid stretch of 7–20 hydrophobic residues (preferentially leucines or alanines) usually flanked by charged residues (*4*). SRP54 is a basic protein and two-domain GTPase (*4*). The G-domain contains all the consensus elements of a GTPase. One conspicuous difference is found in the GTP binding motif G-4, which reads TKXD and deviates from the consensus NKXD (*5*). The C-terminal M-domain is rich in methionines and positively charged residues. The M-domain is generally less conserved between different SRP homologues and also varies in length. M-domain has the potential to form amphipathic helices that build the signal sequence binding pocket (*4*). It also contains a conserved and positively charged sequence motif that participates in the binding of SRP54 to the SRP RNA (*4*).

Sponges (Porifera) are an excellent model organism for molecular evolutionary studies. They represent the lowest metazoan phylum that existed prior to the Cambrian explosion *(6)*, at least 580 million years ago *(7)* and can therefore be considered as living fossils *(8)*. Here, we report the characterization of cDNA coding for the signal recognition particle 54 kD protein (SRP54) from the marine sponge *Geodia cydonium*.

Materials and Methods

Isolation and characterization of SRP54 cDNA

Specimens of *G. cydonium* (Porifera, Demospongiae, Tetractinomorpha, Astrophorida, Geodiidae) were collected in the Northern Adriatic Sea near Rovinj, Croatia. The preparation of *G. cydonium* cDNA library in Lambda ZAP Express™ vector (Stratagene, La Jolla, USA) was already described (*9*). During the screening of the sponge library for lambda phages encoding Ras-related small GTPases (*10*), a recombinant phage carrying a cDNA for SRP54 was also identified as weak positive. Phagemid pBK-CMV with SRP54 cDNA insert was excised *in vivo* from lambda vector using 704 helper virus and *E. coli* XLORL cells as already described *(9)*. The nucleotide sequence of the cDNA insert was determined on the automated DNA sequencer ALF Express (Pharmacia, Uppsala, Sweden) using ThermoSequenase Cy5 Terminator Kit (Amersham-Pharmacia, Uppsala, Sweden). The universal primers for sequencing from the insert ends as well as two additional oligonucleotides complementary to the deduced sequences were used to obtain the complete sequence of SRP54 cDNA. All sequences were obtained at least twice.

Sequence analysis

Homology searches and sequence retrieval were performed via the Internet server at the National Center for Biotechnology Information, National Institute of Health, Bethesda, MD, USA (http://www.ncbi.nlm.nih.gov). BLASTP and BLASTX programs were used. BLASTP searches for homologies between protein sequences, while BLASTX compares the six-frame conceptual translation products of a nucleotide query sequence (both strands) against a protein sequence database *(11)*. Nucleotide and protein sequences were stored and analyzed using PC/GENE 14.0 programs from IntelliGenetics (Mountain View, CA, USA). CLUSTAL X *(12)* was used for multiple sequence alignments (MSA) and construction of the phylogenetic tree from the MSA. The degree of support for internal branches was further assessed by bootstrap analysis. Programs GeneDoc *(13)* and Tree view, version 1.6.6. *(14)* were used for graphic presentation of the results.

Results and Discussion

Signal recognition particle 54 kD protein (SRP54) from G. cydonium

The analyzed sponge cDNA is 1586 nucleotides (nt) long, excluding the polyA tail; the sequence has been deposited in the GenBank under accession number AY100457. Homology searches, performed by the BLASTX program, identified the protein encoded by this cDNA from *G. cydonium* as signal recognition particle 54 kD protein (SRP54). The protein was therefore named GCSRP54. The cDNA contains 29 untranslated nt at the 5'-end, followed by the open reading frame (nt 30–1529) that codes for a SRP54 protein 499 amino acids long with a calculated M_r of 55175. 57 unique nt are located between the stop codon and the poly(A) tail. Like in the other cDNAs from *G. cydonium*, the typical signal for polyadenylation, AATAAA *(15)*, is not present at the 3'-end and it is unclear which sequence corresponds to that signal in sponges.

The amino acid sequence of SRP54 from *G. cydonium* is shown in Fig. 1. Four GTP binding regions in the G-domain, the SRP54 specific signature and the SRP RNA binding motif in M-domain are indicated. GCSRP54 was compared with all SRP54 proteins found in protein databases using BLASTP. The highest homology was found with SRP54 from mammals (human, dog and mouse). Unfortunately, from the animals belonging to lower metazoan taxa, only SRP54 from *Drosophila melanogaster* (AAD46831) and *Ceaenorhabditis elegans* (Z54271) are known and they are less similar to sponge SRP54 (Table 1). Multiple sequence alignment of four metazoan SRP54s, produced by CLUSTAL X, is shown in Fig. 2. SRP54 proteins are two-domain GTPases and in comparison with human/mammalian SRP54s, the first 294 aa in sponge SRP54 form the G-domain. This domain shows 79 % identity (92 % overall similarity) with hu-

Table 1. The percentage of identity and overall similarity (in parenthesis) between the four metazoan SRP54 proteins aligned in Fig. 2. For abbreviations see Fig. 2.

	HUMAN	GEOCY	DROME	CAEEL
HUMAN	100 %	77 % (90 %)	75 % (89 %)	69 % (84 %)
GEOCY	77 % (90 %)	100 %	74 % (88 %)	68 % (82 %)
DROME	75 % 89 %	74 % (88 %)	100 %	66 % (82 %)
CAEEL	69 % (84 %)	68 % (82 %)	66 % (82 %)	100 %

```
MVLADLGRKITTALLPRQRTVINEEVLQAMLKEICTALLEADVNVKLVGKLRQNVRAAID  60
FEDMGAGLSKRRIIQTSVFNELCKLLDPGVPVWHPTKGHSNVIMFVGLQGSGKTTTCTKL 120
AYHYQKKGWKTCLVCADTFRAGAFDQLKQNATKARVPFYGSYTEMDPVVIAQEGVEKFKE 180
DSFEVIIVDTSGRHKQEESLFEEMLQVSQAIDPDNIIFVMDGTIGQACESQARAFKEKVD 240
VASVIVTKLDGHAKGGGALSAVAATRSPIIFIGTGEHIDEMEPFKTKPFVSKLLGMGDLE 300
GLMEKVSDLKLDENEELMDKLKHGQFTLRDMYEQFQNIMKMGPFNQIIGMIPGFSPDFMS 360
KGNERESMAKLKRLMTMMDSMNDGELDHPNGAKLFSKQPGRAARVARGSGTSVREVNELL 420
KQYSNFSATVKKMGGIKGLFKGGDLGKNVNPSQMAKLNQQMAKMMDPRVLQQMGGMSGLQ 480
NMMRQFQQGASNMPGFKGK 499
```

Fig. 1. The amino acid sequence of *G. cydonium* SRP54 protein is shown in one letter code. Four GTP binding regions in the G-domain as well as the SRP RNA binding sequence in M-domain are underlined. The SRP54 specific signature is boxed while the G-domain is shown in bold and the M-domain in regular letters.

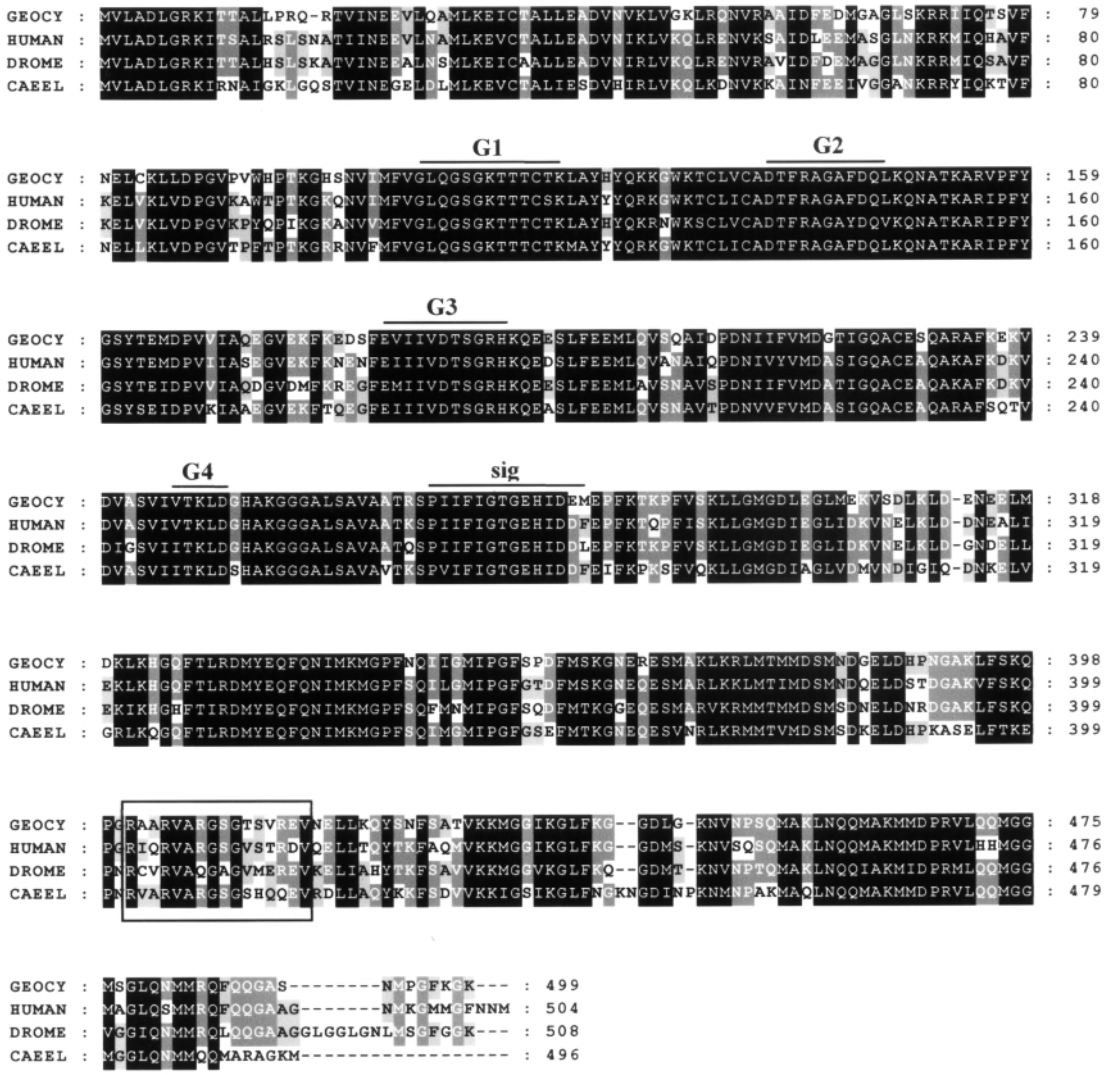

Fig. 2. Multiple alignment of four SRP54 aa sequences from multicellular organisms: *Geodia cydonium* (GEOCY, AY100457), *Homo sapiens* (HUMAN, U51920), *Drosophila melanogaster* (DROME, AAD46831) and *Caenorhabditis elegans* (CAEEL, Z54271). G1, G2, G3, G4: regions involved in GTP/GDP binding; sig: SRP54 specific signature. RNA binding region is boxed. 100 % conserved aa (identical + similar) are shown in white on black and 75 % conserved aa in white on gray.

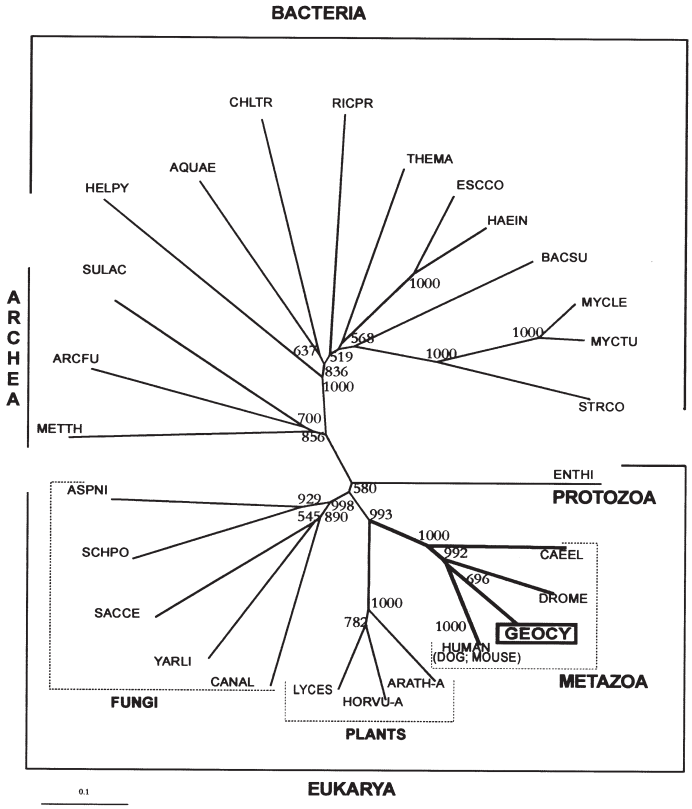

Fig. 3. The unrooted phylogenetic tree of selected SRP54 proteins from all three domains of life. SRP54 sequences from BACTERIA: *Aquifex aeolicus* (AQUAE, 32984037), *Thermotoga maritima* (THEMA, AAD36632), *Chlamydia trachomatis* (CHLTR, 3328415), *Rickettsia prowazekii* (RICPR, AJ235269), *Escherichia coli* (ESCCO, P07019), *Haemophilus influenzae* (HAEIN, P44518), *Heliobacter pylori* (HELPY, P56005), *Mycobacterium leprae* (MYCLE, Z97369), *Mycobacterium tuberculosis* (MYCTU, Q10963), *Bacillus subtilis* (BACSU, D14356) and *Streptomyces coelicolor* (STRCO, 3191979). SRP54 from ARCHEA: *Sulfolobus acidocaldarius* (SULAC, Y12702), *Archaeoglobus fulgidus* (ARCFU, 3334342) and *Methanobacterium thermoautotrophicum* (METTH, 3334341). SRP54 from EUKARYA: Protozoa – *Entamoeba histolytica* (ENTHI, Y12515); Plants – *Hordeum vulgare* (HORVU-A, P49968), *Arabidopsis thaliana* (ARATH-A, L19997) and *Lycopersicon esculentum* (LYCES, P49971); Fungi – *Yarrowia lipolytica* (YARLI, Q99150), *Schizosaccharomyces pombe* (SCHPO, P21565), *Candida albicans* (CANAL, 3334344), *Saccharomyces cerevisiae* (SACCE, P20424) and *Aspergillus niger* (ASPNI, 2119054); Metazoa – *Drosophila melanogaster* (DROME, AAD46831), *Caenorhabditis elegans* (CAEEL, Z54271), *Mus musculus* (MUSMU, P14576), *Canis species* (CANSP, P13624) and *Homo sapiens* (HUMAN, U51920). The numbers at the nodes refer to the level of confidence as determined by bootstrap analysis. Scale bar indicates an evolutionary distance of 0.1 aa substitutions per position in the sequence.

man (dog, mouse) SRP54 and only one amino acid deletion was found at the N-terminus (Fig. 2). Four GTP binding regions are perfectly conserved in all known metazoan proteins (Fig. 2). M-domain (aa 295–499) in GCSRP54 contains 23 methionines (11.2 %) and is also highly conserved (75 % identity or 88 % similarity with human M-domain). The majority of changes are located close to, or at the C-terminus of the protein. Interestingly, the amino acid sequence motif that mediates the binding of SRP54 to the SRP RNA (Fig. 1) is not fully conserved in metazoan SRP54s (see Fig. 2). Similarity of GCSRP54 with SRP54 from organisms belonging to other major kingdoms of life is considerably lower, for example GCSRP54 shows only 29 % identity (47 % overall similarity) with the *E. coli* homologue of SRP54, 48 kD protein Ffh (P07019).

The unrooted phylogenetic tree of selected SRP54 proteins from organisms representing all three domains (five kingdoms) of life is shown in Fig. 3. Clustering of SRP54s into major groups, according to their origin, is clearly seen and is supported by high bootstrap values. Metazoan proteins form one cluster, with sponge SRP54 lying closest to the human homologue. SRP54s from dog and mouse are (almost) identical to human protein and were therefore excluded from the picture. Interestingly, SRP54 proteins from multicellular animals are more closely related to plants' SRP54s than to SRP54s from fungi. This is usually not the case for the majority of phylogenetic trees based on the protein sequences, because plants branched off earlier in the evolution, before the split of Metazoa and Fungi. However, contrary to the SRP54s in plants and metazoan animals, SRP54s from fungi have extra long C-terminuses (M-domains) and this major difference influences their position in the phylogenetic tree of SPR54s.

High homology between sponge and mammalian homologous proteins was often noticed during our study of ancient proteins from sponges and furthermore, some sponge/human orthologues are even not encoded in the genome of *C. elegans* (16), although sponges (Porifera) branched off first from the common ancestor of all Metazoa more than 580 million years ago

(7). Sponges, as well as human/mammals (vertebrates), show on average slower molecular evolutionary rates in protein coding sequences than *D. melanogaster* and especially *C. elegans* (invertebrates) (*16*). Unfortunatelly, apart from mammals, nematodes and insects, very limited information about protein/gene sequences is currently available from animals belonging to other metazoan taxa. Consequently, when compared with mammalian, *D. melanogaster* and *C. elegans* proteins, sponge and mammalian homologues/orthologues often remain the most similar protein pairs that best reflect the structure of ancient proteins (genes) in the ancestral progenitor, common to all Metazoa.

Acknowledgements

This work was supported by grants from the Croatian Ministry of Science and Technology (P00981003 to V. G.) and the Deutsche Forschungsgemeinschaft (Mü 348/12–5 to W.E.G.M.).

References

1. R. J. Keenan, D. M. Freymann, R. M. Stroud, P. Walter, *Annu. Rev. Biochem. 70* (2001) 755–775.
2. C. Zwieb, T. Samuelsson, *Nucleic Acids Res. 28* (2000) 171–172.
3. K. Wild, O. Weichenrieder, K. Strub, I. Sinning, S. Cusack, *Curr. Opin. Struct. Biol. 12* (2002) 72–81.
4. H. Lütcke, *Eur. J. Biochem. 228* (1995) 531–550.
5. H. R. Bourne, D. A. Sanders, F. McCormick, *Nature, 349* (1991) 117–127.
6. A. H. Knoll, *Proc. Natl. Acad. Sci. 91* (1994) 6743–6750.
7. C. W. Li, J. Y. Chen, T. E. Hua, *Science, 279* (1998) 879–882.
8. W. E. G. Müller, *Naturwissenschaften, 82* (1998) 11–25.
9. K. Pfeifer, M. Haasemann, V. Gamulin, H. Bretting, F. Fahrenholz, W. E. G. Müller, *Glycobiology, 3* (1993) 179–184.
10. V. Gamulin, A. Krasko, A. Mikoč, S. Durajlija–Žinić, W. E. G. Müller, *Period. Biol. 103* (2001) 229–234.
11. S. F. Altschul, T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, *Nucleic Acids Res. 25* (1997) 3389–3402.
12. J. D. Thompson, D. G. Higgins, T. J. Gibson, *Nucleic Acids Res. 22* (1994) 4673–4680.
13. K. B. Nicholas, H. B. Jr. Nicholas, D. W. Deerfield, *Embnew. News, 4* (1997) 14–14.
14. D. M. Roderic, *Computer Applications in the Biosciences, 12* (1996) 357–358.
15. D. Zarkower, P. Stephenson, M. Sheets, M. Wickens, *Mol. Cell Biol. 6* (1986) 2317–2323.
16. V. Gamulin, I. M. Müller, W. E. G. Müller, *Biol. J. Linnean Soc. 71* (2000) 821–828.

Protein SRP54 iz morske spužve *Geodia cydonium*

Sažetak

U sistematskoj potrazi za filogenetski sačuvanim proteinima u spužava, najjednostavnijih i najstarijih živućih Metazoa, autori su identificirali i analizirali cDNA koja kodira protein SRP54 u morske spužve *Geodia cydonium.* SRP54 je evolucijski najsačuvaniji protein ribonukleoproteinskoga kompleksa SRP (signal recognition particle), odgovornog za translokaciju sekretornih i transmembranskih proteina. SRP je nastao vrlo rano u evoluciji i svugdje je prisutan u živom svijetu, od bakterija do čovjeka. cDNA spužve *G. cydonium* kodira protein SRP54 dug 499 aminokiselina, izračunate molekularne mase 55175, koji pokazuje najviši stupanj sličnosti (90 %) s ljudskim proteinom SRP54, više nego sa SRP54 iz kukca *Drosophila melanogaster* (88 %) ili oblića *Caenorhabditis elegans* (82 %). Velika sličnost proteina spužava s homolozima u sisavaca uočena u ovom radu, kao i u našim prijašnjim istraživanjima, upućuje na sporije evolucijske promjene u genima spužava i sisavaca (kralješnjaka) u usporedbi s ubrzanijim promjenama u genima kukaca i oblića (beskralješnjaka). Geni/proteini spužava su stoga posebno korisni za rekonstrukciju strukture proteina i kompleksnosti genoma/proteoma u ancestralnom organizmu koji je bio zajednički predak svih višestaničnih životinja.