Aspects of Poriferan (*Suberites domuncula*) Apoptosis and Innate Immune System and Evolutionary Implications

Dissertation zur Erlangung des Grades
Doktor der Naturwissenschaften

Am Fachbereich Biologie
der Johannes Gutenberg-Universität in Mainz

Bérengère Luthringer

Geb. am 30 März 1979 in Plœmeur (Frankreich)

Mainz, 2009

Dekan:

1. Berichterstatter:

2. Berichterstatter:

Tag der mündlichen Prüfung: 23/04/2009

TABLE OF CONTENTS

I. ABSTRACT

Survivin, a unique member of the family of inhibitors of apoptosis (IAP) proteins, orchestrates intracellular pathways during cell division and apoptosis. Its central regulatory function in vertebrate molecular pathways as mitotic regulator and inhibitor of apoptotic cell death has major implications for tumor cell proliferation and viability, and has inspired several approaches that target survivin for cancer therapy. Analyses in early-branching Metazoa so far propose an exclusive role of survivin as a chromosomal passenger protein, whereas only later during evolution the second, complementary antiapoptotic function might have arisen, concurrent with increased organismal complexity. To lift the veil on the ancestral function(s) of this key regulatory molecule, a survivin homologue of the phylogenetically oldest extant metazoan taxon (phylum Porifera) was identified and functionally characterized. SURVL of the demosponge *Suberites domuncula* shares significant similarities with its metazoan homologues, ranging from conserved exon/intron structures to the presence of localization signal and protein-interaction domains, characteristic of IAP proteins. Whereas sponge tissue displayed a very low steady-state level, SURVL expression was significantly up-regulated in rapidly proliferating primmorph cells. In addition, challenge of sponge tissue and primmorphs with cadmium and the lipopeptide Pam3Cys-Ser-(Lys)4 stimulated *SURVL* expression, concurrent with the expression of newly discovered poriferan caspases (*CASL* and *CASL2*). Complementary functional analyses in transfected HEK-293 revealed that heterologous expression of poriferan survivin in human cells not only promotes cell proliferation but also augments resistance to cadmium-induced cell death. Taken together, these results demonstrate both a deep evolutionary conserved and fundamental dual role of survivin, and an equally conserved central position of this key regulatory molecule in interconnected pathways of cell cycle and apoptosis.

Additionally, SDCASL, SDCASL2, and SDTILRc (TIR-LRR containing protein) may represent new components of the innate defense sentinel in sponges. SDCASL and SDCASL2 are two new caspase-homolog proteins with a singular structure. In addition to

their CASc domains, SDCASL and SDCASL2 feature a small prodomain NH_2-terminal (effector caspases) and a remarkably long COOH-terminal domain containing one or several functional double stranded RNA binding domains (dsrm). This new caspase prototype can characterize a caspase specialization coupling pathogen sensing and apoptosis, and could represent a very efficient defense mechanism. SDTILRc encompasses also a unique combination of domains: several leucine rich repeats (LRR) and a Toll/IL-1 receptor (TIR) domain. This unusual domain association may correspond to a new family of intracellular sensing protein, forming a subclass of pattern recognition receptors (PRR).

II. INTRODUCTION

A. Innate immunity

Organisms are always subjected to foreign, potentially harmful microbes. Their survival depends on effective protection system(s). Two kinds of defense strategies can be distinguished based on the type of receptors involved in the pathogen recognition. The first mechanism is mediated by the antigen receptors (B-cell and T-cell receptors). To face up to the wide range of antigens, the collection of receptors has to evolve constantly via somatic recombinations of their genes (reviewed in Schatz *et al.* 1992), Due to these features, this defense type is named adaptive immunity. The second defense mechanism is mediated by pattern recognition receptors or PRR. In opposition to adaptive immunity, this defense is based on non-genetically evolving receptors and is termed innate immunity. PRR sense conserved and invariant features of pathogens or pathogens associated patterns (PAMP; reviewed in Janeway 1989) but also danger signals released by damaged, activated (by inflammatory molecules), and dying cells (alarmins, reviewed in Oppenheim *et al.* 2007). PAMP include components of bacterial and fungal cell walls, flagellar proteins, and viral nucleic acids.

PRR embrace, among others, the transmembrane Toll-like receptors (TLR; located within the plasma membrane and the endosome), the C-type lectin receptors (CLR), the RIG-I-like helicase (RLH), and the cytosolic Nod-like receptors (NLR).

The TLR family is the first and most studied group. TLR are classical transmembrane protein comprising extracellular leucine-rich repeats (LRR, pathogen sensing) and intracellular Toll/IL (interleukin)-1 receptor (TIR) domain, initiating the signal transduction pathway. This cascade relies on two signal transducers, the myeloid differentiation primary response gene 88 (MyD88) and the TIR-domain-containing adapter-inducing interferon-β (TRIF), and culminates in the activation of nuclear factor-kappa B (NF-kB) and the initiation of innate inflammatory immune responses (reviewed in Kawai and Akira 2007).

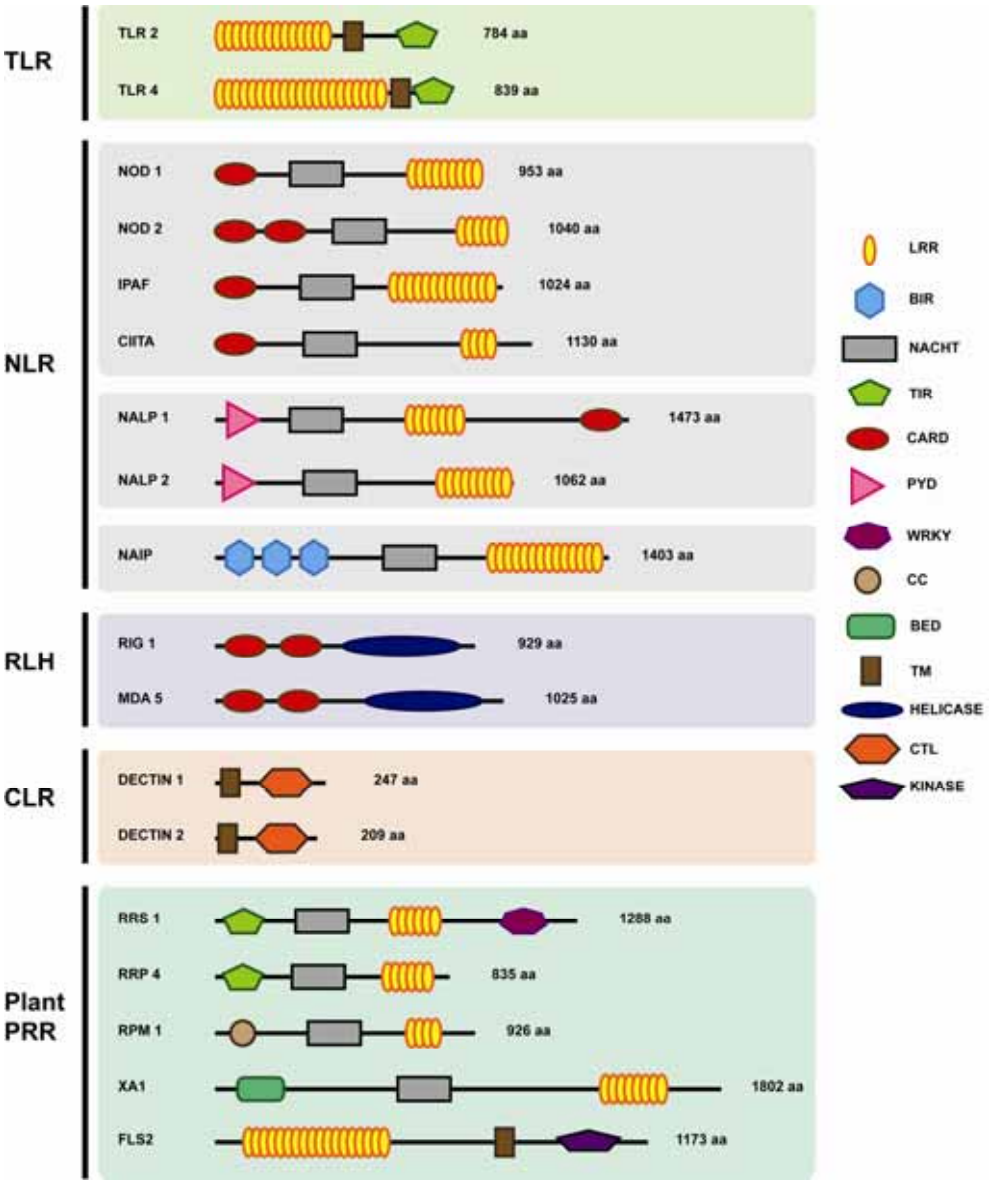


Figure 1: Schematic representation of innate immune sensors or pattern recognition receptors (PRR). TLR are classical transmembrane proteins composed of variable number of LRR (leucine rich repeat), a transmembrane region (TR), and a TIR (Toll/IL-1 receptor) domain. Similarly, the C-type lectin receptors (CLR) encompass a transmembrane region and a C-type lectin domain (CTL). However, the NLR, intracellular sensors, are composed of variable amino-terminal effector-binding domains (CARD, caspase recruitment domain, PYD, pyrin domain, or BIR, baculovirus IAP repeat) a centrally located nucleotide-binding oligomerization domain (NACHT), and a carboxy-terminal ligand-recognition region or LRR. RIG-1 helicases (RLH) contain a helicase and two CARD domains. Correspondingly, PRR molecules from plants are classified in two groups: transmembrane PRR and resistance (R) or nucleotide binding domain and leucine-rich repeat (NB-LRR) proteins. Plant immune defense is primarily mediated by plant-disease-resistance (R) proteins. R proteins have either a TIR, a coiled-coil (CC), or a BED (zinc finger BED-type) domain, followed by a NACHT, and NH_2-terminal LRR.

NLR, also known as NOD-LRR, NACHT-LRR, and CATERPILLER (CARD, transcription enhancer, R(purine)-binding, pyrin, lots of leucine repeats), is a subgroup of PRR of increasing interest (reviewed in Kufer *et al.* 2005 and Meylan *et al.* 2006). NLR is a large family about 20 intracellular proteins with a common protein-domain organization but diverse functions. All members possess variable number of COOH-terminal leucine-rich repeats (LRR) involved in pathogen recognition and a central nucleotide-binding oligomerization domain (NACHT) mediating oligomerization. Finally, the NH_2-termini of NLR are characterized by various effector domains: CARD (caspase recruitment domain), PYD (pyrin), or BIR (baculovirus inhibitor of apoptosis protein repeat) domains (fig. 1). NLR comprise two large sub-classes: the NOD (CARD effector domain) and the NALP (pyrin effector domain) clan. With NAIP (BIR effector domain), more than 20 NLR have been identified so far. The diversity of the effector part allows them to link various downstream signalling pathways. The CARD-containing proteins mediate NF-kB (nuclear factor-kappa B) activation, whereas PYD molecules, such as NALP3 regulate IL-1beta and IL-18 production (reviewed in Rietdijk *et al.* 2008).

Similar molecules are found in plants and lower organisms. Plants lack the adaptive immunity, but two classes of native immune receptors can be distinguished (fig. 1). The first class encompasses the transmembrane PRR which can detect microbes-associated molecular patterns (MAMP; reviewed in Zipfel and Felix 2005). The second comprises the cytosolic resistance (R) or NB-LRR proteins and are related to NLR proteins (reviewed in Ting and Davis 2005). Plant NB-LRR proteins can be categorized into TIR and non-TIR classes. The class of non-TIR molecules is less well defined, but most proteins contain NH_2-terminal α-helical coiled-coil-like sequences (Pan *et al.* 2000). Similarly, the genome of *Strongylocentrotus purpuratus* (sea urchin) includes a vast repertoire of 222 Toll-like receptors, a superfamily of more than 200 proteins similar to NLR, and a large family of scavenger receptor cysteine-rich proteins (reviewed in Rast *et al.* 2006). However, even if TLR-like proteins are found in *Drosophila melanogaster* (mammalian TLR derive their name from the *Drosophila* Toll protein, a transmembrane receptor first identified as an essential component of fly dorsal-ventral embryonic development; Wu and Anderson 1997) and in *Caenorhabditis elegans*, no protein similar to NLR was identified.

NF-kB plays a key role in regulating the immune response to infection. In unstimulated cells, the NF-kB dimers are sequestered an inactive state in the cytoplasm by IkB (Inhibitor of kB; mask the nuclear localization signals NLS). Activation of the NF-kB is initiated by the signal-induced degradation of IkB proteins via activation of a kinase called the IkB kinase (IKK). NF-kB is a versatile transcription factor linked to cell survival and apoptosis: it mainly induces the expression of several genes able to inhibit apoptosis and to promote growth (*e. g.*, caspase-8-c-FLIP (FLICE inhibitory protein), TNFR-associated factor 1 (TRAF1), TRAF2, and survivin; reviewed in Karin and Lin 2002; Tracey *et al.* 2005). On the other hand, apoptosis may be the ultimate resort of infected cells to limit the dissemination of pathogen or of inflammation reaction that would otherwise cause immune pathology but is also implicated in a variety of pathological conditions (*e. g.*, auto-immunity and sepsis; reviewed in Naugler and Karin 2008 and in Uwe 2008).

Furthermore, apoptosis can be directly triggered by TLR and NLR. Depending on the TLR, several activating pathways exist (Aliprantis *et al.* 2000; Han *et al.* 2004). For example, when bacterial lipoproteins activate TLR2, the receptor recruits MyD88, which mediates both apoptosis and NF-kB activation. The MyD88 death domain (DD) can associate with the DD of FADD (FAS-associated death domain protein), which in turn will recruit caspase-8 via its death effector domain (DED), leading to the initiation of the proteolytic activation of the caspases (fig. 2 and 4). Similarly, the NLR family plays diverse roles in apoptosis (reviewed in Kaparakis *et al.* 2007). Nucleotide-binding oligomerization domain containing 1 (Nod1) recognizes intracellular bacteria, primarily through sensing glycopeptides. When activated, the Nod molecules are able to recruit receptor-interacting serine-threonine kinase 2 (RIP2) via their CARD domains. RIP2 is a key component controlling either cytokine release or NF-kB activation (via TGF-beta activated kinase 1, TAK1) or apoptosis (via caspase-8; da Silva Correia *et al.* 2007).

B. Apoptosis

1. Origin and history of apoptosis

In the fifties, developmental biologists already reported some observations of dying cells with different morphological features from necrosis (*e. g.*, no loss of membrane integrity and no inflammation response) and was referred to as "programmed cell death" (Glücksmann 1951). However, the apoptosis concept emerged in the sixties after the disparate histological observations of "shrinkage necrosis" (*e. g.*, Kerr JFR (1965) in ischaemic liver and Searle J (1970) in human skin carcinomas). Then the idea arised of a general homeostatic function, regulating the size of cell populations under normal and pathological conditions and the term apoptosis was suggested (reviewed in Kerr *et al.* 1972 and Kerr 2002). In the eighties, recurring observations of chromatin condensation observed during this mechanism (leading to regular fragmentation of DNA, ladder of oligonucleosomes and nucleosomes) became a hallmark of apoptosis with a specific biochemical origin. The groundbreaking discovery of the *C. elegans* cell death protein 9 (Ced-9) and its close sequence homology to human antiapoptotic B-cell leukaemia/lymphoma 2 protein (Bcl-2) for the first time demonstrated the molecular conservation of the apoptotic cell death program along a broad range of Metazoa, from nematodes to vertebrates (Hengartner *et al.* 1992; Hengartner and Horvitz 1994). Subsequently, this finding was corroborated by the discovery of the *C. elegans* protein Ced-3 (Yuan *et al.* 1993), coding for a homologue of the human interleukin-converting enzyme (IL1-β) that rapidly became the prototypic member for a new family of proteases, the caspases (reviewed in Thornberry and Lazebnik 1998; Cerretti *et al.* 1992). Since then, the fundamental role of apoptosis in various physiological processes has been increasingly settled.

Finally, the major role of apoptosis in countless physiological and pathological processes was recognized in 2002 when the Nobel Prize rewarded three researchers Brenner, Sulston, and Horvitz for their discoveries, and offered the possibility of extrapolation to higher organisms.

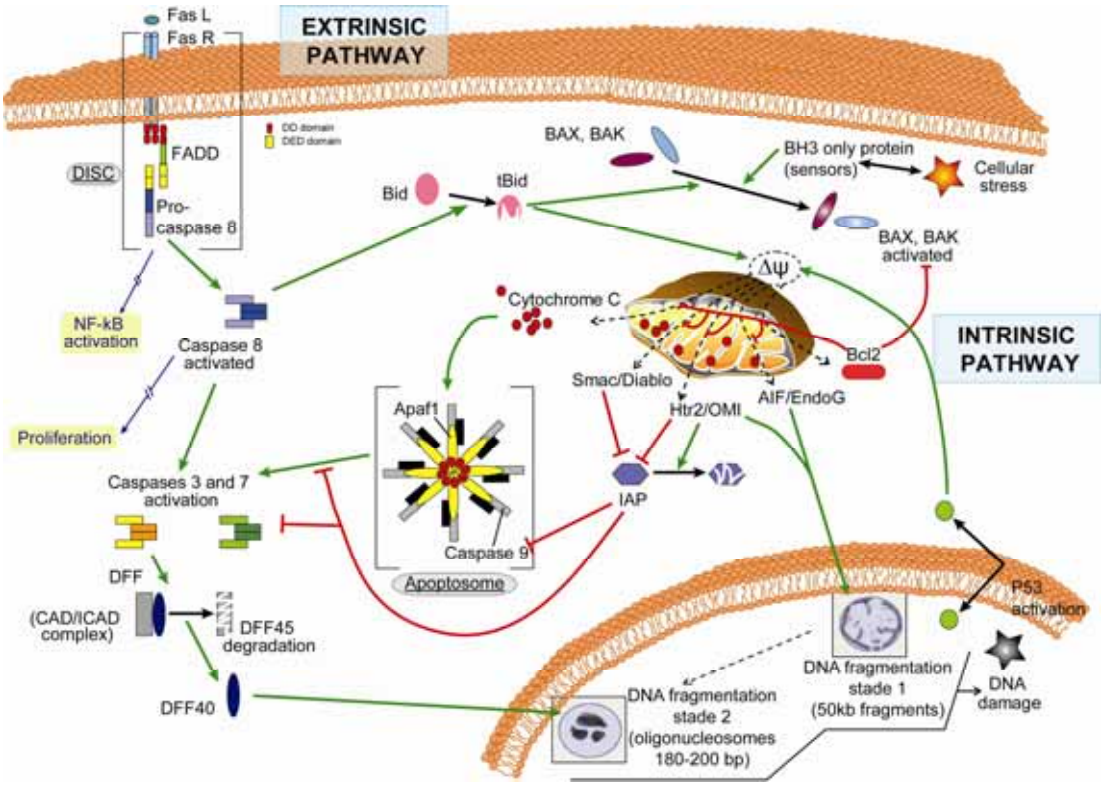


Figure 2: Apoptosis mediating pathways. The extrinsic pathway is mediated by tumor necrosis factor receptor (TNFR) also named death receptor (DR). Subsequently to adaptor proteins and initiator caspases recruitment (formation of a complex death-inducing signalling complex (DISC), effector caspases will be activated. In the intrinsic pathway, the procaspase-9 auto-activates after aggregating in a cytosolic multi-molecular complex or apoptosome. The apoptosome formation is subsequent to cytochrome C release from the mitochondria.

2. Triggering of apoptosis

Various extracellular (*e. g.*, toxins, hormones, growth factors, nitric oxide, and cytokines) or intracellular (*e. g.*, glucocorticoids, radiation, nutrient deprivation, viral infection, and hypoxia) signals or factors can set up apoptosis. Amplification of this integrated initial signal, through extrinsic or intrinsic pathway, will led to the execution phase mediated by specific cysteine protease or caspases, culminating in the death of the cell (fig. 2).

The extrinsic or cytoplasmic pathway involves transmembrane receptor-mediated interactions. These receptors are members of the tumor necrosis factor (TNF) receptor gene superfamily and exhibit a specific intracellular domain or death domain (DD) needed for the signal transduction. The intrinsic pathway is centred on the mitochondria and the

members of B-cell CLL/lymphoma 2 (Bcl-2) protein family. Stimuli will lead to mitochondrial membrane permeabilization (at the origin of the liberation of proapoptotic factors from the inter-membrane space to the cytosol) associated (in most cases) with a loss a mitochondrial membrane potential ($\Delta\psi_m$). For each pathway, signals are then coordinated thanks to multi-proteic complexes: the apoptosome or the death-inducing signalling complex (DISC; for intrinsic and extrinsic pathway respectively), both transducing downstream signal cascade resulting in the execution phase of apoptosis (caspases activation).

The core proteins involved in programmed cell death can be classified in 5 groups: the Bcl-2 family, the tumor necrosis factors super family and their receptors, the soluble inter membrane mitochondria protein (SIMP) family, the caspase family, and the inhibitor of apoptosis protein (IAP) family.

3. Core component of apoptosis

a. The Bcl-2 family

This family includes pro- (*e. g.*, Bcl-2-associated X protein (Bax), Bcl-2-antagonist/killer (Bak), Bcl-2 associated agonist of cell death (Bad), BH3 interacting domain death agonist (Bid), and Bcl-2-interacting mediator of cell death (Bim)) and antiapoptotic (*e. g.*, Bcl-2 and Bcl-xL) proteins (reviewed in Levine *et al.* 2008). Protein alignment revealed 4 highly conserved regions named Bcl-2 homology or BH (1 to 4) domains. Some proapoptotic molecules contain BH1, BH2, and BH3 domains (Bak), whereas some include only a BH3 domain (Bid, Bim, Bad). BH4 domain is specific to antiapoptotic proteins (Bcl-2, Bcl-xL, Bcl-w, Mcl-1, A1, and Diva). These proteins are able to form dimers, thus modulating the sensitivity to death signals. Furthermore, members of the Bcl-2 family may regulate outer mitochondrial membrane integrity, mediating the release of apoptogenic molecules (*e. g.*, cytochrome C and Smac/DIABLO), which leads to caspase activation (most of the members possess a COOH-terminal hydrophobic region allowing their anchorage to

external mitochondrial membrane and may modulate its permeability; García-Sáez *et al.* 2006).

b. The tumor necrosis factors (TNF) super family and their receptors

The superfamily of TNF proteins encompasses 19 ligands and 29 receptors (TNFR; fig. 3). These proteins regulate normal functions, such as immune response, haematopoiesis, and morphogenesis, but they have been also implicated in tumorigenesis, transplant rejection, septic shock, viral replication, bone resorption, rheumatoid arthritis, and diabetes (reviewed in Aggarwal 2003 and Gaur; Aggarwal 2003).

The receptors contain variable number of cysteine-rich domains (CRD) in their extracellular portion. CDR mediates stability of the trimetric/multimeric structures that the receptors form. However, only few TNFR contain intracellular death domain (DD). This motif is needed for the transmission of the "death signal". Seven members are forming the sub-group of "death receptors" (DR): Fas (CD95/APO-1; ligand: FasL), TNFR1 (p55/CD120a; ligands: TNFα and β), DR3 (ligand: VEGI), DR6, p75NGFR (nerve growth factor receptor, ligand: NGF) and the two TRAIL receptors: DR4 (TRAIL-R1) and DR5 (TRAIL-R2).

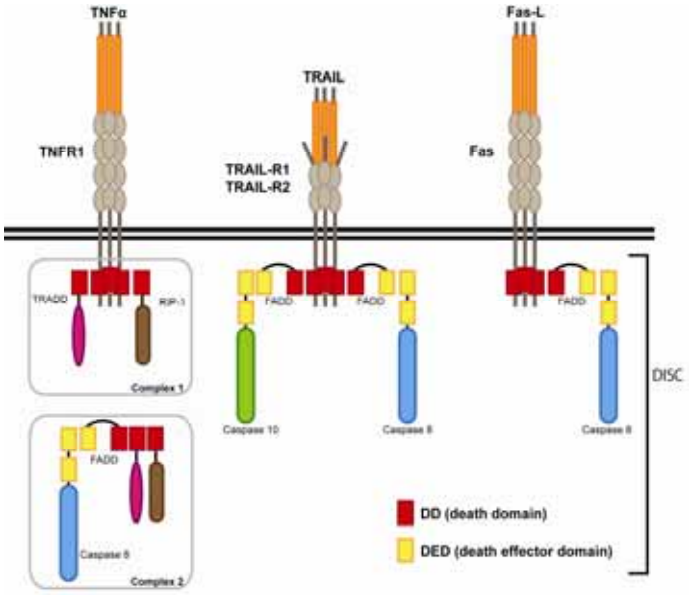


Figure 3: Cellular signalling pathways leading to activation of cell death. TNFR1 binds to TNFα, TRAIL-R1/R2 to TRAIL, and Fas to Fas-L. Through their death domain, the complexes (form with the receptor and its ligand) recruit adaptor proteins (RIP-1 and TRADD, FADD) and initiator caspases (-8 or -10; DISC formation), leading to the effector caspases activation.

As the Fas induction was the first describes, its example will be detailed to illustrate all the DR signalling pathways. After multimerisation of Fas via its ligand Fas-L, an intracellular ubiquitous molecule, FADD, is recruited via its DD. FADD also contains a NH_2-terminal death effector domain (DED), thanks to which it will be able to attract the procaspase-8 (zymogen, initiator caspase). The complex formed with Fas, Fas-L, FADD, and caspase-8 is called death-inducing signalling complex (DISC). The procaspase-8 and DISC proximity will lead to their mutual cleavage and to the liberation of active caspase-8 in the cytosol mediating activation of the effector caspases.

c. The soluble inter-membrane mitochondria protein (SIMP) family

Mitochondria are an essential organelle for the cellular survival, assuring the production of adenosine triphosphate (ATP) via the respiratory chain. It also plays an important role in the apoptotic death, sequestering in the intermembrane space a series of soluble molecules, which when released induce apoptosis. Most of the current data are compatible with the notion that the $\Delta\psi_m$ decrease constitutes an irreversible event of apoptotic process regulated by members of the Bcl-2/Bax family (Deshmukh *et al.* 2000; Goldstein *et al.* 2000). Changes in the mitochondrial inner membrane function are accompanied by an increase in outer membrane permeability, leading to the release of soluble intermembrane proteins (SIMP; reviewed in Bettaieb *et al.* 2003). The most important of these molecules is the cytochrome C, a small heme protein, which is an essential component of the electron transport chain (between complexes III and IV). In its absence, the stream of electrons is interrupted, stopping the production of ATP, and leading to superoxide radical anion accumulation.

During apoptosis, the cytochrome C is released in the cytosol and initiates the caspase cascade (activation of the caspase-9, initiator caspase). Furthermore, high level of cytosolic Ca^{2+} stimulates the generation of factors, including reactive oxygen species (ROS) and free fatty acids, known to contribute to the opening of the mitochondrial permeability transition pore (PTP; reviewed in Jeong and Seol 2008). Concomitantly, several other

mitochondrial molecules are released in the cytosol, mainly apoptotic peptidase activating factor 1 (Apaf1), adenosine kinase (AK), Smac/DIABLO (second mitochondria-derived activator of caspases/direct IAP-binding protein with low pi), and the serine protease Htra2/OMI, neutralizing the inhibitive effect of the inhibitor of apoptosis proteins (IAP) on caspases-3, -7, and -9 activities, thus facilitating their activation by the cytochrome C. Furthermore some procaspases are present in the intermembrane space and their activation may be mediated by the heat shock proteins HSP60 and HSP10.

Finally, the mitochondrial flavoprotein apoptosis-inducing factor (AIF) and the endonuclease G, also members of the SIMP family, when released quickly migrate to the nucleus where they contribute to the chromatin condensation and to the DNA fragmentation, respectively (Kim *et al.* 2005).

d. The caspase family

The first evidence of primordial role of *c*ysteine-*asp*artic acid prote*ases* or caspases in the execution of a death signal came from the identification of the first proapoptotic *C. elegans* Ced-3 gene, coding for a homologue of the human interleukin-1β converting enzyme (ICE/IL1-β) that rapidly became the prototypic member for caspases (reviewed in Thornberry and Lazebnik 1998; Cerretti *et al.* 1992). The caspases are conserved across animal species with functional homologues (*i. e.*, orthologues) in *D. melanogaster*, *C. elegans*, *Xenopus laevis*, and sponges, suggesting evolution from a conserved protease superfamily. The caspases exist within the cell as inactive proforms or zymogens. After induction of apoptosis, these zymogens can be cleaved to form active enzymes (fig. 4). These proteases have a conserved structure comprising an NH_2-terminal prodomain (variable size) and two domains which after cleavage will become the small and large subunit (p10 and p20 respectively). Members of this endoprotease family possess a catalytic site comprising a cysteic residue (directly implicated in the catalytic process) within a QACXG motif (X can be R, Q or G). These enzymes recognize and successively cleave polypeptide chains, after an aspartate residue, on their carboxyl-termini.

The members of the caspase family can be grouped in various ways (reviewed in Fuentes-Prior and Salvesen 2004). They are often grouped according to their substrate specificity, their functional role, or their prodomain structure (NH_2-terminal peptide or prodomain). The latter one will be detailed. Eleven caspases have so far been identified in humans and their prodomain serves to divided caspases in two broad groups. Long prodomains are found in caspase-1, -2, -4, -5, -8, -9, -10, -12, and -13 and form the group of initiator (apical) caspases while short prodomains are present in caspase-3, -6, -7, -11, and -14, and form the group of effector (executioner) caspases. This prodomain is removed during their processing to active form and regulate activity of the caspases. Two unique domains are found within the long prodomains: the death effector domain (DED, which mediates hydrophobic associations) and the caspase recruitment domain (CARD, which mediates electrostatic interaction), both serving to recruit caspases to specific complexes (fig. 2 and 39).

Activation of the caspases represents a pivotal step in the cell death signalling cascade and serves to amplify the apoptotic signal. These caspases are responsible for the cleavage of the key cellular/nuclear proteins (*e. g.*, cytoskeletal proteins), the activation of DNases, and the inhibition of DNA repair enzymes, leading to the typical morphological changes (*e. g.*, cell shrinkage and DNA ladder) observed in cells undergoing apoptosis. However this cascade (sequential activation) can be regulated at many points (*e. g.*, IAP and SIMP).

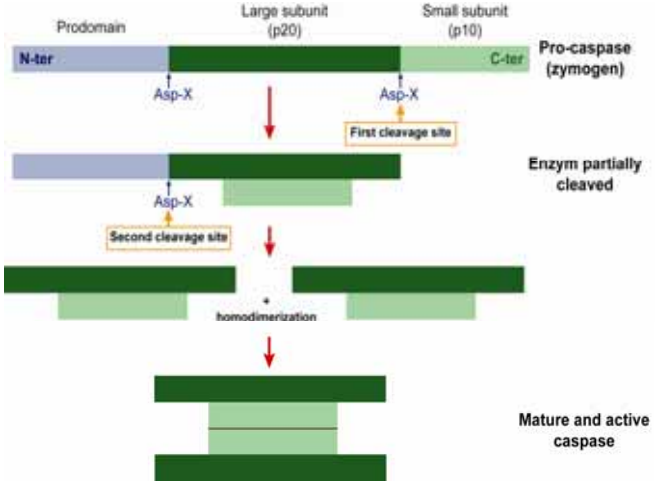


Figure 4: Caspases activation. As zymogen, these enzymes encompass a prodomain, a small (p10), and a large subunit (p20; containing the catalytic sites). Their cleavage (in two subsequent steps, release of the small subunit then of the large one) leads to the homo and heterodimerization thus to the caspases activation.

e. The inhibitor of apoptosis protein (IAP) family

IAP (inhibitor of apoptosis) genes were originally described in baculoviruses (Crook *et al.* 1993) preventing the release of apoptotic process in infected insect cells. IAP bear at least one NH_2-terminal motif termed baculoviral IAP (inhibitor of apoptosis) repeat (BIR).

Some IAP can also bear other motifs, such as a really interesting new gene (RING) finger domain and a caspase activation recruitment domain (CARD), at their COOH-terminal. Many IAP family members have been identified in diverse species ranging from virus to mammals (fig. 6).

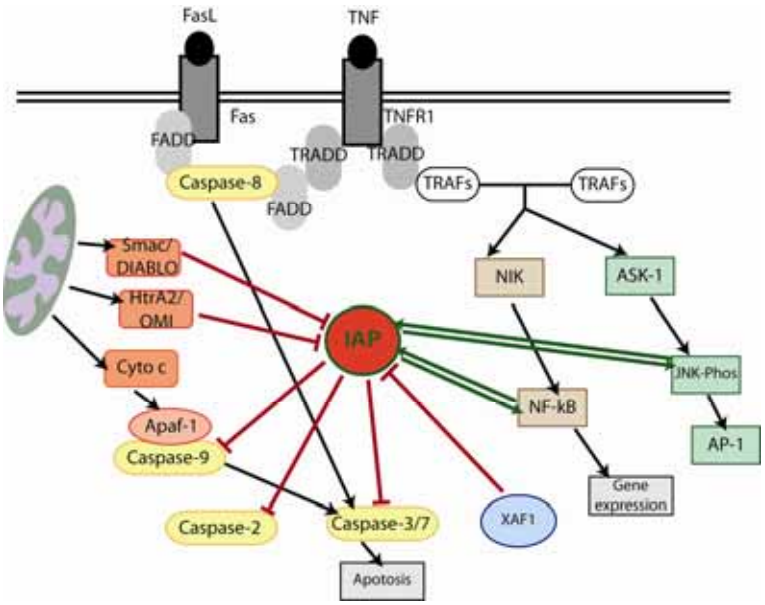


Figure 5: Complex regulatory mechanism of inhibitor of apoptosis protein (IAP) in apoptosis (adapted from the review of Wei *et al.* 2008). IAP are negatively regulated by IAP antagonists (*e. g.*, Smac/DIABLO, HtrA2/OMI, and X-linked IAP-associated factor (XAF1)) and execute positive feedback on NF-kB and c-Jun NH_2-terminal kinases (JNK) pathways.

Although the BIR domain is required for the antiapoptotic functions of the IAP family members, not all BIR-containing proteins have antiapoptotic functions. Studies showed that IAP can directly inhibit the caspases (via their BIR domain) by promoting degradation of their actives forms, or by sequestering caspases away from their substrates

(Tenev *et al.* 2005; fig. 5). In the extrinsic activation pathway of apoptosis, IAP do not bind caspase-8 but inhibit its substrate caspase-3 while in the intrinsic pathway they bind procaspase-9 or directly caspase-9 and -3 thus interfering with the apoptosome formation. However, several endogenous IAP antagonists counteract their antiapoptotic activity (*e. g.*, Smac/DIABLO and HtrA2/Omi).

A subgroup of IAP (including survivin in mammalian cells and BIR-1 and BIR-2 in *C. elegans*) possesses not only a caspase inhibitor activity, but also regulates mitotic spindle formation and cytokinesis (Li *et al.*1998; Uren *et al.* 1999). Due to this unique feature, some scientists raised the suggestion to name this subgroup "BIR domain-containing (BIRC) proteins" (reviewed in Verhagen *et al.* 2001).

C. Survivin

In addition to their cell death regulating activities several IAP have a complementary role in cell division, because of survivin, a prototypic representative of the IAP family. Survivin is the smallest member of the IAP family. It possesses a single BIR domain and no RING domain.

Accordingly, the dual role of human survivin is based upon its function as (**i**) a chromosomal passenger protein that controls segregation of chromatin/regulation of cytokinesis and (**ii**) a caspase inhibitor that prevents processing/activation of caspases (Beltrami *et al.* 2004).

Survivin is a chromosomal passenger protein and physically associates with the inner centromere protein (INCENP), Aurora B (Honda *et al.* 2003) and Borealin/Dasra B (Gassmann *et al.* 2004), thus forming the chromosomal passenger complex (CPC; fig. 7). These interactions are required for a number of key functions during mitosis. Survivin mediates proper targeting of chromosomal passenger proteins to kinetochores (Vagnarelli and Earnshaw 2004). Besides, functional evidence implicates a role for survivin in microtubule dynamics promoting their stability (Giodini *et al.* 2002).

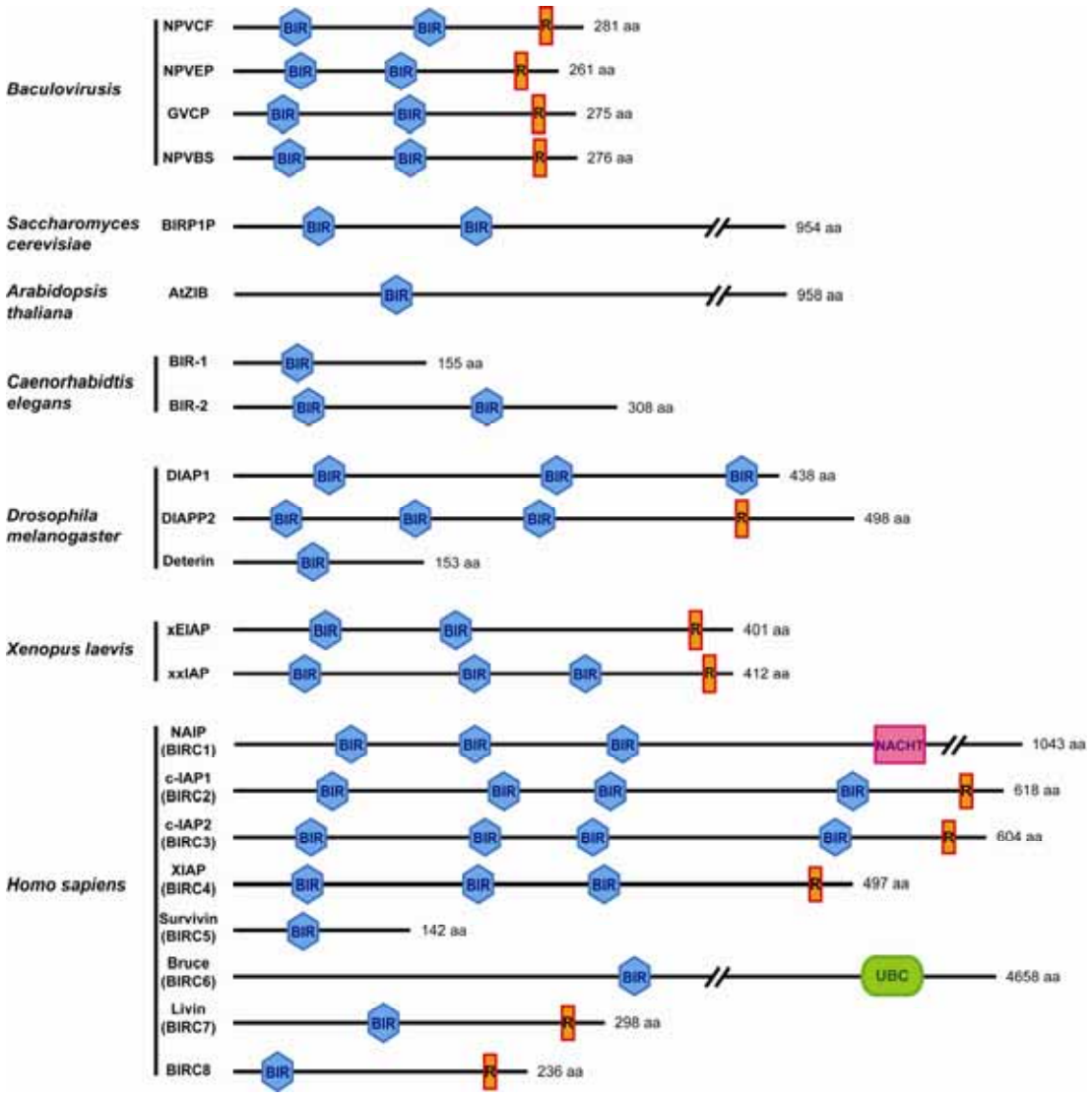


Figure 6: Inhibitor of apoptosis proteins from different species with their respective size (BIR: baculovirus IAP repeat, R: RING domain, UBC: ubiquitin-conjugating catalytic core domain, and NACHT: NTPase domain; drawn to scale). IAP proteins are members of a conserved family of proteins identified in species ranging from virus to mammals. The common structural feature is the presence of at least one BIR domain.

Initially, survivin was described as an inhibitor of caspases, promoting the degradation of active caspase-3, caspase-7, and caspase-9 (Altieri 2003). Survivin lacks the structural motifs present in other IAP that mediate binding to caspases; however, survivin inhibits apoptosis and its mechanism of action may be more sophisticated than direct caspase inhibition. Moreover, survivin can be inhibited by Smac/DIABLO. Survivin thus

occupies a central position between the proapoptotic and the antiapoptotic factors (Dohi *et al.* 2004).

Survivin is usually expressed in embryonic tissues or rapidly dividing cells (*e. g.*, thymus, placenta, and CD34+ stem cells) but is highly expressed in most tumor cell types. Indeed survivin is a well-established analytical marker of neoplasms (reviewed in Li 2003; Ambrosini *et al.* 1997) and is an unfavorable prognostic marker. The survivin is an inhibitor of apoptosis, which promotes cancer by suppressing the body's ability to limit excessive cell growth. These characteristics make survivin a particularly attractive target for selective therapies.

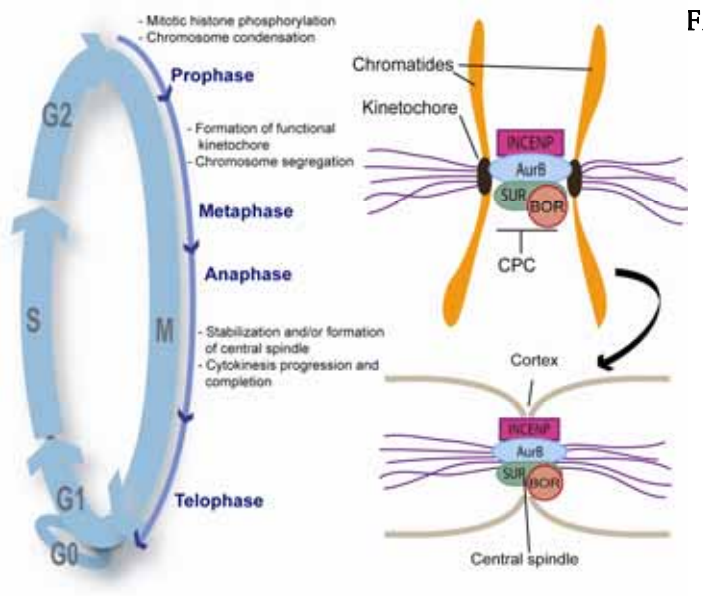


Figure 7: Survivin and regulation of cell division. The chromosomal passenger complex (CPC) includes the inner centromere protein (INCENP), borealin (BOR), survivin (SUR), and Aurora B (AurB). This complex coordinates chromosomal and cytoskeletal events of mitosis. Survivin is an important mediator of centromere and midbody docking of Aurora-B during mitosis.

The role of survivin as a chromosomal passenger protein is evolutionary conserved for survivin-like molecules in *C. elegans* (BIR-1; Fraser *et al.* 1999), *D. melanogaster* (deterin; Jones *et al.* 2000), and yeast (Huang *et al.* 2005). Survivin-homolog expression is correlated with cell division, localized to chromosomes and spindle midzone. In addition, early stage development lacking of survivin-like molecules display chromosomal and spindle defects (Fraser *et al.* 1999; Speliotes *et al.* 2000). Besides, only deterin is able to deter cells from caspase-induced apoptosis.

D. Porifera

At first sight, what differentiates plants and animals is the lack of sensitivity and mobility of plants but both have the ability to feed, grow, and reproduce. That is probably why sponges were originally describes as plants. Thanks to Dujardin (Felix Dujardin, 1801-1860), their animal nature was recognized at the middle of the nineteenth century (Dujardin 1841).

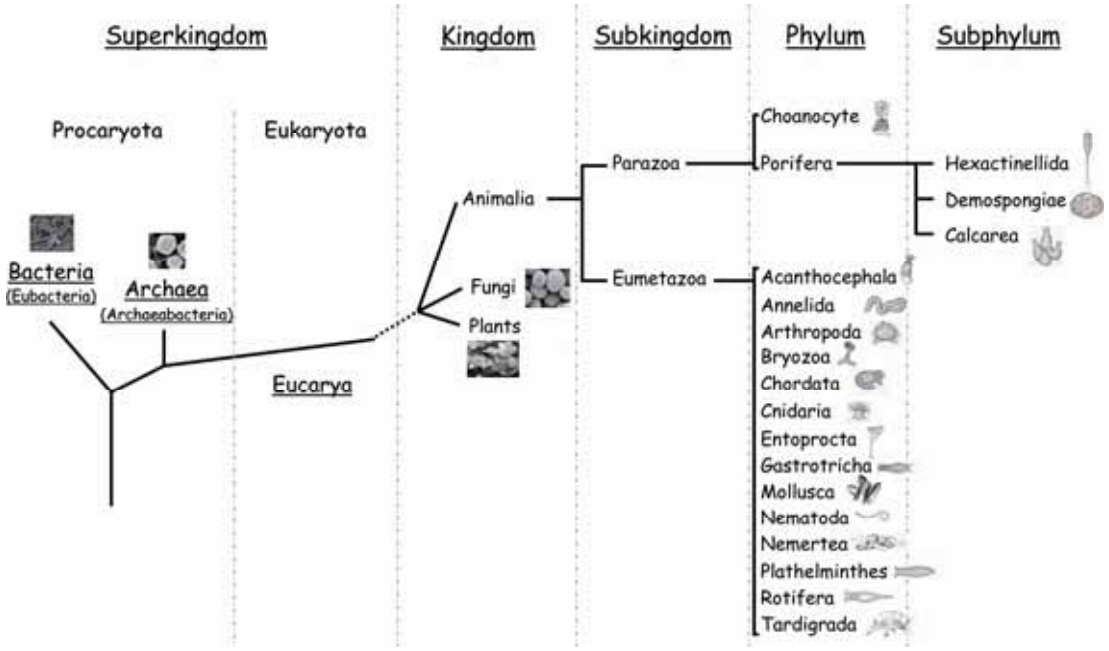


Figure 8: Phylogenetic tree. This evolutionary tree is showing the evolutionary relationships among various biological species that are believed to have a common ancestor.

Prokaryotes = Eubacteria + Archaea

Eubacteria: (now called bacteria) unicellular organisms characterised by the absence of nuclear membrane, by a single naked chromosome (occasionally two), and by a very small range of organelles (generally only plasma membrane and ribosome).

Archaea: in the past they were viewed as an unusual group of bacteria (they possess neither nucleus nor organelle) and named archaebacteria; however they are genetically and metabolically different from all other known bacteria. Archea are believed to be "living fossils" bridging the evolution gap between the bacteria and the eukaryotes.

Eukaryota (or Eukaryotes): possess chromosomes (with nucleosomal structure) separated from the cytoplasm by a nuclear and distinct cytoplasmic organelles allowing compartmentalisation of function.

Sponge cells are differentiated but do not form “real tissue”, thus sponges are ranking in a unique phylum “Porifera” (from Latin *porus* "pore" and *ferre* "to bear"; fig. 8). In fact, as the cells are totipotent, the cellular differentiation is reversible and can re-differentiate in another specialised form.

Archeocyaths are organisms with sponge type morphology, currently considered as an extinct class of demosponge. They appeared during the Early Cambrian (around 10 MY above the Cambrian/Precambrian; see appendix 1). Archeocyaths are believed to have comprised more than half of the marine reefs biomass during Paleozoic. Particularly abundant during Middle Cambrian, they declined and disappeared in the Late Cambrian. During Middle Paleozoic, stromatoporoids (calcareous poriferan) are known to be the main reef builders. They declined during the first big extinction (Ordovician); however, they still participated to the reef construction during Middle and Late Devonian. They became extinct at the end of the Devonian. The possible sponge emergence, in their modern known forms, dates from the early Tommotian (approximately 535 to 530 MY ago).

Until recently, sponge phylogeny has been source of strife. Porifera phylum was previously thought to be monophyletic and divided in two subphyla: Symplasma (containing the Hexactinellida group) and the Cellularia (containing the Calcarea and the Demospongia groups; Mehl *et al.* 1997). Nowadays, the phylum reorganisation and the paraphyletic origin, based on molecular data (*e. g.*, small and large ribosomal subunit (SSU or LSU) data; Schütze *et al.* 1999, Borchiellini *et al.* 2001, and Müller and Müller 2003) are accepted.

> A **monophyletic** group comprises a single common ancestor and all the descendant of that ancestor, in opposition to a **paraphyletic** group which does not include all the descendants of the most recent common ancestor.

Sponge is an organism which seems easy to describe. A sponge is a sedentary, filter-feeding metazoan (*i. e.*, animal) which has a single layer of flagellated cells that drive a unidirectional current of water through its body. Such a brief description does not do them

justice. Sponges are an ancient and highly successful group of animals and can be found in many freshwater, all marine, and even aerial habitats. These sedentary animals occur in rivers and streams, from rock pools to deep ocean (up to 5,000 metres), from frozen arctic seas to warm tropical seas. There are about 10,000 known species, their basic organisation is pretty simple and remains fairly constant, and they manage to show a great variety of forms.

The sponges are, with some exceptions, sessile, *i. e.*, sedentary animals, that live on a substrate.

Figure 9: Examples of sponges from different poriferan classes. (A) *Suberites domuncula*, (Demospongia, the orange mass), lives in symbiosis with a hermit crab (*Paguristes eremita*) (B) *Sycon raphanus*, Calcarea, and (C) *Aphrocallistes vastus*, Hexactinellida.

A sponge is formed by two layers: the pinacoderm and the choanoderm, surrounding an acellular layer: the mesohyl. The mesohyl contains small needles or spicules. The word spicule is a generic term which includes the mineral extracellular secretion that is present in the tissues of various groups of invertebrates. They are formed either by silica crystals (SiO_2/H_2O), or calcite (or calcium carbonate, $CaCO_3$), or are composed of spongin (collagen-like fibres). Those fibres confer stability and rigidity to the structure but allow also the Porifera phylum to be subclassified in three classes (fig. 9 and 10): the Calcarea (spicules made out of calcium carbonate), the Demospongia (90% of all species, spicules consisting of spongin fibres, mineral silica, or both) and the Hexactinellida (siliceous spicules).

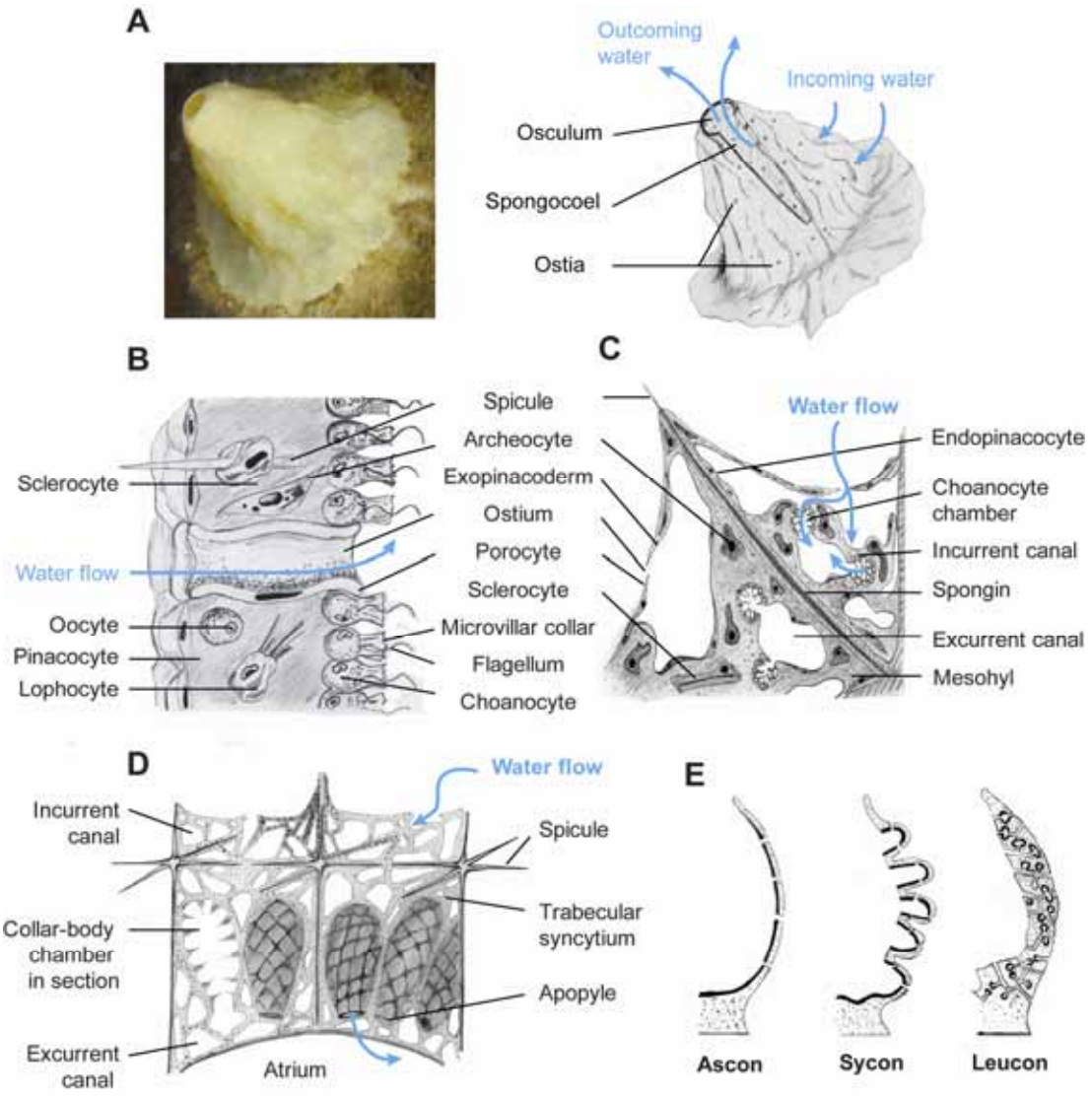


Figure 10: Sponge general and detailed anatomy. Section through Cellularia sponges (B, asconoid sponge, adapted from Rigby 1987 and C, freshwater leuconoid sponge modified from Ax 1996) and through Symplasma sponge (D, redraw from Bergquist 1978). Representation of the three body-types of sponge: ascon, sycon, and leucon (E).

Bodies of demosponges and calcareous sponges (Cellularia group; fig. 10B and C) are composed of cells organized into two types of tissues, epithelioid and connective. Sponge epithelia-like tissues are the pinacoderm and the flagellated choanoderm. The pinacoderm covers the outer surface of the body (exopinacoderme), lines the incurrent and excurrent canals (endopinachoderm), and consists of two types of differentiated cells: pinacocytes (flattened cells) and porocytes (form the ostia of all asconoid as well as many syconoid and leuconoid sponges).

Each porocytes surrounds a pore which diameter is regulated by contraction of cytoplasmic filaments. The choanoderm consists of flagellated cells or choanocytes (*i. e.*, collar cells). Those cells are able to generate the water flow through the sponge. Choanocytes have an apical collar of long microvilli around a single flagellum (fig. 10B). The connective tissue layer between the pinacoderm and the choanoderm is called the mesohyl ("middle wood") because it forms a bushy, fibrous network. The mesohyl is composed of a proteinaceous, gel-like matrix that contains differentiated, as well as undifferentiated cells and spicules. Among the numerous cells present in the mesohyl are the archeocytes (macrophage-like, progenitor cells), which are large ameboid cells bearing a prominent nucleus and numerous large lysosomes. Archeocytes are totipotent and can differentiate into any other sponge cell types. These cells are playing a role in digestion (phagocytic function) and internal transport. Lophocytes (*i. e.*, crest cells) are also ameboid cells, responsible for collagens fibres secretion and maintain through the mesohyl. Spongocytes (only for the Demospongiae taxon) secrete collagen as well but it polymerizes into dense skeletal fibres: the spongin. Sclerocytes (hard cells) secrete the mineral skeletal spicules of many sponges. Myocytes are muscle-like cells containing actin and myosin that aggregate around the oscula of some demosponges, and regulate the oscula aperture size, stepping in water flow control. Lastly, oocytes and spermatocytes are reproductive cells that undergo gametogenesis in the mesohyl to form sperm and eggs.

The symplasma group is composed of the hexactinellids (fig. 10D) and lacks the sheet-like pinacoderm pavement and the cellular choanoderm. Instead, tissue is arranged in a three-dimensional cobweb-like structure or trabecular syncytium. The choanoderm is replaced by a choanosyncytium. Each group of collar bodies (collar and flagellum but lacking a nucleus) occupies a syconoid-like pocket that is supported by the trabecular network. Each strand in the trabecular syncytium surrounds and encloses an axis of mesohyl. The mesohyl contains bundles of collagen fibres, spicules and cells (*e. g.*, sclerocytes, archeocytes, and presumably germ cells).

The sponge thrives by filtering out nutritious particles (*e. g.*, phytoplankton, bacteria, and organic debris; fig. 10A). Huge amounts of water are so refined, reducing water turbidity. A sponge approximately pumps/filtrates a volume of water equal to its

body volume every 10 seconds. The sponge internal cavity or "atrium" communicates with the surroundings via two kinds of orifices: perforated cells or "porocytes" and inhalant pores or "ostia". Three body-plans occur among Porifera: asconoid, syconoid and leuconoid (fig. 10E). The asconoid form, the simplest of the three, is found only in small sponges and is characterized by ostia leading directly to the atrium, which is covered with choanocytes. In the syconoid form, the choanocytes no longer cover the atrium but a multitude of small channels, radial canals, which give the advantage of increasing the contact surface between water and cells, leading to a greater absorption of food. Finally, the leuconoid form is the most complex, the ostia and canals lead to many rooms carpeted with choanocytes. Calcarea displays asconoid and syconoid types, Demospongiae leuconoid type and Hexactinellida syconoid type.

E. Porifera: a model to study apoptosis and immune pathways

Porifera is the evolutionary oldest, still extant, and the closest relative of the hypothetical ancestral animal, the Urmetazoa, from which the metazoan lineages diverged (more than 700 MYA; reviewed in Müller 2001). Besides, it has been proposed that during the evolution from Fungi to the Metazoa, an increase in gene number (gene duplication) and novelties (*e. g.*, novel adhesion molecules, and formation of new domains) occurred (reviewed in Müller *et al.* 2001).

Metazoa have developed complex systems to maintain homeostasis, *e. g.*, avoid unwanted biological invasion and eliminate cells. Immunity involves both specific and non-specific components. While native/innate immunity is described in all metazoans, adaptive (artificially acquired) immunity was first thought to be restricted to vertebrates, and especially mammalians. However some specific molecules (precursor) of acquired immunity have been identified in sponges. Macrophage-derived cytokine-like molecule was cloned (Samal *et al.* 1994), sharing high similarity with the human pre-B-cell colony-enhancing factor, and was up-regulated during the grafting process (reviewed in Müller *et al.* 1999; Müller *et al.* 1999). Additionally, several molecules containing Ig-like domains were identified (*e. g.*, *Geodia cydonium* receptor tyrosine kinase (RTK) possesses two complete

(and highly polymorphic) immunoglobulin [Ig]-like domains, Pancer *et al.* 1998 and Blumbach *et al.* 1999). Likewise, studies revealed that sponges are provided with the key elements of the innate immune system (*e. g.*, TIR and DD (Wiens *et al.* 2007) and antifungal defense, (1->3)-β-D-glucan-binding protein (Perovic´-Ottstadt *et al.* 2004)). Another good example is the antiviral defense via 2', 5' oligoadenylate synthetase (2', 5' OAS; Wiens *et al.* 1999). The 2', 5' OAS/RNase L system is an innate immunity pathway that responds to a pathogen-associated molecular pattern (double stranded RNA) to induce degradation of viral and cellular RNA which thereby block viral infections. Interferons can act as inducers and regulate this reaction. However, even if 2', 5' OAS remained in until humans, it has been lost in *C. elegans* (reviewed in Müller 2001). Corroborated with the non-existence of Nod-like receptors (NLR) in *D. melanogaster* and *C. elegans*, the poriferan phylum appears to be an attractive model to study components and mechanisms of the innate immune system. Furthermore, LPS (lipopolysaccharides) activation pathway was well characterized and the major molecules identified (*e. g.*, receptor and adaptor) in *S. domuncula* (Wiens *et al.* 2005 and 2007).

Similarly, key molecules of the apoptotic pathway were found in poriferan genomes. Demosponges *G. cydonium* and *S. domuncula* Bcl-2 homologues were cloned and their functions assessed (Wiens *et al.* 2000 and 2001). Another core component of the apoptotic machinery is the caspases (see p. 12). Caspases or cysteine-aspartic acid proteases are essential effector and regulator of apoptosis, necrosis, and inflammation. So far, several caspase-like proteins have already been identified in demosponges (Wiens *et al.* 2002, 2003, and 2007). Another argument in favour of using Porifera as a model organism is the existence of death domain (DD) in their genome while no death-domain-containing molecule has been identified in *C. elegans*. DD is a homotypic protein interaction module composed of a bundle of six alpha-helices. Some DD-containing proteins are involved in the regulation of apoptosis and inflammation through their activation of caspases and NF-kB, which typically involves interactions with TNF (tumour necrosis factor) cytokine receptors. The poriferan proapoptotic molecule DD2 (containing two death domains) was identified in *G. cydonium* and revealed sequence homology to human FADD and FAS (Wiens *et al.* 2000).

With completely different consideration, sponge long span-life and easy culture/maintaining conditions are really attractive advantages. Moreover, sponges are not subjected to the same evolution constraint as most of the well-known model organisms, such as *D. melanogaster* and *C. elegans*, which have truncated and highly specialized genomes. The latter animal lineages appear to have undergone an accelerated evolution (Gamulin *et al.* 2000; Kortschak *et al.* 2003) whereas due to a combination of extended life span and extensive niche breadth, Porifera demonstrates a bradytelic evolution (Müller 1998). Therefore, poriferan genes display a closer sequence similarity to their human homologues than to the respective molecules of worm and fly (Gamulin *et al.* 2000). In addition, even though *D. melanogaster* is genetically more complex than *C. elegans* and exhibits a slower evolutionary rate compared to nematodes (Coghlan and Wolfe 2002) various proteins and pathways seem to be fly-specific, establishing only a restricted comparability with the human system. Comparative transcriptomics and proteomics also revealed several examples of poriferan molecules, genetic features, protein domains, etc. that remained conserved throughout evolution, but which have been lost in *C. elegans* and/or *D. melanogaster* (reviewed in Müller 2001 and Müller *et al.* 2001).

Finally, primmorphs (aggregates of proliferating sponge cells, Custodio *et al.* 1998) provide a unique system to study and an easy assess to the main biological functions (*e. g.*, proliferation, apoptosis, and immunity). Primmorphs are aggregates of sponge cells which have recovered their proliferating properties, mimicking an entire and functional sponge but which are easier to deal with.

III. AIM OF THE STUDY

Analyses in early-branching Metazoa so far propose an exclusive role of survivin as a chromosomal passenger protein, whereas only later during evolution the second, complementary anti-apoptotic function might have arisen, concurrent with increased organismal complexity. To lift the veil on the ancestral function(s) of this key regulatory molecule, a survivin homologue (SDSURVL) of the phylogenetically oldest extant metazoan taxon (phylum Porifera) was identified and functionally characterized. Similarly, Nod-like receptors (NLR) are present in Echinodermata, in higher Metazoan, but not in Nematoda or Hexapoda. A poriferan NLR-like, TIR-LRR containing protein (SDTILRc), was cloned and its involvement in innate immunity elucidated. Additionally, two atypical caspases (SDCASL and SDCASL2) were also discovered. These two caspases possess a small prodomain (thus they belong to the subgroup of effector caspases) and a remarkably long COOH-terminal domain containing one or several functional double stranded RNA binding domains (dsrm). A dual role in apoptosis and pathogen sensing was proposed and studied.

IV. MATERIALS AND METHODS

A. Materials

1. Chemicals and ready-to-use solutions

Acetic acid ($C_2H_4O_2$)	Roth, Karlsruhe
Agar	Roth, Karlsruhe
Agarose	Roth, Karlsruhe
Albumin bovine, fraction V	Roth, Karlsruhe
β-mercaptoethanol (C_2H_6OS)	Roth, Karlsruhe
BCIP ($C_8H_4BrClNO_4PK_2$)	Sigma, Taufkirchen
Blocking reagent	Roche, Mannheim
Boric acid (H_3BO_3)	Roth, Karlsruhe
Bromophenol blue ($C_{19}H_{10}Br_4O_5S$)	Serva, Heidelberg
BugBuster™	Merck, Darmstadt
CDP-Star™ ready to use	Roche, Mannheim
Calcium chloride ($CaCl_2$)	Roth, Karlsruhe
Cadmium ($CdCl_2$)	Sigma, Taufkirchen
Chloroform ($CHCl_3$)	Roth, Karlsruhe
Coomassie brilliant-blue R250	Serva, Heidelberg
dNTP-mix	Peqlab, Erlangen
Di-Sodium hydrogen phosphate (Na_2HPO_4)	Roth, Karlsruhe
Diethyl pyrocarbonate (DEPC, $C_6H_{10}O_5$)	AppliChem, Darmstadt
DIG easy Hyb	Roche, Mannheim
Dimethyl formamide (DMF, C_3H_7NO)	Roth, Karlsruhe
Dimethyl pyrocarbonate (DMPC, $C_4H_6O_7$)	Sigma, Taufkirchen
Dimethyl sulphoxide (DMSO, C_2H_6OS)	Roth, Karlsruhe
Dithiothreitol (DTT, Cleland's reagent, $C_4H_{10}O_2S_2$)	Roth, Karlsruhe
Ethylenediaminetetraacetic acid (EDTA, $C_{10}H_{16}N_2O_8$)	Roth, Karlsruhe
Ethanol (C_2H_6O)	Roth, Karlsruhe
Ethidium bromide ($C_{21}H_{20}BrN_3$)	Boehringer, Mannheim
Foetal calf serum (FCS)	Sigma, Taufkirchen
Formaldehyde (CH_2O)	Roth, Karlsruhe
Formamide (CH_3NO)	Roth, Karlsruhe
Gel code® blue stain reagent	Pierce, Bonn
Glucose (αD+, $C_6H_{12}O_6$)	Roth, Karlsruhe
Glycergel™	Dako, Hamburg
Glycerol ($C_3H_8O_3$)	Sigma, Taufkirchen

Glycine ($C_2H_5NO_2$)	Roth, Karlsruhe
Guanidine hydrochloride (GuHCl)	Sigma, Taufkirchen
N-2-Hydroxyethyl piperazine-N'-2-ethane sulphonic acid) (HEPES, $C_8H_{18}N_2O_4S$)	Roth, Karlsruhe
Isopropyl-β-D-thiogalactopyranoside (IPTG, $C_9H_{18}O_5S$)	Roth, Karlsruhe
Isopropanol (C_3H_8O)	Merck, Darmstadt
L-glutamine ($C_5H_{10}N_2O_3$)	Sigma, Taufkirchen
Lipopolysaccharides (LPS) from *Escherichia coli* 026:B6	Sigma, Taufkirchen
Magnesium sulfate ($MgSO_4$)	Merck, Darmstadt
Maleic acid ($C_4H_4O_4$)	Roth, Karlsruhe
Manganese chloride ($MnCl_2$)	Roth, Karlsruhe
MATra-A reagent	IBA, Göttingen
3-(N-morpholino) propanesulfonic acid (MOPS, $C_7H_{15}NO_4S$)	Roth, Karlsruhe
NBT ($C_{40}H_{30}N_{10}O_6$-2Cl)	Sigma, Taufkirchen
NP-40	Sigma, Taufkirchen
Pam_3Cys-Ser-$(Lys)_4$	Alexis, Lörrach
Paraformaldehyde ($(CH_2O)n$)	Sigma, Taufkirchen
Phase lock gel	Eppendorf, Hamburg
Phosphate buffered saline (PBS)	Biochrom AG, Berlin
Pipes (Piperazine-N,N'-bis(2-ethanesulphonic acid))	Roth, Karlsruhe
Polyoxyethylene sorbitan monolaurate (Tween 20)	Roth, Karlsruhe
Poly (cytidylic-inosinic) acid potassium salt (poly(IC))	Sigma, Taufkirchen
Potassium dihydrogen phosphate (KH_2PO_4)	Roth, Karlsruhe
Potassium hydroxide (KOH)	Roth, Karlsruhe
Potassium chloride (KCl)	Roth, Karlsruhe
RNase Away	Fisher Scientific, Schwerte
Roti®-Quant R250	Roth, Karlsruhe
Saccharose (D+; $C_{12}H_{22}O_{11}$)	Roth, Karlsruhe
Sequagel complete buffer and XR monomer	BIOZYM, Hess.Oldendorf
Sodium acetate ($C_2H_3NaO_2$)	Roth, Karlsruhe
Sodium citrate ($Na_3C_6H_5O_7$)	Roth, Karlsruhe
Sodium chloride (NaCl)	Roth, Karlsruhe
Sodium dodecylsulfat (SDS; $C_{12}H_{25}NaO_4S$)	Roth, Karlsruhe
Sodium hydroxide (NaOH)	Roth, Karlsruhe
Sodium phosphate (NaH_2PO_4)	Roth, Karlsruhe
Sodium pyruvate ($C_3H_3NaO_3$)	Biochrom AG, Berlin
Sodium sulfate (Na_2SO_4)	AppliChem, Darmstadt
Supersignal west pico chemiluminescent substrate	Pierce, Bonn
Tris-HCl ($C_4H_{11}NO_3$-HCl)	Roth, Karlsruhe
Tryptone	Roth, Karlsruhe
Triton X-100	Sigma, Taufkirchen
TRIzol®	Invitrogen, Karlsruhe
Urea (CH_4N_2O)	Roth, Karlsruhe
X-gal ($C_{14}H_{15}BrClNO_6$)	Peqlab, Erlangen
Yeast extracts	Roth, Karlsruhe

2. Antibiotics

Ampicillin	Sigma, Taufkirchen
Carbenicillin	Roth, Karlsruhe
Geneticin (G418)	Roth, Karlsruhe
Gentamicin	Roth, Karlsruhe
Kanamycin	Roth, Karlsruhe
Zeocin	Invitrogen, Karlsruhe

3. Media

DMEM (Dulbecco's modified Eagle's medium)	Biochrom AG, Berlin
Luria Bretani (LB) broth	Roth, Karlsruhe
OptiMem	Invitrogen, Karlsruhe
RPMI (Roswell Park Memorial Institute medium)	Biochrom AG, Berlin
Seawater	Sigma, Taufkirchen

4. Enzymes

Benzonase nuclease	Merck, Darmstadt
Lysozyme	Sigma, Taufkirchen
Protease inhibitor cocktail	Roche, Mannheim
Proteinase K	Roth, Karlsruhe
Ribonuclease I (Rnase A)	Sigma, Taufkirchen
T4 DNA ligase	Fermentas, St. Leon-Roth
Trypsin-EDTA	Sigma, Taufkirchen
Apa I	Sigma, Taufkirchen
BamH I	Sigma, Taufkirchen
EcoR I	Sigma, Taufkirchen
Nco I	Sigma, Taufkirchen
Not I	Sigma, Taufkirchen
Sac I	Sigma, Taufkirchen
Xho I	Sigma, Taufkirchen
PCR master mix (2x)	Fermentas, St. Leon-Roth
peqGOLD DNA-polymerase	Peqlab, Erlangen
Platinum® Taq DNA polymerase high fidelity	Invitrogen, Karlsruhe
Taq recombinant DNA polymerase	Fermentas, St. Leon-Roth
Thermo Sequenase DNA polymerase	GE Healthcare, Freiburg

5. Antibodies

Anti-digoxigenin-AP, Fab fragments	Roche, Mannheim
Anti-GFP-horseradish peroxidase (HRP)	Miltenyi Biotec, Bergisch
Anti-human survivin (rabbit)	AbD Serotec, Düsseldorf
Anti-human tubulin (mouse)	AbD Serotec, Düsseldorf
Anti-His (COOH-terminal, mouse)	Invitrogen, Karlsruhe
Anti-rabbit-IgG-HRP (sheep)	AbD Serotec, Düsseldorf
Anti-mouse-IgG-HRP (mouse)	AbD Serotec, Düsseldorf
Anti-mouse-IgG-Alkaline phosphatase (AP, sheep)	Sigma, Deisenhofen

6. Markers

GeneRuler™ DNA ladder mix	Fermentas, St. Leon-Roth
RNA molecular weight marker I, DIG-labelled	Roche, Mannheim
Precision Plus Protein™ Dual Color	Bio-Rad, München
PeqGOLD Protein Marker I	PeqLab, Erlangen

7. Kits

FuGENE®6	Roche, Mannheim
GeneRacer™ kit	Invitrogen, Karlsruhe
High pure PCR product purification kit	Roche, Mannheim
High pure plasmid isolation kit	Roche, Mannheim
MATra reagents	IBA GmbH, Göttingen
Ni-NTA agarose	Qiagen, Hilden
Ni-NTA spin kit	Qiagen, Hilden
NucleoSpin® extract II	Macherey-Nagel, Düren
NucleoSpin® plasmid	Macherey-Nagel, Düren
TOPO TA cloning kit	Invitrogen, Karlsruhe
PCR DIG labelling mix PLUS	Roche, Mannheim
pGEM®-T vector system	Promega, Mannheim
Plasmid maxi kit	Qiagen, Hilden
Plasmid midi kit	Qiagen, Hilden
Pierce® BCA protein assay kit	Perbio, Bonn
pcDNA™3.1/CT-GFP-TOPO® expression kit	Invitrogen, Karlsruhe
pTrcHis TOPO® TA expression kit	Invitrogen, Karlsruhe
QIAprep spin miniprep kit	Qiagen, Hilden
RevertAid™H minus first strand cDNA synthesis kit	Fermentas, St. Leon-Roth
SuperScript™ III first strand synthesis for RT-PCR	Invitrogen, Karlsruhe

Supersignal-west-pico-chemiluminescent-substrate-kit	Pierce, Bonn
Thermo sequenase fluorescent labelled primer Cycle sequencing kit	GE Healthcare, Freiburg

8. Bacteria strains and cell-lines

E. coli XL1-*Blue MRF'* — Stratagene, Heidelberg
Genotype: D (mcrA) 183 D (mcrCB-hsdSMR-mrr) 173endA1 supE44 thi-1 recA1 gyrA96 relA1 lac [F' proAB lacIqZDM15 Tn10 (Tetr)]

E. coli DH5α
Genotype: *supE44*Δ(*lac*U169(φ80d*lacZ*ΔM15 *hsd*R17 *rec*A1 *end*A1 *gyr*A96 thi-1 relA1

E. coli NovaBlue Singles™ Competent Cells — Merck, Darmstadt
Genotype: *endA1 hsdR17* (rK12⁻mK12+) *supE44 thi-1 recA1gyrA96relA1lac*F' [*proA+B+laclqZΔM15*::Tn*10* (TcR)]

E. coli TOP10 — Invitrogen, Karlsruhe
Genotype: F-*mcr*AΔ (*mrr-hsd*RMS-*mcr*BC) Φ80lacZ ΔM15Δ*lac* 74*rec*A1*deo*R*ara*D139Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G

Homo sapiens HEK-293 (ATCC CRL 1573) — American Type Culture Collection (ATCC), Rockville (USA)

Mus musculus RAW-Blue™ — InvivoGen, Toulouse (FR)

9. Vectors (cloning and expression) and libraries (cDNA and genomic)

Gene JET™ PCR cloning kit	Fermentas, St. Leon-Roth
pCR®II TOPO TA cloning kit	Invitrogen, Karlsruhe
pGEM-T	Promega, Mannheim
pTrcHis2 TOPO® TA expression kit	Invitrogen, Karlsruhe
pEGFP-C2	Clontech, Heidelberg
pcDNA™3.1/CT-GFP-TOPO®	Invitrogen, Karlsruhe

pTriplEx2	Clontech, Heidelberg
Lambda Fix II	Stratagene, Heidelberg
S. domuncula genomic DNA library	Seack *et al.* 2001
S. domuncula cDNA library	Kruse *et al.* 1997

10. Laboratory equipment and supplies

ART filter-tips		Biozym, Hess. Oldendorf
Blotting nylon membrane, positively charged		Boehringer Mannheim
Cell culture bottles, 50 and 250 mL Falcon		BD Biosc., Heidelberg
Cell culture dishes, 35 and 90 mm ∅		Nunc, Langenselbold
Cell scraper		Nunc, Langenselbold
Cellophane film		Insula, Mannheim
Centrifuges:	Sorvall RC 5B	DuPont, Bad Nauheim
	Eppendorf centrifuge 5402	Eppendorf, Hamburg
	Heraeus Biofuge fresco	Kendro, Hanau
Centrifuges tube:	15 and 50 mL Falcon	BD Biosc., Heidelberg
Counting chamber Neubauer		Merck, Darmstadt
Cover glasses, 12 mm ∅		Roth, Karlsruhe
Cryotubes		Nunc, Langenselbold
Drigalski spatula		Roth, Karlsruhe
Electrophoresis apparatus		Bio-Rad, München
ELISA reader Titertek Multiskan MK II Plus		Bartolomey, Rheinbach
Filter paper for Western blot		Pharmacia, Freiburg
Fluorescence and light microscope AH3		Olympus, Hamburg
Fluorometer Fluoroskan II		Fisher Scientific, Schwerte
Freezer (basic model, -80°C)		Nunc, Langenselbold
Heatblock Thermostat 5320		Eppendorf, Hamburg
Heatplate IKA		Labortechnik, Staufen
Hybond-NX nylon membrane 50X11		Amersham, Braunschweig
Hyper film cassette tape		Amersham, Braunschweig
Laminar flow		Sterilbank Slee, Mainz
Li-Cor sequencer 4200/4300		MWG-Biotech, Ebersberg
Precellys 24		Bertin-Peqlab, Erlangen
Slides welder		Petra-electric, Burgau
iBlot™ dry blotting system		Invitrogen, Karlsruhe
Incubators:	Queue model QWJ 500	Nunc, Langenselbold
	BK Type 2/56	Ehret, Emmendingen
Mortar		Haldenwanger, Berlin
Multiwell plates ,96, 24, and 6-well plates		Nunc, Langenselbold
	12-well plates	BD Biosc., Heidelberg
Nitrocellulose membrane		Schleicher & Schüll, Dassel
Nylon membranes, positively charged		Roche, Mannheim

Pasteur pipettes		Roth, Karlsruhe
PCR reaction vessels (0.2 mL)		Dianova, Hamburg
Pestle		Haldenwanger, Berlin
pH-meter type CG 840		Schott, Mainz
PhosphoImager		Bio-Rad, Munich
Pipettes (2, 10, 20, 100, 200, and 1000 µL)		Gilson, Limburg-Offheim
Pipette tips		Eppendorf, Hamburg
Plastic bags Plastibrand		Brand, Wertheim
PowerPac 200/300		Bio-Rad, Munich
Reaction Vessels:	1.5 mL	Brand, Wertheim
	2.0 mL	Eppendorf, Hamburg
Slide		Roth, Karlsruhe
Sonifier Branson B-12		Heinemann, Schwaebisch
Spectrophotometers:	UV/Visible GelDoc 2K System PC	Bio-Rad, München
	Pharmacia LKB-Ultrospec III	Pharmacia Co., Freiburg
Spectorphotometer SmartSpec Plus		Bio-Rad, München
Sterile filter, pore size 0.2 µm		Sartorius, Göttingen
Sunrise microplate reader		TECAN, Mainz-Kastel
Trans-Blot® semi-dry electrophoretic transfer cell		Bio-Rad, München
Transilluminator with gel documentation system		Amersham, Braunschweig
Thermocyclers:	Bio-Rad iCycler	Bio-Rad, München
	Bio-Rad MyCycler	Bio-Rad, München
	Primus 96 Plus	MWG, Ebersberg
Thermomixer comfort		Eppendorf, Hamburg
UV-crosslinker Stratalinker 1800		Agilent tech., Waldbronn
Vortexer		Labotech, Wiesbaden
Waterbath		Köttermann, Hängingen
Whatman paper DE 81		Whatman, Göttingen
X-ray Fuji RX		Bechthold, Kelkheim

11. Experimental animals

Live specimens of *Suberites domuncula* (Porifera-Demospongiae-Tetractinomorpha-Hadromerida-Suberitidae-Suberites) were collected by scuba diving or by dragging near Rovinj (North Adriatic, Croatia) from depths between 10 and 35 m, every summer and spring. Fresh collected sponges were either immediately snap frozen in liquid nitrogen or were brought to Mainz (Germany) and kept in aquaria at 16°C, under controlled aeration before being used in the experiments.

12. Computer and online softwares

BIND	http://bond.unleashedinformatics.com/Action?idsearch=145161 Bader and Hogue 2000
BLAST	http://www.ncbi.nlm.nih.gov/BLAST/ Altschul *et al.* 1990
CLUSTAL V1.81	Freeware: ftp.ebi.ac.uk Thompson *et al.* 1994
COILS	http://www.ch.embnet.org/software/COILS_form.html) Lupas *et al.* 1991
CpGProD	http://pbil.univ-lyon1.fr/software/cpgprod_query.html Ponger and Mouchiroud 2001
DNAStar V6	DNAStar Inc.
DNASIS	Hitachi Software Engineering, Yokohama
e-Seq V2.0 DNA	LI-COR, MWG Ebersberg
EMBL-EBI	http://www.ebi.ac.uk/inc/head.html Emmert *et al.* 1994
ExPASy	http://www.expasy.org/tools/protparam.html Gasteiger *et al.* 2003
FASTA	http://www.ebi.ac.uk/htbin/fasta.py?request Pearson and Lipman 1988
GENEDOC	www.cris.com Nicholas and Nicholas 1997
Genescan	http://genes.mit.edu/GENSCAN.html Burge and Karlin 1997
GraBCas	http://wwwalt.med-rz.uniklinik-saarland.de/med_fak/ humangenetik /software Backes *et al.* 2005
iep	http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=iep Alland *et al.* 2005
iHOP	http://www.ihop-net.org Hoffmann and Valencia 2004
NPS	http://npsa-pbil.ibcp.fr/cgi-bin/primanal_hth.pl Combet *et al.* 2000
P2SL	http://www.i-cancer.org/p2sl/ Atalay and Cetin-Atalay 2005
PAST	http://folk.uio.no/ohammer/past/ Ryan *et al.* 1995

Pfam	http://sanger.ac.uk/cgi-bin/pfam/ Finn *et al.* 2006
PHYLIP	http://evolution.genetics.washington.- edu/phylip/html Felsenstein 1993
Primer 3	http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi Rozen and Skaletsky 2000
Prosite	http://www.ebi.ac.uk/searches/prosite.input.html Hulo *et al.* 2006.
PSORTII	http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=psort Nakai and Horton 1999
SMART	http://smart.embl-heidelberg.de/ Letunic *et al.* 2006
Software Li-Cor	e-Seq Release 3.0 DNA Sequencing Software
TESS	www.cbil.upenn.edu/cgi-bin/tess/tess? Schug 2008
TFSEARCH	www.cbrc.jp/research/db/TFSEARCH Heinemeyer *et al.* 1998
Treeview	http://taxonomy.zoology.gla.ac.uk/rod/treeview.html
TUPS	http://sparks.informatics.iupui.edu/Softwares-Services_files/tups.htm Zhou and Zhou 2003
Vector NTI Suite	http://www.invitrogen.com
WebGene	http://www.itb.cnr.it/sun/webgene/ Milanesi *et al.* 1999

13. Primers

Vector-specific primers (pBK-CMV and pGEMT)

BK	5' – ACA GGA AAC AGC TAT GAC CTT G – 3'	Tm = 64°C
CMVF	5' – CTT GAT TAC GCC AAG CTC – 3'	Tm = 54°C
CMVR	5' – TAG GGC GAA TTG GGT ACA C – 3'	Tm = 58°C
M13-20	5' – GTA AAA CGA CGG CCA GT – 3'	Tm = 59°C
SP6	5' – TAG GTG ACA CTA TAG AAT ACT CA – 3'	Tm = 55°C
T3	5' – CGG AAT TAA CCC TCA CTA AAG – 3'	Tm = 60°C
T7	5' – GTA ATA CGA CTC ACT ATA GGG C – 3'	Tm = 64°C

Tm = melting temperature. The Tm is calculated on the basis of the base composition (Rychlik and Rhoads 1989).

Vector-specific sequencing primers (IR labelled)

SP6 $_{700}$ / SP6 $_{800}$	5' – CTA TTT AGG TGA CAC TAT AG – 3'	Tm = 51°C
T7 $_{700}$ / T7 $_{800}$	5' – GTA ATA CGA CTC ACT ATA GG – 3'	Tm = 53°C

Degenerative primers

SDSURVL	Surv_deg_F1	5' – GAT GGC IGA GGC T/CGG CTT – 3'
	Surv_deg_R1	5' – CT/CA TGT CCA GTT TT/CA AG/AA – 3'
SDCASL1 / SDCASL2	Casp_deg_F	5' – T/CTI ATT/C/A CAA/G GCI TGT/C A/T/CG/TI GG – 3'

I = inosine

Gene-specific primers

Tubulin α

	Northern blotting:	
Tub_alpha F1	5' – ATG CGT GAG TGT ATC TCT ATC – 3'	Tm = 56°C
Tub_alpha R1	5' – CTT TCC GAC GGT GTA ATG AC – 3'	Tm = 61°C

Survivin

	Specific primers:	
IAP_sF1	5' – AAT TAC TGA AGC TGT TTG AA – 3'	Tm = 56°C
IAP_sR1	5' – TAA AAT GTC ATG TGC TTG TC – 3'	Tm = 56°C
	Northern blotting:	
IAP_nF1	5' – ATG GCT AAT ACG TAT GAG AGT – 3'	Tm = 54°C
IAP_nR1	5' – CTA ATA TAG GGC TCA TAC GC – 3'	Tm = 55°C
	In-situ:	
IAP_isF1	5' – TAT GAG AGT TGT GAC AGA GTA AA – 3'	Tm = 55°C
IAP_isF1	5' – ACA AGT AGT GCC CCT AAC AGA CT – 3'	Tm = 61°C
	Protein expression, cloning in pEGFP-C2:	
EcoRI_IAP-F	5' – AAGAATTC**ATG**GCTAATACGTATGAGAG – 3'	Tm = 64°C
BamHI_IAP-R	5' – AAGGATCCCG**ATG**TTCGTGTCTTTTCAG – 3'	Tm = 70°C

The Kozak sequence is underlined and the ATG/Met_{start} is shown in bold.

CASL

	Specific primers:	
CASL_sF1	5' – ACT AAC CTA CTG CAC TAA TCC – 3'	Tm = 54°C
CASL_sR1	5' – TAC TGA TAC TCA GAT AAT TTT GA – 3'	Tm = 53°C
	Northern blotting:	
CASL_F1	5' – AAG CGT TGG ATA CCA GTC GT – 3'	Tm = 63°C
CASL_R1	5' – TGC TGG TGG CTT CTT TTC TT – 3'	Tm = 63°C
	Protein expression, cloning in pTrcHis2 TOPO® TA:	
pTRc_C1Rf	5' – CAA CAG GCT GGA ATT GTA CG – 3'	Tm = 63°C
pTRc_C1Er	5' – GTT ATA GAA ACC TAA CTC AAC ATT AGC – 3'	Tm = 59°C

CASL2

	Specific primers:	
CASL2_sF1	5' – GTT TTT GAC GCA GCT GGC AAC – 3'	Tm = 68°C
CASL2_sR1	5' – TGT TGT CAT AAC ATT CAC AAG ACA TGT – 3'	Tm = 67°C
	Northern blotting:	
CASL2_F1	5' – GTG ACG GAA CAA ACA TGA CG – 3'	Tm = 64°C
CASL2_R1	5' – AAT CGA TTC CCC TCG AGA TT – 3'	Tm = 64°C
	In-situ:	
CASL2_isF1	5' – GGT GTG TGG CTT ACA AAA CTT TC – 3'	Tm = 63°C
CASL2_isF1	5' – GAA CAG ACT CAG TTG GAG ATG CT – 3'	Tm = 63°C

TILRc

	Specific primers:	
TL_sF1	5' – GTG ATT TAC CTT ATC AGA AAC AG – 3'	Tm = 56°C
TL_sR1	5' – TAA AAG TGC AAT TCA ATA TTC TAT – 3'	Tm = 56°C
	Northern blotting:	
CT-GFP_TL_F	5' – AAG CGT TGG ATA CCA GTC GT – 3'	Tm = 63°C
CT-GFP_TL_R	5' – TGC TGG TGG CTT CTT TTC TT – 3'	Tm = 65°C
	Protein expression, cloning in pcDNA™3.1/CT-GFP-TOPO®:	
CT-GFP_TL_F	5' – GTT ATG GAC ATC ACC TCC GTC – 3'	Tm = 63°C
CT-GFP_TL_R	5' – TGG AGT TTG CCA ACC CTT C – 3'	Tm = 65°C

B. Methods

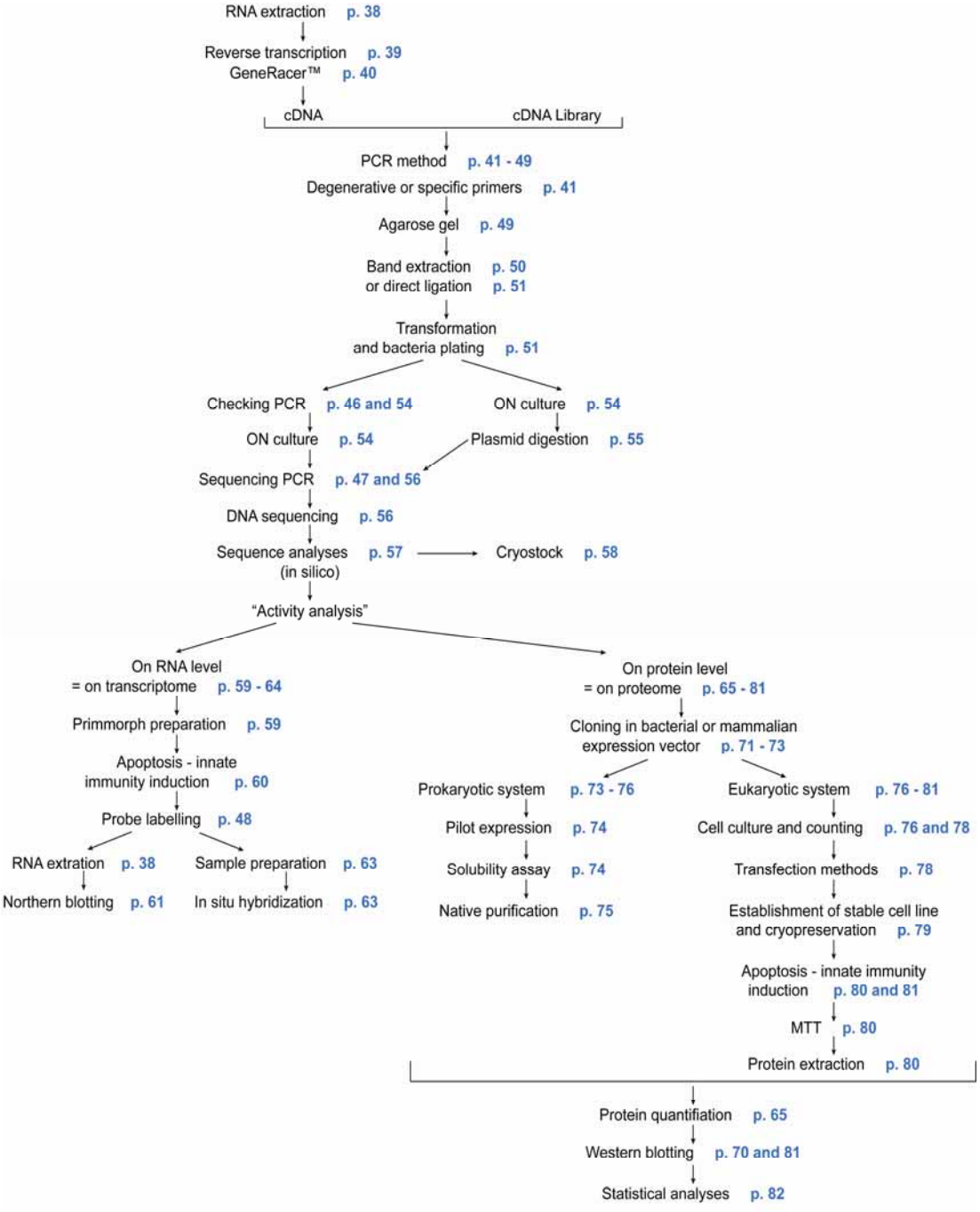


Figure 11: General procedure flowchart.

1. RNA extraction

This isolation, phenol-chloroform separation, is based on differential biochemical properties of DNA, RNA or proteins in aqueous or organic environments.

TRIzol® is a ready-to-use, monophasic solution of phenol and guanidine isothiocyanate, suitable for isolating total RNA, DNA, and proteins. Isolation procedures are based upon improvements to the single-step RNA isolation method developed by Chomczynski and Sacchi (1987). It relies on phase separation upon centrifugation of an upper aqueous phase and a lower organic phase (phenol). Nearly all of the RNA are collected into the aqueous phase (upper and colourless phase), while the DNA and proteins remain respectively in the interphase and the organic phase (lower, red, phenol-chloroform phase). This phase separation was performed in 2 mL phase lock gel tubes. This gel allowed the separation of the aqueous and organic phases with a solid barrier and eases handling. RNA from the aqueous phase was then precipitated with isopropanol and the pellet washed with 75% ethanol. After evaporation of the ethanol, the pellet was re-solubilised in DEPC water. Samples were prepared with Precellys, a bead grinder homogenizer. Manipulation of RNA has to be performed under RNase free conditions. All instruments (*e. g.*, glass vessels, mortar, pestle, and spatulas) were sterilized (autoclaved for one hour at 180°C) or incubate in a SDS solution (1% [w/v]). Potential RNases (in liquids) were inactivated with DEPC (diethyl pyrocarbonate) or DMPC (dimethyl pyrocarbonate). DEPC inactivates RNases by covalent modifications of the histidine residues. DMPC is a safer alternative to DEPC. It reacts with amine, hydroxyl, and thiol groups of RNases inactivating them. DEPC or DMPC was added to solutions (0.1% [v/v]) and inactivated by autoclaving after overnight incubation.

2. Spectrophotometric measurement of concentration and purity of nucleic acids

For quantification of nucleic acids (purines and pyrimidines absorb UV light), readings were taken at 260 nm and 280 nm wavelengths. Absorbance at 260 nm allows estimation

of the nucleic acids in the sample. An OD of 1 corresponds to ≈50 µg/mL for double-stranded DNA, ≈40 µg/mL for single-stranded DNA and RNA, and ≈33 µg/mL for single-stranded oligonucleotides. The ratio between the readings at 260 nm and 280 nm (OD_{260}:OD_{280}) provides an estimate of the purity of the nucleic acid. Pure preparations of DNA and RNA have OD_{260}:OD_{280} values of 1.8 and 2 respectively. Protein or phenol contaminations could foil accurate quantification of nucleic acids, lowering the OD.

- Double-stranded DNA concentration (µg/mL) = 50 × OD_{260}
- Single-stranded RNA or DNA concentration (µg/mL) = 40 × OD_{260}
- Single-stranded oligonucleotides concentration (µg/mL) = 33 × OD_{260}

3. Reverse transcription

Reverse transcriptase is an RNA-dependant DNA polymerase: this enzyme transcribes a single-stranded RNA into a double-stranded DNA. This feature is used by retroviruses and retrotransposons (transposons with RNA intermediates) to integrate their genetic information into host's genome.

Reverse transcriptases used in molecular biology are genetically engineered version. The enzymes generally lack ribonuclease H activity specific to RNA in RNA-DNA hybrids. Therefore, no degradation of RNA occurs during first strand cDNA synthesis, and higher yields of full-length cDNA for long template are obtained.

The type of primer used to synthesize the first strand cDNA has to be determined in accordance with the needs:

- Random hexamer primer (non-specific). In this case, all RNA (*e. g.*, tRNA, mRNA) are templates for cDNA synthesis.
- Oligo $(dT)_{18}$. In this case, only mRNA with 3'- poly(A) tails are templates for cDNA synthesis.
- Sequence/gene-specific primer (GSP): at a specific primer-binding site.

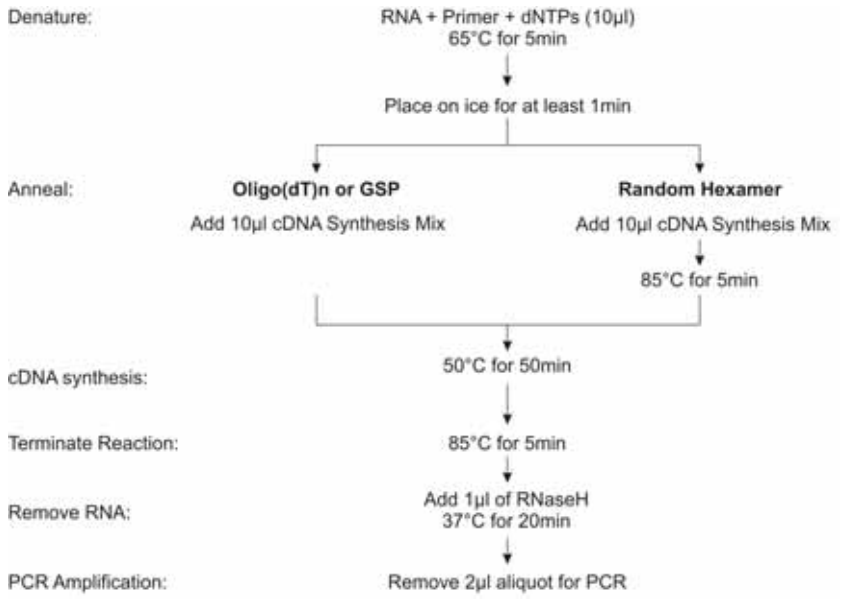


Figure 12: Reverse transcription general procedure flowchart.

Standard conditions for 10 µL cDNA synthesis

10X RT buffer	2 µL
MgCl2 (25 mM)	4 µL
DTT (0.1 M)	2 µL
RNase OUT	1 µL
SuperScript III RT	1 µL
ddH_2O	up to 10 µL

4. GeneRacer™

GeneRacer™ is an advanced rapid amplification of cDNA ends (RACE) technique that improves the efficiency of amplifying full-length, 5'- and 3'- ends of cDNA. The GeneRacer™ kit can be thus used to obtain the complete sequence of a gene (method adapted from Maruyama and Sugano 1994, Schaefer 1995, and Volloch *et al.* 1994).

The advanced protocol starts at the RNA level by specifically targeting only 5'- capped mRNA. In subsequent steps the cap is removed and replaced with the GeneRacer™ RNA Oligo. During reverse transcription, this RNA oligo sequence is incorporated into the cDNA. Only cDNA that is completely reversed transcribed will contain this known sequence. 5'- RACE PCR is then performed using the homologous

GeneRacer™ 5'- primer that is specific to the RNA oligo sequence and a gene-specific primer. The result is amplified DNA that contains the full-length 5'- cDNA sequence.

5. Polymerase chain reaction (PCR)

a. Generation of primers

Primers were ordered from MWG-Biotech (http://www.mwgbiotech. com).They were designed manually and extra checked for hairpin and dimer. To generate gene-specific primers the following rules have been applied:

- length varying from 18 to 25 nucleotides
- Tm (melting temperature) = 4 × (G+C) + 2 × (A+T)

  58°C < Tm < 62°C

  Primer pairs (upstream and downstream primer) should have about the same Tm
- [G+C] content of about 40–60%
- no self-compatibility, no possibility to build primer-dimer nor secondary structure
- no palindrome sequence
- a G or a C should be at the 3´- end of the primer (to ensure a tight binding)

For "basic PCR", 10 pmol/µL working aliquots have been used (working concentration: ≈0.2 pmol/µL).

Degenerative primers were designed on the basis of conserved regions of the target sequences. Regions rich in amino acids that are specified by only one or two codons should be preferred to regions rich in amino-acids with six codons choices (lower degeneracy as possible; fig. 13). For every 1% mismatching of bases in a double stranded DNA, there is a reduction of Tm by 1–1.5%. However, the precise effect of mismatches depends on the [G+C] content of the oligonucleotide and, even more critically, on the distribution of mismatched bases in the double-stranded DNA. Mismatches in the middle of the oligonucleotide are far more deleterious for hybridization than mismatches at the ends. However, a "neutral" base can be used to reduce the degeneracy of oligonucleotides and to

allow the use of more stringent hybridization conditions (*e. g.*, temperature and buffer). The best neutral base is inosine (I). It forms stable base pairs with all four conventional bases, and the strength of the pairing is approximately equal in each case. The working concentration for PCR with degenerative primer should be at least twice higher than with normal primers.

1	2	3	4	6
Met Trp	Phe Tyr His Gln Asn Lys Asp Glu Cys	Ile	Val Pro Thr Ala Gly	Leu Ser Arg

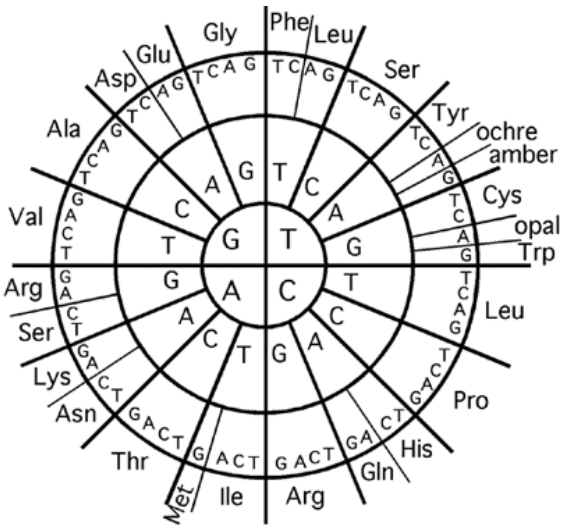


Figure 13: Degeneracy of the genetic code and table of standard genetic code.

Primers for sequencing are labelled with a fluorophore (IRD $_{800}$ or IRD $_{700}$) added to the 5′- end of each primer that emits in the near-infrared region of the spectrum. Sequencing primers' working aliquots of 2 pmol/µL were used

Primers used for cloning in expression vectors were designed following the rules dictated by the manufacturers.

- pEGFP-C2: requires a "restriction sites" or cohesive-end cloning. In this kind of cloning, the vector and the PCR product are digested by restriction enzymes. Cohesive fragments thus generated can be subsequently ligated, allowing the cloning of the insert in a correct position/phase into the vector thanks to compatible overhangs. Several conditions must be checked: *e. g.*, compatibility of restriction enzymes (identical restriction buffer, temperature), unique cleavage site, and absence of restriction site in the sequence to clone. Then, primers were

designed to allow the gene to be cloned into the MCS (multiple cloning site) in frame with the COOH-terminal EGFP, without any stop codon.
The primers contain 2–3 adenosines, the corresponding restriction sites, and 15–20 nucleotides of the gene of interest (5'->3'). The additional adenosines facilitate the positioning of the restriction enzyme.

- pcDNA™3.1/CT-GFP-TOPO® needs the same specifications and the insert should contain a Kozak translation initiation sequence for proper initiation of translation ((G/A)NNATGG, the ATG initiation codon is shown underlined). To express the gene as a recombinant fusion protein, it must be cloned in frame with the COOH-terminal peptide (V5 epitope and polyhistidine tag) without stop codon. As for pTrcHis2 TOPO® TA, the cloning strategy is based on the "TOPO® cloning" method: cloning without the requirement for DNA ligase. No PCR primers containing specific sequences are required, but the DNA must be cloned in frame with the DNA encoding the NH_2-terminal peptide.

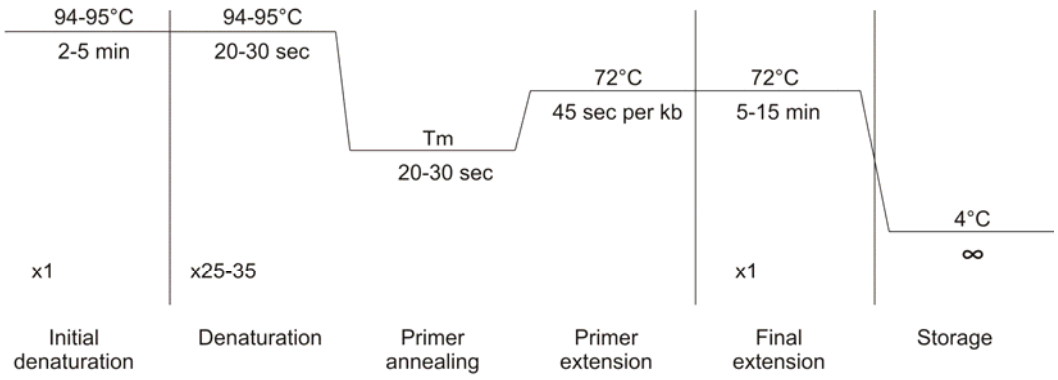
b. Enzymes

DNA-dependant DNA polymerases are used to amplify segments of DNA. DNA polymerase needs a primer and a 3'- OH group to add the first nucleotide and result in elongation of the new strand in a 5'->3' direction. One of the most common or well known enzymes is the Taq polymerase isolated from the bacteria *Thermophilus aquaticus*. This enzyme creates DNA products that have "A" (adenine) overhangs at their 3'- ends. However, the Taq possesses relatively low replication fidelity and lacks a 3'->5' exonuclease proofreading activity (property to correct mistakes in newly-synthesized DNA). Other DNA polymerases possess this 3'->5' exonuclease activity, allowing the incorrect base pair to be excised. Following base excision, the polymerase can re-insert the correct base and replication can continue.

According to the application (standard, high specificity, high fidelity, or difficult template *e. g.*, GC rich) and desired product length, different enzymes can be selected.

Mostly all PCR have been performed using peqGOLD DNA-polymerase (a Taq recombinant DNA polymerase, Peqlab) or PCR master mix (2X) from Fermentas. For more difficult templates or if proofreading/high-fidelity was required, Platinum® Taq DNA polymerase high fidelity (Invitrogen) was used. To generate a 3'- A overhang, an additional step with a Taq DNA polymerase had to be performed before the ligation process.

c. Cycle profile for standard PCR



- Initial denaturation: completes denaturation of the DNA template assuring primer annealing and extension efficiency.
- Denaturation during cycling: allows separation of the DNA dimer formed with the newly synthesized strand.
- Primer annealing: anneals the primers to the single-stranded DNA template. This step is performed at temperature about 3–5 degrees Celsius below the Tm of the primers used.
- Primer extension: DNA polymerase synthesizes a new DNA strand complementary to the DNA template strand by adding dNTP that are complementary to the template in 5'->3' direction (condensing the 5'-phosphate group of the dNTP with the 3'- hydroxyl group at the end of the nascent, extending, DNA strand). A 45 s extension is sufficient for fragments up to 1 kb. In an optimal reaction less than 10 template molecules can be amplified in less than 40 cycles to a product detectable

on gel. Most PCR should include 25–35 cycles. As cycle number increases, non-specific products can accumulate.

- Final extension: promotes the completion of partial extension products and annealing of single-stranded complementary products.

d. Standard PCR

Standard conditions for 50 µl reaction:

10X PCR buffer	5 µL
MgCl2	2.5 mM
dNTP	10 nmol
Forward primer	10 pmol
Reverse primer	10 pmol
Taq-polymerase	1 U
DNA	1 ng
ddH_2O	up to 50 µL

Parameters given above (PCR cycles and standard conditions) have to be optimized for specific needs and template/primer combinations. An example is the touch-down PCR. This PCR variant is used to reduce non-specific background by gradually reducing the annealing temperature. During the first annealing steps, the temperature is slightly (3 to 5°C) above the primer Tm and then decreases to reach few degrees (3 to 5°C) below the last cycles. At high temperature, specificity for primer binding increases while at lower temperatures, amplification of the target sequence improves (Don *et al.* 1991).

In some cases, some reaction additives can enhance the specificity and/or the efficiency of a PCR: betaine (0.5–2 M), bovine serum albumin (BSA; 100 ng per 50 µL reaction), dimethyl sulphoxide (DMSO; 2–10%, [v/v]), glycerol (1–5%, [v/v]), pyrophosphatase (PPP; 0.001–0.1 units/reaction), spermidine, detergents, gelatine, or formamide (2–10%).

e. Checking PCR

Standard conditions for 50 µL reaction / 5 reactions mix:

10X PCR buffer	5 µl
MgCl2	2.5 mM
dNTP	10 nmol
Forward primer	10 pmol
Reverse primer	10 pmol
Taq-polymerase	1 U
DNA/plasmid	5 × 2 µL culture
ddH_2O	up to 50 µL

Checking PCR is one of the methods to inspect if the sequence of interest is correctly inserted in a plasmid after ligation.

After overnight culture of transformed bacteria (on plates, see below), suitable colonies were picked and grown in 12 µL LB medium with the appropriate antibiotic for 30 to 45 min at 37°C. Then, 2 µL of bacterial suspension was added (template) to 8 µL standard PCR reaction (final volume 10 µL).

The selection of primer pairs is orientated by several parameters. If the insert is long, nested primers will be preferred. If the size of the insert has to be verified, vectors-specific primers will be selected. If the orientation of the insert has to be checked, combination of insert-specific and vector-specific primers will be preferred.

After reaction, PCR products were checked on agarose gel, thus only the interesting bands, *i. e.*, only bacteria containing plasmid of interest, were subsequently grown overnight in a larger volume for further investigations.

LB-Medium:

- Add:
  - 10 g/L tryptone
  - 5 g/L yeast extract
  - 5 g/L NaCl
- Adjust pH to 7.0 (with NaOH)
- Autoclave

LB-Agar 1.5%:

- Add:
  - 10 g/L tryptone
  - 5 g/L yeast extract
  - 5 g/L NaCl
  - 15 g/L agar
- Adjust pH to 7.0 (with NaOH)
- Autoclave

f. Sequencing PCR

Plasmids containing insert of interest have been sequenced following the Sanger method (Sanger *et al.* 1977). The Sanger method, or dideoxy sequencing, is the most commonly used method of DNA sequencing. In 1980 Walter Gilbert and Frederick Sanger shared the Nobel Prize in Chemistry for their contribution concerning nucleic acid sequencing. Dideoxy sequencing is based on the same method as basic PCR in 4 separate reactions, each reaction differing in one modified nucleotide, *i. e.*, dideoxynucleotide (ddA, ddT, ddG or ddC), representing only 1% of the total pool of nucleotides. A dideoxynucleotide differs from a deoxynucleotide in that it has a 3'- H rather than a 3'- OH on the deoxyribose, preventing the formation of a phosphodiester bond with an incoming DNA nucleotide thus terminating the DNA synthesis. For a population of DNA molecules in a reaction tube, new DNA fragments will stop at all possible positions where the particular dideoxynucleotide is incorporated, generating all possible lengths.

To perform this reaction, reagents from GE Healthcare were used (ready to use mix, one per tube). The A, C, G, and T reagents contain 40 mM Tris-HCl, pH 9.5, 6 mM MgCl2, 0.04% Tween™-20, 0.04% Triton™ X-100, 0.08 mM β-mercaptoethanol, 1.67 units/µL thermo sequenase DNA polymerase with 0.00028 units/µL *Thermoplasma acidophilum* inorganic pyrophosphatase, 0.4 mM dATP, 0.4 mM dCTP, 0.4 mM 7-deaza-dGTP, 0.4 mM dTTP in addition to the specific dideoxynucleotide terminator, and one of the specific ddNTPs (*e. g.*, ddATP for A reagent, 1.33 µg/µL).

The sequencing procedure can be divided in four steps:

<u>1- Sequencing reaction, Master Mix for one sample</u>

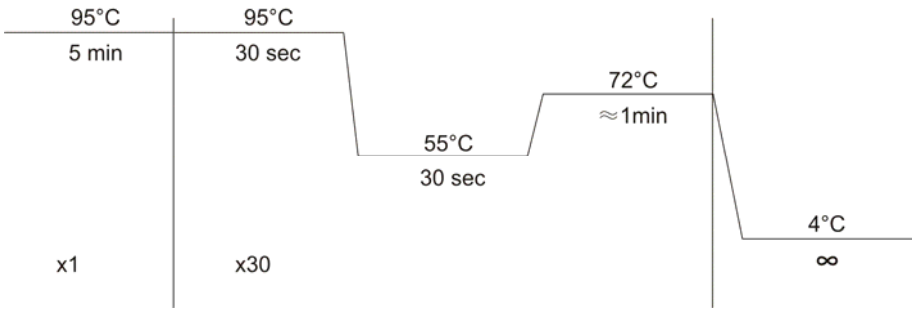
Template	x µL (1.5 µg)
Dye-primer (IRD)	2 pmol
ddH_2O	up to 18.75 µL

<u>2- Preparation of reactions</u>

For each template to be sequenced, four tubes were labelled "A", "C", "G", and "T" and placed on ice. Then, 1.5 µL of A reagent was dispensed into each tube labelled "A".

This process has to be repeated for the "C", "G", and "T" tubes. Subsequently, 4.5 µL of the master mix from step 1 was distributed into each of the four labelled tubes and mixed thoroughly.

3- Cycling parameters



4- Preparation of sequencing samples for loading

Prior loading, a final denaturation step, 95°C for 3 minutes, was performed, immediately placed on ice, and stopped with 3 µL of "stop solution".

Stop solution:

- Add:
  - 95% formamide [v/v]
  - 10 mM EDTA
  - 0.1% xylencyanol [w/v]
  - 0.1% bromophenol blue [w/v]

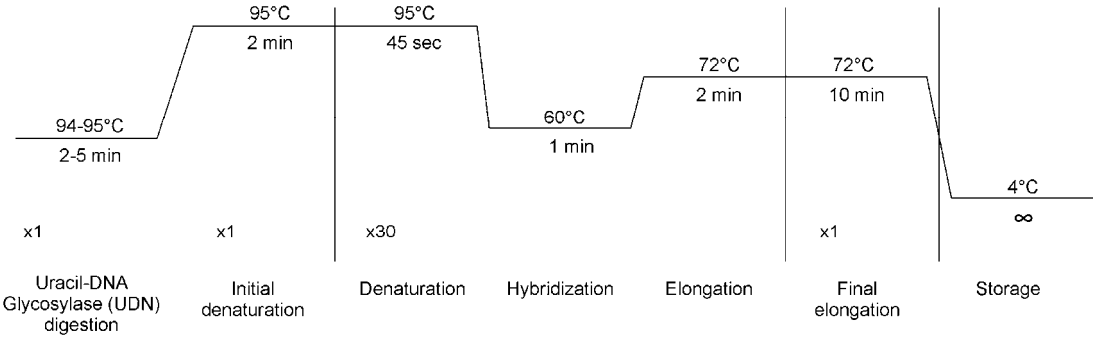
g. DIG labelling PCR

The DIG system is an effective system from Roche for nonradioactive labelling and detection of nucleic acids. Several kits/methods are available for labelling DNA, RNA, or oligonucleotides. However, DNA-DIG labelled probes have been preferred for their stability. Thereby, DNA probes have been generated by PCR using the PCR DIG labelling mix^{PLUS} (Roche), where dNTP have been substituted by a mix of dATP, dCTP, dGTP (each 2 mM), 5.7 mM dUTP, and 0.3 mM DIG-dUTP. DNA probes thus obtained where then cleaned with Roche high pure PCR product purification kit and controlled on a gel. DNA labelled probes must be boiled prior use (10 min at 95°C then cooled on ice).

DNA probes can be used for all types of filter hybridization, from Northern blotting to *in situ* hybridization. Sizes of the probes should encompass ≈500 and ≈250–350 nucleotides for Northern blotting and *in situ* hybridization, respectively.

A first PCR (standard PCR) was generally achieved (with a known template *e. g.*, plasmid) to check the unique resulting band. Then, the first PCR product was used as template to a second PCR performed with the PCR DIG labelling mix^{PLUS}. The probes for Northern blotting are double stranded DNA (second PCR with primer pair) whereas they are single stranded for *in situ* hybridization (PCR executed only with one primer). Accumulation of the PCR product is linear in opposition to normal PCR which is exponential. The probe resulting of the PCR with the reverse primer should not (control).

Cycling parameters for DNA-DIG labelled probe



6. Agarose gel electrophoresis

Negative charges carried by the sugar-phosphate backbone of nucleic acids allow their migration and their size separation on standard agarose gel electrophoresis (electric field; Sambrook *et al.* 1989). However, other factors than length can affect the migration *e. g.*, conformation of the DNA molecule (circular, linearized, or supercoiled plasmid), agarose concentration (low concentration better resolution for big fragments and vice versa for small fragments), voltage, and buffer composition (TAE is preferred to TBE for small fragments migration).

PCR reactions were mixed with loading dye to enable the visualization of their migration. Xylene cyanol and bromophenol blue are common color markers used in loading buffer; they run about the same speed as 5000 bp and 300 bp DNA (respectively) but the precise position can vary with the agarose percentage. DNA molecular weight markers were used to calculate the length of DNA fragments. Gels (usually 1% agarose, 1X TBE) were run at 90 V. After migration, gels were stained with ethidium bromide (EtBr). EtBr is an intercalating agent (between base pairs) When exposed to ultraviolet light, it fluoresces with an orange color, intensifying almost 20-fold after binding to nucleic acids. Therefore, distinct bands of DNA were visible and were documented with digital camera.

<u>6X loading buffer (DNA):</u>

- Add:
  - 40% saccharose [w/v]
  - 0.25% bromophenol blue [w/v]

<u>10X TAE-buffer:</u>

- Add:
  - 0.4 M Tris
  - 0.2 M acetic acid
  - 0.01 M EDTA

<u>10X TBE-buffer:</u>

- Add:
  - 90 mM Tris-HCl pH 8.3
  - 90 mM boric acid
  - 2 mM EDTA

7. Isolation of DNA from agarose gel

To concentrate diluted nucleic acid solutions, to have suitable DNA for downstream processing, or to remove contaminants (salts and soluble macromolecular components), bands can be isolated from agarose gel.

After short exposition of the gel in a fresh EtBr bath, DNA band was carefully excised under UV illuminator. The DNA was purified with NucleoSpin® extract II kit (Macherey-Nagel) or high pure PCR product purification kit (Roche). The sample was then mixed with a chaotropic salt and applied to the glass fibre fleece in filter tube. Under the ethanolic buffer conditions used in the procedure, all nucleic acids in the sample bind to the glass fleece, while contaminating substances (salts, proteins, nucleotides, mineral oil, and other contaminants) were not. Brief wash and spin steps readily removed these

contaminants. Once purified, the DNA was easily eluted in a small volume of low salt buffer.

8. Cloning

According to the employed vector and/or to the polymerase used, several methods exist to clone the PCR-amplified DNA molecules, including cohesive-end cloning (introduction of restriction sites at the 5'- ends of PCR primers, *e. g.*, pEGFP-C2), blunt-end cloning (*e. g.*, pJET1), ligation-independent cloning (TOPO® TA technology), and TA cloning (*e. g.*, pGEMT). TA ligation (with pGEMT) is the most applied technique to clone PCR product. The TA cloning strategy is both simple and much more efficient than blunt-ended ligation for the cloning of PCR products. The single 3′- T overhangs at the insertion site greatly improve the efficiency of ligation, preventing recircularization of the vector, and providing compatible overhangs for PCR products.

The pGEMT ligation was performed over night at 4°C or for 2 hours at room temperature. For most of the vectors from TOPO TA technology, the ligation was efficient after few minutes at room temperature.

9. Transformation of competent cells

Transformation is defined as the transfer of genetic information from a donor to a recipient using naked DNA. Here the recipient is bacteria and the DNA a recombinant plasmid. The endogenous genetic information contained in the plasmid allows the plasmid to replicate by using the host cell's enzymes. Plasmid replication and host cell's multiplication greatly amplify the number of the recombinant plasmids.

Transformation is efficient if the bacteria are said "competent": if they are able to uptake extracellular ("naked") DNA from their environment. Most of the time competent bacteria were purchased from companies, but this ability can be induced by different methods, such as the Inoue method (see below).

Several transformation methods exist; however, introduction of the foreign DNA was completed by thermal destabilisation of the membrane (*i. e.*, "heat-shock"). During the first step, an aliquot of competent bacteria was thawed, DNA added (generally not more than 10% of the bacteria volume), and incubated on ice. The incubation time generally did not exceed 5 min (to allow the adsorption of DNA around the bacteria membrane). The cells were then heat shocked at 42°C for 30 s (the plasmid entered the bacteria) and placed on ice for 2 minutes. LB or SOC medium (250 µL) was added to the cells and bacteria were incubated at 37°C for 1 h (to allow the bacteria to recover and to express the antibiotic resistance gene). Finally 50–200 µL of the cells were plated on plates (containing the appropriate antibiotic in combination or not with IPTG/X-gal) and incubated over night at 37°C.

SOB medium (1L):

- Add:
  - 20 g bacto tryptone
  - 5 g yeast extract
  - 0.5 g NaCl
  - 10 mL of a 250 mM KCl solution
- Adjust pH to 7.0 (with NaOH)
- Autoclave
- Just before use, add 5 mL of a sterile solution of 2 M $MgCl_2$

TB:

- Add:
  - 10 mM pipes
  - 55 mM $MnCl_2$
  - 15 mM $CaCl_2$
  - 250 mM KCl
- Adjust pH to 6.7 (with KOH)
- Autoclave

SOC medium: identical to SOB medium, except that it contains 20 mM glucose.

a. Preparation of "ultra-competent" *E. coli*: the Inoue method (adapted from Inoue H *et al.* 1990)

Preparing competent *E. coli* is an alternative to commercial bacteria. The bacterial culture is grown at 18°C rather than the conventional 37°C (modifications of physical characteristics of bacterial membranes and extension of efficient transformation phases). 250 mL of SOB medium was inoculated with an overnight culture (grown in LB) and grown until OD_{600} = 0.6 (18°C, 250 rpm). Then the culture was chilled on ice for 10 minutes and

subsequently spun at 2,500 g for 10 min. at 4°C. The pellet was resuspended gently in 80 mL of ice cold TB medium and incubated on ice for 10 minutes. The mixture was then centrifuged at 2,500 g for 10 min. at 4°C. The pellet was again resolubilized gently in 20 mL of ice cold TB and DMSO was added to a final concentration of 7%. After 10 min on ice, the bacteria were aliquoted (50 µL), snap-frozen in liquid nitrogen, and stored at -80°C.

i. Monitoring of the bacterial growth via optic density (OD)

Bacterial growth can be divided in 4 phases (LAG, LOG, STATIONARY, and DEATH phases) and be monitored over time via culture turbidity (OD_{600}). When the culture reaches an absorbance of 0.5–0.7, the growth is optimal (in the logarithmic phase).

10. Selection of recombinant vector of interest

a. Antibiotic resistance and white/blue selection

All vectors which were used possessed a specific antibiotic resistance gene (positive selection) conjugated or not with a negative selection (lacZ gene).

When the vectors have this dual selection and the bacteria (*e. g.*, TOP 10 and NovaBlue singles™ competent cells) are plated on specific antibiotic supplemented agar, thus only the cells which have received the vector can survive. The negative selection involves the lacZ gene. This gene encodes the β-galactosidase (β-gal), an intracellular enzyme that cleaves the disaccharide lactose into glucose and galactose. This enzyme can also cleave other substrates, such as the galactoside 5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside, more usually named BCIG or X-gal. When X-gal is hydrolyzed the resulting compound is oxidized into an insoluble blue product. Thus, if X-gal and an inducer of β-galactosidase (usually IPTG which induces activity of β-galactosidase by binding and inhibiting the lac repressor) are contained within an agar medium on a culture

plate, colonies which have a functional *lacZ* gene can easily be distinguished. Usually, the gene is positioned in the MCS, resulting to the gene inactivation in case of efficient ligation. Only colourless colonies receive recombinant plasmids. However, other vectors (generally expression vectors) just provide an antibiotic resistance. In this case all colonies have to be examined.

After this simple or double selection, only the plasmids which contain the insert of interest have to be singled out.

> Ampicillin can be substitute by carbenicillin; both are used for systemic use, causing cell wall disruption (gram-negative coverage but gram-positive limited one). However, carbenicillin is more stable than ampicillin and was preferred.

LB/Amp/IPTG/X-gal plates:

- Add to 1 L LB:
  - 1.5% agar [w/v]
- Autoclave
- Around 50°C, add:
  - 0.01% antibiotic [w/v]
  - 0.5 mM IPTG
  - 0.02% X-gal [v/v]

X-gal stock solution:

- 50% X-gal [w/v] in N, N'-dimethylformamid

b. Checking PCR

Checking PCR is one of the methods to inspect if the sequence of interest is correctly inserted (see the checking PCR section) and will lead to the selection of some colonies for overnight cultures.

c. Overnight cultures

Volume of the overnight cultures was chosen in accordance with the needs (mini-, midi-, or maxi-prep). The cultures were then placed for 18 to 24 hours at 37°C under vigorous agitation (200 rpm).

d. Plasmid digestion

Another possibility to inspect clones of interest is to check the size of the insert after plasmid digestion. Several unique restriction sites are present in the MCS allowing excision of the insert and linearization of the vector. Sizes of the inserts are directly visible after electrophoresis on agarose gel. The reaction was let 1 hour at 37°C to allow complete plasmid digestion. The digestion reaction was terminated via thermal inactivation (65°C or 80°C for 20 min). The reaction was then loaded on agarose gel to check the size of the insert.

Restriction reaction mix:

10x buffer	2 µL
Plasmid elution	5 µL
Restriction enzymes	2 × 0.5 µL
ddH_2O	12 µL

11. Isolation of plasmid DNA

a. Isolation of plasmid DNA (mini-, midi-, and maxi-scale)

Plasmids DNA were isolated from bacterial cultures using high pure plasmid isolation kit from Roche or NucleoSpin® plasmid kit from Macherey-Nagel. Bacteria cultures (0.5 to 4.0 mL) were grown for 12 to 16 h. The procedures were followed as outlined in the manufacturer's protocol. Those extractions rely on an alkaline lysis to free the plasmid DNA from the cell, leaving behind the *E. coli* chromosomal DNA trapped in the cell wall debris. After the solution is cleared of cell debris and chromosomal DNA, the supernatant is retained and passed to the spin filter tube. The nucleic acid binds specifically to the surface of glass fibres in the presence of chaotropic salt (guanidine HCl). Since the binding process is specific for nucleic acids, the bound plasmid DNA is purified from salts, proteins, and other cellular impurities by washing steps followed by elution in low-salt buffer or water.

For larger amount of plasmid DNA, Qiagen plasmid midi kit (25 to 100 mL of bacteria culture) or maxi kit (100 to 500 mL of bacteria culture) were employed. This protocol is also based on a modified alkaline lysis procedure, followed by binding of plasmid DNA on an anion-exchange resin.

b. Boiling method for mini-scale samples

This protocol is an alternative method for the preparation of bacterial plasmids (adapted from Holmes and Quigley 1981). Bacteria from a 2.5–4 mL ON culture were pelleted and redissolved in cold STET (200 µL + 8 µL lysozyme). The cells were then lysed for 50 to 90 s at 95°C. After centrifugation the pellet containing genomic DNA and cell debris was removed with a wooden stick. To prevent RNA degradation, the supernatant was treated with RNase A (1/10 vol.) for 1 h at 37°C. Afterwards, the proteins were precipitated via a phenol/chloroform procedure. The DNA were then precipitated with isopropanol and dissolved in 10 to 200 µL ddH_2O

STET:	Lysozyme:	RNase A:
o Add: - 0.1 µL NaCl - 10 mM Tris-Cl (pH 8.0) - 1 mM EDTA (pH 8.0) - 5% Triton X-100	- Stock 20 mg/mL	- Stock 1 mg/mL

12. DNA sequencing

Once the plasmids have been isolated, sequencing PCR is performed (see sequencing PCR section). Sequencing PCR shows two main differences compared to the standard PCR reaction. Firstly, only one primer is used. Thus the amplification is linear and not exponential. Secondly, incorporation of dideoxynucleotide terminates DNA strand synthesis, resulting in infrared dye (IRD) labelled DNA fragments of varying lengths. These generated fragments are then separated via a polyacrylamide gel electrophoresis and

detected by an infrared sensor (as infrared dyes migrate past the detector window they are both excited by a laser diode and detected). The data are collected in real time and create a raw image data (series of bands). The Li-Cor system can detect IRD $_{700}$ and IRD $_{800}$ dyes. The software creates two independent image files from the same gel. One file contains image information from the 700 channel and the other file contains image information from the 800 channel. The software (e-Seq V2.0 DNA) analyzes the image by finding the lanes for each sample and calling bases. The sequence data can be presented in text (.txt) or in standard chromatogram format (.scf).

13. Sequence analysis (*in silico*)

Analysis and comparison of DNA sequences were performed using Lasergene processing software (DNAStar V6), DNASIS V5 or Vector NTI Suite. Homology searches were performed via the servers at the European Bioinformatics Institute; Hinxton, United Kingdom (EBI; http://www.ebi.ac.uk/inc/head.html) and the National Center for Biotechnology Information (NCBI; Bethesda, MD (http://www.ncbi.nlm.nih.gov/BLAST/)). Multiple alignments were carried out with ClustalW version 1.6 (Thompson *et al.* 1994).

5'- UTR regions were analyzed with various freeware: TESS (Schug 2008; www.cbil.upenn.edu/cgi-bin/tess/tess?), TFSEARCH (Heinemeyer *et al.* 1998; www.cbrc.jp/research/db/TFSEARCH), and WebGene (TATA Signal Prediction in Eukaryotic Genes; Milanesi 1999; www.itb.cnr.it/sun/webgene/). CpG islands were determined using CpGProD (CpG Island Promoter Detection; Ponger and Mouchiroud 2001; http://pbil.univ-lyon1.fr/software/cpgprod_query.html). The coiled-coil conformation was studied via the freeware COILS (Lupas *et al.* 1991; http://www.ch.embnet.org/software/COILS_form.html), and transmembrane helical-protein topology predicted via TUPS (Zhou and Zhou 2003; http://sparks.informatics.iupui.edu/Softwares-Services_files/tups.htm) and PSORTII (Nakai and Horton 1999; http://mobyle.pasteur.fr/cgi-bin/MobylePortal/portal.py?form=psort

Phylogenetic trees were constructed on the basis of amino acid sequence alignments applying the Neighbor-Joining method to the distance matrices that were calculated using the Dayhoff PAM matrix model (Dayhoff *et al.* 1978; Saitou and Nei 1987). The degree of support for internal branches was further assessed by bootstrapping (Felsenstein 1993). The graphical output of the bootstrap figures was produced through the "Treeview" software (R.D.M. Page; University of Glasgow, United Kingdom; http://taxonomy.zoology.gla.ac.uk/rod/treeview.html). Further graphic presentations were prepared with GeneDoc (Nicholas and Nicholas 1997). Potential subunits, domains, patterns, and transmembrane regions were predicted after searching the Pfam database (Finn *et al.* 2006; www.sanger.ac.uk/Software/Pfam/), the SMART database (Letunic *et al.* 2006; http://smart.embl-heidelberg.de/), the ExPASy database (http://www.expasy.org/tools/protparam.html) or according to Kyte and Doolittle (1982) and the ELM database (Emmert *et al.* 1994; http://elm.eu.org/).

For the exhaustive list of softwares used, see section computer and online softwares.

14. Cryopreservation

For long term preservation of bacteria containing a plasmid of interest (-80°C), cells were conserved in glycerol-containing buffer. Glycerol provides protection from freezing and thawing by acting colligatively to reduce the high salt concentration that occurs during freezing. Bacteria suspensions were mixed with the 20% sterile-filtered cryopreservation buffer and snap-frozen in liquid nitrogen before storage at -80°C.

<u>Cryopreserver (glycerol-containing) buffer:</u>

- 25 mM Tris-HCl, pH 8.0
- 100 mM MgSO4
- 65% glycerol [v/v]

After analysing the sequence of interest, several options are available to further study the new molecule. Roughly, two approaches can be distinguished: analysis on transcriptome (RNA) or on proteome (protein). Both will be developed below.

15. Analysis on transcriptome

a. Primmorphs preparation (adapted from Custodio MR *et al.* 1998)

This method is a soft dissociation of sponge tissues. EDTA is widely used to sequester di- and trivalent metal ions. Ca^{2+} and Mg^{2+}, thus sequestrated, are not able to maintain the matrix stability (the collagenase activity require Ca^{2+}), leading to the rupture of intercellular matrices.

Figure 14: Primmorphs preparation and formation. Small pieces of sponge (first left picture) were dissociated and singular cells washed. After few days, aggregates or prim morphs, consisting of proliferating cells, get their nice round shape (right picture).

Small pieces of sponge, (cut in ≈1 mm^3 cubes), were transferred into 50 mL conical tubes with 40 mL of calcium- and magnesium-free synthetic sea-water (CMFSW)+EDTA, and shaken for 30 to 60 min on a rotary shaker (gentle shaking at 16°C, dissociation phase). The mixture was filtered through 40 µm mesh nylon net. The cells were then harvested by centrifugation (500 g for 5 min) and washed twice in CMFSW. After the last centrifugation, the pellet was resuspended in the adequate volume of natural sterile-filtered seawater (Sigma-Aldrich), supplemented with 0.2% [v/v] of RPMI-1640 medium (used as nutrient source), 50 mg/L gentamicin (an aminoglycoside that inhibits bacterial protein synthesis), and 60 µM silicate. Two-thirds of the culture medium was replaced each day during the first week; later the medium was changed only once or twice a week. As soon as the primmorphs get their nice round shape or aggregates, consisting of proliferating cells (minimum one week with shaking) the wanted experiment can be performed. In this study, two sets of experiments were performed. For the first set, primmorphs were exposed to cadmium ($CdCl_2$, 50 µM) or the synthetic lipopeptide (S)-[2,3-Bis(palmitoyloxy)-(2-RS)-

propyl]-N-palmitoyl-(R)-Cys-(S)-Ser-(S)-Lys4-OH·3HCl (Pam_3Cys-Ser-$(Lys)_4$; 40 µg/mL) for 6 and 9 h, or they remained untreated as a control. For the second set of experiments, primmorph samples were taken 0 (control), 7, 14, and 21 d after aggregation.

Total RNA was then extracted and Northern blotting performed.

Calcium Magnesium Free Sea Water (CMFSW, ± EDTA):

- 400 mM NaCl
- 7 mM Na_2SO_4
- 10 mM KCl
- 10 mM HEPES
- ± 2.5 mM EDTA

Solutions were filtrated through 0.2 µm polycarbonate filters

b. Apoptosis and innate immunity induction for transcriptome analysis

In order to prove the involvement of a gene in a physiological process or to follow its expression under special conditions, exposures of sponge or primmorphs to stressors were performed.

For one incubation, an alive sponge was cut in pieces, one piece was directly frozen in liquid nitrogen (control) and the rest was placed in a 50 mL Falcon with filtered, ventilated see water, and the chosen toxic, at 16°C. For a given experiment, it is important to take samples from the same sponge in order to decrease the impact of intra-species changes (phenotypic plasticity/epigenetic variations). For the time-scale experiment, samples were removed from the same sponge, cut in small pieces, and snap-frozen in liquid nitrogen. The following stress exposures were carried out:

- Cadmium exposure: 50 µM for up to 12 h (6 h, 9 h, and 12 h)
- Pam3Cys-Ser-(Lys)4 exposure: 40 µg/mL for up to 12 h (6 h, 9 h, and 12 h)
- Lipopolysaccharide exposure: 40 µg/mL for up to 72 h (6 h, 12 h, 24 h, 48 h, and 72 h)

After exposures, total RNA was extracted and Northern blottings performed.

c. Northern blotting

Figure 15: *Digitalis purpurea*

During Northern blot, total RNA is electrophoretically separated and transferred on a membrane (Alwine *et al.* 1977). After blotting the transcript size can be determined and/or targeted gene expression monitored, via hybridization with digoxigenin (DIG) labelled DNA probe. The DIG (DIG-dUTP) is incorporated during the probe PCR (see the DIG-labelling PCR section). Digoxigenin is a steroid found exclusively in *Digitalis purpurea* and *Digitalis lanata*. For detection of the hybridization signal on the nucleic acid blot the antibody (antidigoxigenin-alkaline phosphatase)/probe hybrids are visualized with chemiluminescent alkaline phosphatase substrates (CDP-Star alkaline phosphatase substrate). More sensitive chemiluminescent substrates produce light that can be conveniently recorded with X-ray film.

During the first day, after RNA denaturation (65°C for 15 min), 5 µg of total RNA was size-separated (gel ran at 60 to 100 V). After several washings (two times 15 min with DEPC H_2O and one time 5 min with 20X saline-sodium citrate (SSC) buffer) the gel was placed in a PCR tank (opening pockets down), on a sheet of Whatman paper (soaking in SSC buffer as both ends). On the gel was placed the adequate sized positively charged nylon membrane and a stack (8 to 10 cm) of lab papers (paper towels) surmounting by a weight. RNA was thus blotted on Hybond N1 membranes over night thanks to capillary forces.

During the second day, the membrane was dried and the RNA fixed by UV radiation (1200 J). The membrane was pre-hybridized with DIG easy hyb (1 mL per 5 cm^2) for 2 h at 42°C. During this time the DNA probe was denaturized (10 min at 95°C, then directly cooled on ice) and subsequently added to the pre-hybridization solution (2 µL

probe per mL of DIG Easy Hyb) to perform over night hybridization at 42°C. Hybridization was performed with a *S. domuncula* survivin–like probe (*SDSURVL*, nt_{1-357}), a caspase–like protease 1 probe (*SDCASL*, nt_{49-822}) and a caspase–like protease 2 probe (*SDCASL2*, $nt_{810-1480}$). Furthermore, a probe was designed to detect the expression of the housekeeping gene ß-tubulin (*SDTUB*, nt_{83-483}; NCBI accession number AJ550806) as internal reference.

During the third day, after low stringency washings, (2 times 15 min at room temperature on shaker) and high stringency washings (2 times 15 min at 42°C with shaking) the membrane was immersed in the blocking solution (P2) at room temperature for 30 min under gentle shaking. This step was directly followed by the incubation with antidigoxigenin antibody (1/10,000 in blocking solution) for 30 min at room temperature under gentle shaking. After this incubation, excess of antibody was washed (1 time 15 min washing with P1 complemented with 0.1% Tween 20 [v/v] and 2 times 15 min washing with P1, at room temperature under vigorous shaking). Afterwards, the alkaline phosphatase was activated by incubation in P3 (5 min). Excess of P3 was removed, CDP star added and incubated with the membrane for approximately 3 min. The membrane was then placed between two plastic sheets, and the CDP star excess removed. Ultimately the hybridized probes were visualised on a radiography film.

Low stringency wash buffer:

- Add:
  - 2X SSC
  - 0.1% SDS

High stringency wash buffer:

- Add:
  - 0.5X SSC
  - 0.1% SDS

20X SSC:

- Add:
  - 3 M NaCl
  - 0.3 M sodium citrate
- Adjust pH to 7.0
- Autoclave

10X MOPS:

- Add:
  - 0.2 M MOPS
  - 0.05 M Na acetate
  - 0.01 M EDTA
- Adjust pH to 7.0
- Autoclave

P1:

- Add:
  - 1 M maleic acid
  - 1.5 M NaCl
- Adjust pH to 7.0
- Autoclave

P2 (blocking buffer):

- Add:
  - 10% blocking-reagent [w/v] in P1

P3:

- Add:
  - 0.1 M Tris-HCl
  - 0.1 M NaCl
- Adjust pH to 9.5
- Autoclave

RNA agarose gel (1%, 140 mL):

- Add:
  - 1,4 g agarose
  - 14 mL MOPS
  - 118.5 mL DEPC H_2O
- Around 65°C add 7.5 mL formaldehyde (acts as a denaturant)

6X loading buffer:

- Add:
  - 250 µL formamide
  - 83 µL formaldehyde
  - 50 µL 10X MOPS
  - 50 µL glycerol
  - 117 µL DEPC H_2O
  - few bromophenol blue crystals

Water: Always use autoclaved, DMDC- or DEPC-treated, (1 mL/1 L ddH_2O)

All solutions are RNase free, thus treated overnight with DMDC or DEPC and then autoclaved.

d. *In situ* hybridization

The *in situ* hybridization method applied is based on a procedure described by Polak and McGee (1998) with modifications recently described (Perovic´ *et al.* 2003).

In short, frozen sections (8 µm) of primmorph samples were fixed in 4% paraformaldehyde, then incubated with proteinase K (1 µg/mL), and fixed again in paraformaldehyde. Hybridization was performed overnight at 45°C in 2X SSC sodium chloride/sodium citrate (SSC supplemented with 50% formamide), either with sense (control) or antisense probes. The single-stranded *SDSURVL* DNA probes (236 nt; forward primer 5' – TAT GAG AGT TGT GAC AGA GTA AA – 3'; reverse primer 5' – TGT TCA TCA CGG GGA TTG TCT GA – 3') had been labelled with the PCR-DIG-probe synthesis kit, using a part of *SDSURVL* as linearized template. In a similar way *SDCASL2* probes were obtained (292 nt; forward primer 5' – GGT GTG TGG CTT ACA AAA CTT TC – 3'; reverse primer 5' – GAA CAG ACT CAG TTG GAG ATG CT – 3').

Following hybridization the sections were washed at 50°C at decreasing salt concentrations (1X SSC to 0.2X SSC), blocked, and then incubated with anti-DIG Fab fragments, conjugated to alkaline phosphatase. Through addition of NBT and BCIP, hybridized probes were visualized with an Olympus light microscope AH3.

BSA stock:

- Add:
  - 10% [w/v] albumin bovine, fraction V in ddH_2O

NBT stock:

- Add:
  - 75 mg/mL in 70% [v/v] DMF

Blocking solution:

- Add:
  - PTW with 2% [w/v] Blocking Reagent

PBT (in1X PBS, 100 mL):

- Add:
  - 0.1% [v/v] Triton X-100
  - 2 mL 10% BSA

PFA:

- Add:
  - 4% paraformaldehyde in 1X PBS or CMFSW
- Adjust to pH 7.4-7.5

NBT/X-phosphate:

- Add:
  - 87.5 µL NBT
  - 112.5 µL X-phosphate
- Adjust to 25 mL with 1X P3

Proteinase K:

- Stock 20 mg/mL in 1X PBS
- Working concentration 1 mg/mL

X-Phosphate:

- Add:
  - 50 mg/mL in 100% DMF

PTW:

- Mix:
  - 0.1% [v/v] Tween 20 in 1X PBS

1X P3:

- Add:
  - 100 mM Tris/HCl pH 9.5
  - 100 mM NaCl
  - 50 mM $MgCl_2$

10X PBS-stock:

- Add:
  - 136 mM NaCl
  - 2.6 mM KCl
  - 6.46 mM $Na_2HPO_42H_20$
  - 1.46 mM KH_2PO_4
- Adjust to pH 7.2

20X SSC:

- Add:
  - 3 M NaCl
  - 0.3 M Sodium citrate
- Adjust to pH 7.0

16. Analysis on proteome

a. Protein quantification

Several methods (OD measurement, Lowry, bicinchoninic acid, Bradford or Biuret methods) are available to measure protein concentrations. However Pierce® BCA protein assay kit (based on bicinchoninic acid or BCA) was mostly used and will be detailed.

i. Assay according to bicinchoninic acid (BCA; adapted from Smith PK *et al.* 1987)

This method combines the well-known reduction of Cu^{++} to Cu^{+} by the protein in an alkaline medium (the Biuret reaction) with the highly sensitive and selective colorimetric detection of the cuprous cation (Cu^{+}) using a unique reagent containing bicinchoninic acid. The purple-coloured reaction product of this assay is formed by the chelation of two molecules of BCA with one cuprous ion. This water-soluble complex exhibits a strong absorbance at 562 nm that is nearly linear with increasing protein concentrations over a broad working range (20–2,000 µg/mL). Protein concentrations are generally determined and reported with reference to standards of a common protein, such as bovine serum albumin (BSA).

Working reagent was prepared by mixing 50 parts of BCA reagent A with 1 part of BCA reagent B. 1 mL of this mix was added to 50 µL of the sample (1/10 diluted) and was incubated at 60°C for 30 min. Then 100 µL of each samples were placed in 96-well plate and absorbances were measured in duplicate at 562 nm.

To correlate the measured absorbances to protein concentration, a linear regression was calculated with absorbances of known concentrations of BSA (BSA calibration curve; OD_{595} = f (concentration in µg/µL), see below).

ii. Preparation of diluted albumin (BSA) standards

From a BSA stock solution (2 mg/mL), dilute 1/2 in ddH_2O, C_f = 1 mg/mL and:

BSA (µL)	0	5	10	15	20	25
ddH_2O (µL)	50	45	40	35	30	25
Concentration (µg/mL)	0	100	200	300	400	500

iii. Experion™ Pro260

Experion™ is an automated electrophoresis system which allows analysis of protein and RNA samples. The Experion system combines electrophoresis, staining, destaining, detection, basic analysis and digital result documentation into a single automated step. The Experion technology is based on microfluidic chips which contain a series of plastic wells bonded over a small glass plate. The glass plate is etched with an optimized network of micro-channels, one of which intersects with each plastic wells. Once these channels have been primed with a matrix and the samples applied to the appropriate wells, the electrophoresis station directs the samples through these micro-channels by specifically controlling the voltages and currents that are applied.

The Experion™ Pro260 analysis kit can separate and quantify protein samples ranging from 10 to 260 kDa in mass. The assay automatically determines the relative concentration of protein samples using a single-point calibration, in other words, the peak area of the protein of interest is compared to the peak area of a 260 kDa internal upper marker, which is present in each sample at a known concentration. Additionally, the user has the option to obtain an absolute protein concentration by using known concentrations of a purified protein to create a calibration curve on the chip. Experion software generated an electropherogram, simulated gel image and results table for each sample.

b. Polyacrylamide gel electrophoresis (PAGE)

i. Denaturing PAGE (SDS-PAGE)

SDS (sodium dodecyl sulphate) is an anionic detergent which denatures secondary and non-disulfide-linked tertiary structures, and applies a negative charge to each protein in proportion to its mass. Consequently, the distance of migration through the gel is directly related to the size of the protein (Weber and Osborn 1969).

Both agarose and acrylamide can be used to separate DNA and protein. Agarose gel is much cheaper and easier to prepare compared to acrylamide gel; however, the higher resolution is obtained with acrylamide gel.

Polyacrylamide gel is a matrix formed by copolymerisation of monomers of acrylamide and bisacrylamide (N, N'-methylene-bisacrylamide). The polymerisation reaction is a vinyl addition catalysed by free radicals. The reaction is initiated by TEMED (tetramethylethylenediamine), which induces free radical formation from ammonium persulphate (APS). The free radicals transfer electrons to the acrylamide/bisacrylamide monomers, radicalizing them and causing them to react with each other to form the polyacrylamide chain (Shi and Jackowski 1998). In the absence of bisacrylamide, the acrylamide would polymerise into long strands, not a porous gel. Therefore, the pore size is determined by the concentration of acrylamide and bisacrylamide, pore size decreases with higher concentration.

Electrophoresis was performed using gels of different pore sizes. These discontinuities in the electrophoretic matrix were created using two layers of gel, namely stacking gel (spacer gel, upper one, large pores) and resolving gel (separating gel, lower one, smaller pores). Percentage of acrylamide/bisacrylamide can be adapted according to the wanted resolution (higher percentage will bring better resolution of small proteins and vice versa).

SDS PAGE gels were run in vertical electrophoresis tanks (Mini Protean II) from Bio-Rad (mini-gels: 7 cm × 8 cm × 0.75 or 1 mm). The gels were used directly or stored at 4°C wrapped in wet paper towel.

Samples intended to be separated on SDS-PAGE gel have to be mixed with loading buffer and boiled, 10 min at 95°C, before being loaded. After loading of the samples, the tank is filled with SDS-running buffer and the voltage set to 120–140.

	Resolving gel				Staking gel
	7.5%	10%	12%	15%	
H_2O	4.8	4.0	3.3	2.3	5.3
Acrylamide/bis-acrylamide (30%)	2.5	3.3	4.0	5.0	2.0
1.5 M Tris HCl (pH 8.8)	2.5	2.5	2.5	2.5	
1.0 M Tris HCl (pH 6.8)					2.5
SDS (10%)	0.1	0.1	0.1	0.1	0.1
TEMED	0.01	0.01	0.01	0.01	0.01
APS (10%)	0.1	0.1	0.1	0.1	0.1

SDS-loading buffer (5X):

- Add:
  - 250 mM Tris-HCl, pH 6.8
  - 50% saccharose [w/v]
  - 10% SDS [w/v]
  - 0.2% bromophenol blue [w/v]
  - 12.5% β-mercaptoethanol [v/v]

SDS-running buffer (4X):

- Add:
  - 25 mM Tris
  - 192 mM Glycine
  - 0.1% SDS [w/v]

ii. Native/semi-native PAGE - shift assay

"Native" or "non-denaturing" gel electrophoresis is run in the absence of SDS. While in SDS-PAGE the electrophoretic mobility of proteins depends primarily on their molecular mass, in native PAGE the mobility depends on both the protein's charge and its hydrodynamic size. The intrinsic charge on the protein (depends on the amino acid composition) at the pH of the running buffer governed the migration. Furthermore, the protein conformation can modify the mobility on the gel. While carried out near neutral pH, native gel can be used to study conformation, self-association or aggregation, and the binding of other proteins or compounds.

These properties make them good tools to detect double-stranded RNA binding, leading to a change either in the charge or in the conformation of a protein, thus a "gel-shift".

	Resolving gel				Staking gel	
	Native		Semi-native		Native	Semi-native
	10%	12%	10%	12%		
H_2O	Up to 10 mL					
Polyacrylamide (40%)	2.5	3	2.5	3	1.4	1.4
Tris pH 8.8	2.5	2.5	2.5	2.5		
Tris pH 6.8					2.5	2.5
SDS 10%			0.01	0.01		0.01
Glycerol	0.2	0.2	0.2	0.2		
APS	0.1	0.1	0.1	0.1	0.1	0.1
TEMED	0.01	0.01	0.01	0.01	0.01	0.01

Loading buffer (4X):	Running buffer (4X):	
Native and Semi-native	Native	Semi-native
o Add: - 0.5 M Tris pH 6.8 - 40% glycerol - 0.2% bromophenol blue [w/v]	o Add: - 25 mM Tris - 192 mM glycine - 0.2% bromophenol blue [w/v]	o Add: - 25 mM Tris - 192 mM glycine - 0.1% SDS [w/v] - 0.2% bromophenol blue [w/v]

iii. Coomassie staining of polyacrylamide gel

Coomassie blue (also known as Coomassie brilliant blue) staining is a simple and reasonably sensitive method (50–100 ng detection limit compared to 1–10 ng limit of silver staining) for visualizing protein bands in polyacrylamide gels.

Gel was soaked in GelCode® blue stain reagent (Pierce) for up to 1 h. After staining, gel was washed with distilled water until disappearance of non-specific staining (background).

c. Western blotting

Western blotting or immunoblotting is a technique used to identify and locate proteins based on their ability to bind to specific antibodies. This technique can give information about the size of your protein (with comparison to a size marker or ladder in kDa), and also information on protein expression (with comparison to a control, such as untreated sample and another cell type or tissue). Western blottings were performed after SDS polyacrylamide gel electrophoresis to separate proteins according to their masses. Separate proteins were transferred to polyvinylidene fluoride (PVDF) membrane (beforehand activated in methanol and wetted in blotting buffer; nitrocellulose membranes are cheaper than PVDF, but are more fragile and do not stand up well to repeat probing) by semi-dry blotting, using Trans-Blot® semi-dry electrophoretic transfer cell (Bio-Rad), applying a current of 0.56 mA/cm^2 of the gel for 1h. After the transfer step, non-specific binding sites were blocked by 1% [w/v] blocking solution in TBST for 1 h at room temperature or at 4°C overnight. After, washing (2 times 5 min) with Tris-buffered saline Tween-20 (TBST), the blots were incubated with the first antibody with the suitable dilution (in 1% [w/v] blocking solution in TBST) for 90 min at room temperature with gentle shaking. Unbound reagents were removed and background was reduced by 5 washing steps of 5 min in TBST. The second antibody (alkaline phosphatase or horse radish peroxidase conjugated) with the suitable dilution (in 1% [w/v] blocking solution in TBST) was incubated for 1 h at room temperature with gentle shaking, to detect first antibody bounded to the proteins of interest. After the incubation time, the membrane was washed three times with TBST and two times with TBS (5 min, at room temperature with vigorous shaking). After equilibration in P3 for 3 min the antibody-protein complexes are visualized with colorimetric substrates NBT/BCIP or Supersignal West pico chemiluminescent substrate kit (Pierce). However, anti-GFP (first antibody) is directly coupled with a horse radish peroxidise. In this case, washings were performed after the ON incubation and directly followed by revelation step. An alternative of the semi-dry transfer is the iBlot™ dry blotting system from Invitrogen. iBlot™ completes the transfer of proteins to a nitrocellulose or PVDF membrane in approximately 5–7 minutes.

Blotting (transfer) buffer:

- Add:
  - 25 mM Tris
  - 192 mM Glycine
  - 20% Methanol [v/v]

TBS:

- Add:
  - 10 mM Tris-HCl, pH 8.0
  - 150 mM NaCl

TBST:

- Add:
  - 0.1% Tween 20 in TBS [v/v]

Blocking / antibody solution:

- Add:
  - 5% milk powder in TBST [v/v]

d. General considerations on expression

i. Vector choice

According to the cell use (*e. g.*, yeast, mammalian cells, and bacteria) and the needs, the gene of the protein of interest has to be cloned in a vector-specific for eukaryotic or prokaryotic systems. For prokaryotic (bacteria) system, pTrcHis2 TOPO® TA expression kit has been used. For eukaryotic system, pcDNA™3.1/CT-GFP-TOPO® and pEGFP-C2 have been used. Primers were designed following the specifications indicated by manufacturers (mainly depending on the ligation method, see section generation of primers). TA cloning has already been presented above. Cohesive-end cloning is a bit more complex because it requires an additional step: digestion of the PCR product and the vector by restriction enzymes. However, insertion of the digested PCR product is orientated (only one possibility due to restriction sites specificities) whereas the TA cloning allows introduction of the insert in both directions.

ii. Restriction endonuclease reaction

Several key factors have to be considered (*e. g.*, proper amounts of DNA, enzyme, and buffer). Correct and optimal parameters will allow achievement of the digestion without any star activity. Star activity refers to altered or relaxed specificity of the enzymes (*i. e.*,

cleaving sequences which are similar but not identical to their defined recognition sequence). By definition, 1 unit of restriction enzyme will completely digest 1 µg of substrate DNA in a 50 µL reaction in 60 minutes. This ratio, [enzyme:DNA:reaction volume], can be used as a guide when designing reactions. A unit of restriction enzyme is defined as the quantity of restriction enzyme required to cut 1 microgram of DNA of the bacteriophage lambda in 1 hour.

Standard conditions for 50 µL reaction:

Restriction Enzymes	10 units is sufficient (generally 1 µL is used)
DNA	1 µg
10X restriction buffer	5 µL (1X)
BSA	add to a final concentration of 100 µg/mL (1X, if necessary)
ddH_2O	up to 50 µL

- Incubation time 1 hour
- Incubation temperature enzyme dependent

iii. Ligation, transformation, and screening

Except the TOPO® TA technology (Invitrogen; where the cloning is ligation-independent), TA, and cohesive-end require DNA-ligases. The T4 DNA ligase catalyzes the formation of a phosphodiester bond between juxtaposed 5'- phosphate and 3'- hydroxyl termini in duplex DNA (or RNA) with blunt or cohesive-end termini. The enzyme repairs single-strand nicks in duplex DNA, RNA or DNA/RNA hybrids but has no activity on single-stranded nucleic acids. The T4 DNA ligase requires ATP as cofactor.

Most of cloning kits provide T4 DNA ligase and protocol for ligation. However, for cohesive-end cloning, ratio [insert:vector] can be adapted in order to increase ligation efficiency. Generally, a ratio [2× insert:1× vector] provides good results.

The mixture was then incubated for 1 hour at 22°C or ON at 4°C and was directly used for transformation. After transformation, identification of the recombinant clones, and production of larger quantity of plasmid was performed, following the same protocols as presented above.

Standard conditions for 20 µL reaction:

Linear vector DNA	5–10 µL (50–400 ng)
Insert DNA	use a [1:1] up to a [3:1] molar (ratio of insert DNA termini to vector DNA)
10X ligation buffer	2 µL
50% PEG 4000 solution	2 µL (for blunt ends only)
T4 DNA Ligase	0.2–0.4 µL (1–2 u) for sticky ends 1 µL (5u) for blunt ends
ddH_2O	up to 20 µL

e. Prokaryotic system: bacterial cells

i. Expression of recombinant protein in *E. coli*

Once the recombinant vector produced, it was transformed in special bacteria strains, such as Top10 and Bl21.

Expression of recombinant proteins encoded by pTrcHis2 TOPO® TA vector is rapidly induced by the addition of isopropyl-β-D-thiogalactoside (IPTG) which binds to the lac repressor protein and inactivates it. Once the lac repressor is inactivated, the host cell's RNA polymerase can transcribe the sequences downstream from the promoter. The transcripts produced are then translated into the recombinant protein (Jacob and Monod 1961; Müller-Hill *et al.* 1968).

In addition, pTrcHis2-TOPO® encodes a COOH-terminal peptide containing the *c-myc* epitope and a 6×His tag for detection and purification of the recombinant protein.

For biochemical and structural studies (goal of this study), it is important to optimize conditions for the expression of soluble, functionally active protein, thus to extract the protein under native conditions (for antigen production, the protein can be expressed either in native or denatured form). Conditions for optimal expression of individual proteins must be determined empirically. Optimal growth and expression conditions for the protein of interest should be established with small-scale cultures before large-scale protein purification is attempted.

ii. Time-course analysis of protein expression (pilot expression)

To optimize the expression of a given protein construct, a time-course analysis of the level of protein expression was performed (pilot expression), in other words, the optimal induction period was established by checking the 6×His-tagged protein present at various times after induction.

To start the pilot expression, 2 mL of SOB or LB containing 50 µg/mL ampicillin (or carbenicillin) was inoculated with a single recombinant *E. coli* colony, and grown overnight at 37°C with shaking (225–250 rpm). The next day, 10 mL SOB or LB medium (containing the appropriate antibiotic, 50 µg/mL) was inoculated with 0.2 mL of the overnight culture. The culture was grown at 37°C with vigorous shaking to an OD600 = 0.6 (the cells should be in mid-log phase). 1 mL of the culture was removed (control, zero time point sample), centrifuged at maximum speed in a microcentrifuge for 30 seconds, and the supernatant aspirated. The cell pellet was frozen at -20°C. IPTG was added to a final concentration of 1 mM (9 µL of a 1 M IPTG stock to 9 mL) to initiate the expression. The culture was grown for 24 h at 37°C with shaking, and aliquots were taken at various times (1 mL after and prepared like the control sample.

Once all samples collected, pellets were dissolved in 1X SDS loading buffer and sonicated if needed. After 10 min at 95°C for denaturation, 12 to 15 µL samples were loaded on SDS gel. Finally, protein expression was analysed on Western blot.

iii. Determination of target protein solubility

Intracellular protein content is often a balance between the amount of soluble protein in the cells, the formation of inclusion bodies, and protein degradation. Thus, solubility and insolubility were also verified with small-scale cultures.

Samples were collected in a similar manner to that described in the time-course experiment. Each pellet was resuspended in 5 mL of lysis buffer for native purification. The samples were then frozen and thawed in cold water or alternatively, lysozyme was added to

1 mg/mL, and incubated on ice for 30 min. The mixture was then sonicated 6 times 10 s with 10 s pauses at 200–300 W (lysate always on ice). The lysate was centrifuged at 10,000 g at 4°C for 20–30 min. The supernatant was decanted (soluble protein) and saved on ice. The rest of the pellet was resuspended in 5 mL lysis buffer (suspension of the insoluble matter, insoluble protein).

Soluble and insoluble parts were then analysed on SDS-PAGE. Samples were prepared adding 5 µL of 2X SDS-PAGE sample buffer to 5 µL of the sample fractions. The samples were heated at 95°C for 5 min. After centrifugation at 15,000 g for 1 min, 20 µL samples were loaded on a 12% SDS-PAGE gel, and the gel was run according to standard procedures. Proteins solubility was directly analysed on Western-blot (anti-His tag).

<u>Lysis buffer for native conditions:</u>

- Add:
  - 50 mM NaH_2PO_4
  - 300 mM NaCl
  - 10 mM imidazole
- Adjust to pH 8.0 (NaOH)

<u>Lysis buffer for denaturing conditions:</u>

- Add:
  - 100 mM NaH_2PO_4
  - 10 mM Tris-Cl
  - 8 M urea or 6 M GuHCl
- Adjust to pH 8.0 (NaOH)

iv. Recombinant His-tagged protein purification (native purification)

To increase the protein solubility, culture were grown at 4°c for one week after IPTG induction.

Cells were harvested by centrifugation. The cells were resuspended in BugBuster™ protein extraction reagent (5 mL BugBuster™ per gram of bacteria, complemented with benzonase and protease inhibitors). This additional step increased the probability to get the protein in a soluble fraction. This reagent is formulated for the gentle disruption of the cell wall of *E. coli*, resulting in the liberation of soluble protein. It provides a simple, rapid, low cost alternative to mechanical methods, such as French press and sonication, for releasing expressed target proteins in preparation for functional studies or purification. The proprietary formulation utilizes a mixture of non-ionic detergents that is capable of cell wall perforation without denaturing soluble protein. The cell suspension was incubated on a

shaking platform or rotating mixer at a slow setting for 10–20 min at room temperature. Then the suspension was sonicated 4 to 5 times 10 s, followed by a centrifugation at 16,000 g for 20 min at 4°C. The supernatant which could contain a fraction of the soluble protein of interest was frozen.

The pellet was re-solubilised in lysis buffer for native condition (2 to 5 mL per gram pelleted bacteria), snap-frozen with liquid nitrogen, and thawed on ice. Then the suspension was sonicated 4 to 5 times 10 s and followed by a centrifugation (16,000 g for 20 min at 4°C). The supernatant should contain the main fraction of the soluble protein of interest. The supernatant was mixed with Ni-NTA agarose beads (1 mL per 4 mL supernatant) and incubated on a shaker for 1–2 h. Afterwards, a porous disc was positioned in the disposal polypropylene column and the solution (protein : Ni-NTA agarose beads) loaded.

The flow through was kept and checked on SDS gel. Proteins without histidine residues were then washed out of the matrix via two successive washing steps with 4 mL native washing buffer. The flow through was also conserved and checked. Finally, the protein of interest was eluted (4 times 500 µL).

<u>Wash buffer for native conditions:</u>

- Add:
  - 50 mM NaH_2PO_4
  - 300 mM NaCl
  - 20 mM imidazole
- Adjust to pH 8.0 (NaOH)

<u>Elution buffer for denaturing conditions:</u>

- Add:
  - 50 mM NaH_2PO_4
  - 300 mM NaCl
  - 250 mM imidazole
- Adjust to pH 8.0 (NaOH)

f. Eukaryotic system: mammalian cells

i. Cell culture

Human embryonic kidney cells, (HEK or HEK-293 cells) originally derive from an embryonic human kidney and were generated by transformation of normal cells with adenovirus type 5 DNA (Graham *et al.* 1977).

HEK-293 cells should be grown in a monolayer, preferably in plastic Petri dishes or flasks. Under optimum growth conditions (37°C and 5% CO_2), HEK-293 cells double about every 36 hours. To maintain consistency, cells should not be splitted indefinitely, not be allowed to become overly confluent, and not be seeded too sparsely. Cells should be split every 2–4 days when they reach 80–90% confluency. Cells are cultured in Dulbecco's modification of Eagle's medium (DMEM, high glucose and with L-glutamine) supplemented with 100 units/mL penicillin G sodium, 100 mg/mL streptomycin, 4 mM L-glutamine, and 10% foetal bovine/calf serum.(FCS).

RAW-Blue cells are derived from RAW 264.7 mouse macrophages with chromosomal integration of a secreted embryonic alkaline phosphatase (SEAP) reporter construct inducible by NF-kB and AP-1. These cells should be grown in a similar way to HEK-293. RAW-Blue cells were maintained and subcultured in growth medium with 200 g/mL zeocin. The growth medium was renewed two times a week. When the cultures were reaching 70 to 80% confluency, the cells were passed. There as well, high confluency was avoided. Cells are cultured in DMEM high glucose supplemented with 10% FCS (growth medium) and are maintained in growth medium supplemented with 200 µg/mL zeocin.

Both cell-lines were splitted as follows. First, the medium was removed and the cells washed once with sterile PBS (containing no Ca^{2+} or Mg^{2+}). To detach cells, 1–2 mL of trypsin-EDTA solution was added for 1–3 min (cells should not be exposed to trypsin for extended periods). Then, few mL of complete growth medium or of BSA were added in order to stop trypsinization. After a gentle centrifugation, the cells were softly resuspended, but thoroughly, in fresh and complete media. The desired number of cells was finally transferred in 50 mL cell culture flask containing 10 mL of complete medium; by gentle rocking, the cells were distributed homogenously.

Of course, contaminations had to be avoided. All possibly virus-contaminated materials, including fluids, were autoclaved or disinfected with 10% bleach or with chemical disinfectant before use.

1X PBS buffer:

- Add:
  - 137 mM NaCl
  - 2.7 mM KCl
  - 8.1 mM NaH_2PO_4
  - 1.8 mM KH_2PO_4
- Adjust to pH 7.2
- Autoclave

ii. Cell counting

Cell counting was performed using a Neubauer chamber (cell counting chamber or haemocytometer) and with Trypan blue staining. Trypan blue is a vital stain used to selectively colour dead tissues or cells (in a viable cell Trypan blue is not absorbed; however, it traverses the membrane of a dead cell).

The Neubauer chamber consists of nine 1.0 mm × 1.0 mm large squares separated from one another by the triple lines. The area of each is 1 mm^2 and represents a volume of 0.1 mm^3 or 10^{-4} mL. Since the volume of one square is 10^{-4} mL, the average number of cells counted (the total number of cells in the upper left and right and bottom left and right and centre, divided by 5, the number of squares counted) multiplied by 10^4 give the concentration (cells/mL) in the suspension used to charge the haemocytometer. If the suspension is diluted prior to counting, the concentration has to be multiplied by the dilution factor.

iii. Transfection methods

Transfection describes the introduction of foreign material into eukaryotic cells using a virus vector or other means of transfer. Several methods were used; however only MATra (magnet assisted transfection) protocol will be described.

On the eve of transfection, HEK-293 cells (or RAW-Blue cells) were seeded at 1.5×10^5 cells/mL in order to reach the wanted confluency. The day of transfection, nucleic acid (1.2 µg per well for 12-wells plates) was diluted with DMEM serum-free and

added to the appropriate amount of MATra-A reagent. The complexes were incubated at ambient temperature for 20 minutes; while the cell medium was changed. The [DNA:bead] mixture was then added to the cells in a drop-wise manner and the wells (or flasks) swirled to ensure distribution over the entire plate surface. After mixing, the plate or flask was immediately placed on the suitable magnet plate and incubated for 15 minutes before removing the magnet plate. Replacing with fresh medium was not needed. The transfection efficiency was assessed 24 to 48 hours after transfection.

HEK-293 cells were transfected with *SDSURVL* cloned in pEGFP-C2 while RAW-Blue cells were transfected with *SDTILRc* cloned in pcDNA 3.1/CT-GFP-TOPO.

iv. Establishment of a stable cell line

HEK-293 were transfected with pEGFP-C2 construct, containing *Suberites domuncula* survivin–like. For constitutive expression (HEK293-*SDSURVL* stable cell line), the cells were split 48 h post-transfection into selective medium, containing aminoglycoside antibiotic G418, at the appropriate concentration (200 µg/mL). Medium was replenished every 3–4 d until G418-resistant colonies were detected. The resistant colonies were then isolated and propagated. *SDSURVL* expression was confirmed microscopically (thanks to the GFP tag) and immunologically (Western blots).

v. Cryopreservation

For long-term storage of cells, HEK-293 or RAW-Blue cells were trypsinized, pelleted, and resuspended in freezing medium. This suspension was transferred in cryotubes for 12 h at -80°C before storage in liquid nitrogen.

<u>Freezing Medium</u>:

- DMEM
- 4.5 g/l glucose
- 10% fetal bovine serum
- 10% DMSO

vi. Apoptosis induction in transfected and mock-transfected HEK-293

To determine the effect of *SDSURVL* expression, transfected (HEK293-*SDSURVL*) and mock-transfected HEK-293 cells were seeded into 96-well plates (0.5×10^4 cells/mL) and exposed to Pam_3Cys-Ser-$(Lys)_4$ (20 ng/mL, 6 h) or cadmium (50 µM, 6 h).

vii. MTT assays

The MTT assay was first described by Mosmann in 1983 and some modifications suggested by Denizot and Lang in 1986. In MTT assays the activity of mitochondrial succinate dehydrogenase is measured as indicator of cellular metabolic rate and viability.

After exposure treatments, 10 µL of an MTT solution (5 mg/mL in PBS, yellow) were added to each well for 4 h (37°C). The resulting formazan crystals were dissolved in a stop solution (40% [v/v] DMF, 10% [w/v] SDS; purple). Absorbances were determined at 570 nm with a Bio-Rad 3550 enzyme linked immunosorbent assay (ELISA) plate reader.

viii. Protein extraction

After exposure treatments, cells were washed once with PBS, harvested from their culture dishes, and centrifuged at 400 g for 10 minutes. After centrifugation, supernatant was removed, the pelleted cells resuspended in the appropriate volume of cold NP-40 cell lysis buffer containing 1X protease inhibitor cocktail (just added before use), and the tube placed on ice fro 30 min. The volume of lysis has to be adapted to the number of harvested cells, count ≈ 50 µL per well of a 6-well plate. The cell lysate was span at 10,000 g for 15 minutes at 4°C. The supernatant was carefully collected, without disturbing the pellet, and transferred to a clean tube.

Total protein can be conserved and frozen at -20°C.

NP-40 cell lysis buffer:

- 50 mM Tris-HCl pH 8.0
- 150 mM NaCl
- 1% NP-40

ix. Western blot analyses

After protein extraction, the total protein concentration was determined in the supernatant with BCA protein assay reagent (see protein quantification section). 20 µg of total proteins were subjected to electrophoresis through 14% [v/v] polyacrylamide gels, containing 0.1% [w/v] SDS. Following electroblotting, PVDF membranes were incubated either with rabbit anti-GFP (1:1,000), with anti-survivin (1:500; directed against $aa_{1\text{-}12}$ of human survivin, a region without homology to poriferan survivin–like) or with anti-tubulin (1:1,000).

x. Innate immunity induction in transfected and mock-transfected RAW-Blue cells

SDTILRc expression was confirmed microscopically (thanks to the GFP tag) and immunologically (Western blots). To determine the effect of *SDTILRc* expression, transfected (RAW-Blue-*SDTILRc*) and mock-transfected RAW-Blue cells were seeded into 96-well plates (180 µL, 5.5×10^5 cells/mL) and exposed to different concentrations of LPS (100, 10, 1, 0.1 and 0 µg/mL) for 24 h.

xi. SEAP quantification

Following LPS exposure, the secreted embryonic alkaline phosphatase (SEAP) reporter construct expression was detected and quantified. In 96-well plate, 160 µL of resuspended QUANTI-Blue™ was loaded and 40 µL of induced RAW-Blue cells supernatant was added. After 1 h incubation at 37°C, the SEAP levels were measured using a spectrophotometer at 620 nm.

17. Statistical analyses

Statistical analyses were performed according to the Mann-Whitney U test, a non-parametric statistical test for assessing whether the difference in medians between two samples of observations is statistically significant. The level of significance was set at $P \leq 0.05$.

V. RESULTS

A. Characterization of the poriferan survivin–like SDSURVL

1. Survivin–like complementary DNA (*SDSURVL*)

The complete *SDSURVL* cDNA (fig. 16) was isolated from a *S. domuncula* cDNA library (Kruse *et al.* 1997) by means of PCR and degenerate primers (forward primer: 5' – GAT GGC IGA GGC T/CGG CTT – 3', and reverse primer: 5' – CT/CA TGT CCA GTT TT/CA AG/AA – 3') that were directed against a conserved region within the BIR domain.

SDSURVL consists of 620 nt (excluding the poly(A) tail) that encompass a CDS (coding sequence) of 447 nt (excluding the first stop codon), beginning at $nt_{43–45}$ (ATG/Met_{start}). During Northern blot analyses, a DIG-labeled probe detected a transcript whose size was consistent with that of the cDNA (≈ 650 nt; fig. 30).

The Kozak sequence encompasses three start codon bases upstream and one start codon base downstream, it plays a major role in the initiation of translation process (reviewed in Kozak 1987). *S. domuncula* Kozak sequence follows the consensus sequence (**G/A**)NNATGG where the G or A (bold) and the G (downstream of the underlined ATG/Met_{start}) are the most critical (reviewed in Kozak 1991). The ATG initiation codon is shown underlined and N represents any nucleotide.

The stop codon starts at the position 490 and is followed by a polyadenylation signal (AGTACA, $nt_{524-529}$). This motif follows the consensus sequence defined as NNUANA (Beaudoing *et al.* 2000). This nucleotide motif is located downstream the coding part of a gene and defines the position where has to be added the poly(A) tail.

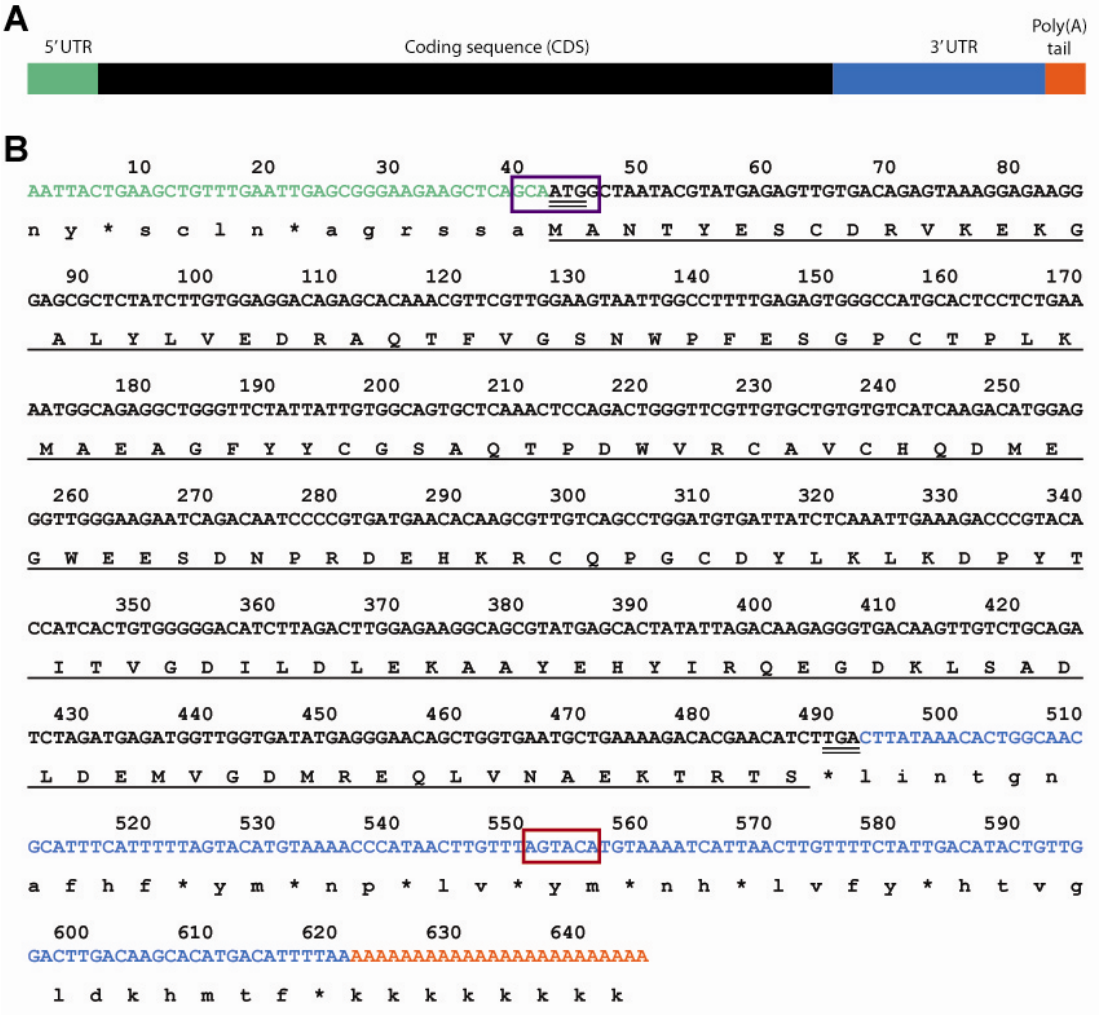


Figure 16: cDNA and deduced amino acid sequence of SDSURVL. The structure of cDNA ,including the untranslated regions (UTR), is depicted (drawn to scale, A). Translation of the survivin–like nucleotide sequence into a protein sequence (B). The ATG/Met_{start} and the TGA/stop codon are double underlined. Deduced amino acid sequence is given below the nucleotide sequence (underlined capital letters). The nucleotides are in the 5′ to 3′ direction. The 5′, the 3′ untranslated regions, and the poly(A) tail are in green, blue, and orange letters, respectively. The Kozak sequence and the polyadenylation signal are marked in violet and red boxes, respectively.

2. Survivin–like genomic sequence (*SDSURVL*)

Genomic sequences were isolated from a *S. domuncula* genomic DNA library (Seack *et al.* 2001) via PCR.

The *S. domuncula* gene comprises four exons and three introns (fig. 17). Each exon/intron border displays typical splicing signals (GT–intron–AG; reviewed in Venables 2004).

The phases of introns vary: the first and third introns are in phase 0, while the second intron is in phase 2. The human survivin gene shares the same features (fig. 18) except for the borders of the third intron (AT–intron–GG).Parts of the first and third exons and the complete second exon are coding for the BIR domain (fig. 18).

Analyses of the 5'- flanking region (fig. 19) of human (1) and murine (2) survivin genes reveal the presence of a TATA-less promoter, containing a canonical CpG island and several transcription factor binding sites (Li and Altieri 1999a,b). However, *in silico* analyses of the putative poriferan promoter region (3) predicted the presence of a TATA box (nt_{-64} to $_{-73}$) in addition to other cis-regulatory elements, CAAT box (nt_{-109} to $_{-113}$), and ribosomal binding site (nt_{-5} to $_{-8}$). Moreover, similar to the human and murine promoters two putative Sp1 sites (nt_{-51}, $_{-292}$) have been predicted, although in contrast to the former ones no cell cycle dependent elements (CDE) or cell cycle homology regions (CHR) have been found in the poriferan non-coding region.

Figure 17: Exon/intron architecture. Each exon/intron border is marked (boxed) as well as introns, start and stop codon (underlined). The splice donor sites (GT) are depicted in green letters and the splice acceptor sites (AG) in orange.

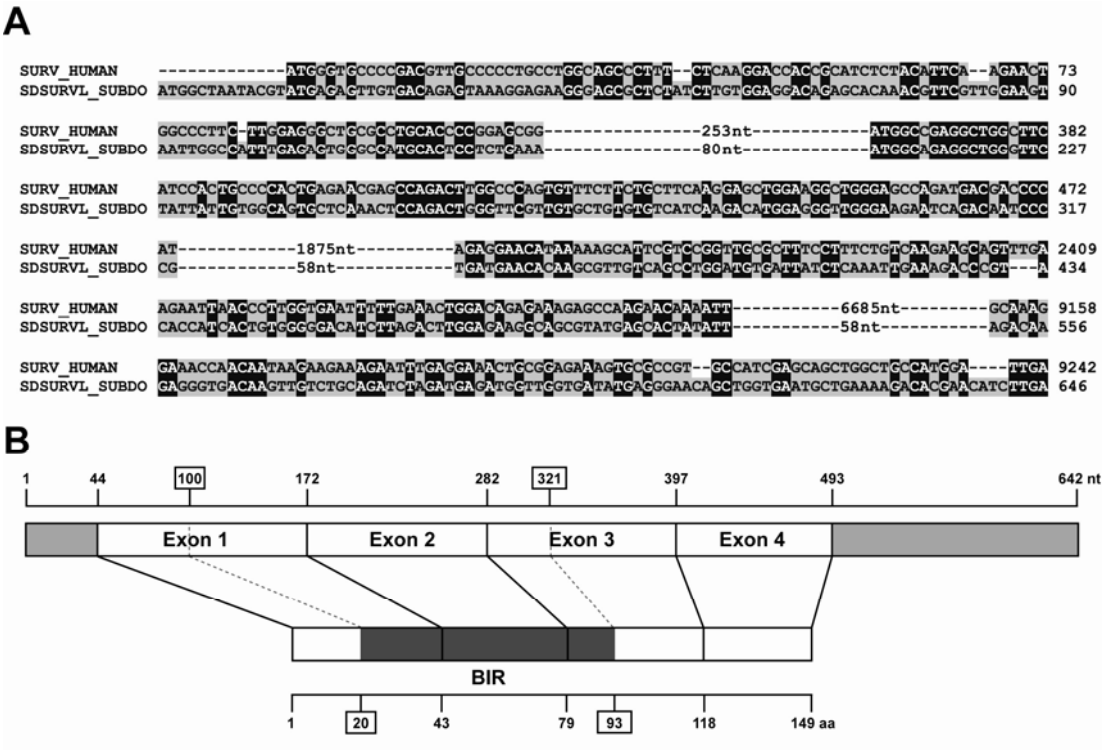


Figure 18: (A) Alignment of *SDSURVL* and human survivin exons with correlating intron phases. Sizes of introns are indicated. (B) Schematic representation of *SDSURVL* mRNA and its translational product, drawn to scale. The BIR domain and its contributing exons are marked and untranslated regions as well.

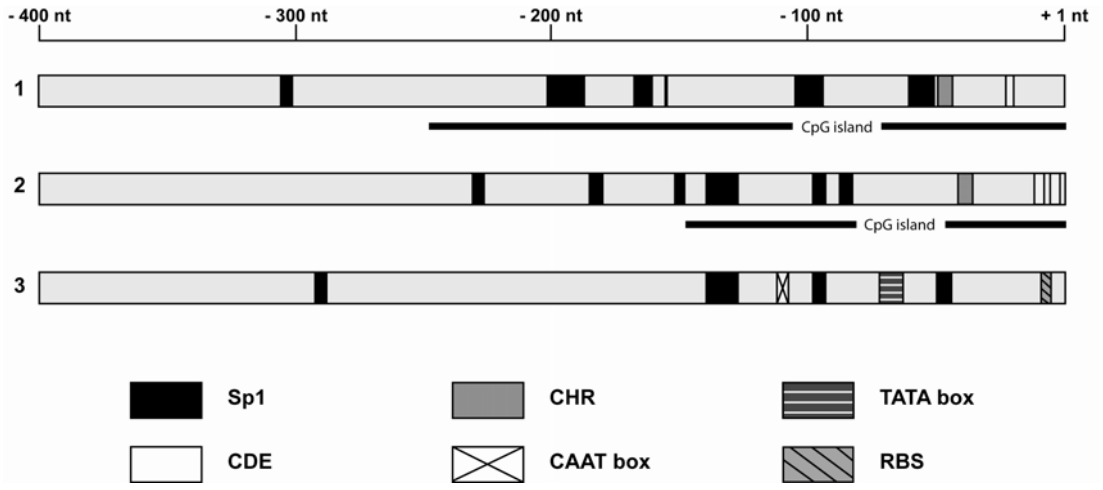


Figure 19: Schematic representation of survivin promoter regions from representative species: murine (1), human (2), and poriferan (3). The different elements: Sp1 (sequence-specific DNA-binding protein-1), CDE (cell cycle dependent element), CHR (cell cycle homology region), CAAT and TATA boxes, and RBS (ribosome binding site) are drawn to scale (first base of Met_{start}, +1 nt). The vertebrate CpG islands are also disclosed. The promoter region is depicted up to -400 nt.

3. Survivin–like protein (SDSURVL)

The deduced protein, SDSURVL, comprises 149 aa with an expected size of 17,003 D. SDSURVL shows significant sequence similarity to vertebrate survivin homologues, such as *Gallus gallus* survivin isoform 1 (with an expect value [*E* value; Coligan *et al.* 2000] of 2 e-20) and human survivin (8 e-18) (fig. 20). In contrast, aa similarities to lower metazoan survivin homologues are considerably inferior, *e. g.*, insect (*D. melanogaster* [7 e-11]), placozoan (*Trichoplax adhaerens* [2 e-11]), and cnidarian (*Nematostella vectensis* [1 e-15]). Viral genomes often encode antiapoptotic or cell cycle promoting proteins, in order to bypass the host defense system. Viral survivin–like proteins generally display lower sequence homology to Metazoa and in particular to *S. domuncula* survivin, *e. g.*, *Epiphyas postvittana nucleopolyhedrovirus* IAP-3 (4 e-07) and *Cydia pomonella granulosis virus* ORF17 IAP-3 (3 e-6).

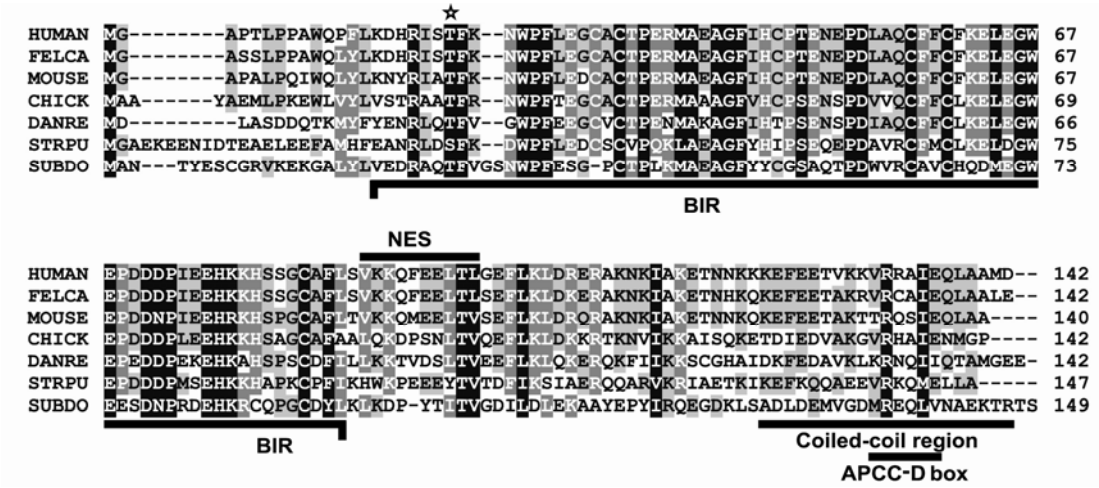


Figure 20: *Suberites domuncula* survivin–like protein SDSURVL. The deduced amino acid sequence (SUBDO) was aligned with survivin homologues of human (HUMAN, NCBI accession number CAG46540), *Felix catus* (FELCA, NP_001009280), murine (MOUSE, NP_033819), *G. gallus* (CHICK, NP_001012318), *D. rerio* (DANRE, NP_919378), *S. purpuratus* (STRPU, XP_796206). Conserved residues (identical or similar with respect to physicochemical properties) in all sequences are shown in white letters on black; those in 80% (60%) are in white on gray (black on light gray). The following features are marked, BIR domain, NES, coiled-coil region, APCC-D box (underlined) as well as the putative phosphorylation site (star).

Since SDSURVL features a single BIR domain, a putative nuclear export signal (NES), and a coiled-coil region, but is missing a RING domain (characteristic of non-survivin IAP) the molecule represents a new member of the survivin subfamily of IAP. The poriferan BIR domain (aa_{20-93}) has an *E* value of 4.85 e-22 and follows the Prosite consensus sequence PS01282 with two exceptions (marked in bold): [HKEPILVY]-x(2)-R-x(3,**8**)-[FYW]-x(11,14)-[STAN]-G-[LMF]-x-[FYHDA]-(x(4) | **x(7)**)-[DESL]-x(2,3)-C-x(2)-C-x(6)-[WA]-x(9)-H-x(4)-[PRSD]-x-C-x(2)-[LIVMA]. The putative NES (aa_{95-103}) loosely fits a consensus sequence proposed by Heger *et al.* (2001), L-x(2,3)-[FILVM**P**]-x(2)-[L**I**]-x-[LI**V**]. *In silico* analysis running P2SL (prediction of protein subcellular localization sites from their amino acid sequences, Atalay and Cetin-Atalay 2005) reveal a cytoplasmic and a nuclear localization. This was corroborated with PSORT version II software (Nakai and Horton 1999), which calculated the following localization probabilities: cytoplasmic 47.8%, nuclear 30.4%, and mitochondrial 21.7%. Furthermore, COILS algorithms (Lupas *et al.* 1991) indicate the typical COOH-terminal coiled-coil conformation ($aa_{127-147}$), usually present in survivin but missing in other IAP family members. Moreover, the conserved threonine residue (whose phosphorylation by cyclin-dependent kinase p34cdc2 mediates human survivin activation; Wall *et al.* 2003), is present in SDSURVL (aa_{26}). Additionally, a ubiquitin ligase complex-binding destruction motif (APCC-D box; $aa_{136-141}$) was predicted via ELM algorithms. APCC-D box represents a target site for the anaphase-promoting ubiquitin ligase complex, resulting in the degradation of cell cycle-regulatory proteins. So far, no such feature has been described for other survivin molecules.

4. Phylogenetic relationships of the poriferan survivin–like protein (SDSURVL)

To determine phylogenetic relationships, molecular phylogenetic analyses were performed by the neighbor-joining method (Saitou and Nei 1987), including survivin sequences of representative members of metazoan taxa and related BIR-containing sequences of viruses and yeast. Accordingly, viruses, Ecdysozoa, and Deuterostomia form well-supported clades

respectively (fig. 21). However, the position of Lophotrochozoa, close to Deuterostomia, was not robustly resolved due to limited taxon sampling (only one platyhelminth survivin-like sequence is known). Nevertheless, the phylogenetic distance tree shows the poriferan SDSURVL protein at the base of two robustly supported monophyletic groups, comprising viral and metazoan sequences respectively. In particular the latter one reveals a distinct pattern of well-supported and established sub-clusters.

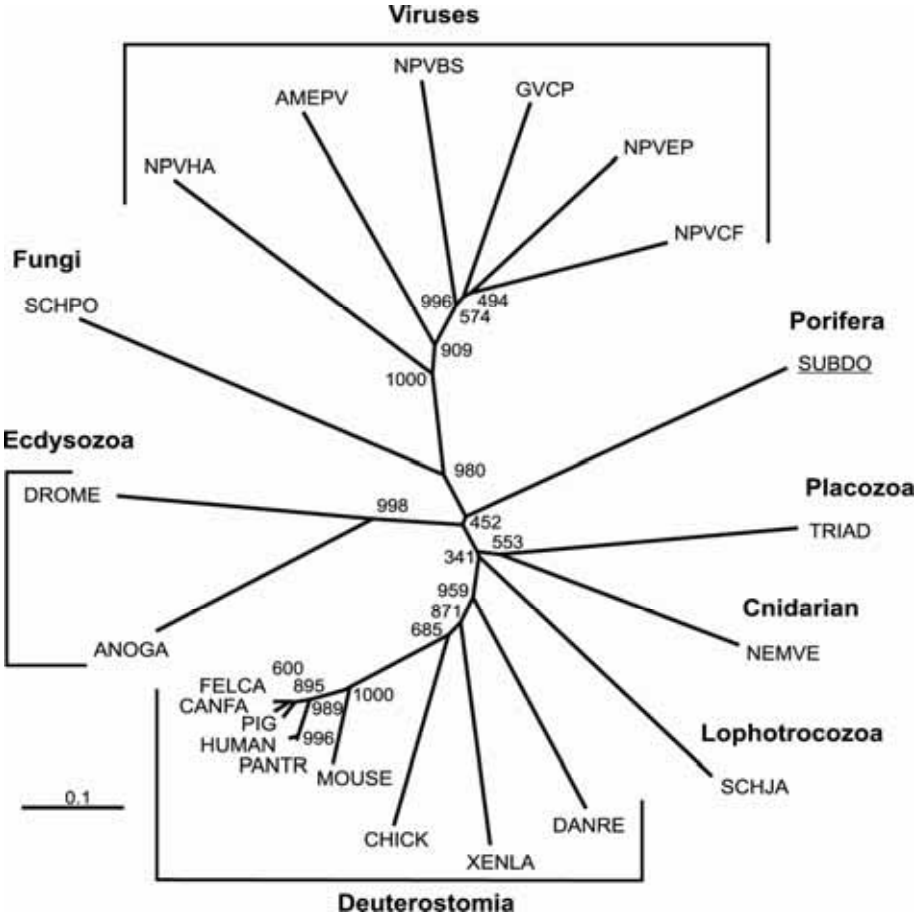


Figure 21: Radial phylogenetic tree depicting the evolutionary relationship of *S. domuncula* SDSURVL to survivin homologues of Metazoa, fungi, and viruses. The tree was generated after alignment of the aforementioned sequences (fig. 20), furthermore integrating sequences from: *T. adhaerens* (TRIAD, EDV24865), *Schistosoma japonicum* (SCHJA, AAW27178), *Nematostella vectensis* (NEMVE, XP_001624746), *Anopheles gambiae* (ANOGA, XP_317026), *D. melanogaster* (DROME, NP_650608), *Xenopus laevis* (XENLA, BAD98266), *Pan troglodytes* (PANTR, XP_512010), *Sus scrofa* (PIG, NP999306), *S. pombe* (SCHPO, NP_587866), *Canis lupus* (CANFA, NP_001003348) as well as BIR-containing viral sequences of *Helicoverpa armigera nuclear polyhedrosis virus* (NPVHA, NP_075172), *Buzura suppressaria NPV* (NPVBS, AAC34373), *Epiphyas postvittana NPV* (NPVEP, NP_203195), *Choristoneura fumiferana NPV* (NPVCF, NP_848342), *Amsacta moorei entomopoxvirus* (AMEPV, NP_064803), and *Cydia pomonella granulosis virus* (GVCP, NP_148801).

B. Characterization of the poriferan caspase–like SDCASL and SDCASL2

1. Caspase–like sequences (*SDCASL* and *SDCASL2*)

Caspases (cysteine-dependent aspartyl-specific proteases) are prominent among apoptotic proteases. They are synthesized as zymogens that become activated by scaffold-mediated transactivation or by cleavage via upstream proteases during an intracellular cascade (fig. 4; Yuan and Horvitz 2004).

After identification of two different sequences via PCR using a forward degenerate primer (5' – T/CTI ATT/C/A CAA/G GCI TGT/C A/T/CG/TI GG – 3'), which targeted the cysteine catalytic site, combined with a library-specific primer, the complete clones were isolated from the *Suberites* cDNA library. Subsequently, their correct sizes were verified on Northern blot (fig. 30). SDCASL was the first published sequence (NCBI accession number CAL36107) followed by SDCASL2 (NCBI accession number FM210332) termed in reference to SDCASL.

SDCASL (1,681 nt excluding poly(A); fig. 22) comprises an ORF between $nt_{148–150}$ (Met_{start}) and $nt_{1579–1581}$ (stop). The sequence surrounding the Met_{start} is optimal for translation initiation (Kozak sequence, AAATGG). The stop codon is followed by a polyadenylation signal (TATAGA, $nt_{1622-1627}$). The ORF encodes a putative protein of 477 aa (SDCASL) with an expected Mr of 53475. This protein reveals greatest sequence homologies to caspase-7, especially to human caspase-7 chain A/B (*E* value 2 e-11), but also to *Drosophila* DCP-1 (*E* value 7e-11) and to caspase-8 of *Bos taurus* (*E* value 5e-9) (fig. 23).

SDCASL2 (1,827 nt - excluding poly(A); fig. 22) comprises an ORF between $nt_{175–177}$ (Met_{start}) and $nt_{1738–1740}$ (stop), coding for a putative protein of 521 aa (calculated Mr 58427). No typical Kozak sequence is observed; however a polyadenylation signal (AGTATA) is located 21 nucleotides downstream the stop codon ($nt_{1762-1767}$). SDCASL2 reveals sequence homology not only to SDCASL (*E* value 7 e-31), but also to members of the caspase-7 subfamily, *e. g.*, *Xenopus laevis* and human representatives (3 e-16 and 3 e-14 respectively), and to *D. melanogaster* caspase DCP-1 (2 e-12) (fig. 23).

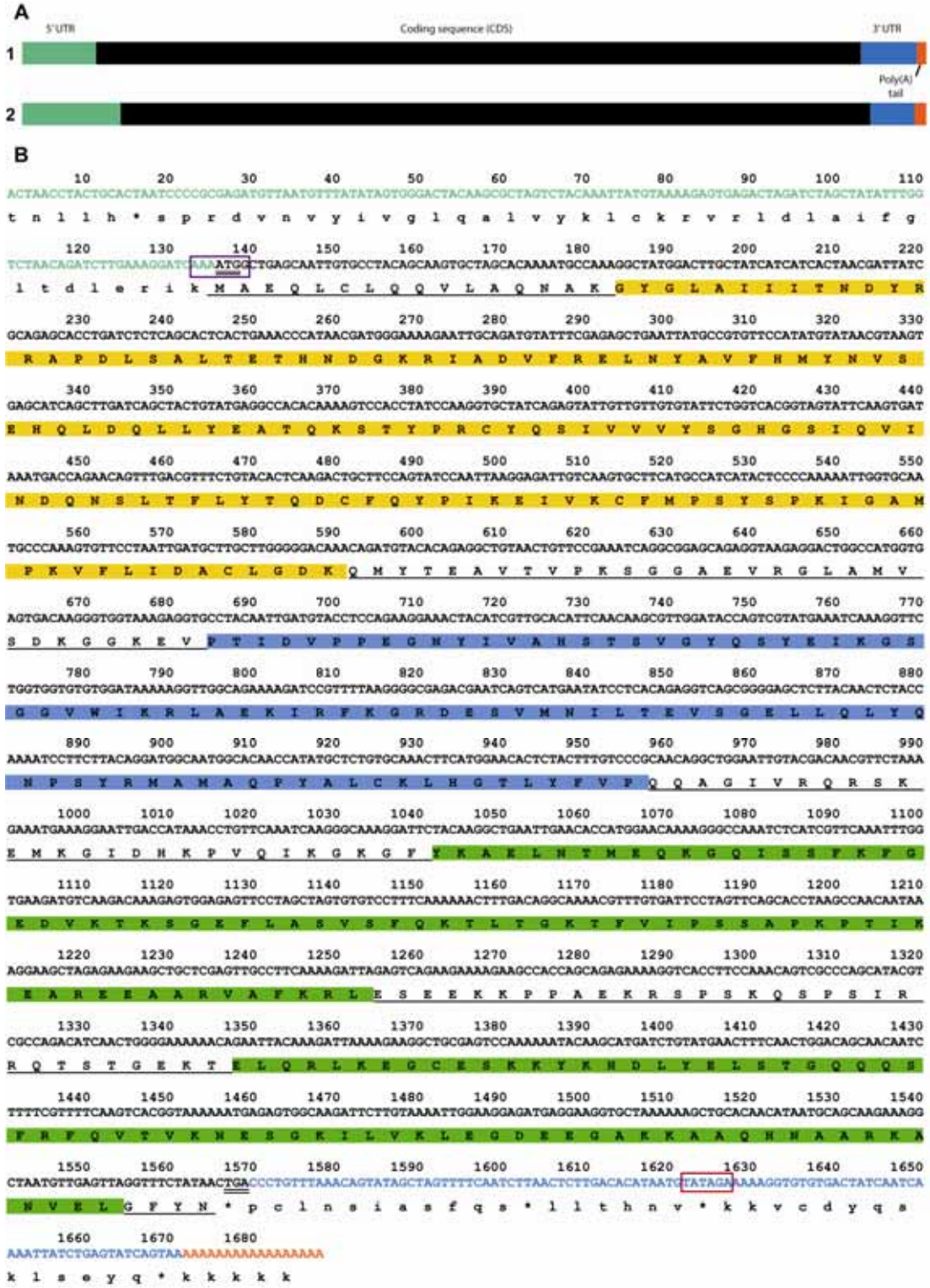
2. Caspase–like proteins (SDCASL and SDCASL2)

Depending on their role in the apoptotic process (either as initiator or effector molecules), caspases reveal different domain architecture, but all contain a CASc domain (Smart accession number SM00115), which has to be cleaved into two subunits (small, p10 and large, p20) during activation.

SDCASL and SDCASL2 putatively represent such a proenzyme comprising the 2 subunits, p10 ($aa_{184–274}$ and $aa_{191-267}$, respectively) and p20 ($aa_{17–152}$ and aa_{22-154}, respectively). P20 contains the characteristic Cys and His active sites (fig. 21 and 22). Regarding the caspase family-specific cysteine active site SDCASL follows the established pattern [K**M**]-P-K-[LIVMF]-[LIVMFY]-[LIVMF](2)-[QP**D**]-[AF]-C-[RQG**L**]-[GE] (Prosite Pattern PS01122) with 3 exceptions (marked in bold). However, the sequence homologous to the histidine active site (PS01121) reveals a surprising number of alterations (marked in bold): [H**P**]-x(2,4)-[SC]-x(2)-{A}-x(**0**,1)-[LIVMF**Y**](2)-[ST]-**G**-H-G. The cysteine active site consensus pattern of SDCASL2 is conserved with two exceptions ([K**L**]-P-K-[LIVMF]-[LIVMFY]-[LIVMF](2)-[QP**D**]-[AF]-C-[RQG]-[GE]), whereas the histidine active site features three alterations, [H**L**]-x(2,4)-[SC]-x(**1**,2)-{A}-x-[LIVMFY](2)-[ST]-**G**-H-G.

The absence of a large NH_2-terminal prodomain ($aa_{1–16}$ and aa_{1-21}, for SDCASL and SDCASL2, respectively) characteristic of initiator caspases, implies a classification into the group of effector caspases. However, both carry an uncharacteristically large COOH-terminal stretch of 203 aa (SDCASL) and 254 aa (SDCASL2) following the p10 subunit. The SDCASL segment ($aa_{275–477}$) reveals no similarity to any published sequence so far except a homology (*E* value 7.5e-2) to a double stranded RNA–binding motif (dsrm, $aa_{405–475}$; Pfam accession number PF00035). Similarly, SDCASL2 COOH-terminal region ($aa_{267–521}$) also shows a weak homology to a dsrm motif ($aa_{441-514}$, *E* value 4.14e00).

In silico analysis running P2SL reveal a cytoplasmic localization for both proteins.


A
5'UTR
Coding sequence (CDS)
3'UTR
1
Poly(A) tail
2
B

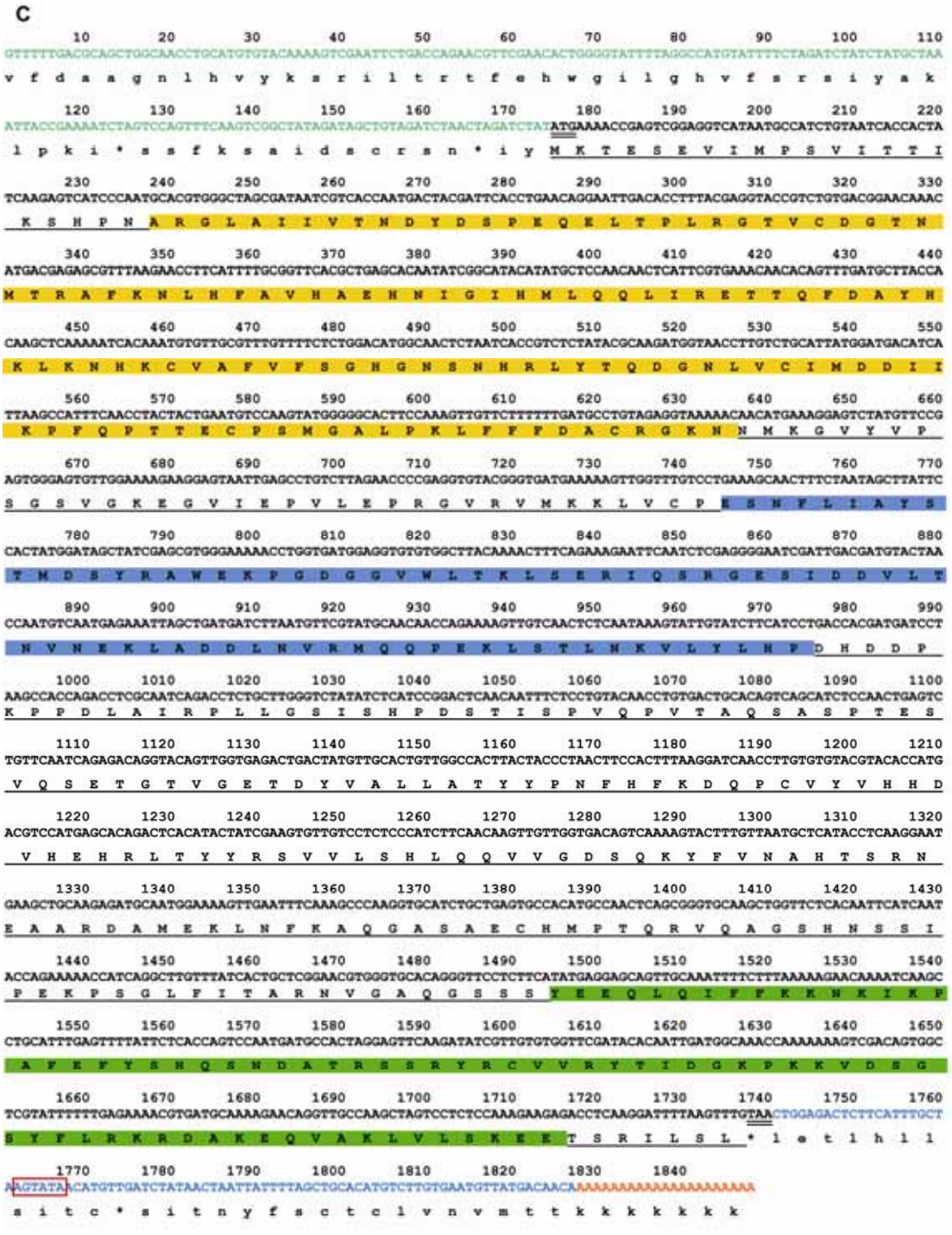


Figure 22: Complementary DNA and deduced amino acid sequences of SDCASL and SDCASL2. The structure of cDNA (1- SDCASL and 2- SDCASL2), including the untranslated regions (UTR) and the poly(A), are depicted (drawn to scale, A). The caspase–like nucleotide sequences is translated into a protein sequence (B- SDCASL and C- SDCASL2). The ATG/Met_{start} and the TGA/stop codon are double underlined. The nucleotides are in the 5′ to 3′ direction. The 5′, the 3′ untranslated regions, and the poly(A) tail are in green, blue, and orange letters, respectively. The Kozak sequence and the polyadenylation signal are marked in violet and red boxes, respectively. Deduced amino acid sequence is given below the nucleotide sequence (underlined capital letters). The p10, p20, and dsrm domains are highlighted in blue, yellow, and green respectively.

3. Phylogenetic relationships of the poriferan caspase–like proteins (SDCASL and SDCASL2)

Due to this unusually large COOH-terminal region and the small prodomain, only the p20– and p10–bearing CASc domains of the poriferan caspases were considered to be of informative value for phylogenetic analyses (fig. 23). Thus, poriferan CASc were aligned with the corresponding domains of representatives of all caspase subfamilies (also including the *Schizosaccharomyces pombe* metacaspase), in order to study phylogenetic relationships.

The alignment reveals highest aa sequence homology to the caspase-7 subfamily. SDCASL shares 21% identical and 37% similar aa with the human protein, 21%/35% with the *Xenopus* molecule, and 20%/35% with that of chicken. In addition, SDCASL2 displays 36% similarities with the chicken and human sequence or 34% with *X. laevis* one. The homology to caspases of other subfamilies was considerably lower, *e. g.*, for SDCASL: chicken caspase-1 (13%/29%) and human caspase-12 (13%/30%) and for SDCASL2: human caspase-12 (12%), chicken caspase-1 (13%), or almost non-existent in case of the yeast metacaspase (8% and 5% for SDCASL and SDCASL2, respectively).

A phylogenetic tree (fig. 23), based on this alignment, was rooted with the distantly related *Nitrobacter sp.* Cys protease C14, exclusively containing the p20 subunit. The resulting tree reveals 6 clusters (numbered according to Wiens *et al.* 2007), comprising caspase-3 and -7 (I), caspase-6 and DCP-1 (II), caspase-8 and -10 (III), caspase-9 (IV), caspase-2 and Ced-3 (V), and a sixth cluster with a greater diversity, consisting of caspase-1, -4, -5, and -12 and the *Danio rerio* putative caspase-2. No resolution as to the positioning of the yeast metacaspase could be obtained, likely due to the mitochondrial origin of metacaspases (Koonin and Aravind 2002).

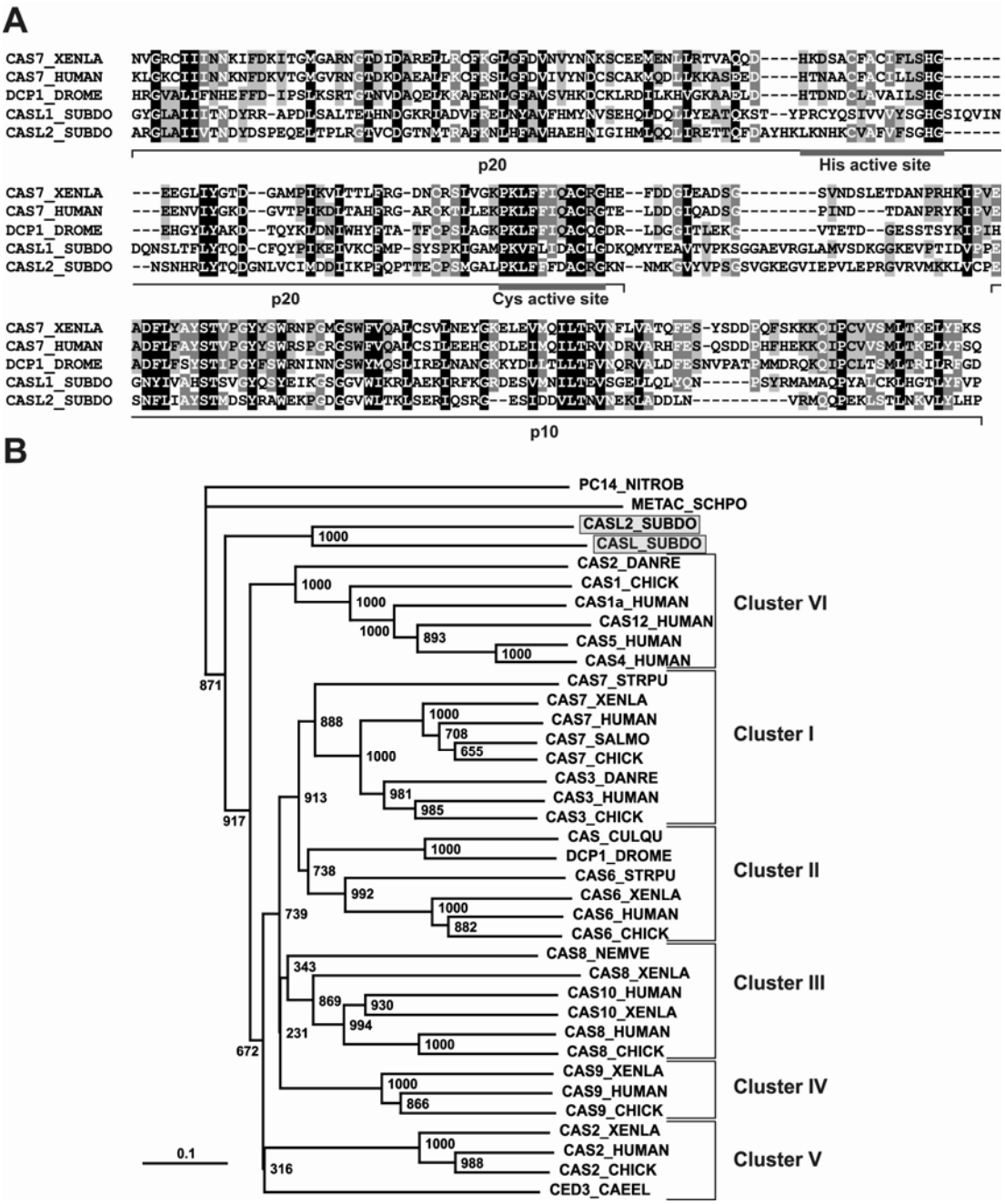


Figure 23: SDCASL and SDCASL2 alignment and evolutionary relationship to CASc homologues of Metazoa, fungi, and viruses. (A) The CASc domain of the deduced poriferan proteins (CASL_SUBDO; $aa_{17–247}$ and CASL2_SUBDO; $aa_{22–267}$) were aligned with the corresponding domain of *D. melanogaster* DCP-1 (DCP1_DROME; AAB58237; $aa_{77–315}$), *Xenopus laevis* caspase-7 (CAS7_XENLA; BAA94748; $aa_{71–316}$), and human caspase-7 (CAS7_HUMAN; BAG62964; $aa_{74–311}$). The putative subunits contained within (p20 and p10) are marked as well as the Cys and His active sites. Conserved residues (identical or similar with respect to physicochemical properties) in all sequences are shown in white letters on black; those in 80% (60%) are in white on gray (black on light gray). (B) Rooted tree generated after alignment of CASc domains of aforementioned sequences, integrating furthermore human caspases: caspase-12 (CAS12_HUMAN; AAQ88589; $aa_{103–338}$), caspase-10 (CAS10_HUMAN; NP_001221; $aa_{232–471}$), caspase-9 (CAS9_HUMAN;

AAC50640; $aa_{152–414}$), caspase-8 (CAS8_HUMAN; NP_001219; $aa_{242–494}$), caspase-6 (CAS6_HUMAN; NP_001217; $aa_{37–289}$), caspase-5 (CAS5_HUMAN; NP_004338; $aa_{166–415}$), caspase-4 (CAS4_HUMAN; NP_150649; $aa_{69–318}$), caspase-3 (CAS3_HUMAN; AAA65015; $aa_{37–275}$), caspase-2 (CAS2_HUMAN; AAH02427; $aa_{191–446}$), and caspase-1a (CAS1a_HUMAN; NP_150634; $aa_{152–401}$); *X. laevis* caspases: caspase-10 (CAS10_XENLA; AAS91709; $aa_{260–496}$), caspase-9 (CAS9_XENLA; BAA94750; $aa_{136–397}$), caspase-8 (CAS8_XENLA; BAA94749; $aa_{252–495}$), caspase-6 (CAS6_XENLA; BAA94747; $aa_{50–299}$), and caspase-2 (CAS2_XENLA; BAA94746; $aa_{160–413}$); *D. rerio* caspases: caspase-3 (CAS3_DANRE; NP_571952; $aa_{39–280}$) and caspase-2 (CAS2_DANRE; AAG45230; $aa_{163–401}$); *G. gallus* caspases: caspase-9 (CAS9_CHICK; AAL23701; $aa_{143–400}$), caspase-8 (CAS8_CHICK; NP_989923; $aa_{228–480}$), caspase-7 (CAS7_CHICK; XP_421764; $aa_{216–460}$), caspase-6 (CAS6_CHICK; NP_990057; $aa_{49–299}$), caspase-3 (CAS3_CHICK; NP_990056; $aa_{45–282}$), caspase–2 (CAS2_CHICK; Q98943; $aa_{162–417}$), and caspase-1 (CAS1_CHICK; NP_990255; aa_{33-280}); *S. purpuratus* caspases: caspase-7–like (CAS7_STRPU; XP_786037; $aa_{133–380}$) and caspase-6–like (CAS6_STRPU; XP_789185; $aa_{73–325}$); *Salmo salar* caspase-7 (CAS7_SALMO; AAY28975; aa_{1-245}); *C. elegans* caspase Ced-3 (CED3_CAEEL; AAG42045; aa235–495); *Nematostella vectensis* caspase-8–like (CAS8_NEMVE; XP_001626560; $aa_{3–241}$); *Culex quiquefasciatus* (CAS_CULQU; XP_001842236 ; aa_{60-304}); metacaspase of *S. pombe* (METAC_SCHPO; AAG38593; $aa_{130–415}$), and the peptidase C14 containing the caspase catalytic subunit p20 of *Nitrobacter sp.* (PC14_NITROB; ZP_01047279; $aa_{29–241}$), used as an outgroup.

C. Characterization of the poriferan TIR-LRR containing protein (SDTILRc)

1. TIR-LRR containing protein sequence (*SDTILRc*)

The complete TIR (toll-interleukin 1)-LRR (leucine-rich repeat) containing protein cDNA (*SDTILRc*) was isolated from a *S. domuncula* cDNA library (Kruse *et al.* 1997). A degenerate forward primer, 5' – AC/TC/T C/AAC/A A/C/TTT G/AG/AT/A GAC/G G/AG/AI GG – 3' (where I = inosine), was designed against leucine-rich repeat domain (SMART accession number SM00370) and was used in combination with a library-specific primer. *SDTILRc* consists of 2,235 nt (excluding the poly(A) tail) that encompass a CDS (coding sequence) of 1,896 nt (excluding the first stop codon), beginning at $nt_{109–111}$ (ATG/Met_{start}). A Kozak sequence ($nt_{106-112}$) and a polyadenylation signal ($nt_{2210-2215}$) are present. One exception is noticeable in the Kozak consensus (marked in bold) G/A/**T**NNATGG. During Northern blot analyses, a DIG-labeled probe detected a transcript whose size was consistent with that of the cDNA (fig. 35).

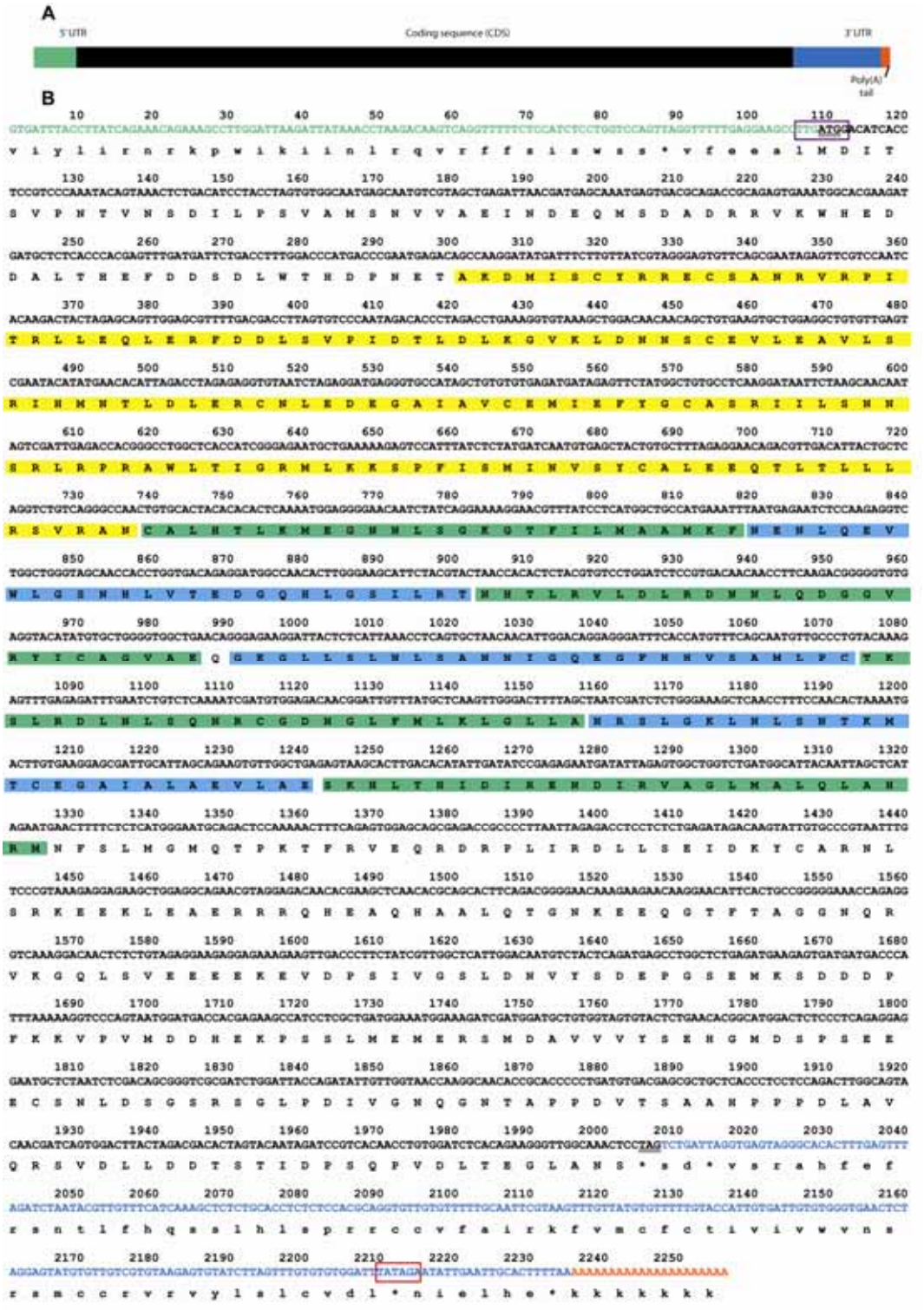


Figure 24: cDNA and deduced amino acid sequence of SDTILRc. The structure of cDNA, including the untranslated regions (UTR), is depicted (drawn to scale, A). The caspase–like nucleotide sequence is translated into a protein sequence (B). The ATG/Met_{start} and the TGA/stop codon are double underlined. The nucleotides are in the 5′ to 3′ direction. The 5′, the 3′ untranslated regions, and the poly(A) tail are in green, blue, and orange letters, respectively. The polyadenylation signal is marked in red box. Deduced amino acid sequence is given below the nucleotide sequence (underlined capital letters). The 7 LRR are highlighted in blue and green; the potential TIR domain in yellow.

2. TIR-LRR containing protein (SDTILRc)

The deduced protein, SDTILRc, comprises 632 aa with an expected size of 72,261 D (fig. 24). SDTILRc shows significant sequence similarity to NLR family, CARD domain containing 3 homologue (NOD3) from *Equus caballus* (*E* value 2 e-18), *B. taurus* (*E* value 4 e-18) or human (*E* value 1e-17). SDTILRc features seven LRR (leucine rich repeat, smart accession number SM00370; fig. 25) and a putative TIR domain (toll interleukin 1-resistance, smart accession number SM00255). LRR consist of 2-45 motifs of 20-30 amino acids. The poriferan LRR (LRR_1 $nt_{210-237}$, LRR_2 $nt_{238-265}$, LRR_3 $nt_{266-293}$, LRR_4 $nt_{295-322}$, LRR_5 $nt_{323-350}$, LRR_6 $nt_{351-378}$, and LRR_7 $nt_{379-406}$) follow the conserved pattern L-x(2)-L-x-L-x(2)-N/C-x-L (Pfam accession number PF00560) with few exceptions marked in bold LRR_1: L-x(2)-L-x-L/**M**-x(2)-N/C-x-L, LRR_2: L-x(2)-L/**V**-x-L-x(2)-N/C-x-L, LRR_4: L-x(2)-L-x-L-x(2)-N/C-x-L/**I**, LRR_5: L-x(2)-L-x-L-x(2)-N/C-x-L/**C**, LRR_6 L-x(2)-L-x-L-x(2)-N/C/**T**-x-L/**M**, and LRR_7: L-x(2)-L/**I**-x-L/**I**-x(2)-N/C-x-L/**I**.

Figure 25: SDTILRc leucine rich repeats alignment. Conserved residues (identical or similar with respect to physicochemical properties) in all sequences are shown in white letters on black; those in 80% (60%) are in white on gray (black on light gray).

The putative TIR (aa_{65-210}) loosely fits these consensus sequences proposed in 2000 by Slack *et al.* (fig. 26). Sequence analyses have revealed the presence of three conserved regions among the different members of the TIR-containing protein family: box 1 (FDAFISY), box 2 (GYKLC-RD-PG), and box 3 (a conserved W surrounded by basic residues). However, deviation from these patterns is not uncommon among this family as presented fig. 26. *In silico* analysis running P2SL reveal a cytoplasmic localization.

Additionally, several major (TRAF2_1) and one minor (TRAF2_2) TRAF2-binding consensus motifs (TRAF2_1, $aa_{194-197}$, $aa_{490-493}$, and $aa_{561-564}$; TRAF2_2,

$aa_{619-624}$) were predicted via ELM algorithms. TRAF2 is a cytosolic protein initiating the TNF (tumor necrosis factor) intracellular signalling.

Further analysis using TUPS (Zhou and Zhou 2003) assume two helical transmembrane segments ($aa_{aa171-188}$ and $aa_{397-413}$).

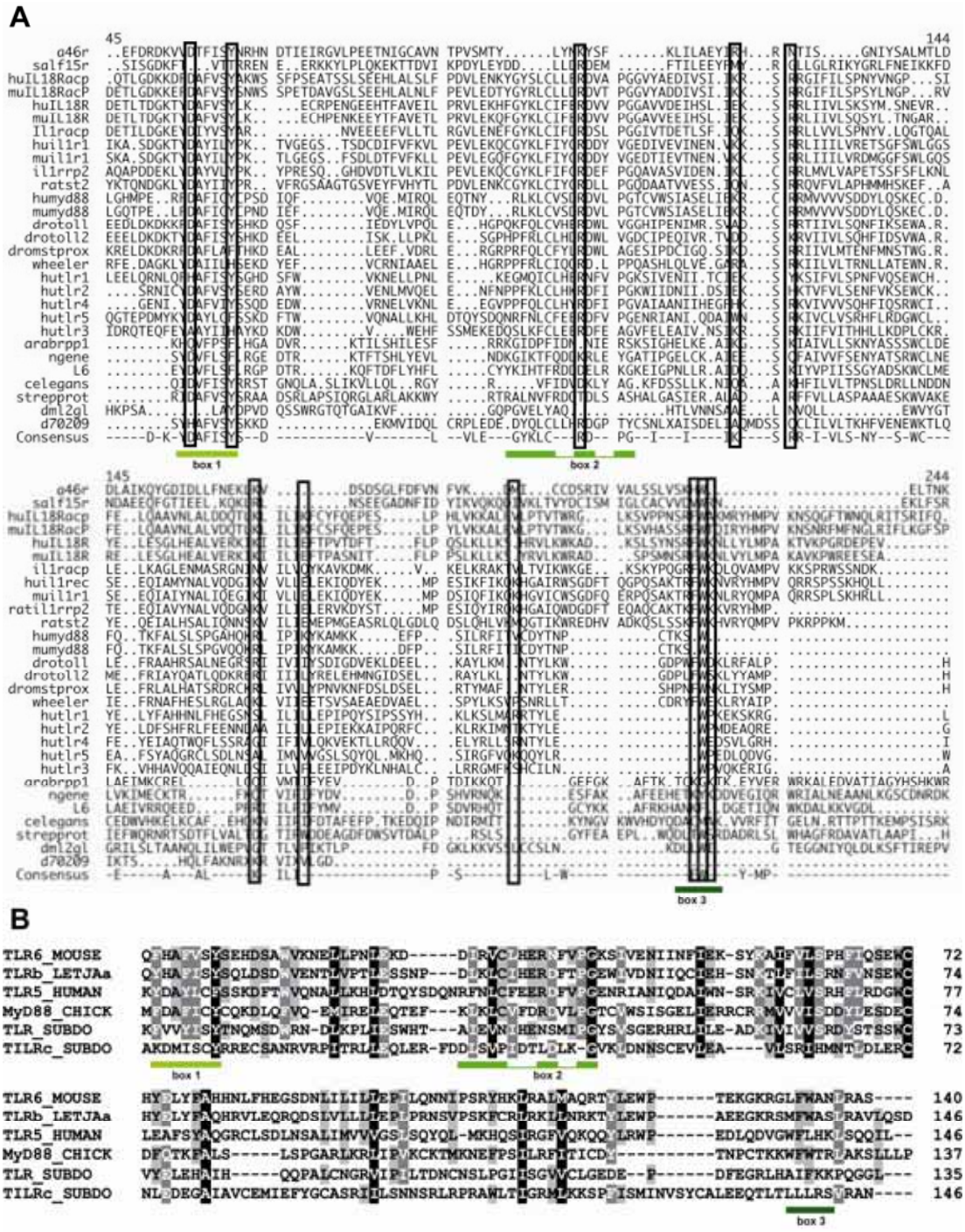


Figure 26: Sequence conservation across the TIR family. (A)-Alignment from Slack *et al.* 2000. Sequences used in the alignment are from: a46R (M72474), salf15R or A52R (Q01220), human IL18RAcP (AAC72196), murine IL18RAcP(AAC72197), human IL18R or IL1Rrp-1 (U43672), murine IL18R or IL1Rrp-1 (U43673), murine IL1RacP (X85999),

human IL1 receptor type I (X16896), murine IL-1 receptor type I (M20658 and M29752), rat IL receptor-related protein 2 (U49066), rat ST2 or Fit-1 or T1 (U04317), human myD88 (U84408), murine myd88 (U84409), *D. melanogaster* toll (P08953), *D. melanogaster* toll2 or Tehao (AF140019), *D. melanogaster* Mstprox (U42425), *D. melanogaster* 18 wheeler (L23171), human Tlr-1(D13637), human Tlr-2(U88878), human Tlr-4(U88880), human Tlr-5 (AF051151), human Tlr-3 (U88879), *Arabidopsis thaliana* disease resistance protein RPP1-WsB (AF098963), tobacco TMV resistance protein N (A54810), *Linum usitatissimum* (flax) rust resistance protein L6 (AF093638), *C. elegans* open reading frame (Z49936), *Streptomyces coelicolor* open reading frame product hypothetical protein SC7H1.23 (AL021411), *Drosophila pseudoobscura* l(2)gl (Q08470), *C. elegans* est d70209 (D70209). The numbers in parentheses are the accession numbers. Slack suggested three conserved regions, named box-1 to -3 (indicated on the alignment). TIR domain of SDTILRc (TILRc_SUBDO; aa_{65-210}) was aligned with the corresponding domain of *Lethenteron japonicum* toll–like receptor b (TLRb_LETJAa; BAE47506; $aa_{651-796}$), *S. domuncula* toll–like receptor (TLR_SUBDO; CAL36105; $aa_{197-331}$), chicken myeloid differentiation primary response factor 88 (MyD88_CHICK; ABQ17966; $aa_{163-299}$), murine toll–like receptor 6 (TLR6_MOUSE; NP_035734; $aa_{652-795}$), and human toll–like receptor 5 (TLR5_HUMAN; NP_003259; $aa_{692-837}$). The three conserved boxes are also indicated. Conserved residues (identical or similar with respect to physicochemical properties) in all sequences are shown in white letters on black; those in 80% (60%) are in white on gray (black on light gray).

3. Phylogenetic relationships of the poriferan TIR-LRR containing protein (SDTILRc)

Due to the unusual structure of SDTILRc protein, the potential TIR domain was considered to be of informative value. Thus, poriferan TIR was aligned with the corresponding domains of representatives of all TIR-containing protein subfamilies (also including an *Escherichia coli* CFT073 TIR domain), in order to study phylogenetic relationships (fig. 27). The alignment has revealed high aa sequence homology to the MyD88 subfamily. SDTILRc shares 13% identical and 30% similar aa with the human protein, 13%/30% with the canine molecule, and 13%/31% with that of chicken. Furthermore, SDTILRc displays 28% similarities with the murine toll–like receptor 6 and the human toll–like receptor 5 or 30% the human toll–like receptor 2. The homology to TIR of other subfamilies is considerably lower, *e. g.*, human toll–like receptor 3 (7%) or *Culex pipiens* MyD88 (7%).

A phylogenetic tree (fig. 27), based on this alignment, was rooted with the distantly related *E. coli* strain. The resulting tree revealed 2 clusters, comprising the MyD88

molecules and a cluster consisting of the toll–like receptors proteins. Both clusters have a common poriferan ancestor.

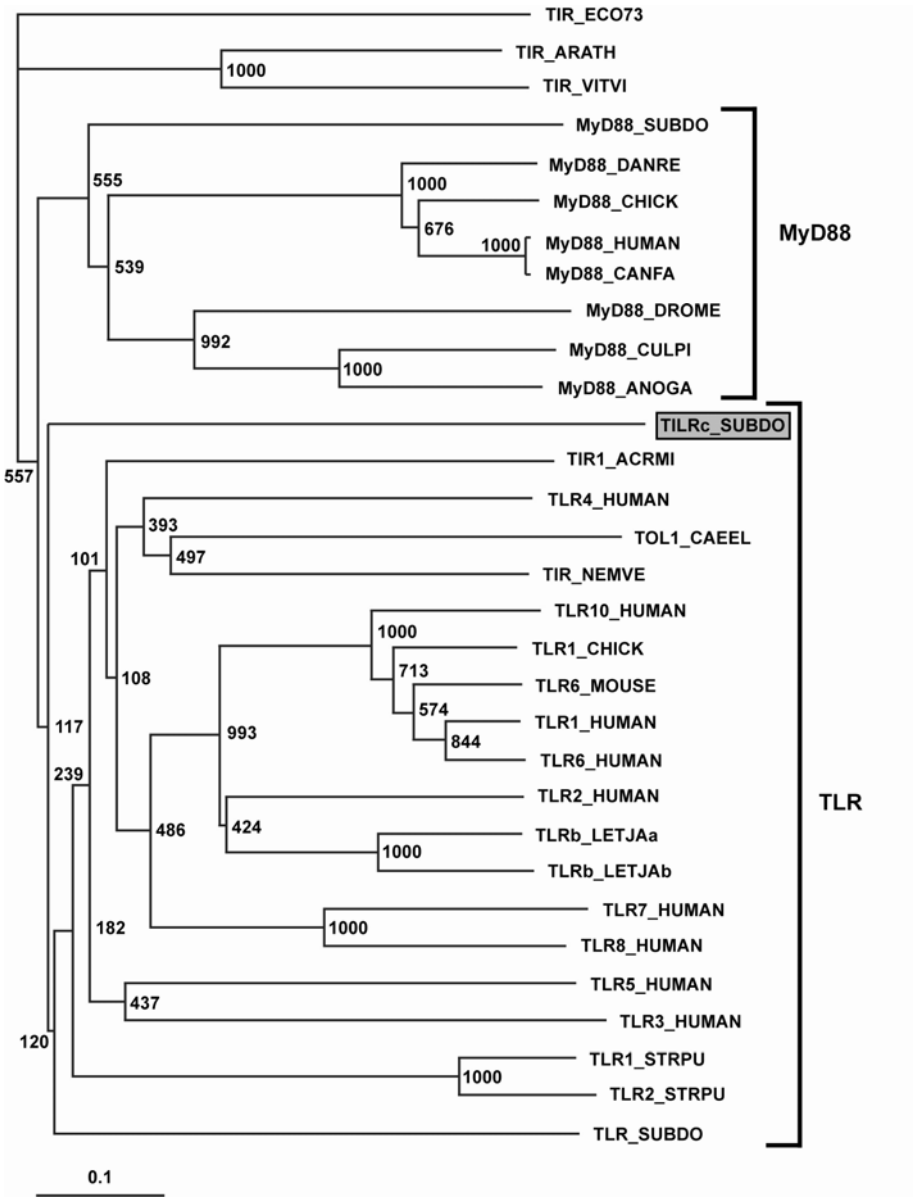


Figure 27: Evolutionary relationship of *S. domuncula* SDTILRC TIR domain to its homologues of Metazoa, fungi, and viruses. Rooted tree generated after alignment of potential TIR domain from *S. domuncula* TILRc and MyD88 TIR domains from human (MyD88_Human; NP_002459; $aa_{173-309}$), chicken (MyD88_CHICK; ABQ17966; $aa_{163-299}$), *Canis familiaris* (MyD88_CANFA; XP_534223; $aa_{160-296}$), *C. pipiens* (MyD88_CULPI; XP_001868621; $aa_{171-304}$), *A. gambiae* (MyD88_ANOGA; XP_314167; aa $_{170-299}$), *D. rerio* (MyD88_DANRE; AAQ90476; $aa_{148-284}$), *S. domuncula* (MyD88_SUBDO; CAI68016; $aa_{115-250}$), and *D. melanogaster* (MyD88_DROME; NP_610479; $aa_{243-375}$); and TIR from toll–like receptor or TIR-containing protein from *S. domuncula* (TLR_SUBDO, CAL36105, $aa_{197-331}$), *Arabidopsis thaliana* (TIR_ARATH; NP_192681; aa_{14-149}), *Vitis vinifera* (TIR_VITVI; CAO16334; aa_{19-159}), human

(TLR1_HUMAN; AAH89403; $aa_{636-779}$- TLR2_HUMAN; NP_003255; $aa_{171-299}$- TLR3_HUMAN; AAH94737; $aa_{755-900}$- TLR4_HUMAN; AAC34135; $aa_{633-778}$- TLR5_HUMAN; NP_003259; $aa_{692-837}$- TLR6_HUMAN; ABY67113; $aa_{641-784}$- TLR7_HUMAN; NP_057646; $aa_{893-1032}$- TLR8_HUMAN; AAF64061; $aa_{897-1043}$- TLR10_HUMAN; AAK26744; $aa_{633-778}$), *S. purpuratus* (TLR1_STROPU; NP_999670; $aa^{740-881}$ and TLR2_STROPU; NP_999671; $aa_{779-913}$), *Lethenteron japonicum* (TLRb_LETJAa; BAE47505; $aa_{660-805}$ and TLRb_LETJAb; BAE47506; $aa_{651-796}$), *C. elegans* (TOL1_CAEEL; NP_001020983; $aa_{1056-1184}$), *Acropora millepora* (TIR1_ACRMI; ABK78770; $aa_{338-475}$), *N. vectensis* (TIR_NEMVE; XP_001637974; aa_{33-150}), mouse (TLR6_MOUSE; NP_035734; $aa_{652-795}$), chicken (TLR1_CHICK; BAD67422; $aa_{647-792}$) and the TIR domain from *E. coli* strain CFT073 (TIR_ECO73; NP_754290; $aa_{171-299}$), used as an outgroup.

D. Evolutionary conserved function(s) of poriferan survivin

So far functional analyses suggest a coevolutionary emergence of the dual role of survivin and metazoan organismal complexity. Thus, in higher Metazoa survivin is both a key regulator of caspases and a chromosomal passenger protein. In order to elucidate the controversially discussed fundamental function of survivin, or if indeed its dual role is conserved through metazoan evolution, functional analyses of survivin were performed in sponges, the taxon closest related to the common metazoan ancestral phylum Urmetazoa. In addition, functional conservation was assessed in a heterologous vertebrate cell model that has been frequently used to explore expression and functional mechanisms of survivin (*e. g.*, Torres *et al.* 2006; Tamm *et al.* 1998).

1. Regulation of proliferation

To elucidate if SDSURVL is involved in cell division, expression of the poriferan molecule was monitored during formation and growth of primmorphs (aggregates of proliferating sponge cells). Concurrently, this function was assessed via heterologous survivin expression in transfected HEK-293 cells.

Primmorphs represent a unique primary sponge cell model that has been used extensively to study poriferan development, proliferation, apoptosis, and cellular aging.

Accordingly, comparative expression analyses were performed, using RNA extracted from primmorphs and sponge tissue. Northern blots illustrate an increasing level of *SDSURVL* transcripts during the first two weeks of primmorph development whereas the control (taken immediately after primmorph formation) reveals a significantly lower expression level. After 14 d, survivin transcription reached a steady-state. On the other hand, in adult sponge tissue survivin transcripts were barely detected. Additionally, *SDCASL2* expression remained at a very low steady state expression level throughout the experiments, in both primmorphs and sponge tissue (fig. 28). Subsequently, *SDSURVL* and *SDCASL2* expression was investigated in aging primmorphs (two weeks of age) via complementary *in situ* hybridization (fig. 28). For this purpose, primmorph cryosections were incubated with labeled sense (control) or antisense probes. Both sense probes did not produce specific signals and generated only low backgrounds. On the other hand, the antisense *SDSURVL* probe detected survivin transcripts abundantly and homogenously throughout the sections, whereas only few *SDCASL2* copies were identified, mainly in the surface cell layer and in cells lining the aquiferous canal system within the mesohyl (internal part of the sponge).

In parallel, a cDNA construct coding for the *S. domuncula* survivin–like protein was introduced into HEK-293 cells. *SDSURVL* expression was confirmed by microscopical inspection (fig. 29) and Western blotting (fig. 31; expected size of 45 kDa corresponding to the SDSURVL-EGFP fusion). Subsequently, the putative regulatory function of SDSURVL in cytokinesis was explored by monitoring the concentrations of *SDSURVL*-transfected (HEK293-*SDSURVL*) and mock-transfected HEK-293 cells, 48 h post-transfection (fig. 31). Whereas mock-transfection provoked no differences in cell concentration compared to the non-transfected control, SDSURVL expression caused a 2.3x increased cell number in HEK293-*SDSURVL* (significance level $P \leq 0.05$). In addition, trypan blue vital staining did not display altered cell viability between control and SDSURVL-expressing populations.

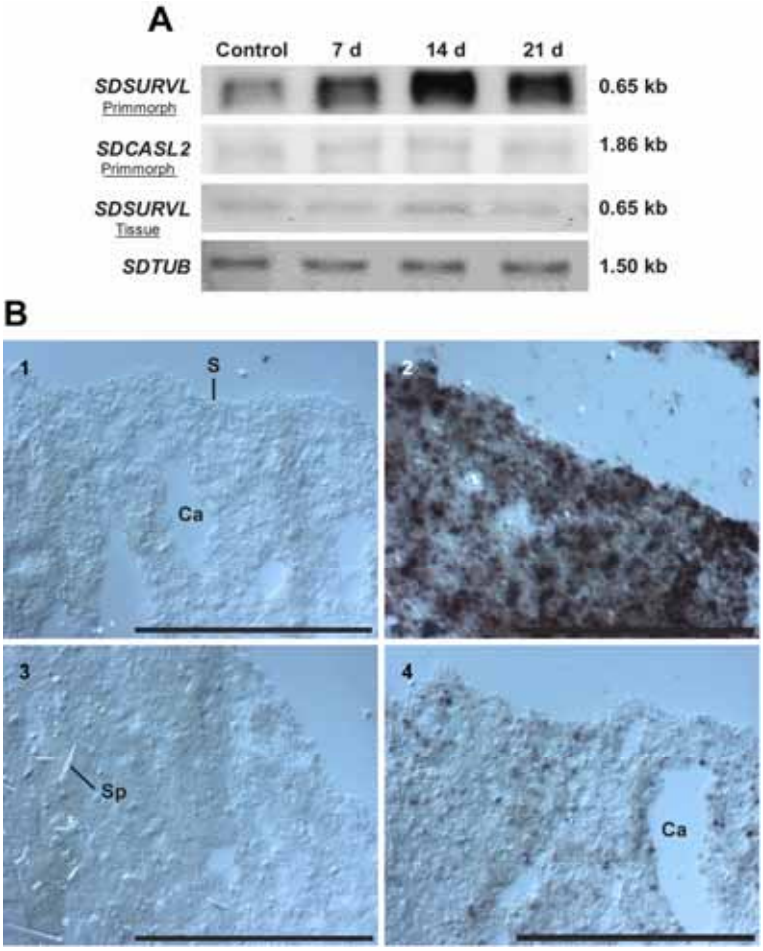


Figure 28: Expression of *SDSURVL* and *SDCASL2* in *Suberites domuncula* primmorphs and adult tissue. (A) Northern blotting analyses. *SDSURVL* and *SDCASL2* transcription was monitored in aging primmorphs and in adult sponge tissue (control, 7, 14, and 21 d). Tubulin transcription was used as an internal control. (B) *In situ* hybridization analyses. Cryosections of primmorphs (two weeks old) were subjected to *in situ* hybridization analyses, using labeled *SDSURVL* (1, 2) or *SDCASL2* (3, 4) probes (1/3, sense (negative control) or 2/4, antisense probe). Hybridized probes were then detected through anti-DIG Fab fragments, conjugated to alkaline phosphatase, and subsequent NBT/BCIP treatment. The sense probes (1, 3) did not produce any signal, whereas the *SDSURVL* antisense probe (2) abundantly and homogenously detected transcripts. The *SDCAL2* antisense probe (4) identified few transcripts in the surface cell layer (**S**) and in cells lining the aquiferous canal system (**Ca**) within the mesohyl, the internal part of the sponge. Sponge spicules (**Sp**) are marked. Scale bar, 1 mm.

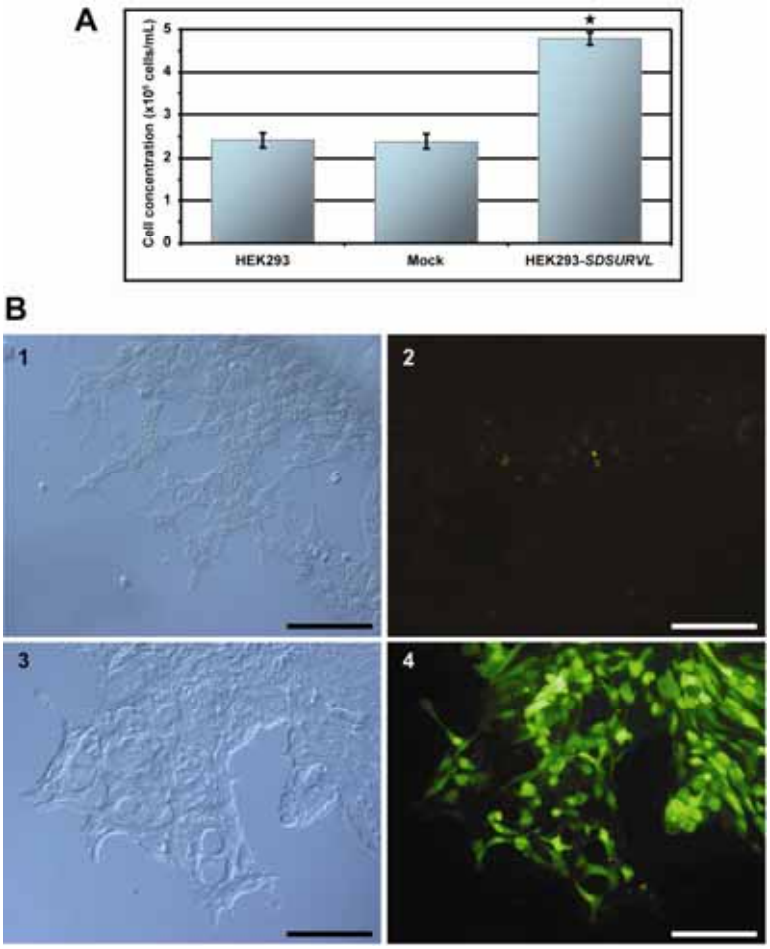


Figure 29: *SDSURVL* expression in HEK-293 cells. (A) Effect of *SDSURVL* expression on cell proliferation. Cells were transfected either with empty vector (Mock) or with a *SDSURVL* construct (HEK293-*SDSURVL*) or remained untransfected (HEK293). HEK293, mock-transfected cells, and HEK293-*SDSURVL* cell concentrations were measured 48 h post-transfection. Proliferation rates of the different controls were comparable, demonstrating that the vector itself had no effect. However, HEK293-*SDSURVL* revealed a higher proliferation rate ($P \leq 0.05$). (B) Histological detection of a GFP-SDSURVL fusion protein in transfected HEK-293 cells. Mock-transfected HEK-293 cells (1, 2) and HEK293-*SDSURVL* (3, 4) under visible (1/3) or UV light (2/4). Scale bar, 100µm.

2. Regulation of apoptosis

To explore the putative implementation of SDSURVL as a prosurvival factor in the programmed cell death signalling cascade, *SDSURVL*, *SDCASL* and *SDCASL2* expression was analyzed in sponge tissue and primmorphs challenged with proapoptotic stimulants. In addition, viability of mock-transfected and HEK293-*SDSURVL* cell lines was assessed following exposure to the same stimuli.

Pam_3Cys-Ser-$(Lys)_4$ and cadmium were employed for their selective ability to trigger extrinsic (the former compound) or intrinsic (later) cell death pathways (Aliprantis *et al.* 2000; Mao *et al.* 2007). Thus, sponge tissue and primmorphs were incubated for 6, 9, and 12 h with either compound. Subsequently, transcripts of *SDSURVL* and *SDCASL2* were detected on Northern blots. Both lipopeptide and cadmium stimulated the expression of poriferan survivin and caspase already after 6 h of challenge (fig. 30). With prolonged incubation time, the expression level of both transcripts increased further. This effect was more pronounced in primmorphs than in tissue samples.

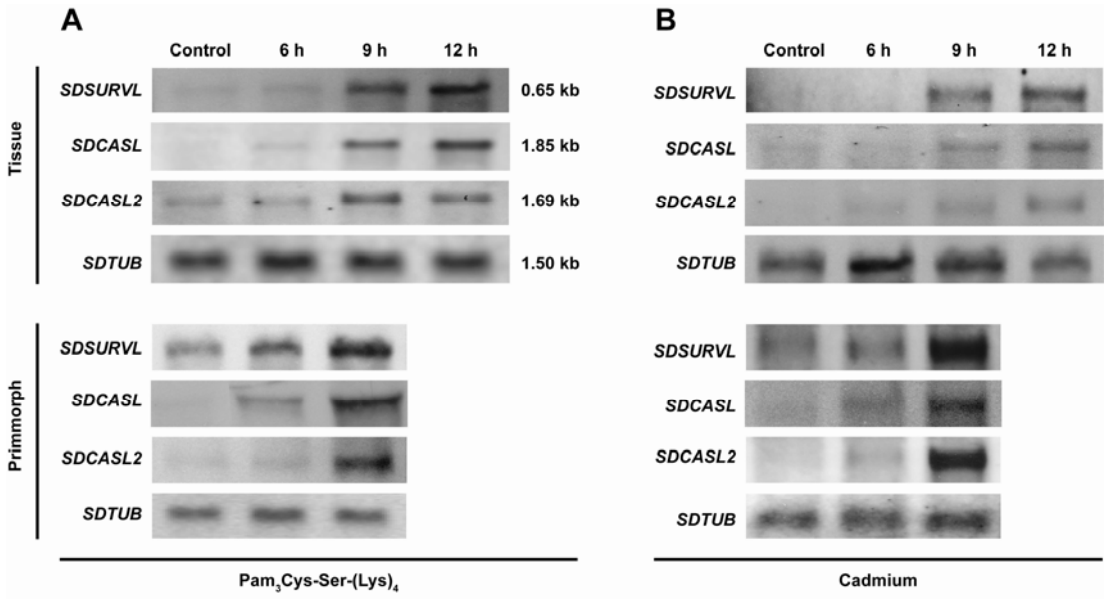


Figure 30: Northern blotting. *SDSURVL* RNA level was followed in primmorph (up to 9 h) and sponge tissues (up to 12 h) exposed to Pam_3Cys-Ser-$(Lys)_4$ (A) or to cadmium (B). RNA was isolated, size separated and blotted. The resulting blots were probed for transcripts of the poriferan survivin–like, caspase–like, caspase–like 2, and tubulin. The latter one was used as an internal control. Following both exposures RNA levels (*SDSURVL*, *SDCASL*, and *SDCASL2*) are enhanced after 6 hours.

To elucidate whether *SDSURVL* expression confers resistance to inducers of apoptosis, mock-transfected HEK-293 and HEK293-*SDSURVL* cell lines were exposed to both apoptosis inducers. MTT assays revealed that the lipopeptide did not alter cell viabilities, neither in mock-transfected nor in HEK293-*SDSURVL* cells, compared to untreated control cells (fig. 31). In addition, the Pam_3Cys-Ser-$(Lys)_4$-treatment did not change the amount of survivin (endogenous as well as poriferan). In contrast, both cell lines were sensitive to cadmium. Thus, MTT assays showed that 46.7% of mock-transfected cells underwent cell death after 6 h of exposure. Nevertheless, *SDSURVL*-expression rescued HEK293-*SDSURVL* since only 25.6% cells were non-viable (fig. 31; significance level $P \leq 0.05$). Furthermore, Western blot analyses revealed an additional effect of cadmium since an increased immunodetection was observed of both, human survivin and its poriferan homologue.

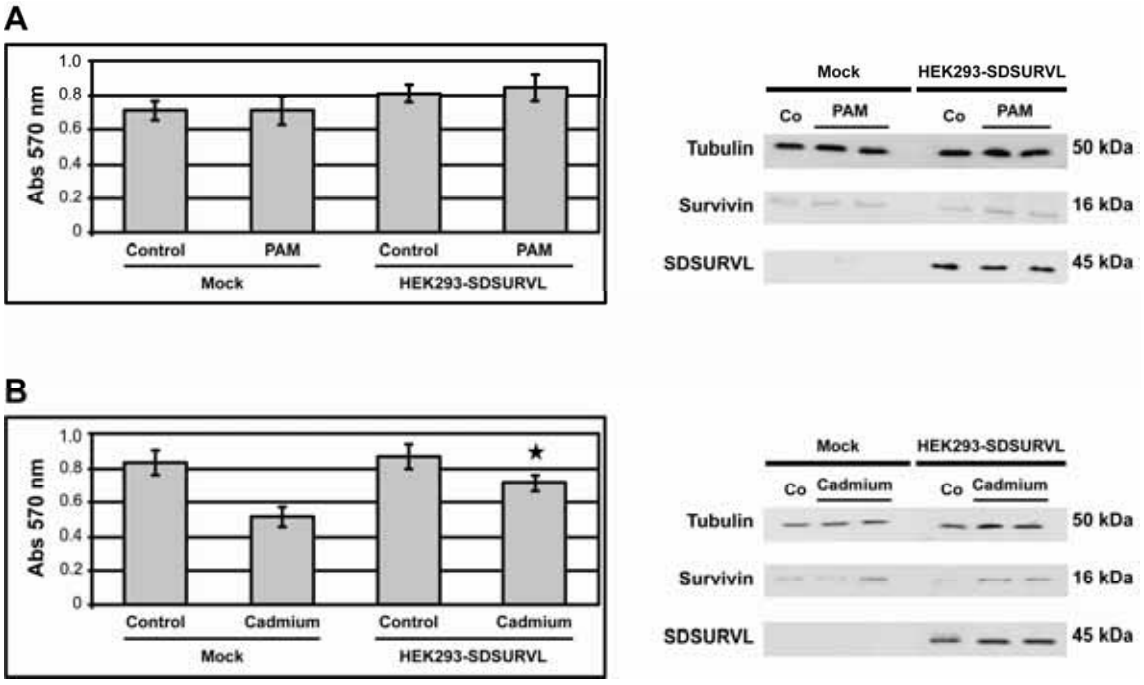


Figure 31: *SDSURVL*-mediated effects on Pam_3Cys-Ser-$(Lys)_4$- and cadmium-challenged HEK-293 cells. Cell viability (MTT assays, left) and protein detection (Western blot analyses, right) following Pam_3Cys-Ser-$(Lys)_4$ (PAM; A) and cadmium exposure (B). Whereas PAM altered neither cell viability nor survivin/SDSURVL concentrations cadmium provoked significant decreased cell viability. However, HEK293-*SDSURVL* cells were statistically less sensitive to cadmium than mock-transfected controls ($P \leq 0.05$). In addition, Western blot analyses registered an increased accumulation of both endogenous and poriferan survivin.

E. Band-shift assay: study of SDCASL double stranded RNA binding domain(s)

In order to establish if the dsrm(s) of poriferan caspase–like is/are functional domain(s) or not, a recombinant protein was produced and studied. To simplify the purification procedure and to avoid hindrance of the CASc domain, only the COOH-terminal extension containing the dsrm(s) was cloned, expressed, and purified under native conditions.

1. Recombinant expression and purification of truncated SDCASL (SDCASLt)

The truncated *SDCASL* cDNA was isolated from a *S. domuncula* cDNA library by means of PCR and degenerate primers (forward primer: 5' – CAA CAG GCT GGA ATT GTA CG – 3', and reverse primer: 5' – GTT ATA GAA ACC TAA CTC AAC ATT AGC – 3'). The sequence thus obtain was cloned into pTrcHis2 TOPO® TA (already containing an initiation codon). Top 10 bacteria were transfected with the recombinant plasmid.

Time-course analysis of SDCASLt (t for truncated) expression (to optimize the expression) and solubility of the target protein were assessed. One prominent band with a molecular mass of about 26.5 kDa was visible 2h after induction. The molecular mass of the recombinant protein confirmed the expected size of 22.8 kDa (deduced from the cDNA) plus 3.7 kDa extra from the expression vector. The optimal induction period was 7h; however, the protein was always found in the insoluble fraction (fig 32).

The recombinant protein needed to be functionally active thus to be purified under native conditions. To increase the amount of soluble protein, bacteria cultures were grown at 4°C for one week. The purification of the recombinant protein was performed under native conditions as presented in the materials and methods section. The presence of the COOH-terminal 6xHis-tag allowed the purification of the recombinant protein using Ni-NTA agarose beads (the protein was purified twice). Purity of the protein was analyzed by 12% SDS-PAGE and additionally with Western blot using anti-His antibody (fig. 33).

Approximately 90 µg/mL of SDCASLt purified under native conditions was used for further analysis.

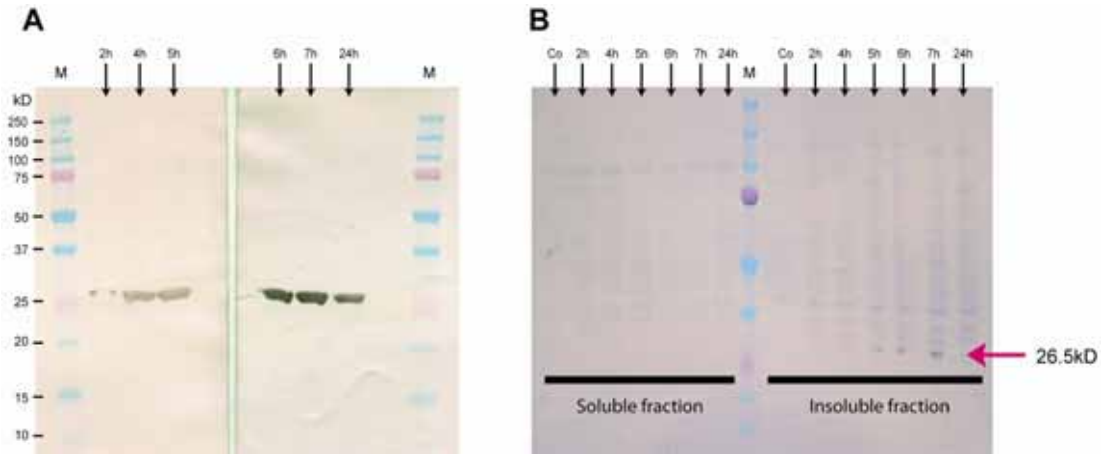


Figure 32: Determination of SDCASLt time-course expression and solubility. To optimize the expression a time-course analysis of the level of protein expression was performed (A). Samples were taken 2, 4, 5, 6, 7, and 24 h after induction. The optimal induction period was established after 7h. However, the recombinant protein was always found in the insoluble fraction (B). The blot was developed using anti-His antibody and alkaline phosphatase conjugated secondary antibody. Precision plus protein™ dual color was used as protein molecular weight marker (M).

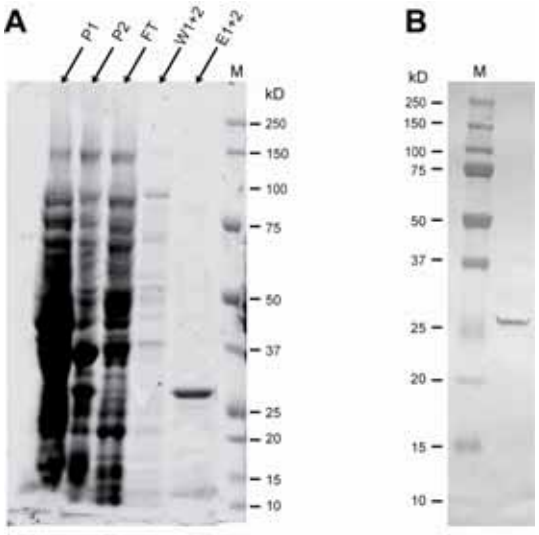


Figure 33: SDCASLt expression under improved conditions. Culture conditions were enhanced and the protein purified under native conditions (A; P1=pellet after BugBuster™ treatment – P2=pellet after lysis buffer for condition treatment – FT=flow-through – W1+2=washing fractions – E1+2=elution fractions – M=precision plus protein™ dual color protein molecular weight marker). Western blot (B, using anti-His antibody) confirmed the correct size and purity of the recombinant protein (about 26.5 kDa).

2. Band-shift assay – study of binding property of SCASL dsrm(s)

The recombinant protein (9 µg) was incubated with either with poly(IC) (polyinosinic-polycytidylic acid; 2 different concentrations: 0.02 and 0.04 µg/µL) or double stranded RNA (27-mer; 3 different concentrations: 0.3, 0.5, and 0.8 pmol) overnight at 4°C. The reaction mixtures were loaded on to a non-denaturing polyacrylamide gel (fig. 34). Two gels were run in parallel: one was stained with Coomassie blue (protein stain) and the other with ethidium bromide (nucleic acid stain).

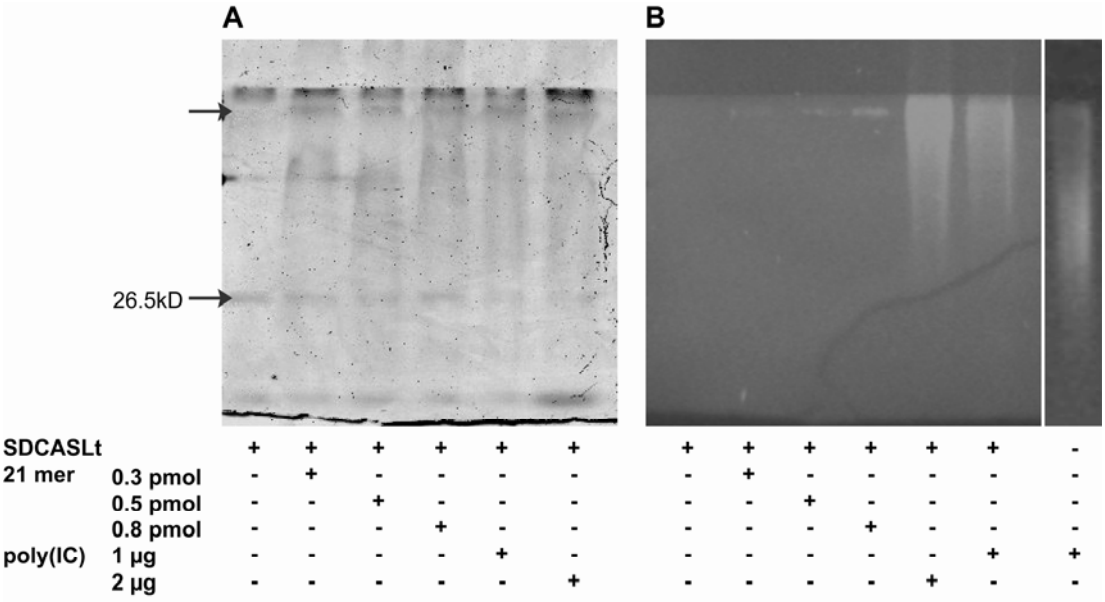


Figure 34: Functional activity of double-stranded RNA binding (dsrm) motif(s) of SDCASLt. Recombinant SDCASLt was incubated either with poly(IC) or a double stranded RNA (27-mer). After electrophoresis (10% resolving gel) one gel was stained with Coomassie blue (A) the other with ethidium bromide (B). An additional band is observable after incubation (upper black arrow).

On the gel stained with Coomassie blue, only one band is visible (26.5 kDa) for SDCASLt without double stranded RNA (lane 1) while two are visible for the other conditions (pink arrows). On the gel stained with ethidium bromide, a band, at the same level as the one observed on the Coomassie stain is observable (additionally, migration of free poly(IC) is faster, last lane on B part). Furthermore the intensity of this band increases

with the amount of double stranded RNA used. Taken together these data suggest that SDCASL double-stranded RNA binding domain(s) is (are) active.

F. Study of TIR-LRR containing protein function

1. *SDTILRc* expression and LPS challenges

In order to elucidate the role of SDTILRc (role in innate immunity or not), *S. domuncula* tissue was challenged with LPS (lipopolysaccharide). LPS are large molecules (lipid + polysaccharide) and are component of the outer membrane of Gram-negative bacteria. LPS are well known molecules which elicit strong immune response (reviewed in Zeytun *et al.* 2007 and Freudenberg *et al.* 2008). Tissues were stressed with two different kind of LPS: from *Escherichia coli* 026:B6 (Sigma) and from isolated bacteria thought to live in symbiosis with *S. domuncula*.

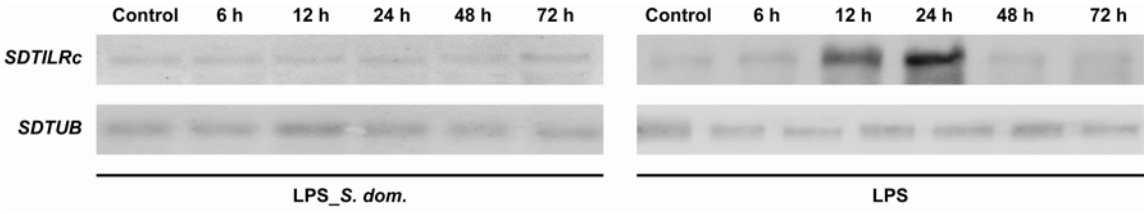


Figure 35: Expression of *SDTILRc* in adult tissue - Northern blotting analyses. *SDTILRc* transcription was monitored in adult sponge tissue (control, 6, 12, 24, 24, 48, and 72 h). Tubulin transcription was used as an internal control. An increase of *SDTILRc* transcription level was observed in the tissues challenged with LPS from *E. coli* (LPS; after 12 h and up to24 h) while the level remained constant in the tissues challenged with LPS from bacteria isolated from *S. domuncula* (LPS_*S. dom*.).

Thus, sponge tissues were incubated for 6, 12, 24, 48, and 72 h with either compound. Subsequently, transcripts of *SDTILRc* were detected on Northern blots. Only LPS from *E. coli* were able to stimulate the expression of poriferan *TILR*c after 12 h of challenge (fig. 35). An increase of transcription level was observable up to 24 h. After 48 h the *SDTILRc* transcription rate came back at its steady-state level. However no variation in

transcription level was observed in the tissues challenged with LPS from bacteria isolated from the sponge.

2. *SDTILRc* expression in RAW-Blue transfected cells

One of the most interesting features of SDTILRc is its Toll/Interleukin-1 receptor (TIR) domain. TIR containing proteins play crucial roles in many host immune responses and ultimately lead to the activation of the NF-kB and AP-1 transcription factors. These factors regulate the inducible expression of the inflammatory cytokines as the TNF (tumor necrosis factor), IL-1 (interleukin) or IL-6.

RAW-Blue cells are mouse macrophage reporter cells which stably express an NF-kB-inducible SEAP reporter gene. The secretion of SEAP is easily detectable and measurable when using QUANTI-Blue™.

These cells provide a rapid and powerful method to monitor the activation of pattern recognition receptors (PRR, such as the toll–like receptors (TLR) and the NOD–like receptor (NLR)). The cells were either transfected with SDTILRc (in pcDNA™3.1/CT-GFP-TOPO®; RAW-*SDTILRc*) or transfected with the empty pcDNA™3.1/CT-GFP-TOPO® vector (mock-transfected). *SDTILRc* expression was confirmed by Western blotting (fig. 36; expected size of 99 kDa corresponding to the SDTILRc-EGFP fusion). Subsequently, the putative function of SDTILRc in innate immunity was explored by monitoring its ability to induce NF-kB activation 24 h after exposure to different concentration of LPS (0 (ddH_2O), 0.1, 1, 10, and 100 µg/mL).

After 24 h hour exposure, the secretion of SEAP was measurable after 1 hour and 30 minutes incubation with QUANTI-Blue™ reaction medium. The SEAP levels were then detected using a spectrophotometer at 620 nm.

No difference in the expression of NF-kB in both cell lines was observable after 24 h exposure neither with ddH_2O nor with 0.1 µg/mL LPS. However, as expected, the other concentrations (1, 10, and 100 µg/mL) revealed a higher expression of NF-kB in RAW-*SDTILRc* than in the mock-transfected cells (fig. 36). With the concentration of 1 µg/mL (10 and 100 µg/mL), the expression of NF-kB is 1.6 times (1.8 times) higher in

the RAW-*SDTILRc* cells than in the mock-transfected cells (statistically different with $P \leq 0.05$).

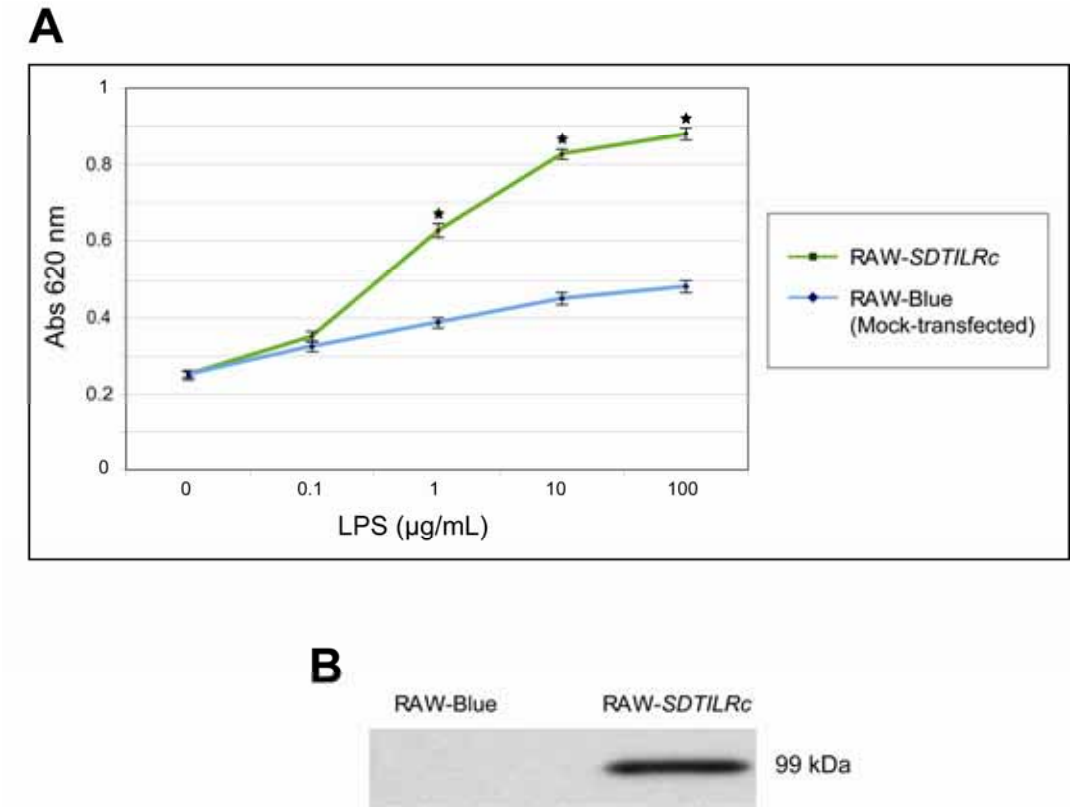


Figure 36: Expression of SDTILRc in RAW-*SDTILRc*. RAW-Blue cells were either transfected with the recombinant pcDNA™3.1/CT-GFP-TOPO® containing SDTILRc (RAW-*SDTILRc*) or the empty vector (RAW-Blue mock-transfected). Expression and the correct size (≈99 kDa) of the recombinant SDTILRc was confirmed via Western blotting (B). Detection and expression of NF-kB-inducible SEAP was followed (A) after different concentration of LPS exposure in both cell lines: in the RAW-*SDTILRc* (green line) and in the mock-transfected RAW-blue cells (blue line). Even if no factual difference was observed after 24 hours of 0.1 µg/mL LPS exposure, expression of NF-kB was around twice higher, and statistically different ($P \leq 0.05$), in transfected cells than the mock-ones with higher LPS concentrations (1, 10, and 100 µg/mL).

VI. DISCUSSION

Apoptosis, or programmed cell death, is a normal component of the development and homeostasis of multicellular organisms. Cells die in a controlled, regulated manner, in response to various stimuli. These features make apoptosis distinct from other forms of death (*e. g.*, necrosis). Necrosis is an uncontrolled cell death, leading to lysis of cells, inflammatory responses and, potentially, to serious health problems. In opposition, apoptosis is a energy dependant process in which cells play an active role in their own death (which is why apoptosis is often referred to as cell suicide).

Inhibitors of apoptosis proteins were initially discovered as baculoviral proteins which had dual function, *i. e.*, inhibition of the host's apoptotic defense system and concurrent enhancement of replication. Since then, cellular members of the ever-growing IAP family have been discovered in a broad range of Metazoa. Within the IAP family survivin represents a unique protein that contains a single BIR domain and a COOH-terminal coiled-coil structure. Thus, it differs from all other family members that comprise 2–3 BIR domains and a canonical COOH-terminal RING finger (fig. 6; Ambrosini *et al.* 1997).

In human, survivin is a singular protein, involved in different pathways (fig. 37). Survivin is implicated in both, extrinsic and intrinsic, pathways of cell death, modulation of p53 cell cycle checkpoint(s), control of spindle formation, and proper kinetochore attachment during cell division. Besides, survivin has other controversial implications: in the endothelial cell viability during tumor angiogenesis and in the participation of cellular stress response via its association with the molecular chaperone Hsp90 resulting in the accumulation of misfolded/damaged proteins. In a short way, in higher Metazoa, survivin has been suggested as a regulator of both apoptosis and cytokinesis (Altieri *et al.* 1999; reviewed in Reed and Bischoff 2000; Zangemeister-Wittke and Simon 2004). Thus, it not only controls cell cycle checkpoints and correct completion of cytokinesis but also regulates the cascade–like activation of proapoptotic proteases (caspases). However, the evolutionary conservation of this dual role remains heavily debated since in lower Metazoa survivin

homologues appear to exhibit an exclusive role in cytokinesis (reviewed in Li 2003). In addition, even though the yeast IAP (BIR1) was proposed to be a close homologue of metazoan survivin (Uren *et al.* 1999; Li *et al.* 2000) it bears two BIR domains and is considerably larger in size than other survivin proteins. BIR1 is required for efficient cell division and exhibits antiapoptotic activity. However, yeast lacks many components of the metazoan apoptotic machinery, including caspases, indicating different regulatory mechanisms.

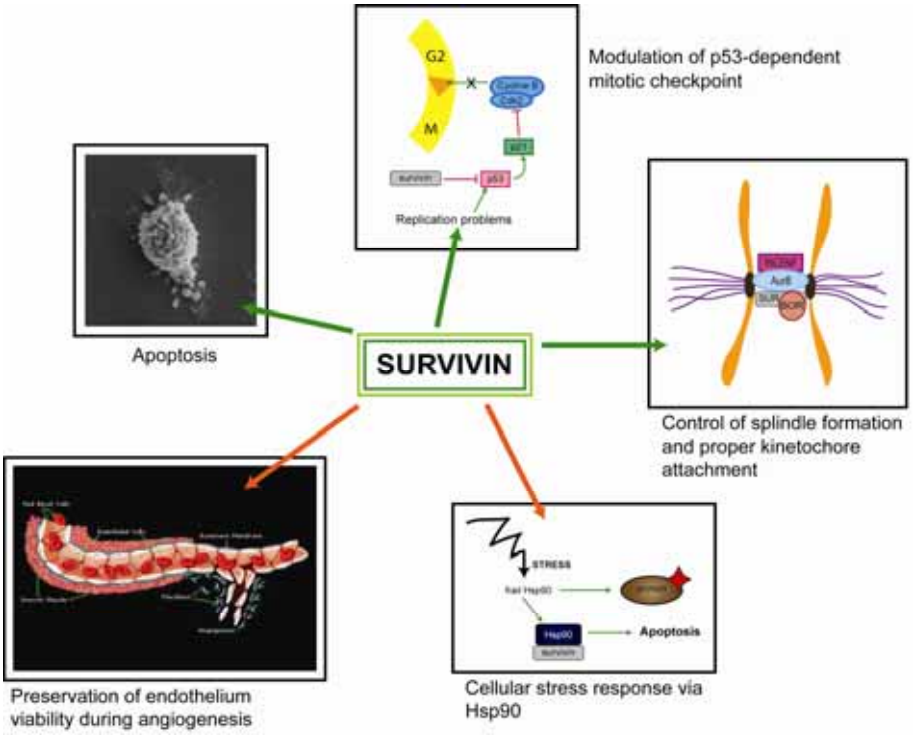


Figure 37: Roles of human survivin. Survivin is an inhibitor of apoptosis implicated in cytokinesis and cell proliferation. Survivin is also involved in other more controversial mechanisms, such as preservation of endothelium viability during cancer angiogenesis and in the accumulation of aberrant proteins.

Key regulators of apoptosis and cell cycle progression have already been discovered in Porifera, the closest animal phylum related to the common metazoan ancestor. However, identification of poriferan survivin provides an important tool not only to investigate the primordial function(s) of this versatile regulator protein but also to elucidate the putative ancient and intimate link between apoptotic cell death and proliferation.

The survivin homologue newly discovered in sponges reveals a significant degree of aa conservation (table 1), in particular to the human sequence (30% identical and 48% similar aa), while *D. melanogaster* deterin displays a lower sequence conservation (22%/40%). Interestingly, the homology to the placozoan putative survivin protein is negligible (9%/15%).

SURVIVIN	Virus NPVEP	Fungi SCHPO	Poriferan SUBDO	Cnidarian NEMVE	Echinoderm STRPU	Arthropod DROME	Chordate HUMAN
Virus *Epiphyas PNPV*		5%	**11%**	11%	10%	12%	11%
Fungi *S. pombe*	8%		**3%**	3%	3%	3%	3%
Poriferan *S. domuncula*	**17%**	**6%**		**25%**	**24%**	**21%**	**26%**
Cnidarian *N. vectensis*	19%	5%	**43%**		41%	35%	40%
Echinoderm *S. purpuratus*	18%	5%	**44%**	60%		31%	40%
Arthropod *D. melanogaster*	21%	5%	**41%**	49%	49%		30%
Chordate *H. sapiens*	21%	5%	**43%**	62%	60%	45%	

Table 1: Comparison of amino-acid survivin sequences from different organisms. Similarity between sequences is quantified via "percent identity" (green) or "percent similarity" (blue).

Before the emergence of bilateral animals Cnidaria and Placozoa split from the main animal lineage, but the order of divergence remains uncertain (*e. g.*, Collins 1998; Halanych 2004 - Appendix 2; Dellaporta *et al.* 2006). Many phylogenetic studies place Placozoa within or sister to Cnidaria (Bridge *et al.* 1995; Siddall *et al.* 1995; Kim *et al.* 1999). Consistently, in the phylogenetic tree of this study Placozoa and Cnidaria are grouped together, although the resulting clade is less supported than most other branches. Furthermore, the position of Ecdysozoa close to the poriferan phylum might be a result of long-branch attraction, due to high nucleotide substitution rates (rapidly evolving lineages), that artificially places this group at a basal metazoan position. The *S. domuncula* (Porifera) sequence is located at the base of the metazoan clade whereas viral sequences form a well-

supported group at the other extremity of the radial phylogenetic tree, followed by the fungal clade. The relative topological relationships confirm the common ancestry of metazoan (including poriferan), fungal, and viral survivin–like proteins, concurrently advocating Porifera as a suitable model to study the functional diversification of survivin during metazoan evolution.

SDSURVL features all characteristics of survivin molecules, *i. e.*, BIR domain, NES, and coiled-coil structure that facilitate protein-interaction and mediate both antiapoptotic activity and proper kinetochore attachment (Johnson *et al.* 2002). However, SDSURVL displays several additional putative protein-interaction domains that are known to be involved in cell cycle regulatory functions: (**i**) APCC-D box, required for ubiquitination and proteasome-mediated degradation via the anaphase-promoting ubiquitin ligase complex (APC/C); (**ii**) LIG_FHA1/2 motifs (aa_{98-104}; $aa_{100-106}$) that bind regulators of the G2/M checkpoint; (**iii**) USP7-binding motif (aa_{33-37}) that is a target of the UPS7 deubiquitinating enzyme, whose substrates are regulators of cell survival pathways. Ubiquitination and deubiquitination regulate numerous cellular processes (*e. g.*, p53; Birks *et al.* 2008). Ubiquitin is a well-conserved 76-amino acid protein and is covalently attached to the amino group of internal Lys residue of target proteins by an enzymatic cascade of ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), ubiquitin ligases (E3), and additional novel ubiquitination factors (E4), if necessary. Ubiquitylated proteins can escape proteasomal degradation via deubiquitinating cysteine protease enzymes, *e. g.*, ubiquitin specific peptidase 7 (USP7; *i. e.*, herpes virus-associated ubiquitin-specific protease or HAUSP). Thus, USP7 gene may play an important role in carcinogenesis regulating the stability or activity of specific cellular proteins.

The gene structure (intron/exon positions, intron phases) is mostly conserved between poriferan and vertebrate survivin with some differences concerning the promoter region. The TATA-less vertebrate promoter includes CpG islands and several Sp1, CDE, and CHR sites. By analogy with other cell cycle-regulated genes it was postulated that Sp1 regulates basal transcriptional expression of the survivin gene, further modulated by CDE/CHR elements, thus imparting cell cycle periodicity of survivin expression (Li and Altieri 1999a). The poriferan promoter features a canonical TATA box, two Sp1 sites, and

misses CDE/CHR elements. Therefore, considering the aforementioned protein interaction motifs, regulation of the poriferan survivin pool might also occur not only on transcriptional level but also on a posttranscriptional/proteasomal level, e. g., via (de)ubiquitination or complex formation with other regulatory proteins (Yang *et al.* 2000). In higher metazoans, regulation of survivin occurs at various levels, including transcription, differential splicing, protein degradation, and intracellular sequestration via different ligands. Transcription of survivin can be directly or indirectly activated via NF-kB (reviewed in Van Antwerp *et al.* 1998), via insulin growth factor I/mTOR pathway (Vaira *et al.* 2007), via the members of the Ras oncogen family, signal transducer and activator of transcription 3, and the antiapoptotic factor Wnt-2 (Sommer *et al.* 2003; Aoki *et al.* 2003). Conversely, survivin is transcriptionaly repressed by wild type p53, p75Rb, and E2F2 (Hoffman *et al.* 2002; Zhou *et al.* 2002; Raj *et al.* 2008). An additional level of complexity in the regulation of survivin is the different expression patterns of survivin alternative splice variants (Noton *et al.* 2006). Ultimately, survivin degradation occurs in the G1 phase (cell cycle) via the ubiquitin-proteasome pathway (Zhao *et al.* 2000), but is stabilized when bound to Hsp90, helping tumor cells to elevate their antiapoptotic threshold and promoting their proliferation (Altieri 2004).

Survivin is expressed in most common human tumours, thus representing a highly suitable prognostic factor and target for cancer therapies. However, survivin is also expressed in fast dividing cells of foetal tissue during the G2/M phase but not detectable in terminally differentiated cells. In addition, survivin is required to maintain proliferation of T cells and hematopoietic cells, concurrently antagonizing apoptosis (Fukuda *et al.* 2003). Survivin exerts its antiapoptotic function mostly via inhibiting caspases, in particular caspase-7 and -3. However, the dual role of survivin in lower Metazoa remains controversial. Thus, *C. elegans* survivin homologue BIR-1 exhibits an exclusive role in cytokinesis whereas *D. melanogaster* deterin deters cells from apoptosis.

To approach this controversy, comparative functional studies of poriferan survivin were performed. Accordingly, the implementation of SDSURVL in cell proliferation and cell death was studied in the most basal metazoan phylum as well as in a heterologous HEK-293 cell model. This strategy had already been successfully adopted to assess the

antiapoptotic potential of deterin (Hawkins *et al.* 1998; Wenzel *et al.* 2000). In contrast to sponge tissue that displayed a very low steady-state expression level, proliferating cells of primmorphs contained a considerable amount of *SDSURVL*, indicative of its involvement in cell division. This assumption was supported by a significantly increased proliferation rate of *SDSURVL*-expressing HEK-293 cells. In addition, application of two inducers of cell death (cadmium, a bivalent metal and Pam3Cys-Ser-(Lys)4, an analog of the NH_2-terminal portion of bacterial lipoprotein) revealed the antiapoptotic potential of SDSURV since both compounds triggered the expression of *SDSURVL* in sponge tissue and primmorphs, probably a countermeasure to the preceding induced expression of the caspase-7 homologue *SDCASL* and *SDCASL2*. Concurrently, cadmium induced apoptosis in HEK-293 cells, though concomitant expression of *SDSURVL* drastically reduced the occurrence of cell death (≈ 50%). However, Pam3Cys-Ser-(Lys)4 incubation did not alter HEK-293 cell viability, probably due to the different repertoire of Toll–like receptors (mediating lipopeptide-induced innate immunity and cell death) found within HEK-293 and sponge cells.

Cytotoxicity mediated by cadmium is associated with the increase of reactive oxygen species (ROS) production (Watanabe *et al.* 2003; Cao *et al.* 2007) and the decrease of mitochondrial transmembrane potential leading to the release of mitochondrial proapoptotic factors into the cytosol (*e. g.*, Smac/DIABLO, cytochrome C, and endonuclease G; Mao *et al.* 2007). These factors concomitantly cause activation of caspase-3/7 (reviewed in Ozben 2007), which is considered as a biochemical hallmark of the apoptotic cascade. However, survivin interacts with various caspases and thus inhibits further progression of the apoptotic cascade. On the other hand, released Smac/DIABLO is a negative regulator of survivin that neutralizes this antiapoptotic effect and consequently restores caspase activity (reviewed in Verhagen and Vaux 2002).

The aforementioned observations in combination with significant similarities of sequence and structure between SDSURVL and its vertebrate counterparts propose a common regulatory function in the interconnected pathways of cell cycle and apoptosis. In HEK-293 (fig. 38) cells poriferan survivin might complement the endogenous molecule by sequestrating its proapoptotic binding partners (such as Smac/DIABLO and caspases),

consequently delaying cell death. In addition, the presence of putative target sites for ubiquitination might explain the elevated levels of human and poriferan survivin detected in challenged transfected cells. Accordingly, this effect does not result of an induced expression but is indicative of decreased proteasomal degradation of either protein. Therefore, the poriferan molecule replenishes the pool of endogenous survivin, concomitantly preventing degradation of its human counterpart by engaging/silencing elements of the proteasome.

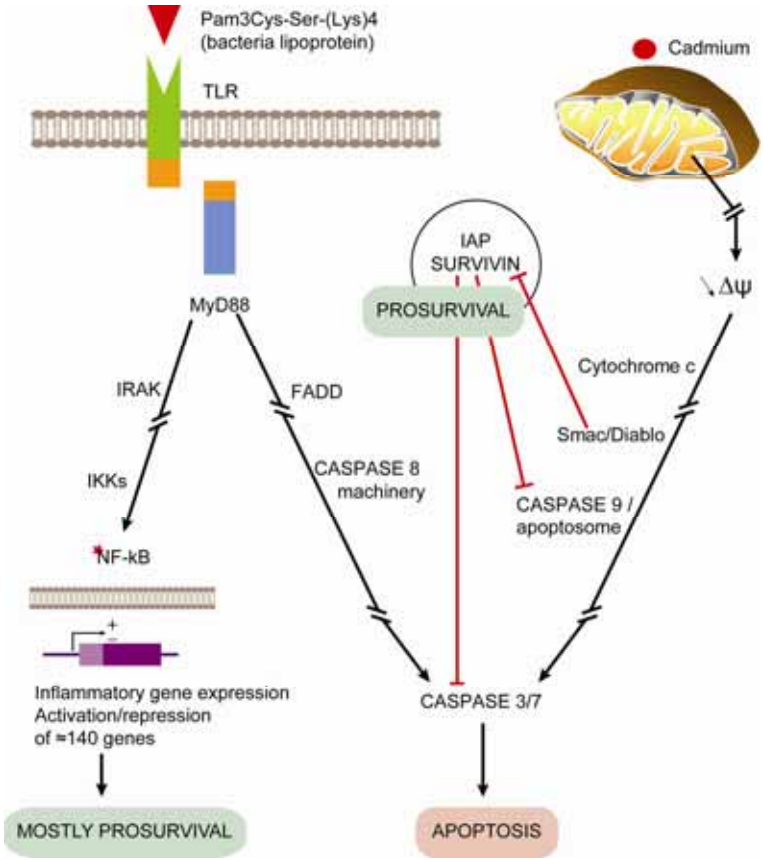
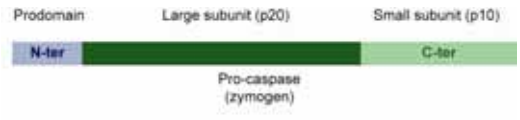


Figure 38: Deduced and proposed function of *S. domuncula* survivin in apoptosis. Survivin exerts its antiapoptotic properties via inhibition of intrinsic and extrinsic pathways of cell death. In case of Pam3Cys-Ser-(Lys)4 challenge, the extrinsic pathway is activated while cadmium mediates the intrinsic pathway. Survivin inhibits caspase processing and activation via direct or indirect interaction. Besides, Smac/Diablo is able to sequester survivin.

Furthermore, concurrent with previous observations during which elevated levels of survivin occur in most human tumor cell types and overexpression of survivin in transfected HEK-293 increases cell proliferation (Tamm *et al.* 1998; Torres *et al.* 2006),

SDSURVL expression stimulates cell division. Consequently, the protein might promote/accelerate completion of various stages of cell proliferation, from proper kinetochore attachment to spindle formation.

Caspases belong to the cysteine protease family and cleave their substrates after aspartic residues (cysteine aspartase). Caspases can be grouped in various ways. They are usually grouped according to their substrate specificity or their functional roles but also to their prodomain structure (fig. 39). Cytoplasmic caspases are initially inactive (zymogen) and called procaspases. These procaspases are then proteolytically activated to form mature proteins. Processing of the procaspases at specific aspartase residues involves the removal of the variable length NH_2-terminal peptide or prodomain and cleavage between the small and large subunits. These cleavages result in a heterodimeric active enzyme (fig. 4).



Caspase n°	Prodomain		Optimal cleavage	
	Structure	Domain	Site	Group
1	X	CARD	W/Y/FEHD	I
2	X	CARD	DEHD	II
3	O	Ø	DEVD	II
4	X	CARD	W/L/FEHD	I
5	X	CARD	W/L/FEHD	I
6	O	Ø	V/T/IEHD	III
7	O	Ø	DEVD	II
8	X	DED	LETD	III
9	X	CARD	LEHD	III
10	X	DED	?	III
12	X	CARD	?	I
13	X	CARD	?	I
14	O	Ø	DEVD	II

Figure 39: Caspases classification. Caspases can be classified according to their prodomain (X=presence/O=absence) or their cleavage site. Two motifs are found within caspase long prodomains (X): CARD (caspase activation and recruitment domain) and DED (death effector domain). Three classes of caspases have been distinguished according to their optimal substrate cleavage sites: the group I (yellow) includes the mediators of inflammation, the groups II (orange) and III (green) the effector and activator caspases (respectively) involved in apoptosis.

To date, thirteen caspases have been isolated in human. Furthermore, caspases are conserved across Metazoa with functional homologue (*i. e.*, orthologues) in *D. melanogaster* (6 members), *C. elegans* (4), *Xenopus*, and *S. purpuratus*. Besides, several proteases, or "caspase–like", have been identified in plants, fungi, and bacteria. They are termed meta- and paracaspases (Uren *et al.* 2000). Even if their functions remain unknown, these data suggest evolution from a conserved protease superfamily.

SDCASL and SDCASL2 are two new caspases which have to be added to the already identified poriferan cysteine proteases (*Suberites domuncula* caspase-related protein, accession number AJ426651 and *Geodia cydonium* caspase 3 AJ417903 and CAC83013). Both caspases putatively represent such proteases, comprising the 2 subunits p20 and p10 and fairly fit the established patterns of the cysteine and histidine active sites.

CASc	SUBDO CASL	SUBDO CASL2	STROPU CAS7	XENLA CAS7	SALMO CAS7	HUMAN CAS7	HUMAN CAS3
SUBDO CASL		33%	18%	20%	18%	19%	16%
SUBDO CASL2	52%		22%	22%	21%	20%	20%
STROPU CAS7	35%	38%		43%	44%	46%	41%
XENLA CAS7	34%	34%	62%		74%	69%	56%
SALMO CAS7	36%	35%	64%	89%		72%	55µ
HUMAN CAS7	36%	36%	64%	83%	84%		54%
HUMAN CAS3	36%	36%	62%	72%	70µ	71%	

Table 2: Comparison of amino-acid CASc sequences from different organisms. Similarity between sequences is quantified via "percent identity" (green) or "percent similarity" (blue).

The phylogenetic tree, obtained after alignment of the poriferan CASc with the corresponding metazoan and yeast corresponding domains, reveals more clusters of caspase subfamilies than real caspases evolution. However, the positioning of the poriferan molecules CASL and CASL2 at the origin of the well-supported clusters reflects the evolutionary ancient character of these apoptotic proteases, thus strongly supporting the monophyly of classic caspases.

Both caspases reveal higher similarity for caspase-7 of various species (table 2) and have small prodomains so differ from the initiator caspases (fig. 22 and 39). Consequently, it can be said that SDCASL and SDCASL2 belong to the group of downstream, effector, caspases. However, SDCASL and SDCASL2 carry an uncharacteristically large COOH-terminal stretch of 203 and 254 aa (for SDCASL and SDCASL2 respectively) following the p10 subunit.

Even more intriguing is the presence of several RNA binding domains in these regions, which, except this feature, reveal no similarity to any published sequence. The COOH-terminus of SDCASL contains two dsrm (double-stranded RNA binding motif, Pfam accession number PF0035, fig. 40) $aa_{303-373}$ (*E* value 2.9 e-01) and $aa_{405-473}$ (*E* value 4.0 e-01). Furthermore, a motif with a weak homology to a K homology RNA-binding domain is also found ($aa_{341-469}$; *E* value 1.72 e+03). Similarly, a single dsrm is noticeable in the COOH-terminal portion of SDCASL ($aa_{441-514}$; *E* value 4.14 e+002).

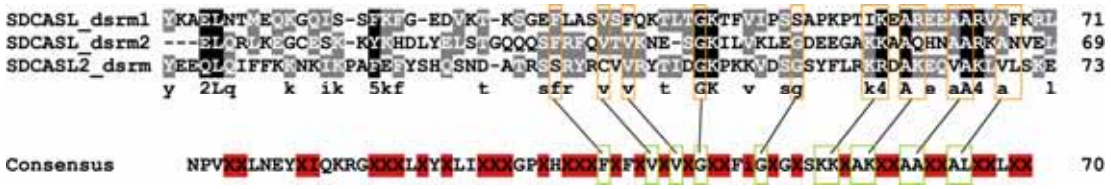


Figure 40: Alignment of SDCASL and SDCASL2 double-stranded RNA binding motifs (dsrm). The consensus sequence from the 3 poriferan dsrm is shown below the alignment (where 2=E or Q, 4=K or R, and 5=F, Y, or W), followed by the general consensus sequence. Accordingly to Krovat and Jantsch (1996), poriferan dsrm can be considered as type B dsrm: they are conserved at their basic COOH-terminal but fit the overall consensus at their NH_2-terminal only poorly.

The role of SDCASL and SDCASL2 in apoptosis had already been highlighted during cadmium and Pam3Cys-Ser-(Lys)4 challenges (upregulation after 6 h exposure). However, potential implementation of SDCASL and SDCASL2 in pathogen sensing (via their dsrm motifs) had to be checked.

Thus, to study poriferan dsrm(s), protein-RNA interactions were explored via gel or band shift assay. SDCASL was preferred to SDCASL2 because of its higher number of domains and *E* values. Furthermore, to simplify the purification procedure and to avoid hindrance of the CASc domain, only the COOH-terminal extension containing the

dsrm(s) was cloned, expressed and purified under native conditions. An efficient binding was observable for poly(IC) (or polyinosinic-polycytidylic acid, a synthetic double-stranded RNA that is used experimentally to model viral infections in vivo) but also for a 27-mer double stranded RNA (siRNA), which is in accordance with the previous reported ability of dsrm to interact with as little as 11 bp of dsRNA (Bycroft *et al.* 1995).

SDCASL dsrm appear not to recognize specific nucleotide sequences, as already put forward by Krovat and Jantsch in 1996. They proposed that dsrm interacts primarily with A-form double helix RNA which differs from the typical dsDNA B-form helix. The dsrm is mainly basic and a helical conformation is predicted at its COOH-terminal which may make direct contact with the negatively charged polynucleotide (Green and Mathews 1992; McCormack *et al.* 1992). Accordingly, SDCASL dsrm are 71 and 69 aa long, encompass 13 strongly basic aa (K and R), their pI are 10.5 and 9.1, and helix regions are predicted at their COOH-termini ($aa_{36\text{-}58}$ and $aa_{38\text{-}59}$).

The dsRNA binding proteins (DRBP) have been identified in eukaryotes, prokaryotes, and viruses and exhibit various but critical functions especially in host defense (reviewed in Saunders and Barber 2003). Nuclear eukaryotic DRBP (*e. g.*, DICER (*S. pombe*), nuclear factors associated with dsRNA (NFAR), and adenosine deaminase acting on RNA (ADAR)) may function in RNA interference (RNAi), mRNA elongation, editing, stability, splicing, and export. In contrast, cytoplasmic DRBP (SDCASL and SDCASL2 location prediction) *e. g.*, double stranded RNA-dependent protein kinase (PKR) and protein kinase, interferon inducible double stranded RNA dependent activator (PACT) have been reported to regulate the translation, and to be involved in dsRNA signalling events, and host defense (reviewed in Clemens and Elia 1997; Patel and Sen 1998).

Double-stranded RNA (dsRNA) forms the genetic material of some viruses (RNA viruses replicating their genomes through complementary strands) and can trigger host's defense mechanisms, such as RNA interference in eukaryotes and interferon response in vertebrates (Schultz *et al.* 2004). Millions of virus–like particles are present in every millilitre of ocean water (Bergh at al. 1989). The abundance of viruses exceeds that of bacteria and Archaea by approximately 15-fold. However, because of their extremely small

size, viruses represent only approximately 5% of the prokaryotic biomass (reviewed in Suttle 2007). Sponges as sessile filter feeders pump a volume of water equal to its body volume once every 10 seconds. Therefore they are in contact with numerous viral particles, have high probability to be infected, and very likely develop sponge-specific mechanisms. Their survival depends on rapidly mounted defense responses. Defense strategies to thwart viral infection, to prevent viral replication, or viral dissemination, usually involve the induction of programmed cell death or apoptosis (reviewed in Barber 2001). These poriferan uncharacteristic caspases may form a new/ancestral class of cysteine protease with sponge-specific function (fig. 41). Thus, SDCASL and SDCASL2 may represent an effective strategy, their dual role owing to a dual complementary role in apoptosis (**i**) and pathogen sensing (**ii**).

Invertebrates embrace an impressive diversity of body forms and lifestyle reflecting independent, disparate, and long evolutionary history. Thus, as invertebrates have colonized specific ecological area, they confronted with specific pathogens which are attempting to disrupt their self-integrity. So it is not fanciful to write that specific environment implicates specific microorganisms attack thus specific defense responses molecules. As an example, approximately eight Toll receptors are present in *D. melanogaster* and one without direct role in pathogen response has been found in *C. elegans* (Kurz and Ewbank 2003). The defense system sharing striking similarities among vertebrates, invertebrates, and plant is the Toll/interleukin-1 receptor family, thus it is surprising that the defense mechanisms are so different between two ecdytozoans. Furthermore, while the defense pathways have the same objective (preservation of integrity, elimination of pathogens) several component of the sensing/defense pathway are different inter and intra-species. Accordingly, several molecules are surprisingly involved in *C. elegans* innate immunity, such as p38MAPK (p38 mitogen-activated protein kinase), programmed cell death, and tissue growth factor-b–like pathways (Alegado *et al.* 2003). Alike, a large number of C-type lectins and some molecules of a family of peptidoglycan-recognition proteins may play a role in pathogen recognition in worm (Dodd and Drickamer 2001; Mallo 2002) and fly (Choe *et al.* 2002) respectively. Moreover, immunity-related genes (*e. g.*, complement–like, lectin, and TLR) were found in the genome of ascidian *Ciona*

intestinalis and the expected protein domains were found in unique combinations (Dehal *et al.* 2002). Therefore this unusual domain association, dsrm and CASc, represents an optimal response to viral infection which would be deleterious not only for the cell but also for the entire organism. Variability of domain combination may have occurred from gene duplication and shuffling of key motifs under pressure of specific ecological circumstances (reviewed in Loker *et al.* 2004) and could explain why heterogeneity in defense's methods is not in contradiction with the phylogenetic relatedness.

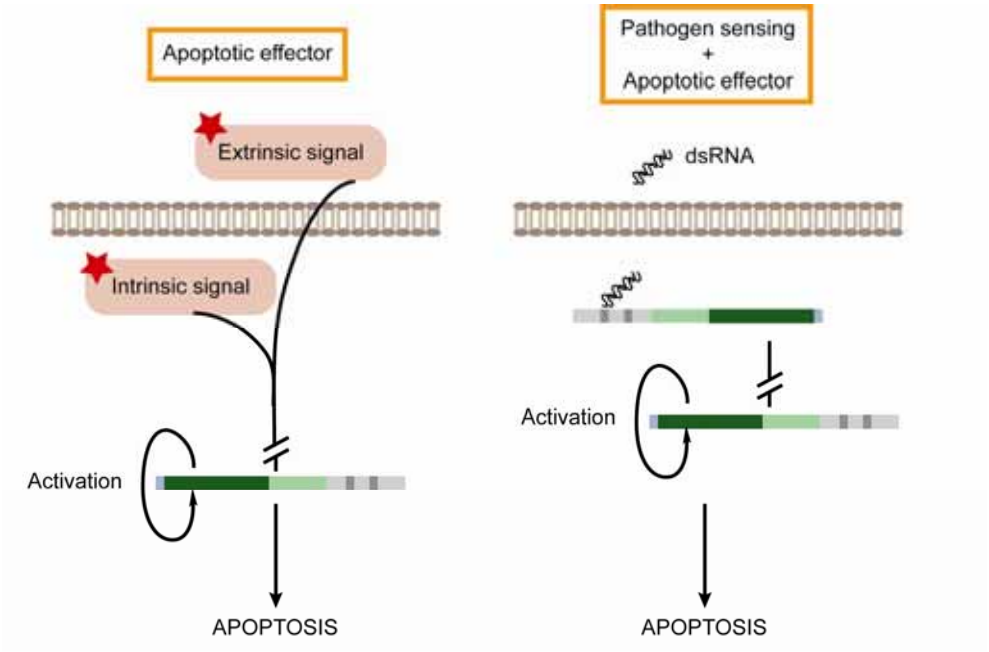


Figure 41: Deduced and proposed function of *S. domuncula* CASL and CASL2. SDCASL and SDCASL2 act as apoptotic effector (close related to caspase-7) but can have also a complementary role in pathogen sensing via their double stranded RNA domain(s). This domain composition may represent the optimal response to sponge viral infection.

Even if the viruses are the most abundant entities in the oceans (94% of the nucleic-acid-containing particles) they represent barely 5% of the biomass while prokaryotes encompass less than 10% of the nucleic-acid-containing particles but more than 90% of the biomass. Thus poriferan should have developed specific mechanisms to counteract bacterial invasion too.

Similar to SDCASL and SDCASL2, TIR-LRR containing protein (SDTILRc) encompasses striking association of domains: several leucine rich repeats (LRR) and a toll-interleukin 1 (TIR)–like domain (table 3) and should represent such a new strategy. Due to these unusual features, only the potential TIR domain of SDTILRc was considered to be of an informative value and was aligned with the corresponding domains of its closest relatives. The resulting tree was rooted with the TIR domain from *E. coli*. A clade formed by the plant sequences was observed at the tree base; however no resolution was obtained. Furthermore, two clusters containing the TIR domain of the MyD88 (myeloid differentiation factor 88) molecules (1) and of the TLR proteins (2) were also spotted. Besides, the position of the TIR domain of *S. domuncula* MyD88 and the TIR domain of SDTILRc at the origin of MyD88 and TLR clusters (respectively) reflects the evolutionary ancient character of these poriferan molecules.

SDTILRc shows significant sequence similarity to NLR family especially to NOD3. NLR (also known as NOD-LRR, NACHT-LRR, and CATERPILLER) is a large family of intracellular proteins with a common protein-domain organization. All members possess variable number of COOH-terminal leucine-rich repeats (LRR) involved in pathogens recognition and a central nucleotide-binding oligomerization domain (NACHT, *i. e.*, NB, NB-ARC, Nod or NBS domain) mediating oligomerization. Finally, the NH_2-termini of NLR are characterized by various effector domains: CARD (caspase recruitment domain), PYD (pyrin) or BIR (baculovirus inhibitor of apoptosis protein repeat) domains (fig. 1). SDTILRc encompasses several COOH-terminal LRR (7) and an NH_2-terminal TIR–like domain, but is missing a NACHT domain.

TIR domain is a key signalling region of all the TLR and IL-1 receptor (IL-1RI), but is also shared by the intracellular signalling adaptors MyD88, MyD88 adaptor–like protein (Mal), TRIF, sterile α-motifs and β-catenin/armadillo repeat motif (SARM), and TRIF-related adaptor molecule (TRAM; reviewed in O'Neill and Bowie 2007), but no TIR domain has been found associated with LRR and NACHT in animals. However, TIR domain is an element of the resistance (R) or NB-LRR plant proteins, molecules related to NLR proteins (reviewed in Ting and Davis 2005). The presence of TIR in many species,

including mammals, plants, and insects, suggests it arose in the common unicellular ancestor of plants and animals.

However the absence of NACHT domain is more intriguing. First of all some domains have been found missing in sponge key molecules, such as the absence of death domain (DD) in *S. domuncula* IRAK-4 (Wiens *et al.* 2007). Secondly, NACHT functions (implication in conformation changes after pathogen sensing and activation) are variable and controversial, thus this domain may be not mandatory. The NACHT domain contains parts of sequence that are conserved in both plant and animal proteins (*e. g.*, Walker's A and B boxes and the kinase 3a motif; van der Biezen and Jones 1998; Takken *et al.* 2006) on the other hand divergent data/function have been reported. Analyses of the NACHT of mammalian apoptotic protease-activating factor 1 (Apaf-1), reveal that adenosine diphosphate (ADP) is preferentially bound in the inactive form and adenosine triphosphate (ATP) seems required for apoptosome formation; however ATP hydrolysis in activation is not clear and interpretation of the ATPase activity variable (Riedl *et al.* 2005; Hu *et al.* 1999; Jiang and Wang 2000). For the NACHT of the worm cell death protein 4 (Ced-4), ATP is preferred in active and inactive conformations, and ATPase activity does not seem to be necessary for its function (Yan *et al.* 2005). Additionally, in plants, data suggest that ATP binding and not hydrolysis is necessary for signalling in plant NB-LRR proteins (Tameling *et al.* 2006). Thirdly, the absence of NACHT domain may be supplemented by a regulatory role of the LRR domains. Positive and negative regulatory roles in signalling have already been highlighted in several studies, for LRR of both plant and animal NB-LRR/NLR. Deletion/truncation of LRR can cause constitutive activation of defense responses (for RPS2, RPS5, and RPP1A; Tao Y *et al.* 2000; Weaver *et al.* 2006), increase in the hypersensitive response (for the potato *R* gene *Rx*, Bendahmane *et al.* 2002), or in the basal activation of transcription factor NF-kB (vertebrate Nod2; Tanabe *et al.* 2004). Paradoxically, opposite effect has also been described: truncation of the LRR domain of the tomato NB-LRR protein I-2 does not seem to result in constitutive activity (Tameling *et al.* 2006) and mutation in the LRR domain of *Arabidopsis* 30S ribosomal protein S5 (RPS5) results in loss of function rather than constitutive activity (Warren *et al.* 1998).

TIR	ARATH TIR	VITVI TIR	SUBDO TILc	SUBDO TLR	SUBDO MyD88	CHICK MyD88	HUMAN TLR5
ARATH TIR		51%	**9%**	21%	15%	15%	15%
VITVI TIR	65%		**10%**	14%	15%	12%	14%
SUBDO TILc	**25%**	**25%**		**10%**	**8%**	**13%**	**12%**
SUBDO TLR	33%	29%	**24%**			15%	16%
SUBDO MyD88	35%	30%	**26%**			25%	13%
CHICK MyD88	34%	29%	**31%**	35%	44%		16%
HUMAN TLR5	35%	34%	**28%**	40%	29%	35%	

Table 3: Comparison of amino-acid TIR sequences from different organisms. Similarity between sequences is quantified via "percent identity" (green) or "percent similarity" (blue).

Expression of SDTILRc was up-regulated after 12 h of challenge with lipopolysaccharide (LPS) from *E. coli* while no effect was observed with LPS of bacteria isolated from *S. domuncula*. LPS is a highly proinflammatory molecule that is a component of the outer envelope of all gram-negative bacteria. In addition to its inflammatory role, LPS induce programmed cell death or apoptosis (Bannerman *et al.* 2001), an event believed to contribute to the pathogenicity and complications of sepsis. LPS is recognized by TLR4, TLR3, TLR7, TLR8, and TLR9 but TLR4 is primarily responsible for inflammatory responses (Fitzgerald *et al.* 2004). Stimulation of TLR4 leads to the activation of NF-kB via TIRAP and MyD88 but additionally activates the transcription factor interferon (IFN) regulatory factor-3 (IRF3), leading to the production of type I IFN (Doyle *et al.* 2002; Kawai *et al.* 2001). Intracellular sensors are also involved in LPS sensing, such as Nalp3 (Kanneganti *et al.* 2006; Sutterwala *et al.* 2006) and NOD2 (Herskovits *et al.* 2007). However, the higher upregulation of NF-kB expression in RAW-blue transfected cells after LPS challenge and also the increase of *SDTILRc* transcripts 12 h after LPS exposure clearly indicate an involvement of SDTILRc in pathogen sensing and signalling pathway mediation.

Thus two possibilities can explain reactivity of SDTILRc to *E. coli* LPS, mainly reposing on SDTILRC/LPS sub-localization (fig. 42). Firstly, alike NOD1/NOD2,

SDTILRc may be found associated with the plasma membrane (Barnich *et al.* 2005; Kufer *et al.* 2008). NOD1 and NOD2 do not possess transmembrane regions which may explain their association to the membrane; however two transmembrane helixes have been predicted in SDTILRc sequence via TUPS (helix 1: $aa_{171-188}$ and helix 2: $aa_{397-413}$). The second possibility is a cytosolic localization, but in this case LPS should be found intracellularly. Intracellular availability of LPS can be explained in various ways, such as, in the simplest case, invasion of the cytoplasm by bacteria and phagocytosis (and subsequent bacterial killing and degradation; Herskovits *et al.* 2007).

Another interesting facet is the non responsiveness of *SDTILRc* to LPS isolated from "sponge-specific bacteria". It has been proposed that cytosolic recognition of microbial factors by NLR proteins appears to be one mechanism whereby the innate immune system is able to discriminate between pathogenic and commensal (members of the host microflora) bacteria (reviewed in Kaparakis *et al.* 2007).

Eventually, an explanation of how SDTILRC can activate immune responses following recognition of pathogen can be attempted. Due to its special features (TIR/LRR), two hypotheses can be raised. Firstly, after pathogen sensing the signalling pathway may be similar to the TLR one (via MyD88, TRIF, TRAM, and TIRAP) thanks to the TIR–like domain, culminating in prosurvival and/or proapoptotic factors expression and/or release. Especially because *S. domuncula* possesses a sponge-specific cell surface known as sponge LPS-interacting protein (SLIP) and a MyD88 related protein involved in LPS mediating downstream signalling pathway, this hypothesis is highly plausible. Thus SDTILRc intracellular sensing pathway may complement defense mechanism against LPS. Lastly, SDTILRc may function in the nucleus. Recent findings suggest that members of the plant TIR-type receptor families (NB-LRR) function in the nucleus via transcriptional reprogramming of host cells for pathogen defense (reviewed in Shen and Schulze-Lefert 2007). *In silico* prediction (Meyers *et al.* 2003) revealed a potential nuclear localization of 71 annotated *Arabidopsis* TIR-containing NB-LRR even if some proteins (20) lack a canonical nuclear localization signal (NLS). SDTILRc does not include a NLS; however 43% of known yeast nuclear proteins enter the nucleus without discernible NLS (reviewed in Lange *et al.* 2007). Furthermore a LIG_NRBOX ($aa_{199-205}$) and a nuclear export signal

(NES; aa_{91-99}; L-x(2,3)-[FILVM]-x(2)-[L]-x-[LIV], exception marked in red, consensus sequence proposed by Heger *et al.* (2001) – NES is generally absent in plant molecules) are predicted in SDTILRc. LIG_NRBOX is a nuclear box motif which confers binding to nuclear receptors. These motifs may assure nuclear transfers/nucleo-cytoplasmic trafficking and a potential regulation. To conclude, SDTILRc may act, like its plants counterparts, as a regulatory signal transduction switch and may be translocated into the nucleus.

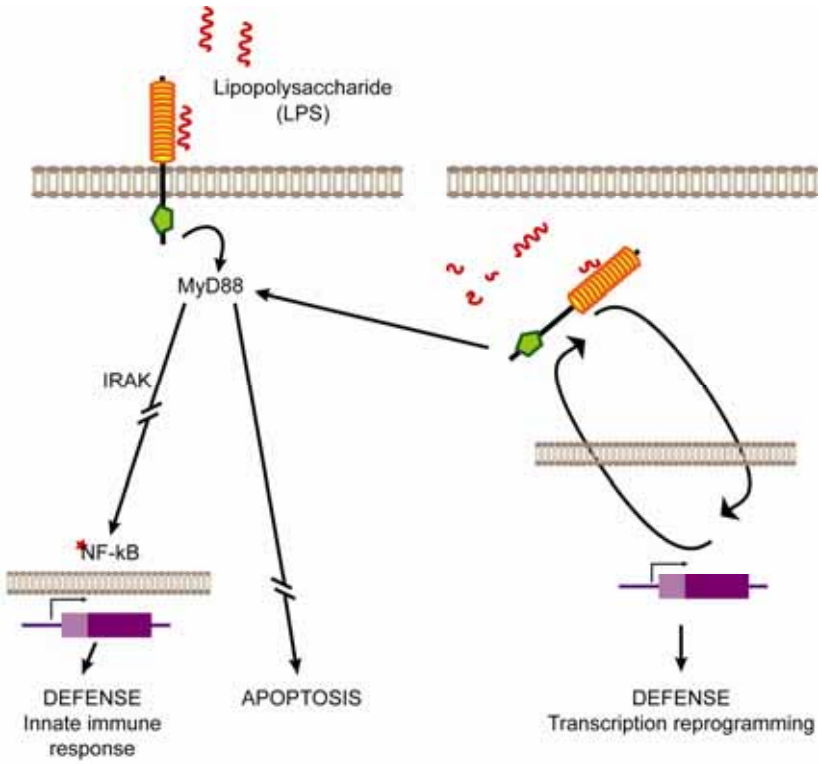


Figure 42: Deduced and proposed function of *S. domuncula* TILRc. SDTILRc (bound to the membrane or free in the cytoplasm) may interact with the canonical NF-kB activating and apoptosis signalling pathways. Moreover SDTILRc may function in the nucleus as its plants counterparts, via transcriptional reprogramming of host cells for pathogen defense.

Several major (TRAF2_1) and one minor (TRAF2_2) TRAF2-binding consensus motifs (TRAF2_1, consensus motif [PSAT]-x-[QE]-E, $aa_{194-197}$, $aa_{490-493}$, and $aa_{561-564}$; TRAF2_2, consensus motif P-x-Q-x-[TSD], $aa_{619-624}$) were predicted via ELM algorithms (consensuses proposed by Ye and Wu in 2000). TRAF2 may regulate SDTILRc effects. TRAF2 is a crucial component of the TNFR superfamily signalling pathway. It initiates important downstream signalling events, such as the activation of the NF-kB family of transcription factors and activation of the mitogen-activated protein kinase (MAPK)

cascade (Song *et al.* 1997). Both TNFRI and TNFRII can signal NF-kB activation via TRAF2, leading to a prosurvival effect (Liu *et al.* 1996; Rothe *et al.* 1994). Besides, SDTILRc carries MAPK docking motifs ($aa_{166-175}$, [KR]-{0,**1**,2}-[KR]-{0,**1**,2}[KR]-{2,4}-[ILVM]-**x**-[ILVF] and $aa_{484-491}$, [KR]-{0,2}-[KR]-{0,2}-[KR]-{2,4}-[ILVM]-**x**-[ILVF] – exceptions shown in bold; consensus sequence proposed by Diella *et al.* in 2008). Thus, SDTILRc may be a MAPK-interacting molecule thus may help to regulate specific interactions in the MAPK cascade or these docking interactions may play an important role in altering protein conformation and allosterically regulating activity (reviewed in Reményi *et al.* 2006)

VII. CONCLUSION

The identification of a bona fide poriferan survivin homologue represents a compelling opportunity to investigate the deep evolutionary original function(s) of survivin while concurrently avoiding limitations of classical invertebrate model organisms. The data presented propose an evolutionary ancient dual regulatory role of survivin, implemented in cell cycle control and apoptosis. Therefore, functional constraints observed in some invertebrates or vertebrate cell lines are probably due to evolutionary divergence from this primordial dual role and are probably the result of adapted genomes and simplified regulatory networks. Consequently, the poriferan survivin provides a unique insight into the genetic complexity of the common metazoan ancestor.

Additionally, the two poriferan caspase-like proteins (SDCASL and SDCASL2), and the TIR-LRR containing protein (SDTILRc), owing to their unusual domain associations, should represent new components of the innate defense sentinel in sponges. These new prototypes should correspond to new families of intracellular sensing, forming a subclass of pattern recognition receptors (PRR) and should represent very efficient defense mechanisms. However, further studies have to be led in order to clarify these defense mechanisms (*e. g.*, activation of SDCASL and SDCASL2 (from a zymogen form to an active form) and potential binding / interacting partners of SDTILRc).

Concurrently, these poriferan molecules offer powerful tools to enhance our comprehension of the interwoven pathways of cell cycle, apoptosis, and innate immunity, thus not only contributing to future anticancer therapeutic strategies but also leading to the discovery of new molecules involved in these major pathways, such as the discoveries of the apoptotic cell death machinery (Hengartner *et al.* 1992; Yuan *et al.* 1993) and the toll molecules (Medzhitov *et al.* 1997) firstly discovered in *C. elegans* and their extrapolation to higher organisms.

VIII. REFERENCES

Aggarwal B. 2003. Signalling pathways of the TNF superfamily: a double-edge sword. Nat Rev Immunol. 3:745–756. *Review*.

Alegado RA, Campbell MC, Chen WC, Slutz SS, Tan MW. 2003. Characterization of mediators of microbial virulence and innate immunity using the *Caenorhabditis elegans* host-pathogen model. Cell Microbiol. 5:435–444.

Aliprantis AO, Yang RB, Weiss DS, Godowski P, Zychlinsky A. 2000. The apoptotic signalling pathway activated by toll-like receptor-2. EMBO J. 19:3325–3336.

Alland C, Moreews F, Boens D, Carpentier M, Chiusa S, Lonquety M, Renault N, Wong Y, Cantalloube H, Chomilier J, Hochez J, Pothier J, Villoutreix BO, Zagury JF, Tufféry P. 2005. RPBS: a web resource for structural bioinformatics. Nucleic Acids Res. 33:44–49.

Altieri DC. 2003. Survivin, versatile modulation of cell division and apoptosis in cancer. Oncogene. 22:8581–8589.

Altieri DC. 2004. Coupling apoptosis resistance to the cellular stress response: the IAP-Hsp90 connection in cancer. Cell Cycle. 3:255–256.

Altieri DC, Marchisio PC, Marchisio C. 1999. Survivin apoptosis: an interloper between cell death and cell proliferation in cancer. Lab Invest. 79:1327–1333.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Alwine JC, Kemp DJ, Stark GR. 1977. Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proc Natl Acad Sci USA. 74:5350–5354.

Ambrosini G, Adida C, Altieri DC. 1997. A novel antiapoptosis gene, BIRC5, expressed in cancer and lymphoma. Nat Med. 3:917–921.

Aoki Y, Feldman GM, Tosato G. 2003. Inhibition of STAT3 signalling induces apoptosis and decreases survivin expression in primary effusion lymphoma. Blood. 101:1535–1542.

Atalay V, Cetin-Atalay R. 2005. Implicit motif distribution based hybrid computational kernel for sequence classification. Bioinformatics. 21:1429–1436.

Ax P. 1996. Multicellular animals: a new approach to the phylogenetic order in nature. Springer-Verlag, Berlin.

Backes C, Kuentzer J, Lenhof HP, Comtesse N, Meese E. 2005. GraBCas: a bioinformatics tool for score-based prediction of Caspase- and Granzyme B-cleavage sites in protein sequences. Nucleic Acids Res. 33:208–213.

Bader GD, Hogue CW. 2000. BIND a data specification for storing and describing biomolecular interactions, molecular complexes and pathways. Bioinformatics. 16:465–477.

Bannerman DD, Tupper JC, Ricketts WA, Bennett CF, Winn RK, Harlan JM. 2001. A constitutive cytoprotective pathway protects endothelial cells from lipopolysaccharide-induced apoptosis. *J Biol Chem*. 276:14924–14932.

Barber GN. 2001. Host defense, viruses and apoptosis. Cell Death Differ. 8:113–126. *Review*.

Barnich N, Aguirre JE, Reinecker HC, Xavier R, Podolsky DK. 2005. Membrane recruitment of NOD2 in intestinal epithelial cells is essential for nuclear factor-{kappa}B activation in muramyl dipeptide recognition. J Cell Biol. 170:21–26.

Beaudoing E, Freier S, Wyatt JR, Claverie JM, Gautheret D. 2000. Patterns of variant polyadenylation signal usage in human genes. Genome Res. 10:1001–1010.

Beltrami E, Plescia J, Wilkinson JC, Duckett CS, Altieri DC. 2004. Acute ablation of survivin uncovers p53-dependent mitotic checkpoint functions and control of mitochondrial apoptosis. J Biol Chem. 279:2077–2084.

Bendahmane A, Farnham G, Moffett P, Baulcombe DC. 2002. Constitutive gain-of function mutants in a nucleotide binding site-leucine rich repeat protein encoded at the Rx locus of potato. Plant J. 32:195–204.

Bergh O, Børsheim KY, Bratbak G, Heldal M. 1989. High abundance of viruses found in aquatic environments. Nature. 340:467–468.

Bergquist PR. 1978. Sponges. Hutchinson, London.

Bettaieb A, Dubrez-Daloz L, Launay S, Plenchette S, Rébé C, Cathelin S, Solary E. 2003. Bcl-2 proteins: targets and tools for chemosensitisation of tumor cells. Curr Med Chem Anticancer Agents. 3:307–318. *Review*.

Birks EJ, Latif N, Enesa K, Folkvang T, Luong le A, Sarathchandra P, Khan M, Ovaa H, Terracciano CM, Barton PJ, Yacoub MH, Evans PC. 2008. Elevated p53 expression is associated with dysregulation of the ubiquitin-proteasome system in dilated cardiomyopathy. Cardiovasc Res. 79:472–480.

Blumbach B, Diehl-Seifert B, Seack J, Steffen R, Müller IM, Müller WEG. 1999. Cloning and expression of new receptors belonging to the immunoglobulin superfamily from the marine sponge *Geodia cydonium*. Immunogenetics. 49:751–763.

Borchiellini C, Manuel M, Alivon E, Boury-Esnault N, Vacelet J, Le Parco Y. 2001. Sponge paraphyly and the origin of Metazoa. J. Evol. Biol. 14:171–179.

Bridge D, Cunningham CW, De Salle R, Buss LW. 1995. Class-level relationships in the phylum Cnidaria: molecular and morphological evidence. Mol. Biol. Evol. 12:679–689.

Burge C, Karlin S. 1997. Prediction of complete gene structures in human genomic DNA. J Mol Biol. 268:78–94.

Bycroft M, Grünert S, Murzin AG, Proctor M, St Johnston D. 1995. NMR solution structure of a dsRNA binding domain from *Drosophila* staufen protein reveals homology to the N-terminal domain of ribosomal protein S5.EMBO J. 14:3563–3571.

Cao XJ, Chen R, Li AP, Zhou JW. 2007. JWA gene is involved in cadmium-induced growth inhibition and apoptosis in HEK-293T cells. J Toxicol Environ Health A. 70:931–937.

Cerretti DP, Kozlosky CJ, Mosley B, *et al.* (12 co-authors). 1992. Molecular cloning of the Interleukin-1 beta converting enzyme. Science. 256:97–100.

Choe KM, Werner T, Stoven S, Hultmark D, Anderson KV. 2002. Requirement for a peptidoglycan recognition protein (PGRP) in relish activation and antibacterial immune responses in *Drosophila*. Science. 296: 359–362.

Chomczynski P, Sacchi N. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem. 162:156–159.

Clemens MJ, Elia A. 1997. The double-stranded RNA-dependent protein kinase PKR: structure and function. J Interferon Cytokine Res. 17:503–524. *Review*.

Coghlan A, Wolfe KH. 2002. Fourfold faster rate of genome rearrangement in nematodes than in *Drosophila*. Genome Res. 12:857–867.

Coligan JE, Dunn BM, Ploegh HL, Speicher DW, Wingfield PT. 2000. Current protocols in protein science. New York: John Wiley & Sons, Inc. p. 2.0.1–2.8.17.

Collins AG. 1998. Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18S rRNA molecular evidence. Proc Natl Acad Sci USA. 95:15458–15463.

Combet C, Blanchet C, Geourjon C, Deléage G. 2000. NPS@: Network Protein Sequence Analysis TIBS. 291:147–150.

Crook NE, Clem RJ, Miller LK. 1993. An apoptosis inhibiting baculovirus gene with a zinc finger like motif. J Virol. 67:2168–2174.

Custodio MR, Prokic I, Steffen R, Koziol C, Borojevic R, Brümmer F, Nickel M, Müller WEG. 1998. Primmorphs generated from dissociated cells of the sponge *Suberites domuncula*: a model system for studies of cell proliferation and cell death. Mech Ageing Dev. 105:45–59.

da Silva Correia J, Miranda Y, Leonard N, Hsu J, Ulevitch RJ. 2007. Regulation of Nod1-mediated signalling pathways. Cell Death Differ. 14:830–839.

Dayhoff MO, Schwartz RM, Orcutt BC. 1978. In: Dayhoff MO, editor. Atlas of protein sequence and structure. Washington (DC): National Biomedical Research Foundation. p.345–352.

Dehal P, Satou Y, Campbell RK, *et al.* (84 co-authors). 2002. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science. 298:2157–2167.

Dellaporta SL, Xu A, Sagasser S, Jakob W, Moreno MA, Buss LW, Schierwater B. 2006. Mitochondrial genome of *Trichoplax adhaerens* supports Placozoa as the basal lower metazoan phylum. Proc Natl Acad Sci USA.103:8751–8756.

Denizot F, Lang R. 1986. Rapid colorimetric assay for cell growth and survival. Modifications to the tetrazolium dye procedure giving improved sensitivity and reliability. J Immunol Methods. 89:271–277.

Deshmukh M, Kuida K, Johnson EMJ. 2000. Caspase inhibition extends the commitment to neuronal death beyond cytochrome C release to the point of mitochondrial depolarization. J. Cell Biol. 150:131–143.

Diella F, Haslam N, Chica C, Budd A, Michael S, Brown NP, Trave G, Gibson TJ. 2008. Understanding eukaryotic linear motifs and their role in cell signaling and regulation. Front Biosci. 13:6580–6603. *Review*.

Dodd RB, Drickamer K. 2001. Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity. Glycobiology. 11:71–79.

Dohi T, Beltrami E,Wall NR, PlesciaJ, Altieri DC. 2004. Mitochondrial survivin inhibits apoptosis and promotes tumorigenesis. J Clin Invest. 114:1117–1127.

Don RH, Cox PT, Wainwright BJ, Baker K, Mattick JS. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. Nucleic Acids Res. 19:4008.

Doyle S, Vaidya S, O'Connell R, Dadgostar H, Dempsey P, Wu T, Rao G, Sun R, Haberland M, Modlin R, Cheng G. 2002. IRF3 mediates a TLR3/TLR4-specific antiviral gene program. Immunity. 17:251–263.

Dujardin F. 1841. Histoire naturelle des Zoophytes De Roret, collection "Nouvelle suites a Buffon, formant, avec les oeuvres de cet auteur, un cours complet d'Histoire naturelle" (Paris).

Emmert DB, Stoehr PJ, Stoesser G, Cameron GN. 1994. The European Bioinformatics Institute (EBI) databases. Nucleic Acids Res. 22:3445–3449.

Felsenstein J. 1993. PHYLIP—phylogeny inference package. Cladistics. 5:164–166.

Finn RD, Mistry J, Schuster-Böckler B, *et al.* (13 co-authors). 2006. Pfam: clans, web tools and services. Nucleic Acids Res. 34:247–251.

Fitzgerald KA, Rowe DC, Golenbock DT. 2004. Endotoxin recognition and signal transduction by the TLR4/MD2-complex. Microbes Infect. 6:1361–1367.

Fraser AG, James C, Evan GI, Hengartner MO. 1999. *Caenorhabditis elegans* inhibitor of apoptosis protein (IAP) homologue BIR-1 plays a conserved role in cytokinesis. Curr Biol. 9:292–301.

Freudenberg MA, Tchaptchet S, Keck S, Fejer G, Huber M, Schütze N, Beutler B, Galanos C. 2008. Lipopolysaccharide sensing an important factor in the innate immune response to Gram-negative bacterial infections: benefits and hazards of LPS hypersensitivity. Immunobiology. 213:193–203. *Review*.

Fuentes-Prior P, Salvesen GS. 2004. The protein structures that shape caspase activity, specificity, activation and inhibition. Biochem J. 384:201–232. *Review*.

Fukuda S, Mantel CR, Pelus LM. 2003. Survivin regulates hematopoietic progenitor cell proliferation through p21WAF1/Cip1-dependent and -independent pathways. Blood. 103:120–127.

Gamulin V, Müller IM, Müller WEG. 2000. Sponge proteins are more similar to those of *Homo sapiens* than to *Caenorhabditis elegans*. Biol. J. Linn. Soc. 71:821–828.

García-Sáez AJ, Coraiola M, Serra MD, Mingarro I, Müller P, Salgado J. 2006. Peptides corresponding to helices 5 and 6 of Bax can independently form large lipid pores. FEBS J. 273:971–981.

Gassmann R, Carvalho A, Henzing AJ, Ruchaud S, Hudson DF, Honda R, Nigg EA, Gerloff DL, Earnshaw WC. 2004. Borealin: a novel chromosomal passenger required for stability of the bipolar mitotic spindle. J Cell Biol. 166:179–191.

Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. 2003. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31:3784–3788.

Gaur U, Aggarwal BB. 2003. Regulation of proliferation, survival and apoptosis by members of the TNF superfamily. Biochem Pharmacol. 66:1403–1408. *Review*.

Giodini A, Kallio M, Wall NR, Gorbsky GJ, Tognin S, Marchisio PC, Symons M, Altieri DC. 2002. Regulation of microtubule stability and mitotic progression by survivin. Cancer Res. 62:2462–2467.

Glücksmann A. 1951. Cell deaths in normal vertebrate ontogeny. Biol Rev. 26:59–86.

Goldstein JC, Waterhouse NJ, Juin P, Evan GI, Green DR. 2000. The coordinate release of cytochrome C is rapid, complete and kinetically invariant. Nat. Cell Biol. 2:156–162.

Graham FL, Smiley J, Russell WC, Nairn R. 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J Gen Virol. 36:59–74.

Green SR, Mathews MB. 1992. Two RNA-binding motifs in the double-stranded RNA-activated protein kinase, DAI. Genes Dev. 6:2478–2490.

Halanych KM. 2004. The new view of animal phylogeny. Annu Rev Ecol Evol Syst. 35:229-256.

Han KJ, Su X, Xu LG, Bin LH, Zhang J, Shu HB. 2004. Mechanisms of the TRIF-induced interferon-stimulated response element and NF-kB activation and apoptosis pathways. J Biol Chem. 279:15652–15661.

Hawkins CJ, Ekert PG, Uren AG, Holmgreen SP, Vaux DL. 1998. Antiapoptotic potential of insect cellular and viral IAPs in mammalian cells. Cell Death Differ. 5:569–576.

Heger P, Lohmaier J, Schneider G, Schweimer K, Stauber RH. 2001. Qualitative highly divergent nuclear export signals can regulate export by the competition for transport cofactors in vivo. Traffic. 2:544–555.

Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, Ignatieva EV, Ananko EA, Podkolodnaya OA, Kolpakov FA, Podkolodny NL, Kolchanov NA. 1998. Databases on transcriptional regulation: TRANSFAC, TRRD, and COMPEL. Nucleic Acids Res. 26:362–367.

Hengartner MO, Horvitz HR. 1994. *C. elegans* cell survival gene Ced-9 encodes a functional homolog of the mammalian proto-oncogene Bcl-2. Cell. 76:665–676.

Hengartner MO, Ellis RE, Horvitz HR. 1992. *Caenorhabditis elegans* gene Ced-9 protects cells from programmed cell death. Nature. 356:494–499.

Herskovits AA, Auerbuch V, Portnoy DA. 2007. Bacterial ligands generated in a phagosome are targets of the cytosolic innate immune system. PLoS Pathog. 3:e51.

Hoffmann R, Valencia A. 2004. A gene network for navigating the literature. Nat Genet. 36:664.

Hoffman WH, Biade S, Zilfou JT, Chen J, Murphy M. 2002. Transcriptional repression of the antiapoptotic survivin gene by wild type p53. J Biol Chem. 277:3247–3257.

Holmes DS, Quigley M. 1981. A rapid boiling method for the preparation of bacterial plasmids. Anal Biochem. 114:193–197.

Honda R, Korner R, Nigg EA. 2003. Exploring the functional interactions between Aurora B, INCENP, and survivin in mitosis. Mol Biol Cell. 14:3325–3341.

Hu Y, Benedict MA, Ding L, Nunez G. 1999. Role of cytochrome C and dATP/ATP hydrolysis in Apaf-1-mediated caspase-9 activation and apoptosis. EMBO J. 18:3586–3595.

Huang HK, Bailis JM, Leverson JD, Gomez EB, Forsburg SL, Hunter T. 2005. Suppressors of Bir1p (survivin) identify roles for the chromosomal passenger protein Pic1p (INCENP) and the replication initiation factor Psf2p in chromosome segregation. Mol Cell Biol. 25:9000–9015.

Hulo N, Bairoch A, Bulliard V, Cerutti L, De Castro E, Langendijk-Genevaux PS, Pagni M, Sigrist CJA. 2006. The PROSITE database. Nucleic Acids Res. 34:227–230.

Inoue H, Nojima H, Okayama H. 1990. High efficiency transformation of *Escherichia coli* with plasmids. Gene. 96:23–28.

Jacob F, Monod J. 1961. Genetic regulatory mechanisms in the synthesis of proteins. J. Mol. Biol. 3:318–328.

Janeway CA. 1989. Approaching the asymptote? Evolution and revolution in immunology.Cold Spring Harb Symp Quant Biol. 1:1–13. *Review*.

Jeong SY, Seol DW. 2008. The role of mitochondria in apoptosis. BMB Rep. 41:11–22. *Review*.

Jiang X, Wang X. 2000. Cytochrome C promotes caspase-9 activation by inducing nucleotide binding to Apaf-1. J. Biol. Chem. 275:31199–31203.

Johnson AL, Langer JS, Bridgham JT. 2002. Survivin as a cell cycle-related and antiapoptotic protein in granulosa cells. Endocrinology. 143:3405–3413.

Jones G, Jones D, Zhou L, Steller H, Chu YX. 2000. Deterin, a new inhibitor of apoptosis from *Drosophila melanogaster*. J Biol Chem. 275:22157–22165.

Kanneganti TD, Ozören N, Body-Malapel M, Amer A, Park JH, Franchi L, Whitfield J, Barchet W, Colonna M, Vandenabeele P, Bertin J, Coyle A, Grant EP, Akira S, Núñez G. 2006. Bacterial RNA and small antiviral compounds activate caspase-1 through cryopyrin/Nalp3. Nature. 440:233–236.

Kaparakis M, Philpott DJ, Ferrero RL. 2007. Mammalian NLR proteins; discriminating foe from friend. Immunol Cell Biol. 85:495–502. *Review*

Karin M, Lin A. 2002. NF-kappaB at the crossroads of life and death. Nat Immunol. 3:221–227. *Review*.

Kawai T, Akira S. 2007. Signalling to NF-kappaB by Toll-like receptors. Trends Mol Med. 13:460–469. *Review*.

Kawai T, Takeuchi O, Fujita T, Inoue J, Mühlradt PF, Sato S, Hoshino K, Akira S. 2001. Lipopolysaccharide stimulates the MyD88-independent pathway and results in activation of IFN-regulatory factor 3 and the expression of a subset of lipopolysaccharide-inducible genes. J Immunol. 167:5887–5894.

Kerr JF. 2002. History of the events leading to the formulation of the apoptosis concept. Toxicology. 182:471–474. *Review*.

Kerr JF, Wyllie AH, Currie AR. 1972. Apoptosis: a basic biological phenomenon with wide-ranging implications in tissue kinetics. Br J Cancer. 26:239–257. *Review*.

Kim J, Kim W, Cunningham CW. 1999. A new perspective on lower metazoan relationships form 18S rDNA sequences. Mol Biol Evol. 16:423–427.

Kim R, Emi M, Tanabe K. 2005. Caspase-dependent and -independent cell death pathways after DNA damage. Oncol Rep. 14:595–599.

Koonin EV, Aravind L. 2002. Origin and evolution of eukaryotic apoptosis: the bacterial connection. Cell Death Differ. 9:394–404.

Kortschak RD, Samuel G, Saint R, Miller DJ. 2003. EST analysis of the cnidarian *Acropora millepora* reveals extensive gene loss and rapid sequence divergence in the model invertebrates. Curr Biol. 13:2190–2195.

Kozak M. 1987. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. 15:8125–8148. *Review*.

Kozak M. 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. Cell Biol. 115:887–903. *Review*.

Krovat BC, Jantsch MF. 1996. Comparative mutational analysis of the double-stranded RNA binding domains of *Xenopus laevis* RNA-binding protein A. J Biol Chem. 271:28112–28119.

Kruse M, Müller IM, Müller WEG. 1997. Early evolution of metazoan serine/threonine- and tyrosine kinases: identification of selected kinases in marine sponges. Mol Biol Evol. 14:1326–1334.

Kufer TA, Fritz JH, Philpott DJ. 2005. NACHT-LRR proteins (NLRs) in bacterial infection and immunity. Trends Microbiol. 13:381–388. *Review*.

Kufer TA, Kremmer E, Adam AC, Philpott DJ, Sansonetti PJ. 2008. The pattern recognition molecule Nod1 is localized at the plasma membrane at sites of bacterial interaction. Cell Microbiol. 10:477–486.

Kurz CL, Ewbank JJ. 2003. *Caenorhabditis elegans*: an emerging genetic model for the study of innate immunity. Nat Rev Genet. 4:380–390.

Kyte J, Doolittle RF. 1982. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 157:105–132.

Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH. 2007. Classical nuclear localization signals: definition, function, and interaction with importin alpha. J Biol Chem. 282:5101–5105. *Review*.

Letunic I, Copley RR, Pils B, Pinkert S, Schultz J, Bork P. 2006. SMART 5: domains in the context of genomes and networks. Nucleic Acids Res. 34:257–260.

Levine B, Sinha S, Kroemer G. 2008. Bcl-2 family members: dual regulators of apoptosis and autophagy. Autophagy. 4:600–606. *Review*.

Li F. 2003. Survivin study: what is the next wave? J Cell Physiol. 197:8–29. *Review*.

Li F, Altieri DC. 1999a. Transcriptional analysis of human survivin gene expression. Biochem J. 2:305–311.

Li F, Altieri DC. 1999b. The cancer anti-apoptosis mouse survivin gene: characterization of locus and transcriptional requirements of basal and cell cycle–dependent expression. Cancer Res. 59:3143–3151.

Li F, Ambrosini G, Chu EY, Plescia J, Tognin S, Marchisio PC, Altieri DC. 1998. Control of apoptosis and mitotic spindle checkpoint by survivin. Nature. 396:580–584.

Li F, Flanary PL, Altieri DC, Dohlman HG. 2000. Cell division regulation by BIR1, a member of the inhibitor of apoptosis family in yeast. J Biol Chem. 275:6707–6711.

Liu ZG, Hsu H, Goeddel DV, Karin M. 1996. Dissection of TNF receptor 1 effector functions: JNK activation is not linked to apoptosis while NF-kB activation prevents cell death. Cell. 87:565–576.

Loker ES, Adema CM, Zhang SM, Kepler TB. 2004. Invertebrate immune systems--not homogeneous, not simple, not well understood. Immunol Rev. 198:10–24. *Review*.

Lupas A, Van Dyke M, Stock J. 1991. Predicting Coiled Coils from protein sequences. Science. 252:1162–1164.

McCormack SJ, Thomis DC, Samuel CE. 1992. Mechanism of interferon action: identification of a RNA binding domain within the N-terminal region of the human RNA-dependent P1/eIF-2 alpha protein kinase. Virology. 188:47–56.

Mallo GV. 2002. Inducible antibacterial defense system in *C. elegans*. Curr Biol. 12:1209–1214.

Mao WP, Ye JL, Guan ZB, Zhao JM, Zhang C, Zhang NN, Jiang P, Tian T. 2007. Cadmium induces apoptosis in human embryonic kidney (HEK) 293 cells by caspase-dependent and -independent pathways acting on mitochondria. Toxicol In Vitro. 21:343–354.

Maruyama K, Sugano S. 1994. Oligo-Capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. Gene. 138:171–174.

Medzhitov R, Preston-Hurlburt P, Janeway CA. 1997. A human homologue of the *Drosophila* Toll protein signals activation of adaptive immunity. Nature. 6640:394–397.

Mehl D, Müller I, Müller WEG. 1997. Molecular biological and paleontological evidence that Eumetazoa, including Porifera (sponges), are of monophyletic origin. Sponge sciences: multidisciplinary perspectives. Eds. Watanabe Y, Fusetani N, Springer-Verlag. Tokyo.

Meyers BC, Kozik A, Griego A, Kuang HH, Michelmore RW. 2003. Genome-wide analysis of NBS-LRR-encoding genes in *Arabidopsis*. Plant Cell. 15: 809–834.

Meylan E, Tschopp J, Karin M. 2006. Intracellular pattern recognition receptors in the host response. Nature. 442:39–44. *Review*.

Milanesi L, D'Angelo D, Rogozin IB. 1999. GeneBuilder: interactive *in silico* prediction of gene structure. Bioinformatics. 15:612–621.

Mosmann T. 1983. Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. J Immunol Methods. 65:55–63.

Müller WEG. 1998. Origin of Metazoa: sponges as living fossils. Naturwissenschaften. 85:11–25.

Müller WEG. 2001. How was metazoan threshold crossed? The hypothetical Urmetazoa. Comp Biochem Physiol A Mol Integr Physiol. 129:433–460. *Review*.

Müller WEG, Müller I. 2003. Analysis of the sponge (Porifera) gene repertoire: implication for the evolution of the Metazoan body plan. Marine Molecular Biotechnology. Ed. Müller WEG, Springer-Verlag. Berlin.

Müller WEG, Blumbach B, Müller IM. 1999. Evolution of the innate and adaptive immune systems: relationships between potential immune molecules in the lowest metazoan phylum (Porifera) and those in vertebrates. Transplantation. 68:1215–1227. *Review*.

Müller WEG, Perovic S, Wilkesman J, Kruse M, Müller IM, Batel R. 1999. Increased gene expression of a cytokine-related molecule and profilin after activation of *Suberites domuncula* cells with xenogeneic sponge molecule(s). DNA Cell Biol. 18:885–893.

Müller WEG, Schröder HC, Skorokhod A, Bünz C, Müller IM, Grebenjuk VA. 2001. Contribution of sponge genes to unravel the genome of the hypothetical ancestor of Metazoa (Urmetazoa). Gene. 276:161–173. *Review*.

Müller-Hill B, Crapo L, Gilbert W. 1968. Mutants that make more *lac* repressor. Proc. Natl. Acad. Sci. USA. 59:1259–1262.

Nakai K, Horton P. 1999. PSORT: a program for detecting sorting signals in proteins and predicting their sub-cellular localization. Trends Biochem Sci. 24:34–36.

Naugler WE, Karin M. 2008. NF-kappaB and cancer-identifying targets and mechanisms. Curr Opin Genet Dev. 18:19–26. *Review*.

Nicholas KB, Nicholas HB Jr. 1997. GeneDoc: a tool for editing and annotating multiple sequence alignments. Version 2.6.002 [Internet]. Distributed by the authors [cited 2007 Jan 22]. Available from: http://www.psc.edu/biomed/genedoc/.

Noton EA, Colnaghi R, Tate S, Starck C, Carvalho A, Ko Ferrigno P, Wheatley SP. 2006. Molecular analysis of survivin isoforms: evidence that alternatively spliced variants do not play a role in mitosis. J Biol Chem. 281:1286–1295.

O'Neill LA, Bowie AG. 2007. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. Nat Rev Immunol. 7:353–364. *Review*.

Oppenheim JJ, Tewary P, de la Rosa G, Yang D. 2007. Alarmins initiate host defense. Adv Exp Med Biol. 601:185–194. *Review*.

Ozben T. 2007. Oxidative stress and apoptosis: impact on cancer therapy. J Pharm Sci. 96:2181–2196. *Review*.

Pan Q, Wendel J, Fluhr R. 2000. Divergent evolution of plant NBS-LRR resistance gene homologues in dicot and cereal genomes. J Mol Evol. 50:203–213.

Pancer Z, Skorokhod A, Blumbach B, Müller WEG. 1998. Multiple Ig-like featuring genes divergent within and among individuals of the marine sponge *Geodia cydonium*. Gene. 207:227–233.

Patel RC, Sen GC. 1998. PACT, a protein activator of the interferon-induced protein kinase, PKR. EMBO J. 17:4379–4390.

Pearson WR, Lipman DJ. 1988. Improved tools for biological sequence comparison. Proc. Natl. Acad. Sci. USA. 85:2444–2448.

Perovic' S, Schröder HC, Sudek S, Grebenjuk VA, Batel R, Stifanić M, Müller IM, Müller WEG. 2003. Expression of one sponge Iroquois homeobox gene in primmorphs from *Suberites domuncula* during canal formation. Evol Dev. 5:240–250.

Perović-Ottstadt S, Adell T, Proksch P, Wiens M, Korzhev M, Gamulin V, Müller IM, Müller WEG. 2004. A (1-->3)-beta-D-glucan recognition protein from the sponge *Suberites domuncula*. Mediated activation of fibrinogen-like protein and epidermal growth factor gene expression. Eur J Biochem. 271:1924–1937.

Polak JM, McGee JD. 1998. *In situ* hybridization. Oxford: Oxford University Press.

Ponger L, Mouchiroud D. 2001. CpGProD: identifying CpG islands associated with transcription start sites in large genomic mammalian sequences. Bioinformatics. 18:631–633.

Raj D, Liu T, Samadashwily G, Li F, Grossman D. 2008. Survivin repression by p53, Rb and E2F2 in normal human melanocytes. Carcinogenesis. 29:194–201.

Rast JP, Smith LC, Loza-Coll M, Hibino T, Litman GW. 2006. Genomic insights into the immune system of the sea urchin. Science. 314:952–956. *Review*.

Reed JC, Bischoff JR. 2000. BIRinging chromosomes through cell division--and survivin' the experience. Cell. 102:545–548. *Review*.

Reményi A, Good MC, Lim WA. 2006. Docking interactions in protein kinase and phosphatase networks. Curr Opin Struct Biol. 16:676–685. *Review*.

Riedl SJ, Li W, Chao Y, Schwarzenbacher R, Shi Y. 2005. Structure of the apoptotic protease-activating factor 1 bound to ADP. Nature. 434:926–933.

Rietdijk ST, Burwell T, Bertin J, Coyle AJ. 2008. Sensing intracellular pathogens-NOD-like receptors. Curr Opin Pharmacol. 8:261–266. *Review*.

Rigby JK. 1987. Phylum Porifera. In Boardman RS, Cheetham AH, Rowell AJ (eds). Fossil Invertebrates. Berlin. Blackwell Science.

Rothe M, Wong SC, Henzel WJ, Goeddel DV. 1994. A novel family of putative signal transducers associated with the cytoplasmic domain of the 75 kDa tumor necrosis factor receptor. Cell. 78:681–692.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 132:365–386.

Ryan PD, Harper DAT, Whalley JS. 1995. PALSTAT, Statistics for palaeontologists. Chapman & Hall (now Kluwer Academic Publishers).

Rychlik W, Rhoads RE. 1989. A computer program for choosing optimal oligonucleotides forfilter hybridization, sequencing and in vitro amplification of DNA. Nucleic Acids Research. 17:8543–8551.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.

Samal B, Sun Y, Stearns G, Xie C, Suggs S, McNiece I. 1994. Cloning and characterization of the cDNA encoding a novel human pre-B-cell colony-enhancing factor. Mol Cell Biol. 14:1431–1437.

Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular Cloning I-III. Cold Spring Harbor Laboratory Press, New York, USA.

Sanger F, Nicklen S, Coulson AR. 1977. DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA. 74:5463–5467.

Saunders LR, Barber GN. 2003. The dsRNA binding protein family: critical roles, diverse cellular functions. FASEB J. 17:961–983. *Review*.

Schaefer BC. 1995. Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. Anal. Biochem. 227:255–273.

Schatz DG, Oettinger MA, Schlissel MS. 1992. V(D)J recombination: molecular biology and regulation. Annu Rev Immunol. 10:359–383. *Review*.

Schug J. 2008. Using TESS to predict transcription factor binding sites in DNA sequence. Curr Protoc Bioinformatics. Mar; Chapter 2:Unit 2.6.

Schütze J, Krasko A, Custodio MR, Efremova SM, Müller IM, Müller WEG. 1999. Evolutionary relationships of Metazoa within the eukaryotes based on molecular data from Porifera. Proc Biol Sci. 266:63–73.

Seack J, Perović S, Gamulin V, Schröder HC, Beutelmann P, Müller IM, Müller WEG. 2001. Identification of highly conserved genes: SNZ and SNO in the marine sponge *Suberites domuncula*: their gene structure and promoter activity in mammalian cells. Biochim Biophys Acta. 1520:21–34.

Shen QH, Schulze-Lefert P. 2007. Rumble in the nuclear jungle: compartmentalization, trafficking, and nuclear action of plant immune receptors. EMBO. 26: 4293–4301.

Shi Q, Jackowski G. 1998. One-dimensional polyacrylamide gel electrophoresis, pp 1–52 in Hames BD (ed) Gel Electrophoresis of Proteins: A Practical Approach, 3rd edn, Oxford University Press, Oxford.

Schultz U, Kaspers B, Staeheli P. 2004. The interferon system of non-mammalian vertebrates. Dev. Comp. Immunol. 28:499–508.

Siddall ME, Martin DS, Bridge D, Desser SS, Cone DK. 1995. The demise of a phylum of protists: phylogeny of Myxozoa and other parasitic Cnidaria. J Parasitol. 81:961–967.

Slack JL, Schooley K, Bonnert TP, Mitcham JL, Qwarnstrom EE, Sims JE, Dower SK. 2000. Identification of two major sites in the type I interleukin-1 receptor cytoplasmic region responsible for coupling to proinflammatory signalling pathways. J Biol Chem. 275:4670–4678.

Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC. 1987. Measurement of protein using bicinchoninic acid. Anal Biochem. 150:76–85.

Sommer KW, Schamberger CJ, Schmidt GE, Sasgary S, Cerni C. 2003. Inhibitor of apoptosis protein (IAP) survivin is upregulated by oncogenic c-H-Ras. Oncogene. 22:4266–4280.

Song HY, Régnier CH, Kirschning CJ, Goeddel DV, Rothe M. 1997. Tumor necrosis factor (TNF)-mediated kinase cascades: bifurcation of nuclear factor-kappaB and c-jun N-terminal kinase (JNK/SAPK) pathways at TNF receptor-associated factor 2. Proc Natl Acad Sci USA. 94:9792–9796.

Speliotes EK, Uren A, Vaux D, Horvitz HR. 2000. The survivin-like *C. elegans* BIR-1 protein acts with the aurora-like kinase AIR-2 to affect chromosomes and the spindle midzone. Mol Cell. 6:211–223.

Sutterwala FS, Ogura Y, Szczepanik M, Lara-Tejero M, Lichtenberger GS, Grant EP, Bertin J, Coyle AJ, Galán JE, Askenase PW, Flavell RA. 2006. Critical role for NALP3/CIAS1/Cryopyrin in innate and adaptive immunity through its regulation of caspase-1. Immunity. 24:317–327.

Suttle CA. 2007. Marine viruses--major players in the global ecosystem. Nat Rev Microbiol. 5:801–812. *Review*.

Takken FL, Albrecht M, Tameling WI. 2006. Resistance proteins: molecular switches of plant defense. Curr. Opin. Plant Biol. 9:383–390.

Tameling WI, Vossen JH, Albrecht M, Lengauer T, Berden JA, Haring MA, Cornelissen BJ, Takken FL. 2006. Mutations in the NB-ARC domain of I-2 that impair ATP hydrolysis cause autoactivation. Plant Physiol. 140:1233–1245.

Tamm I, Wang Y, Sausville E, Scudiero DA, Vigna N, Oltersdorf T, Reed JC. 1998. IAP-family protein survivin inhibits caspase activity and apoptosis induced by Fas (CD95), Bax, caspases, and anticancer drugs. Cancer Res. 58:5315–5320.

Tanabe T, Chamaillard M, Ogura Y, Zhu L, Qiu S, Masumoto J, Ghosh P, Moran A, Predergast MM, Tromp G, Williams CJ, Inohara N, Núñez G. 2004. Regulatory regions and critical residues of NOD2 involved in muramyl dipeptide recognition. EMBO J. 23:1587–1597.

Tao Y, Yuan F, Leister RT, Ausubel FM, Katagiri F. 2000. Mutational analysis of the *Arabidopsis* nucleotide binding site-leucine-rich repeat resistance gene RPS2. Plant Cell. 12:2541–2554.

Tenev T, Zachariou A, Wilson R, Ditzel M, Meier P. 2005. IAPs are functionally non-equivalent and regulate effector caspases through distinct mechanisms. Nat Cell Biol. 7:70–77.

Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.

Thornberry NA, Lazebnik Y. 1998. Caspases: enemies within. Science. 281:1312–1316. *Review*.

Ting JP, Davis BK. 2005. CATERPILLER: a novel gene family important in immunity, cell death, and diseases. Annu Rev Immunol. 23:387–414. *Review*.

Torres VA, Tapia JC, Rodríguez DA, Párraga M, Lisboa P, Montoya M, Leyton L, Quest AF. 2006. Caveolin-1 controls cell proliferation and cell death by suppressing expression of the inhibitor of apoptosis protein survivin. J Cell Sci. 119:1812–1823.

Tracey L, Pérez-Rosado A, Artiga MJ, Camacho FI, Rodríguez A, Martínez N, Ruiz-Ballesteros E, Mollejo M, Martinez B, Cuadros M, Garcia JF, Lawler M, Piris MA. 2005. Expression of the NF-kappaB targets Bcl-2 and BIRC5/Survivin characterizes small B-cell and aggressive B-cell lymphomas, respectively. J Pathol. 206:123–134.

Uren AG, Beilharz T, O'Connell MJ, Bugg SJ, van Driel R, Vaux DL, Lithgow T. 1999. Role for yeast inhibitor of apoptosis (IAP)-like proteins in cell division. Proc.Natl. Acad. Sci. 96:10170–10175.

Uwe S. 2008. Anti-inflammatory interventions of NF-kappaB signalling: potential applications and risks. Biochem Pharmacol. 75:1567–1579. *Review*.

Vagnarelli P, Earnshaw WC. 2004. Chromosomal passengers: the four-dimensional regulation of mitotic events. Chromosoma. 113:211–222.

Vaira V, Lee CW, Goel HL, Bosari S, Languino LR, Altieri DC. 2007. Regulation of survivin expression by IGF-1/mTOR signalling. Oncogene. 26:2678–2684.

Van Antwerp DJ, Martin SJ, Verma IM, Green DR. 1998. Inhibition of TNF-induced apoptosis by NF-kappa B. Trends Cell Biol. 8:107–111. *Review*.

van der Biezen EA, Jones JDG. 1998. The NB-ARC domain: a novel signalling motif shared by plant resistance gene products and regulators of cell death in animals. Curr. Biol. 8:226–227.

Venables JP. 2004. Aberrant and alternative splicing in cancer. Cancer Res. 64:7647–7654. *Review*.

Verhagen AM, Vaux DL. 2002. Cell death regulation by the mammalian IAP antagonist Diablo/Smac. Apoptosis. 7:163–166. *Review*.

Verhagen AM, Coulson EJ, Vaux DL. 2001. Inhibitor of apoptosis proteins and their relatives: IAPs and other BIRPs. Genome Biol. 2-REVIEWS 3009.

Volloch V, Schweitzer B, Rits S. 1994. Ligation-mediated amplification of RNA from murine erythroid cells reveals a novel class of beta-globin mRNA with an extended 5'- untranslated Region. Nucleic Acids Res. 22:2507–2511.

Wall NR, O'Connor DS, Plescia J, Pommier Y, Altieri DC. 2003. Suppression of survivin phosphorylation on Thr34 by flavopiridol enhances tumor cell apoptosis. Cancer Res. 63:230–235.

Warren RF, Henk A, Mowery P, Holub E, Innes RW. 1998. A mutation within the leucine-rich repeat domain of the *Arabidopsis* disease resistance gene *RPS5* partially suppresses multiple bacterial and downy mildew resistance genes. Plant Cell. 10:1439–1452.

Watanabe M, Henmi K, Ogawa K, Suzuki T. 2003. Cadmium-dependent generation of reactive oxygen species and mitochondrial DNA breaks in photosynthetic and non-photosynthetic strains of Euglena gracilis. Comp Biochem Physiol C Toxicol Pharmacol. 134:227–234.

Weaver ML, Swiderski MR, Li Y, Jones JDG. 2006. The *Arabidopsis thaliana* TIRNB-LRR R-protein, RPP1A; protein localization and constitutive activation of defense by truncated alleles in tobacco and *Arabidopsis.* Plant J. 47:829–840.

Weber K, Osborn M. 1969. The reliability of molecular weight determinations by dodecyl sulfatepolyacrylamide gel electrophoresis. J Biol Chem. 244:4406–4412.

Wei Y, Fan T, Yu M. 2008. Inhibitor of apoptosis proteins and apoptosis. Acta Biochim Biophys Sin (Shanghai). 40:278–288. *Review*.

Wenzel M, Mahotka C, Krieg A, Bachmann A, Schmitt M, Gabbert HE, Gerharz CD. 2000. Novel survivin-related members of the inhibitor of apoptosis (IAP) family. Cell Death Differ. 7:682–683.

Wiens M, Kuusksalu A, Kelve M, Müller WEG. 1999. Origin of the interferon-inducible (2', 5') oligoadenylate synthetases: cloning of the (2', 5') oligoadenylate synthetase from the marine sponge *Geodia cydonium*. FEBS Lett. 462:12–18.

Wiens M, Krasko A, Müller CI, Müller WEG. 2000. Molecular evolution of apoptotic pathways: cloning of key domains from sponges (Bcl-2 homology domains and death domains) and their phylogenetic relationships. Mol Evol. 50:520–531.

Wiens M, Diehl-Seifert B, Müller WEG. 2001. Sponge Bcl-2 homologous protein (BHP2-GC) confers distinct stress resistance to human HEK-293 cells. Cell Death Differ. 8:887–898.

Wiens M, Luckas B, Brümmer F, Ammar MSA, Steffen R, Batel R, Diehl-Seifert B, Schröder HC, Müller WEG. 2002. Okadaic acid: a potential defense molecule for the sponge *Suberites domuncula.* Mar Biol. 142:213–223.

Wiens M, Krasko A, Perovic S, Müller WEG. 2003. Caspase mediated apoptosis in sponges: cloning and function of the phylogenetic oldest apoptotic proteases from Metazoa. Biochim Biophys Acta. 1593:179–189.

Wiens M, Korzhev M, Krasko A, Thakur NL, Perović-Ottstadt S, Breter HJ, Ushijima H, Diehl-Seifert B, Müller IM, Müller WEG. 2005. Innate immune defense of the sponge *Suberites domuncula* against bacteria involves a MyD88-dependent signalling pathway. Induction of a perforin-like molecule. J Biol Chem. 280:27949–27959.

Wiens M, Korzhev M, Perović-Ottstadt S, Luthringer B, Brandt D, Klein S, Müller WEG. 2007. Toll-like receptors are part of the innate immune defense system of sponges (Demospongiae: Porifera). Mol Biol Evol. 24:792–804.

Wu LP, Anderson KV. 1997. Related signalling networks in *Drosophila* that control dorsoventral patterning in the embryo and the immune response. Cold Spring Harb Symp Quant Biol. 62:97–103.

Yan N, Chai J, Lee ES, Gu L, Liu Q, He J, Wu JW, Kokel D, Li H, Hao Q, Xue D, Shi Y. 2005. Structure of the Ced-4-Ced-9 complex provides insights into programmed cell death in *Caenorhabditis elegans*. Nature. 437:831–837.

Yang Y, Fang S, Jensen JP, Weissman AM, Ashwell JD. 2000. Ubiquitin protein ligase activity of IAPs and their degradation in proteasomes in response to apoptotic stimuli. Science. 288:874–877.

Ye H, Wu H. 2000. Thermodynamic characterization of the interaction between TRAF2 and tumor necrosis factor receptor peptides by isothermal titration calorimetry. Proc Natl Acad Sci USA. 97:8961–8966.

Yuan J, Horvitz HR. 2004. A first insight into the molecular mechanisms of apoptosis. Cell. 116:53–59.

Yuan J, Shaham S, Ledoux S, Ellis HM, Horvitz HR. 1993. The *C. elegans* cell death gene Ced-3 encodes a protein similar to mammalian interleukin-1 beta-converting enzyme. Cell. 75:641–652.

Zangemeister-Wittke U, Simon HU. 2004. An IAP in action: the multiple roles of survivin in differentiation, immunity and malignancy. Cell Cycle. 9:1121–1123. *Review*.

Zeytun A, van Velkinburgh JC, Pardington PE, Cary RR, Gupta G. 2007. Pathogen-specific innate immune response. Adv Exp Med Biol. 598:342–357. *Review*.

Zhao J, Tenev T, Martins LM, Downward J, Lemoine NR. 2000. The ubiquitin-proteasome pathway regulates survivin degradation in a cell cycle-dependent manner. Cell Sci. 23:4363–4371.

Zhou H, Zhou Y. 2003. Predicting the topology of transmembrane helical proteins using mean burial propensity and a hidden-Markov-model-based method. Protein Sci. 12:1547–1555.

Zhou M, Gu L, Li F, Zhu Y, Woods WG, Findley HW. 2002. DNA damage induces a novel p53-survivin signalling pathway regulating cell cycle and apoptosis in acute lymphoblastic leukemia cells. J Pharmacol Exp Ther. 303:124–131.

Zipfel C, Felix G. 2005. Plants and animals: a different taste for microbes? Curr Opin Plant Biol. 8:353–360. *Review*.

IX. ABBREVIATIONS

Δψm	mitochondrial membrane potential
°C	degree Celsius
µL	microliter
µm	micrometre
µM	micromolar

A

ADAR	adenosine deaminase acting on RNA
ADP	adenosine diphosphate
AIF	apoptosis-inducing factor
AK	adenosine kinase
AMEPV	*Amsacta moorei entomopoxvirus*
ANOGA	*A. gambiae*, *Anopheles gambiae*
AP	alkaline phosphatase
Apaf1	apoptotic peptidase activating factor 1
APS	ammonium persulphate
ARATH	*A. thaliana*, *Arabidopsis thaliana*
ATP	adenosine triphosphate
AurB	Aurora-B

B

Bad	Bcl-2 associated agonist of cell death
Bak	Bcl-2-antagonist/killer
Bax	Bcl-2-associated X protein
BCA	bicinchoninic acid
BCL2	B-cell CLL/lymphoma 2
BED	zinc finger BED-type
BH (1 to4)	Bcl-2 homology
Bid	BH3 interacting domain death agonist
Bim	Bcl-2-interacting mediator of cell death
BIR	baculovirus IAP repeat
BIRC	BIR domain-containing proteins
BOR	Borealin
BOVIN	*B. taurus*, *Bos taurus*
Bp	base pair(s)
BSA	bovine serum albumin

C

CAEEL	*C. elegans*, *Caenorhabditis elegans*
CAM	cell adhesion molecules
CANFA	*C. lupus*, *Canis lupus*
CARD	caspase recruitment domain
CASc	caspase, interleukin-1 beta converting enzyme (ICE) homologues
CATERPILLER	CARD, transcription enhancer, R(purine)-binding, pyrin, lots of leucine repeats
CC	coiled-coil
CD	cluster of differentiation
CDE	cycle dependent element
cDNA	complementary DNA
CDS	coding sequence
Ced-3	cell death protein 3
Ced-4	cell death protein or cell death abnormality 4
Ced-9	cell death protein or cell death abnormality 9
CHICK	*G. gallus*, *Gallus gallus*
CHR	cell cycle homology regions
CIOIN	*C. intestinalis*, *Ciona intestinalis*
CLR	C-type lectin receptors
cm	centimetre
cm^2	square centimetre
CMFSW	calcium-and magnesium-free synthetic sea-water
COOH-terminus	*i. e.*, carboxyl-terminus, carboxy-terminus, and C-terminal end
CPC	chromosomal passenger complex
CRD	cysteine-rich domain
CRD	cysteine-rich domains
CTL	C-type lectin domain
CULPI	*C. pipiens*, *Culex pipiens*

D

DANRE	*D. Rerio*, *Danio rerio*
dATP	deoxyadenine triphosphate
dCTP	deoxycytidine triphosphate
DD	death domain
ddH2O	double distilled water
ddNTP	dideoxyribonucleotide triphosphate
DED	death effector domain
DEPC	diethyl pyrocarbonate

dGTP	deoxyguanosine triphosphate
DIG	digoxigenin
DISC	death-inducing signalling complex
DMEM	Dulbecco's modified Eagle's medium
DMF	dimethyl formamide
DMPC	dimethyl pyrocarbonate
DMSO	dimethyl sulphoxide
DNA	desoxyribonucleic acid
DNase	desoxyribonuclease
dNTP	deoxyribonucleotide triphosphate
DR	death receptor
DROME	*D. Melanogaster*, *Drosophila melanogaster*
dsrm	double stranded RNA binding domains
dsRNA binding proteins	DRBP
dsRNA	double stranded RNA
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
dUTP	deoxyuridine triphosphate

E

ECM	extra cellular matrix
ECO	*E. Coli*, *Escherichia coli*
EDTA	ethylenediaminetetraacetic acid
EGFP	enhance green fluorescence protein
ELISA	enzyme linked immunosorbent assay
EQUPR	*E. caballus*, *Equus caballus*
et al.	et alii
EtBr	ethidium bromide

F

Fab	fragment antigen binding
FADD	FAS-associated death domain protein
Fas	(=APO-1) TNF receptor superfamily, member 6
FCS	foetal calf serum
FELCA	*F. catus*, *Felix catus*
FLICE	FADD-homologous ICE/Ced-3–like protease = caspase 8
FLIP	FLICE inhibitory protein

G

g	g-force (9.81 m/s^2)
g	gramme
GEOCY	*G. cydonium*, *Geodia cydonium*
GFP	green fluorescent protein
GSP	gene-specific primer
GVCP	*Cydia pomonella granulosis virus*
GVCP	*Cydia pomonella granulosis virus*

H

h	hour
H2O	water
HEK-293	human embryonic kidney cells
His	histidine
HRP	horseradish peroxidase
HSP	heat shock protein
Htra2	high temperature requirement protein A2

I

IAP	inhibitor of apoptosis protein
ICE	interleukin-1β converting enzyme
IFN	interferon
IgG	immunoglobulin G
IkB	Inhibitor of kB
IKK	IkB kinase
IL	interleukin
IL-1RI	IL-1 receptor
INCENP	inner centromere protein
IPTG	isopropyl-β-D-thiogalactoside
IRD	infrared dye
IRF3	interferon regulatory factor-3

J

J	joule
JNK	*i. e.*, c-Jun N-terminal kinase 1, and mitogen activated protein kinase 8

K

kb	kilobase(s)
kDa	kiloDalton(s)

L

L	litre
LB	Luria Bretani (broth)
LETJA	*L. japonicum*, *Lethenteron japonicum*
log	logarithm
LPS	lipopolysaccharides
LRR	leucine rich repeat
LSU	large ribosomal subunit

M

m	metre
M	molar (mol/L)
Mal	MyD88 adaptor–like protein
MAMP	microbes-associated molecular patterns
MATra	magnet assisted transfection
MCS	multiple cloning site
min	minute
mL	millilitre
mm	millimetre
mM	millimolar
MOUSE	*M. musculus*, *Mus musculus*
MY	million years
MyD88	myeloid differentiation primary response gene 88

N

NB-LRR	nucleotide binding domain and leucine-rich repeat protein
NACHT	*i. e.*, NB, NB-ARC, Nod, NBS, and nucleotide-binding oligomerization domain
NEMVE	*N. Vectensis*, *Nematostella vectensis*
NES	nuclear export signal
NFAR	nuclear factors associated with dsRNA
NF-kB	nuclear factor-kappa B
ng	nanogramme

NH2-terminus	*i. e.*, amino-terminus, amine-terminus, and N-terminal end
NITROB	*Nitrobacter sp.*
NLR	*i. e.*, NOD-LRR, NACHT-LRR, CATERPILLER, and Nod–like receptors
NLS	nuclear localization signal
nm	nanometre
Nod1	nucleotide-binding oligomerization domain containing 1
NPVBS	*Buzura suppressaria nuclear polyhedrosis virus*
NPVCF	*Choristoneura fumiferana nuclear polyhedrosis virus*
NPVEP	*Epiphyas postvittana nucleopolyhedrovirus*
NPVEP	*Epiphyas postvittana nucleopolyhedrovirus*
NPVHA	*Helicoverpa armigera nuclear polyhedrosis virus*
NRBOX	nuclear box motif

O

OD	optic density
ON	over-night

P

p38MAPK	p38 mitogen-activated protein kinase
PACT	protein kinase, interferon inducible double stranded RNA dependent activator
PAGE	polyacrylamide gel electrophoresis
PAM	pathogens associated patterns
PANTR	*P. troglodytes*, *Pan troglodytes*
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pH	potentia hydrogenii
PIG	*S. scrofa*, *Sus scrofa*
PKR	double stranded RNA-dependent protein kinase
PPP	pyrophosphatase
PRR	pattern recognition receptors
PTP	permeability transition pore
PVDF	polyvinylidene fluoride
PYD	pyrin domain

R

RACE	rapid amplification of cDNA ends
RBS	ribosome binding site

RIP2	*i. e.*, receptor-interacting serine-threonine kinase 2, and CARD-containing interleukin-1 beta-converting enzyme-associated kinase
RLH	RIG-I–like helicase
RLM-RACE	RNA ligase-mediated and oligo-capping rapid amplification of cDNA ends
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
ROS	reactive oxygen species
rpm	revolutions per minute
RPMI	Roswell Park Memorial Institute medium
RPS5	30S ribosomal protein S5
RT	reverse transcription
RT	room temperature
RTK	receptor tyrosine kinase

S

s	second
SALMO	*S. salar*, *Salmo salar*
SARM	sterile α-motifs and β-catenin/armadillo repeat motif
SCHJA	*S. japonicum*, *Schistosoma japonicum*
SCHPO	*S. Pombe*, *Schizosaccharomyces pombe*
SDCASL	*Suberites domuncula* caspase–like protease
SDCASL2	*Suberites domuncula* caspase–like protease 2
SDS	sodium dodecyl sulphate
SDSURVL	*Suberites domuncula* survivin–like
SDTILRc	*Suberites domuncula* TIR-LRR containing protein
SEAP	secreted embryonic alkaline phosphatase
SIMP	soluble inter membrane mitochondria protein
SLIP	sponge LPS-interacting protein
Smac/DIABLO	second mitochondria-derived activator of caspases/direct IAP-binding protein with low pi
SSC	saline-sodium citrate
SSU	small ribosomal subunit
STRPU	*S. purpuratus*, *Strongylocentrotus purpuratus*
SUBDO	*S. domuncula*, *Suberites domuncula*
SUR	survivin

T

TAE	Tris-Acetate-EDTA

TAK1	TGF-beta activated kinase 1
TBE	Tris-Borate-EDTA
TBST	Tris-buffered saline
TBST	Tris-buffered saline Tween-20
TEMED	tetramethylethylenediamine
TIR	Toll/IL-1 receptor
TLR	Toll–like receptors
Tm	melting temperature
TNF	tumor necrosis factor
TNFR	TNF receptor
TR	transmembrane region
TRAF1	TNFR-associated factor 1
TRAF2	TNFR-associated factor 2
TRAIL	TNF-related apoptosis inducing ligand
TRAM	TRIF-related adaptor molecule
TRIAD	*T. adhaerens*, *Trichoplax adhaerens*
TRIF	TIR-domain-containing adapter-inducing interferon-β

U

UBC	ubiquitin-conjugating catalytic core domain
UV	ultraviolet

V

V	volt
v/v	volume per volume
VITVI	*V. vinifera*, *Vitis vinifera*

W

w/v	weight per volume

X

XAF1	X-linked IAP-associated factor
XENLA	*X. laevis*, *Xenopus laevis*

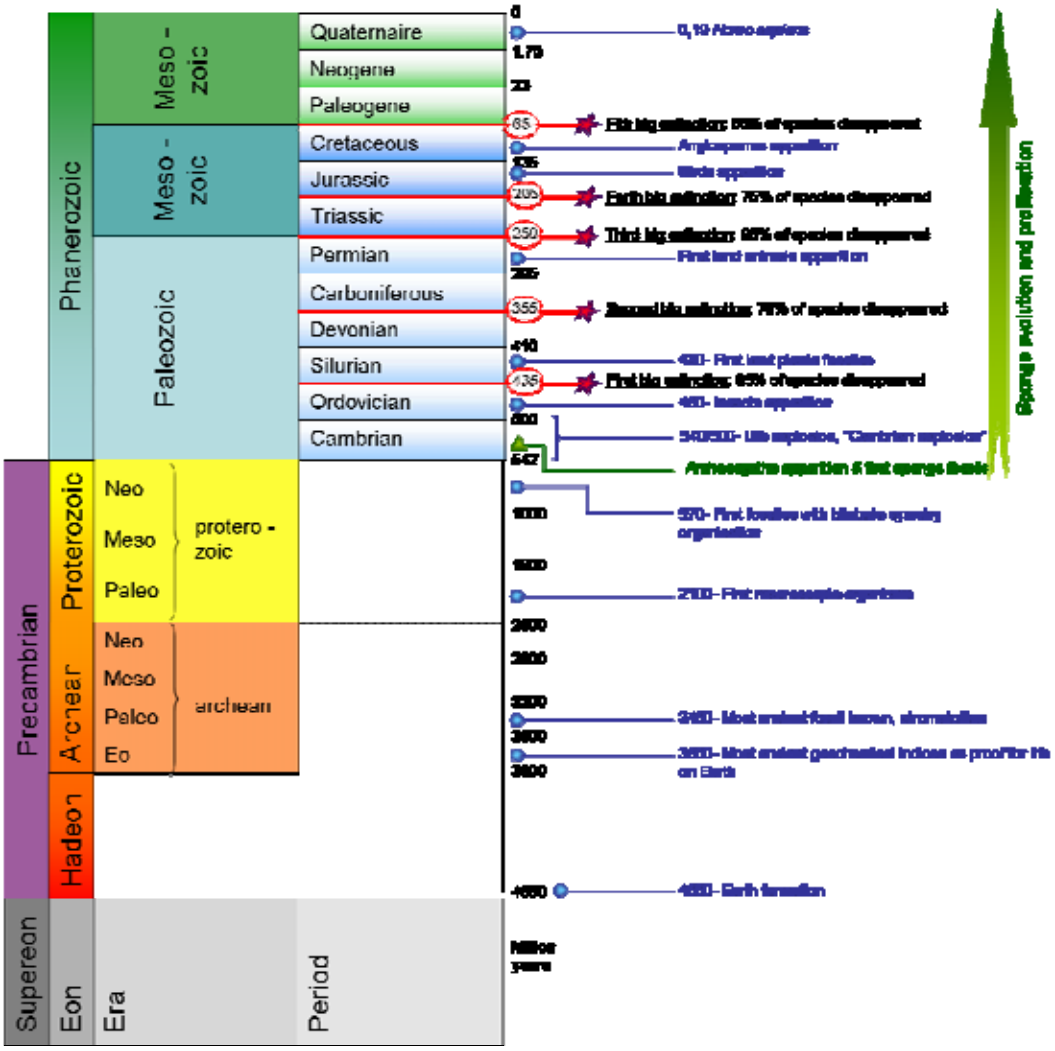
X. INDEX OF FIGURES AND TABLES

A. Index of figures

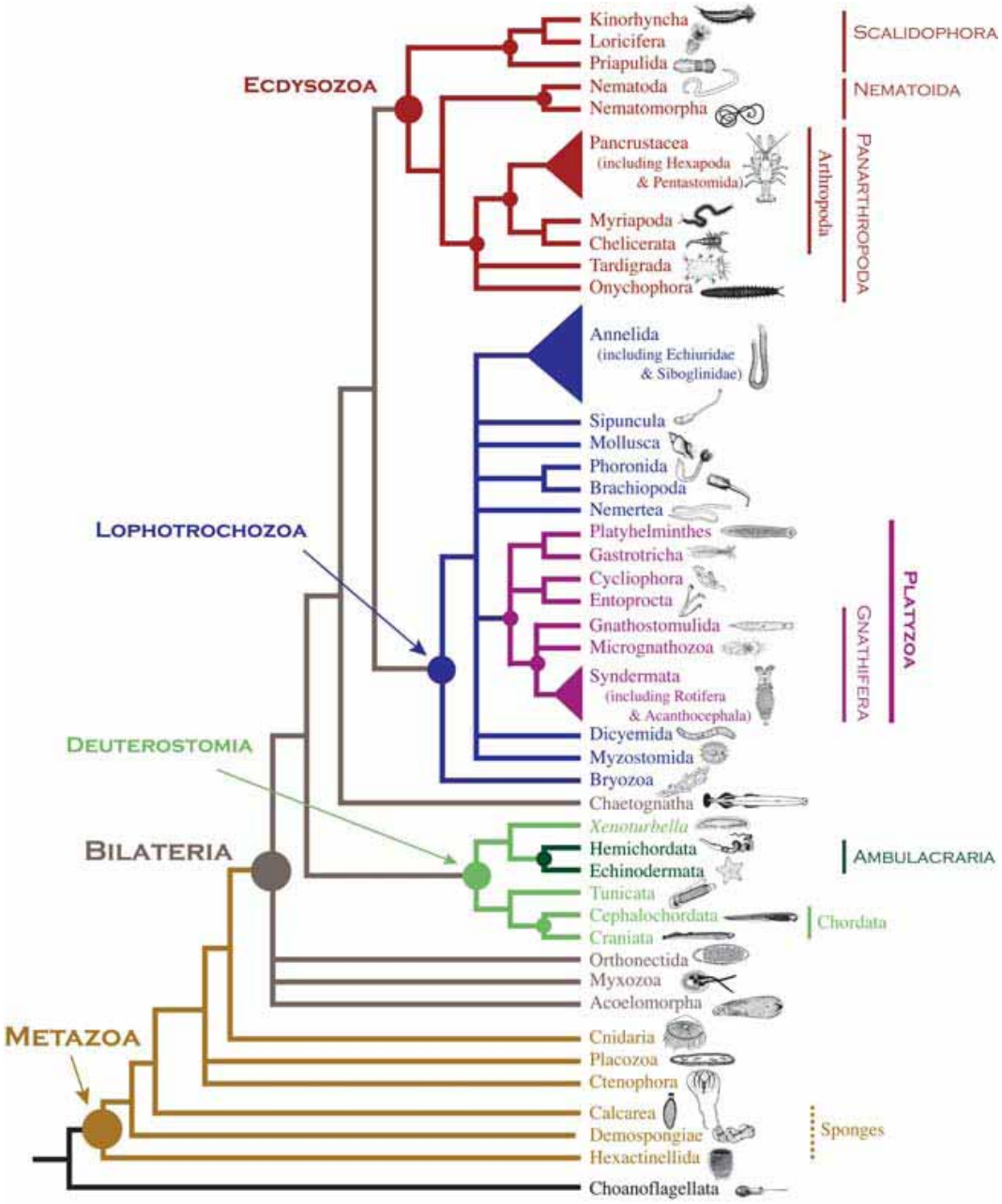
B. Index of tables

XI. APPENDIXES

A. Appendix 1: geologic time-scale



B. Appendix 2: modern view of animal phylogeny based largely on molecular data (from Halanych KM 2004)



C. Appendix 3: legal notice

„Ich erkläre, dass ich die vorgelegte Thesis selbständig, ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt habe, die ich in der Thesis angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der Satzung der Johannes Gutenberg Universität Mainz zur Sicherung guter wissenschaftlicher Praxis niedergelegt sind, eingehalten."

D. Appendix 4: publications

Wiens M, Korzhev M, Perović-Ottstadt S, Luthringer B, Brandt D, Klein S, Müller WEG. 2007. Toll–like receptors are part of the innate immune defense system of sponges (Demospongiae: Porifera). Mol Biol Evol. 24:792–804.

Luthringer B, Isbert S, Thakur NL, Kelve M, Müller WEG, Wiens M. 2009. Poriferan survivin exhibits a conserved regulatory role in the interconnected pathways of cell cycle and apoptosis. Submitted.