THESIS ON NATURAL AND EXACT SCIENCES B178

Characterization of the Oligoadenylate Synthetase Subgroup from Phylum Porifera

MAILIS PÄRI

TUT
PRESS

TALLINN UNIVERSITY OF TECHNOLOGY
Faculty of Science
Department of Gene Technology

This dissertation was accepted for the defence of the degree of Doctor of Philosophy in Gene Technology on 28 October, 2014.

Supervisors: Anne Kuusksalu, PhD,
Department of Gene Technology, Tallinn University of Technology, Estonia

Annika Lopp, PhD,
Department of Gene Technology, Tallinn University of Technology, Estonia

Opponents: Assoc. Prof. Rune Hartmann, PhD,
Department of Molecular Biology, University of Aarhus, Denmark

Prof. Mart Saarma, PhD,
Institute of Biotechnology, University of Helsinki, Finland

Defence of the thesis: 10 December, 2014

Declaration:
I hereby declare that this doctoral thesis, my original investigation and achievement, submitted for the doctoral degree at Tallinn University of Technology has not been previously submitted for doctoral or equivalent academic degree.

/Mailis Päri/

ISSN 1406-4723
ISBN 978-9949-23-697-8 (publication)
ISBN 978-9949-23-698-5 (PDF)

LOODUS- JA TÄPPISTEADUSED B178

Hõimkonna Porifera oligoadenülaadi süntetaaside alamrühma iseloomustamine

MAILIS PÄRI

TTÜ
KIRJASTUS

CONTENTS

LIST OF PUBLICATIONS

This thesis is based on the following publications that are referred to in the text by their Roman numerals:

I **Päri M**, Kuusksalu A, Lopp A, Reintamm T, Justesen J, Kelve M. (2007). Expression and characterization of recombinant 2',5'-oligoadenylate synthetase from the marine sponge *Geodia cydonium*. *FEBS J*, 274(13):3462-3474.

II Reintamm T, Kuusksalu A, Metsis M, **Päri M**, Vallmann K, Lopp A, Justesen J, Kelve M. (2008). Sponge OAS has a distinct genomic structure within the 2-5A synthetase family. *Mol Genet Genomics*, 280(5):453-466.

III **Päri M**, Kuusksalu A, Lopp A, Kjaer K, Justesen J, Kelve M. (2014). Enzymatically active 2',5'-oligoadenylate synthetases are widely distributed among Metazoa, including protostome lineage. *Biochimie*, 97:200-209.

INTRODUCTION

2',5'-oligoadenylate synthetases (2-5A synthetases, OASs, EC 2.7.7.84) are the enzymes that catalyse the polymerization of ATP into 2',5'-linked oligoadenylates. They belong to the nucleotidyl transferase superfamily subclass of template-independent nucleotidyl transferases, together with poly(A) and poly(U) polymerases and CCA-adding enzymes. These enzymes are believed to have derived from a common ancestor. In the course of evolution, each of them has gained its specific enzymatic properties and functions to carry out in the cell. The OASs became 2'-specific while the majority of nucleotidyl transferases are 3'-specific, catalysing the formation of 3',5'-phosphodiester linkages. Moreover, the OASs acquired the ability to initiate synthesis without requiring a primer. In contrast to constitutively active nucleotidyl transferases, OASs evolved to be regulated by a specific cofactor, double-stranded RNA.

Nucleotidyl transferases catalyse the transfer of nucleotides to a variety of substrates and they are involved in many important biological processes. For OASs, their main role is described in the innate immune system, where they mediate the interferon-induced response to several viruses. However, other functions for these proteins have been proposed. The preservation of multiple OAS genes in mammalian genomes that encode different types and isoforms of OAS proteins, refers to their diverse functions in cells.

In addition to mammals, where the OAS family has been well studied, the genes encoding these proteins have been found in the genomes of animals from many phyla over the phylogenetic tree of multicellular animals. These phyla comprise vertebrates and invertebrates, including the evolutionarily lowest animals, sponges, as well as lineages from the Protostomia branch of animal evolution. Moreover, the presence of OAS genes has been predicted in some unicellular organisms. However, the OAS activity has only been demonstrated in higher vertebrates (mammals, birds and reptiles) and in marine sponges.

This thesis focuses on describing the structure and enzymatic properties of OASs from different marine sponges. It demonstrates the diversity of OASs from marine sponges. The enzymatic characteristics and genomic structures of sponge OASs were elucidated and compared to those of mammalian OASs. In addition, the presence of active OAS proteins in several distantly related animal species was studied.

The results obtained in this work contribute to a better understanding of the evolution of 2',5'-oligoadenylate synthetases as well as their emergence from the common nucleotidyl transferase ancestor.

ABBREVIATIONS

2-5A	2',5'-linked oligoadenylates
AP-1	activating protein 1
CCAtr	CCA-adding enzyme
cGAS	cyclic GMP-AMP synthase
dNTP	deoxynucleotide triphosphate
dsDNA	double-stranded DNA
dsRNA	double-stranded RNA
EMCV	Encephalomyocarditis virus
EST	expressed sequence tag
ets-1	E26 transformation-specific 1
GAS	IFN-γ activation sequence
GST	glutathione S-transferase
HIV	Human immunodeficiency virus
HTLV	Human T-lymphotropic virus
IFN	interferon
IMAC	immobilized metal ion affinity chromatography
IRF	IFN regulatory factor
IRF-E	IFN regulatory factor element
ISRE	IFN-stimulated response element
kanNT	kanamycin nucleotidyl transferase
MALDI-TOF MS	matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry
NF-κB	nuclear factor-kappaB
NRAP	nucleolar RNA-associated protein
NRM	nucleotide recognition motif
NT	nucleotidyl transferase
NTP	nucleoside triphosphate
OAS	2',5'-oligoadenylate synthetase
OASL	2',5'-oligoadenylate synthetase-like protein

PAP	poly(A) polymerase
PDB	Protein Data Bank
pol β	DNA polymerase β
poly(A)·poly(U)	polyadenylic:polyuridylic acid
poly(I)·poly(C)	polyinosinic:polycytidylic acid
PUP	poly(U) polymerase
rDNA	ribosomal DNA
RICS	RNA-induced catalytic structure
RLI	RNase L inhibitor
RNase L	ribonuclease L
rNT	ribonucleotidyl transferase
RP-HPLC	reverse-phase high-performance liquid chromatography
RSV	Respiratory syncytial virus
SEC	size exclusion chromatography
SNP	single nucleotide polymorphism
SP1	specificity protein 1
TdT	terminal deoxynucleotidyl transferase
TUT	terminal uridylyltransferase
UbL	ubiquitin-like sequence
WNV	West Nile virus
wt	wild type

REVIEW OF THE LITERATURE

1. Nucleotidyl transferases

Nucleotidyl transferases (NTs, EC 2.7.7.-) catalyse the transfer of a nucleotide to an acceptor hydroxyl group of nucleic acids, proteins or antibiotics, a reaction important in many biological processes, such as DNA replication and repair, transcription, RNA processing and viral replication (Aravind and Koonin, 1999). According to their sequence and structural homologies, nucleotidyl transferases are grouped into distinct superfamilies (Ito and Braithwaite, 1991). The most widespread superfamily of nucleotidyl transferases is typified by the eukaryotic DNA polymerase β and sometimes referred to as the pol β superfamily (Aravind and Koonin, 1999; Betat et al., 2010).

Members of this family include DNA polymerases β, μ and λ, terminal deoxynucleotidyl transferases (TdT), antibiotic nucleotidyl transferases, protein nucleotidyl transferases, poly(A) polymerases (PAP), poly(U) polymerases (PUP), CCA-adding enzymes (CCAtr), 2',5'-oligoadenylate synthetases (OAS) and cyclic GMP-AMP synthases (cGAS) (Holm and Sander, 1995; Aravind and Koonin, 1999; Kwak and Wickens, 2007; Yamtich and Sweasy, 2010; Civril et al., 2013).

1.1. Classification

Nucleotidyl transferases of the pol β superfamily are classified into Class I and Class II enzymes. The classification is based on the presence of 4 additional amino acid residues that form an extra turn in the active site α-helix in Class II NTs. Class I consists of DNA pol β, TdT, eukaryotic PAP, OAS and arhaeal CCAtr while Class II is a small group comprising eubacterial PAPs and both eukaryotic and eubacterial CCAtrs, as well as A-adding and CC-adding enzymes (Yue et al., 1996; Martin et al., 2008). Class II enzymes share a well-conserved 25-kDa N-terminal domain, including the highly conserved active site helix and a RRD-motif which contains the third active site carboxylate and participates in substrate selection (Yue et al., 1996; Martin and Keller, 2007). Class I enzymes are more heterogenous and have little sequence homology outside the active site signature motif (Yue et al., 1996).

Class I ribonucleotidyl transferases (rNTs) are further divided into canonical and non-canonical rNTs based on the characteristic nucleotide recognition motif (NRM). Canonical rNTs are classical nuclear PAPs, which polyadenylate the pre-mRNAs. Their signature motif MPIITPAYPQQN (NRM 1) is almost invariant in canonical rNTs of all eukaryotes. The consensus sequence of the NRM of non-canonical rNTs (NRM 2) is hx[I/L/V][E/Q][E/D/N]Phx$_4$N (where x is any and h is a hydrophobic amino acid). Non-canonical rNTs include PAPs and PUPs from the Cid1 family, terminal uridylyltransferases (TUTs) and OASs

(Stevenson and Norbury, 2006; Martin and Keller, 2007; Martin et al., 2008).

It has been proposed that most of the eukaryotic canonical and non-canonical rNTs evolved from the archaeal CCAtrs, whereas the eukaryotic CCAtrs are the descendants of eubacterial Class II CCAtrs (Martin and Keller, 2007; Martin et al., 2008).

Since the OAS studied in this work is a Class I rNT, the following review focuses mainly on describing the properties of Class I enzymes.

1.2. Template-independent rNTs

Template-independent rNTs transfer ribonucleotides to the 3'-end of the primer without using a nucleic acid template for nucleotide selection. Some rNTs may add a single nucleotide to the primer molecule while others synthesise a homopolymeric tail of several hundred nucleotides in length. Most rNTs catalyse the transfer of one specific ribonucleotide, usually ATP or UTP, but unusual specificity is displayed by the CCAtrs, which catalyse the addition of the trinucleotide CCA to the 3'-ends of tRNAs (Martin et al., 2008).

1.2.1. Overall structure

The NTs of the pol β superfamily share a common core fold and conserved active site residues. However, their sequence similarity is often low (Kuchta et al., 2009). Based on their common architecture, the template-independent rNTs can be defined as a separate subclass in the polβ superfamily. Their catalytic site is located between the classic β-sheet domain and the α-helical bundle domain (Fig. 1). The latter is unique to the template-independent rNTs (Torralba et al., 2008). The β-sheet domain is homologous to the palm domain in template-dependent NTs (Torralba et al., 2008) and it is found in all members of the pol β superfamily (Martin et al., 2000; Li et al., 2002; Xiong et al., 2003; Augustin et al., 2003; Lunde et al., 2012).

The first published crystal structure of a template-independent polymerase was that of bovine PAP bound to an analog of ATP, 3'-dATP (Martin et al., 2000). It was shown that PAP was a U-shaped molecule composed of three domains (Fig 1A). The N-terminal domain was homologous to the catalytic domain of nucleotidyl transferases and comprised a five-stranded mixed β-sheet and two α-helices. The central domain, however, was not homologous to the pol β fingers domain, but instead it shared a structural similarity with the allosteric activity domain of ribonucleoside reductase R1, which comprises a four-helix bundle and a three-stranded mixed β-sheet. The C-terminal domain of PAP acquired a compact globular structure consisting of a four-stranded antiparallel β-sheet, flanked on one side by two helices. This domain was similar to the RNA-binding domains of several RNA-binding proteins. The catalytic and RNA-binding domains were connected to the central domain through a central β-strand and via a hinge-like region, respectively (Martin et al., 2000).

The Class I CCAtr from *Archaeoglobus fulgidus* consists of four domains named head, neck, body and tail (Fig. 1C). The former three domains form a U-

shape structure similar to that of PAP while the tail domain is loosely connected to the rest of the enzyme and is involved in dimer formation by interacting with the other subunit. The head domain is constructed of five-stranded β-sheet flanked by α-helices and corresponds to the catalytic domain of PAP and other NTs. The neck domain consisting of mixed β-sheets and α-helices is the equivalent of the central domain of PAP. The body domain is homologous to the C-terminal RNA-binding domain of PAP and is involved in tRNA binding. However, the Class I CCAtr and PAP differ in the orientation of the RNA-binding domain relative to the central domain by a rotation of about 30° (Xiong et al., 2003; Okabe et al., 2003).

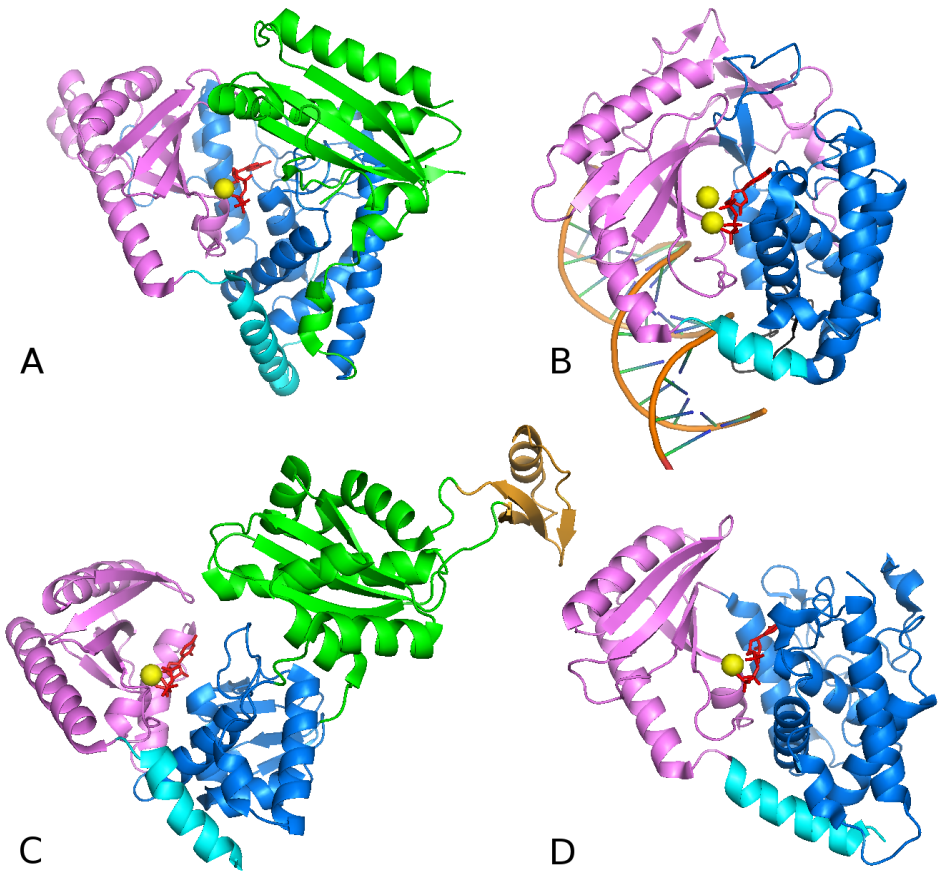


Figure 1. Examples of three-dimensional structures of template-independent rNTs: bovine PAP in complex with 3'-dATP (A; PDB ID: 1Q78), murine cGAS with dsDNA and ATP (B; PDB ID: 4K97), archaeal CCAtr in complex with CTP (C; PDB ID: 1R89) and yeast Cid1 with bound UTP (D; PDB ID: 4FH5). The N-terminal extension (part of the non-contiguous central domain) is coloured cyan, the NT catalytic domain violet, the central domain blue and the C-terminal domain green. The tail domain of CCAtr is shown in orange and the Zn thumb of cGAS in black. The divalent metal ions are shown as yellow spheres and the nucleotides are shown in red. The figure was created using the PyMOL™ Molecular Graphics System.

The non-canonical polymerases Trf4p (for DNA topoisomerase related function, a poly(A) polymerase) and Cid1 (for caffeine-induced death suppressor, a poly(U) polymerase) adopt similar structures to the previously described enzymes (Fig. 1D; Hamill et al., 2010; Lunde et al., 2012). However, like the majority of other non-canonical PAPs and PUPs of the Cid1 family, they consist only of two domains and lack the RNA-binding domain. Therefore they may require additional RNA-binding proteins for interacting with their substrates (Stevenson and Norbury, 2006; Hamill et al., 2010).

The two primer-independent rNTs of the polβ superfamily, the OAS and the cGAS, also consist of only two domains – the N-terminal NT domain and the C-terminal domain corresponding to the central domain of PAP and CCAtr (Hartmann et al., 2003; Donovan et al., 2013; Kranzusch et al., 2013). As primer-independent enzymes they do not need the RNA-binding domain which in primer-dependent rNTs interacts with the substrate (e.g. mRNA for PAP and tRNA for CCAtr). The OAS and cGAS are unique among the NTs as they require binding of a cofactor (dsRNA or dsDNA, respectively) and subsequent conformational rearrangements in the N-terminal catalytic domain for the assembly of their active sites (Donovan et al., 2013; Civril et al., 2013). In cGAS the activating DNA is bound to the positively charged cleft between the N-terminal extension and the C-terminal domain (Fig. 1B). A zinc-binding motif essential for DNA binding is located in the C-terminal domain (Civril et al., 2013; Kranzusch et al., 2013). This motif is not found in other NTs and is proposed to preclude dsRNA from binding to cGAS due to steric hindrance (Xiao and Fitzgerald, 2013). The structure and activation mechanism of OAS will be described in Sections 2.1.2 and 2.2.3, respectively.

1.2.2. Active site

The active site of Class I template-independent rNTs is located on one side of the U-shaped molecule, between the N-terminal catalytic domain and the central domain and is composed of residues from both domains (Martin et al., 2000). The active site consensus signature is $hG[GS]x_{7\text{-}13}[D/E]h[D/E]h$, where h is a hydrophobic and x is any amino acid residue (Martin and Keller, 2007; Kuchta et al., 2009). In the crystal structures, the two carboxylate residues are located in the beginning of the β-strand 2 and together with the third aspartate or glutamate residue, located on the spatially adjacent β-strand, form the catalytic triad (Fig. 2). Their side chains are directed towards the central cleft where they coordinate two catalytic divalent metal ions, Mg^{2+} or Mn^{2+}. Conserved residues from the helical turn motif (P-loop) between the β-strands 1 and 2 as well as from two central domain α-helices interact with the phosphate groups of the nucleotide (Martin et al., 2000; Okabe et al., 2003; Martin et al., 2004; Lunde et al., 2012). Residues from the nucleotide recognition motif, located in the central domain, are responsible for interacting with the base of the nucleotide (Martin and Keller, 2007).

1.2.3. Catalytic mechanism

Despite adding different nucleotides and using a variety of substrates, the NTs appear to generally utilize the same mechanism of nucleotidyl transfer, involving a two-metal-coordinated attack by the 3'-OH of the primer on the α-phosphate of the incoming nucleoside triphosphate (Martin and Keller, 1996; Kwak and Wickens, 2007). The exceptions are OAS and kanamycin nucleotidyl transferase (kanNT), that are 2'- and 4'-specific, respectively. It is believed, that these enzymes also use the same catalytic mechanism and place their different nucleophilic atoms in similar positions to enable the attack of the donor α-phosphate (Hartmann et al., 2003).

The catalytic mechanism of nucleotide transfer by rNTs has been well studied, using crystal structures of mammalian and yeast PAPs in their complexes with catalytic metal ions and ATP analogs (Martin et al., 2000; Martin et al., 2004; Balbo and Bohm, 2007; Balbo and Bohm, 2009; Yang et al., 2014). One of the divalent metal ions binds as a co-substrate while the other coordinates the 3'-OH of the primer and the α-phosphate group of the incoming nucleotide (Balbo and Bohm, 2009). The protein contacts with the phosphates of the nucleotide are mediated by the conserved serine residue (Ser102 in bovine PAP) in the P-loop preceding the catalytic triad as well as lysine and tyrosine residues (Lys228 and Tyr237 in bovine PAP) from the central domain. The ribose is located between the Phe from the P-loop and Val247 (bovine PAP). The latter also interacts with the adenine base together with Thr317 from the NRM. 2'-OH of the ribose is bound by a water molecule coordinated by Thr207 (Fig. 2; Martin et al., 2000; Martin et al., 2004; Yang et al., 2014).

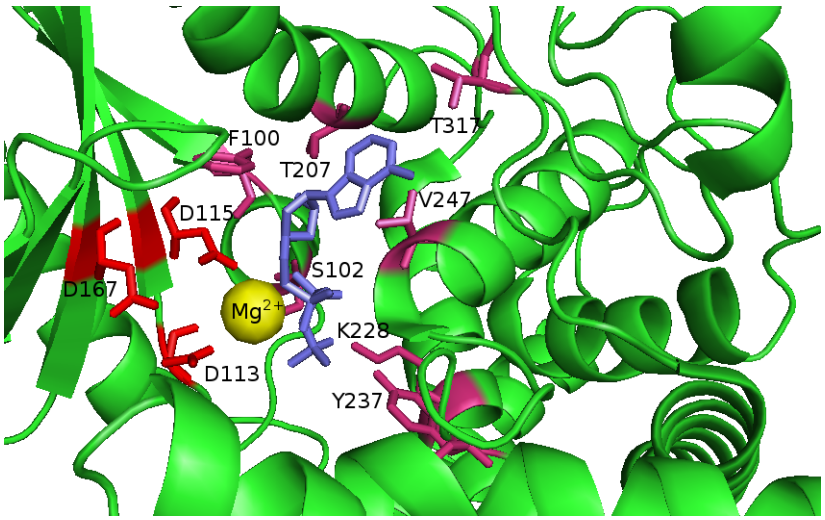


Figure 2. The structure of the active site of bovine PAP in complex with an ATP analog, 3'-dATP. The three catalytic aspartate residues are coloured red, other residues interacting with ATP are shown in purple, the Mg^{2+}-ion is represented by a yellow sphere and 3'-dATP is coloured blue. The figure was created from PDB file 1Q78 (Martin et al., 2004) using the PyMOL™ Molecular Graphics System.

The binding of both, ATP and primer, induce a conformational change resulting in the closed enzyme state where the nucleotidyl transfer reaction is catalysed (Balbo and Bohm, 2007). The 3'-OH of the primer is deprotonated and attacks the α-phosphate of the nucleotide. One active site metal ion promotes the phosphodiester bond formation in a transesterification reaction producing the pyrophosphate leaving group that departs with the other metal ion in the form of $MgPP_i^{2-}$ (Hartmann et al., 2003; Martin et al., 2008; Balbo and Bohm, 2009). After each round of catalysis the domain opening is induced and the product is released (Balbo and Bohm, 2007).

1.2.4. Specificity

Based on the structural similarity it is proposed that the catalytic domains of the enzymes of the polβ superfamily share a common ancestor (Okabe et al., 2003). The presence of various additional structural elements and domains with distinct enzymatic and regulatory activities is dependent on the enzyme family (Aravind and Koonin, 1999; Kuchta et al., 2009). Evidently, these differences are the basis for the distinct substrate specificity displayed by these enzymes and determine the choice of nucleotide to be transferred.

Class I CCAtr recognizes the substrate tRNA molecule with its body and tail domains. For efficient tRNA binding the enzyme dimerizes via its tail domain and the dimer binds two tRNA molecules. The acceptor stem and the T-stem loop of the tRNA are inserted into a cleft in the enzyme molecule while the 3'-terminus is located in the active site pocket. The tRNA molecule remains fixed during the CCA addition (Xiong and Steitz, 2004; Tomita et al., 2006; Tomita and Yamashita, 2014). The C-terminal domain of PAP, responsible for the substrate binding, interacts with the mRNA in a sequence non-specific manner. Only a few base-specific interactions occur between the protein and the mRNA (Balbo and Bohm, 2007).

In CCAtr, the protein-tRNA complex forms a template for the correct nucleotide addition. In the absence of the tRNA, the nucleotide binding was observed to be nonspecific (Xiong et al., 2003). The incoming CTP forms hydrogen bonds with the phosphate groups of the tRNA and with the specific arginine residue of the central domain of the enzyme. These interactions exclude GTP and UTP from binding (Xiong and Steitz, 2004). The ATP addition is catalysed by one-metal-ion mechanism, since the binding of the other metal ion is sterically hindered (Pan et al., 2010). In PAP, steric differences exclude CTP and UTP from binding to the active site. ATP is selected over GTP by the hydrogen bonding of N1 atom in ATP and a tyrosine residue in PAP (Martin et al., 2004). In TUTs, the binding of ATP and GTP to the active site is suppressed due to steric hindrance. The specificity of UTP over CTP is determined by a water molecule between the conserved carboxylates, which detects the presence of a hydrogen atom in the N3 position of uridine base (Deng et al., 2005). In the Cid1 family PUPs, a histidine residue in the NRM was shown to be important for the UTP selection; in Cid1 family PAPs, an asparagine residue is found in this position (Lunde et al., 2012).

2. 2',5'-oligoadenylate synthetases

2',5'-oligoadenylate synthetases are interferon-induced and dsRNA activated enzymes that catalyse the polymerization of ATP into 2',5'-linked oligoadenylates. They belong to the nucleotidyl transferase superfamily subclass of template-independent NTs together with PAPs and CCAtrs (Aravind and Koonin, 1999; Torralba et al., 2008). It has been hypothesised that OASs evolved from Trf proteins (Rogozin et al., 2003), the non-canonical PAPs belonging to Cid1 family, members of which also lack the RNA-binding domain as OASs do (Stevenson and Norbury, 2006; Hamill et al., 2010). OASs became 2'-specific, catalysing the formation of 2',5'-phosphodiester bond instead of the 3',5'-bond catalysed by most rNTs. Together with the ability to initiate synthesis of oligoadenylates from ATP without requiring any primer, OAS also evolved from constitutively active NT to a latent enzyme that needs dsRNA binding to induce conformational changes in order to assemble its active site (Kristiansen et al., 2011; Donovan et al., 2013). In mammals OAS is an important component of the IFN-induced antiviral system. OAS and the dsDNA sensor cGAS are proposed to be a part of a second-messenger system that rapidly signals upon pathogen recognition (Kranzusch et al., 2013).

2.1. Structure of OAS

2.1.1. Gene structure

The human OAS family consists of four genes encoding four types of proteins with different isoforms – the small form (p42/p44/p46/p48, OAS1), the medium form (p69/p71, OAS2), the large form (p100, OAS3) and the oligoadenylate synthetase-like protein (p59, OASL), which contains C-terminal ubiquitin-like domain and does not exhibit 2-5A synthesising activity (Justesen et al., 2000).

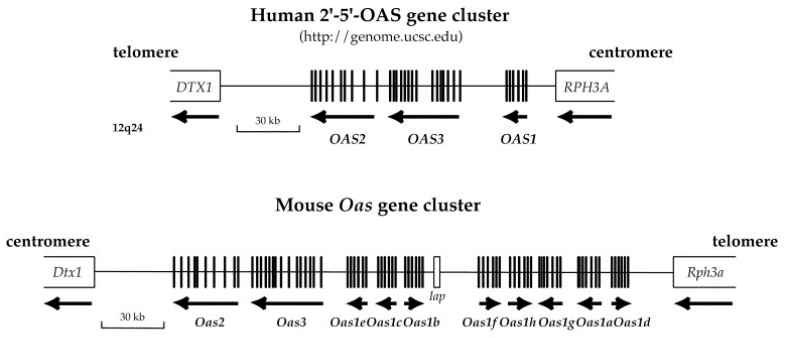


Figure 3. Schematic organization of *OAS* genes in human and mouse genomes. Exons are represented by black boxes and the direction of transcription is shown by arrows. Reprinted from Mashimo et al., 2003, Copyright (2003), with permission from Elsevier.

The three genes (*OAS1*, *OAS2* and *OAS3*) that encode OASs with enzymatic activity are clustered within a region of 105 kb on chromosome 12q24.13. They are arranged in the order centromere-5'-*OAS1*-*OAS3*-*OAS2*-3'-telomere and have the same orientation of transcription (Fig. 3; Rebouillat and Hovanessian, 1999; Hovanessian and Justesen, 2007). The distance between *OAS1* and *OAS3* is approximately 20 kb and between *OAS3* and *OAS2* about 5 kb (Rebouillat et al., 2000). The *OASL* is located 8 Mb towards telomere, in the region 12q24.31 (Hovanessian and Justesen, 2007).

Several transcripts are transcribed from human *OAS1* gene, generated by the alternative splicing of exons 5, 6 and 7 (Fig. 4). The encoded protein isoforms – p42, p44, p46 and p48 – have identical N-termini of 346 residues but different C-termini (Mashimo et al., 2008). *OAS2* is spliced in two isoforms, p69 and p71, which share a common N-terminus of 683 residues and different C-termini of 4 and 44 residues, respectively (Mashimo et al., 2008). *OAS3* is composed of 16 exons and encodes only one protein isoform, p100 (Rebouillat et al., 2000; Hovanessian and Justesen, 2007). *OASL* gene encodes a two-domain protein, p56/p59, consisting of an OAS unit and a 164 amino acid C-terminal domain of two ubiquitin-like repeats (Rebouillat et al., 1998; Hartmann et al., 1998b; Eskildsen et al., 2003).

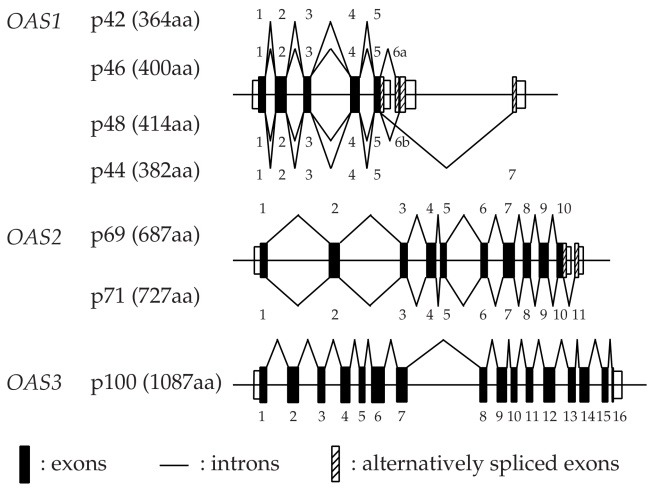


Figure 4. Exon/intron structure and splice variants of the human *OAS* genes. The scale for exons is five times larger than that for introns. Reprinted from Mashimo et al., 2003, Copyright (2003), with permission from Elsevier.

The analysis of the coding sequences of the OAS genes reveals the presence of a conserved domain of five exons referred to as the OAS unit. This domain occurs once in OAS1 and OASL proteins (the first 346 aa), is repeated twice in OAS2 and three times in OAS3 proteins (Rebouillat et al., 2000). There is a high degree of amino acid identity between all members of each of the five exon groups (Fig. 5). Furthermore, the splice sites and reading frames across the exon/intron boundaries are also conserved. In contrast, the introns show no

similarities and vary in size between the OAS genes (Fig. 4). This exon conservation suggests that the duplications of an ancestral gene are the origin of the OAS gene family. However, the intron size variations indicate that this occurred rather early during evolution (Justesen et al., 2000).

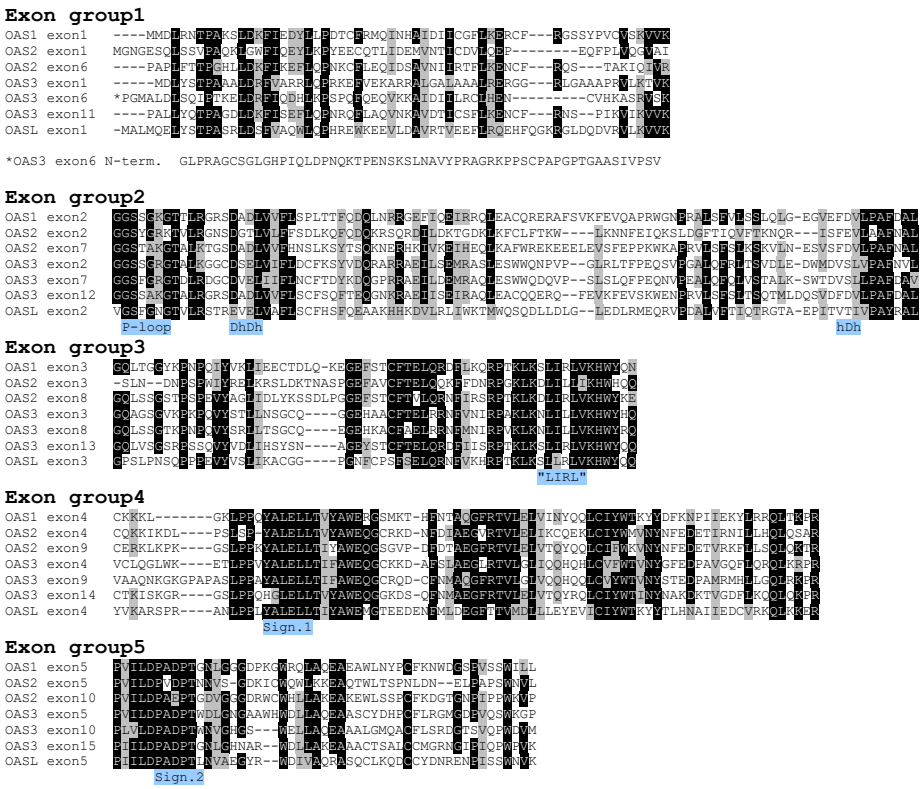


Figure 5. Alignment of the protein sequences encoded by the exons of the OAS units in human *OAS1*, *OAS2*, *OAS3* and *OASL* genes (GenBank accession NC_000012.12 bases 112906934-112919907, 112978469-113011723, 112938352-112973249 and 121020292-121039242, respectively). The alignment was created using ClustalX (Thompson et al., 1997) and modified manually. The identical amino acids are indicated with black boxes and similar amino acids with grey boxes, the OAS-specific motifs are shown. Based on (Justesen et al., 2000).

The 72 bp region (named E-IRS, -159 to -87) upstream of *OAS1* contains two enhancer-like DNA elements, A (-159 to -122) and B (-109 to -87). They are both required for the IFN-induced *OAS1* promoter activity. However, their activity is independent of their orientation and position and they do not have to be physically close to each other. The element A, carrying an IRF-E and GAS site, has been proposed to function as a constitutive enhancer and element B, which carries an ISRE site and is located downstream of A, has been suggested to be the specific target of IFN-β_1 action. Downstream of the E-IRS, the presence of sequences which bind a putative repressor was suggested (Benech

et al., 1987; Cohen et al., 1988; Justesen et al., 2000).

The structures of the *OAS2* and *OAS3* promoters are more complex than that of the *OAS1* promoter. The *OAS2* promoter is located in a region 972 bp upstream from the translational start site. It contains an ISRE consensus sequence, two IRF-1-like elements (I1 and I2), a putative GAS element, an element resembling a NF-κB site, an AP-1 binding site and an ets-1 site (Wang and Floyd-Smith, 1998). The *OAS3* promoter region was identified as a 780 bp region upstream of the translation start site and it is composed of at least one ISRE, two IRF1-binding motifs, two GAS elements and one NF-κB site. In addition to the classical elements of IFN-induced genes, the promoter of *OAS3* also contains binding motifs for transcription factors responsible for a basal level of transcription, including SP1 and AP1-like site (Rebouillat et al., 2000).

Despite the presence of common motifs, the promoters of the three human *OAS* genes do not possess a strong sequence homology. Similarly to the evolution of the genomic sequences (conservation of exonic sequences and divergence of intronic sequences), the evolution of the regulatory sequences seems to be related to the conservation of the motifs indispensable for the transcriptional induction by IFN (Rebouillat et al., 2000). None of the human *OAS* genes contain a TATA or a CAAT box in their promoter sequences (Cohen et al., 1988; Wang and Floyd-Smith, 1997; Rebouillat et al., 2000).

The mouse *OAS* gene family contains one *Oas2*, one *Oas3* gene and eight *Oas1* genes (*Oas1a – Oas1h*) (Eskildsen et al., 2002; Kakuta et al., 2002). They are located in a ~200 kb region of the mouse chromosome 5F (Fig. 3). The two *Oasl* (*Oasl1* and *Oasl2*) genes are located 6 Mb towards centromere (Mashimo et al., 2008). Interestingly, the protein encoded by *Oasl2* possesses 2',5'-oligoadenylate synthetase activity while the protein encoded by *Oasl1* is enzymatically inactive like human OASL (Eskildsen et al., 2003). The mouse genomic segment harboring the *OAS* genes together with a few other flanking genes is inverted when compared to the human *OAS* locus (Mashimo et al., 2008). *Oas2* and *Oas3* share similar exon/intron structure with corresponding human genes and are transcribed in the same orientation as well. The *Oas1* genes also have a similar exon/intron structure, but they are ordered differently – *Oas1a*, *Oas1c*, *Oas1e* and *Oas1g* are in the same orientation as human *OAS1*, while *Oas1b*, *Oas1d*, *Oas1f* and *Oas1h* are transcribed in the opposite direction (Fig. 3; Eskildsen et al., 2002; Mashimo et al., 2003). In contrast to human *OAS1* which encodes several alternative transcripts, the murine *Oas1* genes are not alternatively spliced. The exception is *Oas1a*, which encodes two transcripts differing in exon 6 (Mashimo et al., 2003). However, most of the murine *Oas1* genes encode enzymatically inactive proteins (Kakuta et al., 2002). Two pseudogenes, *Oas1i* and *Oas1j*, were identified 5' of the *Oas1d* and between *Oas1b* and *Oas1f*, respectively. They both have adjacent retrovirus-like elements, suggesting that a retroelement integration has led to gene duplication in this region (Kakuta et al., 2002; Perelygin et al., 2006).

The presence of OAS genes in other animals is discussed in Section 2.4.1.

2.1.2. Protein structure

To date, three-dimensional structures of two mammalian OAS1 proteins have been determined (PDB IDs: 1PX5 and 4IG8; Hartmann et al., 2003; Donovan et al., 2013). They show that the OAS monomer consists of two lobes, the N-terminal catalytic domain and the C-terminal domain, connected by a helix-loop-helix linker. The overall structure of OAS (Fig. 6) has a similar fold to that of poly(A) polymerase in its catalytic and central domains. OAS does not have the equivalent of PAP C-terminal domain that binds to the mRNA. The N-terminal catalytic domain of OAS is structurally highly similar to the catalytic domains of PAP (Fig. 1A), CCAtr (Fig. 1C) and DNA polymerase β. The C-terminal domain of OAS has structural similarity to the central domain of PAP (Hartmann et al., 2003; Kristiansen et al., 2011).

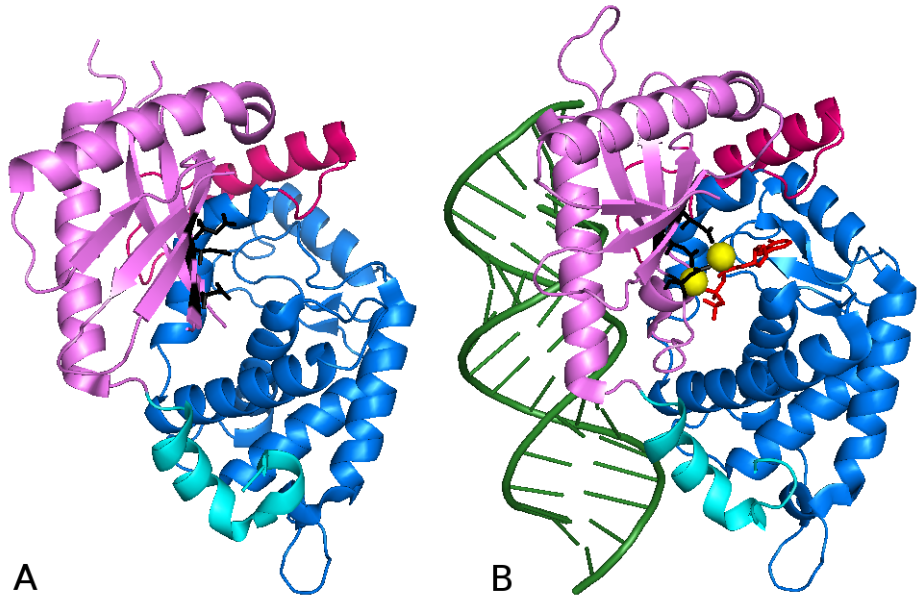


Figure 6. Crystal structures of the porcine OAS1 in apo form (A, PDB ID: 1PX5) and of human OAS1 in complex with dsRNA (green), substrate analog 3'-dATP (red) and catalytic Mg^{2+}-ions (yellow spheres) (B, PDB ID: 4IG8). The N-terminal extension (cyan) leads to the N-terminal catalytic domain (violet), interdomain linker (purple) and C-terminal domain (blue). The catalytic aspartate residues are coloured black. The figure was created using the PyMOL™ Molecular Graphics System.

The N-terminal catalytic domain consists of the αββαβββ motif forming a twisted five-stranded antiparallel β-sheet. The C-terminal domain of OAS is formed by a four helix bundle and a three-stranded mixed β-sheet with an additional interstrand connecting α-helix. The two domains are secured together by a short N-terminal extension that extends across to the C-terminal domain (Fig. 6; Hartmann et al., 2003; Kristiansen et al., 2011).

The active site of OAS is located on the interface between the N- and C-

terminal domains and consists of residues from both domains (Hartmann et al., 2003; Kristiansen et al., 2011). The catalytic triad conserved in the nucleotidyl transferase superfamily is in human OAS1 formed by Asp75-Asp77 in the strand β2 and Asp148 in the strand β5. The three aspartate residues are highly conserved in all enzymatically active OASs. Single or double mutants of different OAS forms, where the aspartate residues were replaced by alanine, were enzymatically inactive (Sarkar et al., 1999b; Yamamoto et al., 2000; Donovan et al., 2013). These acidic residues in OAS coordinate two catalytic divalent Mg^{2+} ions similarly to the corresponding residues in other members of the NT family (Yamamoto et al., 2000; Hartmann et al., 2003; Donovan et al., 2013).

Between strands β1 and β2 is the helical turn P-loop, where the conserved Ser62 in porcine OAS interacts with the substrate ATP phosphate groups (Hartmann et al., 2003). The P-loop motif in porcine OAS is ^{60}GGSSGKG^{66} which is well conserved among mammalian OAS proteins. The mouse OAS1a P-loop double mutant Gly62Ala/Gly63Ala (corresponding to Gly60 and Gly61 in porcine OAS) was inactive. Another P-loop mutant of the same enzyme, Lys67Arg (Lys65 in porcine OAS), had reduced enzymatic activity of about 35% of the wt enzyme. However, when the lysine was mutated to methionine or alanine, the activity was further reduced to a substantial degree (Yamamoto et al., 2000; Torralba et al., 2008; Donovan et al., 2013).

The substrate binding sites of OAS were identified by cross-linking radiolabeled ppp-azidoA2'p*5'dA to human OAS2 (p69). As a result, two peptides ^{420}SYTSQK^{428} and ^{541}DLIRLVK^{547} were identified. Mutational analysis of potential substrate binding residues Tyr421, Arg544 and Lys547 confirmed that there are indeed two independent substrate binding sites in OAS and binding to each site is required for enzyme activity. Using dATP as a donor and A5'ppp5'A as an acceptor molecule and measuring the K_m values of different substrates in the case of mutant proteins, the Tyr421(Phe88 in OAS1) and the Arg544 (Arg210 in OAS1) were proposed to be the acceptor and the donor binding sites, respectively. The Lys547Ala mutant had also reduced activity (about 8% of wt), but the K_m for dATP was only slightly elevated if compared to wt enzyme (Sarkar et al., 2002). However, the position of the corresponding residues Arg210 and Lys213 in the OAS1 crystal structure suggests that the Lys213 is in better position to interact with the substrate than the Arg210 (Donovan et al., 2013). Both residues are located in the highly conserved sequence motif of mammalian OASs (LIRLVKHW) and are essential for enzymatic activity as determined by mutational analyses (Hartmann et al., 2003; Torralba et al., 2008). The Arg210 may be involved in stabilization of the active site by interacting with the flexible backbone of the P-loop (Torralba et al., 2008) or in dsRNA binding as it is located in the protein/RNA interface in the cocrystal structure (Donovan et al., 2013).

Two OAS-specific conserved motifs YALELLT and RPVILDPADP are referred to as Signatures 1 and 2, respectively. They contain essential amino acid residues for OAS activity. Residues Tyr230, Glu233 and Asp300

(numbered according to human OAS1) are implicated in substrate binding, based on the structural modelling of the active site (Torralba et al., 2008; Donovan et al., 2013). The Signature 2 is similar to the proposed nucleotide recognition motif of polβ type NTs (Martin and Keller, 2007; Kjaer et al., 2009). The conserved elements in OASs are shown on Fig. 5.

It has been proposed that human OAS1 and OAS2 need oligomerization for their enzymatic activities. According to gel filtration experiments, the native form of OAS2 might be a dimer and OAS1 likely exists in tetrameric form. OAS3 eluted as a monomer (Marié et al., 1990b). In the C-terminus of the p48 OAS1 a stretch of 3 amino acid residues, Cys-Phe-Lys (the CFK motif) was shown to be essential for enzymatic activity of the protein. When mutated, the enzyme was able to bind dsRNA and ATP, but was not able to oligomerize. The CFK motif was found to be conserved in other small isoforms of human OAS, as well as in p69 and was proposed to be the oligomerization motif (Ghosh et al., 1997). However, this sequence is not conserved in OAS proteins from other animals (Justesen et al., 2000; Torralba et al., 2008; Donovan et al., 2013). Currently there is no firm evidence that oligomerization is required for OASs (Hornung et al., 2014).

2.2. Induction and activation of OAS proteins

The basal levels of OAS expression are low and vary between different OAS isoforms and cell types (Witt et al., 1990; Witt et al., 1993). After treatment with interferons, the levels of OAS increase significantly. The rate and fold of induction are dependent on the OAS isoform, the cell type and the type of interferon (Baglioni and Maroney, 1980; Witt et al., 1990; Witt et al., 1993; Liu et al., 2012).

The OASs are synthesised as inactive proteins and they need to be activated by dsRNA. However, they do not possess a typical RNA-binding motif (Kodym et al., 2009). OAS proteins are able to bind different single-stranded and double-stranded RNAs but there is no correlation between binding affinity and the ability of the RNA to activate an OAS (Hartmann et al., 1998a; Kristiansen et al., 2011). The crystal structure analysis shows that the dsRNA binds to the protein surface at the junction of the N- and C-terminal lobes and induces a major conformational change in the N-terminal lobe of the protein. The rearrangements narrow the active site and bring together the catalytically important aspartate residues to enable substrate binding (Donovan et al., 2013).

2.2.1. Induction of the OAS gene expression

The expression of OAS genes is induced by both, Type I (IFN-α and IFN-β) and Type II (IFN-γ) interferons. However, Type I IFNs are more effective in inducing OAS than Type II IFN. The fold of induction is several times higher and the increase in OAS activity is more rapid after the treatment of cells with IFN-α or IFN-β than with IFN-γ (Rosenblum et al., 1988; Witt et al., 1990; Der et al., 1998; Rebouillat et al., 2000; Liu et al., 2012). The induction mechanism

by different types of interferons is obviously different, since Type I interferons induce the OAS mRNA synthesis in the presence of cycloheximide, when the protein synthesis is inhibited, whereas Type II interferon does not (Baglioni and Maroney, 1980).

Both the basal OAS levels and the scale of induction also seem to be cell type specific (Hovanessian et al., 1987; Chebath et al., 1987a; Witt et al., 1990; Marié et al., 1990b; Witt et al., 1993). For example, after treatment with IFN-β, monocytes had greater OAS activity than lymphocytes or macrophages (Witt et al., 1990). Furthermore, one isoform of OAS1, the p42, was not expressed in Daudi cells (Chebath et al., 1987a).

Considering the structure of the promoters of the *OAS* genes (see Section 2.1.1), their expression is likely regulated by several different factors. Glucocorticoids, various growth factors (e.g. epidermal growth factor, nerve growth factor), protein kinase C and interleukin (IL-6) have been shown to induce OAS alone or in combinations with interferons (reviewed in Player and Torrence, 1998). The OAS proteins are induced differentially. The levels of OAS1 were the most elevated after the treatment of cells with Type I IFNs if compared to the levels of OAS2 and OAS3 (Rosenblum et al., 1988; Der et al., 1998). This may be achieved due to the presence of distinct sets of regulatory elements in their promoters.

2.2.2. Activators of OAS proteins

In the absence of the dsRNA activator, the enzymatic activity of the OAS is often below the detection level. Binding of the suitable dsRNA activator enhances the enzyme activity more than 1000-fold (Justesen et al., 2000). The potential to activate OAS depends on various features of the dsRNA, including length, sequence composition, nucleoside modifications and deviations from the perfect duplex (Meng et al., 2012).

The most commonly used activator for studying various aspects of OAS enzymatic activity *in vitro* is the synthetic double-stranded RNA poly(I)·poly(C). The amount of poly(I)·poly(C) required for the activation of OAS proteins varies in the range of two orders of magnitude. However, the comparisons are somewhat unreliable, since the commercially available polynucleotide preparations used in the experiments may not be uniform in their molecular size distribution and double-strandedness (Justesen et al., 2000). The optimal concentration of dsRNA (μg/ml) is dependent on the amount of the enzyme in the activity assay and is estimated to be about half as much as the enzyme concentration (μg/ml) in the case of murine OAS3. If the concentration of the dsRNA is increased above the optimal value, the specific activity of the OAS decreases (Samanta et al., 1980).

Several attempts have been made to find the structural features of the RNA important for the activation of OAS proteins. No correlation exists between RNA size, binding affinity and OAS activation (Hartmann et al., 1998a; Molinaro et al., 2006). dsRNAs forming A-helices of different sequences exhibit striking differences in their capability to activate OAS. By testing a pool

of chemically synthesised short dsRNAs with specific sequences, a sequence motif required for the activation of OAS, NNWWNNNNNNNNNWGN, was proposed. Due to the 9 base-pairs between the two critical sequence patterns (WW and WG) they are separated by a full turn of the A-helix making the minor grooves of all four base-pairs accessible from one side of the dsRNA. With certain assumptions, the decisive features of this sequence are in accordance with the activator most commonly used *in vitro*, synthetic poly(I)·poly(C) (Kodym et al., 2009).

The presence of certain modified nucleosides in the RNA reduces its ability to activate OAS. If compared to unmodified RNA, the RNAs with identical sequences containing pseudouridine, N6-methyladenosine or 2-thiouridine were poor activators (Anderson et al., 2011).

No particular RNA species has been identified as a natural activator but different viral and cellular RNAs have been reported to activate OAS (Hartmann et al., 1998a). For example, the HTLV-1 Rex-RE, a viral RNA with highly branched secondary structure containing complex stem-loops, was able to stimulate OAS activity of the interferon-treated HeLa cell extracts and of the highly purified recombinant OAS1 in a dose-dependent manner (Mordechai et al., 1995). HIV-1 TAR RNA and Epstein-Barr virus RNA EBER-1 were equally potent activators of the purified recombinant OAS1 (Mordechai et al., 1995; Sharp et al., 1999). However, in the presence of the same concentration of poly(I)·poly(C) the specific activity of the OAS was about 3-5 times higher than that in the presence of the viral RNAs. The authors suggest that this may be due to the shorter double-stranded regions in these RNAs if compared to the poly(I)·poly(C) (Mordechai et al., 1995; Sharp et al., 1999).

The partially double-stranded adenoviral virus-associated RNA I (VA_I) could also activate OAS, though it was 10-fold less effective than the perfect dsRNA poly(I)·poly(C). The VA_I RNA exists in solution as a highly base-paired molecule with two long imperfectly base-paired stems of 23-25 bp joined at the center by a short stem-loop structure. The mutants of VA_I RNA, where the mismatches in the stem regions were repaired, were much better activators of OAS than the wt VA_I RNA. One of the mutants with 26 bp perfectly paired stem was almost as good activator as poly(I)·poly(C). The mutants, in which the structure of the central stem-loop was disrupted, had reduced capabilities to activate OAS (Desai et al., 1995). Furthermore, recent results indicate that while the full-length VA_I RNA activates OAS1, the biologically relevant Dicer-processed molecule, missing the terminal stem, does not. Since this truncated RNA has greater affinity for OAS1 than the full-length VA_I the virus could use it as an effective mechanism to supress the OAS1-mediated immune response (Meng et al., 2012).

Total cellular RNAs from different cells and tissues have been shown to activate OASs. RNA samples from pancreatic β cells and total RNA from heart tissue had the highest activation abilities whereas the cellular RNA isolated from immune cells and tissues (thymus and spleen) showed the lowest ability to activate OAS (Dan et al., 2012). Two specific mRNAs capable of binding to

and activating recombinant OAS1 in *in vitro* assays were cloned from prostate cancer cells. These mRNAs, encoding poly(rC)-binding protein 2 (PCBP2) and Raf kinase inhibitor protein (RKIP), were able to stimulate OAS activity to 40-60% of the activity obtained with poly(I)·poly(C) (Molinaro et al., 2006).

2.2.3. Mechanism of activation

The recently published crystal structure of human OAS1 in complex with dsRNA and substrate analog 3'-dATP revealed the activation mechanism of mammalian OAS proteins. Two adjacent minor grooves in the dsRNA are recognized by two dsRNA binding sites located ~30 Å apart in the OAS molecule. This interaction distorts the dsRNA from the ideal A-form helix by bending the ribophosphate backbone and also causes major conformational rearrangements in the N-terminal domain of the OAS protein (Donovan et al., 2013).

The helix N3 moves from the position that blocks dsRNA binding to the position that forms a part of the protein/dsRNA binding interface. The C-terminus of helix N5 moves outward from dsRNA and the β-sheet floor slides towards helix N5. Two basic side chains, Lys66 and Arg195, exchange their positions next to Glu233 resulting in the formation of a new helix N4 from a previously existing strand. This conserved element of structure that begins with strand β1, continues with the dsRNA-induced helix N4 and ends with strand β2 is referred to as RNA-induced catalytic structure (RICS) (Donovan et al., 2013).

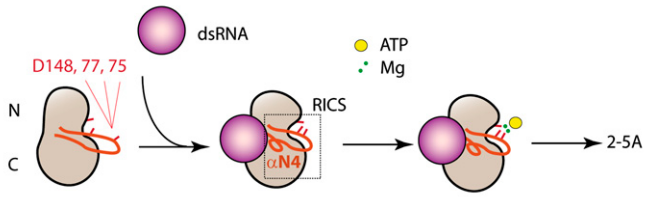


Figure 7. Model for OAS activation by dsRNA. Reprinted from Donovan et al., 2013, copyright (2013), the authors.

In the inactive conformation of the OAS the catalytic residues Asp75, Asp77 in the strand β2 and Asp148 in the strand β5 face divergent directions and are unable to coordinate Mg^{2+} ions or bind substrate (Fig 6A). The conformational changes induced by dsRNA narrow the active site and bring the catalytic aspartate residues together to enable substrate binding and catalysis (Fig. 6B, 7). The constitutively active PAP and CCAtr contain bulky hydrophobic residues in their protein cores that lock their active sites in the RICS-equivalent conformations (Donovan et al., 2013).

2.3. Functions of OAS

It is not clear why so many OAS isoforms exist and what the specific functions of each isoform are. These enzymes are induced differentially by interferons

(see Section 2.2.1). Also, their cellular distribution is divergent. In addition to cytoplasmic localization, the OAS1 enzymes have been detected in the nuclear fraction, the p46 isozyme has also been found in association with ribosomes. OAS2 was found to be distributed in and around nuclei, in mitochondria and cell membranes as well as in ribosome-containing areas of cytoplasm. OAS3 is believed to be mainly associated with rough endoplasmic reticulum. The subcellular location of the OAS isoforms varied slightly depending on the cell line studied (Hovanessian et al., 1987; Chebath et al., 1987a; Marié et al., 1990b; Besse et al., 1998). The myristylation of p69 OAS2 was proposed to account for its association with membranes (Marié et al., 1990b).

In addition to the dissimilar expression pattern, the OASs also exhibited different enzymatic characteristics. The OAS3 was already activated by 1 μg/ml of poly(I)·poly(C), while OAS2 and OAS1 required significantly higher amounts of dsRNA for their maximum activity (Marié et al., 1997; Bandyopadhyay et al., 1998; Eskildsen et al., 2003). OAS2 was able to synthesise the longest 2-5A oligomers, longer than 20-mers have been detected in the activity assays (Marié et al., 1997; Sarkar et al., 1999a). The OAS1 synthesised 2-5A up to 8-mers, the main products being dimer to tetramer (Eskildsen et al., 2003). The OAS3 preferentially catalysed the formation of 2-5A dimers. However, after longer incubation times, higher oligomers were also present in the reaction mixture (Marié et al., 1997; Rebouillat et al., 1999).

The different OAS isoforms are likely involved in a number of cellular processes, each having its particular role and mode of regulation. The best characterized function of mammalian OASs is the RNase L-dependent response to the infection of several viruses (reviewed in Silverman, 2007). Additional roles for the OAS-RNase L system and for different OAS isoforms have been proposed, such as regulation of cell growth and differentiation as well as induction of apoptosis (Rysiecki et al., 1989; Salzberg et al., 1990; Salzberg et al., 1997; Castelli et al., 1997; Ghosh et al., 2000; Ghosh et al., 2001; Domingo-Gil and Esteban, 2006).

The expression pattern of genes associated with various cellular processes was changed when the OAS-RNase L pathway was activated. The up-regulated genes were the genes involved in transcription, translation, cell growth arrest and apoptosis. Down-regulated genes were related to protein synthesis, cell proliferation and maintaining mitochondrial homeostasis (Domingo-Gil et al., 2010).

The C-terminal domain of OASs and other nucleotidyl transferases of the pol β family was found to be structurally homologous to the α-helical ATP-cone, a nucleotide-binding regulatory domain of ribonucleotide reductases and other enzymes and transcription regulators. It was suggested that the pol β C-terminal domain binds RNA and might have nuclease activity. This domain is also present in NRAPs, nucleolar RNA-binding proteins, that do not possess the NT domain. It was proposed that in OASs, devoid of 2-5A synthesising activity, the potentially active C-terminal domain may be the functional domain and the inactive NT domain may have a regulatory role (Rogozin et al., 2003).

The best characterized functions of OASs are described in the following sections.

2.3.1. Antiviral response

The OAS-RNase L antiviral pathway is depicted on Fig. 8. The expression of OAS and RNase L genes is induced by interferons. Both enzymes are synthesised as inactive proteins and they need to be activated by their cofactors. In the case of viral infection, dsRNA, that may originate from the viral genome or be a replication intermediate, binds to and activates OAS, which then catalyses the synthesis of 2-5A oligomers. The 2-5A oligomers bind to the pocket formed between two ankyrin repeats and the pseudo-protein kinase domain in RNase L (Han et al., 2014). RNase L then forms homodimers and becomes active (Dong and Silverman, 1995). The active RNase L cleaves single-stranded RNA 3' of the UpAp and UpUp dinucleotides (Wreschner et al., 1981). Depending on the levels of 2-5A in the cell, the RNase L may preferentially cleave viral RNA. Higher 2-5A levels have been shown to cause also the degradation of rRNA and cellular mRNA (Li et al., 1998).

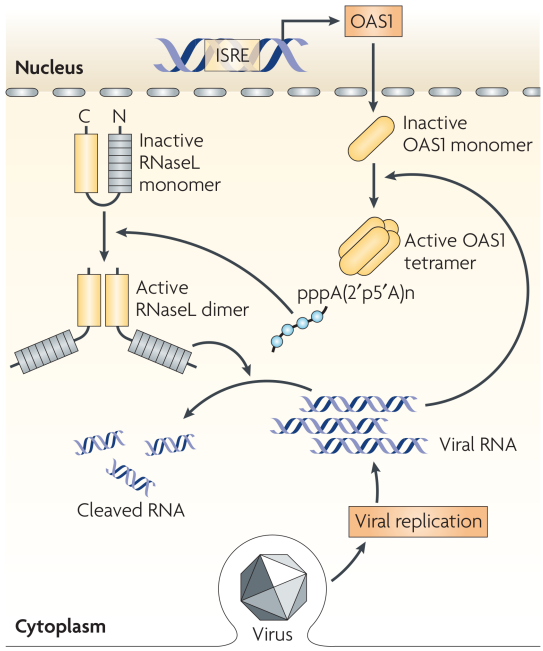


Figure 8. The OAS-RNase L antiviral pathway. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Immunology (Sadler and Williams, 2008), copyright (2008).

To restrict the activation of RNase L and limit the antiviral response, 2-5A oligomers are rapidly degraded by 2'-phosphodiesterase and 5'-phosphatase (Kubota et al., 2004). The activity of RNase L is also regulated by a specific protein RNase L inhibitor (RLI), which interacts with the latent RNase to

prevent it from binding 2-5A and subsequent dimerization (Bisbal et al., 1995). RNase L can suppress viral infections by cleaving different RNA substrates (reviewed in Silverman, 2007). From the perspective of the infected cell, the most preferred scenarios are the cleavage of viral genomic RNA to prevent virus from replicating or the cleavage of viral mRNA to inhibit viral protein synthesis. Degradation of cellular mRNAs and rRNAs that are required for viral replication is another effective strategy against viral infections. Besides, some small host cell RNA cleavage products may initiate a signal pathway that results in amplification of IFN production and achieving a broader antiviral response in the organism. The most severe response to viral infection is the elimination of infected cells by apoptosis (Silverman, 2007).

The antiviral effect of the OAS has been demonstrated against members of several RNA virus families. OAS1 or OAS2 expression in cells caused dose-dependent inhibiton of Encephalomyocarditis virus (EMCV, family *Picornaviridae*) replication (Chebath et al., 1987b; Coccia et al., 1990; Marié et al., 1999; Ghosh et al., 2000). Several OAS proteins were shown to mediate resistance to Dengue virus, a member of the *Flaviviridae* family (Lin et al., 2009). A SNP that leads to the production of OAS1 isoforms with lower enzymatic activity has been correlated with the higher probability of the West Nile virus (WNV, family *Flaviviridae*) infection if compared to patients expressing OAS1 p46 isoform with higher enzymatic activity (Lim et al., 2009). The OAS-RNase L pathway has been implicated in antiviral response to Respiratory syncytial virus (RSV), a member of another RNA virus family, *Paramyxoviridae* (Behera et al., 2002). However, most DNA viruses are resistant to the OAS-RNase L system (Silverman, 2007).

In addition to the traditional antiviral pathway, RNase L-independent responses to viruses have been described for some OAS isoforms (reviewed in Kristiansen et al., 2011). Both, the human OASL and murine Oas1b that lack enzymatic activity and are therefore unable to activate RNase L possess antiviral activity. It was shown that the Oas1b inhibits the accumulation of viral RNA in early stages of the WNV infection (Kajaste-Rudnitski et al., 2006). The antiviral activity of OASL against EMCV requires the ubiquitin-like domain of the protein (Marques et al., 2008). Furthermore, the enzymatically active porcine OAS1 was demonstrated to enter into cells when added exogenously and to possess strong antiviral activity which was independent of its enzymatic activity and RNase L activation (Kristiansen et al., 2010). However, the exact mechanism of these RNase L-independent antiviral actions is unknown.

In the light of these results it has been speculated that the OAS proteins may have two independently acting antiviral pathways, the classical RNase L-dependent and a yet uncharacterized RNase L-independent pathway (Kristiansen et al., 2011).

2.3.2. Apoptosis

After the activation of the OAS-RNase L pathway, several genes related to the induction of apoptosis were up-regulated and the expression of antiapoptotic

genes was decreased (Domingo-Gil et al., 2010). It was found that the rRNA degradation occurring simultaneously with apoptosis resembled the rRNA cleavage induced by RNase L. Furthermore, cells that lacked RNase L activity were more resistant to apoptosis than cells expressing active RNase L (Castelli et al., 1998). One possible mechanism for OAS-RNase L system to induce apoptosis is the RNA degradation which results in protein synthesis inhibition and leads to apoptosis (Díaz-Guerra et al., 1997). In the context of Vaccinia virus infection, the induction of apoptosis by OAS-RNase L system involves rRNA cleavage and subsequent caspase activation. Caspases 2, 8 and 9 are implicated in this process. After caspase activation, morphological and physiological changes occur in mitochondria, including the disruption of mitochondrial potential and cytochrome c release (Domingo-Gil and Esteban, 2006).

RNase L-independent proapoptotic activity has been attributed to one human OAS1 isoform, p48 (Ghosh et al., 2001). The apoptosis inducing function of this protein was discovered, as its expression by transfection of HT1080 cells caused cell death. Further experiments demonstrated that it was due to cellular apoptosis indicated by cytochrome c release from mitochondria, caspase activation and DNA fragmentation. This induction of apoptosis was shown to be isoform-specific and independent of the 2',5'-oligoadenylate synthesising activity, since the enzymatically inactive mutants of p48 were also active in causing apoptosis in the HT1080 cells. The apoptotic effect of p48 was also seen in the RNase L-/- fibroblasts, further demonstrating that the 2',5'-oligoadenylate synthesis and subsequent activation of RNase L was not required. Other OAS1 isoforms, p42 and p46, or OAS2 isoform p69 were not able to induce apoptosis when expressed in cells at similar levels. Finding of the putative Bcl-2 homology 3 (BH3) domain in the C-terminal tail of the p48 protein pointed to a possible mechanism of causing apoptosis by binding to the anti-apoptotic proteins Bcl-2 and Bcl-x_L (Ghosh et al., 2001).

2.4. Evolution of OAS

The remarkable structural similarity between the catalytic core of OAS and that of PAP and Class I CCAtr suggests that these enzymes share a common evolutionary origin (Kristiansen et al., 2011). It has been proposed that the ancestral OAS diverged from an ancient template-independent nucleotidyl transferase at the beginning of metazoan evolution (Torralba et al., 2008). OAS obtained a unique ability to catalyse the formation of oligonucleotides with 2',5'-phosphodiester linkages instead of ordinary 3',5'-linkages and to initiate synthesis from ATP without requiring any primer (Torralba et al., 2008; Kristiansen et al., 2011). In contrast to constitutively active nucleotidyl transferases, OAS evolved to require dsRNA binding for the assembly of its active site (Donovan et al., 2013).

The ancestral OAS was apparently similar to the modern OAS1 proteins. Early in the vertebrate evolution, the ancient *OAS* gene duplicated giving rise to the predecessors of modern *OAS1* and *OASL* genes. Ancestral *OASL* later fused

to the ubiquitin-like domain present in the C-terminus of modern OASL proteins (Kjaer et al., 2009). The mammalian *OAS* gene family has further expanded by gene duplications, domain coupling to form multidomain genes, block duplications to simultaneously produce copies of multiple genes and domain duplications within the genes (Kumar et al., 2000).

2.4.1. Occurrence of OAS

The human and mouse OAS gene families are well characterized (see Section 2.1.1). The OAS genes of the rat (*Rattus norvegicus*) are clustered on Chromosome 12. Their number, organization and exon/intron structure is very similar to the *Oas* locus in the mouse genome indicating common structures of OAS gene families in rodents (Perelygin et al., 2006). The equine and canine OAS gene families are similar to that of humans. In the genomes of pig and cattle the *OAS1* gene is present in three copies and the gene corresponding to human *OAS3* has not been found (Perelygin et al., 2005; Perelygin et al., 2006). It is believed that, with some exceptions, the four types of OAS genes are present in all mammals (Kjaer et al., 2009).

A single OAS gene (*OASL*) was found in chicken. It is composed of six exons. The first five exons encode the catalytic domain of OAS with the exon/intron structure similar to that of human *OAS1* gene and the sixth exon encodes the ubiquitin-like domain of two consecutive sequences (UbL1 and UbL2) homologous to ubiquitin. The chicken OAS gene has two alleles, A and B, encoding two types of OAS proteins. OAS-B lacks 32 amino acid residues from the highly hydrophobic region of the UbL1 sequence if compared to OAS-A. Both proteins are enzymatically active. However, the OAS-B exhibits only 10-15% of the enzymatic activity of the OAS-A protein (Yamamoto et al., 1998; Tatsumi et al., 2000).

OAS mRNAs have also been cloned from evolutionarily lowest animals, marine sponges *Geodia cydonium* (Wiens et al., 1999), *Geodia barretti* (Vallmann et al., 2011), *Suberites domuncula* (Grebenjuk et al., 2002) and *Haliclona panicea* (Müller and Müller, 2003) and a freshwater sponge *Lubomirskia baicalensis* (Schröder et al., 2008). The demonstration of the OAS activity in several other marine sponge extracts indicates that OASs are widely distributed in sponges (Reintamm et al., 2003b).

OAS activity has previously also been detected in reptiles but not in amphibians, fish or insects (Cayley et al., 1982). The OAS genes were not found in the sequenced genome of the frog *Xenopus tropicalis*. However, by analysing the NCBI EST database, OAS genes in two salamander species belonging to genus *Ambystoma* were identified (Kjaer et al., 2009). The sequencing of the elephant shark (*Callorhinchus milii*) genome revealed the presence of the OAS gene in this organism (Venkatesh et al., 2007). Putative OAS sequences were also found among EST data from two other cartilaginous fish species, the dogfish shark (*Squalus acanthias*) and the little skate (*Leucoraja erinacea*) (Kjaer et al., 2009). These results demonstrate the presence of OAS in amphibians and cartilaginous fish, while the three

sequenced teleost fish (two pufferfish species *Takifugu rubripes* and *Tetraodon nigroviridis* and the zebrafish *Danio rerio*) have evidently lost the respective genes (Venkatesh et al., 2007).

Based on a bioinformatic analysis, the presence of OAS genes was predicted in tunicates (sea squirt *Ciona intestinalis*) and in cephalochordates (amphioxus *Branchiostoma floridae*) (Torralba et al., 2008; Kjaer et al., 2009).

OAS genes were also identified in several evolutionary branches of multicellular animals belonging to the Protostomia clade. The presence of predicted OAS genes in the genomes of the leech *Helobdella robusta* and the polychaete worm of genus *Capitella* indicate the occurrence of OAS in annelids. The sequenced genome of the gastropod snail *Lottia gigantea* contains two OAS genes. The search in the EST database confirms the presence of OAS genes in other animals from the phylum Mollusca, as well – the mussel *Mytilus californianus* and the oyster *Crassostrea virginica*. However, no OAS genes have been found in the sequenced genomes of insects *Drosophila melanogaster*, *Aedes aegyptii*, *Anopheles gambiae*, *Apis mellifera* and the nematode *Caenorhabditis elegans* (Kjaer et al., 2009).

In addition to sponges, the OAS genes appear to be present in other divergent marine animals, as well. In the sea anemone *Nematostella vectensis* belonging to phylum Cnidaria, the OAS gene prediction based on the sequenced genome is supported by the EST evidence (Torralba et al., 2008; Kjaer et al., 2009).

According to the bioinformatic analysis of the sequenced genomes and the EST databases, the small OAS proteins – OAS1 and OASL – are widely distributed on the phylogenetic tree of animals. The medium (OAS2) and large (OAS3) forms, on the other hand, appear to be present only in mammals (Kjaer et al., 2009).

Surprisingly, the presence of OAS sequences has also been predicted in the genomes of unicellular organisms, the choanoflagellate *Monosiga brevicollis* and the green algae *Chlamydomonas reinhardtii.* It has been speculated that the choanoflagellate OAS could have been the first protein with 2',5'-oligoadenylate synthesising activity (Kjaer et al., 2009). Genes that show conservation in regions containing OAS-specific motifs have been identified in some bacteria. However, the authors propose that these may have been acquired by the horizontal gene transfer from eukaryotes (Kjaer et al., 2009).

2.4.2. Origin of the OAS gene family

Clustering of OAS genes in a close proximity in the human genome suggests their common ancestry (Rebouillat and Hovanessian, 1999). It has been proposed that the medium and large OAS forms evolved by duplication of the ancestral gene, possibly encoding the small OAS form (Rebouillat et al., 2000). A comparison of the N- and C-terminal OAS units present in the human OAS2 with the first 346 amino acid residues of OAS1 revealed 41% and 53% identity, respectively (Marié and Hovanessian, 1992). The N-terminal, middle and C-terminal OAS units of OAS3 share 42%, 47% and 60% identity with the OAS

unit of OAS1, respectively. Furthermore, in OAS3 the N-terminal and middle units are more similar to one another than to the C-terminal unit (Fig. 5; Rebouillat et al., 1999; Rebouillat et al., 2000). Interestingly, the C-terminal unit in OAS2 and OAS3 has the highest degree of identity to OAS1, especially in the conserved motifs that are considered to be important for OAS enzymatic activity (Hovanessian and Justesen, 2007).

By constructing a phylogenetic tree of all OAS units present in the four types of OAS proteins, two groups of researchers (Kumar et al., 2000; Perelygin et al., 2006) have proposed a similar scenario for the evolution of the OAS gene family (Fig. 9). They suggest that the first event was the duplication of the ancestral *OAS* gene producing the ancestors of the OASL-related and OAS1-related groups. Next the duplication of the cromosomal segment containing those two genes occurred. One OAS1-related gene then evolved into the ancestral *OAS1* gene, while the other evolved into the C-terminal domain of the two-domain ancestral gene, *OAS2*. Similarly, one of the OASL-related genes became the ancestor of modern *OASL* genes and the other evolved into the N-terminus of the ancestral *OAS2*. Subsequently, an internal domain duplication generated the modern three-domain *OAS3* gene (Kumar et al., 2000; Perelygin et al., 2006). This model explains why modern *OAS1* genes are more closely related to the C-terminal OAS units of the *OAS2* and *OAS3* genes, while the N-terminal units are more closely related to the *OASL* genes (Perelygin et al., 2006).

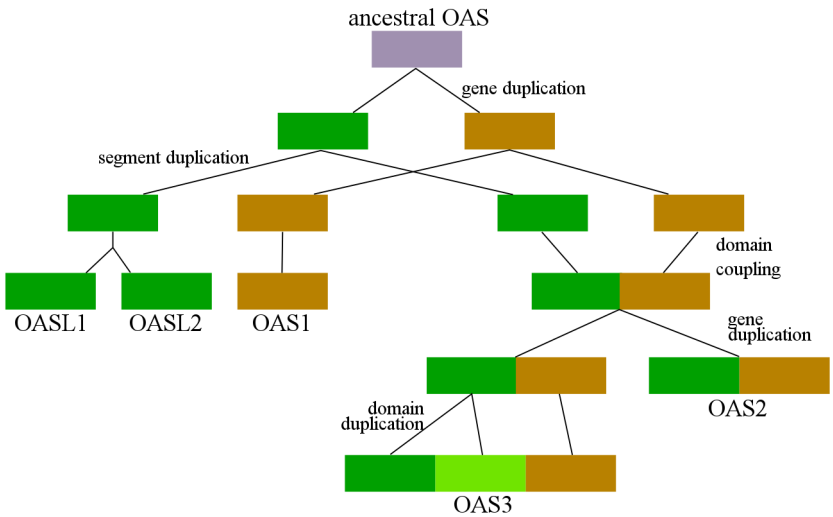


Figure 9. The model for the evolution of the OAS gene family. Based on Kumar et al., 2000; Perelygin et al., 2006.

The *OAS1* gene is found in most animals where the OAS genes are present. Furthermore, it is the only type of OAS found in the evolutionary branches of animals that diverged before the vertebrate lineage. This confirms the hypothesis that the ancestral gene may have been the ancient *OAS1*. The

presence of an *OASL* gene in the little skate (*Leucoraja erinacea*) and the elephant shark (*Callorhinchus milii*) but not in tunicates or cephalochordates (Kjaer et al., 2009) indicates that the aqcuisition of the ubiquitin-like domain occurred before the divergence of cartilaginous fishes. The phylogenetic analysis of *OASL* genes from different animals divides them into three groups which may be a distinction between active and inactive OASL proteins. The proteins grouped together with active OASL from chicken and OASL2 from mice could also be expected to have enzymatic activity, while the proteins grouped with OASL1 proteins are probably not active in the 2-5A synthesis (Kjaer et al., 2009). Based on the phylogenetic tree of known OASL proteins it seems that the duplication of the *OASL* gene to *OASL1* and *OASL2* and the loss of activity of OASL1 proteins occurred in the common ancestor of mammals or early in mammalian evolution. A pseudogene corresponding to mouse *Oasl2* has been found in the human genome (Eskildsen et al., 2003). However, there is too little information about the true number and enzymatic activity of OAS and OASL in animals phylogenetically older than mammals to draw definitive conclusions.

The *OAS2* and *OAS3* genes are found only in mammals. The *OAS2* gene evolved after the *OASL* gene was already present and the *OAS3* gene originated after the *OAS2* gene. The *OAS2* gene was found in the possum genome and it is assumed that the *OAS3* evolved after the divergence of marsupials (Kjaer et al., 2009). However, it is also possible that the *OAS3* gene was lost in possum as it has happened in other genomes during the mammalian evolution. For example, no *OAS3* was found in the cattle or pig genomes (Perelygin et al., 2006).

3. Sponges

Sponges (phylum Porifera) are commonly considered to be the simplest and most ancient, still extant multicellular animals. Their presence has been dated back to more than 700 million years (Erwin et al., 2011). Sponges are aquatic sessile filter-feeding animals with a simple body plan. They have more than 10 types of specialized cells, e.g. pinacocytes, which form the external cover; choanocytes, which are flagellated collar cells that create the water current through the canals; sclerocytes, which secrete spicules, the elements of the sponge skeleton; archeocytes, which are amoeboid cells, that move freely in the mesohyl and are capable of phagocytosis; and bacteriocytes, which contain procaryotic symbionts in their vacuoles. Most sponge cells can de- and redifferentiate. Sponges have no true organs and the presence of tissues is debated (rewieved in Ereskovsky, 2010).

Sponges are classified into 4 classes – Demospongiae, which is the largest class consisting of sponges with siliceous spicules and/or with a skeleton of organic fibers; Hexactinellida or glass sponges, which have syncytial tissues and siliceous spicules with a triaxonic symmetry; Calcarea or calcareous sponges, whose mineral skeleton is composed of calcium carbonate; and Homoscleromorpha which, unlike other sponges, contain a basement membrane (Van Soest et al., 2012).

At present there are over 8500 sponge species described, but it is estimated that the actual number of species is twice as large. Sponges are currently divided into 25 orders, 128 families and 680 genera. The majority of the sponge species (83%) belong to the class Demospongiae. Hexactinellida and Calcarea both include 8% of the sponge species (623 and 681 species, respectively). About 1% (87) of the sponge species is classified into the class Homoscleromorpha (Van Soest et al., 2012). Most of the sponge species live in the marine environment but about 250 species have adapted to life in fresh water (Van Soest et al., 2012).

3.1. Importance to science

The scientific interest towards sponges has rapidly increased during the past decades. Porifera as the earliest branching metazoan phylum is of significant interest in the studies of metazoan genome evolution. Sponges may be the clue to understanding the evolutionary origin of multicellularity. Resolving the question of sponge monophyly or paraphyly is essential for making assumptions about the features of the last common ancestor of the Metazoa (Philippe et al., 2009). Besides, several genes regarded as common to higher animals have been found in the sponge genomes. However, the corresponding processes (e.g. certain developmental and neuronal signaling pathways) are not present or not recognized in sponges. These data raise the question about the complexity of the metazoan ancestor (Riesgo et al., 2014).

The remarkable plasticity of their cellular organization and differentiation is the reason why sponges are of great interest in stem cell studies (Ereskovsky, 2010; Funayama, 2010). The production of secondary metabolites with bioactive properties makes sponges invaluable tools for pharmaceutical research in hope of developing new and more effective drugs (reviewed in Laport et al., 2009). Moreover, sponges are an important component of the benthic fauna; they influence the marine ecosystem by having an impact on the growth substrate and nutrient cycling, as well as by associating with other marine organisms (reviewed in Bell, 2008).

3.1.1. Phylogenetic studies

The multicellular animals are believed to be of monophyletic origin (Müller, 1995) and it is accepted that the choanoflagellates represent the closest living unicellular relatives of multicellular animals. However, the evolutionary distance between choanoflagellates and metazoans is substantial, and evidently few, if any, intermediate lineages have survived (King et al., 2008). The question of the earliest branching metazoan lineage has been debated for a long time. Furthermore, the relationships between the four non-bilaterian animal phyla – Placozoa, Porifera, Cnidaria and Ctenophora – as well as class-level relationships within these phyla are still obscure (Osigus et al., 2013).

Historically, based on their morphological features, sponges are considered to have been the first to diverge from the phylogenetic tree of the multicellular

animals (Nielsen, 2008). This view is supported by recent molecular phylogenies, based either on nucleus-encoded protein or genomic sequences (Srivastava et al., 2008; Philippe et al., 2009; Sperling et al., 2009; Srivastava et al., 2010; Pick et al., 2010; Erwin et al., 2011). However, the size, gene content and phylogenetic analysis of the mitochondrial genomes as well as the analysis of a complex set of data including morphological characters, ribosomal gene sequences and nuclear coding genes place Placozoa at the basis of the Metazoa (Schierwater et al., 2009; Osigus et al., 2013). On the other hand, the sequence of the ctenophore *Mnemiopsis leidyi* genome supports the hypothesis that Ctenophora is the sister group to all other animals (Ryan et al., 2013).

Whereas the monophyly of Placozoa, Cnidaria, Ctenophora and Bilateria has not been questioned by molecular phylogenetic analyses, this is not the case for Porifera (Dohrmann and Wörheide, 2013). Morphological character analysis supports the hypothesis that sponges form a monophyletic group (Peterson and Eernisse, 2001). It means that the sponge body plan (notably featuring an aquiferous system with internalized choanocyte chambers and the pinacoderm) may have evolved in the stem line of the Porifera after divergence from the common ancestor. Rather than reflecting the ancestral animal form, sponges are better considered as highly specialized organisms, possibly having acquired a sedentary life style from a hypothetical pelagic ancestor (Philippe et al., 2009). On the other hand, several phylogenetic analyses based on rDNA, nuclear housekeeping genes or mitochondrial genomes have proposed sponge paraphyly (Peterson and Eernisse, 2001; Sperling et al., 2009; Erwin et al., 2011; Osigus et al., 2013; Nosenko et al., 2013). The mitochondrial genome analysis revealed a sister group relationship between demosponges and the Homoscleromorpha and grouped the Hexactinellida as a sister group to Bilateria (Osigus et al., 2013). The phylogenetic analysis based on nuclear housekeeping genes found three independent sponge lineages, the Demospongiae + Hexactinellida, the Calcarea and the Homoscleromorpha. The latter was found to be more closely related to eumetazoans than to demosponges (Sperling et al., 2009). The presence of the basement membrane in homoscleromorphs (Leys et al., 2009; Leys and Riesgo, 2012) supports this hypothesis. In the case of the sponge paraphyly, the features shared by modern sponges would have already been present in the common ancestor of the Metazoa with a sponge-like body plan, and later been lost in some branches during evolution (Philippe et al., 2009). However, the recent molecular phylogenetic studies have also supported the sponge monophyly hypothesis (Schierwater et al., 2009; Philippe et al., 2009; Pick et al., 2010; Nosenko et al., 2013). It was suggested that the shortness of the branch leading to Porifera reflected the closely spaced splitting events during the emergence of sponge lineages and this was the reason for the difficulty in recovering sponge monophyly in previous studies (Philippe et al., 2009). Increased taxon sampling has been shown to further support the hypothesis of monophyly of the Porifera (Nosenko et al., 2013).

It is clear that better methods and increased taxon sampling are needed for

future studies. Moreover, in addition to molecular sequence data analysis, integrative approaches including analyses of morphology, cytology, development and genome architecture could be helpful in a final clarification of the early evolution of multicellular animals (Dohrmann and Wörheide, 2013). The current consensus view of the relationships between the metazoan lineages is presented on Fig. 10.

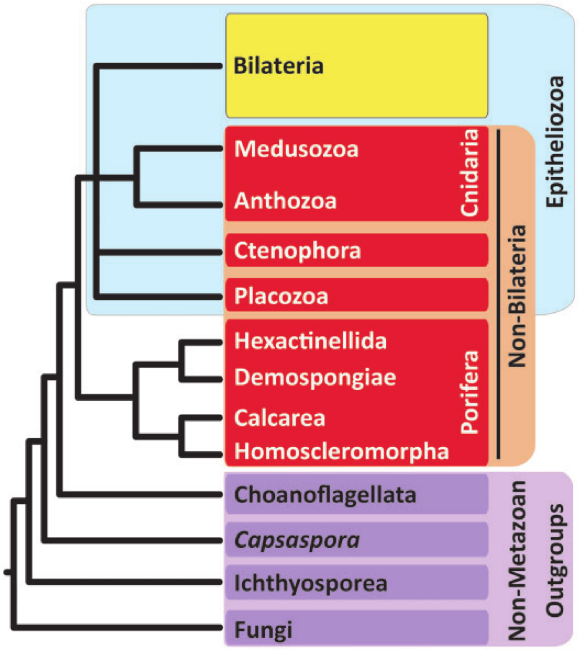


Figure 10. Current consensus view of phylogenetic relationships of the major metazoan lineages (Dohrmann and Wörheide, 2013). Reprinted by permission of Oxford University Press.

In addition to the uncertain divergence of sponges and other multicellular animals, the phylogenetic relationships inside the sponge classes are also unsettled.

Only recently, the fourth sponge class – Homoscleromorpha – was recognized as a separate class and not belonging to the class Demospongiae (Gazave et al., 2010; Gazave et al., 2012). Besides, the division of Homoscleromorpha into two families (Plakinidae consisting of sponges with spicules and Oscarellidae consisting of aspiculate species) was restored (Gazave et al., 2010).

The relationships between demosponge lineages are uncertain, as well. For example, the position of freshwater sponges in the demosponge order Haplosclerida is questioned (Van Soest et al., 2012; Hill et al., 2013).

The data about sponge genomes is scarce. The genome size estimates for the 75 species of sponges, mostly marine demosponges, averaged 0.2 ± 0.01 pg (about 200 Mb) with a remarkable variation – from 0.04 pg in *Tethya actinia* to

0.63 pg in *Mycale laevis* (Jeffery et al., 2013). The genome of only a single sponge, *Amphimedon queenslandica*, has been completely sequenced to date. The assembled genome was ~167 Mb and it encoded 30000 predicted protein-coding loci which was assumed to be an overestimation. The analysis of the genome revealed a considerable conservation of gene structure and genome organization if compared to other animals. 63% of the predicted proteins had identifiable homologues in other organisms (Srivastava et al., 2010). The genomes of few other sponges (e.g. the homoscleromorph *Oscarella carmela* and the hexactinellid *Acanthascus dawsoni*) are presently being sequenced and the results are expected to be published soon (www.genomesonline.org, accessed 18.03.2014).

Recent analyses of transcriptomes of several sponges from all four classes show that several metazoan genes, involved in complex processes common to higher animals, are present in all sponge classes. Since the sponges do not have conventional structures, behaviour or mechanisms of development, it was hypothesised that the genes acquired their known functions later in evolution and may have different functions in sponges. It is also possible that sponges have lost the corresponding structures. Alternatively, the genes may perform similar functions in sponges and in higher animals, but those are difficult to recognize in sponges (Riesgo et al., 2014).

3.1.2. Bioactive substances and novel enzymatic activities

Marine sponges often produce organic compounds which are not directly involved in their normal growth, development or reproduction processes. These substances are termed secondary metabolites. Because of the sessile nature of the sponges, the production of these bioactive compounds has been interpreted as a mechanism by which sponges counter various environmental stress factors such as predation, overgrowth by fouling organisms, or competition for limited space (Grasela et al., 2012).

Sponges are one of the richest sources of bioactive molecules. The pharmaceutical interest in sponges started in the 1950s with the discovery of the nucleosides spongothymidine and spongouridine, which were the basis for the synthesis of anticancer drug ara-C and antiviral drug ara-A (Laport et al., 2009). Approximately 300 new natural compounds are discovered from sponges every year (Blunt et al., 2012; Blunt et al., 2013; Blunt et al., 2014). The chemical diversity of sponge substances is remarkable – in addition to the unusual nucleosides, other classes of compounds such as terpenes, sterols, cyclic peptides, alkaloids, fatty acids, peroxides and amino acid derivatives have been described (Laport et al., 2009; Abbas et al., 2011; Blunt et al., 2014). Those bioactive substances have shown anticancer, antibacterial, antifungal, antiviral, antiprotozoal and anti-inflammatory properties (reviewed in Laport et al., 2009).

In addition to natural substances, previously undescribed enzymatic activities have been identified in sponges. One example is the ATP N-glycosidase activity characterized in the marine sponge *Axinella polypoides*,

that converts adenosine-5'-triphosphate into adenine and ribose-5-triphosphate (Reintamm et al., 2003a). Another novel enzymatic activity, the recently discovered endo-2',5'-ribonuclease from *Tethya aurantium* specifically catalyses the degradation of 2',5'-phosphodiester linkages in a metal-ion-independent manner (Lopp et al., 2012). The enzymes responsible for these novel activities or the synthesis of the bioactive secondary metabolites mostly remain unidentified.

Sponges form close associations with a wide variety of microorganisms, which are generally referred to as symbionts, irrespective of the actual beneficial relationship to either partner. Microbes may comprise up to 40% of the sponge tissue volume. Sponge-associated microbial communities are highly diverse; they comprise bacteria, archaea and eukaryotes (dinoflagellates, diatoms, microalgae, fungi) (reviewed in Taylor et al., 2007; Webster and Taylor, 2012). Many of these microorganisms are specific to the sponge host (Webster et al., 2010). Interestingly, in several cases the isolated bioactive compounds are suspected to be produced by the sponge-associated microorganisms rather than sponges (Taylor et al., 2007). Often the exact microorganisms responsible for the synthesis of these substances are unknown. However, sometimes the candidate genes for the respective enzymes can be identified by sequencing the total metagenomes of the sponges (Hentschel et al., 2012). For example, this method has been used to identify the bacterial genes important for the production of the antitumor polyketides onnamide A isolated from sponge *Theonella swinhoei* (Piel et al., 2004) and psymberin from *Psammocinia* aff. *bulbosa* (Fisch et al., 2009).

Despite the high potential of sponge-derived bioactive substances in pharmacological and biomaterial applications, only a few marine natural compounds have been successfully developed into products. The major obstacles are the insufficient supply of biological material and the excessive structural complexity of the bioactive substances in order to synthesise them chemically. Therefore, many research efforts are directed to the establishment of the sponge cell or tissue cultures and to the development of methods for culturing sponge symbionts, many of which are at present unculturable (reviewed in Schippers et al., 2012).

3.2. OAS in sponges

A surprisingly high concentration of biologically active 2',5'-oligoadenylates was found in the marine sponge *Geodia cydonium*. The 2-5A synthesising activity from the sponge extract was close to that of the interferon-stimulated mouse fibroblast L929 cells, which was considered to be one of the richest sources of OAS activity (Kuusksalu et al., 1995). The product profile of the sponge OAS activity resembled that of the mammalian OASs as it was shown to produce oligomers up to octamer (Kuusksalu et al., 1998). In contrast to mammalian OASs which are active only in the presence of dsRNA, the crude extract of *G. cydonium* exhibited 2-5A synthesising activity without the addition of dsRNA. Furthermore, the nuclease treatments, performed to free the

sponge extract from any nucleic acids that could act as potential activators, had no effect on the oligoadenylate synthesis catalysed by the OAS in *G. cydonium* extract (Lopp et al., 2002).

In addition to *G. cydonium*, the presence of the OAS activity without the addition of any dsRNA has been demonstrated in the crude extracts of various other species of Demospongiae (Kelve et al., 2003; Reintamm et al., 2003b; Saby et al., 2009a; Lopp et al., 2010). However, the 2-5A synthesising capacity varied greatly between different sponge species, in the range of four orders of magnitude. In addition, the product profile of OASs from different sponge species was quite different. Most of the studied sponges were able to catalyse only the synthesis of short 2-5A oligomers (up to tetramers), whereas the extract of *Chondrosia reniformis* was able to synthesise 2-5A oligomers up to 17-mer, the main products being hexamer and heptamer. The variability in 2-5A synthesising capacity and in the product profiles may refer to the occurrence of a variety of OAS isozymes, which may be associated with the evolution and diversification of *OAS* genes in sponges (Reintamm et al., 2003b).

cDNAs encoding OAS proteins have been cloned from several marine sponges – *G. cydonium*, *Geodia barretti*, *Suberites domuncula* and *Haliclona panicea* – and a freshwater sponge *Lubomirskia baicalensis* (Wiens et al., 1999; Grebenjuk et al., 2002; Müller and Müller, 2003; Schröder et al., 2008; Vallmann et al., 2011). The sponge OASs share little sequence similarity with each other and even less with vertebrate OASs (Fig. 11). For example, the OAS from *G. cydonium* shares 18% and 39% of identical and similar amino acid residues, respectively, with the enzymatically active OAS1a from mouse and 17% and 37%, respectively, if compared to chicken OAS-B (Wiens et al., 1999). In comparison, this sponge OAS is 28% identical and 48% similar to the amino acid sequence of OAS from another marine sponge *S. domuncula* (Grebenjuk et al., 2002). OASs from two sponge species, both belonging to the same genus *Geodia*, *G. cydonium* and *G. barretti*, are considerably more similar, they share a 75% identity and a 85% similarity between their amino acid sequences (Vallmann et al., 2011). Despite the low overall conservation in amino acid sequences between OASs from different animals, the motifs characteristic to vertebrate OASs (P-loop, three catalytic Asp residues, signatures 1 and 2) are also conserved in sponge proteins (Fig. 11; Wiens et al., 1999; Grebenjuk et al., 2002; Schröder et al., 2008).

The functions of OAS proteins in sponges remain unclear at present. The modulation of OAS activity in marine sponges has been correlated with the presence of environmental stressors. Compared with the control, a significant increase in 2-5A oligomer content and 2-5A synthesising activity was found in the tissue of marine sponge *G. cydonium* following heat shock, cold shock and exposure to an environment with decreased pH or increased ionic strength. However, the exposure to a hypotonic or alkaline pH environment did not change the OAS activity if compared to the control sample (Schröder et al., 1997).

The OAS in marine sponges has been implicated in immune response,

including the defence against microorganisms (Grebenjuk et al., 2002; Müller and Müller, 2003; Schröder et al., 2008). The samples from sponges, kept in the aquarium for 6 months, had decreased OAS activities if compared to the sponges collected from the sea, where the bacterial content is significantly higher and more diverse. Furthermore, after the exposure of the sponge tissue or primmorphs to lipopolysaccharide, a bacterial endotoxin, or after infecting with bacteria, an increase in OAS activity and also in OAS expression was detected (Grebenjuk et al., 2002; Müller and Müller, 2003). The exact pathways of this response, involving OASs, are not identified in sponges. Several components of the mammalian 2-5A system, including RNase L have not been found, either. However, the activity of one of the key enzymes of this system, 2'-phosphodiesterase, has been reported in marine sponges (Saby et al., 2009b). In sponges, the regulation of 2',5'-oligoadenylate degradation may also be performed by another enzyme, which has not been described in other organisms. A novel endoribonuclease activity that specifically catalysed the degradation of 2',5'-linkages, was found to be present in several marine sponges (Lopp et al., 2012).

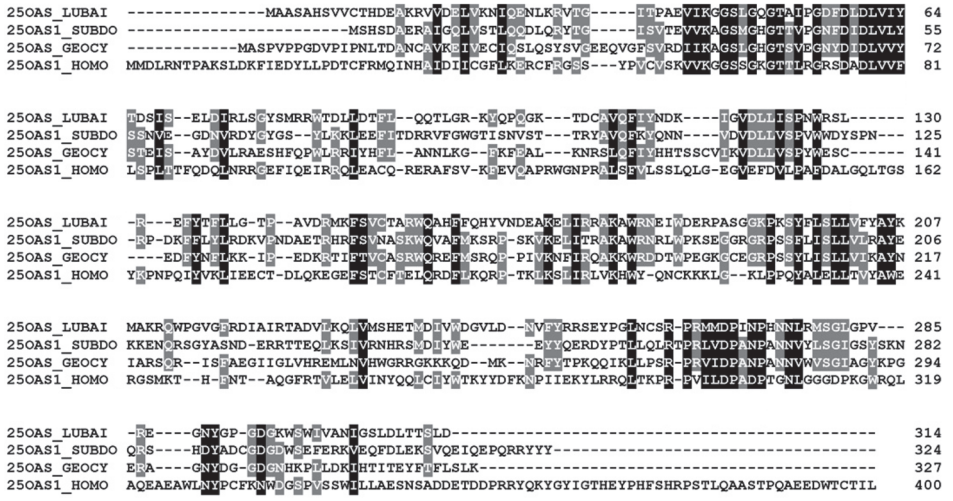


Figure 11. Alignment of amino acid sequences of different sponge and human OAS proteins. Residues conserved (similar with respect to their physico-chemical properties) in all sequences are shown in white on black and those conserved in at least three sequences are shown in white on gray. Reprinted from Schröder et al., 2008, Copyright (2008), with permission from Elsevier.

It has been demonstrated that the OASs from marine sponges have broader substrate specifity than that known for vertebrate enzymes (Lopp et al., 2010). With the crude extracts of the marine sponges *Thenea muricata* and *Chondrilla nucula* a variety of 2',5'-linked heterooligomers was obtained *in vitro* wherein all four ribonucleotides could serve as either acceptor or donor substrates for oligomerization. The remarkable difference if compared to vertebrate OASs is the ability of sponge OASs to use pyrimidine nucleotides as acceptor

molecules. However, the substrate specifity of an OAS differed among the sponge species. ATP was the preferred acceptor for OAS from *T. muricata*, while in the case of *C. nucula*, GTP was comparable with ATP as an acceptor molecule. In addition to the ability to synthesise the 2',5'-linked hetereooligomers *in vitro*, they were also found to occur naturally in the sponge extracts which has never been demonstrated in mammalian cells (Lopp et al., 2010).

Crude extracts of some sponge species from genus *Tedania* are capable of producing also 3',5'-linked oligoadenylates from ATP. In the activity assays the products with 3',5'-phosphodiester linkages, as well as with mixed 3',5'- and 2',5'-linkages were detected in addition to 2',5'-linked oligoadenylates (Lopp et al., 2004; Lopp et al., 2012).

More investigation is needed to establish the role for OASs in sponges as well as in other animals, where all of the components of the interferon-induced antiviral 2-5A system, described in mammals, are not present.

AIM OF THE STUDY

The aim of this study was to characterize 2',5'-oligoadenylate synthetases from marine sponges and other phylogenetically distant animals in order to elucidate the origin and evolution of OASs. To this end, the following objectives were set:

- To optimize the expression and purification procedures in order to produce recombinant OASs for further experiments.
- To verify the enzymatic activity of several bioinformatically predicted OASs from different multicellular animals and a closely related unicellular organism in order to find out the distribution of active OASs on the phylogenetic tree.
- To describe the enzymatic characteristics of recombinantly expressed OASs from different animals in order to compare them with the properties of the well-characterized mammalian OASs.
- To determine the gene structure of sponge OASs in order to ascertain whether it corresponds to that of mammalian OASs.

MATERIALS AND METHODS

The following methods were used during this study:

- RNA and DNA extraction (Publications II and III)
- RT-PCR and PCR (Publications II and III)
- cDNA cloning (Publication III)
- recombinant protein expression (Publications I and III)
- SDS-PAGE and Western blot analysis (Publications I and III)
- native PAGE (Publication I)
- IMAC (Publications I and III)
- GST affinity chromatography (Publication III)
- SEC (Publications I and III)
- OAS activity assay (Publications I and III)
- phosphatase and nuclease treatments (Publications I and III)
- RP-HPLC (Publications I and III)
- MALDI-TOF MS analysis (Publications I and III)
- mutational analysis

Detailed descriptions of materials and methods are provided in the publications of this thesis. Some unpublished data are also presented. The materials and procedures used are described below.

Mutations were introduced into the OASs from the marine sponge *T. ignis* (TigOAS) and the pig *S. scrofa* (SscOAS) as follows:
The I51A mutation was introduced into TigOAS using two complementary mutagenic primers
TigOAS_I51A_F (5'-CAGGTGACTACGACGCTGACATTGTCATCTACT-3')
TigOAS_I51A_R (5'-AGTAGATGACAATGTCAGCGTCGTAGTCACCTG-3'), where the site of the mutation is underlined. First, two separate PCRs were carried out using pET vector specific T7 promoter primer in combination with the mutagenic TigOAS_I51A_R primer and the TigOAS_I51A_F primer in combination with the vector specific T7 terminator primer. The PCR was carried out in a 50 µl final volume containing 1 ng TigOAS in pET24d plasmid DNA as a template, 20 pmol each primer, 0.2 mM each dNTP, 1x Pfu buffer and 1.25 U Pfu DNA Polymerase (Fermentas) with an initial denaturation at 95 °C for 3 min, followed by 35 cycles at 95 °C for 30 s, 52 °C for 30 s, 72 °C for 2 min, and a final extension step at 72 °C for 10 min. The PCR products were excised from gel and purified using GeneJet™ Gel Extraction Kit (Fermentas). Second PCR was performed in a 50 µl final volume containing 1 µl each both purified DNA fragments, 20 pmol each T7 promoter primer and T7 terminator primer, 0.2 mM each dNTP, 1x Pfu buffer and 1.25 U Pfu DNA Polymerase (Fermentas). The thermal cycling was as follows: initial denaturation at 95 °C

for 2 min, followed by 10 cycles at 95 °C for 30 s, 62 °C for 30 s, 72 °C for 2 min where the annealing temperature was decreased by 1 °C in each cycle, followed by 25 cycles at 95 °C for 30 s, 52 °C for 30 s, 72 °C for 2 min, and a final extension step at 72 °C for 10 min. The PCR product was precipitated with isopropanol and its ends were cut with restriction enzymes NcoI and HindIII. Then it was again excised from gel and purified using GeneJet™ Gel Extraction Kit (Fermentas). The purified DNA fragment was ligated into pET24d vector that had been restricted with the same enzymes and dephosphorylated.
The same method was used to introduce the A75I mutation into SscOAS using mutagenic primers
SscOAS_A75I_F (5'-AGGGGCCGATCAGATATTGACCTCGTCGTCT-3')
SscOAS_A75I_R (5'-AGACGACGAGGTCAATATCTGATCGGCCCCT-3').
After precipitation, the ends of the PCR product were cut with NcoI and BamHI. The DNA fragment was excised from gel and ligated into pET vector that had been restricted with the same enzymes.
Both constructs were verified by sequencing.

RESULTS AND DISCUSSION

The following results were obtained in this study:

Publication I:

- The recombinant OAS from the marine sponge *Geodia cydonium* (GcyOAS) could be produced as a soluble protein in the bacterial expression system.
- The recombinant GcyOAS bound to bacterial RNA with high affinity and purified as a set of heterogenous RNA-protein complexes.
- The recombinant GcyOAS-RNA complex was able to catalyse the formation of 2',5'-linked oligoadenylates.
- Compared to the 2-5A synthesising activites of the sponge crude extracts, the activity of the sponge recombinant OAS preparation was low.
- The addition of the commonly used activator of mammalian OASs, poly(I)·poly(C), caused only small modulations of the existing activity of the recombinant GcyOAS protein preparation.
- In addition to 2',5'-linked oligoadenylates, oligomers with 3',5'- and mixed linkages were identified among the reaction products.

Publication II:

- The exon/intron structure of sponge *OAS* genes was completely different from that of mammalian *OAS* genes.
- The *OAS* gene in *G. cydonium* had eight translated exons instead of the five in the mammalian OAS unit.
- The intron positions in sponge *OAS* genes differed from those in mammalian *OAS* genes.
- Introns in sponge *OAS* genes were considerably shorter than those in mammalian *OAS* genes.
- In the genome of *G. cydonium*, *OAS* was presented by at least two gene versions that are tandemly organized.
- Both *OAS* gene versions were transcribed in *G. cydonium*.
- In the genome of the marine sponge *Amphimedon queenslandica*, three *OAS* gene types, whose coding sequences differ at the same level as those between different sponge species, were found bioinformatically.
- In contrast to mammals, where *OAS* genes are clustered in a single locus, the evolutionary events shaping the *A. queenslandica OAS* gene family have resulted in the localization of *OAS* genes in different loci.

Publication III:

- The OAS from another marine sponge, *Tedania ignis*, was cloned.

- Based on the amino acid sequences, the OAS from *T. ignis* (TigOAS) shared a 29% identity and a 43% similarity with the OAS from *G. cydonium*.
- The exon/intron structure of the *OAS* gene from *T. ignis* was similar to that of the *OAS* genes of other marine sponges.
- The recombinant TigOAS was enzymatically active, catalysing the synthesis of 2',5'-oligoadenylates.
- Similarly to the GcyOAS, also TigOAS was able to catalyse the formation of oligonucleotides with 3',5'-linkages as well as with mixed 2',5'- and 3',5'-linkages.
- The ratio of 2-5A oligomers to 3-5A oligomers synthesised by TigOAS was even more in favor of the formation of 3-5A than in the case of GcyOAS.
- In contrast to GcyOAS which formed complexes with bacterial RNA, TigOAS purified as a pure protein and required activation by dsRNA.
- The putative OASs from the unicellular choanoflagellate *Monosiga brevocollis* (MbrOAS) and from the salamander *Ambystoma mexicanum* (AmeOAS) did not show 2',5'-oligoadenylate synthesising activity in our assays.
- The recombinant OASs from the cartilaginous fish *Leucoraja erinacea* (LerOAS) and the mollusk *Mytilus californianus* (McaOAS) were able to catalyse the formation of 2',5'-oligoadenylates in the presence of poly(I)·poly(C).

Unpublished results:

- At least two types of OAS genes with different genomic structures exist in multicellular animals. Both types have spread across the phylogenetic tree of animal evolution.
- The ratios of 2-5A oligomers to 3-5A oligomers synthesised by the I51A mutant of TigOAS and the A75I mutant of porcine OAS (SscOAS) were identical to those synthesised by the corresponding wt enzymes.

1. The structure of OASs from marine sponges

1.1. OAS proteins from marine sponges are diverse (Publications II and III)

The OAS activity in marine sponges was first detected in *G. cydonium* and later the cDNA containing the signature motifs of known OAS proteins was cloned from this sponge (Kuusksalu et al., 1995; Wiens et al., 1999). By now, the wide occurrence of OASs in sponges has been demonstrated both by showing the presence of 2-5A synthesising activity in their crude extracts and by cloning the corresponding cDNAs. Additionally, in some sponge species the presence of OASs is predicted on the basis of the available sequencing data (Grebenjuk et al., 2002; Reintamm et al., 2003b; Schröder et al., 2008; Vallmann et al., 2011; Publications II and III).

The comparison of the amino acid sequences of sponge OASs to those of mammalian OAS proteins reveals their low sequence similarity (Wiens et al., 1999; Grebenjuk et al., 2002). For example, the OAS from marine sponge *T. ignis* shares 17% identity and 26% similarity with the human OAS1 p42 isoform and 19% identity and 29% similarity with the mouse Oas1a. The OAS proteins from different sponge species share almost as little sequence similarity among one another as with OASs from mammals. The inter-species comparison of sponge OAS protein sequences shows that they are only about 20-40% identical to one another and their similarity is between 30-55% (Publication II, Table 4).

This similarity is evidently due to the few conserved regions and amino acid residues essential for the enzymatic activity and dsRNA binding. The protein regions between those conserved signatures are higly variable (Publication II, Fig. 7). The variability in these regions may be restricted only by the ability of the respective amino acid chains to form the necessary secondary structure elements in order to contribute to the correct folding of the entire protein molecule. The conserved regions are implicated in substrate binding and catalysis and thus the possibilities for the respective amino acids to be replaced by those with different side chains are limited.

At least two OAS1 gene versions are present in the genome of *G. cydonium*. The encoded proteins are very similar, their amino acid sequence identity is about 90% (Publication II, Fig. 1, Table 2). The same is true for the proteins encoded by the three known OAS gene versions in *S. domuncula* (Grebenjuk et al., 2002; Müller and Müller, 2003). However, the analysis of the *A. queenslandica* genome revealed the presence of three main types of OAS1 with several variants (Publication II, Table 3). The coding sequences of the three types, AmpOAS1A, AmpOAS1B and AmpOAS1C, differed from one another at the same level as those of OAS proteins from different sponge species. It is not yet clear whether the multitude of OAS genes and the presence of different types of OASs is a unique feature of this particular sponge or whether it is

common among sponges. It is possible that the similar number and variability of OASs is to be found in other sponge species, as well.

As in mammals, the OAS genes in sponges have duplicated and diverged. The low similarity between sponge OAS proteins is probably due to their independent evolution after the divergence from their common ancestor. The reason for the prevalence of gene duplication events over the loss of gene duplicates is presently unknown. Some of the sponge OAS genes may have been transformed into pseudogenes. At least two OAS pseudogene variants have been identified in *G. barretti* (Vallmann et al., 2011). Alternatively, the products of the gene duplicates may have acquired somewhat specific roles instead of being functionally redundant.

1.2. Sponge OASs have a distinct genomic structure (Publications II, III and unpublished)

To elucidate the common origin of the sponge and vertebrate OASs, the genomic structures of the OAS genes from *G. cydonium* and *T. ignis* were experimentally determined. Surprisingly, the exon/intron structure of sponge OASs was completely different if compared to the structure of vertebrate OAS genes. The OAS gene from *G. cydonium* had eight translated exons instead of the five present in the mammalian OAS unit. Furthermore, all intron positions relative to the conserved amino acid motifs were completely different (Publication II, Fig. 1-3). However, a similarity was observed in the order of the intron phases – all four introns in respective phases are shifted upstream in the sponge OAS gene in comparison with the human gene.

The exon/intron structure of the OAS gene from *T. ignis* was similar to that from *G. cydonium* with the exception that the first intron was missing (Publication III, Fig. 1). The three types of OAS genes found in the sequenced genome of *A. queenslandica* shared an exon/intron structure close to those in *G. cydonium* and *T. ignis* (Publication II, Fig. 5). No OAS with a vertebrate-like gene structure was identified in the genome of *A. queenslandica*.

The true orthologous genes in sponges and higher animals have highly conserved genomic structures including the positions and phases of introns. However, introns in sponge genes are significantly shorter if compared to those in mammalian genes (Müller et al., 2002; Srivastava et al., 2010). Short introns were also characteristic of the sponge OAS genes although the overall genomic structure was different from that of the mammalian OAS genes (Publication II, Fig. 1).

Also, the arrangement of the OAS genes in *A. queenslandica* genome was different from that in mammalian genomes where the OAS genes are clustered in a single locus (Mashimo et al., 2008). In *A. queenslandica* 2 types of OAS loci (one containing OAS1A and the other OAS1B and OAS1C) could be identified (Publication II, Fig. 6). The two loci were separated by at least 5 other genes. This is in contrast with the observation that the *A. queenslandica* gene structures and genome organization in general show marked conservation relative to other animals (Srivastava et al., 2010).

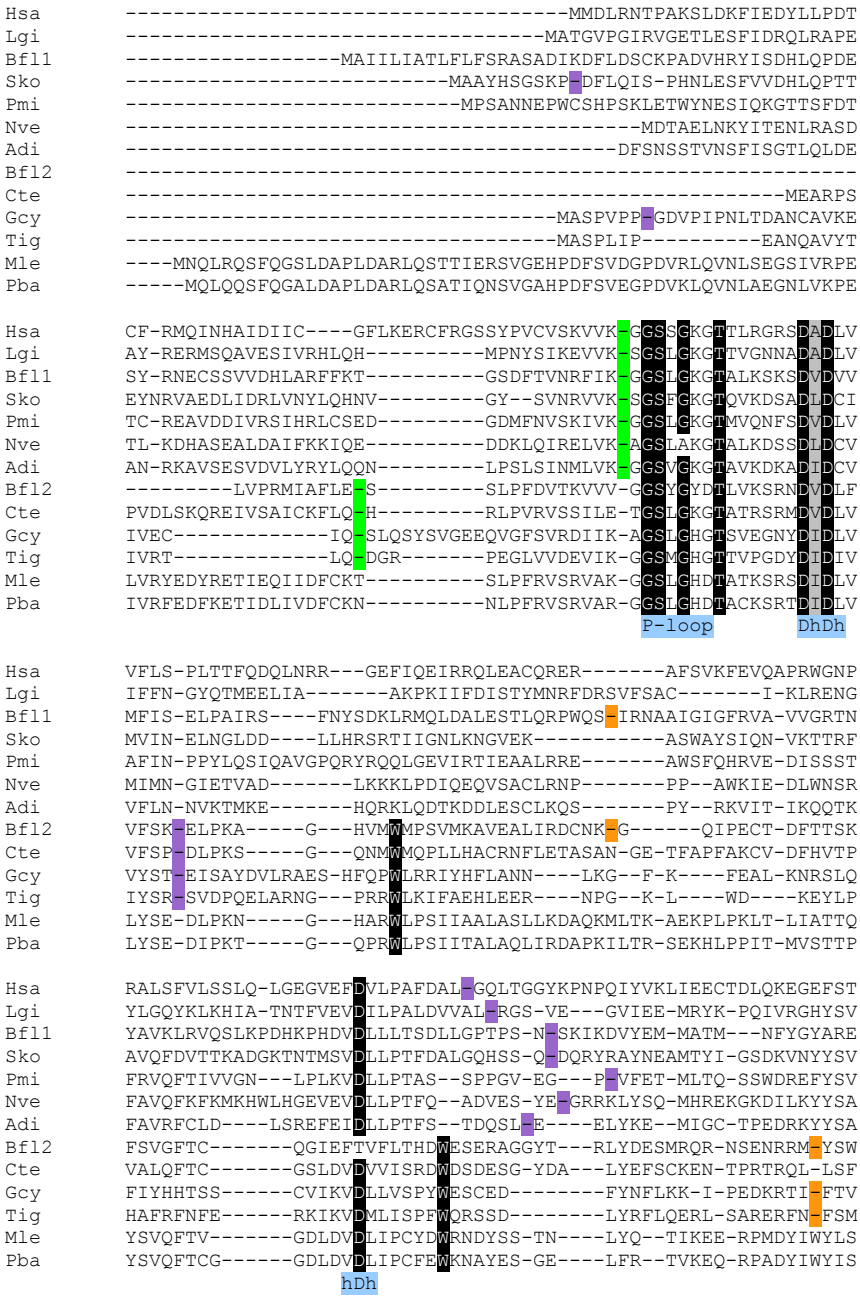
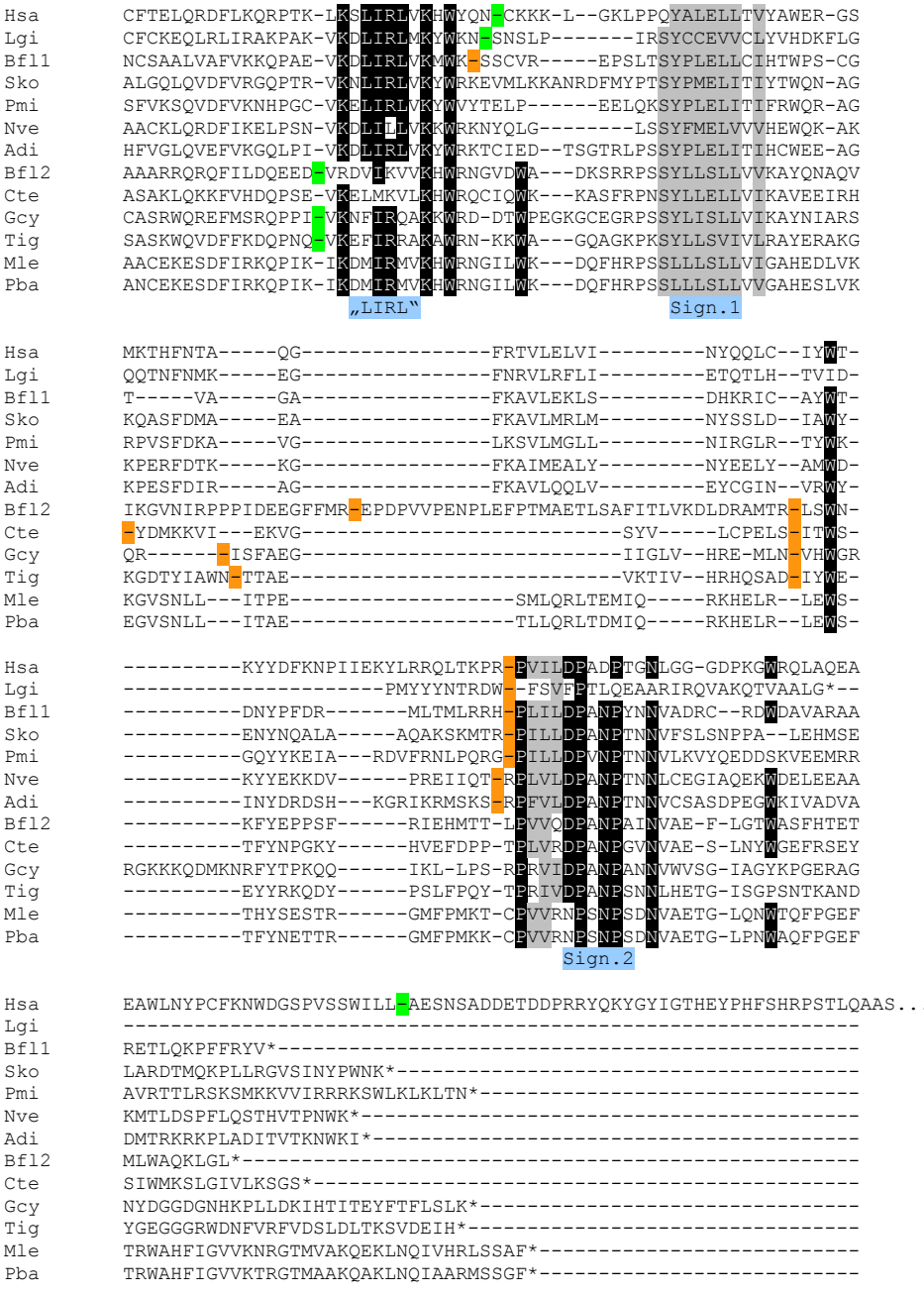


Figure 12. Alignment of OAS protein sequences from different animals. Hsa – *Homo sapiens* (phylum Chordata), Lgi – *Lottia gigantea* (Mollusca), Bfl – *Branchiostoma floridae* (Chordata), Sko – *Saccoglossus kowalevskii* (Hemichordata), Pmi – *Patiria miniata* (Echinodermata), Nve – *Nematostella vectensis* (Cnidaria), Adi – *Acropora digitifera* (Cnidaria), Cte – *Capitella teleta* (Annelida), Gcy – *Geodia cydonium*



(Porifera), Tig – *Tedania ignis* (Porifera), Mle – *Mnemiopsis leidyi* (Ctenophora), Pba – *Pleurobrachia bachei* (Ctenophora). The alignment was created using ClustalX and modified manually. The alignment is shaded on the basis of the position conservation. Intron positions in the corresponding genes are indicated as follows: green – phase 0 introns, violet – phase 1 introns, orange – phase 2 introns. In the case of OASs from *A. digitifera* and *B. floridae*, the beginning of the coding sequence remained undetermined.

The homology of amino acid sequences of sponge OAS proteins is low. Besides, minor variations exist between exon/intron structures of different OAS genes. Despite that, the genomic structure common to sponge OASs could be defined (Publication II, Fig. 7).

Previously, it had been suggested that the sponge OASs share a direct common ancestor with the vertebrate OASs (Schröder et al., 2008). However, the entirely different genomic structures of sponge and mammalian OASs are in conflict with the idea of a hypothetical direct intron-containing common ancestor for OASs. Though intron sliding has been described (Rogozin et al., 2012), it is highly improbable that in vertebrate OAS genes all introns migrated downstream if compared to sponge OAS genes. If intron gain and loss, the major mechanism of intron evolution, had given rise to the differences between human and sponge OAS genes, at least some of the intron positions should be conserved. Furthermore, it would not explain the observed similarity in the intron phase pattern.

To elucidate the evolution of the exon/intron pattern of the OAS genes, some of the available genome sequencing data from various representative species belonging to different animal phyla were analysed for the presence and structure of OAS genes. The alignment of the putative and also some previously known OAS sequences is presented on Fig. 12 (unpublished).

To date, the genomes of only two ctenophores, *Mnemiopsis leidyi* and *Pleurobrachia bachei*, have been sequenced (Ryan et al., 2013; Moroz et al., 2014). In the genome assembly of the sea walnut *M. leidyi* a putative OAS gene with no introns was found (GenBank accession AGCP01019774, bases 10130-11242). A similar intronless OAS gene was also present in the genome of *P. bachei* (GenBank accession AVPN01000435, bases 38556-39671). The predicted genomic structure of the OAS genes in these two ctenophores was supported by the presence of an EST sequence from another *Pleurobrachia* species, *P. pileus*, in the NCBI database (accession FP997205) and the transcriptome sequencing data from several other ctenophores (Moroz et al., 2014).

The OAS gene from a cnidarian (Torralba et al., 2008; Kjaer et al., 2009), sea anemone *Nematostella vectensis*, had 3 introns, which were in the same positions and phases as the corresponding introns in mammalian OAS genes. The genome of another representative of the phylum Cnidaria, the coral *Acropora digitifera*, contained an OAS gene with a similar genomic structure (GenBank accession BACK01020290, bases 1-3776). Putative OAS genes with introns in the same positions relative to conserved amino acid motifs could also be found in the genome assemblies of the bat star, *Patiria miniata*, a member of the phylum Echinodermata and of an acorn worm, *Saccoglossus kowalevskii*, from phylum Hemichordata (GenBank accessions AKZP01127852, bases 3331-7202 and ACQM01042178, bases 3815-7095, respectively).

Surprisingly, the exon/intron structure of the gene (GenBank accession AMQN01017518, bases 21712-22814) encoding the OAS protein in the annelid worm *Capitella teleta* (Kjaer et al., 2009), bore a resemblance to the genomic

structure of the OAS proteins from sponges. It contained 4 introns, the positions and phases of which coincided with those of the respective introns in sponge OAS gene sequences.

The gene (GenBank accession AMQO01003210, bases 49037-51279) encoding the predicted OAS from the owl limpet, *Lottia gigantea* (Kjaer et al., 2009), belonging to the phylum Mollusca, had an exon/intron structure similar to that of mammalian OAS genes.

Interestingly, the genome of the Florida lancelet, *Branchiostoma floridae*, contained at least two OAS genes (Kjaer et al., 2009) with different genomic structures. One of them (GenBank accession ABEP02019184, bases 25671-27886) had a genomic structure similar to that of mammalian OASs while the other one (GenBank accession ABEP02010054, bases 25017-29871 and further) resembled sponge OASs in this respect.

The phylogenetic tree of OAS protein sequences presented by Kjaer et al. (Kjaer et al., 2009) proposed the division of metazoan OASs into two large groups. One of them consists of OAS sequences from mammals and other vertebrates as well as from evolutionarily lower animals from phylums Cnidaria and Mollusca. The other group is mainly composed of OAS sequences from sponges and annelids. Interestingly, genomes of animals from subphyla Cephalochordata (*B. floridae*) and Tunicata (*C. intestinalis*) of phylum Chordata contain at least two divergent OAS sequences, one of which grouped with sponges and the other one with mammalian OASs (Kjaer et al., 2009). This division which is based on the amino acid sequences also seems to be supported by the exon/intron structure of the respective OAS genes (Fig. 12).

Analysed together, these data suggest that in animal evolution the early radiation of OAS genes has occurred. The last common ancestor of metazoa probably already had several different intron-containing OAS genes. Their subsequent evolution has led to their deletions, duplications and/or modifications while retaining the elements of the ancient exon/intron patterns.

2. The enzymatic activity of sponge OASs

2.1. The enzymatic activity of sponge OASs (Publications I and III)

Though the cDNAs encoding putative OASs had been cloned from several marine sponges (Wiens et al., 1999; Grebenjuk et al., 2002; Vallmann et al., 2011), it had not been demonstrated before whether they encode enzymatically active proteins or inactive OAS1 isoforms.

In this work, OASs from two marine sponges, *G. cydonium* and *T. ignis*, were produced as recombinant proteins in the bacterial expression system and subjected to activity assays. Both of them were able to catalyse the synthesis of 2',5'-linked oligoadenylates from ATP, proving that they were indeed active 2',5'-oligoadenylate synthetases.

In the reaction conditions used, the recombinant enzymes mainly catalysed

the synthesis of 2-5A dimers and trimers, the longest oligomer observed was a tetramer (Publication I, Fig. 8; Publication III, Fig. 4G-H). However, the 2-5A dimer remained the main product even at a high ATP conversion. Furthermore, the specific activity of the recombinant OASs was low if compared to the activity observed in crude extracts of the sponges (Reintamm et al., 2003b). The amount of ATP converted to the products was in the same range for the purified recombinant protein and for the sponge crude extract.

The low activity of the recombinant proteins may be due to several reasons. Most of the polypeptide produced in bacterial expression system may be incorrectly folded and enzymatically inactive. In addition, the specific dsRNA, poly(I)·poly(C), was obviously not the proper activator of the studied sponge OASs. The OAS from *G. cydonium* was already in complex with bacterial RNA of heterogenous composition. Evidently, this RNA was not the proper activator for this protein. Furthermore, the bound RNA pool could include components that are inhibitory or activate the OASs poorly. It is also likely that the OASs expressed here are not the ones responsible for the high 2-5A synthesising activities observed in sponge extracts. Sponge genomes contain several different OAS genes (see Results and Discussion, Section 1.1), the products of which contribute to the total product pool catalysed by the oligoadenylate synthesising activity of the specific sponge studied. It is possible that the high enzymatic activity in sponge extracts is due to the presence of a multitude of OAS proteins. Alternatively, the existence of (a) highly active OAS(s) responsible for the synthesis of the majority of the 2-5A oligomers could be speculated. Evidently, it would be different from the ones studied in this work.

2.2. Sponge OASs are differentially activated by RNA (Publications I and III)

All known vertebrate OASs are produced as latent enzymes and require activation by their cofactor, dsRNA. For studying the enzymatic activity of OASs in *in vitro* assays, the synthetic dsRNA, poly(I)·poly(C) is commonly used, although it is not the native cofactor for these enzymes.

Surprisingly, the recombinant OAS from *G. cydonium* was able to catalyse the formation of 2-5A oligomers from ATP *per se* and the addition of poly(I)·poly(C) only modified the activity less than 2-fold (Publication I, Fig. 2). The UV-spectrum of the protein preparation was not characteristic of that of the pure protein. Instead, it resembled the spectrum of nucleic acid indicating that the recombinant protein preparation could be contaminated with nucleic acids. The HPLC analysis revealed that the anomalous UV-spectrum for a protein was caused by the presence of RNA (Publication I, Fig. 3). The RNA evidently copurified from the bacterial lysate in complex with the protein.

Characterization of the preparation by using electrophoresis in a native polyacrylamide gel and by size exclusion chromatography demonstrated that the recombinant protein preparation consisted of a set of heterogenous complexes of RNA and recombinant GcyOAS (Publication I, Fig. 4 and 5). Analysis of the size exclusion chromatography fractions showed that the

specific activity of the protein was related to the number of bound nucleotides per protein monomer. The preparations with larger amounts of nucleotides per protein molecule had higher specific activities (Publication I, Fig. 6).

Several approaches were launched in order to free the recombinant protein preparation from the bound RNA of bacterial origin. Nuclease treatments were performed during the purification of the recombinant GcyOAS from bacterial lysate. However, the resulting protein preparation still contained RNA and was capable of catalysing the synthesis of oligoadenylates without the addition of any poly(I)·poly(C) to the reaction mixture. Nuclease treatments performed during activity assays only slightly modified the enzymatic activity of the recombinant GcyOAS preparation (Publication I, Fig. 7). The low efficiency of these treatments suggested that RNA in these complexes was not readily accessible to the action of nucleases. On the other hand, when high doses of the nuclease were added to the protein preparation, the rapid formation of a precipitate was observed. The nuclease apparently degraded unprotected regions of the RNA in the negatively charged GcyOAS-RNA complex and caused its precipitation when the complex became electrically neutral. Thus, an efficient nuclease treatment of the RNA-protein complex resulted in a certain critical point in its precipitation, which was likely related to the pI of the complex.

Alternatively, alkaline buffers were used in the protein purification procedures in order to separate the protein from RNA. The obtained recombinant GcyOAS preparation lost its OAS activity though it still contained RNA. The loss of activity could be explained by the presence of some alkali-labile minor component in the activating RNA that is essential for the activation of GcyOAS.

The experiments suggest that the RNA derived from *E. coli* was bound to the recombinant GcyOAS with a high affinity, being partially protected from RNase degradation in these complexes.

The recombinant OAS from *T. ignis* did not bind nucleic acid during its expression in *E. coli* cells and was purified as a pure protein. Like mammalian OASs, it needed activation by poly(I)·poly(C) for its enzymatic activity (Publication III, Fig. 3C). However, the amount of poly(I)·poly(C) needed for the activation was rather high (at least 1 mg/ml) if compared to the requirements of the non-sponge OASs studied (Publication III, Fig. 3). A high concentration of the poly(I)·poly(C) required may refer to the poor suitability of this synthetic RNA for the activation TigOAS. For comparison, the necessary amount of dsRNA to achieve a half-maximum activity of different mammalian OAS forms (OAS1, OAS2 and OAS3) varies greatly, ranging from less than 1 μg/ml to more than 500 μg/ml poly(I)·poly(C) (Justesen et al., 2000).

The sponge OASs evidently have low affinity to poly(I)·poly(C), indicated by the inability of this synthetic dsRNA to compete with the bacterial RNA bound to GcyOAS and by the poor activation of TigOAS. It is likely that the RNA recognition by the sponge OASs differs from that exhibited by vertebrate OASs.

The crystal structure of the human OAS1 in complex with dsRNA reveals that the OAS interacts with two adjacent minor grooves of the activating dsRNA. The amino acid residues making contact with the RNA are predominantly located in the N-terminal part of the protein – in the helices αN2 and αN3, strand β1 and in the helix αN4 that is formed during the conformational change upon dsRNA binding (Donovan et al., 2013). The residues in human OAS1 shown to make contact with dsRNA are not all conserved in sponge OASs. The region corresponding to αN2 and αN3 in human OAS1 is highly variable in length, as well as in amino acid sequence among sponge proteins. However, based on the secondary structure predictions, in OASs from *T. ignis* and *G. cydonium* this region most likely consists of structural elements with helical conformation similarly to that in mammalian OAS proteins (http://bioinf.cs.ucl.ac.uk/psipred/, accessed 08.04.2014; Hartmann et al., 2003; Donovan et al., 2013).

The sponge OAS may need the RNA with specific primary and secondary structure elements for its activation. Poly(I)·poly(C) as a synthetic dsRNA may meet these requirements only partially. Various viral and cellular RNA molecules containing double-stranded regions have been shown to activate mammalian OASs (see Section 2.2.2). It has been demonstrated that the OASs are able to bind DNA or single-stranded RNA, as well. However, these molecules were not able to activate the enzyme (Marié et al., 1990a). Still, some ssRNA aptamers with only a few double-stranded regions were potent activators for OASs (Hartmann et al., 1998a). In experiments with GcyOAS, in addition to commonly used poly(I)·poly(C), the effect of double-stranded poly(A)·poly(U) as well as some homopolymeric ssRNAs on the activity of the recombinant protein was tested. No significant increase in the OAS activity was observed, indicating that the tested polynucleotides, like poly(I)·poly(C), were not able to compete with the bound RNA of bacterial origin (Publication I, Fig. 2).

The sponge crude extracts exhibit OAS activity without the addition of any nucleic acid activator (Reintamm et al., 2003b). Furthermore, it has been demonstrated that the OAS activity exhibited by crude extracts of *G. cydonium* was not abolished by nuclease treatments (Lopp et al., 2002). This may refer to the existence of a strong endogenous nucleic acid-protein complex in the crude extracts of *G. cydonium*. It could also be the case in other sponges where the OAS activity is present. Thus it can be hypothesised, that though the activation mechanism of sponge OASs may be similar to that of vertebrate ones, the affinities and structural requirements for the RNA component are probably quite different.

2.3. Sponge OASs catalyse the formation of both 2',5'- and 3',5'-linkages (Publications I, III and unpublished)

Surprisingly, in addition to the oligoadenylates with 2',5'-linkages, products with 3',5'-phosphodiester and mixed linkages were detected among the reaction products synthesised by the OASs from *G. cydonium* and *T. ignis*. The synthesis

of oligoadenylates containing 3',5'-linkages has not been described in enzymatic assays with mammalian OASs. Due to the high sensitivity of the detection method used in this work, the presence of oligoadenylates with 3',5'-linkages was also detected among the reaction products synthesised by the porcine OAS (SscOAS). However, the SscOAS was able to catalyse the synthesis of 3',5'-linkages at a very low level – the yield of 2-5A was about three orders of magnitude higher than that of 3-5A. The ratios of 2-5A oligomers to 3-5A oligomers in the reactions catalysed by the GcyOAS-RNA complex and by the poly(I)·poly(C)-activated TigOAS were 5.4 and 2.4, respectively (Publication I, Fig. 8, Publication III, Fig. 4G-H).

The 2'- and 3'-specificities of different OASs probably depend on the structures of their active sites. The distinct assembly and structure of the active sites and consequently the choice of linkage to be catalysed is presumably influenced by the dissimilar amino acid context surrounding the active site residues. Based on the crystal structures of enzymes from the nucleotidyl transferase family, the 2'- and 3'-specificities are believed to be determined through an orientation of the acceptor nucleotide so that the respective hydroxyl of the ribose would be in a favourable position to react (Hartmann et al., 2003). It is possible, that in sponge OASs the active site is more flexible than that in 2'-specific mammalian OASs and thus it is able to accommodate the nucleotide in both orientations.

From the amino acid sequence alignments of mammalian OAS active sites it is evident that, apart from a small number of exceptions, all enzymatically active 2'-specific mammalian OASs contain an alanine residue between the two conserved aspartate residues (Rogozin et al., 2003; Donovan et al., 2013). In sponge OAS sequences, however, isoleucine, leucine or valine residue can be found in this position (Justesen et al., 2000; Rogozin et al., 2003; Schröder et al., 2008; Publication II, Fig. 7, Publication III, Fig. 1). A similar selection of amino acid residues appears in the active sites of different 3'-specific rNTs, such as PAPs, PUPs and CCAtrs (Kuchta et al., 2009). None of these enzymes seems to have alanine residue in this position.

Based on this observation and the ability of sponge OASs to catalyse the synthesis of the 3',5'-linkages, it was hypothesised that this residue between the two Mg^{2+}-coordinating aspartates may have a role in determining the substrate orientation and therefore the linkage to be catalysed. From the crystal structures of porcine and human OAS proteins (Hartmann et al., 2003; Donovan et al., 2013) it is seen that the side chain of this residue is pointed to the opposite side of the β-strand and not towards the active site, where it could directly influence the position of the aspartate residues or interact with the nucleotide coordinating residues. However, we considered it possible that the side chain of this residue may still have an effect on the β-strand structure and thus on the active site if the relatively small side chain of alanine is changed to somewhat larger side chains of valine, isoleucine or leucine.

To test this hypothesis, we mutated the alanine residue in the active site of the 2'-specific porcine OAS to isoleucine generating the SscOAS A75I mutant.

Similarly, the corresponding isoleucine was mutated to alanine (TigOAS I51A mutant) in the OAS from *T. ignis* that catalysed the synthesis of 2',5'- as well as 3',5'-linkages. Both mutant enzymes, as well as their wild type counterparts were expressed as histidine-tagged proteins in the bacterial system and purified via binding to Ni-NTA agarose. The activity assays revealed that in the case of both mutants the specificities of the mutant enzymes were the same as those of the corresponding wild type enzymes. The ratios of 2-5A oligomers to 3-5A oligomers synthesised by the I51A mutant of TigOAS and the A75I mutant of SscOAS were identical to those synthesised by the wild type TigOAS and SscOAS, respectively (our unpublished data).

Evidently, changing this single amino acid residue was not sufficient to modify the active site structure enough to allow the enzyme, evolved to catalyse only the 2',5'-linkages, to become more flexible and to be able to catalyse the formation of 3',5'-linkages, as well. Similarly, the enzyme synthesising both linkages could not be made specific to only one linkage by this small modification. It is likely that the structure and flexibility of the active sites of OASs is determined by several (other) amino acid residues surrounding the highly conserved motifs. Recently, an enzyme closely related to OASs, the cGAS, was reprogrammed to catalyze the formation of 3',5'-phosphodiester linkage instead of the 2',5'-linkage it usually synthesises as the first step in the production of cGAMP. It was shown that in the human cGAS a recessed side chain (T211) at the bottom of the active site pocket and an extended side chain (R376) at the top of the pocket induce the orientation of the substrate nucleotides that favours the formation of the 2',5'-linkage between them. When the recessed and extended side chains were exchanged, the enzyme catalysed the formation of 3',5'-linked cGAMP (Kranzusch et al., 2014).

In the case of OASs, the 2'- and 3'- specificities may also be determined by the activating RNA species since the RNA participates in the assembly of the active site (Donovan et al., 2013). The results of the present study indicate that the poly(I)·poly(C) directs the sponge OAS to form an active site which favours the catalysis of 2',5'-linkage over 3',5'-linkage. The ratio of 2-5A to 3-5A was slightly shifted in favour of 2-5A oligomers when the poly(I)·poly(C) was added to the activity assays with the GcyOAS-RNA complex. It is possible that the bacterial RNA bound to recombinant GcyOAS contains components that direct the enzyme to synthesise 2',5'-linkages, as well as components that induce the catalysis of 3',5'-linkages. In the cell extracts of *T. ignis* mostly the synthesis of 3',5'-linked oligoadenylates is detected (Lopp et al., 2012). It could be hypothesised that the recombinant TigOAS would behave as a 3-5A synthetase in the presence of a suitable activator. The poly(I)·poly(C) used to activate the recombinant protein may guide the active site to take the conformation that prefers the catalysis of 2',5'-linkages. Explaining the structural determinants of the RNA for favouring the synthesis of one or the other linkage would be most interesting. However, the synthesis of both linkages in the presence of a single activator suggests that the active sites of sponge OASs are, in any case, rather flexible.

The sponge genomes encode different types of OAS proteins with several variants which potentially differ in their catalytic properties. It is reasonable to assume that the activities seen in the crude extract of a sponge, are combinations of activities exhibited by the different OASs present. For the respective sponge species, *G. cydonium* or *T. ignis*, the cloned OAS protein could be one of several OASs with different activation and enzymatic characteristics.

Based on their amino acid sequences, as well as the genomic structures of the respective genes, the OAS proteins can be divided into two groups (Kjaer et al., 2009; Fig. 12). Considering the different specificities of the studied examples from both groups (the 2'-specifity of the mammalian OASs and the ability of sponge OASs to catalyse the formation of 3',5'-linkages, as well) it is interesting to speculate whether the putative OASs from annelids and tunicates, grouped together with sponge OASs on the basis of their genomic structures, may catalyse the formation of both linkages like sponge enzymes do.

Mammalian OASs have been shown to be able to use different NTPs as donor substrates *in vitro*. However, the addition of a nucleoside triphosphate other than ATP terminates the chain elongation (Justesen et al., 1980; Marié et al., 1997). Surprisingly, this was not the case for sponge OASs. It has been shown that the OASs present in sponge crude extracts were able to polymerize different nucleotides, surprisingly even CTP and UTP, into 2',5'-phospodiester linked oligonucleotides. The ability to use different nucleotides as substrates and the assortment of products synthesised was dependant on the sponge species (Lopp et al., 2010). Since other substrates besides ATP were not tested in this work, the capability of the studied recombinant OASs to accept different nucleotides as substrates remains unknown. Still, it is clear that the active site structure of sponge OASs differs from that of mammalian enzymes. Determining the three-dimensional structure of a sponge OAS in complex with the activating RNA and a substrate analog would be of great interest and a valuable tool for elucidating the nature of the activation mechanism and specificity of sponge OASs. Moreover, it would offer additional guidelines for engineering novel enzymes with specific properties.

3. Enzymatically active OASs are widely distributed among Metazoa (Publication III)

Rapid increase in the number of animals whose genomes have been sequenced and growing databases of EST and transcriptome sequences have provided valuable tools for determining the presence of OAS genes in the genomes of different multicellular animals throughout the evolutionary tree. Therefore, it is now believed that the OAS proteins are widely distributed among Metazoa. However, losses of OAS genes have occurred in some branches, e.g. in teleost fish (Kjaer et al., 2009).

Herein, in addition to the sponge OASs, some bioinformatically predicted OASs from several other distantly related multicellular animals, as well as their

unicellular relative were examined. The OASs from the salamander *A. mexicanum* and the little skate *L. erinacea* were chosen to represent the evolutionary branches of vertebrates where the OAS activity had not been detected before. The OAS from the mollusk *M. californianus* was of great interest since no OAS activity had been previously demonstrated in the whole Protostomia branch. Furthermore, several protostome lineages have probably lost OAS genes thus making it even more interesting to show OAS activity in this branch of evolution. Since the OAS activity had been previously believed to be related only to Metazoa, the prediction of a putative OAS gene in the genome of the unicellular choanoflagellate *M. brevicollis* (Kjaer et al., 2009) was unexpected. Therefore it was particularly interesting to test the enzymatic activity of the choanoflagellate OAS. All proteins were expressed in the bacterial expression system and subsequently the oligoadenylate activity assays were performed to test their capability to synthesise oligoadenylates from ATP. The OASs from the little skate, *L. erinacea*, and the mollusk, *M. californianus*, proved to be active oligoadenylate synthetases (Publication III, Fig. 4). Two of the tested proteins, OASs from the salamander *A. mexicanum* and from the choanoflagellate *M. brevicollis* did not exhibit oligoadenylate synthesising activity in our assays. Possible reasons for this were discussed in Publication III. Thus, based on the current data, it is not possible to confirm or dismiss the hypothesis that the OAS activity may be a property restricted only to Metazoa.

The demonstration of the enzymatically active OAS from *M. californianus* was the first case reported for the whole animal clade Protostomia. It is increasingly clear that the active OASs are indeed present in many animal phyla throughout the evolutionary tree and no evolutionary gap exists between vertebrate and sponge OASs (Fig. 13). This is further confirmed by the 2',5'-oligoadenylate synthesizing activity exhibited by the corals *Alcyonium acaule* and *Gersemia fruticosa*, representatives of the phylum Cnidaria (A. Lopp, unpublished data).

The other studied OAS with enzymatic activity, that from *L. erinacea*, was the evolutionarily oldest vertebrate species for which OAS activity has been shown so far. The particular OAS contains the ubiquitin-like repeats in its C-terminus (Kjaer et al., 2009) and is therefore considered to be an OAS-like protein (OASL). Originally the OASL proteins were believed not to be able to produce 2-5A oligomers (Rebouillat et al., 1998; Hartmann et al., 1998b). However, the cDNAs encoding OASL proteins with 2-5A synthesising activity have been cloned before (Yamamoto et al., 1998; Eskildsen et al., 2003).

Both OASs, those from *M. californianus* and *L. erinacea*, resembled mammalian OASs in respect to their activation requirements and enzymatic activity. They were expressed as inactive proteins and needed activation by dsRNA. Moreover, the required amount of poly(I)·poly(C) was significantly lower if compared to the amount poly(I)·poly(C) needed for the activation of sponge OASs (Publication III, Fig. 3). Also in resemblance to mammalian OAS, the enzymes from the mollusk and the little skate catalysed only the formation of 2',5'-oligoadenylates; no products with 3',5'-linkages were detected among

the reaction products (Publication III, Fig. 4).

The genomes of *L. erinacea* or *M. californianus* are not sequenced yet and thus the exon/intron structure of their OAS genes in unknown. However, the OAS gene in the genome of another member of class Chondrichthyes, the elephant shark, *C. milii*, has an exon/intron structure similar to that of mammalian OAS genes. Therefore it is reasonable to assume that the OAS gene in *L. erinacea* may also have the mammalian type exon/intron organization. The same is true for the OAS gene from *M. californianus*, since the respective genes in the genome of the owl limpet, *L. gigantea*, and of the Mediterranean mussel *Mytilus galloprovincialis*, have the genomic structure homologous with that of mammalian OAS genes, as well (Fig. 12 and data not shown).

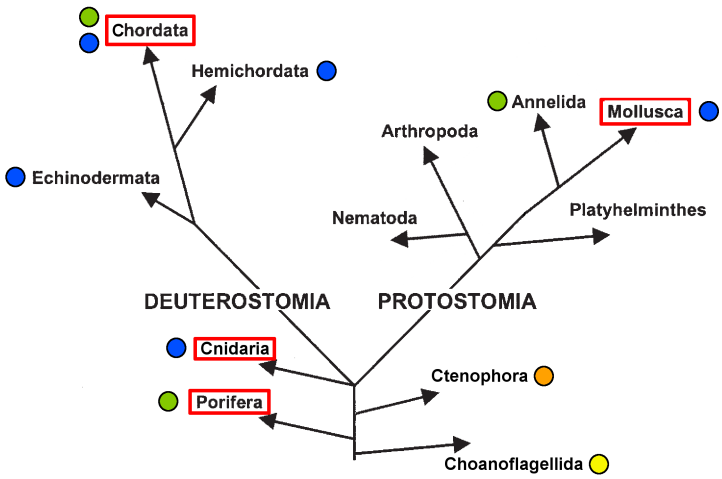


Figure 13. A schematic phylogenetic tree of selected animal phyla showing the distribution of OAS genes with sponge-like (green dot) and vertebrate-like (blue dot) genomic structures, as well as intronless OAS genes (orange dot). The predicted OAS gene from choanoflagellate *M. brevicollis* (marked with yellow dot) has an exon/intron structure different from both, sponge and vertebrate, OAS gene structures. The animal phyla, where the OAS activity has been shown, are indicated with red borders.

To conclude, the current data show that the OASs are proteins of ancient origin that have been preserved in the majority of modern animal lineages. At least two different types of OASs with distinct genomic structures and possibly distinct enzymatic characteristics have arisen early in the animal evolution. Both the sponge type OASs and the mammalian type OASs have spread across the evolutionary tree of animals, being present in deuterostome, as well as in the protostome branch (Fig. 13). However, it appears that the choice to preserve either one or the other type has been made at the phylum level. The exception is the phylum Chordata, some members of which contain the sponge type OASs in their genomes in addition to the mammalian type OASs. Most animals

belonging to this phylum have only the mammalian type OASs. Still, more data are needed to make certain conclusions, since only few representatives from each lineage could be analysed in this respect. Both types of OASs have acquired different characteristics during evolution. The mammalian type OASs appear to be highly 2'-specific while the sponge type OASs may be more linkage-tolerant. The activation properties also differ between the two types of OAS proteins. However, OAS proteins with sponge type genomic structures from other phyla (e.g. Annelida) should be examined to substantiate this hypothesis. Since the sponge OAS proteins are more flexible in their catalytic properties, it may be hypothesised that these enzymes represent the link between the 3'-specific nucleotidyl transferases and the 2'-specific OASs.

CONCLUSIONS

Based on the results obtained in this study, the following conclusions can be drawn:

- Enzymatically active 2',5'-oligoadenylate synthetases are widely distributed among Metazoa, including the protostome lineage, where the OAS activity had not been shown earlier.
- Sponge 2',5'-oligoadenylate synthetase genes do not share a direct intron-containing ancestor with mammalian OAS genes.
- At least two subgroups (the sponge type and the mammalian type) of 2',5'-oligoadenylate synthetases exist, both of which are represented in the deuterostome as well as in the protostome branch of animal evolution.
- Sponge 2',5'-oligoadenylate synthetases are diverse proteins; their primary structures as well as enzymatic properties vary between OASs from different sponge species.
- The requirements for the activation of sponge 2',5'-oligoadenylate synthetase proteins are different from those of mammalian enzymes.
- Sponge 2',5'-oligoadenylate synthetases are more linkage-tolerant than the highly 2'-specific mammalian OASs; this could be related to the dissimilar structure of their active sites. The dsRNA could have a role in determining the 2'- and 3'-specificities in the case of sponge OASs as well.

REFERENCES

Abbas S, Kelly M, Bowling J, Sims J, Waters A, Hamann M. (2011). Advancement into the Arctic region for bioactive sponge secondary metabolites. *Mar Drugs*, 9(11):2423-2437.

Anderson BR, Muramatsu H, Jha BK, Silverman RH, Weissman D, Karikó K. (2011). Nucleoside modifications in RNA limit activation of 2'-5'-oligoadenylate synthetase and increase resistance to cleavage by RNase L. *Nucleic Acids Res*, 39(21):9329-9338.

Aravind L, Koonin EV. (1999). DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res*, 27(7):1609-1618.

Augustin MA, Reichert AS, Betat H, Huber R, Mörl M, Steegborn C. (2003). Crystal structure of the human CCA-adding enzyme: insights into template-independent polymerization. *J Mol Biol*, 328(5):985-994.

Baglioni C, Maroney PA. (1980). Mechanisms of action of human interferons. Induction of 2'5'-oligo(A) polymerase. *J Biol Chem*, 255(18):8390-8393.

Balbo PB, Bohm A. (2007). Mechanism of poly(A) polymerase: structure of the enzyme-MgATP-RNA ternary complex and kinetic analysis. *Structure*, 15(9):1117-1131.

Balbo PB, Bohm A. (2009). Proton transfer in the mechanism of polyadenylate polymerase. *Biochem J*, 420(2):229-238.

Bandyopadhyay S, Ghosh A, Sarkar SN, Sen GC. (1998). Production and purification of recombinant 2'-5' oligoadenylate synthetase and its mutants using the baculovirus system. *Biochemistry*, 37(11):3824-3830.

Behera AK, Kumar M, Lockey RF, Mohapatra SS. (2002). 2'-5' Oligoadenylate synthetase plays a critical role in interferon-gamma inhibition of respiratory syncytial virus infection of human epithelial cells. *J Biol Chem*, 277(28):25601-25608.

Bell JJ. (2008). The functional roles of marine sponges. *Estuarine, Coastal and Shelf Science*, 79():341-353.

Benech P, Vigneron M, Peretz D, Revel M, Chebath J. (1987). Interferon-responsive regulatory elements in the promoter of the human 2',5'-oligo(A) synthetase gene. *Mol Cell Biol*, 7(12):4498-4504.

Besse S, Rebouillat D, Marie I, Puvion-Dutilleul F, Hovanessian AG. (1998). Ultrastructural localization of interferon-inducible double-stranded RNA-activated enzymes in human cells. *Exp Cell Res*, 239(2):379-392.

Betat H, Rammelt C, Mörl M. (2010). tRNA nucleotidyltransferases: ancient catalysts with an unusual mechanism of polymerization. *Cell Mol Life Sci*, 67(9):1447-1463.

Bisbal C, Martinand C, Silhol M, Lebleu B, Salehzada T. (1995). Cloning and characterization of a RNAse L inhibitor. A new component of the interferon-regulated 2-5A pathway. *J Biol Chem*, 270(22):13308-13317.

Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. (2012). Marine natural products. *Nat Prod Rep*, 29(2):144-222.

Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. (2013). Marine natural products. *Nat Prod Rep*, 30(2):237-323.

Blunt JW, Copp BR, Keyzers RA, Munro MHG, Prinsep MR. (2014). Marine natural products. *Nat Prod Rep*, 31(2):160-258.

Castelli JC, Hassel BA, Maran A, Paranjape J, Hewitt JA, Li XL, Hsu YT, Silverman RH, Youle RJ. (1998). The role of 2'-5' oligoadenylate-activated ribonuclease L in apoptosis. *Cell Death Differ*, 5(4):313-320.

Castelli JC, Hassel BA, Wood KA, Li XL, Amemiya K, Dalakas MC, Torrence PF, Youle RJ. (1997). A study of the interferon antiviral mechanism: apoptosis activation by the 2-5A system. *J Exp Med*, 186(6):967-972.

Cayley PJ, White RF, Antoniw JF, Walesby NJ, Kerr IM. (1982). Distribution of the ppp(A2'p)nA-binding protein and interferon-related enzymes in animals, plants, and lower organisms. *Biochem Biophys Res Commun*, 108(3):1243-1250.

Chebath J, Benech P, Hovanessian A, Galabru J, Revel M. (1987a). Four different forms of interferon-induced 2',5'-oligo(A) synthetase identified by immunoblotting in human cells. *J Biol Chem*, 262(8):3852-3857.

Chebath J, Benech P, Revel M, Vigneron M. (1987b). Constitutive expression of (2'-5') oligo A synthetase confers resistance to picornavirus infection. *Nature*, 330(6148):587-588.

Civril F, Deimling T, de Oliveira Mann CC, Ablasser A, Moldt M, Witte G, Hornung V, Hopfner K. (2013). Structural mechanism of cytosolic DNA sensing by cGAS. *Nature*, 498(7454):332-337.

Coccia EM, Romeo G, Nissim A, Marziali G, Albertini R, Affabris E, Battistini A, Fiorucci G, Orsatti R, Rossi GB, et al.. (1990). A full-length murine 2-5A synthetase cDNA transfected in NIH-3T3 cells impairs EMCV but not VSV replication. *Virology*, 179(1):228-233.

Cohen B, Peretz D, Vaiman D, Benech P, Chebath J. (1988). Enhancer-like interferon responsive sequences of the human and murine (2'-5') oligoadenylate synthetase gene promoters. *EMBO J*, 7(5):1411-1419.

Dan M, Zheng D, Field LL, Bonnevie-Nielsen V. (2012). Induction and activation of antiviral enzyme 2',5'-oligoadenylate synthetase by in vitro

transcribed insulin mRNA and other cellular RNAs. *Mol Biol Rep*, 39(7):7813-7822.

Deng J, Ernst NL, Turley S, Stuart KD, Hol WGJ. (2005). Structural basis for UTP specificity of RNA editing TUTases from Trypanosoma brucei. *EMBO J*, 24(23):4007-4017.

Der SD, Zhou A, Williams BR, Silverman RH. (1998). Identification of genes differentially regulated by interferon alpha, beta, or gamma using oligonucleotide arrays. *Proc Natl Acad Sci U S A*, 95(26):15623-15628.

Desai SY, Patel RC, Sen GC, Malhotra P, Ghadge GD, Thimmapaya B. (1995). Activation of interferon-inducible 2'-5' oligoadenylate synthetase by adenoviral VAI RNA. *J Biol Chem*, 270(7):3454-3461.

Díaz-Guerra M, Rivas C, Esteban M. (1997). Activation of the IFN-inducible enzyme RNase L causes apoptosis of animal cells. *Virology*, 236(2):354-363.

Dohrmann M, Wörheide G. (2013). Novel scenarios of early animal evolution--is it time to rewrite textbooks?. *Integr Comp Biol*, 53(3):503-511.

Domingo-Gil E, Esteban M. (2006). Role of mitochondria in apoptosis induced by the 2-5A system and mechanisms involved. *Apoptosis*, 11(5):725-738.

Domingo-Gil E, González JM, Esteban M. (2010). Identification of cellular genes induced in human cells after activation of the OAS/RNaseL pathway by vaccinia virus recombinants expressing these antiviral enzymes. *J Interferon Cytokine Res*, 30(3):171-188.

Dong B, Silverman RH. (1995). 2-5A-dependent RNase molecules dimerize during activation by 2-5A. *J Biol Chem*, 270(8):4133-4137.

Donovan J, Dufner M, Korennykh A. (2013). Structural basis for cytosolic double-stranded RNA surveillance by human oligoadenylate synthetase 1. *Proc Natl Acad Sci U S A*, 110(5):1652-1657.

Ereskovsky AV. (2010). The Comparative Embryology of Sponges. Dordrecht, Heidelberg, London, New York : Springer.

Erwin DH, Laflamme M, Tweedt SM, Sperling EA, Pisani D, Peterson KJ. (2011). The Cambrian conundrum: early divergence and later ecological success in the early history of animals. *Science*, 334(6059):1091-1097.

Eskildsen S, Hartmann R, Kjeldgaard NO, Justesen J. (2002). Gene structure of the murine 2'-5'-oligoadenylate synthetase family. *Cell Mol Life Sci*, 59(7):1212-1222.

Eskildsen S, Justesen J, Schierup MH, Hartmann R. (2003). Characterization of the 2'-5'-oligoadenylate synthetase ubiquitin-like family. *Nucleic Acids Res*, 31(12):3166-3173.

Fisch KM, Gurgui C, Heycke N, van der Sar SA, Anderson SA, Webb VL, Taudien S, Platzer M, Rubio BK, Robinson SJ, Crews P, Piel J. (2009).

Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting. *Nat Chem Biol*, 5(7):494-501.

Funayama N. (2010). The stem cell system in demosponges: insights into the origin of somatic stem cells. *Dev Growth Differ*, 52(1):1-14.

Gazave E, Lapébie P, Ereskovsky AV, Vacelet J, Renard E, Cárdenas P, Borchiellini C. (2012). No longer Demospongiae: Homoscleromorpha formal nomination as a fourth class of Porifera. *Hydrobiologia*, (687):3-10.

Gazave E, Lapébie P, Renard E, Vacelet J, Rocher C, Ereskovsky AV, Lavrov DV, Borchiellini C. (2010). Molecular phylogeny restores the supra-generic subdivision of homoscleromorph sponges (Porifera, Homoscleromorpha). *PLoS One*, 5(12):e14290.

Ghosh A, Sarkar SN, Guo W, Bandyopadhyay S, Sen GC. (1997). Enzymatic activity of 2'-5'-oligoadenylate synthetase is impaired by specific mutations that affect oligomerization of the protein. *J Biol Chem*, 272(52):33220-33226.

Ghosh A, Sarkar SN, Rowe TM, Sen GC. (2001). A specific isozyme of 2'-5' oligoadenylate synthetase is a dual function proapoptotic protein of the Bcl-2 family. *J Biol Chem*, 276(27):25447-25455.

Ghosh A, Sarkar SN, Sen GC. (2000). Cell growth regulatory and antiviral effects of the P69 isozyme of 2-5 (A) synthetase. *Virology*, 266(2):319-328.

Grasela JJ, Pomponi SA, Rinkevich B, Grima J. (2012). Efforts to develop a cultured sponge cell line: revisiting an intractable problem. *In Vitro Cell Dev Biol Anim*, 48(1):12-20.

Grebenjuk VA, Kuusksalu A, Kelve M, Schütze J, Schröder HC, Müller WEG. (2002). Induction of (2'-5')oligoadenylate synthetase in the marine sponges Suberites domuncula and Geodia cydonium by the bacterial endotoxin lipopolysaccharide. *Eur J Biochem*, 269(5):1382-1392.

Hamill S, Wolin SL, Reinisch KM. (2010). Structure and function of the polymerase core of TRAMP, a RNA surveillance complex. *Proc Natl Acad Sci U S A*, 107(34):15045-15050.

Han Y, Donovan J, Rath S, Whitney G, Chitrakar A, Korennykh A. (2014). Structure of human RNase L reveals the basis for regulated RNA decay in the IFN response. *Science*, 343(6176):1244-1248.

Hartmann R, Justesen J, Sarkar SN, Sen GC, Yee VC. (2003). Crystal structure of the 2'-specific and double-stranded RNA-activated interferon-induced antiviral protein 2'-5'-oligoadenylate synthetase. *Mol Cell*, 12(5):1173-1185.

Hartmann R, Norby PL, Martensen PM, Jorgensen P, James MC, Jacobsen C, Moestrup SK, Clemens MJ, Justesen J. (1998a). Activation of 2'-5'

oligoadenylate synthetase by single-stranded and double-stranded RNA aptamers. *J Biol Chem*, 273(6):3236-3246.

Hartmann R, Olsen HS, Widder S, Jorgensen R, Justesen J. (1998b). p59OASL, a 2'-5' oligoadenylate synthetase like protein: a novel human gene related to the 2'-5' oligoadenylate synthetase family. *Nucleic Acids Res*, 26(18):4121-4128.

Hentschel U, Piel J, Degnan SM, Taylor MW. (2012). Genomic insights into the marine sponge microbiome. *Nat Rev Microbiol*, 10(9):641-654.

Hill MS, Hill AL, Lopez J, Peterson KJ, Pomponi S, Diaz MC, Thacker RW, Adamska M, Boury-Esnault N, Cárdenas P, Chaves-Fonnegra A, Danka E, De Laine B, Formica D, Hajdu E, Lobo-Hajdu G, Klontz S, Morrow CC, Patel J, Picton B, Pisani D, Pohlmann D, Redmond NE, Reed J, Richey S, Riesgo A, Rubin E, Russell Z, Rützler K, Sperling EA, di Stefano M, Tarver JE, Collins AG. (2013). Reconstruction of family-level phylogenetic relationships within Demospongiae (Porifera) using nuclear encoded housekeeping genes. *PLoS One*, 8(1):e50437.

Holm L, Sander C. (1995). DNA polymerase beta belongs to an ancient nucleotidyltransferase superfamily. *Trends Biochem Sci*, 20(9):345-347.

Hornung V, Hartmann R, Ablasser A, Hopfner K. (2014). OAS proteins and cGAS: unifying concepts in sensing and responding to cytosolic nucleic acids. *Nat Rev Immunol*, 14(8):521-528.

Hovanessian AG, Justesen J. (2007). The human 2'-5'oligoadenylate synthetase family: unique interferon-inducible enzymes catalyzing 2'-5' instead of 3'-5' phosphodiester bond formation. *Biochimie*, 89(6-7):779-788.

Hovanessian AG, Laurent AG, Chebath J, Galabru J, Robert N, Svab J. (1987). Identification of 69-kd and 100-kd forms of 2-5A synthetase in interferon-treated human cells by specific monoclonal antibodies. *EMBO J*, 6(5):1273-1280.

Ito J, Braithwaite DK. (1991). Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res*, 19(15):4045-4057.

Jeffery NW, Jardine CB, Gregory TR. (2013). A first exploration of genome size diversity in sponges. *Genome*, 56(8):451-456.

Justesen J, Ferbus D, Thang MN. (1980). Elongation mechanism and substrate specificity of 2',5'-oligoadenylate synthetase. *Proc Natl Acad Sci U S A*, 77(8):4618-4622.

Justesen J, Hartmann R, Kjeldgaard NO. (2000). Gene structure and function of the 2'-5'-oligoadenylate synthetase family. *Cell Mol Life Sci*, 57(11):1593-1612.

Kajaste-Rudnitski A, Mashimo T, Frenkiel M, Guénet J, Lucas M, Desprès P. (2006). The 2',5'-oligoadenylate synthetase 1b is a potent inhibitor of West Nile virus replication inside infected cells. *J Biol Chem*,

281(8):4624-4637.

Kakuta S, Shibata S, Iwakura Y. (2002). Genomic structure of the mouse 2',5'-oligoadenylate synthetase gene family. *J Interferon Cytokine Res*, 22(9):981-993.

Kelve M, Kuusksalu A, Lopp A, Reintamm T. (2003). Sponge (2',5')oligoadenylate synthetase activity in the whole sponge organism and in a primary cell culture. *J Biotechnol*, 100(2):177-180.

King N, Westbrook MJ, Young SL, Kuo A, Abedin M, Chapman J, Fairclough S, Hellsten U, Isogai Y, Letunic I, Marr M, Pincus D, Putnam N, Rokas A, Wright KJ, Zuzow R, Dirks W, Good M, Goodstein D, Lemons D, Li W, Lyons JB, Morris A, Nichols S, Richter DJ, Salamov A, Sequencing JGI, Bork P, Lim WA, Manning G, Miller WT, McGinnis W, Shapiro H, Tjian R, Grigoriev IV, Rokhsar D. (2008). The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. *Nature*, 451(7180):783-788.

Kjaer KH, Poulsen JB, Reintamm T, Saby E, Martensen PM, Kelve M, Justesen J. (2009). Evolution of the 2'-5'-oligoadenylate synthetase family in eukaryotes and bacteria. *J Mol Evol*, 69(6):612-624.

Kodym R, Kodym E, Story MD. (2009). 2'-5'-Oligoadenylate synthetase is activated by a specific RNA sequence motif. *Biochem Biophys Res Commun*, 388(2):317-322.

Kranzusch PJ, Lee AS, Berger JM, Doudna JA. (2013). Structure of human cGAS reveals a conserved family of second-messenger enzymes in innate immunity. *Cell Rep*, 3(5):1362-1368.

Kranzusch PJ, Lee ASY, Wilson SC, Solovykh MS, Vance RE, Berger JM, Doudna JA. (2014). Structure-Guided Reprogramming of Human cGAS Dinucleotide Linkage Specificity. *Cell*, 158(5):1011-1021.

Kristiansen H, Gad HH, Eskildsen-Larsen S, Despres P, Hartmann R. (2011). The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. *J Interferon Cytokine Res*, 31(1):41-47.

Kristiansen H, Scherer CA, McVean M, Iadonato SP, Vends S, Thavachelvam K, Steffensen TB, Horan KA, Kuri T, Weber F, Paludan SR, Hartmann R. (2010). Extracellular 2'-5' oligoadenylate synthetase stimulates RNase L-independent antiviral activity: a novel mechanism of virus-induced innate immunity. *J Virol*, 84(22):11898-11904.

Kubota K, Nakahara K, Ohtsuka T, Yoshida S, Kawaguchi J, Fujita Y, Ozeki Y, Hara A, Yoshimura C, Furukawa H, Haruyama H, Ichikawa K, Yamashita M, Matsuoka T, Iijima Y. (2004). Identification of 2'-phosphodiesterase, which plays a role in the 2-5A system regulated by interferon. *J Biol Chem*, 279(36):37832-37841.

Kuchta K, Knizewski L, Wyrwicz LS, Rychlewski L, Ginalski K. (2009).

Comprehensive classification of nucleotidyltransferase fold proteins: identification of novel families and their representatives in human. *Nucleic Acids Res*, 37(22):7701-7714.

Kumar S, Mitnik C, Valente G, Floyd-Smith G. (2000). Expansion and molecular evolution of the interferon-induced 2'-5' oligoadenylate synthetase gene family. *Mol Biol Evol*, 17(5):738-750.

Kuusksalu A, Pihlak A, Müller WE, Kelve M. (1995). The (2'-5')oligoadenylate synthetase is present in the lowest multicellular organisms, the marine sponges. Demonstration of the existence and identification of its reaction products. *Eur J Biochem*, 232(2):351-357.

Kuusksalu A, Subbi J, Pehk T, Reintamm T, Müller WE, Kelve M. (1998). Identification of the reaction products of (2'-5')oligoadenylate synthetase in the marine sponge. *Eur J Biochem*, 257(2):420-426.

Kwak JE, Wickens M. (2007). A family of poly(U) polymerases. *RNA*, 13(6):860-867.

Laport MS, Santos OCS, Muricy G. (2009). Marine sponges: potential sources of new antimicrobial drugs. *Curr Pharm Biotechnol*, 10(1):86-105.

Leys SP, Nichols SA, Adams EDM. (2009). Epithelia and integration in sponges. *Integr Comp Biol*, 49(2):167-177.

Leys SP, Riesgo A. (2012). Epithelia, an evolutionary novelty of metazoans. *J Exp Zool B Mol Dev Evol*, 318(6):438-447.

Li F, Xiong Y, Wang J, Cho HD, Tomita K, Weiner AM, Steitz TA. (2002). Crystal structures of the Bacillus stearothermophilus CCA-adding enzyme and its complexes with ATP or CTP. *Cell*, 111(6):815-824.

Li XL, Blackford JA, Hassel BA. (1998). RNase L mediates the antiviral effect of interferon through a selective reduction in viral RNA during encephalomyocarditis virus infection. *J Virol*, 72(4):2752-2759.

Lim JK, Lisco A, McDermott DH, Huynh L, Ward JM, Johnson B, Johnson H, Pape J, Foster GA, Krysztof D, Follmann D, Stramer SL, Margolis LB, Murphy PM. (2009). Genetic variation in OAS1 is a risk factor for initial infection with West Nile virus in man. *PLoS Pathog*, 5(2):e1000321.

Lin R, Yu H, Chang B, Tang W, Liao C, Lin Y. (2009). Distinct antiviral roles for human 2',5'-oligoadenylate synthetase family members against dengue virus infection. *J Immunol*, 183(12):8035-8043.

Liu S, Sanchez DJ, Aliyari R, Lu S, Cheng G. (2012). Systematic identification of type I and type II interferon-induced antiviral factors. *Proc Natl Acad Sci U S A*, 109(11):4239-4244.

Lopp A, Kuusksalu A, Reintamm T, Müller WEG, Kelve M. (2002). 2',5'-oligoadenylate synthetase from a lower invertebrate, the marine sponge Geodia cydonium, does not need dsRNA for its enzymatic activity.

Biochim Biophys Acta, 1590(1-3):140-149.

Lopp A, Reintamm T, Kuusksalu A, de Rosa S, Kelve M. (2004). Non-traditional enzyme pathways of ATP utilization by the marine sponge Tedania sp. *Proceedings of XI Int. Symp. on Marine Natural Products, Sorrento, Italy*: 65.

Lopp A, Reintamm T, Kuusksalu A, de Rosa S, Kelve M. (2012). A novel endoribonuclease from the marine sponge Tethya aurantium specific to 2',5'-phosphodiester bonds. *Biochimie*, 94(8):1635-1646.

Lopp A, Reintamm T, Kuusksalu A, Tammiste I, Pihlak A, Kelve M. (2010). Natural occurrence of 2',5'-linked heteronucleotides in marine sponges. *Mar Drugs*, 8(2):235-254.

Lunde BM, Magler I, Meinhart A. (2012). Crystal structures of the Cid1 poly (U) polymerase reveal the mechanism for UTP selectivity. *Nucleic Acids Res*, 40(19):9815-9824.

Marié I, Blanco J, Rebouillat D, Hovanessian AG. (1997). 69-kDa and 100-kDa isoforms of interferon-induced (2'-5')oligoadenylate synthetase exhibit differential catalytic parameters. *Eur J Biochem*, 248(2):558-566.

Marié I, Hovanessian AG. (1992). The 69-kDa 2-5A synthetase is composed of two homologous and adjacent functional domains. *J Biol Chem*, 267(14):9933-9939.

Marié I, Rebouillat D, Hovanessian AG. (1999). The expression of both domains of the 69/71 kDa 2',5' oligoadenylate synthetase generates a catalytically active enzyme and mediates an anti-viral response. *Eur J Biochem*, 262(1):155-165.

Marié I, Svab J, Hovanessian AG. (1990a). The binding of the 69- and 100-kD forms of 2',5'-oligoadenylate synthetase to different polynucleotides. *J Interferon Res*, 10(6):571-578.

Marié I, Svab J, Robert N, Galabru J, Hovanessian AG. (1990b). Differential expression and distinct structure of 69- and 100-kDa forms of 2-5A synthetase in human cells treated with interferon. *J Biol Chem*, 265(30):18601-18607.

Marques J, Anwar J, Eskildsen-Larsen S, Rebouillat D, Paludan SR, Sen G, Williams BRG, Hartmann R. (2008). The p59 oligoadenylate synthetase-like protein possesses antiviral activity that requires the C-terminal ubiquitin-like domain. *J Gen Virol*, 89(Pt 11):2767-2772.

Martin G, Doublié S, Keller W. (2008). Determinants of substrate specificity in RNA-dependent nucleotidyl transferases. *Biochim Biophys Acta*, 1779(4):206-216.

Martin G, Keller W. (1996). Mutational analysis of mammalian poly(A) polymerase identifies a region for primer binding and catalytic domain, homologous to the family X polymerases, and to other

nucleotidyltransferases. *EMBO J*, 15(10):2593-2603.

Martin G, Keller W. (2007). RNA-specific ribonucleotidyl transferases. *RNA*, 13(11):1834-1849.

Martin G, Keller W, Doublié S. (2000). Crystal structure of mammalian poly(A) polymerase in complex with an analog of ATP. *EMBO J*, 19(16):4193-4203.

Martin G, Möglich A, Keller W, Doublié S. (2004). Biochemical and structural insights into substrate binding and catalytic mechanism of mammalian poly(A) polymerase. *J Mol Biol*, 341(4):911-925.

Mashimo T, Glaser P, Lucas M, Simon-Chazottes D, Ceccaldi PE, Montagutelli X, Desprès P, Guénet JL. (2003). Structural and functional genomics and evolutionary relationships in the cluster of genes encoding murine 2',5'-oligoadenylate synthetases. *Genomics*, 82(5):537-552.

Mashimo T, Simon-Chazottes D, Guénet J. (2008). Innate resistance to flavivirus infections and the functions of 2'-5' oligoadenylate synthetases. *Curr Top Microbiol Immunol*, 321():85-100.

Meng H, Deo S, Xiong S, Dzananovic E, Donald LJ, van Dijk CW, McKenna SA. (2012). Regulation of the interferon-inducible 2'-5'-oligoadenylate synthetases by adenovirus VA(I) RNA. *J Mol Biol*, 422(5):635-649.

Molinaro RJ, Jha BK, Malathi K, Varambally S, Chinnaiyan AM, Silverman RH. (2006). Selection and cloning of poly(rC)-binding protein 2 and Raf kinase inhibitor protein RNA activators of 2',5'-oligoadenylate synthetase from prostate cancer cells. *Nucleic Acids Res*, 34(22):6684-6695.

Mordechai E, Kon N, Henderson EE, Suhadolnik RJ. (1995). Activation of the interferon-inducible enzymes, 2',5'-oligoadenylate synthetase and PKR by human T-cell leukemia virus type I Rex-response element. *Virology*, 206(2):913-922.

Moroz LL, Kocot KM, Citarella MR, Dosung S, Norekian TP, Povolotskaya IS, Grigorenko AP, Dailey C, Berezikov E, Buckley KM, Ptitsyn A, Reshetov D, Mukherjee K, Moroz TP, Bobkova Y, Yu F, Kapitonov VV, Jurka J, Bobkov YV, Swore JJ, Girardo DO, Fodor A, Gusev F, Sanford R, Bruders R, Kittler E, Mills CE, Rast JP, Derelle R, Solovyev VV, Kondrashov FA, Swalla BJ, Sweedler JV, Rogaev EI, Halanych KM, Kohn AB. (2014). The ctenophore genome and the evolutionary origins of neural systems. *Nature*, 510(7503):109-114.

Müller WE. (1995). Molecular phylogeny of Metazoa (animals): monophyletic origin. *Naturwissenschaften*, 82(7):321-329.

Müller WEG, Böhm M, Grebenjuk VA, Skorokhod A, Müller IM, Gamulin V. (2002). Conservation of the positions of metazoan introns from sponges to humans. *Gene*, 295(2):299-309.

Müller WEG, Müller IM. (2003). Origin of the metazoan immune system:

identification of the molecules and their functions in sponges. *Integr Comp Biol*, 43(2):281-292.

Nielsen C. (2008). Six major steps in animal evolution: are we derived sponge larvae?. *Evol Dev*, 10(2):241-257.

Nosenko T, Schreiber F, Adamska M, Adamski M, Eitel M, Hammel J, Maldonado M, Müller WEG, Nickel M, Schierwater B, Vacelet J, Wiens M, Wörheide G. (2013). Deep metazoan phylogeny: when different genes tell different stories. *Mol Phylogenet Evol*, 67(1):223-233.

Okabe M, Tomita K, Ishitani R, Ishii R, Takeuchi N, Arisaka F, Nureki O, Yokoyama S. (2003). Divergent evolutions of trinucleotide polymerization revealed by an archaeal CCA-adding enzyme structure. *EMBO J*, 22(21):5918-5927.

Osigus H, Eitel M, Bernt M, Donath A, Schierwater B. (2013). Mitogenomics at the base of Metazoa. *Mol Phylogenet Evol*, 69(2):339-351.

Pan B, Xiong Y, Steitz TA. (2010). How the CCA-adding enzyme selects adenine over cytosine at position 76 of tRNA. *Science*, 330(6006):937-940.

Perelygin AA, Lear TL, Zharkikh AA, Brinton MA. (2005). Structure of equine 2'-5'oligoadenylate synthetase (OAS) gene family and FISH mapping of OAS genes to ECA8p15-->p14 and BTA17q24-->q25. *Cytogenet Genome Res*, 111(1):51-56.

Perelygin AA, Zharkikh AA, Scherbik SV, Brinton MA. (2006). The mammalian 2'-5' oligoadenylate synthetase gene family: evidence for concerted evolution of paralogous Oas1 genes in Rodentia and Artiodactyla. *J Mol Evol*, 63(4):562-576.

Peterson KJ, Eernisse DJ. (2001). Animal phylogeny and the ancestry of bilaterians: inferences from morphology and 18S rDNA gene sequences. *Evol Dev*, 3(3):170-205.

Philippe H, Derelle R, Lopez P, Pick K, Borchiellini C, Boury-Esnault N, Vacelet J, Renard E, Houliston E, Quéinnec E, Da Silva C, Wincker P, Le Guyader H, Leys S, Jackson DJ, Schreiber F, Erpenbeck D, Morgenstern B, Wörheide G, Manuel M. (2009). Phylogenomics revives traditional views on deep animal relationships. *Curr Biol*, 19(8):706-712.

Pick KS, Philippe H, Schreiber F, Erpenbeck D, Jackson DJ, Wrede P, Wiens M, Alié A, Morgenstern B, Manuel M, Wörheide G. (2010). Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships. *Mol Biol Evol*, 27(9):1983-1987.

Piel J, Hui D, Wen G, Butzke D, Platzer M, Fusetani N, Matsunaga S. (2004). Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge Theonella swinhoei. *Proc Natl Acad Sci U S A*, 101(46):16222-16227.

Player MR, Torrence PF. (1998). The 2-5A system: modulation of viral and cellular processes through acceleration of RNA degradation. *Pharmacol Ther*, 78(2):55-113.

Rebouillat D, Hovanessian AG. (1999). The human 2',5'-oligoadenylate synthetase family: interferon-induced proteins with unique enzymatic properties. *J Interferon Cytokine Res*, 19(4):295-308.

Rebouillat D, Hovnanian A, David G, Hovanessian AG, Williams BR. (2000). Characterization of the gene encoding the 100-kDa form of human 2',5' oligoadenylate synthetase. *Genomics*, 70(2):232-240.

Rebouillat D, Hovnanian A, Marié I, Hovanessian AG. (1999). The 100-kDa 2',5'-oligoadenylate synthetase catalyzing preferentially the synthesis of dimeric pppA2'p5'A molecules is composed of three homologous domains. *J Biol Chem*, 274(3):1557-1565.

Rebouillat D, Marié I, Hovanessian AG. (1998). Molecular cloning and characterization of two related and interferon-induced 56-kDa and 30-kDa proteins highly similar to 2'-5' oligoadenylate synthetase. *Eur J Biochem*, 257(2):319-330.

Reintamm T, Lopp A, Kuusksalu A, Pehk T, Kelve M. (2003a). ATP N-glycosidase - a novel ATP-converting activity from a marine sponge Axinella polypoides. *Eur J Biochem*, 270(20):4122-4132.

Reintamm T, Lopp A, Kuusksalu A, Subbi J, Kelve M. (2003b). Qualitative and quantitative aspects of 2-5A synthesizing capacity of different marine sponges. *Biomol Eng*, 20(4-6):389-399.

Riesgo A, Farrar N, Windsor PJ, Giribet G, Leys SP. (2014). The Analysis of Eight Transcriptomes from All Poriferan Classes Reveals Surprising Genetic Complexity in Sponges. *Mol Biol Evol*, ():.

Rogozin IB, Aravind L, Koonin EV. (2003). Differential action of natural selection on the N and C-terminal domains of 2'-5' oligoadenylate synthetases and the potential nuclease function of the C-terminal domain. *J Mol Biol*, 326(5):1449-1461.

Rogozin IB, Carmel L, Csuros M, Koonin EV. (2012). Origin and evolution of spliceosomal introns. *Biol Direct*, 7():11.

Rosenblum MG, Cheung L, Kessler D. (1988). Differential activity of the 30-kD and the 100-kD forms of 2'-5'An synthetase induced by recombinant human interferon-alpha and interferon-gamma. *J Interferon Res*, 8(3):275-282.

Ryan JF, Pang K, Schnitzler CE, Nguyen A, Moreland RT, Simmons DK, Koch BJ, Francis WR, Havlak P, , Smith SA, Putnam NH, Haddock SHD, Dunn CW, Wolfsberg TG, Mullikin JC, Martindale MQ, Baxevanis AD. (2013). The genome of the ctenophore Mnemiopsis leidyi and its implications for cell type evolution. *Science*, 342(6164):1242592.

Rysiecki G, Gewert DR, Williams BR. (1989). Constitutive expression of a 2',5'-oligoadenylate synthetase cDNA results in increased antiviral activity and growth suppression. *J Interferon Res*, 9(6):649-657.

Saby E, Justesen J, Kelve M, Uriz MJ. (2009a). In vitro effects of metal pollution on Mediterranean sponges: species-specific inhibition of 2',5'-oligoadenylate synthetase. *Aquat Toxicol*, 94(3):204-210.

Saby E, Poulsen JB, Justesen J, Kelve M, Uriz MJ. (2009b). 2'-phosphodiesterase and 2',5'-oligoadenylate synthetase activities in the lowest metazoans, sponge [porifera]. *Biochimie*, 91(11-12):1531-1534.

Sadler AJ, Williams BRG. (2008). Interferon-inducible antiviral effectors. *Nat Rev Immunol*, 8(7):559-568.

Salzberg S, Hacohen D, David S, Dovrat S, Ahwan S, Gamliel H, Birnbaum M. (1990). Involvement of interferon-system in the regulation of cell growth and differentiation. *Scanning Microsc*, 4(2):479-489.

Salzberg S, Hyman T, Turm H, Kinar Y, Schwartz Y, Nir U, Lejbkowicz F, Huberman E. (1997). Ectopic expression of 2-5A synthetase in myeloid cells induces growth arrest and facilitates the appearance of a myeloid differentiation marker. *Cancer Res*, 57(13):2732-2740.

Samanta H, Dougherty JP, Lengyel P. (1980). Synthesis of (2'-5')(A)n from ATP. Characteristics of the reaction catalyzed by (2'-5')(A)n synthetase purified from mouse Ehrlich ascites tumor cells treated with interferon. *J Biol Chem*, 255(20):9807-9813.

Sarkar SN, Bandyopadhyay S, Ghosh A, Sen GC. (1999a). Enzymatic characteristics of recombinant medium isozyme of 2'-5' oligoadenylate synthetase. *J Biol Chem*, 274(3):1848-1855.

Sarkar SN, Ghosh A, Wang HW, Sung SS, Sen GC. (1999b). The nature of the catalytic domain of 2'-5'-oligoadenylate synthetases. *J Biol Chem*, 274(36):25535-25542.

Sarkar SN, Miyagi M, Crabb JW, Sen GC. (2002). Identification of the substrate-binding sites of 2'-5'-oligoadenylate synthetase. *J Biol Chem*, 277(27):24321-24330.

Schierwater B, Eitel M, Jakob W, Osigus H, Hadrys H, Dellaporta SL, Kolokotronis S, Desalle R. (2009). Concatenated analysis sheds light on early metazoan evolution and fuels a modern "urmetazoon" hypothesis. *PLoS Biol*, 7(1):e20.

Schippers KJ, Sipkema D, Osinga R, Smidt H, Pomponi SA, Martens DE, Wijffels RH. (2012). Cultivation of sponges, sponge cells and symbionts: achievements and future prospects. *Adv Mar Biol*, 62():273-337.

Schröder HC, Natalio F, Wiens M, Tahir MN, Shukoor MI, Tremel W, Belikov SI, Krasko A, Müller WEG. (2008). The 2'-5'-oligoadenylate synthetase in the lowest metazoa: isolation, cloning, expression and functional

activity in the sponge Lubomirskia baicalensis. *Mol Immunol*, 45(4):945-953.

Schröder HC, Wiens M, Kuusksalu A, Kelve M, Müller WEG. (1997). Modulation of 2'-5'oligoadenylate synthetase by environmental stress in the marine sponge Geodia cydonium. *Environmental Toxicology and Chemistry*, 16(7):1403-1409.

Sharp TV, Raine DA, Gewert DR, Joshi B, Jagus R, Clemens MJ. (1999). Activation of the interferon-inducible (2'-5') oligoadenylate synthetase by the Epstein-Barr virus RNA, EBER-1. *Virology*, 257(2):303-313.

Silverman RH. (2007). Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response. *J Virol*, 81(23):12720-12729.

Sperling EA, Peterson KJ, Pisani D. (2009). Phylogenetic-signal dissection of nuclear housekeeping genes supports the paraphyly of sponges and the monophyly of Eumetazoa. *Mol Biol Evol*, 26(10):2261-2274.

Srivastava M, Begovic E, Chapman J, Putnam NH, Hellsten U, Kawashima T, Kuo A, Mitros T, Salamov A, Carpenter ML, Signorovitch AY, Moreno MA, Kamm K, Grimwood J, Schmutz J, Shapiro H, Grigoriev IV, Buss LW, Schierwater B, Dellaporta SL, Rokhsar DS. (2008). The Trichoplax genome and the nature of placozoans. *Nature*, 454(7207):955-960.

Srivastava M, Simakov O, Chapman J, Fahey B, Gauthier MEA, Mitros T, Richards GS, Conaco C, Dacre M, Hellsten U, Larroux C, Putnam NH, Stanke M, Adamska M, Darling A, Degnan SM, Oakley TH, Plachetzki DC, Zhai Y, Adamski M, Calcino A, Cummins SF, Goodstein DM, Harris C, Jackson DJ, Leys SP, Shu S, Woodcroft BJ, Vervoort M, Kosik KS, Manning G, Degnan BM, Rokhsar DS. (2010). The Amphimedon queenslandica genome and the evolution of animal complexity. *Nature*, 466(7307):720-726.

Stevenson AL, Norbury CJ. (2006). The Cid1 family of non-canonical poly(A) polymerases. *Yeast*, 23(13):991-1000.

Tatsumi R, Hamada K, Sekiya S, Wakamatsu M, Namikawa T, Mizutani M, Sokawa Y. (2000). 2',5'-oligoadenylate synthetase gene in chicken: gene structure, distribution of alleles and their expression. *Biochim Biophys Acta*, 1494(3):263-268.

Taylor MW, Radax R, Steger D, Wagner M. (2007). Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev*, 71(2):295-347.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. (1997). The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res*, 25 (24):4876-4882.

Tomita K, Ishitani R, Fukai S, Nureki O. (2006). Complete crystallographic analysis of the dynamics of CCA sequence addition. *Nature*, 443(7114):956-960.

Tomita K, Yamashita S. (2014). Molecular mechanisms of template-independent RNA polymerization by tRNA nucleotidyltransferases. *Front Genet*, 5():36.

Torralba S, Sojat J, Hartmann R. (2008). 2'-5' oligoadenylate synthetase shares active site architecture with the archaeal CCA-adding enzyme. *Cell Mol Life Sci*, 65(16):2613-2620.

Vallmann K, Aas N, Reintamm T, Lopp A, Kuusksalu A, Kelve M. (2011). Expressed 2-5A synthetase genes and pseudogenes in the marine sponge Geodia barretti. *Gene*, 478(1-2):42-49.

Van Soest RWM, Boury-Esnault N, Vacelet J, Dohrmann M, Erpenbeck D, De Voogd NJ, Santodomingo N, Vanhoorne B, Kelly M, Hooper JNA. (2012). Global diversity of sponges (Porifera). *PLoS One*, 7(4):e35105.

Venkatesh B, Kirkness EF, Loh Y, Halpern AL, Lee AP, Johnson J, Dandona N, Viswanathan LD, Tay A, Venter JC, Strausberg RL, Brenner S. (2007). Survey sequencing and comparative analysis of the elephant shark (Callorhinchus milii) genome. *PLoS Biol*, 5(4):e101.

Wang Q, Floyd-Smith G. (1997). The p69/71 2-5A synthetase promoter contains multiple regulatory elements required for interferon-alpha-induced expression. *DNA Cell Biol*, 16(12):1385-1394.

Wang Q, Floyd-Smith G. (1998). Maximal induction of p69 2', 5'-oligoadenylate synthetase in Daudi cells requires cooperation between an ISRE and two IRF-1-like elements. *Gene*, 222(1):83-90.

Webster NS, Taylor MW. (2012). Marine sponges and their microbial symbionts: love and other relationships. *Environ Microbiol*, 14(2):335-346.

Webster NS, Taylor MW, Behnam F, Lücker S, Rattei T, Whalan S, Horn M, Wagner M. (2010). Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ Microbiol*, 12(8):2070-2082.

Wiens M, Kuusksalu A, Kelve M, Müller WE. (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(1-2):12-18.

Witt PL, Marié I, Robert N, Irizarry A, Borden EC, Hovanessian AG. (1993). Isoforms p69 and p100 of 2',5'-oligoadenylate synthetase induced differentially by interferons in vivo and in vitro. *J Interferon Res*, 13(1):17-23.

Witt PL, Spear GT, Helgeson DO, Lindstrom MJ, Smalley RV, Borden EC.

(1990). Basal and interferon-induced 2',5'-oligoadenylate synthetase in human monocytes, lymphocytes, and peritoneal macrophages. *J Interferon Res*, 10(4):393-402.

Wreschner DH, McCauley JW, Skehel JJ, Kerr IM. (1981). Interferon action--sequence specificity of the ppp(A2'p)nA-dependent ribonuclease. *Nature*, 289(5796):414-417.

Xiao TS, Fitzgerald KA. (2013). The cGAS-STING pathway for DNA sensing. *Mol Cell*, 51(2):135-139.

Xiong Y, Li F, Wang J, Weiner AM, Steitz TA. (2003). Crystal structures of an archaeal class I CCA-adding enzyme and its nucleotide complexes. *Mol Cell*, 12(5):1165-1172.

Xiong Y, Steitz TA. (2004). Mechanism of transfer RNA maturation by CCA-adding enzyme without using an oligonucleotide template. *Nature*, 430(7000):640-645.

Yamamoto A, Iwata A, Koh Y, Kawai S, Murayama S, Hamada K, Maekawa S, Ueda S, Sokawa Y. (1998). Two types of chicken 2',5'-oligoadenylate synthetase mRNA derived from alleles at a single locus. *Biochim Biophys Acta*, 1395(2):181-191.

Yamamoto Y, Sono D, Sokawa Y. (2000). Effects of specific mutations in active site motifs of 2',5'-oligoadenylate synthetase on enzymatic activity. *J Interferon Cytokine Res*, 20(3):337-344.

Yamtich J, Sweasy JB. (2010). DNA polymerase family X: function, structure, and cellular roles. *Biochim Biophys Acta*, 1804(5):1136-1150.

Yang Q, Nausch LWM, Martin G, Keller W, Doublié S. (2014). Crystal structure of human poly(A) polymerase gamma reveals a conserved catalytic core for canonical poly(A) polymerases. *J Mol Biol*, 426(1):43-50.

Yue D, Maizels N, Weiner AM. (1996). CCA-adding enzymes and poly(A) polymerases are all members of the same nucleotidyltransferase superfamily: characterization of the CCA-adding enzyme from the archaeal hyperthermophile Sulfolobus shibatae. *RNA*, 2(9):895-908.

ACKNOWLEDGEMENTS

This work was carried out in the Department of Gene Technology at the Tallinn University of Technology. I am glad to have had the chance to work in such a friendly and inspirational environment.

I am grateful to my supervisors Anne Kuusksalu and Annika Lopp for their guidance, encouragement and support not only during my PhD studies but also during the previous stages of my education at the university. I thank Anne for taking me as her apprentice already in my first year of bachelor studies. I appreciate her patience to teach me basic laboratory skills.

I thank Merike Kelve for holding together our friendly and united research group. Her care for us is highly valued. I appreciate the inspirational discussions with Tõnu Reintamm and thank him for the technical assistance in the laboratory.

I am thankful to all the past and present members of our research group for being such wonderful colleagues. I especially value Kerli's friendship during all the years we have worked and studied together.

I thank Lagle Kasak and Karin Valmsen for valuable advice and for sharing their experience in recombinant protein expression. I also thank Karin for pre-reviewing this thesis. I thank Mari Kelve-Liivsoo for proofreading my publications, as well as this thesis, and for making linguistic corrections.

I am thankful to all the coauthors of my publications. I am grateful to Just Justesen for the opportunity to work in his laboratory.

I owe much to my parents and sisters for their support during my studies. I thank Allan and our children, Harri and Marvin, for their love and support. I am greatly indebted to Allan's mother Kersti for taking care of our children whenever I needed.

PUBLICATION I

Päri M, Kuusksalu A, Lopp A, Reintamm T, Justesen J, Kelve M. (2007). Expression and characterization of recombinant 2',5'-oligoadenylate synthetase from the marine sponge *Geodia cydonium*. *FEBS J*, 274(13):3462-3474

the FEBS Journal

Expression and characterization of recombinant 2′,5′-oligoadenylate synthetase from the marine sponge *Geodia cydonium*

Mailis Päri[1], Anne Kuusksalu[2], Annika Lopp[2], Tõnu Reintamm[2], Just Justesen[3] and Merike Kelve[1,2]

1 Department of Gene Technology, Tallinn University of Technology, Estonia
2 Department of Molecular Genetics, National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
3 Department of Molecular Biology, Aarhus University, Denmark

Keywords
Geodia cydonium; marine sponge; oligoadenylates; recombinant 2-5A synthetase; RNA binding

Correspondence
M. Kelve, Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15, Tallinn 12618, Estonia
Fax: +372 6204401
Tel: +372 6204432
E-mail: merike.kelve@ttu.ee

(Received 20 December 2006, revised 7 May 2007, accepted 11 May 2007)

doi:10.1111/j.1742-4658.2007.05878.x

2′,5′-oligoadenylate (2-5A) synthetases are known as components of the interferon-induced cellular defence mechanism in mammals. The existence of 2-5A synthetases in the evolutionarily lowest multicellular animals, the marine sponges, has been demonstrated and the respective candidate genes from *Geodia cydonium* and *Suberites domuncula* have been identified. In the present study, the putative 2-5A synthetase cDNA from *G. cydonium* was expressed in an *Escherichia coli* expression system to characterize the enzymatic activity of the recombinant polypeptide. Our studies reveal that, unlike the porcine recombinant 2-5A synthetase, the sponge recombinant protein associates strongly with RNA from *E. coli*, forming a heterogeneous set of complexes. No complete dissociation of the complex occurs during purification of the recombinant protein and the RNA constituent is partially protected from RNase degradation. We demonstrate that the sponge recombinant 2-5A synthetase in complex with *E. coli* RNA catalyzes the synthesis of 2′,5′-phosphodiester-linked 5′-triphosphorylated oligoadenylates from ATP, although with a low specific activity. Poly(I)·poly(C), an efficient artificial activator of the mammalian 2-5A synthetases, has only a minimal effect (an approximate two-fold increase) on the sponge recombinant 2-5A synthetase/bacterial RNA complex activity.

The 2′,5′-oligoadenylate synthetases (2-5A synthetases; OAS; EC 2.7.7.–) were discovered as a part of the interferon antiviral pathway in mammals [1,2]. In higher animals (vertebrates), when activated by dsRNA, 2-5A synthetases catalyze the polymerization of ATP into unusual 2′,5′-linked oligoadenylates, with the general structure $pppA(2'p5'A)_n$ where $n \geq 1$, commonly abbreviated as 2-5A. 2-5A binds to and activates a latent endoribonuclease, RNase L [3]. Activated RNase L catalyzes the degradation of viral and cellular RNAs, including ribosomal RNA, suppressing protein synthesis and viral growth. Some evidence suggests that 2-5A synthetases are also involved in other cellular processes, such as regulation of cell growth, differentiation, tumorigenesis and apoptosis [4–6].

There are three different size classes of 2-5A synthetases: the small (OAS1), medium (OAS2) and large (OAS3) isoforms, consisting of one, two or three conserved OAS units, respectively [7–12]. Within the classes of 2-5A synthetases, alternative splicing produces multiple isozymes with different C-terminal regions [8]. The 2-5A synthetase family also contains a fourth member, oligoadenylate synthetase-like protein, which is made up of a single OAS unit and two C-terminal ubiquitin-like repeats [13–15].

Abbreviations
2-5A, 2′,5′-oligoadenylate; Ni-NTA, nickel–nitrilotriacetic acid; OAS, 2′,5′-oligoadenylate synthetases; SEC, size exclusion chromatography.

It is known that all vertebrate 2-5A synthetases are expressed as latent proteins and require dsRNA for their activation [16]. However, different from many other dsRNA-binding proteins, 2-5A synthetases are among the few proteins that bind dsRNA without having a dsRNA binding motif [17,18]. As has emerged from studies of the crystal structure of porcine 2-5A synthetase, a distinct positively charged groove on the surface embracing N- and C-terminal domains of the protein mediates dsRNA binding [19].

The 2-5A synthetases, which belong to the DNA polymerase β-like nucleotidyl transferase superfamily, are classified into the same group with CCA-adding enzymes, eukaryotic poly(A) polymerase and TRF4/5 polymerases [20]. The mammalian 2-5A synthetases are highly conserved proteins that share little sequence similarity with nucleotidyl transferases of other families; however, the catalytic domain features of 2-5A synthetases and other polymerases (e.g. DNA polymerase β) are conserved [19,21]. The total fold of a mammalian 2-5A synthetase, porcine OAS1, shows the highest structural similarity with 3′-specific poly(A) polymerase [19]. On the basis of a detailed sequence signature analysis, Rogozin *et al.* [22] proposed that the 2-5A synthetase family has evolved from the more ancient poly(A) polymerase or TRF4/5 families.

In addition to mammals and birds, the 2-5A synthesis has also been found in reptilian tissues but not in amphibians and fish [23]. We have demonstrated the presence of a high 2-5A synthesizing activity in the extracts of a number of marine sponges, the simplest multicellular animals [24,25], and identified the reaction products as authentic 2′,5′-linked oligoadenylates [26]. To date, cDNAs encoding the putative oligoadenylate synthetase have been cloned from two sponges: one from *Geodia cydonium* and two from *Suberites domuncula* [27,28]. By contrast to the high sequence similarity among vertebrate 2-5A synthetase proteins, the *S. domuncula* and *G. cydonium* enzymes share 28% identity and 48% similarity with each other [28]. Moreover, the amino acid sequence deduced from the *G. cydonium* cDNA shares only 18% identity and 39% similarity with the mouse 2-5A synthetase [27]. Despite the low sequence similarity, the motifs known to be essential for the 2-5A synthesizing activity [21] are present in the sponge polypeptides [27,28]. Interestingly, although this enzyme has been found in sponges, in the oldest extant metazoan phylum, it is absent (evidently through gene loss) in some branches of the evolutionary tree of life. Sequence comparison data have not revealed the 2-5A synthetase gene either in insect (*Drosphila melanogaster*), nematode (*Caenorhabditis elegans*), yeast (*Saccharomyces cerevisiae*), plant (*Arabidopsis thaliana*) or fish (*Danio rerio*, *Fugu rubripes*) [8,11,27,28].

With regard to the role of 2-5A synthetase in sponges, the participation of this enzyme in responses to environmental stressors and to bacterial infection has been suggested [28–30]. Whether the 2-5A synthetase in the lowest multicellular animals, similar to the higher Metazoa, is involved in host-defence reactions against viruses remains unknown. To date, the 2-5A synthetase as a single component of the whole mammalian 2-5A/RNase L system has been identified. Considering the long evolutionary distance between sponges and vertebrate lineages, the elucidation of the function of the 2-5A synthetase in these invertebrates, particularly in the innate immune system, would be of considerable interest.

Before the present study was started, only vertebrate 2-5A synthetases had been expressed in heterologous systems for use in detailed studies of the structural and functional properties of the enzyme. In the present study, the putative 2-5A synthetase cDNA from the marine sponge *G. cydonium* (EMBL accession number Y18497) was expressed in a bacterial expression system and the histidine-tagged recombinant protein was purified by affinity chromatography. As previous data have indicated differences in the activation features between the sponge and mammalian enzymes [19,31], the enzyme of invertebrate origin needs to be properly characterized by means of a recombinant protein technique.

Results

Expression and purification of His-tagged proteins

N- and C-terminally hexahistidine tagged constructs of the 2-5A synthetase cDNA from *G. cydonium* were expressed in a bacterial expression system and the recombinant proteins were purified by affinity chromatography on a nickel–nitrilotriacetic acid (Ni-NTA) column. Two different sponge cDNA constructs were chosen for studies investigating whether modification of either the N- or C-terminus of the protein could affect the properties of the enzyme. For comparison, a mammalian recombinant enzyme, C-terminally hexahistidine tagged porcine 2-5A synthetase, was produced under the same conditions.

The sponge and porcine recombinant proteins were expressed as soluble proteins and bound well to the affinity beads. However, the expression level of the C-terminally tagged sponge 2-5A synthetase was much lower than that of the N-terminally tagged protein. The highest expression level was observed in the case

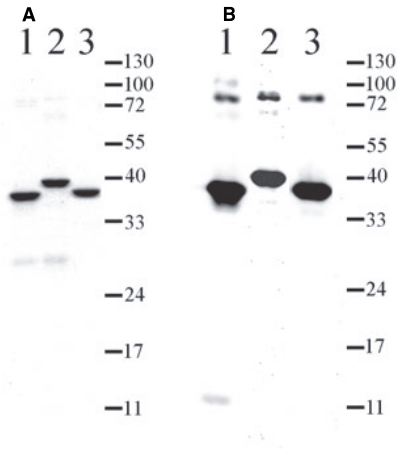


Fig. 1. SDS/PAGE (A) and western blot analysis (B) of the affinity purified C-terminally and N-terminally His-tagged recombinant 2-5A synthetase from *G. cydonium* (lanes 1 and 2, respectively) and C-terminally His-tagged recombinant porcine 2-5A synthetase (lane 3). The amount of the protein loaded to the gel was 1 μg. (A) Gel was stained with Coomassie Blue. (B) Proteins were detected with anti-His serum as described in Experimental procedures.

of the porcine 2-5A synthetase (data not shown). Figure 1 demonstrates the results of the purification of the recombinant proteins. The occurrence of dominant bands of the recombinant proteins provides evidence of a high degree of purification obtained by affinity chromatography. Additionally, some fainter bands of higher and lower molecular weight could be seen in the preparations (Fig. 1A). Bands of higher molecular weight, which were also recognized by anti-His serum (Fig. 1B), may correspond to the aggregates of the recombinant proteins. A faint band of a lower molecular weight (approximately 30 kDa) was visible in the sponge (but not in the porcine) recombinant protein preparations (Fig. 1A). This band was not recognized by monoclonal anti-His serum even under the conditions of the overloaded recombinant protein (Fig. 1B, lanes 1 and 2). Most probably it represents an impurity present in the sponge recombinant 2-5A synthetase preparations.

RNA binding of the sponge recombinant 2-5A synthetase

All known vertebrate 2-5A synthetases are known to be activated by their cofactor, dsRNA. Therefore, we performed activity assays of the purified enzyme preparations by adding poly(I)·poly(C), the synthetic dsRNA, usually used in *in vitro* assays of the enzymatic activity of 2-5A synthetases. As expected, the porcine recombinant protein was practically inactive [specific activity of 0.05 nmol ATP polymerized·(μg protein·h)$^{-1}$] in the absence of poly(I)·poly(C), but its specific activity was increased more than 1000-fold in the presence of the activator (Fig. 2A). Another dsRNA, poly(A)·poly(U), was also capable of activating the porcine enzyme, but to a lesser extent than poly(I)·poly(C) (Fig. 2A).

Surprisingly, the recombinant 2-5A synthetase preparations from *G. cydonium* were able to catalyze the formation of 2-5A oligomers from ATP *per se* and the addition of poly(I)·poly(C) only managed to double

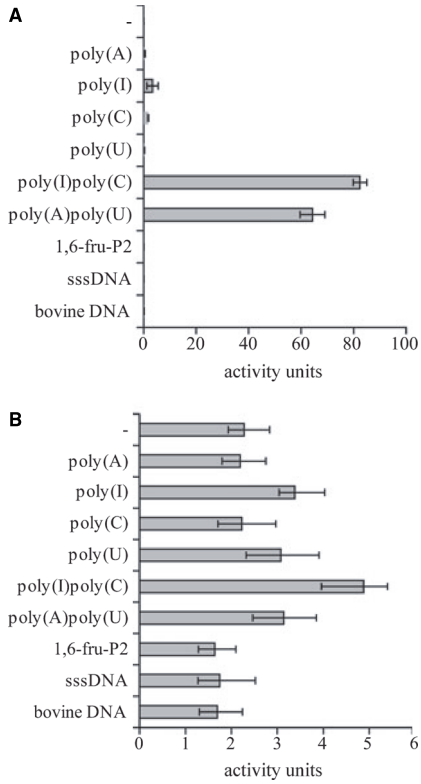


Fig. 2. The effect of various potential activators on the 2-5A synthesizing activity of the recombinant porcine 2-5A synthetase (A) and N-terminally His-tagged recombinant 2-5A synthetase from *G. cydonium* (B) during 1 h of incubation in the presence of 100 μg·mL^{-1} of the indicated substance. The activity units are expressed as nmol ATP polymerized·(μg protein·h)$^{-1}$. Error bars indicate the highest and lowest values of the activity from three independent experiments.

the enzymatic activity (Fig. 2B). No difference was found between N- and C-terminally tagged proteins in that respect. The poly(I)·poly(C) concentration of 0.1 mg·mL^{-1} used in the present study for activation proved to be the most effective one in the studied range of the concentrations (0.001–1 mg·mL^{-1}). The other potential activators, various single-stranded or dsRNAs and DNAs and the only known non-nucleic acid activator of 2-5A synthetases, fructose 1,6-diphosphate [32], also caused small modulations of the existing activity of the sponge recombinant protein (Fig. 2B).

The ability of the sponge protein preparation to catalyze the formation of oligoadenylates *per se* referred to the possibility that the preparation could be contaminated with nucleic acids. Indeed, the UV-spectrum of the recombinant 2-5A synthetase from *G. cydonium*, enriched on the Ni-NTA column and dialyzed thereafter, had a maximum at 260 nm. By contrast, the UV-spectrum of the analogously purified porcine recombinant 2-5A synthetase corresponded to that of a pure protein.

HPLC analysis of a sponge recombinant protein preparation showed that it contained small amounts of four different 2′,3′-cyclic ribonucleotides. The incubation of the preparation at room temperature for longer periods increased the quantities of the cyclic nucleotides (Fig. 3). The relative molar amounts of 2′,3′-cCMP, 2′,3′-cUMP, 2′,3′-cGMP and 2′,3′-cAMP were 1.2 : 1.0 : 1.6 : 1.1, respectively. Also, the total alkaline hydrolysis of the preparation gave similar ratios for the four nucleotides (data not shown). These products could arise from RNA degradation by trace amounts of a nonspecific endoribonuclease of *E. coli*, RNase I [33] which, possibly via binding to RNA, could be copurified with the recombinant protein.

Thus, RNA, obviously copurified in complex with the protein, was present in the sponge recombinant protein preparations.

Based on the amino acid sequence, the calculated pI of the recombinant 2-5A synthetase from *G. cydonium* is 9.6 [27]. Therefore, the protein should be positively charged at neutral pH. However, the analysis of the protein preparation in basic (pH 8.8, for acidic proteins) as well as in acidic (pH 4.5, for basic proteins) native gels showed that the protein was negatively charged and migrated only in basic gel where several distinct bands could be observed (Fig. 4A, lanes 4 and 5). The distinct bands in the gel seen in lanes 4 and 5 could correspond to different complexes of nucleic acid and protein because they were stained with ethidium bromide (Fig. 4B) and recognized by anti-His serum (data not shown). The only exception was the fast moving band in the gel (Fig. 4, fraction X), which was neither stained with ethidium bromide nor recognized by anti-His serum; this band likely represents the same 30 kDa impurity which had been detected by SDS/PAGE analysis (Fig. 1A). The porcine recombinant protein (the calculated pI is 9.05) behaved in a predicted manner, not migrating towards anode in the basic gel (Fig. 4A, lane 6).

For further characterization of the sponge recombinant 2-5A synthetase complex with RNA, size exclusion chromatography (SEC) was performed. As shown in Fig. 5, the absorbance registered at 260 nm was constantly higher than at 280 nm, demonstrating the elution of the RNA component in a wide range of molecular masses. The recombinant protein eluted as a broad peak starting from the column void volume (Fig. 5, inset). This suggests that the protein eluted as a set of heterogeneous complexes containing RNA and evidently more than one polypeptide molecule. A protein of lower molecular weight (approximately 30 kDa) that eluted in later fractions (Fig. 5, fractions 18–20) obviously corresponded to a minor component, which had copurified together with the recombinant protein (Fig. 1A).

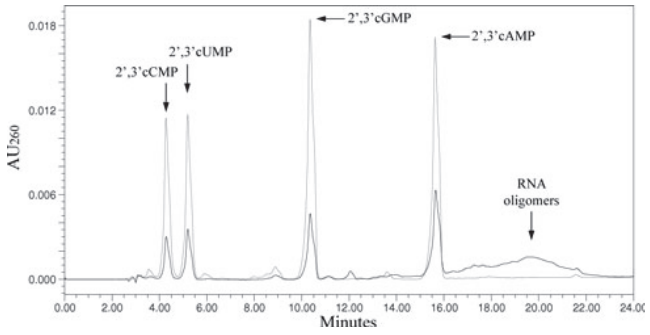


Fig. 3. HPLC chromatogram of a sponge recombinant N-terminally His-tagged 2-5A synthetase preparation (0.8 μg of protein) before (black line) and after (gray line) incubation at room temperature for 99 h.

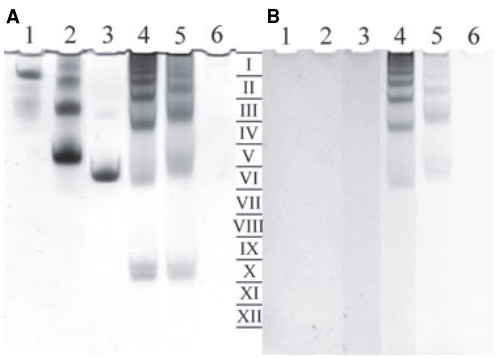


Fig. 4. The basic native polyacrylamide gels stained with Coomassie Blue (A) and EtBr (B). 1, catalase (5 μg); 2, BSA (5 μg); 3, pepsin (5 μg); 4, C-terminally His-tagged recombinant 2-5A synthetase from *G. cydonium* (13 μg); 5, N-terminally His-tagged recombinant 2-5A synthetase from *G. cydonium* (9 μg); 6, porcine recombinant 2-5A synthetase (20 μg).

Figure 6 depicts the calculated specific activities of the recombinant protein of the SEC fractions plotted against the number of nucleotides per protein molecule from the corresponding fractions. Although the accuracy of determining the absolute values of these parameters may be low, the data show an increasing trend in the specific activity depending on the number of nucleotides per protein molecule.

In order to obtain an RNA-free recombinant protein, the enzyme capable of hydrolyzing single-stranded and double-stranded nucleic acids, Benzonase® nuclease (Novagen, Merck KGaA, Darmstadt, Germany), was used. Figure 7 demonstrates the results of the nuclease treatment, which was carried out during the 2-5A activity assays. As can be seen, the added amount of the nuclease effectively inactivated the porcine 2-5A synthetase [by degrading poly(I)·poly(C)] (Fig. 7A), but it had only a modest effect on the 2-5A synthesizing activity of the recombinant protein from *G. cydonium* (Fig. 7B).

The nuclease was also added at different steps of the sponge protein purification: during cell lysis and protein binding as well as during the column washing steps. A less viscous lysate was observed in the presence of the nuclease. Inspection of UV-spectra of the nuclease-treated and untreated preparations revealed that both of them were contaminated with nucleic acids. Calculation of RNA content showed that the nuclease treatment reduced the number of nucleotides per protein molecule from 34 to 23. Thus, the nuclease treatment at this step was of low efficiency. Evidently,

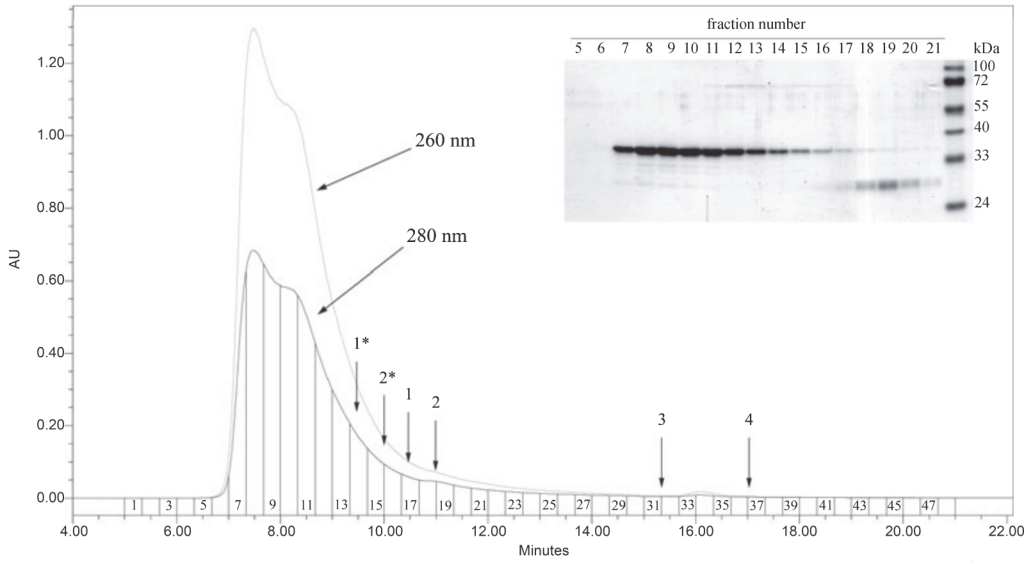


Fig. 5. Fractionation of the C terminally His tagged recombinant 2 5A synthetase preparation by size exclusion chromatography. The collected fractions are shown. The following proteins or substances were used for the calibration of the column: 1, BSA (66.4 kDa); 2, albumin from chicken egg (45.0 kDa); 3, cytochrome *c* (12.5 kDa); 4, tryptophan (0.2 kDa). *Dimer. Inset: SDS/PAGE analysis of fractions collected during SEC of the recombinant protein preparation.

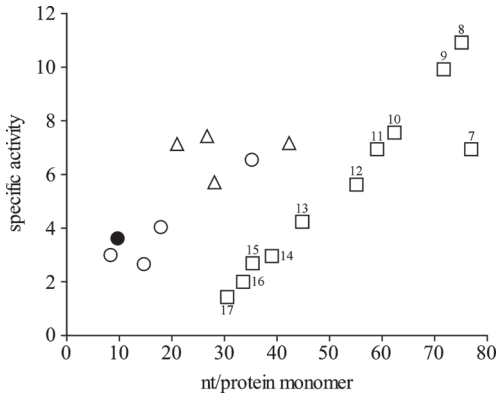


Fig. 6. The relationship between the number of nucleotides per protein monomer and the specific activity of the protein. □, fractions collected during size exclusion chromatography (fraction numbers correspond to those in Fig. 5); ○, different recombinant protein preparations; △, different recombinant protein preparations, where the number of nucleotides was increased by adding 0.1 mg·mL^{-1} poly(I)·poly(C); ●, recombinant protein preparation, where the number of nucleotides was decreased by nuclease treatment. The number of nucleotides per protein monomer was estimated as described in Experimental procedures.

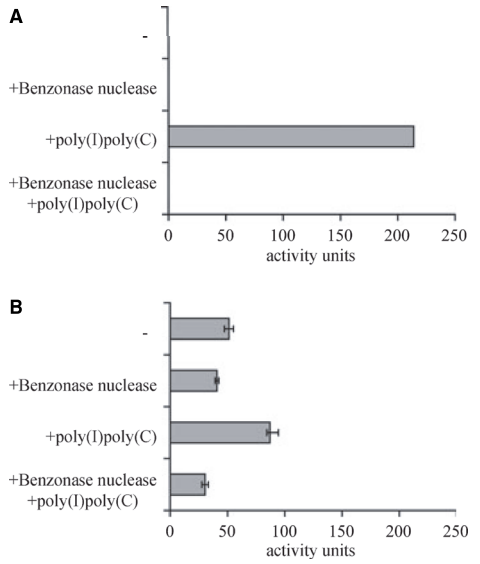


Fig. 7. The effect of the Benzonase nuclease and/or poly(I)·poly(C) on the 2-5A synthesizing activity of the recombinant porcine 2-5A synthetase (A) and N-terminally His-tagged recombinant 2-5A synthetase from *G. cydonium* (B). The products formed from ATP during a 5 h synthesis in the presence or absence of Benzonase nuclease and/or poly(I)·poly(C) were dephosphorylated and analyzed by the HPLC method. The activity units are expressed as nmol ATP polymerized·(μg protein)$^{-1}$. Error bars indicate the highest and lowest values of the activity from three independent experiments.

the conditions of the purification were not optimal for the nuclease (high NaCl and phosphate concentrations and the absence of Mg^{2+}).

Another nuclease treatment was carried out after purification and dialysis of the recombinant protein (i.e. under conditions optimal for the nuclease digestion). The results showed that the addition of the nuclease caused the precipitation of the material in a concentration-dependent manner. Formation of the precipitate in the solution containing the highest amount of the nuclease (0.5 U·μL^{-1}) was visible already after 5 min. Attempts to solubilize the formed pellet by decreasing the pH of the medium, or by adding poly(I)·poly(C), poly(A)·poly(U), ATP, NaCl or combinations of them, were not successful. Finally, the pellet was dissolved in alkaline conditions (pH 10.4), but the UV-spectrum indicated the presence of nucleic acids. The precipitated material was estimated to contain approximately ten nucleotides per polypeptide molecule and it was still enzymatically active (Fig. 6).

In an alternative approach, we tried to modify the purification conditions of the recombinant protein by means of changing pH of the lysis, wash and elution buffers. Finally, protein purification was carried out under conditions in which cell lysis and binding to affinity beads was performed at pH 8.0, but the wash and elution buffers were both alkaline (pH 10.5). In that case, the protein remained soluble and eluted from the affinity column. At this pH value, RNA–protein ionic complexes should dissociate; nevertheless, the UV-spectrum of the resulting protein preparation revealed that nucleic acids (28 nucleotides per protein molecule) were still present. However, in this case, the 2-5A synthesizing activity of the protein was negligible [specific activity of 0.008 nmol ATP·(μg protein·h)$^{-1}$] and the addition of poly(I)·poly(C) did not increase it.

In summary, the sponge 2-5A synthetase expressed in *E. coli* bound some bacterial RNA with high affinity, forming complexes that were partially protected against nuclease degradation of the bound RNA.

Enzymatic characterization of the sponge recombinant protein preparation purified by Ni-NTA chromatography

Searching for optimal conditions for the activity of the affinity purified recombinant enzyme preparation, we

found that they were similar to the conditions of 2-5A activity assays often used for the proteins of this family [34,35]. The increase in specific activity was achieved by rather high ATP (5 mM) and $MgCl_2$ (25 mM) and low salt concentrations (no salt added). In the chosen reaction conditions (see Experimental procedures), the enzyme-RNA complex catalyzed the formation of 2-5A oligomers with the specific activity of approximately 1–10 nmol ATP polymerized·(μg protein·h)$^{-1}$. Variations in the specific activity depended upon the obtained protein batch irrespective of the His-tag localization in the molecule; the specific activity was likely related to the nucleotide content of the preparation (Fig. 6).

The products of the sponge 2-5A synthetase-catalyzed ATP oligomerization assay are presented in Fig. 8. The oligomerization yielded in 2-5A dimer, 2-5A trimer and 2-5A tetramer but, even at high conversion percentages of ATP, the dinucleotide was the main product. Interestingly, in addition to typical 2-5A products, oligomers containing 3′,5′-internucleotide bond (the dimer and minute amounts of the trimer) were identified among reaction products. Also, the products with mixed linkages (i.e. 2′,5′- and 3′,5′-linked trimers) were detected (Fig. 8). All these oligomers were verified by their HPLC retention times, alkaline hydrolysis, RNase T_2 treatment and MALDI-MS analysis.

The ability to catalyze both 2′,5′- and 3′,5′-linked products was also characteristic of the recombinant protein–RNA complexes separated by electrophoresis in native gel (Fig. 4, fractions I–VI) and by size exclusion chromatography (Fig. 5, fractions 7–17).

The products with 3′,5′-linkage have not been described before in enzymatic assays of mammalian 2-5A synthetases. We were able to detect 3′,5′-oligoadenylates (considering the retention time of faint HPLC signals) also in the assays of poly(I)·poly(C) activated porcine recombinant 2-5A synthetase, but the lower limit of the calculated 2-5A/3-5A product ratio was approximately 2000. With regard to the sponge 2-5A synthetase, this ratio was 5.4 ± 0.5 (n = 16) for different recombinant protein batches. The addition of poly(I)·poly(C) to those preparations increased the ratio of 2-5A oligomers to 3-5A oligomers only slightly in favour of 2-5A products. Thus, the sponge recombinant 2-5A synthetase in complex with *E. coli* RNA oligomerized ATP with an apparent loss of isomeric purity of the products.

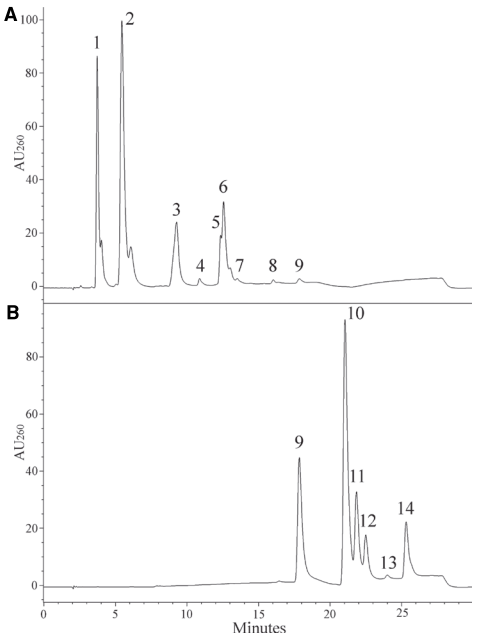


Fig. 8. The product profile of the C-terminally His-tagged recombinant 2-5A synthetase from *G. cydonium*. HPLC chromatograms of products, formed from ATP during a 6 h synthesis, in their phosphorylated (A) or dephosphorylated ('core') (B) forms. In brackets, *m/z* obtained from MALDI-MS analysis are shown. 1, ATP; 2, $p_3A2'p5'A$; 3, $p_3A2'p5'A2'p5'A$; 4, $p_3A2'p5'A2'p5'A2'p5'A$ (*m/z* 1493.5); 5, $p_3A2'p5'A3'p5'A$; 6, $p_3A3'p5'A$; 7, $p_3A3'p5'A2'p5'A$; 8, $p_3A3'p5'A3'p5'A$; 9, adenosine; 10, mixture of A2′p5′A and A2′p5′A2′p5′A2′p5′A; 11, mixture of A2′p5′A2′p5′A and A3′p5′A2′-p5′A (*m/z* 924.6); 12, A2′p5′A3′p5′A (*m/z* 924.7); 13, putative A2′p5′A2′p5′A3′p5′A (*m/z* 1253.9); 14, mixture of A3′p5′A and A3′p5′A3′p5′A (*m/z* 595.4 and 925.4, respectively).

Discussion

His-tagged recombinant proteins of vertebrate 2-5A synthetases produced in *E. coli* and purified by affinity chromatography have been successfully used in the studies of the respective proteins [34,35]. Applying this approach for the production of the first recombinant protein of invertebrate origin, the 2-5A synthetase from the sponge *G. cydonium*, quite unexpected results were obtained. By contrast to analogously produced porcine recombinant 2-5A synthetase, the UV-spectrum of the affinity purified preparation indicated that it was contaminated with nucleic acids. Further, HPLC analysis revealed that the anomalous for a protein UV-spectrum was caused by RNA, which was evidently copurified from the bacterial lysate in complex with the protein. However, such a preparation was able to catalyze oligomerization of ATP into 2′,5′-linked products *per se* and the added dsRNA was unable to improve the activation parameters substan-

tially. These results highlight two significant features: first, the 'putative' 2-5A synthetase cDNA from *G. cydonium* codes for a protein that has oligoadenylate synthetase activity, thus being the 'true' 2-5A synthetase, and, second, the recombinant protein spontaneously forms enzymatically active complexes with heterologous RNA.

Characterization of the preparation by native gel analysis and by size exclusion chromatography demonstrated that the recombinant protein preparation consisted of a set of heterogeneous complexes of RNA and the protein, which did not dissociate under particular separation conditions. Analysis of the size exclusion chromatography fractions showed that the specific activity of the protein was related to the number of bound nucleotides per protein monomer. Generally, the preparations with larger amounts of nucleotides per protein molecule had higher specific activities.

In order to free the recombinant protein preparation from the bound RNA of bacterial origin, nuclease treatments were undertaken under a variety of conditions. The low efficacy of these treatments suggested that RNA in these complexes was not readily accessible to the action of nucleases. On the other hand, the addition of high doses of the nuclease quickly resulted in the protein precipitation. Such a treatment evidently degraded unprotected regions of the RNA in the negatively charged protein–RNA complex and caused its precipitation when the complex became electrically neutral. Thus, an efficient nuclease treatment of the RNA–protein complex resulted in a certain critical point in its precipitation, which was likely related to pI of the complex.

In an alternative approach we tried to obtain an RNA-free protein by using alkaline buffers (pH > 10) in purification procedures. This experiment provided further evidence for the formation of a tight protein–nucleic acid complex, although this complex had lost its 2-5A synthesizing activity. One of the explanations might be that the activation of the recombinant protein could be achieved by RNA containing some alkali-labile minor component (such as dihydrouridine or *N*7-methylguanosine).

Thus, the obtained results suggest that the RNA derived from *E. coli* was bound to the recombinant protein with a high affinity, being partially protected from RNase degradation in these complexes. Besides, our earlier study showed that the 2-5A synthetase activity exhibited by crude extracts of *G. cydonium* depended neither on the addition of exogenous dsRNA, nor on nuclease treatments [31]. Considering the results of the present study, the existence of a strong endogenous nucleic acid–protein complex in the sponge crude extracts can be presumed.

2-5A synthetases, unlike other nucleotidyl transferases, catalyze 2′-5′, not 3′-5′, phosphodiester bond formation between substrates bound to the acceptor and donor sites. The 2′- and 3′-specificities of the enzymes of nucleotidyl transferase superfamily are believed to be achieved through an orientation of the acceptor nucleotide molecule so that the ribose 2′- or 3′-hydroxyl would be in a favourable position to react [19]. Surprisingly, our results demonstrated a low regioselectivity exhibited by the sponge recombinant protein preparation because we identified 3′,5′-linked adenylates as minor reaction products. Although the reason for this phenomenon is unclear, we can speculate that the particular features of different RNA–protein complexes could be involved in determining the unusual product profile of the preparation.

The specific activity of the recombinant protein was rather low, being in the same range as that of a sponge tissue extract per μg of total protein [25]. There are several interpretations for the low activity of the recombinant protein produced in bacteria. The tightly bound bacterial RNA was obviously not a proper activator for the recombinant protein. It is also possible that, despite its ability to bind RNA, most of the polypeptide produced in *E. coli* was in enzymatically inactive conformation. Besides, the bound RNA was of heterogeneous composition and could include inhibitory or poorly activating components.

The RNA binding site for 2′,5′-oligoadenylate synthetases is poorly defined. These enzymes are thought to interact with RNA in a sequence unspecific manner. In addition to dsRNA, the 2-5A synthetases are able to bind to DNA and ssRNA as well, but those polynucleotides have not been shown to activate the enzyme [36]. However, some ssRNA aptamers with little secondary structure, containing only few base-paired regions, activate the 2-5A synthetase as strongly as dsRNA [37]. Recently, the activation of 2-5A synthetase in prostate cancer cells by certain cellular mRNAs was demonstrated [38].

Hartmann *et al.* [19] have demonstrated that the dsRNA binding domain in the porcine OAS1 involves several positively charged residues localized on the surface of the protein. Only two of the five basic residues, which have been shown to be important for dsRNA binding and enzymatic activity in porcine 2-5A synthetase, are conserved in the *G. cydonium* sequence [19]. This may bring about an RNA recognition by the sponge enzyme that differs from that exhibited by vertebrate 2-5A synthetases. Our data

demonstrate a much higher affinity of RNA to the recombinant enzyme from *G. cydonium* than to the porcine one. Moreover, the sponge 2-5A synthetase may need an RNA with special primary and secondary structure elements for its activation. Poly(I)·poly(C) as a synthetic dsRNA may meet these requirements only partially.

Further studies will be required to clarify the structure of the activator of 2-5A synthetases in the sponges as well as the nature of the RNA binding site in this protein molecule. This knowledge would shed light on the function(s) of this ancient form of the enzyme in the multicellular animals that are evolutionarily most distant from humans. The general significance of the study of 2-5A synthetase as one of the key components of the mammalian 2-5A system will be its contribution to our understanding of the evolution of the innate immune system in Metazoa.

Experimental procedures

Expression and purification of the recombinant 2-5A synthetase from *G. cydonium*

N-terminally 6xHis-tagged construct

The coding region of the putative 2-5A synthetase cDNA (EMBL accession number Y18497) was cloned into pQE30 expression vector (Qiagen GmbH, Hilden, Germany). The resulting polypeptide contained additional N-terminal amino acids MRGS**HHHHHH**GSACELGTPIRFYAA KGD, including the hexahistidine affinity tag (in bold) and the anti-RGS-$(His)_4$ antibody (Qiagen) binding site (underlined), relative to the published polypeptide sequence (UniProt accession number O97190).

With some modifications, the QIAExpress™ protocol (Qiagen) for the expression of the histidin-tagged proteins was used. The insert-containing plasmid was transformed into the *E. coli* strain M15 (pREP4) (Qiagen). The transformed bacteria were grown in 2xYT media, containing appropriate antibiotics, on a rotary shaker at 200 r.p.m. at 37 °C until the cell density of $A_{600\ nm} = 0.6$ was reached. Then the expression of recombinant plasmid was induced by adding isopropyl-β-D-thiogalactoside (Sigma, St Louis, MO, USA) at a final concentration of 0.5 mM. After overnight incubation at room temperature, cells were harvested by centrifugation and lysed in lysis buffer (50 mM Na_2HPO_4, pH 8.0, 500 mM NaCl, 10% glycerol, 20 mM imidazole) by sonication on ice. The lysate was clarified by centrifugation and the supernatant was mixed with Ni^{2+}-NTA-agarose beads and rotated at 4 °C for 1 h. The beads were washed with wash buffer (50 mM Na_2HPO_4, pH 8.0, 500 mM NaCl, 10% glycerol, 50 mM imidazole), applied to a column and eluted with elution buffer (50 mM Na_2HPO_4, pH 6.8, 500 mM NaCl, 10% glycerol, 250 mM imidazole) in 0.75–1.5 mL fractions. The fractions were analyzed by 12.5% SDS/PAGE.

In a separate experiment the wash and elution buffers used were alkaline, containing 50 mM $NaHCO_3$, pH 10.5 instead of 50 mM Na_2HPO_4.

C-terminally 6xHis-tagged construct

The bacterial expression vector pET9d (Novagen, Merck, Darmstadt, Germany) containing the *G. cydonium* 2-5A synthetase cDNA with a C-terminal hexahistidine affinity tag was constructed by Signe Eskildsen (University of Aarhus, Denmark). The resulting polypeptide incorporated additional C-terminal amino acids and hexahistidine affinity tag (GS**HHHHHH**) relative to the published polypeptide sequence. Following transformation into BL21 (DE3) *E. coli* cells, the C-terminally tagged recombinant protein was expressed and purified as described above. Both N- and C-terminally tagged recombinant proteins contain an amino acid substitution F32L compared to the published sequence.

Expression and purification of the porcine recombinant 2-5A synthetase

The recombinant BL21 (DE3) *E. coli* bacteria containing the expression vector pET9d with the porcine 2-5A synthetase cDNA were a gift from Rune Hartmann (University of Aarhus, Denmark). The recombinant protein having a C-terminal hexahistidine affinity tag was produced and purified as described above.

SDS/PAGE and western blot analysis

The proteins were separated in 12.5% SDS-polyacrylamide gel [39]. To visualize proteins, the gel was stained with PageBlue™ Protein Staining Solution (Fermentas, Burlington, ON, Canada) and scanned to produce a digital image.

For the Western blot analysis, the separated proteins were transferred to a Hybond C Extra membrane (Amersham, Little Chalfont, UK). The membrane was blocked for 1 h with a solution of 5% (w/v) nonfat dry milk in phosphate-buffered saline (NaCl/Pi), pH 7.4 containing 0.1% (v/v) Tween 20 (NaCl/Pi-Tween). The membrane carrying N-terminally tagged protein was incubated for 1 h with 1 : 5000 (v/v) dilution in NaCl/Pi of mouse anti-[RGS-$(His)_4$] serum (Qiagen). For C-terminally tagged proteins, mouse monoclonal antibody to $(His)_6$ tag (Quattromed, Tartu, Estonia) was used (dilution 1 : 2500, v/v). Then the membranes were incubated for 1 h with 1 : 5000 (v/v) dilution in NaCl/Pi of goat anti-mouse serum $F(ab')_2$ fragment conjugated to HRP (Santa Cruz Biotechnology, Santa Cruz, CA, USA). Between the incubations, the membrane was washed three

times with NaCl/Pi-Tween and, after the last incubation, twice more with NaCl/Pi. The proteins were visualized using ECL method (SuperSignal® West Pico Chemiluminescent Substrate; Pierce, Rockford, IL, USA).

Dialysis of the recombinant 2-5A synthetase

To remove imidazole, fractions containing recombinant protein were pooled and dialysed against buffer A (10 mM Hepes, pH 7.5, 1 mM Mg-acetate, 90 mM KCl, 2 mM β-mercaptoethanol, 10% glycerol). Alternatively, pooled fractions were concentrated and the imidazole containing buffer was exchanged against buffer A or buffer N (20 mM Tris/HCl, pH 7.5, 1 mM Mg-acetate, 20 mM NaCl, 2 mM β-mercaptoethanol, 10% glycerol) using Amicon® Ultra Centrifugal Filter Devices (10 kDa MWCO, Millipore, Bedford, MA, USA).

When alkaline buffers were used for protein purification, the imidazole buffer was exchanged against buffer B (50 mM $NaHCO_3$, pH 10.5, 1 mM Mg-acetate, 20 mM NaCl, 10% glycerol) or buffer N at pH 10.5, adjusted with NaOH.

Nuclease treatments

To ensure a recombinant protein preparation free from nucleic acids, several nuclease treatments during or after purification of the protein were undertaken.

First, for nuclease treatment during protein purification, 12.5 $U \cdot mL^{-1}$ of Benzonase® nuclease (Novagen) were added into the lysis and/or wash buffer.

Second, for nuclease treatment in the 2-5A synthetase activity assay, 0.2 $U \cdot \mu L^{-1}$ of the Benzonase® nuclease were added to the reaction mixture.

Finally, for nuclease treatment after protein purification, 200 μL of the dialyzed protein solution in buffer N (optimal conditions for the nuclease) were incubated at room temperature in the presence of 0, 0.005, 0.05 or 0.5 $U \cdot \mu L^{-1}$ of the Benzonase® nuclease for different time periods. The formation of the precipitate was monitored visually. After formation of the precipitate, the protein suspension was centrifuged at 2300 ***g*** using an Eppendorf centrifuge 5415D, rotor F-45-24-11 (Eppendorf AG, Hamburg, Germany) at room temperature for 1 min. The pellet was washed several times with buffer N and dissolved in buffer N containing approximately 3.7 mM NaOH (final pH 10.4). The protein suspension, as well as the supernatant and dissolved protein solution, was tested for its 2-5A synthesizing activity.

2-5A synthetase activity assay

Under optimized conditions, 2-5A synthetase activity was assayed by incubating the recombinant protein in the reaction mixture containing 20 mM Tris/HCl, pH 8.0, 25 mM $MgCl_2$ and 5 mM ATP as a substrate, in a final volume of 50 μL, at 37 °C for different time periods. The reaction was stopped by heating at 95 °C for 5 min and centrifuged at 16 000 ***g*** for 5 min using an Eppendorf 5415D. (In some experiments, varying concentrations of poly(I)·poly(C), poly(A)·poly(U), poly(I), poly(C), poly(U), D-fructose 1,6-diphosphate, bovine high molecular weight DNA, sonicated DNA from salmon sperm (all from Sigma), poly(A) (Reanal, Budapest, Hungary) and/or Benzonase® nuclease were added to the reaction mixture.

The analysis of reaction products was performed as previously described [31]. Briefly, the reaction products were subjected to a C_{18} reverse-phase column (Supelcosil™ LC-18, 250 × 4.6 mm, 5 μm, Supelco, Bellefonte, PN, USA) at 40 °C. Eluent A was 50 mM ammonium phosphate pH 7.0 and eluent B was 50% methanol in water. The products were separated and analysed in a linear gradient of eluent B (0–40%, 20 min); the column was equilibriated with eluent A before the next injection (10 min). The absorption was measured at 260 nm. The retention times of ATP, adenosine and oligoadenylates, in either their phosphorylated or dephosphorylated ('core') forms were estimated by comparing them with those of authentic compounds. The quantification of the products was performed by measuring the relative peak areas (Millenium^{32}, version 3.05 software, Waters Corporation, Milford, MA, USA). The 2-5A synthesizing activity was expressed as a specific activity [nmol ATP polymerized·(μg protein·h)$^{-1}$].

For dephosphorylation of the products, the reaction mixture was treated with shrimp alkaline phosphatase (SAP, Fermentas). SAP in a final concentration of 0.04 $U \cdot \mu L^{-1}$ was added to the reaction mixture and incubated at 37 °C for 1 h.

Identification of the reaction products

RNase T_2 treatment

The fractions corresponding to the individual peaks were collected from the HPLC outlet and treated with 0.4–1.6 units of RNase T_2 (Invitrogen, Carlsbad, CA, USA) overnight at 37 °C. The reaction was stopped by heating at 95 °C for 5 min and the products were analyzed by HPLC as described above.

Alkaline hydrolysis

HPLC fractions were treated with 0.3 M NaOH at 95 °C for 10 min. After neutralization, the products were analyzed by HPLC.

MALDI-MS analysis

HPLC fractions were directly subjected to mass spectrometric analysis. The analysis was carried out with a matrix-assisted

laser-desorption/ionization time-of-flight (MALDI-TOF) mass spectrometer, as previously described [25].

Native polyacrylamide gel electrophoresis

The acidic native gels were composed of 10% acrylamide:bis-acrylamide (39 : 1), 80 mM β-alanine, 40 mM acetic acid, pH 4.4 and 12.5% glycerol. The gels were polymerized with 0.075% *N*,*N*,*N*′,*N*′-tetramethylethylene diamine and 0.3% ammonium persulfate. The running buffer was 80 mM β-alanine, 40 mM acetic acid, pH 4.4. The gels were run at 20 mA for 1 h.

The basic native gels were composed of 10% acrylamide:bis-acrylamide (39 : 1), 0.375 M Tris/HCl, pH 8.8 and 12.5% glycerol. The gels were polymerized with 0.025% *N*,*N*,*N*′,*N*′-tetramethylethylene diamine and 0.15% ammonium persulfate. The protein samples were mixed with appropriate amounts of 5 × sample buffer (50% glycerol, 0.15% bromophenol blue) and loaded to the gel. The gels were run in Tris-glycine buffer (pH 8.3) at the constant current of 20 mA for 1–1.5 h.

The gels were stained with PageBlue™ Protein Staining Solution (Fermentas). For visualizing nucleic acids, the gels were soaked in 1 μg·mL^{-1} EtBr solution for few minutes.

The basic gel was cut to 0.5 cm strips and the enzymatic activity assays were performed as described.

Size exclusion chromatography of the recombinant protein preparation

Size exclusion chromatography was performed using the HPLC system and software described above. The recombinant protein preparation was loaded onto a SEC column (BioSep-SEC-S3000, 300 × 7.8 mm, 5 μm, Phenomenex, Torrance, CA, USA) at room temperature and the elution was performed with buffer N at a flow rate of 0.75 mL·min^{-1} for 40 min. The column was calibrated with 100 μg of each of the following substances: bovine serum albumin, chicken egg albumin, cytochrome *c* and tryptophan. After washing the column with buffer N, 48 fractions (250 μL each) were collected. The collected fractions were analyzed in 12.5% SDS-polyacrylamide gel and tested for their enzymatic activity.

Estimation of protein and RNA concentration in recombinant protein preparation

Protein concentrations in recombinant protein preparations were measured by a modified Bradford method [40].

Protein concentrations in size exclusion chromatography fractions were estimated by the absorbances at 260 and 280 nm using the formula C_p (mg·mL^{-1}) = $1.55A_{280} - 0.76A_{260}$ [41]. RNA concentration in the protein preparation was estimated using the formula C_{RNA} (mg·mL^{-1}) = $(A_{260} - 0.5\ C_p)$·0.04. The molar concentration of nucleotides was calculated by dividing the RNA concentration C_{RNA} (mg·mL^{-1}) by the average nucleotide molecular weight of 339.5 g·mol^{-1}.

The number of nucleotides per protein molecule was calculated by dividing the molar concentration of nucleotides by the molar concentration of protein in the preparation.

Acknowledgements

We are grateful to J. Subbi from the National Institute of Chemical Physics and Biophysics, Tallinn, Estonia, for performing MALDI-MS experiments. This work was supported by the European Comission (project COOP-CT-2005, contract number 017800) and the Estonian Science Foundation (grant number 5932).

References

1 Kerr IM & Brown RE (1978) pppA2′p5′A2′p5′A: an inhibitor of protein synthesis synthesized with an enzyme fraction from interferon-treated cells. *Proc Natl Acad Sci USA* **75**, 256–260.
2 Sarkar SN & Sen GC (2004) Novel functions of proteins encoded by viral stress-inducible genes. *Pharmacol Ther* **103**, 245–259.
3 Dong B & Silverman RH (1995) 2-5A-dependent RNase molecules dimerize during activation by 2-5A. *J Biol Chem* **270**, 4133–4137.
4 Player MR & Torrence PF (1998) The 2-5A system: modulation of viral and cellular processes through acceleration of RNA degradation. *Pharmacol Ther* **78**, 55–113.
5 Ghosh A, Sarkar SN & Sen GC (2000) Cell growth regulatory and antiviral effects of the P69 isozyme of 2-5 (A) synthetase. *Virology* **266**, 319–328.
6 Chawla-Sarkar M, Lindner DJ, Liu YF, Williams BR, Sen GC, Silverman RH & Borden EC (2003) Apoptosis and interferons: role of interferon-stimulated genes as mediators of apoptosis. *Apoptosis* **8**, 237–249.
7 Rebouillat D & Hovanessian AG (1999) The human 2′,5′-oligoadenylate synthetase family: interferon-induced proteins with unique enzymatic properties. *J Interferon Cytokine Res* **19**, 295–308.
8 Justesen J, Hartmann R & Kjeldgaard NO (2000) Gene structure and function of the 2′-5′-oligoadenylate synthetase family. *Cell Mol Life Sci* **57**, 1593–1612.
9 Kakuta S, Shibata S & Iwakura Y (2002) Genomic structure of the mouse 2′,5′-oligoadenylate synthetase gene family. *J Interferon Cytokine Res* **22**, 981–993.
10 Eskildsen S, Hartmann R, Kjeldgaard NO & Justesen J (2002) Gene structure of the murine 2′-5′-oligoadenylate synthetase family. *Cell Mol Life Sci* **59**, 1212–1222.

11 Mashimo T, Glaser P, Lucas M, Simon-Chazottes D, Ceccaldi PE, Montagutelli X, Despres P & Guenet JL (2003) Structural and functional genomics and evolutionary relationships in the cluster of genes encoding murine 2′,5′-oligoadenylate synthetases. *Genomics* **82**, 537–552.
12 Perelygin AA, Zharkikh AA, Scherbik SV & Brinton MA (2006) The mammalian 2′-5′ oligoadenylate sythetase gene family: evidence for concerted evolution of paralogous Oas1 genes in Rodentia and Artiodactyla. *J Mol Evol* **63**, 562–576.
13 Hartmann R, Olsen HS, Widder S, Jorgensen R & Justesen J (1998) p59OASL, a 2′-5′ oligoadenylate synthetase like protein: a novel human gene related to the 2′-5′ oligoadenylate synthetase family. *Nucleic Acids Res* **26**, 4121–4128.
14 Rebouillat D, Marie I & Hovanessian AG (1998) Molecular cloning and characterization of two related and interferon-induced 56-kDa and 30-kDa proteins highly similar to 2′-5′ oligoadenylate synthetase. *Eur J Biochem* **257**, 319–330.
15 Eskildsen S, Justesen J, Schierup MH & Hartmann R (2003) Characterization of the 2′-5′-oligoadenylate synthetase ubiquitin-like family. *Nucleic Acids Res* **31**, 3166–3173.
16 Minks MA, West DK, Benvin S & Baglioni C (1979) Structural requirements of double-stranded RNA for the activation of 2′,5′-oligo (A) polymerase and protein kinase of interferon-treated HeLa cells. *J Biol Chem* **254**, 10180–10183.
17 Bandyopadhyay S, Ghosh A, Sarkar SN & Sen GC (1998) Production and purification of recombinant 2′-5′ oligoadenylate synthetase and its mutants using the baculovirus system. *Biochemistry* **37**, 3824–3830.
18 Fierro-Monti I & Mathews MB (2000) Proteins binding to duplexed RNA: one motif, multiple functions. *Trends Biochem Sci* **25**, 241–246.
19 Hartmann R, Justesen J, Sarkar SN, Sen GC & Yee VC (2003) Crystal structure of the 2′-specific and double-stranded RNA-activated interferon-induced antiviral protein 2′-5′-oligoadenylate synthetase. *Mol Cell* **12**, 1173–1185.
20 Aravind L & Koonin EV (1999) DNA polymerase beta-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. *Nucleic Acids Res* **27**, 1609–1618.
21 Sarkar SN, Ghosh A, Wang HW, Sung SS & Sen GC (1999) The nature of the catalytic domain of 2′-5′-oligoadenylate synthetases. *J Biol Chem* **274**, 25535–25542.
22 Rogozin IB, Aravind L & Koonin EV (2003) Differential action of natural selection on the N and C-terminal domains of 2′-5′ oligoadenylate synthetases and the potential nuclease function of the C-terminal domain. *J Mol Biol* **326**, 1449–1461.
23 Cayley PJ, White RF, Antoniw JF, Walesby NJ & Kerr IM (1982) Distribution of the ppp (A2′p)nA-binding protein and interferon-related enzymes in animals, plants, and lower organisms. *Biochem Biophys Res Commun* **108**, 1243–1250.
24 Kuusksalu A, Pihlak A, Muller WE & Kelve M (1995) The (2′-5′) oligoadenylate synthetase is present in the lowest multicellular organisms, the marine sponges. Demonstration of the existence and identification of its reaction products. *Eur J Biochem* **232**, 351–357.
25 Reintamm T, Lopp A, Kuusksalu A, Subbi J & Kelve M (2003) Qualitative and quantitative aspects of 2-5A synthesizing capacity of different marine sponges. *Biomol Eng* **20**, 389–399.
26 Kuusksalu A, Subbi J, Pehk T, Reintamm T, Muller WE & Kelve M (1998) Identification of the reaction products of (2′-5′) oligoadenylate synthetase in the marine sponge. *Eur J Biochem* **257**, 420–426.
27 Wiens M, Kuusksalu A, Kelve M & Muller WE (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.
28 Grebenjuk VA, Kuusksalu A, Kelve M, Schutze J, Schroder HC & Muller WE (2002) Induction of (2′-5′) oligoadenylate synthetase in the marine sponges *Suberites domuncula* and *Geodia cydonium* by the bacterial endotoxin lipopolysaccharide. *Eur J Biochem* **269**, 1382–1392.
29 Schröder HC, Wiens M, Kuusksalu A, Kelve M & Müller WEG (1997) Modulation of 2′-5′ oligoadenylate synthetase by environmental stress in marine sponge *Geodia cydonium*. *Environ Toxicol Chem* **16**, 1403–1409.
30 Müller WEG & Müller IM (2003) Origin of the metazoan immune system: identification of the molecules and their functions in sponges. *Integr Comp Biol* **43**, 281–292.
31 Lopp A, Kuusksalu A, Reintamm T, Muller WE & Kelve M (2002) 2′,5′-oligoadenylate synthetase from a lower invertebrate, the marine sponge *Geodia cydonium*, does not need dsRNA for its enzymatic activity. *Biochim Biophys Acta* **1590**, 140–149.
32 Suhadolnik RJ, Li SW, Sobol RW & Varnum JM (1990) 2′,5′-A synthetase: allosteric activation by fructose 1,6-bisphosphate. *Biochem Biophys Res Commun* **169**, 1198–1203.
33 Cannistraro VJ & Kennell D (1991) RNase I*, a form of RNase I, and mRNA degradation in *Escherichia coli*. *J Bacteriol* **173**, 4653–4659.
34 Desai SY, Patel RC, Sen GC, Malhotra P, Ghadge GD & Thimmapaya B (1995) Activation of interferon-inducible 2′-5′ oligoadenylate synthetase by adenoviral VAI RNA. *J Biol Chem* **270**, 3454–3461.
35 Sarkar SN & Sen GC (1998) Production, purification, and characterization of recombinant 2′,5′-oligoadenylate synthetases. *Methods* **15**, 233–242.

36 Marie I, Svab J & Hovanessian AG (1990) The binding of the 69- and 100-kD forms of 2′,5′-oligoadenylate synthetase to different polynucleotides. *J Interferon Res* **10**, 571–578.
37 Hartmann R, Norby PL, Martensen PM, Jorgensen P, James MC, Jacobsen C, Moestrup SK, Clemens MJ & Justesen J (1998) Activation of 2′-5′ oligoadenylate synthetase by single-stranded and double-stranded RNA aptamers. *J Biol Chem* **273**, 3236–3246.
38 Molinaro RJ, Jha BK, Malathi K, Varambally S, Chinnaiyan AM & Silverman RH (2006) Selection and cloning of poly (rC) -binding protein 2 and Raf kinase inhibitor protein RNA activators of 2′,5′-oligoadenylate synthetase from prostate cancer cells. *Nucleic Acids Res* **34**, 6684–6695.
39 Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
40 Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**, 248–254.
41 Warburg O & Christian W (1941) Isolierung und Kristallisation des Gärungsferments Enolase. *Biochem Z* **310**, 384–421.

PUBLICATION II

Reintamm T, Kuusksalu A, Metsis M, **Päri M**, Vallmann K, Lopp A, Justesen J, Kelve M. (2008). Sponge OAS has a distinct genomic structure within the 2-5A synthetase family. *Mol Genet Genomics*, 280(5):453-466

Mol Genet Genomics (2008) 280:453–466
DOI 10.1007/s00438-008-0379-5

ORIGINAL PAPER

Sponge OAS has a distinct genomic structure within the 2-5A synthetase family

Tõnu Reintamm · Anne Kuusksalu · Madis Metsis · Mailis Päri · Kerli Vallmann · Annika Lopp · Just Justesen · Merike Kelve

Received: 20 June 2008 / Accepted: 31 August 2008 / Published online: 17 September 2008

Abstract 2′,5′-Oligoadenylate synthetases (2-5A synthetases, OAS) are enzymes that play an important role in the interferon-induced antiviral defense mechanisms in mammals. Sponges, the evolutionarily lowest multicellular animals, also possess OAS; however, their function is presently unclear. Low homology between primary structures of 2-5A synthetases from vertebrates and sponges renders their evolutionary relationship obscure. The genomic structure of vertebrate OASs has been thoroughly examined, making it possible to elucidate molecular evolution and expansion of this gene family. Until now, no OAS gene structure was available from sponges to compare it with the corresponding genes from higher organisms. In the present work, we determined the exon/intron structure of the OAS gene from the marine sponge *Geodia cydonium* and found it to be completely different from the strictly conserved exon/intron pattern of the OAS genes from vertebrates. This finding was corroborated by the analysis of OAS genes from another sponge, *Amphimedon queenslandica*, whose genome was recently sequenced. Our data suggest that vertebrate and sponge OAS genes have no direct common intron-containing ancestor and two (sub)types of OAS may be discriminated. This study opens new perspectives for understanding the phylogenesis and evolution of 2-5A synthetases as well as functional aspects of this multigene family.

Keywords 2′,5′-Oligoadenylate synthetase · OAS · Genomic structure · Sponge · *Geodia cydonium* · *Amphimedon queenslandica*

Abbreviations The nomenclature of particular OASs in the present paper is based on the level of their sequence similarities. OAS1 refers to the OAS consisting of a single OAS domain. Capital letters (OAS1X) are used to differentiate between OAS1 types with sequence homologies of less than 50%; their variants are additionally marked in small letters (OASXx, sequence homology ∼70 to ∼95%). Labels prime and double prime denote sponge OAS1Xx gene haplotypes (sequence homology close to 100%, a few amino acid substitutions).

Communicated by S. Hohmann.

Electronic supplementary material The online version of this article (doi:10.1007/s00438-008-0379-5) contains supplementary material, which is available to authorized users.

T. Reintamm · A. Kuusksalu · M. Metsis · M. Päri · K. Vallmann · A. Lopp · M. Kelve (✉)
Department of Gene Technology,
Tallinn University of Technology,
Akadeemia tee 15, 12618 Tallinn, Estonia
e-mail: merike.kelve@ttu.ee

M. Metsis
Centre for Biology of Integrated Systems,
Tallinn University of Technology,
Akadeemia tee 15a, 12618 Tallinn, Estonia

J. Justesen
Department of Molecular Biology,
Aarhus University, 8000 Århus C, Denmark

Introduction

2′,5′-Oligoadenylate synthetases (2-5A synthetases; OAS; EC2.7.7.-) belong to the ancient group of nucleotidyltransferases (Holm and Sander 1995; Aravind and Koonin 1999; Martin and Keller 2007), being unique in catalyzing the formation of 2′,5′-phosphodiester linkage instead of the usual 3′,5′-linkage. The 2-5A synthetases comprise a distinct group within the DNA polymerase β-like nucleotidyltransferase superfamily and share little sequence similarity with enzymes of other families. On the basis of a detailed

sequence signature analysis, Rogozin et al. (2003) proposed that the 2-5A synthetase family has evolved from the more ancient poly(A)polymerase or TRF4/5 families.

The 2′,5′-oligoadenylate synthetases were discovered as a part of the interferon antiviral pathway in mammals (rev. by Hovanessian 2007). When activated by double-stranded RNA (dsRNA), the enzymes catalyze the polymerization of ATP into 2′,5′-linked oligoadenylates (2-5A) with the general formula $pppA(2'p5'A)_n$, where $n \geq 1$. The only known function of 2-5A in higher vertebrates is to activate latent ribonuclease (RNase L) that degrades single-stranded RNA of viral and cellular origin. These enzymes together with 2′-phosphodiesterase, which degrades 2-5A, constitute an RNA degradation pathway referred to as the 2-5A system. The 2-5A system in mammals has been extensively characterized (rev. by Player and Torrence 1998). Though described initially as important components of the antiviral action of interferon, 2-5A synthetases have now been demonstrated to also be involved in other cellular processes, such as regulation of cell growth, differentiation, pre-mRNA splicing, tumorigenesis and apoptosis (Sperling et al. 1991; Salzberg et al. 1997; Ghosh et al. 2000; Chawla-Sarkar et al. 2003).

Studies on OAS gene distribution among metazoans have shown their wide occurrence in vertebrates; however, in certain vertebrate lineages (e. g., in some teleost fishes and in the amphibian *Xenopus tropicalis*) they were not found (evidently lost) (Venkatesh et al. 2007). Besides, the genomic analysis of the model organisms has not revealed the presence of 2-5A synthetase either in the insect (*Drosophila melanogaster*), nematode (*Caenorhabditis elegans*), yeast (*Saccharomyces cerevisiae*) or the plant (*Arabidopsis thaliana*).

We have demonstrated the presence of OAS in the evolutionarily lowest multicellular animals, sponges (Kuusksalu et al. 1995, 1998). As sponges reside close to the metazoan stem, their genomic analysis will give us an insight into the ancient genomic conditions from which metazoan emerged. This will also promote our understanding of the evolution of particular genes of interest. Up to now, protein primary structures of four 2-5A synthetases from the following sponge species are available: *Geodia cydonium* (Wiens et al. 1999), *Suberites domuncula* (Grebenjuk et al. 2002), *Haliclona panicea* (Müller and Müller 2003) and *Lubomirskia baicalensis* (Schröder et al. 2008). The constructed phylogenetic tree reveals that the known sponge OAS sequences fall into the same branch; vertebrate OASs are more distantly related (Schröder et al. 2008). All these proteins from sponges consist of a single OAS domain. The overall homology of sponge OAS proteins compared to those from vertebrates is weak, indeed. For example, the amino acid sequence of the OAS from *G. cydonium* is 18% identical to the mouse OAS1a/L3 isoform and 17% identical to the chicken OAS (Wiens et al. 1999). Considering the low homology between primary structures and the long evolutionary distance between vertebrates and sponges, the detailed evolutionary and functional relationships of OASs from these metazoans remain obscure.

It would be of considerable interest to elucidate the function of 2-5A synthetases in sponges, having in mind either the possible role of these enzymes in the innate immune system, particularly in the antiviral protection system of these invertebrates (Schröder et al. 2008) or other roles that might be connected with cell growth or gene regulation. Our previous studies have revealed significant differences in the enzymatic properties of the enzymes from sponges and vertebrates. Differently from higher animals, the OAS from sponges acts in a dsRNA-independent manner in vitro and the enzymatic activity is resistant to nuclease treatments (Lopp et al. 2002; Reintamm et al. 2003). When expressed in *E. coli*, the recombinant OAS from *G. cydonium* has high affinity RNA binding properties; however, the structural requirements for the RNA constituent of the enzyme/nucleic acid complexes remain to be established (Päri et al. 2007). Further studies for the characterization of the sponge OAS are necessary to understand the function of the 2-5A synthesis in these lower metazoans.

The repertoire of OAS genes in vertebrates is surprisingly diverse (Perelygin et al. 2006). However, the members of the 2-5A synthetase gene family from higher animals (from human to chicken) share a common build-up, consisting of a basic exon/intron unit of five translated exons. This unit is found once in OAS1 and OAS-like genes, repeated twice in OAS2 and three times in OAS3 and followed by one or two isozyme-specific terminal exon(s) (rev. in Justesen et al. 2000; Eskildsen et al. 2002). The positions of introns in the homologous regions of the OAS genes are strictly conserved between the vertebrates analyzed so far (Perelygin et al. 2006). Both the human and mouse OAS genes are clustered in a single locus, whereas the OAS-like gene(s) is (are) located in a close but separate locus (Justesen et al. 2000; Eskildsen et al. 2002).

In the current study, we present the genomic structure of the OAS from the evolutionarily lowest metazoa, a sponge, for the first time. The presence of at least two OAS gene versions, as well as the tandem arrangement of OAS genes in *G. cydonium* is demonstrated. The coding sequence of the *G. cydonium* OAS gene is interrupted by seven comparatively short introns and their positioning substantially differs from the pattern found in vertebrates. The exon/intron structure of the OAS gene from *G. cydonium* was confirmed by the analysis of genomic data of another sponge, *Amphimedon queenslandica*, whose genome was recently sequenced. Unlike in vertebrates, the OAS genes from *A. queenslandica* do not form a single contiguous cluster but are grouped in separate loci containing one or

two tandemly arranged OAS genes. Taken together, our results define sponge OAS as a clearly distinct subgroup of the OAS family. These data suggest that the last common ancestor of vertebrates and sponges already had several, most probably intron-containing, OAS genes which were further subjected to independent evolutionary scenarios.

Materials and methods

Sponges

The specimens of *Geodia cydonium* Jameson 1811 (Porifera, Demospongiae, Astrophorida, Geodiidae) were collected in two locations of the Mediterranean: the Tyrrhenian Sea (Napoli Bay, individual 1) and the Adriatic Sea (Rovinj, individual 2) (gifts from S. de Rosa and W. E. G. Müller, respectively). The animal samples were cut into pieces, frozen in liquid nitrogen and kept at −70°C until their further use.

DNA isolation

One hundred to two hundred milligrams of the tissue was homogenized in liquid nitrogen and digested in 700 µl of the solution containing 50 mM Tris–HCl, pH 8.0, 100 mM EDTA, 100 mM NaCl, 1% SDS and 0.5 mg/ml proteinase K at 55°C overnight, followed by treatment with 50 µg/ml of RNase A at 37°C for 1 h. The mixture was centrifuged to remove debris, the supernatant collected, treated with phenol/chloroform and precipitated.

RNA isolation

The tissue was mechanically broken in liquid nitrogen and RNA was extracted using the procedure described by Chomczynski and Sacchi (1987).

Primers

The primers for establishing the exon/intron structure of the *G. cydonium* OAS gene were designed on the basis of the corresponding cDNA sequence (EMBL/GenBank Accession No. Y18497). The primers were named according to this sequence, the numbers indicating the 5′ positions of the primers. The primer set was defined using Primer 3.0 software (MIT). The primers were then re-screened to minimize the presence of those containing the AGG triplet (and reverse primers containing the CCT triplet) that represents the most common splice junction consensus sequence. Additionally, the primers R1119 and R1140 were used to obtain the longest possible non-coding sequence. The primer sequences are presented in Table 1.

Table 1 The primers for the analysis of the OAS gene from *G. cydonium*

5′ Primer		3′ Primer	
L4	GGTTTAGCCTCGGACCAA	R245	CCCAGTGAGCCAGCCTTA
L30	TCGTTCTGCTGACCCAGA	R255	GGTTCCGTGACCCAGTGA
L78	TATGGCTAGCCCGGTTCC	R256	AGGTTCCGTGACCCAGTG
L84	TAGCCCGGTTCCTCCAG	R485	GGGCTCACCAGGAGATCA
L93	TCCTCCAGGTGATGTTCCA	R493	CCCAGTAAGGGCTCACCA
L130	GCAAACTGTGCCGTAAAGG	R495	CTCCCAGTAAGGGCTCACC
L136	TGTGCCGTAAAGGAGATCG	R605	GGGGGCTGTCGACTCATA
L232	GCTGGCTCACTGGGTCAC	R692	GAGCTAGGTCGCCCTTCA
L239	CACTGGGTCACGGAACCT	R693	TGAGCTAGGTCGCCCTTC
L240	ACTGGGTCACGGAACCTC	R816	ACGTCGTCCCCAGTGAAC
L243	GGGTCACGGAACCTCAGTC	R820	TACCACGTCGTCCCCAGT
L482	GCCCTTACTGGGAGAGCTG	R892	GTGGTCTGCTTGGCAGGA
L563	GTGCTTCCAGATGGCAGAG	R914	GGGTTGGCAGGGTCAATA
L591	GAGTCGACAGCCCCCAAT	R949	AGCCAGCAATCCCAGAGA
L593	GTCGACAGCCCCCAATAG	R950	TAGCCAGCAATCCCAGAGA
L594	TCGACAGCCCCCAATAGT	R995	TTTCCATCTCCACCGTCA
L645	GGATGACACATGGCCTGA	R1119	ACCACATCATAGCCATGACACA
L673	TGTGAAGGGCGACCTAGC	R1140	TACTGTAAAACAAAA
L801	TCACTGGGGACGACGTG		
L804	CTGGGGACGACGTGGTAA		
L879	GCCAAGCAGACCACGAGT		
L69	CAAAGGTGATATGGCTAGCCCGGTT		

The following primers were designed to discriminate the GeoOAS1Aa/GeoOAS1Ab sequences, which were found during the study:

L283 CTTGTGGTCTATTCAAGA
R788 TCTCTATGAACTAATCCAATAA

The cloned PCR amplicon L879-R245 containing the GeoOAS1A intergenic region was analyzed using the following additional primers:

AGCTCTACCAGTGGGCTCAA
GCAGAACGAAGAAAGCCTTG

PCR

The genomic sequences were amplified using HotStarTaq Master Mix (Qiagen) according to the manufacturer's instructions. The PCR temperature profile was as follows: initial denaturation at 95°C for 15 min, followed by 35 cycles at 58°C for 30 s, 72°C for 2 min and 95°C for 30 s with a final extension at 72°C for 10 min. The reaction was carried out in 10 µl of mixture containing 5 ng of genomic DNA and 5 pmoles of each primer. For comparison, 5 pg of the plasmid expressing OAS cDNA of *G. cydonium* (Päri et al. 2007) was analyzed under the same conditions. PCR products were analyzed in agarose gels stained with ethidium bromide.

RT-PCR

Total RNA weighing 3.6 µg was reversely transcribed using RevertAid™ M-MuLV Reverse Transcriptase (Fermentas) either with the primer oligoTA (5′-T_{17}A-3′) or R1140 following the manufacturer's protocol. PCR was carried out with primers L283/R788 or L78/R1140 (20 pmol each). The PCR was performed using 2× PCR Master Mix (Fermentas). Thermal cycling was as follows: initial denaturation at 95°C for 2 min, 35 cycles at 94°C for 1 min, 44°C for 1 min, 72°C for 1.5 min and a final extension at 72°C for 10 min. The PCR products were separated in a 1% agarose gel and visualized under UV-light after staining with ethidium bromide. Bands were excised from the gel and DNA was purified using Jetquick Gel Extraction Spin Kit (Genomed) and subjected to direct sequencing from both ends, using the same primers that were used in PCR.

5′- and 3′-RACE procedures were performed using FirstChoice® RLM-RACE kit (Ambion) according to manufacturer's instructions except for OAS1Ab 5′-region cDNA that was synthesized using primer R788. For the amplification of OAS1Aa 5′-region, primers R820 and R257 were combined with 5′-RACE outer and inner primers, respectively. The OAS1Ab 5′-region was amplified using primers R788 and R256 together with 5′-RACE outer and inner primers, respectively. The 3′-RACE was performed using primer L69.

Sequencing

Homogenous PCR products were purified, using Jetquick PCR Purification Spin Kit (Genomed), and subjected to straight sequencing, using the same primers that were used in PCR. Alternatively, amplicons were first cloned into the pTZ57R/T vector using InsTAclone PCR Cloning Kit™ (Fermentas) following the manufacturer's instructions. The vector specific primers were used in the latter case.

The sequencing reaction was performed using BigDye Terminator v3.1 Cycle Sequencing chemistry (Applied Biosystems) with 25 cycles of denaturation at 95°C for 30 s, annealing at 50°C for 15 s and extension at 60°C for 4 min. ABI PRISM 3130 Genetic Analyzer (Applied Biosystems) was used for the analysis.

The analysis of sponge DNA databases

Known protein sequences of sponge OAS were used for the search of cDNA and genomic sequences of OAS from the databases of the marine sponges *Oscarella carmela* Muricy and Parse 2004 (Porifera, Demospongiae, Homoscleromorpha, Homosclerophorida, Plakinidae, Oscarella) and *Amphimedon queenslandica* Hooper and van Soest 2006 (Porifera, Demospongiae, Haplosclerida, Haplosclerina, Niphatidae, Amphimedon) using the tBLAST program (Altschul et al. 1997). *O. carmela* has 11,176 entries in the NCBI EST database (http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucest) and *A. queenslandica* has both EST and WGS data (83,040 and 28,34,852 raw single-pass reads of DNA sequences) (ftp://ftp.ensembl.org/pub/traces/reniera_sp_jgi_2005). While the EST matches for OAS were found in *O. carmela*, no ESTs were detected in the *A. queenslandica* database. Multiple matches coding for several putative OAS genes were present in the *A. queenslandica* WGS database. Three main types of OAS could be defined; genomic contigs were assembled for each OAS gene. To resolve the genomic arrangement of OAS genes, the flanking sequences were assembled and analyzed. Nucleotide sequence manipulations were performed using the BioEdit program (http://www.mbio.ncsu.edu/BioEdit/).

Results

OAS is presented by at least two gene versions in the genome of *Geodia cydonium*

A set of primer pairs based on the OAS1 cDNA sequence from *G. cydonium* was used to design the experiment to

define the genomic structure of the corresponding gene. The genomic PCR products with chain lengths longer than those from the corresponding cDNAs were subjected to sequencing. The results showed that several amplicons that should have had overlapping regions were derived from highly similar yet different targets. Therefore the sequencing strategy was modified by amplifying the longest possible genomic fragments. The usage of the primer pairs L4-R995 and L78-R1119 resulted in homogenous amplicons derived from two different target sequences, respectively. The target corresponding to the amplicon L4-R995 (1899 bp), the sequence of which was closer to that in GenBank (Y18497) was termed GeoOAS1Aa. The predicted cDNA sequence of GeoOAS1Aa differed from that in the database (Y18497) by 7 transitions (3 amino acid substitutions), including the mismatch in the 3′-end of the primer L78 (position 94).

The sequence of the longest product obtained using primer L78 (78–1119) defined the second version of the OAS gene (hereinafter referred to as GeoOAS1Ab) (Fig. 1). Both GeoOAS1Aa and GeoOAS1Ab genomic fragments were deposited in the GenBank database (Acc. No. EU856087 and EU856088, respectively). Differences between the two versions of the OAS gene in the region 78–995 include: (1) insertion/deletion of one triplet in the coding region (Ser125), (2) 44 mismatches in the exons resulting in 26 amino acid substitutions, and (3) minor variations of the lengths and sequences of the introns (Fig. 1; Supplementary Table 1). The overall nucleotide sequence identity between the two gene versions is 94.2%. Co-amplification of the fragments from GeoOAS1Aa and GeoOAS1Ab exhaustively explains the heterogeneity of the shorter amplification products observed by us.

OAS in *Geodia cydonium* has a unique exon/intron structure different from that in vertebrates

The comparison of the genomic OAS sequences, elucidated from *G. cydonium*, with the original cDNA sequence (Y18497) (Wiens et al. 1999) showed the presence of eight coding exons in the GeoOAS1A gene plus the non-coding 5′-leader sequence and a long intron following it (Fig. 2). The gene has two types of introns, mainly of short length but three of them are of longer length. The translational start codon is located in exon 1 in position 79 and the translational stop is in the last exon in position 1060. The characteristic 2-5A synthetase regions, described by Wiens et al. (1999), are localized in three exons (3, 6, 8) and two functionally important amino acids are found in exon 4.

All the exon/intron boundaries comply with the GT/AG rule for splice junctions (Fig. 1).

The genomic structure of an OAS unit has been conserved in all human 2-5A synthetases and also in other mammalian OASs (Kumar et al. 2000; Perelygin et al. 2006). The results of our study showed that the number of translated exons in the *G. cydonium* OAS gene is different from that in the mammalian OAS unit (eight and five, respectively). Moreover, the comparison of a representative of mammalian OAS, the OAS1 gene from human, with that from *G. cydonium* conducted on the basis of the amino acid alignment revealed that the intron positions relative to the conserved amino acid stretches are completely different. However, in the order of intron phases there is a similarity which can be followed in the OAS gene from *G. cydonium* and a mammalian OAS gene. In comparison with the human OAS gene, all four introns in respective phases are shifted upstream in the sponge gene (Fig. 3).

Both OAS gene versions are expressed in *Geodia cydonium*

The original OAS cDNA from *G. cydonium* (Y18497) has the sequence close to that of GeoOAS1Aa. Thus the expression of GeoOAS1Aa has already been demonstrated. We now addressed the question of the expression of the other gene version, GeoOAS1Ab. The RNA was isolated from a sample from individual 1. The presence of a cDNA fragment of GeoOAS1Ab in this sample was confirmed by sequencing the amplicon obtained with the primer pair L283-R788. The following 5′-RACE resulted in 5′-terminus of GeoOAS1Aa transcript lacking exon 0; the transcription initiation site was in the position 55. The 5′-UTR of GeoOAS1Ab transcript was 42 nucleotides longer than that in GeoOAS1Aa′s shorter transcript, attributable to either exon 1 or exon 0. Based on this result, the universal primer L69 was designed, which in combination with the 3′-RACE outer primer yielded a mixture containing the full-length coding cDNAs for both GeoOAS1Aa and GeoOAS1Ab. The mRNA amplicons were deposited in the GenBank (Acc. Nos. EU856090–EU856093).

OAS1 gene versions in G. cydonium are tandemly organized

The existing primers were used to test the possibility of tandem arrangement of GeoOAS1Aa and GeoOAS1Ab by PCR. We succeeded in amplifying the intergenic area using primers L801 and R256 (Fig. 4).

The 3′-terminus of the fragment matched the 5′-terminus of the GeoOAS1Aa, including the first two exons and the first two introns. Thus the downstream OAS version is most probably GeoOAS1Aa. The 5′-terminus of this fragment had a short match with the amplicon from the 3′-terminus of the GeoOAS1Aa. Still the 5′-partner of this tandem repeat can be GeoOAS1Ab, as well, as this region does not contain any definitive descriptor that discriminates between possible GeoOAS1A versions. The overall length of the intergenic stretch is 1620 nucleotides and it contains

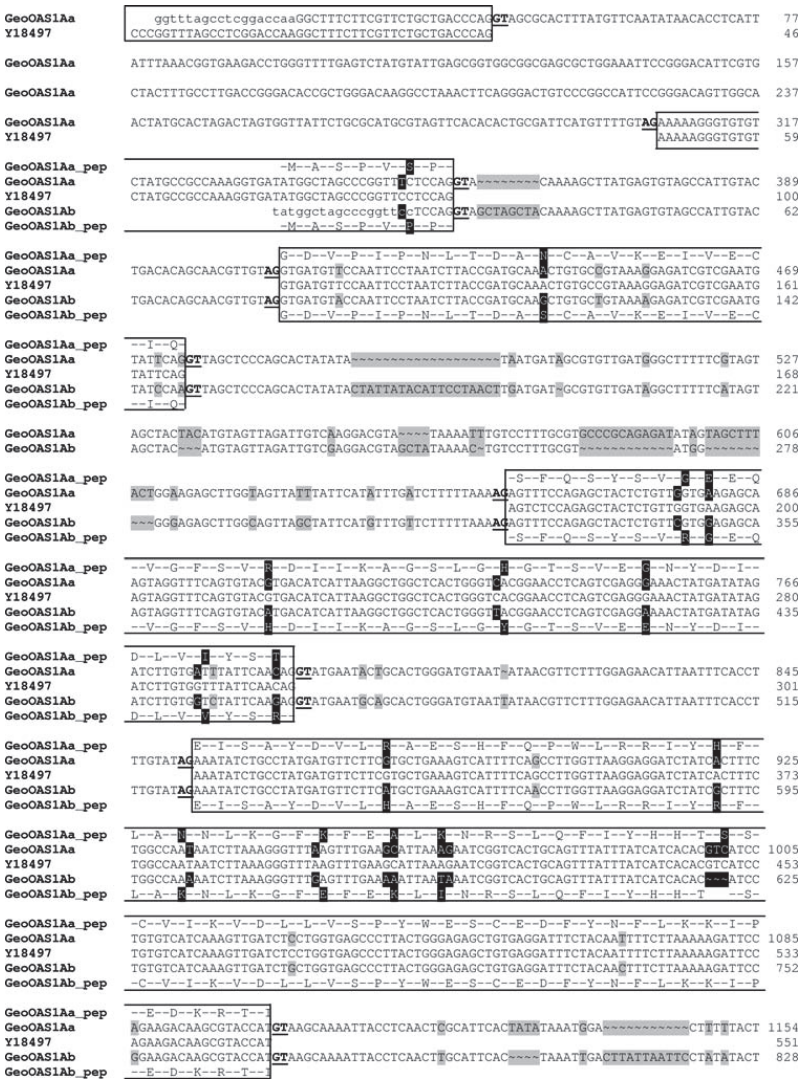


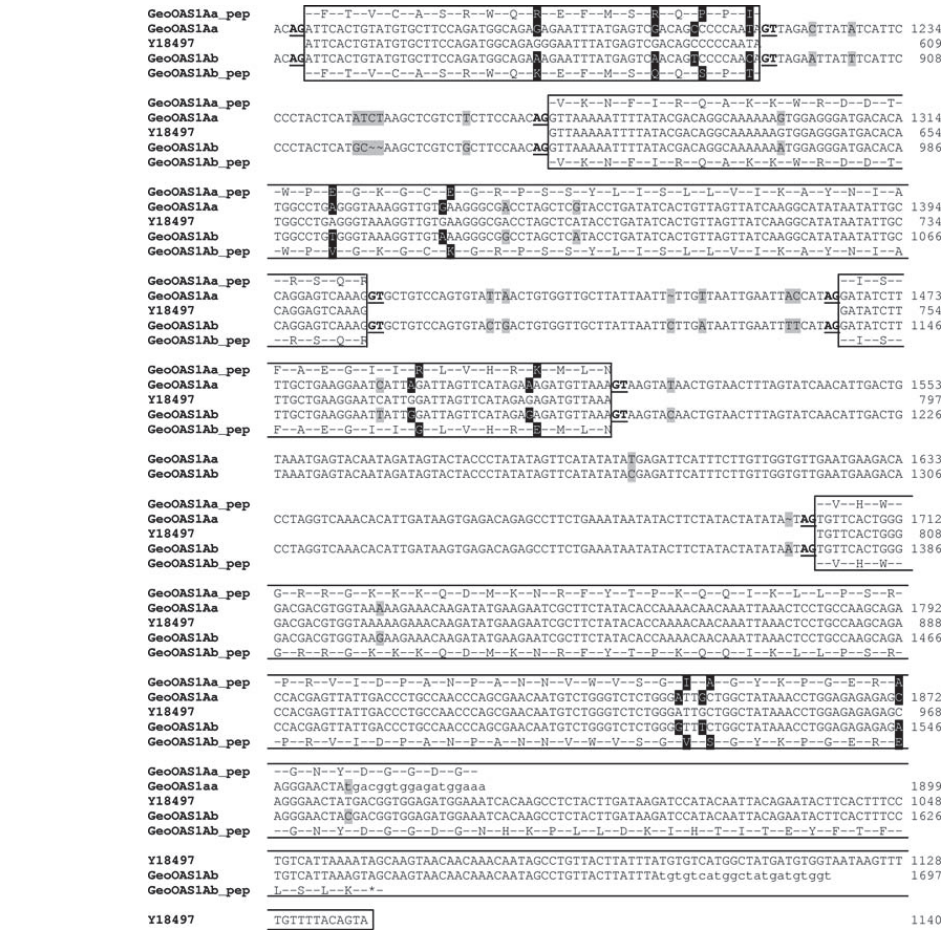
Fig. 1 Alignment of GeoOAS1Aa and GeoOAS1Ab. The genomic sequences of GeoOAS1Aa and GeoOAS1Ab from individual 1 were aligned with GeoOAS1 cDNA sequence from the database (Y18497) and then compared to each other. The exons are shown in *boxes*. *Black* residues indicate both the amino acid changes and the nucleotide mismatches causing these changes. *Gray* areas indicate the changes which do not alter the coded protein. The primer-forced sequences are given in *small letters*. The beginnings and ends of introns are indicated in *bold* and by means of the *underlined letters*

eight repeats of 44 nucleotides after the end of mRNA. This sequence was deposited in the GenBank database (Acc No. EU856089).

OAS gene versions in *G. cydonium* are genetically linked and may have more than two haplotypes

Both GeoOAS1Aa and GeoOASAb were also identified in a sample from individual 2 (Acc. Nos. EU856094 to EU856096), which refers to their genetic linkage; they are not allelic variants of the same gene. GeoOAS1Aa sequences of two individuals were highly similar (Table 2A), however, in the region 4–949 still a number of single nucleotide substitutions were detected, 8 of them being in exons and 9 in introns (where additionally two deletions were observed). Five substitutions were non-synonymous, leading to the amino acid substitutions. In contrast to GeoOAS1Aa, the genomic sequence of GeoOAS1Ab was almost identical to that in the first individual in the analyzed region 78–892 (Table 2A); only a single nucleotide insertion/deletion was detected in intron 7. The occurrence of individual differences of GeoOAS1A genes in particular animals demonstrates the presence of the GeoOAS1A haplotypes in *G. cydonium.*

Fig. 1 continued



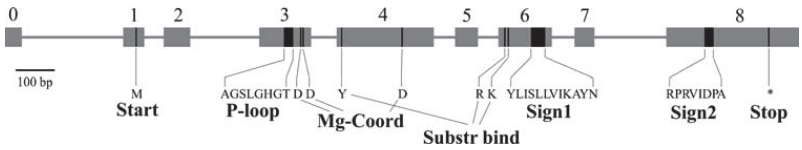


Fig. 2 The GeoOAS1A genomic structure. Exon numbers are indicated on the *top*. The start and stop codons are marked as "*M*" and *, respectively. The distribution of important residues in exons is given under the gene. The contiguous structural motifs *P-loop*, *Sign1* and *Sign2* (conserved signatures) are given according to Wiens et al. (1999); three aspartic acid residues participating in Mg^{2+}-coordination are as proposed by Hartmann et al. (2003); substrate binding residues are presented according to Sarkar et al. (2002)

The comparison of the OAS from *G. cydonium* with homologous sequences found in sponge databases

To place the experimental data obtained in this study into a wider context, available sponge datasets were analyzed for the structure and organization of OAS genes in other sponges. *Amphimedon queenslandica* served as a model sponge for its genome being extensively sequenced and annotated. In the WGS database, a number of tBLAST hits were found against two partly conserved amino acid stretches in the GeoOAS1A molecule, one of them being in the central part of the protein and another one close to its carboxy-terminus. The nucleotide sequences were extended and complete structures of the hypothetical OAS genes for

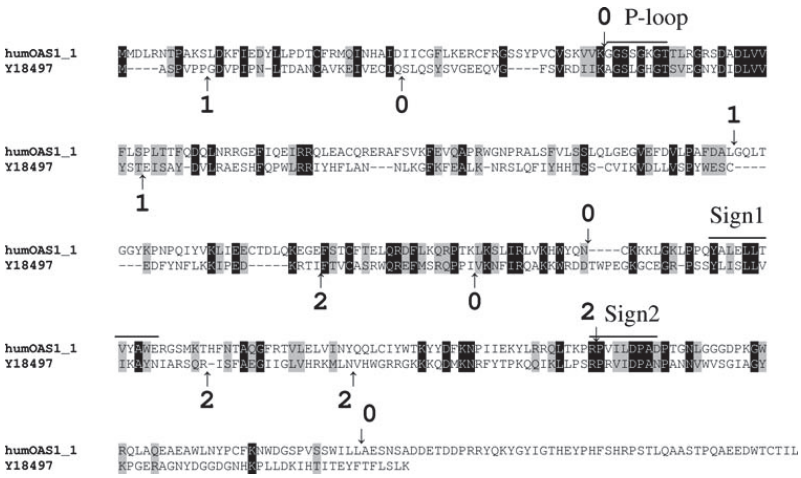


Fig. 3 Comparison of mammalian and sponge OAS structures. Amino acid sequences are derived from human 2′,5′-oligoadenylate synthetase 1 isoform 1 (NP_058132.2) and from the OAS of *G. cydonium* (Y18497). The alignment of sequences is shaded according to the identity (in *black*) and similarity (*gray*) of amino acids. Intron locations of the corresponding genes are indicated by *arrows*; the *numbers* indicate intron phases

Fig. 4 The intergenic region between two copies of GeoOAS1A. Exons are shown in *black* and numbered according to Fig. 2. The *gray box* indicates the localization of a tandem repeat of 44 nucleotides. PCR primers are indicated on the *top*

A. queenslandica were predicted (Fig. 5). Furthermore, the common genomic structure for sponge OAS could be defined despite minor variations between exon/intron structures of different genes from *G. cydonium* and *A. queenslandica* (Fig. 5).

Three main types of OAS (AmpOAS1A, AmpOAS1B and AmpOAS1C) could be defined with several variants and numerous subvariants for each of them (Table 3). The differences in the sequences of AmpOAS1Aa and AmpOAS1Ab were more pronounced than those of GeoOAS1Aa and GeoOAS1Ab (Tables 2A, 3), as we could not predict the first two exons at the equivalent positions for AmpOAS1A (Fig. 5). AmpOAS1B also had at least two different versions where the end of exon 6 and the beginning of the next exon could not be predicted in the equivalent positions. The alternative versions of the third AmpOAS type, AmpOAS1Ca and AmpOAS1Cb had identical exon/intron structures; the third version AmpOAS1Cc

Table 2 Intra-species differences between individual OAS protein sequences from *G. cydonium* (*A*) and *S. domuncula* (*B*)

A

	GeoOAS1Aa	GeoOAS1Aa′	GeoOAS1Aa″	GeoOAS1Ab	GeoOAS1Ab″
GeoOAS1Aa	1	0.983	0.983	0.910	0.910
GeoOAS1Aa′	0.990	1	0.993	0.921	0.921
GeoOASAa″	0.990	0.993	1	0.921	0.921
GeoOAS1Ab	0.945	0.948	0.948	1	1.000
GeoOAS1Ab″	0.945	0.948	0.948	1.000	1

B

	SubOAS1Aa	SubOAS1Ab	SubOAS1Ac
SubOAS1Aa	1	0.951	0.969
SubOAS1Ab	0.972	1	0.950
SubOAS1Ac	0.975	0.972	1

GeoOAS1Aa′ sequence corresponds to the database sequence Y18497; GeoOAS1Aa is derived from individual 1 and GeoOAS1Aa″ is derived from individual 2. SubOAS1Aa and SubOAS1Ab correspond to database entries AJ301652 and AJ301653, respectively; SubOAS1Ac was taken from Müller and Müller (2003). The numbers above the diagonal indicate identities of the protein pairs, the numbers below the diagonal represent similarities of protein pairs according to BLOSUM62

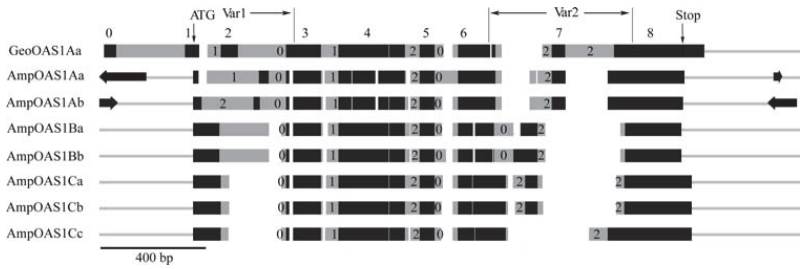


Fig. 5 Schematic alignment of the OAS genomic structure from *G. cydonium* with predicted genomic OAS structures from *A. queenslandica*. The illustration is based on amino acid alignment. Only the examples with non-identical exon/intron structures are presented. Exon numbers are given on the *top*. Exons are marked in *black*; introns are *gray* and not aligned; the gaps inside introns indicate the different lengths of introns; the gaps in exons indicate the insertions/deletions (indels) of amino acids. Intron phases are written on the introns. *Thick black arrows* indicate the exons of the neighboring genes, described in Fig. 6. The two variable regions without any predictable common structure are marked as "*Var 1*" and "*Var 2*", respectively

Table 3 Differences between individual OAS protein sequences predicted from genomic data for *A. queenslandica*

	AmpOAS1Aa	AmpOAS1Ab	AmpOAS1Ba	AmpOAS1Bb	AmpOAS1Ca	AmpOAS1Cb	AmpOAS1Cc
AmpOAS1Aa	1	0.894	0.219	0.222	0.215	0.227	0.229
AmpOAS1Ab	0.928	1	0.221	0.224	0.215	0.227	0.235
AmpOAS1Ba	0.362	0.365	1	0.887	0.456	0.468	0.409
AmpOAS1Bb	0.367	0.371	0.934	1	0.476	0.482	0.426
AmpOAS1Ca	0.382	0.357	0.626	0.644	1	0.935	0.741
AmpOAS1Cb	0.385	0.360	0.629	0.644	0.970	1	0.738
AmpOAS1Cc	0.394	0.382	0.577	0.597	0.860	0.857	1

See also Figs. 5, 6 and 7. The numbers above the diagonal indicate the identities of the protein pairs, the numbers below the diagonal represent the similarities of protein pairs according to BLOSUM62

differed from them by lacking the sixth exon and having a longer last exon (Fig. 5).

The characteristic heterogeneity of OAS1A genes from *G. cydonium* and *A. queenslandica* was also observed in *Oscarella carmela*. Three ESTs (G840P312L18, G840P313A21, G840P310A16) are the products of the same gene, one of them (G840P310A16) containing a single non-spliced intron. The intron position and phase matches those of the intron between exons 5 and 6 in GeoOAS1Aa. The fourth EST (G840P37P5) is derived from a separate gene as deletion of one amino acid has occurred in it.

Continuing the analysis of BAC clones presented in WGS database for *A. queenslandica*, we could locate the AmpOAS1 genes into different loci. In the first locus, two AmpOAS1A copies are tandemly arranged (clone BAYA372634) with about 200 nucleotides between them. Some AmpOAS1A versions may also exist as single copies between the same flanking regions. Interestingly, these flanking regions belong to an intron of another gene (Fig. 6a). The AmpOAS1B and AmpOAS1C genes are localized in a separate locus. Their mutual arrangement is tandemic, as well, but the repeat includes a part of the neighboring gene. AmpOAS1C is located about 6 kb downstream from AmpOAS1B (Fig. 6b).

AmpOAS1Ba and AmpOAS1Bb, as well as AmpOAS1Ca, AmpOAS1Cb and AmpOAS1Cc represent only some of the numerous alternative locus variants. The general structure of a locus may be considered to be a consensus sequence that is not attributable to the individual sequences.

Summarizing the results of the analysis of genomic data presently available for *A. queenslandica*, we can conclude that the arrangement of the OAS1 genes in *A. queenslandica* is different from that in vertebrates. Instead of a single cluster of genes, two types of AmpOAS1 loci can be observed. These loci containing AmpOAS1A and AmpOAS1B/AmpOAS1C, respectively, are separated by at least five other genes.

Discussion

The aim of this work was to establish whether the 2-5A synthetases from the evolutionarily lowest multicellular

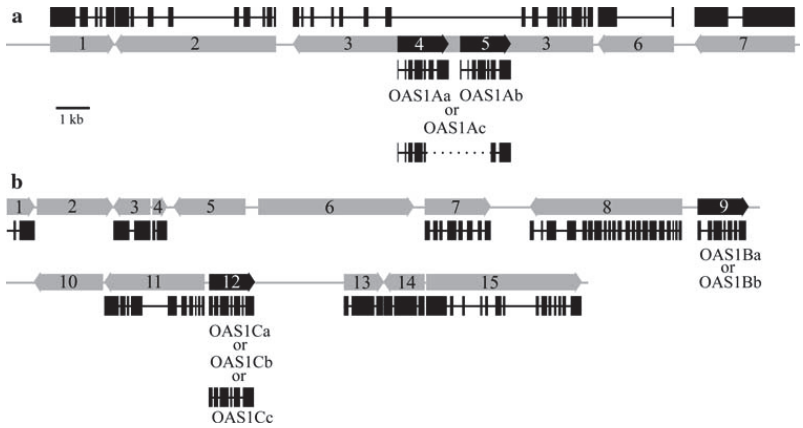


Fig. 6 Schematic presentation of two loci in *A. queenslandica* containing OAS genes. **a** The AmpOAS1A locus (23 kb). **b** The locus containing AmpOAS1B and AmpOAS1C (39 kb). The placement and direction of genes is shown with *thick arrows*. The exons in the genes, which have been elucidated from EST sequences or which are predicted by strong homologies of their ORFs with known protein structures, are given in separate *lines*. Elucidated genes and ORFs are *numerated*. Predicted neighboring genes are specified in Supplementary Table 2. Evidence for the tandem arrangement of OAS1Aa and OAS1Ab in **a** comes from the genomic clone BAYA372634. The *dotted line* in the structure of OAS1Ac indicates the genomic deletion, which includes the 3′-end of OAS1Aa and 5′-end of OAS1Ab

animals, sponges, whose primary structure is distantly related to those of mammalian OAS share a common genomic structure with corresponding vertebrate genes. The results of our study indicate that the exon/intron structure of the sponge OAS is completely different from the pattern found in higher animals (generalized by human OAS1, Fig. 3). Evidently, due to the low evolutionary conservation, the studied case, 2′,5′-oligoadenylate synthetases, is in contrast with the more conserved genes in sponges, whose architecture is close to that in higher animals, including humans (Müller et al. 2002). The OAS gene in the marine sponge *G. cydonium* has eight translated exons instead of five found in the mammalian OAS unit, and the intron positions relative to the conserved amino acid motifs are completely different. Also, the genomic analysis showed that the introns in the OAS genes from *G. cydonium* are shorter than that in mammals; this seems to be typical of sponges in general (Müller et al. 2002). The available set of genomic sequences from another marine sponge, *A. queenslandica*, was used to enlighten the question whether the OAS gene structure from *G. cydonium* is a unique one or characteristic of other sponge species as well. The surprising result of the analysis of the genomic data of *A. queenslandica* was the presence of three single-domain OAS gene types whose coding sequences differ from one another at the same level as we can observe between different sponge species (Fig. 7; Table 4). Generally, the interspecies comparison of known sponge OAS protein sequences in Table 4 refers to very low similarity between these proteins, most probably due to their independent evolution after the divergence from their common ancestor. All three OAS genes have a structure close to that in *G. cydonium* (Fig. 5); no OAS with a vertebrate-like gene structure was identified. Our findings demonstrate a common overall genomic structure of sponge OASs in spite of the low homology of their amino acid sequences. The single OAS intron found in the EST dataset of *O. carmela* also coincides with the consensus structure proposed by us.

The results of the elucidated exon/intron structure do not support the idea of a hypothetical direct intron-containing common ancestor for vertebrate and sponge OAS. If that was the case, all the conserved motifs in human OASs should have significantly migrated to the neighboring exons, which is highly improbable (Rogozin et al. 2005). If we take the major mechanism of intron evolution, intron gain and loss, to be responsible for the differences between human and sponge OAS genes, at least some of the intron positions should be conserved (Rogozin et al. 2005). Furthermore, the reason of the observed similarity of intron phase patterns (Fig. 3) would then remain elusive. Thus, our findings with the sponge OAS genes underline the complexity of the evolution of intron–exon structure of eukaryotic genes. Anyway, the proposition of a direct common ancestor of sponge OAS with the respective vertebrate enzymes that was emphasized earlier (Schröder et al. 2008) has been rendered improbable in the light of the obtained results. When more OAS genomic sequences from different eukaryotic lineages become available, the detailed phylogenetic relationships in the evolutionary history of OAS genes could be established. Our data are in accordance with the

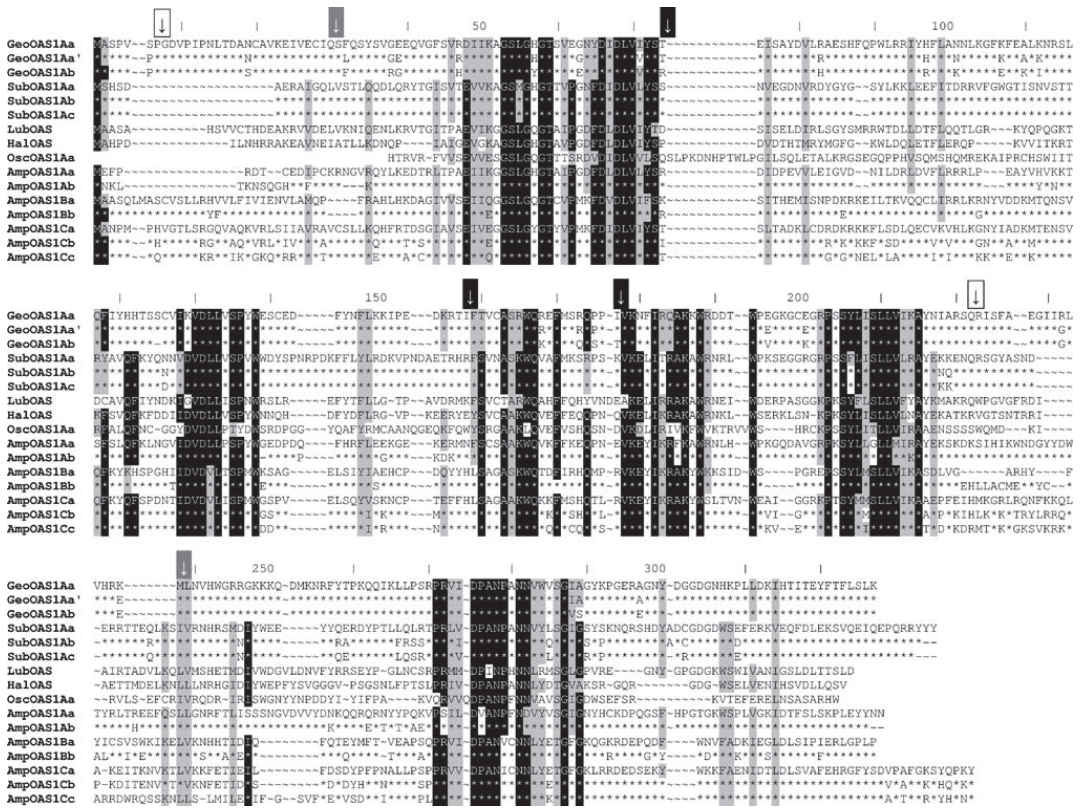


Fig. 7 Alignment of sponge OAS proteins. GeoOAS1Aa (EU856090) and GeoOAS1Ab (EU856092) were isolated from individual 1. The database sequence Y18497, termed here as GeoOAS1Aa′, was used to illustrate the variability of GeoOAS1Aa. SubOAS1Aa and SubOAS1Ab correspond to database entries AJ301652 and AJ301653, respectively; SubOAS1Ac and HalOAS were taken from Müller and Müller (2003). LubOAS sequence was published by Schröder et al. (2008). OASs from *O. carmela* and *A. queenslandica* were deduced from the database analysis. For *O. carmela*, only OscOAS1Aa is shown. *Lub L. baicalensis*, *Hal H. panicea*, *Osc O. carmela.* One representative from each OAS type is fully presented; for other representatives only variable positions are given by amino acids. Identical amino acids in the other representatives of the OAS type are marked by *asterisks*. Missing amino acids in protein variants are marked as "-"; other gaps are marked as "∼". The alignment is *shaded* on the basis of the position conservation. *Arrows* indicate the positions of introns: *arrows* in *black boxes* indicate the intron positions that are conserved through OASs from *G. cydonium* and *A. queenslandica; arrows* in *gray boxes* indicate the introns in conserved phases; *arrows* in *white boxes* indicate non-conserved intron positions

early radiation of the OAS genes, on the basis of each of which new characters have evolved independently of the others.

As in higher organisms, the OAS genes in sponges have duplicated and diverged. In contrast to mammals, where OAS paralogs are clustered together in a single locus (Eskildsen et al. 2002; Mashimo et al. 2003; Perelygin et al. 2006), the repertoire of OAS family in *A. queenslandica* has resulted from several independent gene duplication/diversification/deletion events of a common intron-containing ancestor. The protein sequence identity analysis shows that the most ancient one has apparently led to the localization of AmpOAS1A and AmpOAS1B/AmpOAS1C in different loci. These AmpOAS1 genes of different loci are the most diverged ones in their primary structures, as well (amino acid sequence identity of about 25%) (Table 3). Thus, contrary to mammals (Perelygin et al. 2006), the multiplicity of sponge OAS genes has not arisen from the expansion of gene-wise *cis*-duplications in a single locus. In other sponge species, analyzed in this respect, we presently know of only more recent duplication events (e.g., GeoOAS1Aa and GeoOAS1Ab, amino acid sequence identity of about 90%; OAS sequences from *S. domuncula* in GenBank with amino acid sequence identity of about 95%) (Table 2A, B). Taking into consideration the case of *A. queenslandica*, we can also expect the presence of more

Table 4 Inter-species comparison of sponge OAS protein sequences

	GeoOAS1Aa′	SubOAS1Aa	LubOAS1A	HalOAS1A	OscOAS1Aa	AmpOAS1Aa	AmpOAS1Ba	AmpOAS1Cb
GeoOAS1Aa′	1	0.280	0.268	0.283	0.187	0.286	0.229	0.219
SubOAS1Aa	0.441	1	0.332	0.378	0.245	0.322	0.224	0.254
LubOAS1A	0.403	0.504	1	0.404	0.199	0.330	0.248	0.251
HalOAS1A	0.442	0.551	0.552	1	0.225	0.360	0.244	0.280
OscOAS1Aa	0.286	0.352	0.302	0.347	1	0.222	0.197	0.214
AmpOAS1Aa	0.434	0.509	0.471	0.505	0.336	1	0.219	0.227
AmpOAS1Ba	0.393	0.422	0.389	0.390	0.342	0.362	1	0.482
AmpOAS1Cb	0.372	0.434	0.390	0.403	0.319	0.385	0.629	1

The abbreviations are given in the legend to Fig. 7. The numbers above the diagonal indicate identities of the protein pairs, the numbers below the diagonal represent similarities of protein pairs according to BLOSUM62

diverged OAS sequences in the genomes of other sponges. The data obtained so far do not refer to the possible expansion of the OAS gene family in sponges by domain couplings to produce the multidomain genes as it has occurred in mammals (Kumar et al. 2000).

The OAS1 gene is presented in *G. cydonium* in at least two genetically linked versions, both being expressed in the sponge. We demonstrated the tandemly duplicated arrangement of the OAS1 genes in the genome of *G. cydonium* (Fig. 4). However, the analysis of the intergenic region amplified in this study did not unambiguously reveal the presence of both gene versions in the tandem repeat; the GeoOAS1Ab gene could reside in another locus, as well. These data taken together with the cDNA data (at least two subvariants of GeoOAS1Ab in the same individual; data not shown) may indicate that the total copy number of OAS1 genes in the genome of a single individual of *G. cydonium* may be higher than that revealed in this study.

In addition to the presence of at least two GeoOAS1 versions in the same individual, the inter-individual differences in each gene version was observed (Table 2A). A similar small heterogeneity on the cDNA level has also been demonstrated for OAS from some other sponges (Table 2B; Fig. 7 for *S. domuncula* and "Results" for *O. carmela*, respectively). However, there is no data indicating whether the three different OAS sequences from *S. domuncula* belong to the same or different individual(s). The data for *O. carmela* were obtained from the analysis of an individual specimen of *O. carmela* bearing densely brooded embryos and larvae of all stages (Nichols et al. 2006). Contrary to these cases, our study allows for differentiation between the heterogeneities present in the same sponge individual (GeoOAS1Aa and GeoOAS1Ab) and the inter-individual heterogeneities (e.g., GeoOAS1Aa from individuals 1 and 2).

For the assembled genome of *A. queenslandica*, Degnan (2007) predicted 15–20,000 gene models which should also include AmpOAS1A, AmpOAS1B and AmpOAS1C. The analysis of AmpOAS1A, AmpOAS1B and AmpOAS1C genes revealed high heterogeneity of each of them which could not be explained by clustering of corresponding genes, as not more than two OAS genes are tandemly arranged in *A. queenslandica*.

Two alternative hypotheses may be proposed: (1) either the observed sequence variability is derived from a multiplicity of genomes present in genomic DNA derived from embryos and larvae from one adult (Degnan 2007) or (2) OAS loci in the sponge genome are repeated several times. If we propose that every alternative copy of OAS versions in *A. queenslandica* belongs to a different individual, we can estimate the genome size to be <200–400 Mb (on the basis of OAS loci described in this paper, as well as of the independent loci chosen from BAC clones). At the moment, the dataset of *A. queenslandica* is interpreted as a 100–120 Mb genome with more than 12-fold coverage, represented by more than four haplotypes (Degnan 2007). We can get similar (about 12-fold) coverage of assembled loci only by summing up all loci variants observed, i.e., considering the variability of an OAS locus solely as a set of haplotypes. In this case, the inter-individual variability of OAS genes is very high, including the presence/absence of tandemly arranged OAS duplicates, altered splice sites etc.

So far, very few experimental data are available about the size of sponge genomes. Mirsky and Ris (1951) gave the value of 0.11 pg/cell of DNA for *Dysidea crawshagi*. Imsiecke et al. (1995) estimated the value of 3.7 pg DNA per cell for *S. domuncula* (genome size 1,809 Mb) and 3.3 pg DNA per cell for *G. cydonium* (genome size 1,614 Mb). The latter values are in the same range with that of the human genome. Müller et al. (2003) predicted the presence of approximately 300,000 short genes in the high-dense 1,670 Mb genomes of both *S. domuncula* and *G. cydonium*; in this case the sponge genome should be highly redundant. Accordingly, we can suppose that a particular OAS gene locus in the genome of *A. queenslandica* occurs

more than four times per genome (Degnan 2007), whereas the size of a repeat may be even larger than those assembled by us for AmpOAS1A and AmpOAS1B/AmpOAS1C (23 kb and 39 kb, respectively). Ascribing the copies with small mutations, e.g., point mutations in GeoOAS1Aa, to inter-individual polymorphisms, the more variable individual loci could be assigned to genomic repeats, which have resulted from large internal segmental duplications, well described in *Arabidopsis thaliana* (Cannon et al. 2004). Thus, the enormous individual variability is probably not the case and together with OAS, a multitude of other genes is repeated in the sponge genome. This concept does not contradict our experimental results which demonstrated that both GeoOAS1Aa and GeoOAS1Ab reside in the genomes of two different individuals.

GeoOAS1Aa gene versions in different individuals of *G. cydonium* can be interpreted as different haplotypes. Furthermore, we have treated a single sample (piece) from *G. cydonium* as belonging to a sponge individual, which may be an oversimplification. Following the view that any anatomically isolated and functionally autonomous sponge is an individual (Lieberkühn 1856), a sponge individual can be morphologically defined simply as a mass of sponge substance bounded by continuous pinacoderm (Borojevič et al. 1968). On the other hand, most sponges can be considered to be colonial organisms (Schmidt 1864; Hartman and Reiswig 1973), each one of them having its own osculum. The application of genetic criteria in defining the individual is also complicated in view of frequent fusion of larvae during their development and of neighboring sponges during their later growth, well demonstrated by Gauthier and Degnan (2008).

The functional contribution of the two OAS1 versions in *G. cydonium* is not clear at the moment. Gene duplicates are generally assumed to be functionally redundant at the time of origin (Lynch and Conery 2000). It is possible that some (or most) of sponge OAS genes are just pseudogenes with only one/several OAS version(s) being functional. The occurrence of pseudogenes could explain the changes in the sponge genome, leading to the observed divergence of OAS gene types/versions. The opposite hypothesis about the equivalence of the genomic copies of sponge OAS in terms of functionality is also attractive. The polymorphic nature of sponge OAS could be directly related to its high-affinity RNA binding properties (Päri et al. 2007). Hence, the particular versions of a sponge OAS may be pertinent to bind corresponding specific RNAs (e.g., tRNAs, bacterial and/or viral RNAs etc.), which may be present in the sponges thereby participating in the innate immune responses. However, knowledge of the pathways involved in the innate immune system of invertebrates, and more particularly of sponges, is imperfect at present.

Taken together, the different OAS gene structures in sponges and those in higher vertebrates are concurrent with differences in the primary structures and functioning of the corresponding proteins. Thus, the sponge OASs cannot be considered as true orthologs with the mammalian genes. When combined with the earlier studies of enzymatic characterization of the proteins, these findings allow us to consider the sponge OASs as a clearly distinct subgroup of 2-5A synthetases. From the evolutionary point of view, our results suggest that the last common ancestor of vertebrates and sponges already had several OAS genes which were further subjected to independent evolutionary scenarios leading to their deletions, duplications (including domain coupling) and/or modifications, keeping the elements of the ancient exon/intron structures. On the other hand, our results predict the presence of OAS ancestors in unicellular and colonial progenitors of metazoans, the genomic structures of which would give us valuable information for a further understanding of the evolution of this unique enzyme.

Acknowledgments This work was supported by the European Comission (project COOP-CT-2005, contract number 017800) and the Estonian Science Foundation (grant No. 5932).

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Aravind A, Koonin E (1999) DNA polymerase β-like nucleotidyltransferase superfamily: identification of three new families, classification and evolutionary history. Nucleic Acids Res 27:1609–1618

Borojevič R, Fry WG, Jones WC, Levi C, Rasmont R, Sara M, Vacelet J (1968) A reassessment of the terminology for sponges. Bull Mus Nat Hist Nat Paris 39:1124–1135

Cannon SB, Mitra A, Baumgarten A, Young ND, May G (2004) The roles of segmental and tandem gene duplication in the evolution of large gene families in *Arabidopsis thaliana*. BMC Plant Biol 4:10

Chawla-Sarkar M, Lindner DJ, Liu YF, Williams BR, Sen GC, Silverman RH, Borden EC (2003) Apoptosis and interferons: role of interferon-stimulated genes as mediators of apoptosis. Apoptosis 8:237–249

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

Degnan B (2007) The origin of animal development: insights from sponge evolutionary and functional genomics. In: Fourth international symposium on networks in bioinformatics (ISNB), 16–19 April 2007, Amsterdam, The Netherlands. http://www.nbic.nl/content/documents/10920/Degnan_ISNB07.pdf

Eskildsen S, Hartmann R, Kjeldgaard NO, Justesen J (2002) Gene structure of the murine 2′-5′-oligoadenylate synthetase family. Cell Mol Life Sci 59:1212–1222

Gauthier M, Degnan BM (2008) Partitioning of genetically distinct cell populations in chimeric juveniles of the sponge *Amphimedon queenslandica*. Dev Comp Immunol 32(11):1270–1280. doi:10.1016/j.dci.2008.04.002

Ghosh A, Sarkar SN, Sen GC (2000) Cell growth regulatory and antiviral effects of the P69 isozyme of 2–5 (A) synthetase. Virology 266:319–328

Grebenjuk VA, Kuusksalu A, Kelve M, Schütze J, Schröder HC, Müller WE (2002) Induction of (2′-5′)oligoadenylate synthetase in the marine sponges *Suberites domuncula* and *Geodia cydonium* by the bacterial endotoxin lipopolysaccharide. Eur J Biochem 269:1382–1392

Hartman WD, Reiswig HM (1973) The individuality of sponges. Animal colonies. Dowden, Hutchinson and Ross Inc, Stroudsburg, pp 567–584

Hartmann R, Justesen J, Sarkar SN, Sen GC, Yee VC (2003) Crystal structure of the 2′-specific and double-stranded RNA-activated interferon-induced antiviral protein 2′-5′-oligoadenylate synthetase. Mol Cell 12:1173–1185

Holm L, Sander C (1995) DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily. Trends Biochim Sci 20:345–347

Hovanessian AG (2007) On the discovery of interferon-inducible, double-stranded RNA activated enzymes: the 2′-5′ oligoadenylate synthetases and the protein kinase PKR. Cytokine and Growth Factor Rev 18:351–361

Imsiecke G, Custodio M, Borojevic R, Steffen R, Moustafa A, Müller WEG (1995) Genome size and chromosomes in marine sponges [*Suberites domuncula, Geodia cydonium*]. Cell Biol Int 19:995–1000

Justesen J, Hartmann R, Kjeldgaard NO (2000) Gene structure and function of the 2′, 5′-oligoadenylate synthetase family. Cell Mol Life Sci 57:1593–1612

Kumar S, Mitnik C, Valente G, Floyd-Smith G (2000) Expansion and molecular evolution of the interferon-induced 2′-5′ oligoadenylate synthetase gene family. Mol Biol Evol 17:738–750

Kuusksalu A, Pihlak A, Müller WE, Kelve M (1995) The (2′-5′)oligoadenylate synthetase is present in the lowest multicellular organisms, the marine sponges. Demonstration of the existence and identification of its reaction products. Eur J Biochem 232:351–357

Kuusksalu A, Subbi J, Pehk T, Reintamm T, Müller WE, Kelve M (1998) Identification of the reaction products of (2′-5′)oligoadenylate synthetase in the marine sponge. Eur J Biochem 257:420–426

Lieberkühn N (1856) Zusätze zur Entwicklungsgeschichte der Spongilliden. Arch Anat Physiol, pp 496–514

Lopp A, Kuusksalu A, Reintamm T, Müller WEG, Kelve M (2002) 2′, 5′-oligoadenylate synthetase from a lower invertebrate, the marine sponge *Geodia cydonium*, does not need dsRNA for its enzymatic activity. Biochim Biophys Acta 1590:140–149

Lynch M, Conery JS (2000) The evolutionary fate and consequences of duplicate genes. Science 290:1151–1155

Martin G, Keller W (2007) RNA-specific ribonucleotidyl transferases. RNA 11:1834–1849

Mashimo T, Glaser P, Lucas M, Simon-Chazottes D, Ceccaldi PE, Montagutelli X, Desprès P, Guénet JL (2003) Structural and functional genomics and evolutionary relationships in the cluster of genes encoding murine 2′,5′-oligoadenylate synthetases. Genomics 82:537–552

Mirsky AE, Ris H (1951) The desoxyribonucleic acid content of animal cells and its evolutionary significance. J Gen Physiol 34:451–462

Müller WEG, Müller IM (2003) Origin of the metazoan immune system: identification of the molecules and their functions in sponges. Integr Comp Biol 43:281–292

Müller WE, Böhm M, Grebenjuk VA, Skorokhod A, Müller IM, Gamulin V (2002) Conservation of the positions of metazoan introns from sponges to humans. Gene 295:299–309

Müller WEG, Brümmer F, Batel R, Müller IM, Schröder HC (2003) Molecular biodiversity. Case study: Porifera (sponges). Naturwissenschaften 90:103–120

Nichols SA, Dirks W, Pearse JS, King N (2006) Early evolution of animal cell signaling and adhesion genes. Proc Natl Acad Sci USA 103:12451–12456

Päri M, Kuusksalu A, Lopp A, Reintamm T, Justesen J, Kelve M (2007) Expression and characterization of recombinant 2′, 5′-oligoadenylate synthetase from the marine sponge *Geodia cydonium*. FEBS J 274:3462–3474

Perelygin AA, Zharkikh AA, Scherbik SV, Brinton MA (2006) The mammalian 2′-5′ oligoadenylate synthetase gene family: evidence for concerted evolution of paralogous Oas1 genes in Rodentia and Artiodactyla. J Mol Evol 63:562–576

Player MR, Torrence PF (1998) The 2-5A system: modulation of viral and cellular processes through acceleration of RNA degradation. Pharmacol Ther 78:55–113

Reintamm T, Lopp A, Kuusksalu A, Subbi J, Kelve M (2003) Qualitative and quantitative aspects of 2-5A synthesizing capacity of different marine sponges. Biomol Eng 20:389–399

Rogozin IB, Aravind L, Koonin EV (2003) Differential action of natural selection on the A and C-terminal domains of 2′-5′ oligoadenylate synthetases and the potential nuclease function of the C-terminal domain. J Mol Biol 326:1449–1461

Rogozin IB, Sverdlov AV, Babenko VN, Koonin EV (2005) Analysis of evolution of exon-intron structure of eukaryotic genes. Brief Bioinform 6:118–134

Salzberg S, Hyman T, Turm H, Kinar Y, Schwartz Y, Nir U, Lejbkowicz F, Huberman E (1997) Ectopic expression of 2-5A synthetase in myeloid cells induces growth arrest and facilitates the appearance of a myeloid differentiation marker. Cancer Res 57:2732–2740

Sarkar SN, Miyagi M, Crabb JW, Sen GC (2002) Identification of the substrate-binding sites of 2′-5′-oligoadenylate synthetase. J Biol Chem 277:24321–24330

Schmidt O (1864) Supplement der Spongien des Adriatischen Meeres. Leipzig 23

Schröder HC, Natalio F, Wiens M, Tahir MN, Shukoor MI, Tremel W, Belikov SI, Krasko A, Müller WEG (2008) The 2′-5′-oligoadenylate in the lowest metazoa: isolation, cloning, expression and functional activity in the sponge *Lubomirskia baicalensis*. Mol Immunol 45:945–953

Sperling J, Chebath J, Arad-Dann H, Offen D, Spann P, Lehrer R, Goldblatt D, Jolles B, Sperling R (1991) Possible involvement of (2′-5′)oligoadenylate synthetase activity in pre-mRNA splicing. Proc Natl Acad Sci USA 88:10377–10381

Venkatesh B, Kirkness EF, Loh YH, Halpern AL, Lee AP, Johnson J, Dandona N, Viswanathan LD, Tay A, Venter JC, Strausberg RL, Brenner S (2007) Survey sequencing and comparative analysis of the elephant shark (Callorhinchus milii) genome. PLoS Biol 5:e101

Wiens M, Kuusksalu A, Kelve M, Müller WE (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

ELECTRONIC SUPPLEMENTARY MATERIAL

Supplementary Table 1. Statistics of GeoOAS1Aa and GeoOAS1Ab differences

Exon No	Length		Mutations								
	GeoOAS1Aa	GeoOAS1Ab	Point mutations in codon position			Syn	Indels	Ts	Tv	All	All / Length (%)
			1	2	3						
1*	23	23	1	0	0	0	0	1	0	1	4.3
2	68	68	0	1	5	5	0	5	1	6	8.8
3	133	133	3	4	1	1	0	6	2	8	6.0
4	250	247	2	4	6	4	1	8	4	13	5.2
5	58	58	1	3	0	0	0	4	0	4	6.9
6	137	137	1	1	3	3	0	4	1	5	3.6
7	51	51	2	0	1	1	0	3	0	3	5.9
8*	198	198	2	1	2	2	0	3	2	5	2.5
Total	918	915	12	14	18	16	1	34	10	45	4.9

Intron No	Length			Mutations					
	GeoOAS1Aa	GeoOAS1Ab	Length**	Indels	Ts	Tv	All	All / Length** (%)	Ts+Tv / Length** (%)
1	50	42	43	1	0	0	1	2.3	0
2	178	174	158	7	10	2	19	12.0	7.6
3	66	67	67	1	1	1	3	4.5	3.0
4	55	62	53	2	1	3	6	11.3	7.5
5	53	51	52	1	2	3	6	11.5	9.6
6	59	60	60	1	3	2	6	10.0	8.3
7	186	187	187	1	2	0	3	1.6	1.1
Total	647	643	620	14	19	11	44	7.1	4.8

* - statistics based on the partial primary structure

** - contiguous stretches of nucleotides missing in the other gene (according to the alignment in Fig.1) are treated as single indels and added to the number of aligned nucleotides

Ts, transitions; Tv, transversions

Supplementary Table 2. Specification of genes in two OAS loci in *A. queenslandica*

Locus A

Gene number	Orientation	Title	Method	EST	Protein homology
1	+	Iqcd	EST initiated gene prediction	CABF10427	Score 199 bits to XP_001636992 predicted protein [*Nematostella vectensis*], NEMVEDRAFT_v1g93943, containing IPR000048 IQ calmodulin-binding region
2	-	Gats	EST	CAYI10006 CAYI10007 CABF7938 CABF13842 CAYI1437 CAYI3490 CAYI3848 CAYI8152	Score 232 bits to XP_001636192 predicted protein [*Nematostella vectensis*], NEMVEDRAFT_v1g163533, 213 bits to NP_109644 opposite strand transcription unit to Stag3 (GATS) [*Mus musculus*]
3	-	TYW1_ Radical_SAM+ Wytosine_form	EST initiated gene prediction	CABF20485 CAYI2711	Score 713 bits to XP_001188746 PREDICTED: similar to radical S-adenosyl methionine and flavodoxin domains 1 [*Strongylocentrotus purpuratus*]
4	+	AmpOAS1Aa	Gene prediction	-	Score 190 bits to CAC82933 putative (2-5)A-1 synthetase [*Suberites domuncula*]
5	+	AmpOAS1Ab	Gene prediction	-	Score 190 bits to CAC82933 putative (2-5)A-1 synthetase [*Suberites domuncula*]
3	-	TYW1_flavodoxin		CAYI5477	
6	-	Protein_117_aa	EST	CAYI3012	No evident homology
7	-	ORF_839_aa	ORF	-	Three non-contiguous hits (Score 76, 38 and 71 bits, respectively) to XP_001642127 predicted protein [*Nematostella vectensis*], NEMVEDRAFT_v1g237807, containing BTB/POZ domain in C-terminal region

>Iqcd
IGSYEESRVILYPSQKKLSSVESQRVMTVAKETRRKAEAALLLIPHLYDHLDRFAVTMGTELVARIREYHYLSEEYE
RCYESLIENEIEPDFATSRSKSSLPDHKLEPISSNMSPEELQIRFFSLQQQLRDLTRTLLRELQKCPSLDAIAKEVKGPV
PPRVDNVIGQLGEVVRLMEETLLTTHEEEVKRNEYLSIISERRQVAETEIASLQSELQHARNDLNKEVQKREDHRKK
LIHEIRVVDKLAKEVQKKIEDDTNKQEANSQKLHDGQVSKANEEIGILKKTLLEAKAANKEIESQLRKKSFKRETEA
DNWIQKYDNEIGDLQNELETLTEEFTEEKKQVKEIGERLEILSIEYDAVVKEQEEERKKAEEAQQRLASMVRAATQI
QSCWRSYKIRKMFSKDAKKGGKKGKKGGKKGGKKGSAKKKKA
>Gats
MSVRPALHLLPLELRIASIGKDDLPSFTYPLLQMTLGLLGQQSSCRQNHFFSFTETKNDYSLILDNELFKELIQFPGSE
ALQVAPESWRPLTIEVGAFGSLTGISKLVASVIGPLADEGVSVFCLSTNQEDYVMVKEKDLYKAMTCLHPCFKLLAE
LEQDVQLIENFSHDNVTNQIKNDLNYIELPSPRSITHPFTCTASKFHVCSILPSTLPSIAQPLLQLMFYNTRSDDVFLSL
SIISDDISMVLDARDIDKFPEDSVYSSEAYWKVITIGDGPLGFDECGIVAQVAAPLAQADISTYYICTFYNDHTLVPEG
SVTKALQLLNKYLMSSKVDNSSSPVSVGSLNGLTHPPSQFRTEPLPRLPLTNPVA
>TYW1_Flafodoxin+Radical_SAM+wytosine_form
MATTPYLFVDVFLDISKEINLSLLFICFIVLLASFYWLYRRPSFPSKVKRTGPLSVRVLYASQTGMGKRFAEQLMSEIK
SRGFHSSVSDVSHIDPDEEFTTPELSPCTLYVFLISTYSEGFPPDSARWVFQWLRESVDDCRVSKSLLSGLCFAVFGCG
NSLYKDHYNTAAKNLFDWLVQLSAKSVCQLGLGDANVSQSEHGGLSNDFAAWKNQFLESLERLRNESDENRTKA
SDHNSKFYSNQNGNESRQDGSESLRGVSYEEDLDGSSTDDEGVGSEDDVVDVEDIGDMMVKMKKASADNKRQNE
KKKREMVTPEIRQSLTKQGYRIIGSHSGVKLCRWTKSMLRGRGGCYKHSFYGIESHRCMETTPSLACANKCVFCWR
HHTNPVGTEWVWPMDGPERIAQEALANHYQMIKEYRGVPGVRGERLAEGMKVKHCALSLVGEPIMYPEINKLIGL
LHEKKISTFLVTNAQFPDAIRTLSPVTQLYVSVDASPKESLKKIDRPLFKDYWSRFIESLKALPDKGQRTVYRLTLVK

QWNTDELEGYAELVSLGQPDFIEIKGVTYCGESKASSLTMANVPWHDEVLTFVKELADLLPDYEMACEHEHSNCV
LLSHRKFFIDGKWHTWIDYDKFHELYNDTAGSKSFTALDYMAVTPSWSLIGSAERGFDPGETRWTRKSKRKAESGC

>AmpOAS1Aa

MEFPRDTCEDIPCKRNGVRQYLKEDTRLTPAEIIKGGSLGHGTAIPGDFDLDLVLYSRDIDPEVVLEIGVDNILDRLD
VFLRRRLPEAYVHVKKTSFSLQFKLNGVIDVDLLPSPYWGEDPDQFHRFLEEKGEKERMNFSCSAAKWQVKFFKEQ
PNEVKEYIKRFKAWRNLHWPKGQDAVGRPKSYLLGLLMIRAYEKSKDKSIHIKWNDGYYDWTYRLTREEFQSLLG
NRFTLISSSNGVDVVYDNKQQRQRNYYPQKVPSILDVANPFNDVYVSGIGNYHCKDPQGSFHPGTGKWSPLVGKID
TFSLSKPLEYYNN

>AmpOAS1Ab

MNKLTKNSQGHEDFPCKRKQYLKEDTRLTPAEIIKGGSLGHGTAIPGDFDLDLVLYSRDIDPEVVLEIGVDNILDRLD
VFLRRRLPEAYVYVNKTSFSLQFKLNGVIDVDLLPSPYWEDPDPFHRFLEGKDKKDKMNFSCSAAKWQVKFFKEQP
NEVKEYIKRFKAWRNLHWPKGQDAVGRPKSYLLGLLMIKAYEKSKDKSIHIKWNDGYYDWTYRLTHEEFQSLLG
NRFTLISSSNGVDVVYDKKQQRERTYTPAEVPSILDVANPFNDVYVSGIGNYHCKDPQGSFHPGTGKWSPLVGKIDT
FSLSKPLEYY

>Protein_117_aa

MNNTTNGDFLQDVESFLYTPLGLGALAGATALLTSCLWLICCSVACCCIKWRNKRRESGTIEAGADLKYMTCTEFR
PPHYSSATSTENGYHSNAGTLSSLNRHDALAIHSSMDAILN

>ORF_839_aa

NINGKAVLSRAGGIELLTQCLAEQENKGEKEKGQSSLIIDALCHCCRDVHGRQKVRECGALQVLINLLQSKDHSSFH
HDILSALICYYFDEQTLQFMIRRLGLIKALLYHLKKGIKEEKDEGRREEREHNEEGEVRKMRKRKKRDQKGGRLVK
RLRQQDDPDIFLKNEESFGSVEKSPGSSTKGSEDDINIDEGSGVDIEGDSLDEASDCNEEEHPIVKKSKRKKNEAKKK
LSDQEKPSNSNNDLPTGTYHKDTEPPSKKLSRTISDEESTSSLYYLSPPYNMEDSGGGRGSFSPTRETDLAHSPLFEKL
SPLLNINEDEASSYSATSPLSKGTFSPSCQAALEANNTLSSPVPTNFIDSLLSPTSCSPVAPSKTIKRSSSHEIGSNHSSSP
HSKVLLLLSRVSHLHDCQPILASVDVLSVIMDYYLTMNLKDGHCFKVLSRLFSNPHCFQDCLVTLAPSMCFHHMIM
TDDNGGVTNAQCNATSYDRRIATSNTQSTSYQRAMGYPSPPIGDLSHSICHKMLTEYGGLHNLLCLSKGQLLTTQCS
CSQNTSTIEDTELKELTPSPPLPSSKVSDQETLTQLMAIDTLLYLGSQYRKCPSRKRRHSAIERRSSATLPLPIKAGTISP
PPLSSPLCLYRDFDHWPFDTCIILDDGTMFPVHRRVLVESSEVFSVMLSGCYKESTDSSVSLHSVCPSSFRSLVHHAY
GCGSHCWGRVSEKEAIDEHYDDEVIDEIVASIKEKEGQERGKHLLRVLISANQFFMCSLLTAAQESFMLYLSPKNVT
PLFLFSQLHQCSLLSDATVQLLVGRGPGFCQSESFLRLVRSIECDEFLSTITTVFHQK

Locus B

Gene number	Orientation	Title	Method	EST	Protein homology
1	+	PDHE1alpha	EST	CAYI10060 CAYI3314 CAYI5926 CAYI10023 CAYI5668 CAYI2776 CABF10208 CABF8701	Score 495 bits to XP_001175746 PREDICTED: similar to Probable pyruvate dehydrogenase E1 component alpha subunit, mitochondrial precursor (PDHE1-A) [*Strongylocentrotus purpuratus]*
2	+	ORF_714_aa_ Ank_rep	EST and ORF	CAYH3434	Score 134 bits to XP_420641 PREDICTED: similar to ankyrin B (440 kDa) [*Gallus gallus*]
3	-	DUSL	Gene prediction	-	Score 400 bits to XP_00169015 predicted protein [Nematostella vectensis], NEMVEDRAFT_v1g199771, containing Dihydrouridine synthase-like (DUS-like) FMN-binding domain. 371 bits to XP_001231915 DUS4L [*Gallus gallus*]
4	+	Protein_84_aa	EST	CABF15249 CABF11028 CABF16022 CABF26945 CABF11133 CABF19932 CABF22767 CABF10631 CABF17790 CABF26711 CABF6815	Score 48.9 bits to ABK26429 unknown protein [*Picea sitchensis*], 40 bits to NP_565726 expressed unknown protein [*Arabidopsis thaliana*]
5	-	ORF_707_aa_ Ank_rep	ORF	-	Score 151 bits to XP_001184164 PREDICTED: similar to ankyrin 2,3/unc44 [*Strongylocentrotus purpuratus*]
6	+	ORF_702_aa_ Ank_rep	ORFs	-	Several ORFs with ankyrin-related matches, the longest ORF - 702 aa, Score

					251 bits to XP_001649474 ankyrin 2,3/unc44 [*Aedes aegypti*]
7	+	Acat2	Gene prediction	-	Score 537 bits to XP_001630395 predicted protein [*Nematostella vectensis*] NEMVEDRAFT_v1g169195, acetyl-CoA acetyltransferase; PRK07491
8	-	PITRM1	EST initiated gene prediction	CABF23656	Score 838 bits to XP_001638709 predicted protein [*Nematostella vectensis*], COG1026 Predicted Zn-dependent peptidases
9	+	AmpOAS1B (variants)	Gene prediction	-	Score 140-141 bits to CAC82934 putative (2-5)A-2 synthetase [*Suberites domuncula*]
10	-	ORF_678_aa	ORF	-	CDD hit 38.8 bits to smart00256, FBOX, A Receptor for Ubiquitination Targets (PssmId 47585)
11	-	PITRM1_ truncated	EST initiated gene prediction	CAYI2358 CABF1823 CABF15589	Score 197 bits to XP_001181181 PREDICTED: similar to Pitrilysin metalloproteinase 1 [*Strongylocentrotus purpuratus*] (stop codon in exon 6 leave protease domain incomplete; 3'-terminal parts of ESTs are PITRM1 non-related
12	+	AmpOAS1C (variants)	Gene prediction	-	Score 159-162 bits to CAC82934 putative (2-5)A-2 synthetase [*Suberites domuncula]*
13	+	Aldo/ketolase 2	Gene prediction	-	Score 306 bits to XP_001608172 PREDICTED: similar to CG18547-PA [*Nasonia vitripennis*], Aldo/keto reductase superfamily member
14	-	Aldo/ketolase 1	EST initiated gene prediction	CAYI9950	Score 301 bits to XP_001195433 REDICTED: similar to ENSANGP00000008370 [*Strongylocentrotus purpuratus*], Aldo/keto reductase superfamily member
15	+	Coiled-coil domain ATPase	EST initiated gene prediction	CAYH6884 CAYH1726 CAYH4066	Score 105 bits to EDL30030 coiled-coil domain containing 21, isoform CRA_b [*Mus musculus*], containing conserved domain Smc - Chromosome segregation ATPases [Cell division and chromosome partitioning]

>PDHE1alpha
MAALIRGVRNLLPRGLVVARCASSSAEFSIPEYKLHKLDSGPSTTTTITRDELMKLFRDMTVIREMETEAKNLYLKKI
IKGFCHLYIGQEACAVGMEASITKDDPVITAYRAHGWTYTRGIEPLGILAELTGKKSGCAKGKGGSMHMYAPNFYG
GNGIVGAQVPLGAGIAFELKYNNKDNICITLYGDGAANQGQVFEAYNMAALWKLPVIFACENNRYGMGTAVHRSS
ASTDYYTRGDYIPGVYVDGMDVLAVREATRWAKEYILAGNGPLVMELDTYRYYGHSMSDPGKSYRKTEEVQQFR
KEKDPITTATRYLIQGDLATEEELKEIKKSVQADIKKAVANAISDTELPLEEMYTDIYTSTPEFMVRGCDPFTWGKSE
RA
>ORF_714_aa_Ank_rep
IATPPLPLSMSVSGSSSPVLSVSLLLSKLLPVVDWEVLGLHLNIPKHELEKIRRQFFHEGVERCKAEMLDLWLKTSNE
EALHWNVISDALTKMRGLEELAQQIRTMGPSQEERMSQEEATLSSEESPVTKVEFQFKKLDLKLFSKLESKFAKLVS
DILQELKSRGVSVNELHSYVKIRLELKSSSSPQSLDELLLDKLKPYYCLTNLTLIDNIICEFLDGSKIQENLEEYEKQLE
SFKTSAKMEELVDHMTSSLEGKGEEGTVSVVLKLEGSWLGVTLKHFQQLVEEIFLERSCCFNHIQVKRGCICVSMIV
PEAMTSSLVSLAKQRIEFINTIGIIRLAIGGNEIMDKEEEKDEASFSQYLMRGVLQEKIDVVRFILSLDGLGLADTISPCI
EEAVESGNIQILSMLLESGGDPNFIGNKGLTPLQTASKQGHVTIVNKLLQYKADPNQDGVSGWTALMIATEKNNTSII
ELLIKAGANLNAKMVHGWTPLMLACAYNFPALVSILLKAGADPNIRTIKWAAVHSAASKYEDPSILIDLAESGVNL
NAVTGAGATPLILACQVNFELTINVLLNYKVSLDSQTRSGYTALMAAVCNKSVSIVESLLSAGANVHIADANGQTA
LDWALDSGHPEITELLLVYMGDMPATPLKAPSSAPSVTLVPEDIEVHIPSSSSAILELRDALENPLPPQGTSLKHKSSK
EEKEKSRNTTAY
>DUSL
MAEREEENAKNIMDLFQPGAVVRMCAPMVRYSKLPFRKLVRRYGTQVAFTPMIVSESFVRSQKSRDVEFTTDPLDD
RPLVAQFAANNVEDFSRASQLVSPFVDAVDLNCGCPQRWAMEEGYGSYLITRPELVCDMVRGAKRVAGIPVSIKIR
IHEDLRETVELARRCEVIGLSWLTVHGRTVKQRAEPVNLSAIKLIKESVHIPVIANGDVRNESDINRTVQETGVDGV
MAARGMLQNPAMFGGYDETPIECLEDWVDISLRSGLTFTSFHHHCMYMLERVQAKTERRVFNTLNSTPAVLEFLK
HHYNIQHK

>Protein_84_aa
MASQKWISRPEKPPEKWATIDRSNAIEAWGRMRETTNQHFKFTPRTTAYCVIWALLVPIGIWKIIKWERQYKDRAK
GRPPKDLF
>ORF_707_aa_Ank_rep
SYKRMSEVTKVSLPLLLSKLLPVVDWEVLGLHLNIPKHELEKIRRQFFHEGVERCKAEMFDLWLKTSNEEALHWNV
ISDALTKMRGLEELAQQIRTMNPSQEEMKDRVDSVASKSIQQATLVNLSVKVQKPCLKRFSQLESKFAKLVSDILQE
LKSRHVSVNEFHSYVKIRLELKLSSSPQSLDELLLDRLKPYYCLTNLTLIDNIICEFLDGSKVQENLEEYEKQLESFKT
STKMEELVDHMTSSLEGEGEEGTVSVVLKLEGSWLGVTLKHFQQLVEEIFLERSCCFNHIQVKRGCICVSMIVPEAM
ISSLVSLAKQRIEFINSIGIIRLAIGGNEIINKEEEKDEASFSQYLMRGVLQEKIDVVHFMLSLGGLGLADTISPCIEEAV
ESGNIQILSMLLESGGDPNFIGNKGLTPLQRASKQGHVPIVNKLLQYKADPNQDGIDGATALMLASNFDHFGIVEQL
LSSGASVNKTMFDGWTALMFACDQNHLQIVKRLLQSGANPNVLTHSSKCSPVYFASSKNNNPSMLLTLIETGADINI
ATETGVTPLIIAAQFGFEAIVSVLLSHKALLDCQTHHGFTALMAAAHKGHVDITESLLLAGASTHIADNNGRTALEW
SLQSEQHEITQLLLMNVGDVIDESDHAYNAEYQYGGKRISEDISSPKLHPSIIKLREAMRKPPLSEDTGQKHRHAIRW
NNEHLVKK
>Acat2
MSLERGVFIIGAARTPVGSLNGSLSSLSAHQLGSTAITAVLERAKVLPEKVSEVLVGQILTAAQGQNPARQAAMLSGI
PKEVPSTSINMLCGSGLRTIAMGYQSILVGDADIVVAAGQESMSQAPHAVHLRSGVKFGDASLTDTMMKDGLTDAF
HNYHMGQTAENVAKQYGISREEQDRFAVSSQNKTEESQKNGVFKDEIVSVEIKTRKGPVTVDTDEFPRHSTTIEGLA
KLKPAFVKDGSGTVTAGNTSGINDGAAAVVLASKEAIDGLSCSLPLARIVSWAQCGVDPQIMGMGPVPAVSKALQ
KAGWNKDDVDLFELNEAFAAQSLAVVKELNIDEKKVNVGGGAIALGHPIGASGCRILVTLLYALKRVGGKRGVAS
LCVGGGMGIALCIEMC
>PITRM1
MAAQLFSAPRLLSRRRRLIHLNNTVFNSTRREDVSKLSVAINDRISGYRVTGVTEVSDFNLKAIELLHESTKAQHLHL
SRPESNNVFGVGFRTTPMDSTGISHILEHTTLCGSAHYPVRDPFFKMLTRPLATFMNAFTANDWTFYPFSTQNYNDY
RNLLSVYCDCVFHPNLKKMDFHQEGWRLEHEDPNDRNTPLVFKGVVFNEMKGMFSNAEYLFDTELQRKLLPSHTY
SHVSGGAPLNILDLTWDSLKEFHAKHYHPSNSYFYTYGDIPLSDHLSFIDENVLSKFSVSTPAAAIPLEARWNQPRSH
EIYCAPDPMAPDPNKQTTVGVAYLLGPITDSYEGLLLSILTKLLIDGAASPFYQSLIDSNIGSDYSPSTGYNDGTKDSY
FSVGLQGISTEDAEKVKQIITETFHKVYQEGFPEERVKAILHQVEIGLKHQSSNFGLALLMHFITPWIHGASPTSLLQI
NQNVENFKDHMANDPNYLKEKIKEYFIDNKHNLTLVEYNKKQKELEEEKLRLLVSSLSEDDKQRVYDEGLQLQED
QNKHQTSDCLPSLSVADVDKTTPITELIHEDHGSIPIQYCVQPTNGISYLQFLSSCRDLPDDLKPYLPLFCSVITKMGT
QDLDYRQLSQQIDLYTGGSVCVGTHVASHHSIPNQFEQTFSTSNKLPQLFQVLSLSLILCSRPQLQDEQRLQTLIAME
ASSLAVNVSRSGHRYAMTASASSLSPAAMMAEKYGGISQVLIHRSIIAFNVTEPQFTVASEAAEGFLDNLPGSITDTP
LFMEEGTFTPTLSHTHYQLPFQVYYVGQSYRTVPYTHEDSPSLRVLAQLMGWKYLHKEIRYSLVVSCFVLNKQFLE
GDPHSLDTVNRFNDSIDWVTSRSFTDEDVEEAKLSVFSQ
>AmpOAS1Ba
MAASQLMASCVSLLRHVVLFIVIENVLAMQPFRAHLHKDAGIVVSEIIQGGSLGQGTCVPMKFDVDLVIFSKSITHE
MISNPDKRKEILTKVQQCLIRRLKRNYVDDKMTQNSVQFKYKHSPGHIIDVDVLTSPMWKSAGELSIYIAEHCPDQY
YHLSAGASKWQTDFIRHQMPRVKEYIKRAKYWKKSIDWSPGREPSSYLMSLLVIKASDLVGARHYFYICSVSWKIK
ELVKNHHTIDIQFQTEYMFTVEAPSQPRVIDPANVCNNLYETGFGKQGKRDEPQDFWNVFADKIEGLDLSIPIERLGP
LP
>AmpOAS1Bb
MAASQLMASCVSLLRYFVLFIVIENVLAMQPFRAHLHKDAGIVVSEIIEGGSLGQGTCVPMKFDVDLVIFSRSITHEM
ISNPDERKEILTKVQQCLIRRLKGNYVDDKMTQNSVQFKYKHSPGHTIDVDVLTSPMWESAGELSIYISEHCPDQYY
HLSAGASKWQTDFIRHQMPRVKEYIKRAKYWKKSIDWSPGREPSSYLMSLLVIKASDLVGEHLLACMEHYYCFAL
CSISEKIKELVKSHHAIEIQFQTQYMFTTEAPAQPRVIDPANVCNNLYETGFGKQGRRDEPGDYWDVFAKKIESLDLS
IPFERLGPLP
>ORF_678_aa
KHVYSRVSGHRHCLDYRNPHHTMAEAEPCCKRLKFDVSFFDDESDDDFYHSEPPIDQSTANQLPLIEDEEYTNPFDS
LPNEVIQEICMMLSLRDKMSLRLVNRRLYMICSDPYLWRNVVIDDAYHKTNAPFIKSALQTCRPHVQSLSLRGELPF
SKYQRMILTCKNIHTLNLYGFQISFNALEKIHSGLHHLQFLSLTYIDQIPLSNDRFFIIIAKLKKVALVYNNFEAEKLFE
KWLLNNCLPQIFIIISAHGKEYWSRSRNIHHLPNITHSAYFAAYRKFCQPLDFDFYDVPEYSLEIGPNSSESVAVTANG
KLMITMRDMITPINNDVNKRYAVFKDEAVTPGLSVYTSQYGVNITVLGFSSVTVSLESFRMIVKETPNILEVSLKGSII
SDSLDAYIVPLFVHCLKLRGLNVSCFSTSSRDSINFDMEHFWNLLSKMKYLESLSVDCCSLSPLTPNSQHDGRSARV
LTDRELAVKERVIGHIKRMTRLKSLYVIETIEIEELDYFINQHLLSMISNLESLLYLEINISYAADIASLYSCNMNFVEG
LERILQKCQKLSFLVINFTSVKAIELPVDPALYGNLTHLWLGCQHGPITVEFGSALAINSKKKLKHLLLGHRYFDIEL
DVLDKINEGNLITFYELGSDKCFYRHDRNTCKKRGKLSELDFYIGKFPFLH
>PITRM1_truncated
MAAQLFSLPYVRRGIRLIHLSNPVFNSARREDVSKLSVAVNDRINGYRVTGVTDVPDFNLKAIELLHESTKAQHLHL
ARQDSNNVFGVGFRTTPMDSTGICHILEHTTLCGSAHYPVRDPFFKMLTRSLATFMNAFTANDWTFYPFSTQNYND
YRNLLSVYCDCMFHPNLKNMDFHHSIHHQ
>AmpOAS1Ca
MANPMPHVGTLSRGQVAQKVRLSIIAVRAVCSLLKQHFRTDSGIAVSEIVEGGSLGYGTYVPMKFDIDLVIYSTSLT
ADKLCDRDKRKKFLSDLQECVKVHLKGNYIADKMTENSVQFKYQFSPDNTIDVDVLISPMWGSPVELSQYVSKNC
PTEFFHLSAGAAKWQKKFMSHQTLRVKEYIKRAKYWSLTVNWEAIGGRKPTSYMMSLLVIKAAEPFEIHMKGRLR
QNFKKQLAKEITKNVKTLVKKFETIEILFDSDYPFPNALLPSPPRVVDPANICNNLYETGFGKLRRDEDSEKYWKKFA
ENIDTLDLSVAFEHRGFYSDVPAFGKSYQPKY

>AmpOAS1Cb
MANPMPHVGTLSRGQVAQKVRLSIVAVRAVCSLLKQHFRTDSGIAVSEIVEGGSLGYGTYVPMKFDIDLVIYSISLT
ADKLCDRDKRKKFLSDLQECVKVHLKGNYIADKMTENSVQFKYQFSPDNTIDVDVLISPMWGSPVELSQYISKNCP
MEFFHLSAGAAKWQKKFMSHQTLRVKEYIKRAKYWSLTVNWEVIGGRKPTSYMMSLLVIKAAEPFKIHLKGKLTR
YLRRQLPKDITENVKTLVKNFETIDISFDSDYHFPNALLPSPPRVVDPANICNNLYETGFGKLRRDEDSEKYWKKFAE
NIDTLDLSVAFEHRGFYSDVPAFGKSHQPKY
>AmpOAS1Cc
MANPMPQVGTLSKRQVIKKGKQSRRAVRTVCSLLKEHFRADCGIAVSEIVQGGSLGYGTYVPMKFDIDLVIYSTSLT
ADKLCDGDGRNELLLALQECIKIHLKKKYIEDKKTENSVQFKYQFSPDNTIDVDVLISPMWDDPVELSQYISRNCPN
EFFHLSAGAAKWQKQFMCQQTSRVKEYIKRAKYWSLTVNWEKVEGRKPTSYMISLLVIKATEDFKDRMTGKLGKS
VKRKLARRDWRQSSKNLLSLMILEIIFGSVFFESVSDFPIALLPPLPRVVDPANICNNLYKTGFGKLRRDEDSEKYWK
KFAENIDTLDLSVAFEHRGFYSDAPTFGRSYHPNY
>Aldo/ketolase 2
MEASCKLLYKEIGSTGLRLCCLSLGGSGYGGIYGEYDEKKAEEALRSSLDKGVNYIDTSPFYGEGRSESFLGRVLSR
VPRDKYYIGTKVGRYFWDVKKRFDFSRETILKGFEESLKRLQLEHVDILQLHDVEFAPSLDIIMNEALPAIQELKDKG
FCRAIGITGYPIGPLKEIIAHSHHIKIDSVLSYARLTLNDNSLKDHFDFFKSHGVSIINASPVSMGLLTSEGPQWWNPAL
PYIKDKCKEAVEYCNSKGVDITRLAVNYSTSFDEVVTTLVAISDTQMLDRNLSAVLSPMTEKEKKVAGDLEKFFSRL
PQRHWEGVELKEYWETNWKL
>Aldo/ketolase 1
MATASRSLPMLYRDLGRTGMRLSIVGLGGSGWGKNYGPYDEKEAVRAFHHALDRGVNYIDTAYWYAQGHSEKFL
GQHLLPTVSRDRYFIATKVGRYELDLPNMFNFTAERVTKSLEESLRRLQLSEVDIIQIHDVEFSLSVNQILHETLPALE
RLKEKGLCRHIGITGYPLGPLKEVISQSSVQIDSILSYCRLSLNDWTLKNEFQFFKENKVPIINACPVSMGLLTEGPVQ
DWHPSLPWIRDACNRAVQYCREQGVDITKLALHYSTSFEEAVTTLVSMCDTEMVDKNLDSALSPLTPHEEGVMQH
LLDNVFGDMTQRHWEGVELVSYWNATRK
>Coiled-coil domain ATPase
KEDKEKDTRPICPFSSPPSSLLLFEIIVQSSVPSLFRNCPSTEGIMSSQEESGSSLNNPSLHFPVPVTSKDATTPDEETRSL
QPPTTREQPQAVHESIPHKPPPTAKVAPLVGGSSSAAIITSQRSTTPHEGTSVVGTPRRKGNKLDVLSSSFSCVSLSPTR
TGGGSERSSPVKELLHADRVRNLMSLLTEKSQIIEKQKNTITRLQQDCKKKDSQLRLLTLPQANEGNVGGEANVGLS
APVKLDILSRKLDEAYYELEKAKSESDVMRDKISDKDSHISQLETDADRYKAKIKLLEDKIVKLEGYFGNTPTMEEH
YSMKEQVTELHSQNILLSERLREKEDKIMAQIDELRDKETTLKEMSINIRDLKQDNSDLRLQCETSDRRLVELHSDIS
QYKIELMKSKQQITDLEEKNRQDLRDMRERHKLRFQEATAKLKEQYKTVTKRNRQLELQLLKNSDVMTALNNELK
LSQGTVADLKSSFGDLVKQNQILMEKNYGLQEPPEVGRSGIVSVSPVSDGSDESLQALSQEVSQCVEEVDSLVSLAS
SIFEGKEPDMSLLLGGGATGNGDHKTDSELDPARKYEQQLTQVQKNRKEIERLREKVTDHFASKVANECHLQ

PUBLICATION III

Päri M, Kuusksalu A, Lopp A, Kjaer K, Justesen J, Kelve M. (2014). Enzymatically active 2',5'-oligoadenylate synthetases are widely distributed among Metazoa, including protostome lineage. *Biochimie*, 97:200-209

Biochimie 97 (2014) 200–209

Contents lists available at ScienceDirect

Biochimie

journal homepage: www.elsevier.com/locate/biochi

Research paper

Enzymatically active 2′,5′-oligoadenylate synthetases are widely distributed among Metazoa, including protostome lineage

Mailis Päri [a], Anne Kuusksalu [a], Annika Lopp [a], Karina Hansen Kjaer [b], Just Justesen [c], Merike Kelve [a,*]

[a] Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15, Tallinn 12618, Estonia
[b] Department of Molecular Biology, Aarhus University, C.F. Møllers Allé 3, 8000 Aarhus, Denmark
[c] Department of Molecular Biology and Genetics, Aarhus University, C.F. Møllers Allé 130, 8000 Aarhus C, Denmark

ARTICLE INFO

Article history:
Received 6 September 2013
Accepted 22 October 2013
Available online 31 October 2013

Keywords:
2′,5′-Oligoadenylate synthetase
Nucleotidyl transferase
dsRNA
Choanoflagellate
Porifera
Protostomia

ABSTRACT

2′,5′-Oligoadenylate synthetases (OASs) belong to the nucleotidyl transferase family together with poly(A) polymerases, CCA-adding enzymes and the recently discovered cyclic-GMP-AMP synthase (cGAS). Mammalian OASs have been thoroughly characterized as components of the interferon-induced antiviral system. The OAS activity and the respective genes were also discovered in marine sponges where the interferon system is absent. In this study the recombinant OASs from several multicellular animals and their closest unicellular relative, a choanoflagellate, were expressed in a bacterial expression system and their enzymatic activities were examined. We demonstrated 2-5A synthesizing activities of OASs from the marine sponge *Tedania ignis*, a representative of the phylogenetically oldest metazoan phylum (Porifera), from an invertebrate of the protostome lineage, the mollusk *Mytilus californianus* (Mollusca), and from a vertebrate species, a cartilaginous fish *Leucoraja erinacea* (Chordata). However, the expressed proteins from an amphibian, the salamander *Ambystoma mexicanum* (Chordata), and from a protozoan, the marine choanoflagellate *Monosiga brevicollis* (Choanozoa), did not show 2-5A synthesizing activity. Differently from other studied OASs, OAS from the marine sponge *T. ignis* was able to catalyze the formation of oligomers having both 2′,5′- and 3′,5′-phosphodiester linkages. Our data suggest that OASs from sponges and evolutionarily higher animals have similar activation mechanisms which still include different affinities and possibly different structural requirements for the activating RNAs. Considering their 2′- and 3′-specificities, sponge OASs could represent a link between evolutionarily earlier nucleotidyl transferases and 2′-specific OASs from higher animals.

1. Introduction

2′,5′-Oligoadenylate synthetases (2-5A synthetases, OAS, EC 2.7.7.84) belong to the nucleotidyl transferase superfamily, being grouped into a subclass of template-independent polymerases together with poly(A) polymerases (PAP) and CCA-adding enzymes. It has been proposed that OAS diverged from PAP and CCA-adding enzymes at the beginning of metazoan evolution [1]. OAS became 2′-specific, catalyzing the formation of 2′,5′-phosphodiester linkages instead of 3′,5′-linkages, and acquired the ability to initiate synthesis from ATP without requiring a primer [1,2]. In contrast to constitutively active nucleotidyl transferases OAS needs dsRNA binding for the assembly of its active site [3]. Recently a novel member of nucleotidyl transferases, able to catalyze the formation of 2′,5′-linkages, cGAMP synthase was identified [4,5] and its activation by DNA-induced structural switch was demonstrated [6].

Since their discovery mammalian 2′,5′-oligoadenylate synthetases have been extensively characterized. In mammals they comprise a part of the interferon-induced antiviral system. Upon binding to dsRNA that may be of viral origin, OAS proteins undergo a conformational change and become activated [3]. The enzymatically active OAS polymerizes ATP into 2′,5′-phosphodiester linked oligoadenylates (2-5A) with a general formula $pppA(2'p5'A)_n$, where $1 \leq n \leq 30$. The 2-5A oligomers that have been synthesized activate, in their turn, the latent RNase (RNaseL), which then degrades viral and cellular RNA and thus limits the viral infection [7].

Unlike many RNA-binding proteins, OASs do not possess a typical RNA-binding motif [8]. They are able to bind different RNAs

Abbreviations: 2-5A, 2′,5′-oligoadenylate; cGAS, cyclic-GMP-AMP synthase; dsRNA, double-stranded RNA; OAS, 2′,5′-oligoadenylate synthetase; OASL, OAS-like protein; PAP, poly(A) polymerase; RNaseL, latent RNase.

* Corresponding author. Tel.: +372 6204432; fax: +372 6204401.
E-mail address: merike.kelve@ttu.ee (M. Kelve).

0300-9084/$ – see front matter © 2013 Elsevier Masson SAS. All rights reserved.
http://dx.doi.org/10.1016/j.biochi.2013.10.015

but there is no correlation between binding affinity and the ability of the dsRNA to activate OAS [9]. By testing a pool of short dsRNA molecules with specific sequences, a sequence motif essential for the activation of OAS, NNWWNNNNNNNNNWGN, was proposed by Kodym et al. [8]. With certain assumptions, the decisive features of this sequence are in accordance with the most commonly used *in vitro* activator, synthetic poly(I)·poly(C) [8]. Recently published crystal structure of human OAS1 in complex with dsRNA and dATP revealed the activation mechanism of 2′,5′-oligoadenylate synthetases. dsRNA binds to the protein surface at the junction of the N- and C-terminal lobes and induces a major conformational change in the N-terminal lobe of the protein. The rearrangements narrow the active site and bring together the catalytically important aspartate residues to enable substrate binding [3].

The 2′,5′-oligoadenylate synthesis was first described in interferon-treated mouse cells [10]. A few years later, partial cDNA corresponding to 2′,5′-oligoadenylate synthetase was cloned from human cells [11]. The human oligoadenylate synthetase family consists of four genes encoding four types of proteins with different isoforms – the small form (p42/p46/p48, OAS1), the medium form (p69/p71, OAS2), the large form (p100, OAS3) and the oligoadenylate synthetase-like protein (p59, OASL), which contains C-terminal ubiquitin-like domain [7]. The murine OAS gene family was shown to include at least 11 genes encoding the four different types of OAS [12,13]. Thereafter 2′,5′-oligoadenylate synthetase gene families from other mammals (horse, rat, cow, pig and dog) were characterized [14,15]. A single 2′,5′-oligoadenylate synthetase gene (OASL) was found in chicken [16]. OAS genes have also been cloned from evolutionarily lowest animals, marine sponges *Geodia cydonium* [17], *Geodia barretti* [18] and *Suberites domuncula* [19] and a freshwater sponge *Lubomirskia baicalensis* [20]. However, the medium (OAS2) and large (OAS3) isoforms are only present in mammals, while small (OAS1) isoform and also OAS-like proteins (OASL) are widely distributed [21].

Continuously growing genomic and EST databases give us a chance to identify new OAS genes in various organisms while the fully sequenced genomes allow us to conclude in which organisms the OAS genes have been lost during the evolution. As an example, the sequencing of the elephant shark (*Callorhinchus milii*) genome revealed the presence of the OAS gene in this organism [22]. Putative OAS genes are also present in genomes of two other cartilaginous fish species, dogfish shark (*Squalus acanthias*) and little skate (*Leucoraja erinacea*) [21]. Three sequenced teleost fish (two pufferfish species *Takifugu rubripes* and *Tetraodon nigroviridis* and zebrafish *Danio rerio*) have evidently lost the respective genes [22]. No OAS genes have been found in the sequenced genomes of insects *Drosophila melanogaster*, *Aedes aegyptii*, *Anopheles gambiae*, *Apis mellifera*, the nematode *Caenorhabditis elegans* and the frog *Xenopus tropicalis* [21]. Based on bioinformatic analysis, the presence of OAS genes was predicted in several other evolutionary branches of multicellular animals (phylums Chordata, Annelida, Mollusca, Cnidaria) as well as in unicellular organisms, the choanoflagellate *Monosiga brevicollis* and the green algae *Chlamydomonas reinhardtii* [21]. Search for oligoadenylate synthesizing activity has previously also detected 2-5A synthesis in reptiles but not in amphibians, fish or insects [23]. To date, the OAS proteins with enzymatic activity have only been cloned from mammals [12,24,25], a bird [16] and a sponge [26].

In this study several previously bioinformatically predicted 2′,5′-oligoadenylate synthetase genes from various organisms throughout the tree of evolution [21] were expressed in the bacterial expression system to investigate their enzymatic properties and to verify their belonging to the oligoadenylate synthetase family of proteins. The salamander and little skate OASs were chosen to represent the evolutionary branches of vertebrates where the OAS activity had not been detected earlier. In the protostome lineage some branches have lost OAS genes and the mollusk as a representative from this lineage with the OAS genes preserved in the genome was of high interest. Furthermore, no OAS activity had been previously demonstrated in this branch of the evolution. An OAS from the marine sponge *Tedania ignis* was specifically selected to represent the phylum Porifera as it is the synthesis of both 2′,5′- and 3′,5′-linked oligoadenylates from the substrate ATP that is observed in the crude extracts of several species of the genus Tedania [27]. The unicellular eukaryote *M. brevicollis* was of particular interest since it is the closest relative to Metazoa and the OAS activity has been believed to be related only to Metazoa. The enzymatic properties of the expressed recombinant proteins were compared to those of the similarly produced porcine OAS which was selected as a representative of the mammalian OAS1 proteins.

2. Materials and methods

2.1. Materials

The specimen of the marine sponge *T. ignis* collected from the Bahamas was generously provided by Dr. E. Richelle-Maurer (Université Libre de Bruxelles, Belgium) in the form of frozen cell pellets.

The cDNA clone containing the OAS sequence (see Supplementary material) from the choanoflagellate *M. brevicollis* (MbrOAS) was from Dr. Nicole King (University of California, Berkeley, USA), the OAS cDNA clone from the mollusk *Mytilus californianus* (McaOAS) was from Dr. Andrew Gracey (University of Southern California, Los Angeles, USA), the OAS cDNA clone from the salamander *Ambystoma mexicanum* (AmeOAS) was from Dr. John A. Walker and Dr. Randal Voss (University of Kentucky, Lexington, USA) and the OAS cDNA clone from the little skate *L. erinacea* (LerOAS) was from Christine M. Smith (Mount Desert Island Biological Laboratory, Maine, USA).

The recombinant BL21 (DE3) *Escherichia coli* bacteria containing the expression vector pET9d with the porcine 2-5A synthetase cDNA were from Rune Hartmann (University of Aarhus, Denmark).

2.2. Cloning the putative OAS cDNA from the marine sponge T. ignis

The RNA was extracted from the cells of the marine sponge *T. ignis* with TRIzol® Reagent (Invitrogen) according to manufacturer's protocol. The cDNA was synthesized with the RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas) using oligo(dT)$_{18}$ primer.

A part of the coding sequence of putative OAS was amplified using degenerate primers SpOAS_uniF1 (5′-GTAAAACGACGGCCAGT**GAATTC**GGHTCNHTRGGDYANGGWAC-3′) and SpOAS_uniR1 (5′-CAGGAAACAGCTATGAC**AAGCTT**YGYHGGRTYIGCIGGRTC-3′) containing a restriction site (in bold, EcoRI in forward and HindIII in reverse primer) and the M13/pUC forward and reverse sequencing primer sequences (underlined) to allow further specific amplification and directional cloning of the PCR product. The degenerate part of the forward primer was designed on the basis of the published amino acid sequences of OAS proteins from the marine sponges *G. cydonium*, *S. domuncula*, *L. baicalensis* and *Amphimedon queenslandica* [28]. The degenerate part of the reverse primer was from Ref. [20]. Initial PCR using degenerate primers was carried out as follows: initial denaturation at 95 °C for 2 min, followed by 35 cycles at 94 °C for 40 s, 40 °C for 40 s, 72 °C for 1.5 min, and a final extension step at 72 °C for 10 min. Secondary PCR for amplification of the product using M13/pUC sequencing primers was carried out

under the same cycling conditions, except the annealing temperature that was 50 °C. The PCR product was excized from gel, purified using GeneJET™ Gel Extraction Kit (Fermentas) and sequenced. The sequencing reaction was performed using BigDye Terminator v3.1 Cycle Sequencing chemistry (Applied Biosystems) and the ABI PRISM 3130 Genetic Analyzer (Applied Biosystems) was used for the analysis.

Specific primers were designed on the basis of sequencing results and the 3′-end of the cDNA was amplified using the FirstChoice® RLM-RACE Kit (Ambion) and the gene specific primer TigOAS_F1 (5′-AACAGTGCCAGGTGACTACGAC-3′). For establishing the sequence of the 5′-end, the protocol from Ref. [29] was adapted. First, the cDNA was synthesized with the RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas) using the gene specific primer TigOAS_R1 (5′-TGGATAGTCCTGCTTCCTGTAG-3′). The cDNA was purified with the PCR Product Purification Kit (Qiagen) followed by RNaseA treatment. Poly(A) tail was added to the cDNA with Terminal Deoxynucleotidyl Transferase (Fermentas) according to the manufacturer's protocol and the tailed cDNA was purified again with the PCR Product Purification Kit (Qiagen). The second strand was then synthesized using the 3′-RACE adapter from the FirstChoice® RLM-RACE Kit (Ambion). The PCR was carried out in 50 μl final volume containing 5 μl tailed and purified cDNA, 2 μl 3′-RACE adapter, 0.2 mM each dNTP, 1× Pfu buffer and 1.25 U Pfu DNA Polymerase (Fermentas) with initial denaturation at 95 °C for 2 min, followed by 15 cycles at 95 °C for 40 s, 42 °C for 40 s, 72 °C for 3 min, and a final extension step at 72 °C for 10 min. For amplification of the products, nested PCR was performed. Outer PCR was carried out in 50 μl final volume containing 1 μl previous PCR mixture, 20 pmol each 3′-RACE adapter specific outer primer and TigOAS_R1 primer, 0.2 mM each dNTP, 1× Pfu buffer and 1.25 U Pfu DNA Polymerase (Fermentas) with initial denaturation at 95 °C for 2 min, followed by 35 cycles at 95 °C for 40 s, 64 °C for 40 s, 72 °C for 3 min, and a final extension step at 72 °C for 10 min. Inner PCR was performed in 50 μl final volume containing 1 μl outer PCR mixture, 20 pmol each 3′-RACE adapter specific inner primer and TigOAS_R2 primer (5′-ATGTCGTAGTCACCTGGCACTG-3′) and 1× PCR Master Mix (Fermentas) containing Taq DNA polymerase. Thermal cycling conditions were as in outer PCR. The PCR products were cloned using InsTAclone™ PCR Cloning Kit (Fermentas) and sequenced with vector-specific primers. All sequencing results were analyzed and edited using GAP4 software [30]. The protein sequences were aligned with Clustal Omega software [31], identities and similarities between protein sequences were calculated using SIAS software [32].

The sponge genomic DNA was extracted using the CTAB method as described earlier [33].

In order to explore the intron/exon structure of the gene encoding the OAS in *T. ignis*, PCR was carried out in 50 μl final volume containing 1 μg of DNA from *T. ignis*, 20 pmol each TigOAS_F3 (5′-ATCCATGGCCTCACCCCTCATTC-3′) and TigOAS_R4 (5′-ATAGATCTGTGTATCTCGTCCACTGAC-3′) primer and 1× PCR Master Mix (Fermentas) with initial denaturation at 95 °C for 2 min, followed by 35 cycles at 95 °C for 40 s, 55 °C for 40 s, 72 °C for 2.5 min, and a final extension step at 72 °C for 10 min. The PCR product was excized from gel, cloned and sequenced as described above.

2.3. Cloning into expression vectors

The coding regions of the putative OAS cDNAs were cloned into the pET24d expression vector (Novagen). The resulting polypeptides contained additional C-terminal amino acids Arg and Ser in the case of TigOAS and Leu and Glu in the case of MbrOAS, McaOAS and LerOAS before histidine affinity tag. AmeOAS contained no extra amino acids between the protein sequence and the C-terminal histidine affinity tag. The second amino acid was mutated to Val in MbrOAS and to Ala in McaOAS and AmeOAS in order to introduce a NcoI restriction site for cloning (see Supplementary material for protein sequences).

To obtain GST-fusion proteins, the same inserts were cloned into pET42a expression vector (Novagen) and thus the resulting recombinant proteins contained the same extra amino acids and histidine affinity tag in their C-termini.

In order to add a second C-terminal histidine-tag to LerOAS, a reverse primer containing SalI restriction site was designed against the histidine-tag coding sequence. The sequence previously cloned into pET24d was amplified using vector specific T7 promoter primer and the reverse primer against histidine-tag. The ends of the PCR product were cut with NcoI and SalI and the resulting insert was ligated into the pET24d vector that had been restricted with NcoI and SalI.

All constructs were verified by sequencing.

The calculated molecular weights of the recombinant proteins were 61 kDa, 72 kDa and 78 kDa for GST-tagged MbrOAS, McaOAS and AmeOAS, respectively. The calculated molecular weight of the recombinant LerOAS with two histidine-tags was 59 kDa and in the case of the TigOAS with one histidine-tag it was 36 kDa and with GST-tag 67 kDa [34].

2.4. Expression and purification of the recombinant OAS proteins

The insert-containing plasmids were transformed into *E. coli* BL21-CodonPlus(DE3)-RP strain (Stratagene). The bacteria transformed with pET24d vectors were grown in optM9 media, containing 1xM9 salts [35], 0.2 mM $CaCl_2$, 5 mM $MgCl_2$, 1% glucose, 40 μg/ml chloramphenicol and 100 μg/ml kanamycin, on a rotary shaker at 180 rpm at 37 °C until the cell density of $A_{600\ nm} = 0.6–0.8$ was reached. Then the cultures were cooled to 20 °C and after 30 min adaptation period the expression of recombinant protein was induced by adding IPTG (Fermentas) at a final concentration of 0.5 mM. After overnight incubation at 20 °C the cells were harvested by centrifugation and frozen until lysis.

The bacteria transformed with pET42a vectors were grown in LB media, containing 2 mM $MgCl_2$, 1% glucose, 40 μg/ml chloramphenicol and 30 μg/ml kanamycin. The growth conditions were as described above, except that the protein expression was induced by adding IPTG at a final concentration of 0.1 mM.

Cells expressing histidine-tagged recombinant proteins were lysed in the lysis buffer (50 mM Na_2HPO_4, pH 6.8, 500 mM NaCl, 10% glycerol, 1 mg/ml lysozyme) by sonication on ice. The lysate was clarified by centrifugation and the supernatant was applied to a Ni-NTA agarose (Qiagen) column. The column was washed with wash buffer (50 mM Na_2HPO_4, pH 6.8, 500 mM NaCl, 10% glycerol, 50 mM imidazole) and the bound proteins were eluted with elution buffer (50 mM Na_2HPO_4, pH 6.8, 500 mM NaCl, 10% glycerol, 250 mM imidazole) in 0.5–1 ml fractions.

Cells expressing GST-tagged recombinant proteins were lysed in the lysis buffer (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, 10% glycerol, 1 mg/ml lysozyme) and the cleared lysate was prepared as described above. The cleared lysate was applied to a glutathione agarose (GE Healthcare) column, the column was washed with lysis buffer (without lysozyme) and the bound proteins were eluted with the elution buffer (50 mM Tris–HCl, pH 8.0, 150 mM NaCl, 10% glycerol, 10 mM reduced glutathione) in 0.5–1 ml fractions.

The fractions were analyzed by SDS-PAGE and western blot as described previously [26]. Fractions containing recombinant protein were pooled and concentrated using Amicon® Ultra Centrifugal Filter Devices (10 kDa MWCO, Millipore).

The porcine recombinant OAS was expressed and purified as described earlier [26].

2.5. Size exclusion chromatography

Size exclusion chromatography was performed using Waters HPLC system (Waters Corporation). The recombinant protein preparation was loaded onto an SEC column (BioSep-SEC-S3000, 300 × 7.8 mm, 5 μm, Phenomenex) at room temperature and the elution was performed with 50 mM Tris–HCl, pH7.5, 150 mM NaCl, 10% glycerol at a flow rate of 0.75 ml/min for 30 min. After washing the column with elution buffer for 6 min, 28–42 fractions were collected (250 μl each). The collected fractions were analyzed in 10% SDS-polyacrylamide gel and those containing pure recombinant protein were pooled and concentrated using Amicon® Ultra Centrifugal Filter Devices (10 kDa MWCO, Millipore).

2.6. OAS activity assay

The 2-5A synthesizing activity of recombinant proteins was assayed by incubating the recombinant protein in the reaction mixture containing 20 mM Tris–HCl, pH 8.0, 5 mM $MgCl_2$, 1 mM ATP as a substrate and varying concentrations (0–1.0 mg/ml) poly(I)·poly(C) (from Sigma) as a potential activator of OAS proteins, in a final volume of 50 μl, at 37 °C for different time periods. The reactions were stopped and the products were dephosphorylated as described earlier [26].

The reaction products were analyzed on the Agilent 1100 Series HPLC instrument (Agilent Technologies) equipped with a C18 reverse-phase column (Phenomenex Luna C18(2), 250 × 4.6 mm, 5 mm; pre-column SecurityGuard C18, 4 × 3 mm). Eluent A was 100 mM ammonium phosphate, pH 7.0 and eluent B was 50% methanol in water. The oligomers were separated at 40 °C in a 3-step linear gradient of eluent B in water (0%, 1 min, then 0–28%, 9 min followed by 28–50%, 16 min and 50–60%, 1 min) with eluent A constantly at 20%, then the column was eluted at 60% of eluent B for 1 min and after lowering its proportion to 0% (1 min), the column was equilibrated with 20% eluent A for 7 min. The absorption was measured at 260 nm. ATP, adenosine and oligoadenylates in either their phosphorylated or dephosphorylated forms were identified by comparing their retention times with those of authentic compounds, by their UV spectra (using previously created UV spectra library) and MALDI-MS analysis [26]. The identity of the oligomers containing 3′,5′-linkages was further confirmed by means of RNase T_2 treatment.

2.7. RNase T_2 treatment of the reaction products

The whole dephosphorylated reaction mixture from TigOAS activity assay or fractions corresponding to the individual peaks were treated at 37 °C with 0.012 U/μl of RNase T_2 (Invitrogen) in a reaction mixture containing 50 mM Na-acetate pH 4.0 for 1.5 h. The reaction was stopped by heating at 95 °C for 5 min and the products were analyzed by HPLC as described above.

3. Results

3.1. OAS gene and protein from the marine sponge T. ignis

Degenerate primers designed against conserved motifs in known 2-5A synthetases from several marine sponges were used to amplify a fragment of the 2-5A synthetase cDNA from the marine sponge *T. ignis*. Based on the obtained sequence, specific primers were designed and the 5′- and 3′-ends were amplified and sequenced. Due to the partial degradation of the RNA, the complete sequence of the 5′-end could not be established with the RLM-RACE Kit, where only cap-containing mRNAs are amplified. Therefore alternative approach was used (see Section 2.2.) and the longest amplified fragments were cloned and sequenced. The obtained cDNA was 1028 nt long and contained one ORF which ranged from nt 35–37 (start) to nt 932–934 (stop). The deduced polypeptide was 299 aa long with calculated molecular weight of 34.7 kDa and with the estimated pI of 9.13 [34]. The sequence was deposited in the GenBank database (accession no. KF362123).

The deduced protein sequence analysis revealed that all residues considered to be essential for 2-5A synthetase activity were present as well as the conserved sequence motifs with unknown function (Fig. 1).

The exon/intron structure of the OAS gene from *T. ignis* was similar to that of the OAS genes from other marine sponges [28] and the encoded protein shared 29% identity and 43% similarity with the 2-5A synthetase from another marine sponge *G. cydonium* (GcyOAS, Uniprot accession number O97190) (Fig. 1).

3.2. Expression and purification of putative OAS proteins

Putative oligoadenylate synthetases from different organisms [choanoflagellate *M. brevicollis* (MbrOAS), sponge *T. ignis* (TigOAS), mollusk *M. californianus* (McaOAS), salamander *A. mexicanum* (AmeOAS) and skate *L. erinacea* (LerOAS)] were expressed in a bacterial expression system. TigOAS was expressed as a C-terminally histidine-tagged protein and LerOAS contained two C-terminal histidine tags, separated from one another by a 9 amino acid linker (VDKLAAALE). The additional histidine-tag was introduced to LerOAS in order to increase the affinity to Ni-NTA agarose to which the protein having only one tag bound weakly and eluted in the wash fractions together with a number of contaminating bacterial proteins. MbrOAS, McaOAS and AmeOAS were GST-tagged to enable purification by binding to the glutathione agarose. The attempts to purify them as histidine-tagged proteins using Ni-NTA agarose had failed because the proteins did not bind to the affinity resin.

All recombinant protein preparations purified by affinity chromatography contained impurities, including several proteins of bacterial origin (Fig. 2). In the recombinant protein preparations of histidine-tagged proteins (LerOAS and TigOAS) the major contaminant was <25 kDa bacterial protein. In the GST-tagged recombinant protein preparations, in addition to bacterial proteins, some of the contaminants were probably prematurely terminated recombinant proteins containing GST but no part or only a small part of the OAS protein (Fig. 2, ~25–30 kDa). For further purification of the preparations size exclusion chromatography was used. The recombinant proteins MbrOAS, McaOAS and AmeOAS eluted as broad peaks starting from the column void volume, indicating multimeric nature of the proteins in their native states. It allowed us to separate the recombinant proteins from their contaminants (see Supplementary material). LerOAS eluted in two peaks: 1) monomer, whose peak overlapped partially with those of the contaminants and which was therefore unpurifiable by the method we used and 2) oligomer (at least trimer), that eluted earlier than contaminating proteins. The exception was TigOAS, which eluted only as a monomer and was unseparable from the major contaminating protein due to the small difference in their molecular weights. To test if it was possible to obtain a TigOAS protein preparation with smaller amount of contaminating bacterial proteins, the TigOAS was also expressed as a GST-tagged protein. The resulting protein preparation was significantly purer and did not contain the major contaminant, the <25 kDa bacterial protein, observed in the histidine-tagged TigOAS preparation. The purity of protein

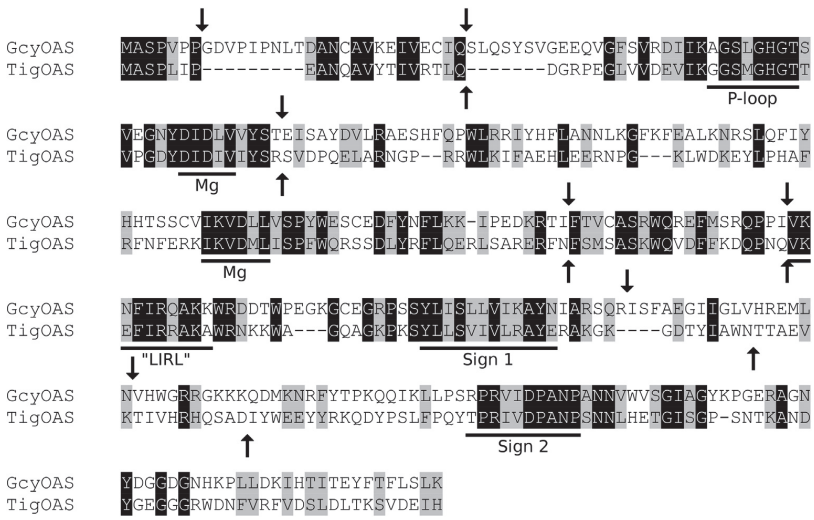


Fig. 1. Comparison of OAS structures from two marine sponges, *Tedania ignis* (TigOAS) and *Geodia cydonium* (GcyOAS). The amino acids are shaded according to the identity (in black) and similarity (in gray). Intron locations in the corresponding genes are indicated by arrows. OAS-specific motifs are underlined.

preparations used for further experiments is demonstrated in Fig. 2. The highest purity was achieved in the case of LerOAS.

3.3. Enzymatic activities of recombinant OASs

The 2′,5′-oligoadenylate synthesizing activity of recombinant proteins was assayed in a reaction mixture containing ATP and varying concentrations of poly(I)·poly(C). Under the conditions tested, we could not detect any oligoadenylate synthesizing activity of the recombinant proteins from *M. brevicollis* or from *A. mexicanum*.

McaOAS, LerOAS and TigOAS were able to polymerize ATP into 2′,5′-linked oligoadenylates in the presence of commonly used synthetic dsRNA poly(I)·poly(C) as it had been shown earlier for mammalian OAS proteins [24]. Without the dsRNA, the enzymatic activity of the proteins was just above the detection limit. McaOAS and LerOAS were activated already by the lowest poly(I)·poly(C) concentration (1 μg/ml) tested, achieving 33% and 53%, respectively, of the maximum activation observed (Fig. 3). In the case of TigOAS, a similar level of activation (27%) was observed in the presence of much higher poly(I)·poly(C) concentration (100 μg/ml). LerOAS exhibited highest oligoadenylate synthesizing activity when the poly(I)·poly(C) concentration in the reaction mixture was 100 μg/ml. For McaOAS and TigOAS the highest enzymatic activity was measured at poly(I)·poly(C) concentration of 1 mg/ml. It should be noted that this was also the highest poly(I)·poly(C) concentration tested and therefore the real poly(I)·poly(C) concentration required for maximal activity remains unknown.

In our standard 2-5A synthetase assay conditions [1 mM ATP, 5 mM $MgCl_2$ and 100 μg/ml poly(I)·poly(C)], LerOAS had the highest specific activity among the recombinant proteins studied: 8–22 nmol of ATP polymerized to oligoadenylates per 1 μg of protein in 1 h. McaOAS and TigOAS exhibited significantly lower specific activity in these conditions: 1.2–1.7 and 0.05–0.06, respectively. Under the same conditions, a mammalian OAS1, the recombinant porcine OAS, had still about an order of magnitude higher specific activity if compared to the most active recombinant protein studied herein (see also Ref. [26]).

In order to describe the products synthesized by the recombinant OAS proteins, the activity assays were carried out for different time periods (up to 31 h). The porcine recombinant protein which was used as a mammalian OAS1 standard polymerized ATP into 2-5A oligomers

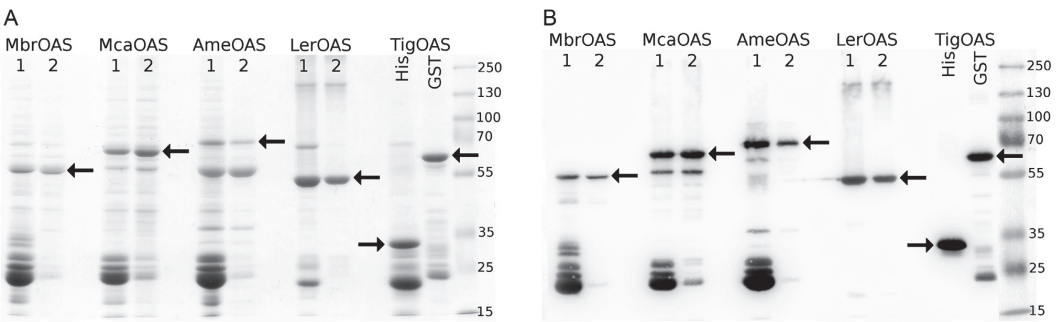


Fig. 2. SDS-PAGE (A) and western blot (B) analysis of the recombinant proteins before (1) and after (2) purification by size exclusion chromatography. Protein preparations of histidine-tagged and GST-tagged TigOAS were not purified by size exclusion chromatography. The gel was stained with ProSieve® Blue Protein Staining Solution (Lonza) and western blot was probed with anti-His_6 antibody. The recombinant OASs are indicated by arrows.

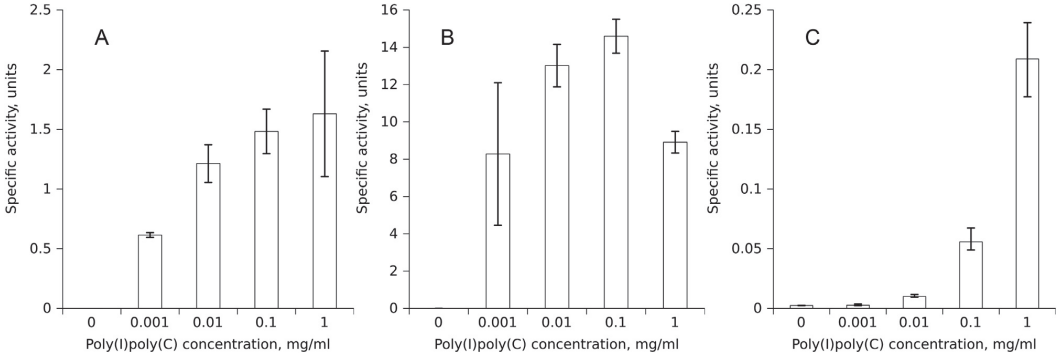


Fig. 3. The effect of increasing concentrations of poly(I)·poly(C) on the enzymatic activity of the recombinant protein preparations from *M. californianus* (A), *L. erinacea* (B) and *T. ignis* (C). The calculated specific activity units are expressed as the amount of ATP (nmol) polymerized per 1 μg of protein in 1 h. In the case of TigOAS, the calculations of specific activity also include the products with 3′,5′-linkages in addition to the 2′,5′-linked oligoadenylates. Error bars indicate the highest and lowest values of the specific activity from two independent experiments.

ranging from dimer to pentamer (Fig. 4A, B). The recombinant proteins McaOAS, LerOAS and TigOAS were able to catalyze the synthesis of 2-5A dimer, trimer and tetramer with the prevalence of dimer among reaction products. In the case of McaOAS, the presence of pentamer in the reaction mixture was also detected (Fig. 4C–H).

Similarly to the OAS previously characterized from another marine sponge *G. cydonium* [26], the sponge recombinant OAS from *T. ignis*, histidine-tagged TigOAS, was able to catalyze the formation of oligoadenylates with 3′,5′-phosphodiester bond. 2′,5′-linked oligoadenylates were found in the reaction mixture together with the 3′,5′-linked di- and trinucleotides. Furthermore, the presence of triadenylates with the first linkage being a 2′,5′-phosphodiester bond and the second one being a 3′,5′-bond was observed (Fig. 4G, H). Ribonuclease T2 treatment of the dephosphorylated reaction mixture of TigOAS resulted in the degradation of all the products containing 3′,5′-linkages, leaving the products with 2′,5′-linkages intact (Fig. 5). 3′-AMP seen additionally on the chromatogram derived from 3–5A dimer and trimer (Fig. 5B), whereas A2′p5′A3′p was a degradation product of the triadenylate A2′p5′A3′p5′A (Fig. 5B and Fig. S6).

The ratio of products containing 2′,5′-linkage to the products with 3′,5′-linkage was 2.4 ± 0.2 in the presence of 100 μg/ml poly(I)·poly(C) and 2.8 ± 0.2 in the presence of 1 mg/ml poly(I)·poly(C) ($n = 4$). To confirm that both enzymatic activities were exhibited by the recombinant protein and not by the <25 kDa bacterial protein, size exclusion chromatography fractions 19 and 23 (Fig. S5) were analyzed for oligoadenylate synthesizing activity. Fraction 19 which contained mostly the <25 kDa bacterial protein and only traces of recombinant TigOAS protein showed minimal enzymatic activity, though both 2′,5′- and 3′,5′-linked diadenylates were detected in the reaction mixture. Fraction 23, where the amount of recombinant TigOAS was higher than that of the contaminating protein, exhibited higher enzymatic activity and the ratio of products with 2′,5′-linkage to the products with 3′,5′-linkage was 2.3 (data nor shown). The GST-tagged TigOAS protein preparation exhibited significantly lower enzymatic activity if compared to the histidine-tagged TigOAS preparation. However, it was still able to catalyze the formation of both, 2′,5′- and 3′,5′-linked oligoadenylates like the histidine-tagged TigOAS.

4. Discussion

Based on the EST and genomic analysis, the OAS proteins are now believed to be represented throughout the animal evolutionary tree with losses in some of its branches [21]. In the present work the cDNAs of four putative oligoadenylate synthetases from distantly related animal species as well as the cDNA of a choanoflagellate, a close relative of animals, were expressed in a bacterial system and their capabilities to synthesize oligoadenylates were tested. Three of the expressed proteins proved to be active 2-5A synthetases, while two proteins, namely the OAS from the marine choanoflagellate *M. brevicollis* and the amphibian *A. mexicanum* were not able to catalyze the formation of 2-5A oligomers from ATP in our assay conditions.

Since 2-5A synthesizing activity is generally believed to be limited to Metazoa, the prediction of a putative OAS gene in the genome of *M. brevicollis* [21] challenged us to revisit this belief. Based on *M. brevicollis* genome sequencing data, a 634 aa protein was predicted by Kuo et al. (GenBank accession XM_001750778). Using the NCBI Conserved Domains Search tool three conserved domains in this predicted protein sequence can be found on the basis of similarity. In the N-terminal part there is a sequence similar to the ATP-binding site of histidine kinase-like ATPases. The nucleotidyl transferase domain of OAS (or class I CCA-adding enzymes) and a domain similar to the C-terminus of OAS1, homologous to a tandem ubiquitin repeat were found in the C-terminal part of the predicted protein [21]. In the present work only the 249 aa C-terminal part of this hypothetical protein containing all the motifs characteristic to OAS proteins was expressed as a recombinant protein. As we could not detect any 2-5A synthesis catalyzed by the recombinant MbrOAS, we hypothesize that if the predicted protein exists, it may need either both domains to acquire the correct fold or to require the additional ATP-binding site in the N-terminal part for being enzymatically active. However, there are no EST sequences available to completely cover the entire predicted protein sequence. Of the two ESTs found for this protein one covers the C-terminal part only and the second, longer one has not been completely sequenced, hence the information on the middle part of the EST is missing. Though the existence of this protein is confirmed by the longer EST, the genome-based prediction of its sequence/length needs to be established by future studies.

According to the phylogenetic tree of OAS proteins, the OAS from *M. brevicollis* has early diverged from the rest of the OAS proteins [21]. The protein encoded by the *M. brevicollis* genome, although it contains the sequence motifs characteristic to OAS, may not have obtained the structural properties that result in enzymatically active protein. If enzymatically active at all, it might catalyze the formation of 3′,5′-phosphodiester linkages instead of

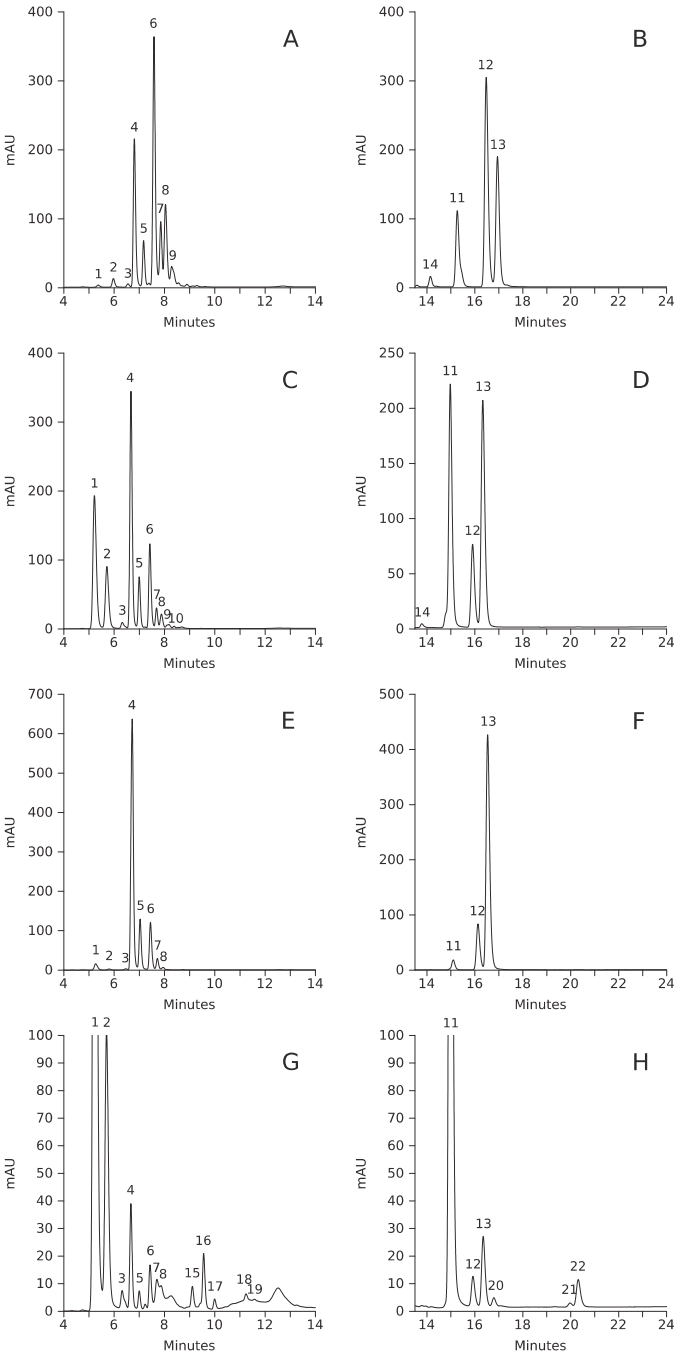


Fig. 4. HPLC analysis of products synthesized from ATP by the recombinant porcine OAS (A, B), McaOAS (C, D), LerOAS (E, F) and TigOAS (G, H) in their phosphorylated (left column) or dephosphorylated (right column) forms. 1, ATP; 2, ADP; 3, AMP; 4, p3A2′p5′A; 5, p_2A2′p5′A; 6, p_3A2′p5′A2′p5′A; 7, p_2A2′p5′A2′p5′A; 8, p_3A2′p5′A2′p5′A2′p5′A; 9, mixture of p_2A2′p5′A2′p5′A2′p5′A and p3A2′p5′A2′p5′A2′p5′A2′p5′A; 10, p2A2′p5′A2′p5′A2′p5′A2′p5′A; 11, mixture of adenosine and A2′p5′A2′p5′A2′p5′A; 12, A2′p5′A2′p5′A; 13, A2′p5′A; 14, A2′p5′A2′p5′A2′p5′A2′p5′A; 15, p_3A2′p5′A3′p5′A 16, p_3A3′p5′A; 17, p_2A3′p5′A; 18, p_3A3′p5′A3′p5′A; 19, p_2A3′p5′A3′p5′A; 20, A2′p5′A3′p5′A; 21, A3′p5′A3′p5′A; 22, A3′p5′A.

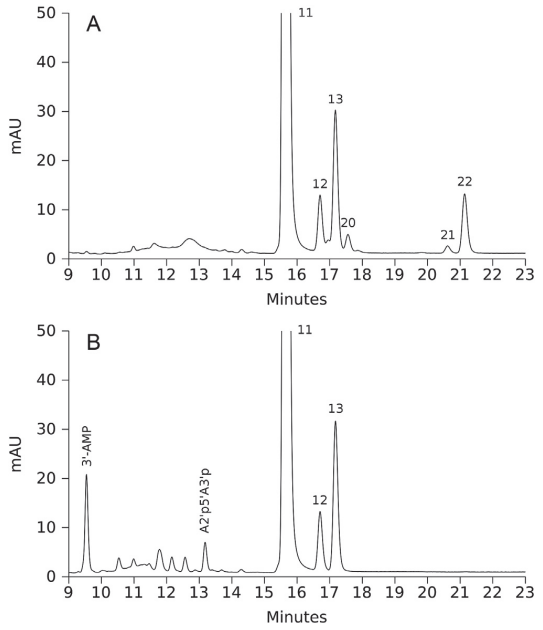


Fig. 5. RNase T_2 treatment of products synthesized from ATP by TigOAS. The dephosphorylated reaction mixture containing oligoadenylates with 2′,5′- and 3′,5′-linkages (A) was treated with RNase T_2 yielding intact 2′,5′-linked oligoadenylates and degradation products of oligoadenylates with 3′,5′- and mixed linkages (B). The peak numbers correspond to those on Fig. 4.

2′,5′-linkages; the optimal conditions for this catalysis being different from those used here. It is also possible that the enzyme still possesses the properties of an ancestral enzyme, being for example incapable of primer-independent synthesis.

The proteins of OAS family lacking enzymatic activity have been described earlier. For example, the majority of the eight murine OAS1 proteins do not possess enzymatic activity [12]. The OAS gene from the amphibian *A. mexicanum* was inactive in our assay; that may be related to the particular cDNA clone which had a linker sequence at the incomplete 5′-terminus, though all the necessary sequence motifs of the functional OAS protein were present. On the other hand, the real number of putative OAS genes in the genome of *A. mexicanum* is not known as the whole genome sequence is not available yet. For a close relative of *A. mexicanum*, *Ambystoma tigrinum*, two ESTs have been identified that encode putative OAS proteins [21]. It is likely that in *A. mexicanum* there are other OAS genes, some of which may encode proteins with 2-5A synthesizing activity. However, Cayley et al. [23] were not able to detect any OAS activity in any of the tested amphibians. Two of them, *Necturus maculosus* and *Salamandra salamandra*, belong to the same order Caudata as *A. mexicanum*; *Bufo woodhousei* is a more distant relative. Up to now, none of these three genomes have been fully sequenced. Obviously, more data are needed to establish the presence of OAS genes and their activities in amphibians.

Three expressed proteins, the recombinant OASs from the sponge *T. ignis*, the mollusk *M. californianus* and the little skate *L. erinacea* were able to polymerize ATP into 2′,5′-linked oligoadenylates in the presence of poly(I)·poly(C), thus being true oligoadenylate synthetases (Fig. 4). The demonstration of the OAS activity in *M. californianus* was the first case for the whole animal clade Protostomia where some of its branches have even lost the OAS genes.

Recently, a putative OAS1 gene was also identified for a representative of the phylum Cnidaria, the sea anemone *Nematostella vectensis* [21]. Though there are no data about OAS activity in this species, another cnidarian of the class Anthozoa, the soft coral *Alcyonium acaule*, exhibited 2-5A synthesizing activity (our unpublished results). These results confirm that active OAS proteins are indeed present in various animal phyla and no evolutionary gap exists between vertebrate and sponge OASs (Fig. 6).

The OAS from the little skate, *L. erinacea*, contains the ubiquitin-like repeats at the C-terminus and is therefore considered to be an OASL protein [21]. Originally the OASL proteins were believed to be enzymatically inactive or to have another catalytic activity not yet identified [36]. Since then cDNAs encoding OASL proteins with 2′,5′-oligoadenylate synthesizing activity have been cloned from chicken and mouse [16,37] in addition to the little skate. Hence it is not possible to predict if the OAS is able to catalyze the formation of oligoadenylates or not just on the basis of it having the ubiquitin-like C-terminus. More likely the structure of the active site and its surrounding sequence is important in determining the protein's ability to produce oligoadenylates. The function of this ubiquitin-like domain is yet to be determined. However, it has been shown that the antiviral activity of the human p59 OASL protein was independent of the 2-5A synthesis and required the ubiquitin-like domain [38].

Various dsRNA molecules, produced during viral infections as well as dsRNA-like cellular structures have been shown to activate OAS proteins *in vitro* [39,40]. In practice, the most commonly used activator of OAS is the synthetic dsRNA, poly(I)·poly(C). The necessary amount of dsRNA to achieve a half-maximum activity of different OAS forms (OAS1, OAS2 and OAS3) ranges from less than 1 μg/ml to more than 500 μg/ml poly(I)·poly(C) [7]. A considerable

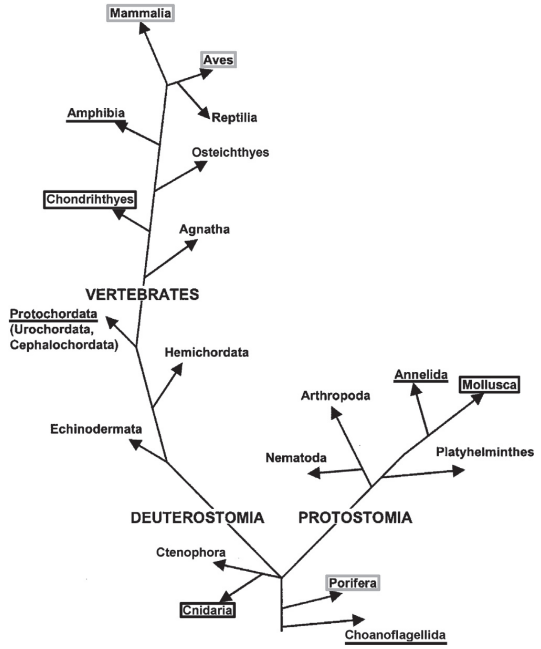


Fig. 6. Schematic representation of the presence of OASs on the phylogenetic tree of animals. The branches where the OAS activity has been shown earlier are marked with gray borders and the novel data are shown by black borders. Animal groups where the presence of OAS genes has been only bioinformatically predicted are underlined. The phylogenetic tree was modified from Ref. [46].

variation was also observed in our experiments. LerOAS reached half-maximum activity already at 1 μg/ml poly(I)·poly(C), while for McaOAS the required concentration was between 1 and 10 μg/ml poly(I)·poly(C) and TigOAS needed more than 100 μg/ml poly(I)·poly(C). We have demonstrated earlier that the recombinant OAS from another marine sponge *G. cydonium*, which binds to RNA already during its expression in the bacterial cell and forms an enzymatically active protein–RNA complex, does not require any additional poly(I)·poly(C) for its enzymatic activity [26]. During its synthesis, the GcyOAS binds certain unidentified bacterial RNAs with very high affinity and poly(I)·poly(C) is unable to compete with them. The recombinant OAS from the marine sponge *T. ignis*, TigOAS also had a low affinity towards poly(I)·poly(C). A high concentration of the polynucleotide required may refer to the poor suitability of this synthetic RNA for the activation of sponge OAS proteins. Though we hypothesize that the activation mechanism of sponge oligoadenylate synthetases is similar to that of the vertebrate ones, the affinities and structural requirements for the RNA component may be quite different.

In addition to the catalysis of the formation of 2′,5′-phosphodiester bonds, the recombinant OAS from *T. ignis* was also able to catalyze the synthesis of 3′,5′-phosphodiester bonds. The same phenomenon was observed earlier when testing the enzymatic activity of the recombinant OAS from *G. cydonium* [26]. However, the ratio of 2-5A oligomers to 3-5A oligomers in a reaction catalyzed by TigOAS compared to that catalyzed by GcyOAS was even more in favor of the formation of 3-5A: 2.4 for TigOAS *vs* 5.4 for GcyOAS. Differently from TigOAS, McaOAS and LerOAS were only able to catalyze the formation of a 2′,5′-linkage, thereby resembling the mammalian OAS proteins.

The primary sequences of sponge OASs show moderate similarity and identity amongst one another [28] and to vertebrate OASs [19]. 2′- and/or 3′-specificities of OASs from different sources presumably depend on the structures of their active sites. The dissimilar amino acid context surrounding the active sites of the different OASs may contribute to the distinct assembly and structure of them and therefore influence the choice of linkage to be catalyzed.

The 2′- and 3′-specificities of OAS may also be determined by the activating RNA species since the addition of poly(I)·poly(C) shifted the ratio of 2′,5′-linkages to 3′,5′-linkages slightly in favor of 2′,5′-linkages in the activity assays with TigOAS as well as in the previous experiments with GcyOAS [26]. Could it be conceivable that TigOAS would behave as a 3-5A synthetase in the presence of the appropriate activator similarly to the activity observed in the crude extract of *T. ignis* [27]? Regarding this issue, it should be mentioned that though sponge crude extracts do not require exogenously added RNA for their characteristic OAS activities [41], the observed activity may actually be mediated by certain endogenously formed nucleic acid–protein complex(es). However, the cloned TigOAS may be one of several OASs with different activation and enzymatic characteristics in *T. ignis* and the activities seen in the crude extract are a combination of them. Useful information would be provided by the analysis of the crystal structure of the sponge OAS crystallized together with its substrate and activator. It would enable us to compare the structure and the flexibility of the active site with those of 2′-specific OASs from higher animals and deduce the features that are important in determining the 2′- and 3′-specificities of nucleotidyl transferases in general.

The OAS proteins from sponges form a distinct group of 2′,5′-oligoadenylate synthetases – their genomic structure differs from the conserved exon/intron structure of vertebrate OAS genes [28]. Also, their enzymatic activities have some characteristic features including their capability to catalyze the formation of 3′,5′-linkages and to utilize all four ribonucleotides (ATP, GTP, UTP, CTP) at the acceptor site of OAS for 2′,5′-oligomerization [42]. The sponge OASs may have preserved some characteristic features of the common ancestor of OASs and other members of the same nucleotidyl transferase family. We suggest that the sponge OASs represent a link between earlier polymerases and specific 2′,5′-oligoadenylate synthetases found in higher animals.

The progress in genomic research has enabled researchers to establish the presence of OAS encoding genes throughout the animal kingdom while evolutionary loss of them in some branches of the phylogenetic tree has also been revealed. In the present study these genomic data for some multicellular animals are complemented with the knowledge about the enzymatic activity of the encoded OAS proteins.

The role(s) of OAS in the animals which lack the antiviral interferon-inducible 2-5A system remain difficult to specify. It has been suggested that OAS may also be involved in the antiviral innate immune defense system of sponges since the OAS gene in *L. baicalensis* could be induced with dsRNA [20]. However, the intracellular signaling pathways leading to the antiviral response or high affinity proteinaceous targets of 2-5A (like RNaseL in vertebrates) in lower animals are not known to date. The interest for 2-5A binding proteins in lower animals is starting only now, after the elucidation of functional OAS genes in these organisms. Until now, sponges were the only invertebrates where OAS activities have been demonstrated. Our attempts to identify any specific 2-5A binding proteins in sponges have been unsuccessful so far. Yet other biochemical targets for 2-5A have been found in mammals and several alternative functions for OASs have been proposed, including the regulation of cell proliferation, gene expression, apoptosis or involvement in RNA metabolism [43,44, reviewed in 45]. Some of these functions preserved in mammals might derive from the original functions of OAS proteins in the evolutionarily older metazoans.

Acknowledgments

This work was supported by the Estonian Ministry of Education and Research (grant No. 0140108) and the Estonian Science Foundation (grant No. 9185). We thank Tõnu Reintamm and Kerli Vallmann for technical assistance and for critical reading of the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.biochi.2013.10.015.

References

[1] S. Torralba, J. Sojat, R. Hartmann, 2′-5′ Oligoadenylate synthetase shares active site architecture with the archaeal CCA-adding enzyme, Cell. Mol. Life Sci. 65 (2008) 2613–2620.
[2] H. Kristiansen, H.H. Gad, S. Eskildsen-Larsen, P. Despres, R. Hartmann, The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities, J. Interferon Cytokine Res. 31 (2011) 41–47.
[3] J. Donovan, M. Dufner, A. Korennykh, Structural basis for cytosolic double-stranded RNA surveillance by human oligoadenylate synthetase 1, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) 1652–1657.
[4] L. Sun, J. Wu, F. Du, X. Chen, Z.J. Chen, Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway, Science 339 (2013) 786–791.
[5] A. Ablasser, M. Goldeck, T. Cavlar, T. Deimling, G. Witte, I. Röhl, K. Hopfner, J. Ludwig, V. Hornung, cGAS produces a 2′-5′-linked cyclic dinucleotide second messenger that activates STING, Nature 498 (2013) 380–384.
[6] F. Civril, T. Deimling, C.C. de Oliveira Mann, A. Ablasser, M. Moldt, G. Witte, V. Hornung, K. Hopfner, Structural mechanism of cytosolic DNA sensing by cGAS, Nature 498 (2013) 332–337.
[7] J. Justesen, R. Hartmann, N.O. Kjeldgaard, Gene structure and function of the 2′-5′-oligoadenylate synthetase family, Cell. Mol. Life Sci. 57 (2000) 1593–1612.

[8] R. Kodym, E. Kodym, M.D. Story, 2′-5′-Oligoadenylate synthetase is activated by a specific RNA sequence motif, Biochem. Biophys. Res. Commun. 388 (2009) 317–322.
[9] R. Hartmann, P.L. Norby, P.M. Martensen, P. Jorgensen, M.C. James, C. Jacobsen, S.K. Moestrup, M.J. Clemens, J. Justesen, Activation of 2′-5′ oligoadenylate synthetase by single-stranded and double-stranded RNA aptamers, J. Biol. Chem. 273 (1998) 3236–3246.
[10] I.M. Kerr, R.E. Brown, pppA2′p5′A2′p5′A: an inhibitor of protein synthesis synthesized with an enzyme fraction from interferon-treated cells, Proc. Natl. Acad. Sci. U. S. A. 75 (1978) 256–260.
[11] G. Merlin, J. Chebath, P. Benech, R. Metz, M. Revel, Molecular cloning and sequence of partial cDNA for interferon-induced (2′-5′)oligo(A) synthetase mRNA from human cells, Proc. Natl. Acad. Sci. U. S. A. 80 (1983) 4904–4908.
[12] S. Kakuta, S. Shibata, Y. Iwakura, Genomic structure of the mouse 2′,5′-oligoadenylate synthetase gene family, J. Interferon Cytokine Res. 22 (2002) 981–993.
[13] S. Eskildsen, R. Hartmann, N.O. Kjeldgaard, J. Justesen, Gene structure of the murine 2′-5′-oligoadenylate synthetase family, Cell. Mol. Life Sci. 59 (2002) 1212–1222.
[14] A.A. Perelygin, T.L. Lear, A.A. Zharkikh, M.A. Brinton, Structure of equine 2′-5′oligoadenylate synthetase (OAS) gene family and FISH mapping of OAS genes to ECA8p15->p14 and BTA17q24->q25, Cytogenet. Genome Res. 111 (2005) 51–56.
[15] A.A. Perelygin, A.A. Zharkikh, S.V. Scherbik, M.A. Brinton, The mammalian 2′-5′ oligoadenylate synthetase gene family: evidence for concerted evolution of paralogous Oas1 genes in Rodentia and Artiodactyla, J. Mol. Evol. 63 (2006) 562–576.
[16] A. Yamamoto, A. Iwata, Y. Koh, S. Kawai, S. Murayama, K. Hamada, S. Maekawa, S. Ueda, Y. Sokawa, Two types of chicken 2′,5′-oligoadenylate synthetase mRNA derived from alleles at a single locus, Biochim. Biophys. Acta 1395 (1998) 181–191.
[17] M. Wiens, A. Kuusksalu, M. Kelve, W.E. Müller, 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 (1999) 12–18.
[18] K. Vallmann, N. Aas, T. Reintamm, A. Lopp, A. Kuusksalu, M. Kelve, Expressed 2-5A synthetase genes and pseudogenes in the marine sponge *Geodia barretti*, Gene 478 (2011) 42–49.
[19] V.A. Grebenjuk, A. Kuusksalu, M. Kelve, J. Schütze, H.C. Schröder, W.E.G. Müller, Induction of (2′-5′)oligoadenylate synthetase in the marine sponges *Suberites domuncula* and *Geodia cydonium* by the bacterial endotoxin lipopolysaccharide, Eur. J. Biochem. 269 (2002) 1382–1392.
[20] H.C. Schröder, F. Natalio, M. Wiens, M.N. Tahir, M.I. Shukoor, W. Tremel, S.I. Belikov, A. Krasko, W.E.G. Müller, The 2′-5′-oligoadenylate synthetase in the lowest metazoa: isolation, cloning, expression and functional activity in the sponge *Lubomirskia baicalensis*, Mol. Immunol. 45 (2008) 945–953.
[21] K.H. Kjaer, J.B. Poulsen, T. Reintamm, E. Saby, P.M. Martensen, M. Kelve, J. Justesen, Evolution of the 2′-5′-oligoadenylate synthetase family in eukaryotes and bacteria, J. Mol. Evol. 69 (2009) 612–624.
[22] B. Venkatesh, E.F. Kirkness, Y. Loh, A.L. Halpern, A.P. Lee, J. Johnson, N. Dandona, L.D. Viswanathan, A. Tay, J.C. Venter, R.L. Strausberg, S. Brenner, Survey sequencing and comparative analysis of the elephant shark (*Callorhinchus milii*) genome, PLoS Biol. 5 (2007) e101.
[23] P.J. Cayley, R.F. White, J.F. Antoniw, N.J. Walesby, I.M. Kerr, Distribution of the ppp(A2′p)nA-binding protein and interferon-related enzymes in animals, plants, and lower organisms, Biochem. Biophys. Res. Commun. 108 (1982) 1243–1250.
[24] Y. Mory, B. Vaks, J. Chebath, Production of two human 2′,5′-oligoadenylate synthetase enzymes in *Escherichia coli*, J. Interferon Res. 9 (1989) 295–304.
[25] R. Hartmann, J. Justesen, S.N. Sarkar, G.C. Sen, V.C. Yee, Crystal structure of the 2′-specific and double-stranded RNA-activated interferon-induced antiviral protein 2′-5′-oligoadenylate synthetase, Mol. Cell 12 (2003) 1173–1185.
[26] M. Päri, A. Kuusksalu, A. Lopp, T. Reintamm, J. Justesen, M. Kelve, Expression and characterization of recombinant 2′,5′-oligoadenylate synthetase from the marine sponge *Geodia cydonium*, FEBS J. 274 (2007) 3462–3474.
[27] A. Lopp, T. Reintamm, A. Kuusksalu, S. de Rosa, M. Kelve, A novel endoribonuclease from the marine sponge *Tethya aurantium* specific to 2′,5′-phosphodiester bonds, Biochimie 94 (2012) 1635–1646.
[28] T. Reintamm, A. Kuusksalu, M. Metsis, M. Päri, K. Vallmann, A. Lopp, J. Justesen, M. Kelve, Sponge OAS has a distinct genomic structure within the 2-5A synthetase family, Mol. Genet. Genomics 280 (2008) 453–466.
[29] M.A. Frohman, M.K. Dush, G.R. Martin, Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer, Proc. Natl. Acad. Sci. U. S. A. 85 (1988) 8998–9002.
[30] J.K. Bonfield, K.F. Smith, R. Staden, A new DNA sequence assembly program, Nucleic Acids Res. 23 (1995) 4992–4999.
[31] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011) 539.
[32] P. Reche, Protein Identities and Similarities. imed.med.ucm.es/Tools/sias.html (accessed 27.05.13).
[33] A. Lopp, T. Reintamm, K. Vallmann, M. Päri, V. Mikli, E. Richelle-Maurer, M. Kelve, Molecular identification, characterization and distribution of freshwater sponges (Porifera: Spongillidae) in Estonia, Fundam. Appl. Limnol. 168 (2007) 93–103.
[34] C. Putnam, Protein Calculator. http://www.scripps.edu/~cdputnam/protcalc.html (accessed 28.05.13).
[35] J. Sambrook, E.F. Fritsch, T. Maniatis, Molecular Cloning: A Laboratory Manual, second ed., Cold Spring Harbor Laboratory Press, New York, 1989.
[36] D. Rebouillat, I. Marié, A.G. Hovanessian, Molecular cloning and characterization of two related and interferon-induced 56-kDa and 30-kDa proteins highly similar to 2′-5′ oligoadenylate synthetase, Eur. J. Biochem. 257 (1998) 319–330.
[37] S. Eskildsen, J. Justesen, M.H. Schierup, R. Hartmann, Characterization of the 2′-5′-oligoadenylate synthetase ubiquitin-like family, Nucleic Acids Res. 31 (2003) 3166–3173.
[38] J. Marques, J. Anwar, S. Eskildsen-Larsen, D. Rebouillat, S.R. Paludan, G. Sen, B.R.G. Williams, R. Hartmann, The p59 oligoadenylate synthetase-like protein possesses antiviral activity that requires the C-terminal ubiquitin-like domain, J. Gen. Virol. 89 (2008) 2767–2772.
[39] R.H. Silverman, Viral encounters with 2′,5′-oligoadenylate synthetase and RNase L during the interferon antiviral response, J. Virol. 81 (2007) 12720–12729.
[40] M. Dan, D. Zheng, L.L. Field, V. Bonnevie-Nielsen, Induction and activation of antiviral enzyme 2′,5′-oligoadenylate synthetase by in vitro transcribed insulin mRNA and other cellular RNAs, Mol. Biol. Rep. 39 (2012) 7813–7822.
[41] T. Reintamm, A. Lopp, A. Kuusksalu, M. Kelve, Qualitative and quantitative aspects of 2-5A synthesizing capacity of different marine sponges, Biomol. Eng. 4-6 (2003) 389–399.
[42] A. Lopp, T. Reintamm, A. Kuusksalu, I. Tammiste, A. Pihlak, M. Kelve, Natural occurrence of 2′,5′-linked heteronucleotides in marine sponges, Mar. Drugs 8 (2010) 235–254.
[43] D. Ferbus, J. Justesen, F. Besançon, M.N. Thang, The 2′,5′ oligoadenylate synthetase has a multifunctional 2′,5′ nucleotidyl-transferase activity, Biochem. Biophys. Res. Commun. 100 (1981) 847–856.
[44] A. Ghosh, S.N. Sarkar, T.M. Rowe, G.C. Sen, A specific isozyme of 2′-5′ oligoadenylate synthetase is a dual function proapoptotic protein of the Bcl-2 family, J. Biol. Chem. 276 (2001) 25447–25455.
[45] M.R. Player, P.F. Torrence, The 2-5A system: modulation of viral and cellular processes through acceleration of RNA degradation, Pharmacol. Ther. 78 (1998) 55–113.
[46] A. Beschin, M. Bilej, E. Torreele, P. De Baetselier, On the existence of cytokines in invertebrates, Cell. Mol. Life Sci. 58 (2001) 801–814.

Supplementary material

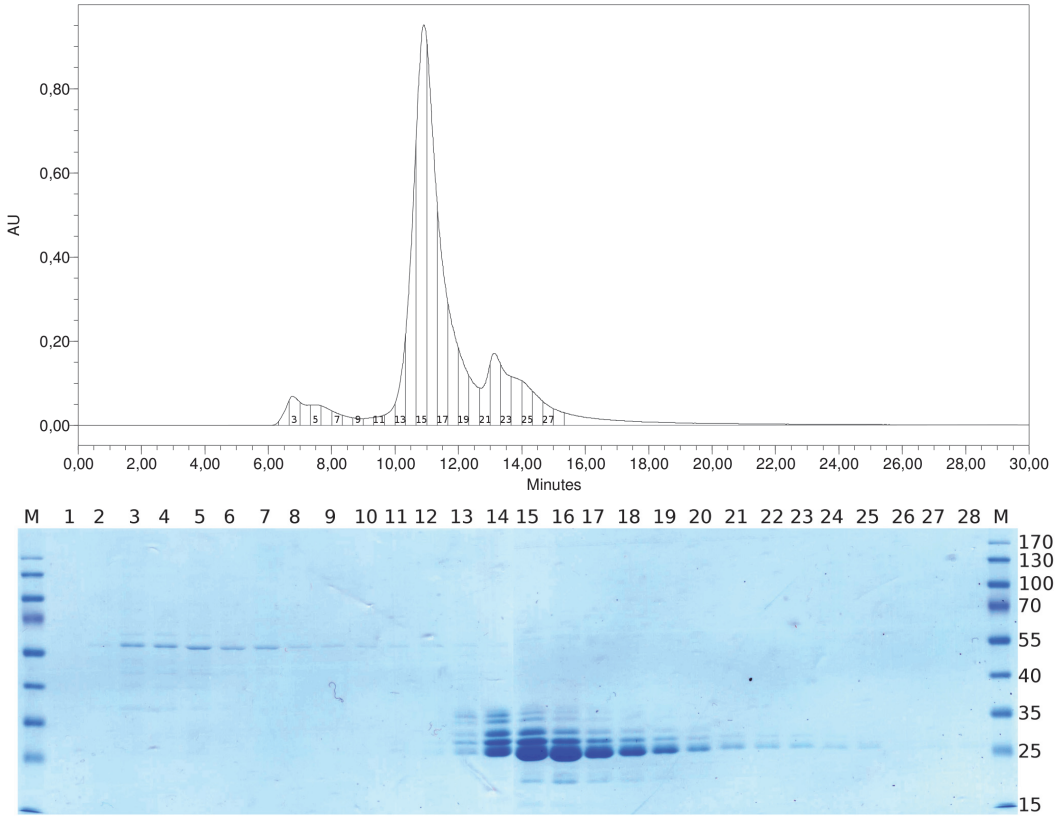


Fig. S1. Fractionation of the GST-tagged MbrOAS preparation by size exclusion chromatography and the SDS-PAGE analysis of the collected fractions. Fractions 2 – 11 were pooled and concentrated for further experiments.

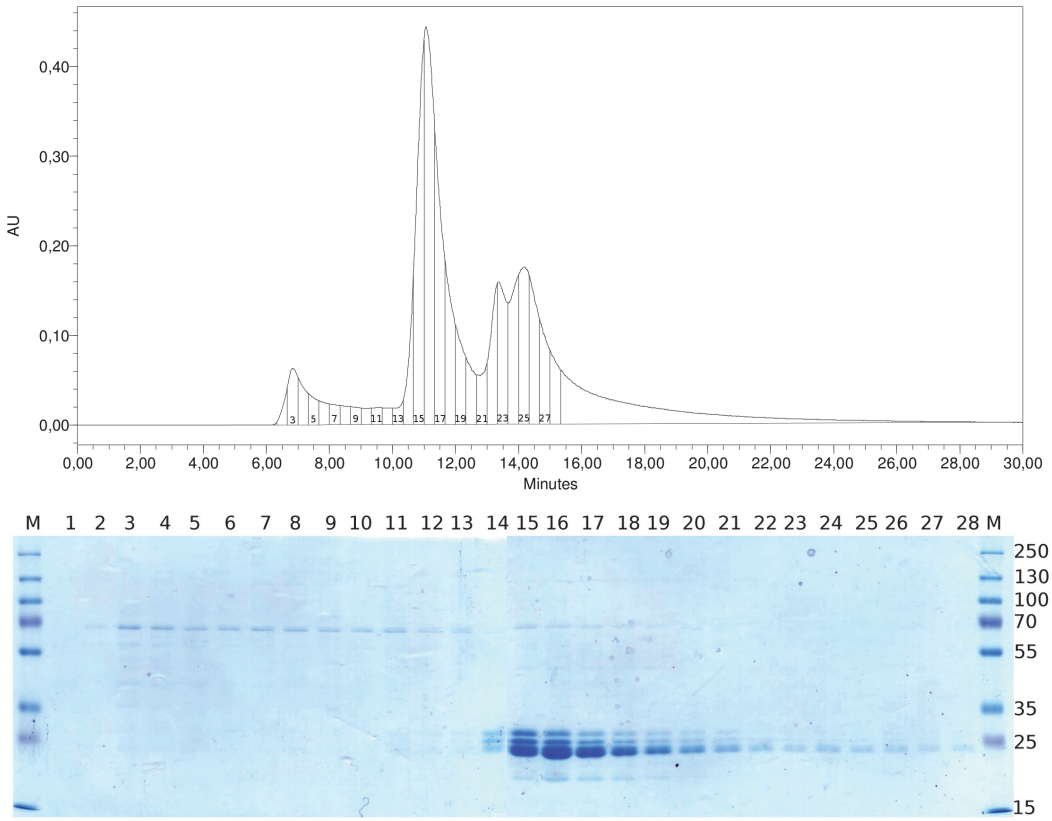


Fig. S2. Fractionation of the GST-tagged McaOAS preparation by size exclusion chromatography and the SDS-PAGE analysis of the collected fractions. Fractions 2 – 10 were pooled and concentrated for further experiments.

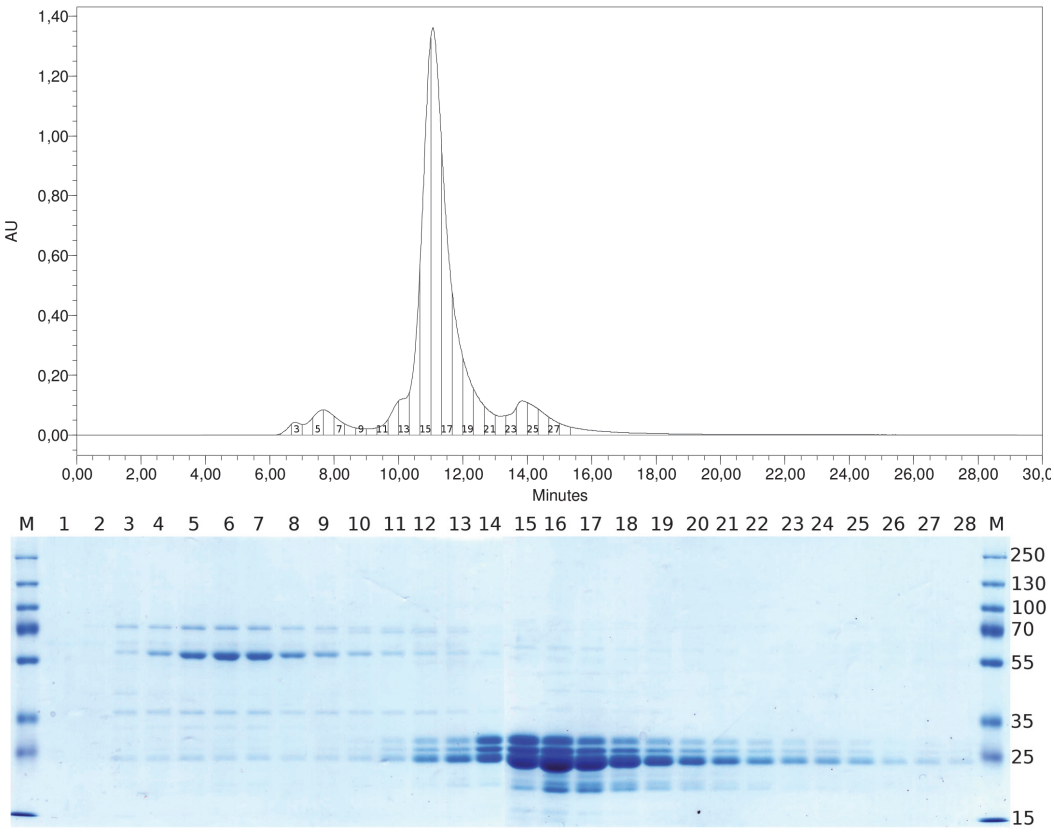


Fig. S3. Fractionation of the GST-tagged AmeOAS preparation by size exclusion chromatography and the SDS-PAGE analysis of the collected fractions. Fractions 2 – 10 were pooled and concentrated for further experiments.

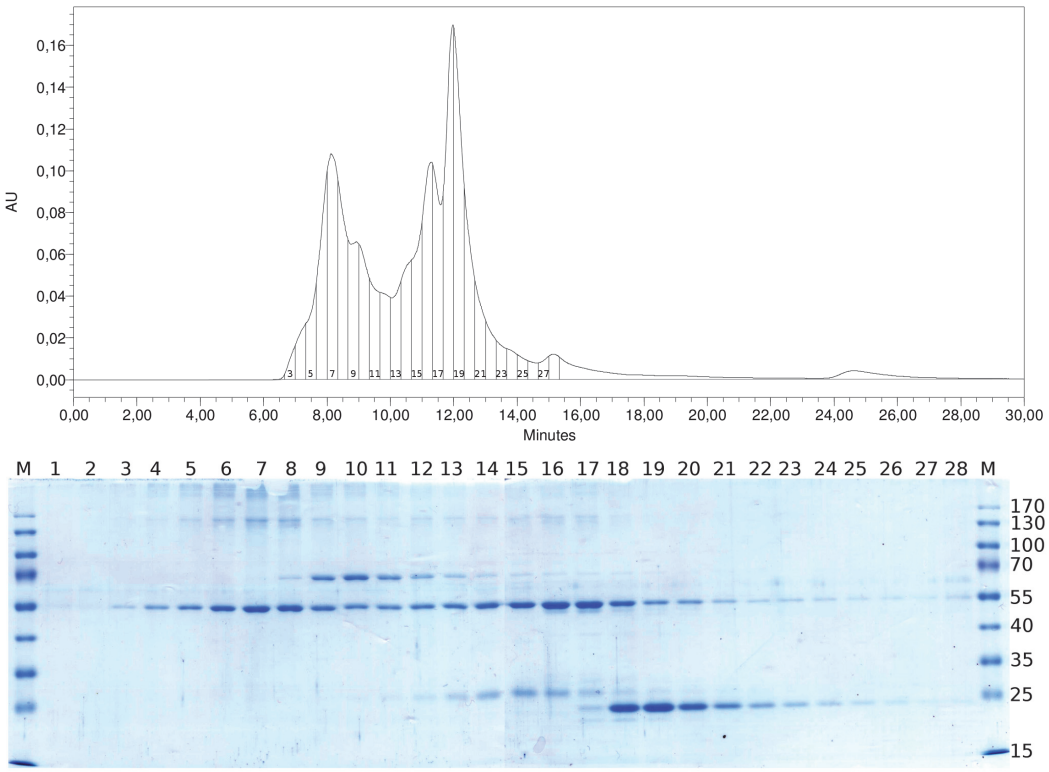


Fig. S4. Fractionation of the histidine-tagged LerOAS preparation by size exclusion chromatography and the SDS-PAGE analysis of the collected fractions. Fractions 3 – 7 were pooled and concentrated for further experiments.

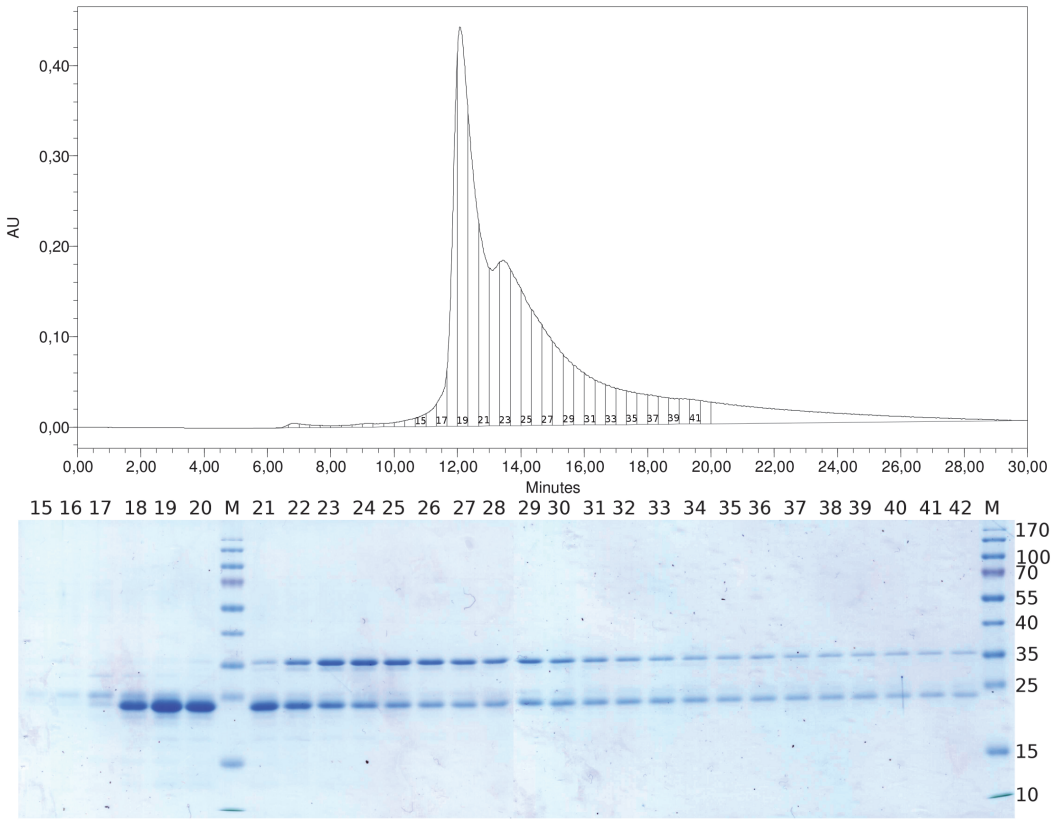


Fig. S5. Fractionation of the histidine-tagged TigOAS preparation by size exclusion chromatography and the SDS-PAGE analysis of the collected fractions. Unpurified protein preparation was used for further experiments.

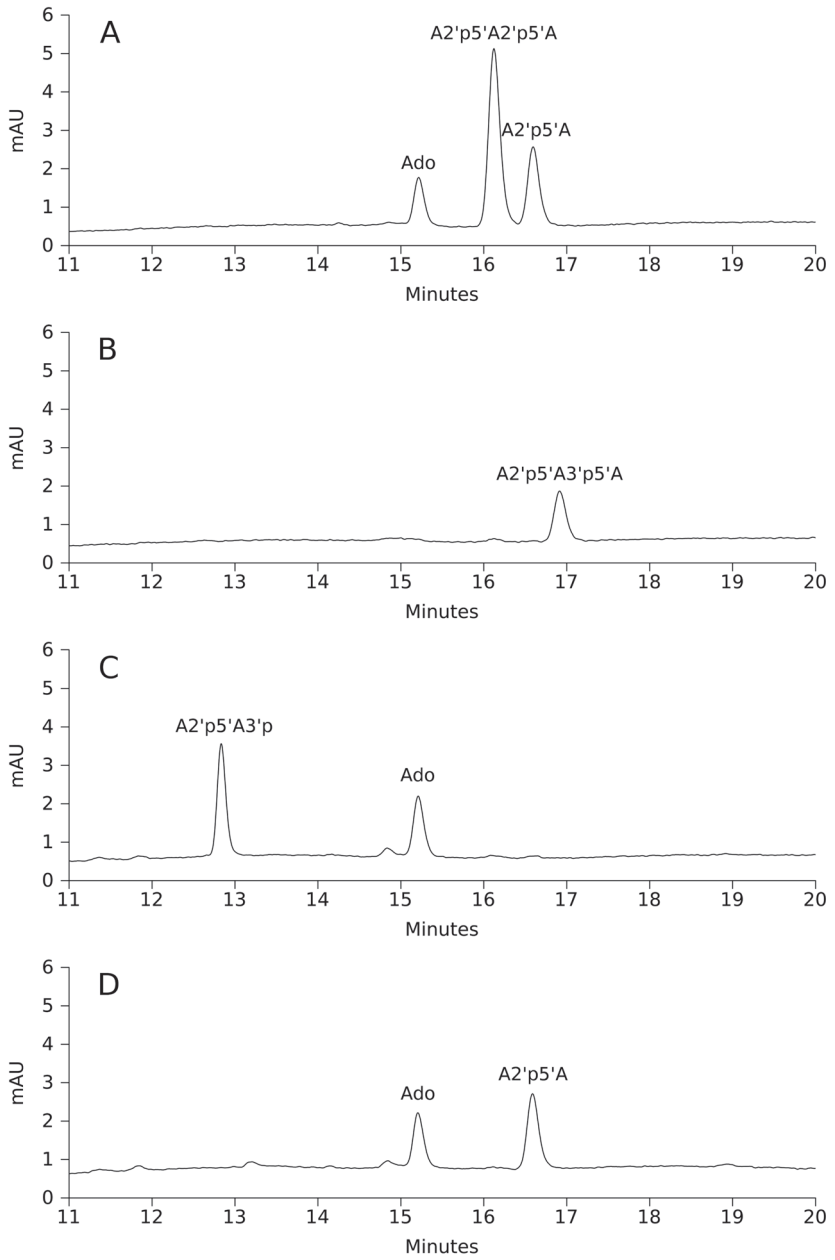


Fig. S6. Verification of A2'p5'A3'p5'A by enzymatic treatments. (A) HPLC analysis of the standard mixture of Ado, 2-5A dimer core (m/z 595.6) and trimer core (m/z 924.6), (B) HPLC analysis of a sample of A2'p5'A3'p5'A (m/z 924.7), (C) HPLC analysis of degradation products of A2'p5'A3'p5'A after RNaseT_2 treatment, (D) HPLC analysis of the RNase T_2 degradation products of A2'p5'A3'p5'A after dephosphorylation. The m⁄z values obtained from MALDI-MS analysis are shown in brackets.

Protein sequences

>MbrOAS
MLSVDEVYPGGSYAKGTSTGAQSDVDLLVVLNDIPTENHSRWLYMVLTELRRVVELGLQGNCQDVKVTR
FAVQLRVDNVNVDLLPVPAALWRSTDQRMKIVNSAALEDRRWLSVAFAKEQVKYVREKKETAGVTGIIR
LIKRWLATQKWRSKPPSFLIELMVLEAKSRLPEADHLALLNKFWELVRHHETLCIIHGSAVHAEIAHER
PLVVDPANPTNNVAGSFQWDEFVGFAREIHATRLRKFVRDLS

>TigOAS
MASPLIKEANQAVYTIVRTLQDGRPEGLVVDEVIKGGSMGHGTTVPGDYDIDIVIYSRSVDPQELARNG
PRRWLKIFAEHLEERNPGKLWDKEYLPHAFRFNFERKIKVDMLISPFWQRSSDLYRFLQERLSARERFN
FSMSASKWQVDFFKDQPNQVKEFIRRAKAWRNKKWAGQAGKPKSYLLSVIVLRAYERAKGKGDTYIAWN
TTAEVKTIVHRHQSADIYWEEYYRKQDYPSLFPQYTPRIVDPANPSNNLHETGISGPSNTKANDYGEGG
GRWDNFVRFVDSLDLTKSVDEIH

>McaOAS
MPGYLFLVYVFSCLSITSADLYHGPETVESHKVDKDDFVGRHHGETLSKFITRIVDPKKDYLLTLGRTI
DKVADLMRSKLAPHKVAATIKGGSLGKGTAVKQMADADLLFPLGGIGSVKELSAKLPEILKTLQSALDR
SEHKVTSIKKTPFTIQFQISIDGVLQDVDLLPIVDLGIQHLTKENLKPVYTEMKKHENLRDYYMKCLSF
LQVNFVKQQPPKVKSVIRLLKYWIKTKQHKLKSYGAELLVIKAYEDLGSPSSVREEDMAIKVFEKLTDL
KSLKVSWTKYFDPKNFNVPSAPYILDPACPYHNLIRQKSGGIDEKYVYLERDAKKILKVLKDQDYRSKR
GEL

>AmeOAS
MTMITPSSEIWTSFFFFSREARHCVGTREFGHYGRGQVKAAVDTICAFLKENCYRGPNSKIKVLKVVKG
GSSAKGTALKVGSDADLVVFLDCFHSYNDQYRMRKEIIEEISKKLEECKREKQFEVQIEKTKWKNPRVL
SFCMRSTKVPDFIDFDVLPAYNALGQLISGYKPPPEVYLRLIRDASPGENFSSCFTELQREFIISRPTK
VKSFIRLVKHWYKTKLHGNRNVPPKYAFELLSIYAWEKGSNGQPRFNMAQSFRNFLELICNYCELCVWW
TENYNLYDNAELGEFLEEQIEKPRPVILDPADPTGNFGVGYNWELVAIEAESCLSTMRYRTSAEQIRDF
RDLPPTLPDSLISRRPRYVPQSKSWSTDADNESSPPPTQQEASGTYCTIL
The linker sequence is underlined.

>LerOAS
MDLYQVLPGDLDTYIYKYLQPDERFLNQVAETIDKICTFLKEQCPYKISKAVKGGSLGKGTAVKNGSDA
DLVVFLNNFKSFKDQTGNRTEILRDICGLLLKYEKNIGHQIKMTPPKQNSRSLNFQFKSTRSSDFVEVD
VLPAFDALGQVTRTRAQVYEALIAAGESGGEFSPSFTELQRDFVKERPTKLKSLIRLVKYWYTEHVQRP
YKPQLRSDKDLPPKYALELLTIYAWESAGKGEDFSTAEGFCTVLELIVQYKHLCIFWTTNYSVDSSCIA
KCLKNKLKEPKPIILDPADPTGNVAAAARWDLVAAAAKCCLQQICVKDVNSWDVKPVKPFEITVTALGG
NSLRLTANIYFEVSELKRYIQQNWTIPVSQQRLVFSGTNLDDGKTLLDAMIFFDATIQLLVTIGIFVNY
NGSKVPIQVLLNDTVSSLKTIEPLKTLQPSQYYLTFQSRPLEDDRTLDYYNINQQSIIDVNLRLRGGSE
LQCFNANSMFLPSPHIM

ABSTRACT

2',5'-oligoadenylate synthetases (2-5A synthetases, OASs) are the enzymes that belong to the nucleotidyl transferase family. In several aspects OASs are distinct from other members of this large enzyme family. They catalyse the synthesis of 2',5'-phosphodiester linkages instead of the formation of the 3',5-linkages catalysed by most of the nucleotidyl transferases. Furthermore, they are able to initiate the synthesis of oligoadenylates from ATP without requiring any primer. In the case of mammalian OASs it has been shown that they require the binding of a cofactor, dsRNA, for the assembly of their active sites. It is in contrast with the constitutively active nucleotidyl transferases, which are locked in the enzymatically active conformation by the amino acid residues fixed in their protein cores.

Originally, the OASs were believed to be present only in higher vertebrates. Surprisingly, the OAS activity was later discovered in marine sponges and the genes containing OAS-specific conserved regions were cloned from sponges. To date, the presence of OAS genes has been bioinformatically predicted in the genomes of animals from many phyla. However, the OAS activity had been demonstrated only in higher vertebrates and sponges. No information was available about the presence of the OAS activity in animals belonging to the phyla that diverged between sponges and vertebrates.

In this work, bioinformatically predicted OASs from several distantly related multicellular animals as well as a unicellular organism were produced as recombinant proteins. The activity assays proved that the OASs from the marine sponges *Geodia cydonium* and *Tedania ignis*, the mollusk *Mytilus californianus* and the little skate *Leucoraja erinacea* were capable of catalysing the synthesis of 2',5'-oligoadenylates from ATP. These results show that enzymatically active OASs are widely distributed among multicellular animals, even in the Protostomia branch. The recombinant OAS protein from the unicellular organism, the choanoflagellate *Monosiga brevicollis* did not exhibit enzymatic activity.

The characterization of enzymatic properties of the recombinant OASs demonstrated that the sponge OASs differ from the mammalian enzymes. Their activation requirements, as well as product pattern were different from those of mammalian OASs. Surprisingly, the sponge OASs catalysed the formation of 3',5'-phosphodiester linkages in addition to 2',5'-linkages. The linkage tolerance of the sponge OASs may be due to the differences in their active site structures if compared to those of mammalian OASs. Some evidence was found that the cofactor, dsRNA may also have a role in determining the 2'- and 3'-specifities of sponge OASs. The variation in the enzymatic characteristics, observed between the OASs from two sponge species, *G. cydonium* and *T. ignis*, was not unexpected considering the diversity of their primary structures. The other OASs studied, those from the mollusk and the little skate, resembled mammalian OASs in their activation properties and linkage specificities.

The genomic structure of OAS genes from sponges was determined. It was substantially different from that of mammalian OASs, since their numbers of introns as well as the intron positions relative to the conserved amino acid motifs did not coincide. These data do not support the idea of a direct intron-containing common ancestor of mammalian and sponge OASs, but rather indicate that the early radiation of OAS genes has occurred.

Based on the obtained results, the existence of two subgroups of OASs, the sponge type OAS and the mammalian type OAS, was proposed. The analysis of genomes of distantly related animals from several animal phyla showed that both types of OAS genes are represented throughout the evolutionary tree of multicellular animals.

KOKKUVÕTE

2',5'-oligoadenülaadi süntetaasid (2-5A süntetaas, OAS) on ensüümid, mis kuuluvad nukleotidüültransferaaside perekonda. Nad erinevad teistest selle ensüümiperekonna liikmetest mitmete omaduste poolest. Nimelt katalüüsib OAS 2',5'-fosfodiestersideme, aga mitte 3',5'-sideme teket nagu enamik nukleotidüültransferaase. OAS on võimeline alustama oligoadenülaatide sünteesi ATPst ega vaja selleks praimerit. Imetajate OAS puhul on näidatud, et need ensüümid vajavad oma aktiivsaitide kokkupakkimiseks kofaktori – kaheahelalise RNA – juuresolekut. See eristab neid konstitutiivselt aktiivsetest nukleotidüültransferaasidest, mis on püsivalt lukustatud aktiivsesse konformatsiooni.

Algselt peeti 2-5A süntetaase omaseks ainult kõrgematele selgroogsetele loomadele (imetajatele, lindudele ja roomajatele), kuid hiljem avastati 2',5'-oligoadenülaate sünteesiv aktiivsus ka merekäsnades – evolutsiooniliselt kõige alamates hulkraksetes loomades. Merekäsnadest kloneeriti geenid, millelt kodeeritavad valgud sisaldasid oligoadenülaadi süntetaasidele iseloomulikke konserveerunud piirkondi ja motiive. Praeguseks on OAS geenide olemasolu ennustatud mitmesse erinevasse hõimkonda kuuluvate loomade genoomides, aga vastavat ensümaatilist aktiivsust on seni detekteeritud vaid kõrgemates selgroogsetes loomades ja käsnades.

Käesolevas töös toodeti bioinformaatiliselt ennustatud 2',5'-oligoadenülaadi süntetaasid rekombinantsete valkudena mitmest, üksteisest evolutsiooniliselt kaugest hulkraksest loomast, samuti ainuraksest organismist. Rekombinantsete valkudega läbi viidud ensümaatilise aktiivsuse katsed tõestasid, et merekäsnadest *Geodia cydonium* ja *Tedania ignis*, molluskist *Mytilus californianus* ning raist *Leucoraja erinacea* kloneeritud OAS oli võimeline katalüüsima 2',5'-oligoadenülaatide teket ATPst. Need tulemused näitasid, et ensümaatiliselt aktiivne OAS on hulkraksete loomade hulgas laialt levinud ning esineb isegi esmassuuste harus. Ainurakse organismi *Monosiga brevicollis* OAS ensümaatilist aktiivsust ei õnnestunud tuvastada.

Rekombinantsete OAS valkude iseloomustamine näitas, et käsnade oligoadenülaadi süntetaasid erinevad oluliselt imetajate vastavatest ensüümidest oma aktivatsiooniomaduste ja sünteesitavate produktide poolest. Selgus, et käsnade OAS katalüüsib lisaks 2',5'-sideme tekkele ka 3',5'-fosfodiestersideme teket. Katalüüsitava reaktsiooni mittespetsiifilisus võib olla tingitud sellest, et käsna OAS aktiivtsentri struktuur erineb imetajate omast. Katsed näitasid, et ka aktivaatorina kasutatav RNA võib osaleda käsna oligoadenülaadi süntetaaside 2'- või 3'-spetsiifilisuse määramises. Erinevatest käsnaliikidest pärinevatel OAS valkudel on erinevad ensüümatilised omadused, mis ei ole üllatav, arvestades käsnade OAS primaarstruktuuride varieeruvust. Samas sarnanevad molluskist ja raist kloneeritud OAS valgud oma aktivatsiooniomaduste ja katalüüsitava sideme spetsiifilisuse poolest imetajate ensüümidega.

Töö käigus määrati käsna OAS geenide ekson/intron struktuur ning näidati,

et see erineb oluliselt imetajate OAS geenide struktuurist. Käsnade ja imetajate OAS geenid erinevad omavahel intronite arvu poolest; intronite paiknemine konserveerunud aminohappejärjestuse motiivide suhtes on samuti erinev. Need tulemused ei kinnita varem välja pakutud hüpoteesi, et imetajate ja käsnade oligoadenülaadi süntetaasid on arenenud ühisest introneid sisaldavast eellasest, vaid viitavad pigem sellele, et juba varase evolutsiooni käigus on toimunud OAS geenide lahknemine.

Käesolevas töös saadud tulemused näitavad kahe OAS alamrühma (käsna-tüüp ja imetaja-tüüp) olemasolu. Erinevatesse hõimkondadesse kuuluvate loomade genoomide analüüsi põhjal võib väita, et mõlemad eespool nimetatud OAS tüübid on hulkraksete loomade seas levinud.

CURRICULUM VITAE

Personal data

Name:	Mailis Päri
Date and place of birth:	31.07.1981, Haapsalu, Estonia

Contact information

Address:	Department of Gene Technology, Tallinn University of Technology (TUT), Akadeemia tee 15, 12618
Phone:	+3726204435, +3726204436
E-mail:	mailis.pari@gmail.com

Education

2005-2014	TUT, Faculty of Science, Department of Gene Technology, PhD student
2003-2005	TUT, Faculty of Science, Department of Gene Technology, MSc
1999-2003	TUT, Faculty of Science, Department of Gene Technology, BSc
-1999	Haapsalu Wiedemanni Gümnaasium

Proffessional employment

2011-	researcher, TUT, Faculty of Science, Department of Gene Technology, Chair of Molecular Diagnostics

Supervised theses

Kaidi Kolk, MSc, 2014, (sup) Merike Kelve, Mailis Päri, Tõnu Reintamm, "Nucleoside phosphorylase/hydrolase domain containing protein from the freshwater sponge *Ephydatia muelleri*: cloning and determination of the genomic structure", TUT, Faculty of Science, Department of Gene Technology.

Publications

Päri M, Kuusksalu A, Lopp A, Kjaer K, Justesen J, Kelve M. (2014). Enzymatically active 2',5'-oligoadenylate synthetases are widely distributed among Metazoa, including protostome lineage. *Biochimie*, 97:200-209.

Reintamm T, Kuusksalu A, Metsis M, Päri M, Vallmann K, Lopp A, Justesen J, Kelve M. (2008). Sponge OAS has a distinct genomic structure within the 2-5A synthetase family. *Mol Genet Genomics*, 280(5):453-466.

Päri M, Kuusksalu A, Lopp A, Reintamm T, Justesen J, Kelve M. (2007). Expression and characterization of recombinant 2',5'-oligoadenylate

synthetase from the marine sponge *Geodia cydonium*. *FEBS J*, 274(13):3462-3474.

Lopp A, Reintamm T, Vallmann K, Päri M, Mikli V, Richelle-Maurer E, Kelve M. (2007). Molecular identification, characterization and distribution of freshwater sponges (Porifera: Spongillidae) in Estonia. *Fundamental and Applied Limnology (Archiv für Hydrobiologie)*, 168(1):93-103.

ELULOOKIRJELDUS

Isikuandmed

Nimi:	Mailis Päri
Sünniaeg ja -koht:	31.07.1981, Haapsalu, Eesti

Kontaktandmed

Aadress:	Geenitehnoloogia Instituut, Tallinna Tehnikaülikool (TTÜ), Akadeemia tee 15, 12618
Telefon:	+3726204435, +3726204436
E-mail:	mailis.pari@gmail.com

Hariduskäik

2005-2014	TTÜ, Matemaatika-loodusteaduskond, Geenitehnoloogia Instituut, doktorant
2003-2005	TTÜ, Matemaatika-loodusteaduskond, Geenitehnoloogia Instituut, magistrikraad
1999-2003	TTÜ, Matemaatika-loodusteaduskond, Geenitehnoloogia Instituut, bakalaureusekraad
-1999	Haapsalu Wiedemanni Gümnaasium

Teenistuskäik

2011-	teadur, TTÜ, Matemaatika-loodusteaduskond, Geenitehnoloogia Instituut, Molekulaardiagnostika õppetool

Juhendatud lõputööd

Kaidi Kolk, magistrikraad, 2014, (juh) Merike Kelve, Mailis Päri, Tõnu Reintamm, "Nukleosiidi fosforülaasi/hüdrolaasi domeeni sisaldav valk mageveekäsnast *Ephydatia muelleri*: kloneerimine ja genoomse struktuuri määramine", TTÜ, Matemaatika-loodusteaduskond, Geenitehnoloogia Instituut.

Publikatsioonid

Päri M, Kuusksalu A, Lopp A, Kjaer K, Justesen J, Kelve M. (2014). Enzymatically active 2',5'-oligoadenylate synthetases are widely distributed among Metazoa, including protostome lineage. *Biochimie*, 97:200-209.

Reintamm T, Kuusksalu A, Metsis M, Päri M, Vallmann K, Lopp A, Justesen J, Kelve M. (2008). Sponge OAS has a distinct genomic structure within the 2-5A synthetase family. *Mol Genet Genomics*, 280(5):453-466.

Päri M, Kuusksalu A, Lopp A, Reintamm T, Justesen J, Kelve M. (2007). Expression and characterization of recombinant 2',5'-oligoadenylate

synthetase from the marine sponge *Geodia cydonium*. *FEBS J*, 274(13):3462-3474.

Lopp A, Reintamm T, Vallmann K, Päri M, Mikli V, Richelle-Maurer E, Kelve M. (2007). Molecular identification, characterization and distribution of freshwater sponges (Porifera: Spongillidae) in Estonia. *Fundamental and Applied Limnology (Archiv für Hydrobiologie)*, 168(1):93-103.

DISSERTATIONS DEFENDED AT TALLINN UNIVERSITY OF TECHNOLOGY ON *NATURAL AND EXACT SCIENCES*

1. **Olav Kongas**. Nonlinear Dynamics in Modeling Cardiac Arrhytmias. 1998.
2. **Kalju Vanatalu**. Optimization of Processes of Microbial Biosynthesis of Isotopically Labeled Biomolecules and Their Complexes. 1999.
3. **Ahto Buldas**. An Algebraic Approach to the Structure of Graphs. 1999.
4. **Monika Drews**. A Metabolic Study of Insect Cells in Batch and Continuous Culture: Application of Chemostat and Turbidostat to the Production of Recombinant Proteins. 1999.
5. **Eola Valdre**. Endothelial-Specific Regulation of Vessel Formation: Role of Receptor Tyrosine Kinases. 2000.
6. **Kalju Lott**. Doping and Defect Thermodynamic Equilibrium in ZnS. 2000.
7. **Reet Koljak**. Novel Fatty Acid Dioxygenases from the Corals *Plexaura homomalla* and *Gersemia fruticosa*. 2001.
8. **Anne Paju**. Asymmetric oxidation of Prochiral and Racemic Ketones by Using Sharpless Catalyst. 2001.
9. **Marko Vendelin**. Cardiac Mechanoenergetics *in silico*. 2001.
10. **Pearu Peterson**. Multi-Soliton Interactions and the Inverse Problem of Wave Crest. 2001.
11. **Anne Menert**. Microcalorimetry of Anaerobic Digestion. 2001.
12. **Toomas Tiivel**. The Role of the Mitochondrial Outer Membrane in *in vivo* Regulation of Respiration in Normal Heart and Skeletal Muscle Cell. 2002.
13. **Olle Hints**. Ordovician Scolecodonts of Estonia and Neighbouring Areas: Taxonomy, Distribution, Palaeoecology, and Application. 2002.
14. **Jaak Nõlvak**. Chitinozoan Biostratigrapy in the Ordovician of Baltoscandia. 2002.
15. **Liivi Kluge**. On Algebraic Structure of Pre-Operad. 2002.
16. **Jaanus Lass**. Biosignal Interpretation: Study of Cardiac Arrhytmias and Electromagnetic Field Effects on Human Nervous System. 2002.
17. **Janek Peterson**. Synthesis, Structural Characterization and Modification of PAMAM Dendrimers. 2002.
18. **Merike Vaher**. Room Temperature Ionic Liquids as Background Electrolyte Additives in Capillary Electrophoresis. 2002.
19. **Valdek Mikli**. Electron Microscopy and Image Analysis Study of Powdered Hardmetal Materials and Optoelectronic Thin Films. 2003.
20. **Mart Viljus**. The Microstructure and Properties of Fine-Grained Cermets. 2003.
21. **Signe Kask**. Identification and Characterization of Dairy-Related *Lactobacillus*. 2003
22. **Tiiu-Mai Laht**. Influence of Microstructure of the Curd on Enzymatic and Microbiological Processes in Swiss-Type Cheese. 2003.
23. **Anne Kuusksalu**. 2–5A Synthetase in the Marine Sponge *Geodia cydonium*. 2003.
24. **Sergei Bereznev**. Solar Cells Based on Polycristalline Copper-Indium

Chalcogenides and Conductive Polymers. 2003.
25. **Kadri Kriis**. Asymmetric Synthesis of C_2-Symmetric Bimorpholines and Their Application as Chiral Ligands in the Transfer Hydrogenation of Aromatic Ketones. 2004.
26. **Jekaterina Reut**. Polypyrrole Coatings on Conducting and Insulating Substracts. 2004.
27. **Sven Nõmm**. Realization and Identification of Discrete-Time Nonlinear Systems. 2004.
28. **Olga Kijatkina**. Deposition of Copper Indium Disulphide Films by Chemical Spray Pyrolysis. 2004.
29. **Gert Tamberg**. On Sampling Operators Defined by Rogosinski, Hann and Blackman Windows. 2004.
30. **Monika Übner**. Interaction of Humic Substances with Metal Cations. 2004.
31. **Kaarel Adamberg**. Growth Characteristics of Non-Starter Lactic Acid Bacteria from Cheese. 2004.
32. **Imre Vallikivi**. Lipase-Catalysed Reactions of Prostaglandins. 2004.
33. **Merike Peld**. Substituted Apatites as Sorbents for Heavy Metals. 2005.
34. **Vitali Syritski**. Study of Synthesis and Redox Switching of Polypyrrole and Poly(3,4-ethylenedioxythiophene) by Using *in-situ* Techniques. 2004.
35. **Lee Põllumaa**. Evaluation of Ecotoxicological Effects Related to Oil Shale Industry. 2004.
36. **Riina Aav**. Synthesis of 9,11-Secosterols Intermediates. 2005.
37. **Andres Braunbrück**. Wave Interaction in Weakly Inhomogeneous Materials. 2005.
38. **Robert Kitt**. Generalised Scale-Invariance in Financial Time Series. 2005.
39. **Juss Pavelson**. Mesoscale Physical Processes and the Related Impact on the Summer Nutrient Fields and Phytoplankton Blooms in the Western Gulf of Finland. 2005.
40. **Olari Ilison**. Solitons and Solitary Waves in Media with Higher Order Dispersive and Nonlinear Effects. 2005.
41. **Maksim Säkki**. Intermittency and Long-Range Structurization of Heart Rate. 2005.
42. **Enli Kiipli**. Modelling Seawater Chemistry of the East Baltic Basin in the Late Ordovician–Early Silurian. 2005.
43. **Igor Golovtsov**. Modification of Conductive Properties and Processability of Polyparaphenylene, Polypyrrole and polyaniline. 2005.
44. **Katrin Laos**. Interaction Between Furcellaran and the Globular Proteins (Bovine Serum Albumin β-Lactoglobulin). 2005.
45. **Arvo Mere**. Structural and Electrical Properties of Spray Deposited Copper Indium Disulphide Films for Solar Cells. 2006.
46. **Sille Ehala**. Development and Application of Various On- and Off-Line Analytical Methods for the Analysis of Bioactive Compounds. 2006.
47. **Maria Kulp**. Capillary Electrophoretic Monitoring of Biochemical Reaction Kinetics. 2006.
48. **Anu Aaspõllu.** Proteinases from *Vipera lebetina* Snake Venom Affecting Hemostasis. 2006.
49. **Lyudmila Chekulayeva**. Photosensitized Inactivation of Tumor Cells by

Porphyrins and Chlorins. 2006.
50. **Merle Uudsemaa**. Quantum-Chemical Modeling of Solvated First Row Transition Metal Ions. 2006.
51. **Tagli Pitsi**. Nutrition Situation of Pre-School Children in Estonia from 1995 to 2004. 2006.
52. **Angela Ivask**. Luminescent Recombinant Sensor Bacteria for the Analysis of Bioavailable Heavy Metals. 2006.
53. **Tiina Lõugas**. Study on Physico-Chemical Properties and Some Bioactive Compounds of Sea Buckthorn (*Hippophae rhamnoides* L.). 2006.
54. **Kaja Kasemets**. Effect of Changing Environmental Conditions on the Fermentative Growth of *Saccharomyces cerevisae* S288C: Auxo-accelerostat Study. 2006.
55. **Ildar Nisamedtinov**. Application of ^{13}C and Fluorescence Labeling in Metabolic Studies of *Saccharomyces* spp. 2006.
56. **Alar Leibak**. On Additive Generalisation of Voronoï's Theory of Perfect Forms over Algebraic Number Fields. 2006.
57. **Andri Jagomägi**. Photoluminescence of Chalcopyrite Tellurides. 2006.
58. **Tõnu Martma**. Application of Carbon Isotopes to the Study of the Ordovician and Silurian of the Baltic. 2006.
59. **Marit Kauk**. Chemical Composition of $CuInSe_2$ Monograin Powders for Solar Cell Application. 2006.
60. **Julia Kois**. Electrochemical Deposition of $CuInSe_2$ Thin Films for Photovoltaic Applications. 2006.
61. **Ilona Oja Açik**. Sol-Gel Deposition of Titanium Dioxide Films. 2007.
62. **Tiia Anmann**. Integrated and Organized Cellular Bioenergetic Systems in Heart and Brain. 2007.
63. **Katrin Trummal**. Purification, Characterization and Specificity Studies of Metalloproteinases from *Vipera lebetina* Snake Venom. 2007.
64. **Gennadi Lessin**. Biochemical Definition of Coastal Zone Using Numerical Modeling and Measurement Data. 2007.
65. **Enno Pais**. Inverse problems to determine non-homogeneous degenerate memory kernels in heat flow. 2007.
66. **Maria Borissova**. Capillary Electrophoresis on Alkylimidazolium Salts. 2007.
67. **Karin Valmsen**. Prostaglandin Synthesis in the Coral *Plexaura homomalla*: Control of Prostaglandin Stereochemistry at Carbon 15 by Cyclooxygenases. 2007.
68. **Kristjan Piirimäe**. Long-Term Changes of Nutrient Fluxes in the Drainage Basin of the Gulf of Finland – Application of the PolFlow Model. 2007.
69. **Tatjana Dedova**. Chemical Spray Pyrolysis Deposition of Zinc Sulfide Thin Films and Zinc Oxide Nanostructured Layers. 2007.
70. **Katrin Tomson**. Production of Labelled Recombinant Proteins in Fed-Batch Systems in *Escherichia coli*. 2007.
71. **Cecilia Sarmiento**. Suppressors of RNA Silencing in Plants. 2008.
72. **Vilja Mardla**. Inhibition of Platelet Aggregation with Combination of Antiplatelet Agents. 2008.
73. **Maie Bachmann**. Effect of Modulated Microwave Radiation on Human Resting Electroencephalographic Signal. 2008.
74. **Dan Hüvonen**. Terahertz Spectroscopy of Low-Dimensional Spin Systems.

2008.
75. **Ly Villo**. Stereoselective Chemoenzymatic Synthesis of Deoxy Sugar Esters Involving *Candida antarctica* Lipase B. 2008.
76. **Johan Anton**. Technology of Integrated Photoelasticity for Residual Stress Measurement in Glass Articles of Axisymmetric Shape. 2008.
77. **Olga Volobujeva**. SEM Study of Selenization of Different Thin Metallic Films. 2008.
78. **Artur Jõgi**. Synthesis of 4'-Substituted 2,3'-dideoxynucleoside Analogues. 2008.
79. **Mario Kadastik**. Doubly Charged Higgs Boson Decays and Implications on Neutrino Physics. 2008.
80. **Fernando Pérez-Caballero**. Carbon Aerogels from 5-Methylresorcinol-Formaldehyde Gels. 2008.
81. **Sirje Vaask**. The Comparability, Reproducibility and Validity of Estonian Food Consumption Surveys. 2008.
82. **Anna Menaker**. Electrosynthesized Conducting Polymers, Polypyrrole and Poly(3,4-ethylenedioxythiophene), for Molecular Imprinting. 2009.
83. **Lauri Ilison**. Solitons and Solitary Waves in Hierarchical Korteweg-de Vries Type Systems. 2009.
84. **Kaia Ernits**. Study of In_2S_3 and ZnS Thin Films Deposited by Ultrasonic Spray Pyrolysis and Chemical Deposition. 2009.
85. **Veljo Sinivee**. Portable Spectrometer for Ionizing Radiation "Gammamapper". 2009.
86. **Jüri Virkepu**. On Lagrange Formalism for Lie Theory and Operadic Harmonic Oscillator in Low Dimensions. 2009.
87. **Marko Piirsoo**. Deciphering Molecular Basis of Schwann Cell Development. 2009.
88. **Kati Helmja**. Determination of Phenolic Compounds and Their Antioxidative Capability in Plant Extracts. 2010.
89. **Merike Sõmera**. Sobemoviruses: Genomic Organization, Potential for Recombination and Necessity of P1 in Systemic Infection. 2010.
90. **Kristjan Laes**. Preparation and Impedance Spectroscopy of Hybrid Structures Based on $CuIn_3Se_5$ Photoabsorber. 2010.
91. **Kristin Lippur**. Asymmetric Synthesis of 2,2'-Bimorpholine and its 5,5'-Substituted Derivatives. 2010.
92. **Merike Luman**. Dialysis Dose and Nutrition Assessment by an Optical Method. 2010.
93. **Mihhail Berezovski**. Numerical Simulation of Wave Propagation in Heterogeneous and Microstructured Materials. 2010.
94. **Tamara Aid-Pavlidis**. Structure and Regulation of BDNF Gene. 2010.
95. **Olga Bragina**. The Role of Sonic Hedgehog Pathway in Neuro- and Tumorigenesis. 2010.
96. **Merle Randrüüt**. Wave Propagation in Microstructured Solids: Solitary and Periodic Waves. 2010.
97. **Marju Laars**. Asymmetric Organocatalytic Michael and Aldol Reactions Mediated by Cyclic Amines. 2010.
98. **Maarja Grossberg**. Optical Properties of Multinary Semiconductor

Compounds for Photovoltaic Applications. 2010.
99. **Alla Maloverjan**. Vertebrate Homologues of Drosophila Fused Kinase and Their Role in Sonic Hedgehog Signalling Pathway. 2010.
100. **Priit Pruunsild**. Neuronal Activity-Dependent Transcription Factors and Regulation of Human *BDNF* Gene. 2010.
101. **Tatjana Knjazeva**. New Approaches in Capillary Electrophoresis for Separation and Study of Proteins. 2011.
102. **Atanas Katerski**. Chemical Composition of Sprayed Copper Indium Disulfide Films for Nanostructured Solar Cells. 2011.
103. **Kristi Timmo.** Formation of Properties of $CuInSe_2$ and $Cu_2ZnSn(S,Se)_4$ Monograin Powders Synthesized in Molten KI. 2011.
104. **Kert Tamm**. Wave Propagation and Interaction in Mindlin-Type Microstructured Solids: Numerical Simulation. 2011.
105. **Adrian Popp**. Ordovician Proetid Trilobites in Baltoscandia and Germany. 2011.
106. **Ove Pärn**. Sea Ice Deformation Events in the Gulf of Finland and This Impact on Shipping. 2011.
107. **Germo Väli**. Numerical Experiments on Matter Transport in the Baltic Sea. 2011.
108. **Andrus Seiman**. Point-of-Care Analyser Based on Capillary Electrophoresis. 2011.
109. **Olga Katargina**. Tick-Borne Pathogens Circulating in Estonia (Tick-Borne Encephalitis Virus, *Anaplasma phagocytophilum*, *Babesia* Species): Their Prevalence and Genetic Characterization. 2011.
110. **Ingrid Sumeri**. The Study of Probiotic Bacteria in Human Gastrointestinal Tract Simulator. 2011.
111. **Kairit Zovo**. Functional Characterization of Cellular Copper Proteome. 2011.
112. **Natalja Makarytsheva**. Analysis of Organic Species in Sediments and Soil by High Performance Separation Methods. 2011.
113. **Monika Mortimer**. Evaluation of the Biological Effects of Engineered Nanoparticles on Unicellular Pro- and Eukaryotic Organisms. 2011.
114. **Kersti Tepp**. Molecular System Bioenergetics of Cardiac Cells: Quantitative Analysis of Structure-Function Relationship. 2011.
115. **Anna-Liisa Peikolainen**. Organic Aerogels Based on 5-Methylresorcinol. 2011.
116. **Leeli Amon**. Palaeoecological Reconstruction of Late-Glacial Vegetation Dynamics in Eastern Baltic Area: A View Based on Plant Macrofossil Analysis. 2011.
117. **Tanel Peets**. Dispersion Analysis of Wave Motion in Microstructured Solids. 2011.
118. **Liina Kaupmees**. Selenization of Molybdenum as Contact Material in Solar Cells. 2011.
119. **Allan Olspert**. Properties of VPg and Coat Protein of Sobemoviruses. 2011.
120. **Kadri Koppel**. Food Category Appraisal Using Sensory Methods. 2011.
121. **Jelena Gorbatšova**. Development of Methods for CE Analysis of Plant Phenolics and Vitamins. 2011.
122. **Karin Viipsi**. Impact of EDTA and Humic Substances on the Removal of Cd

and Zn from Aqueous Solutions by Apatite. 2012.
123. **David Schryer**. Metabolic Flux Analysis of Compartmentalized Systems Using Dynamic Isotopologue Modeling. 2012.
124. **Ardo Illaste**. Analysis of Molecular Movements in Cardiac Myocytes. 2012.
125. **Indrek Reile**. 3-Alkylcyclopentane-1,2-Diones in Asymmetric Oxidation and Alkylation Reactions. 2012.
126. **Tatjana Tamberg**. Some Classes of Finite 2-Groups and Their Endomorphism Semigroups. 2012.
127. **Taavi Liblik**. Variability of Thermohaline Structure in the Gulf of Finland in Summer. 2012.
128. **Priidik Lagemaa**. Operational Forecasting in Estonian Marine Waters. 2012.
129. **Andrei Errapart**. Photoelastic Tomography in Linear and Non-linear Approximation. 2012.
130. **Külliki Krabbi**. Biochemical Diagnosis of Classical Galactosemia and Mucopolysaccharidoses in Estonia. 2012.
131. **Kristel Kaseleht**. Identification of Aroma Compounds in Food using SPME-GC/MS and GC-Olfactometry. 2012.
132. **Kristel Kodar**. Immunoglobulin G Glycosylation Profiling in Patients with Gastric Cancer. 2012.
133. **Kai Rosin**. Solar Radiation and Wind as Agents of the Formation of the Radiation Regime in Water Bodies. 2012.
134. **Ann Tiiman**. Interactions of Alzheimer's Amyloid-Beta Peptides with Zn(II) and Cu(II) Ions. 2012.
135. **Olga Gavrilova**. Application and Elaboration of Accounting Approaches for Sustainable Development. 2012.
136. **Olesja Bondarenko**. Development of Bacterial Biosensors and Human Stem Cell-Based *In Vitro* Assays for the Toxicological Profiling of Synthetic Nanoparticles. 2012.
137. **Katri Muska**. Study of Composition and Thermal Treatments of Quaternary Compounds for Monograin Layer Solar Cells. 2012.
138. **Ranno Nahku**. Validation of Critical Factors for the Quantitative Characterization of Bacterial Physiology in Accelerostat Cultures. 2012.
139. **Petri-Jaan Lahtvee**. Quantitative Omics-level Analysis of Growth Rate Dependent Energy Metabolism in *Lactococcus lactis*. 2012.
140. **Kerti Orumets**. Molecular Mechanisms Controlling Intracellular Glutathione Levels in Baker's Yeast *Saccharomyces cerevisiae* and its Random Mutagenized Glutathione Over-Accumulating Isolate. 2012.
141. **Loreida Timberg**. Spice-Cured Sprats Ripening, Sensory Parameters Development, and Quality Indicators. 2012.
142. **Anna Mihhalevski**. Rye Sourdough Fermentation and Bread Stability. 2012.
143. **Liisa Arike**. Quantitative Proteomics of *Escherichia coli*: From Relative to Absolute Scale. 2012.
144. **Kairi Otto**. Deposition of In_2S_3 Thin Films by Chemical Spray Pyrolysis. 2012.
145. **Mari Sepp**. Functions of the Basic Helix-Loop-Helix Transcription Factor TCF4 in Health and Disease. 2012.
146. **Anna Suhhova**. Detection of the Effect of Weak Stressors on Human Resting

Electroencephalographic Signal. 2012.
147. **Aram Kazarjan**. Development and Production of Extruded Food and Feed Products Containing Probiotic Microorganisms. 2012.
148. **Rivo Uiboupin**. Application of Remote Sensing Methods for the Investigation of Spatio-Temporal Variability of Sea Surface Temperature and Chlorophyll Fields in the Gulf of Finland. 2013.
149. **Tiina Kriščiunaite**. A Study of Milk Coagulability. 2013.
150. **Tuuli Levandi**. Comparative Study of Cereal Varieties by Analytical Separation Methods and Chemometrics. 2013.
151. **Natalja Kabanova**. Development of a Microcalorimetric Method for the Study of Fermentation Processes. 2013.
152. **Himani Khanduri**. Magnetic Properties of Functional Oxides. 2013.
153. **Julia Smirnova**. Investigation of Properties and Reaction Mechanisms of Redox-Active Proteins by ESI MS. 2013.
154. **Mervi Sepp**. Estimation of Diffusion Restrictions in Cardiomyocytes Using Kinetic Measurements. 2013.
155. **Kersti Jääger**. Differentiation and Heterogeneity of Mesenchymal Stem Cells. 2013.
156. **Victor Alari**. Multi-Scale Wind Wave Modeling in the Baltic Sea. 2013.
157. **Taavi Päll**. Studies of CD44 Hyaluronan Binding Domain as Novel Angiogenesis Inhibitor. 2013.
158. **Allan Niidu**. Synthesis of Cyclopentane and Tetrahydrofuran Derivatives. 2013.
159. **Julia Geller**. Detection and Genetic Characterization of *Borrelia* Species Circulating in Tick Population in Estonia. 2013.
160. **Irina Stulova**. The Effects of Milk Composition and Treatment on the Growth of Lactic Acid Bacteria. 2013.
161. **Jana Holmar**. Optical Method for Uric Acid Removal Assessment During Dialysis. 2013.
162. **Kerti Ausmees**. Synthesis of Heterobicyclo[3.2.0]heptane Derivatives *via* Multicomponent Cascade Reaction. 2013.
163. **Minna Varikmaa**. Structural and Functional Studies of Mitochondrial Respiration Regulation in Muscle Cells. 2013.
164. **Indrek Koppel**. Transcriptional Mechanisms of BDNF Gene Regulation. 2014.
165. **Kristjan Pilt**. Optical Pulse Wave Signal Analysis for Determination of Early Arterial Ageing in Diabetic Patients. 2014.
166. **Andres Anier**. Estimation of the Complexity of the Electroencephalogram for Brain Monitoring in Intensive Care. 2014.
167. **Toivo Kallaste**. Pyroclastic Sanidine in the Lower Palaeozoic Bentonites – A Tool for Regional Geological Correlations. 2014.
168. **Erki Kärber**. Properties of ZnO-nanorod/In_2S_3/$CuInS_2$ Solar Cell and the Constituent Layers Deposited by Chemical Spray Method. 2014.
169. **Julia Lehner**. Formation of Cu_2ZnSnS_4 and $Cu_2ZnSnSe_4$ by Chalcogenisation of Electrochemically Deposited Precursor Layers. 2014.
170. **Peep Pitk**. Protein- and Lipid-rich Solid Slaughterhouse Waste Anaerobic Co-digestion: Resource Analysis and Process Optimization. 2014.

171. **Kaspar Valgepea**. Absolute Quantitative Multi-omics Characterization of Specific Growth Rate-dependent Metabolism of *Escherichia coli*. 2014.
172. **Artur Noole**. Asymmetric Organocatalytic Synthesis of 3,3'-Disubstituted Oxindoles. 2014.
173. **Robert Tsanev**. Identification and Structure-Functional Characterisation of the Gene Transcriptional Repressor Domain of Human Gli Proteins. 2014.
174. **Dmitri Kartofelev**. Nonlinear Sound Generation Mechanisms in Musical Acoustic. 2014.
175. **Sigrid Hade**. GIS Applications in the Studies of the Palaeozoic Graptolite Argillite and Landscape Change. 2014.