

Contents lists available at ScienceDirect

Data in Brief

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

Data Article

# Data on four apoptosis-related genes in the colonial tunicate *Botryllus schlosseri*



Nicola Franchi, Francesca Ballin, Lucia Manni\*, Filippo Schiavon, Loriano Ballarin

Department of Biology, University of Padova, Italy

## A R T I C L E I N F O

Article history: Received 13 April 2016 Received in revised form 3 May 2016 Accepted 11 May 2016 Available online 20 May 2016 Keywords: Sequence databases Alignments **BLAST** analysis Botryllus schlosseri Apoptosis-related transcripts Bax AIF1 PARP1 IAP7

## ABSTRACT

The data described are related to the article entitled "Recurrent phagocytosis-induced apoptosis in the cyclical generation change of the compound ascidian *Botryllus schlosseri*" (Franchi et al., 2016) [1]. Four apoptosis-related genes, showing high similarity with mammalian Bax (a member of the Bcl-2 protein family), AIF1 (apoptosis-inducing factor-1), PARP1 (poly ADP ribose polymerase-1) and IAP7 (inhibitor of apoptosis-7) were identified from the analysis of the trascriptome of *B. schlosseri*. They were named BsBax, BsAIF1, BsPARP1 and BsIAP7. Here, their deduced amino acid sequence were compared with known sequences of orthologous genes from other deuterostome species together with a study of their identity/similarity.

© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

## Specifications Table

Subject area	Biology
More specific sub-	Developmental Biology
ject area	
Type of data	Tables, figures

DOI of original article: http://dx.doi.org/10.1016/j.dci.2016.04.011

\* Corresponding author at: Department of Biology, University of Padova via U. Bassi 58/B 35131 Padova Italy.

Tel.: +39 049 8276252; fax: +39 049 8276199.

E-mail address: lucia.manni@unipd.it (L. Manni).

http://dx.doi.org/10.1016/j.dib.2016.05.017

2352-3409/© 2016 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Bioinformatic analysis, RACE
Raw and analysed data
The partial transcripts present in the transcriptome, identified by BLAST analysis, were elongated through 5'and 3' RACE according to the 2nd generation of 5'/3' RACE kit
Analysis with BLAST, LALIGN, SMART, Clustal Omega
Padova, Italy
Data are available in this article and at GenBank via accession numbers Gen- Bank: KU948200 for BsBAX, GenBank: KU948201 for BsPARP1, GenBank: KU948202 for BsAIF1, GenBank: KU948203 for BsIAP7.

## Value of the data

- The data provide the full-length sequences of four apoptosis-related transcripts from the colonial ascidian *B. schlosseri* useful to study the phylogeny trees of the corresponding proteins in chordates.
- From the data, the protein primary structures can be deduced and, from that, three-dimensional models can be obtained, useful to compare the domain organization of the corresponding chordate proteins.
- Expression studies, exploiting the present data, can contribute to elucidate the dynamics of the cyclical apoptosis, which characterizes the colonial blastogenetic cycle of the ascidian *B. schlosseri*.

## 1. Data

The data reported include supporting information to the phylogenetic analyses of Franchi et al. [1]. They consist of transcript sequences, sequence alignments and comparisons of four apoptosis-related genes identified in the recently-obtained transcriptome of *B. schlosseri* [2]. The sequences show high similarity with mammalian transcripts for Bax, AIF1, PARP1 and IAP7 and were named BsBax, BsAIF1, BsPARP1 and BsIAP7, respectively. The expression of these genes was studied further in the above-reported paper [1].

#### 2. Experimental design, materials and methods

Amplification and cloning of transcripts for BsBax, BsAIF1, BsPARP1 and BsIAP7 was achieved with specific primers designed on sequences found in our collection of transcriptomes [2]. In order to verify and complete the full length cDNA, PCR reactions were carried out with a denaturing step at 94 °C for 2 min, 40 cycles of 30 s at 94 °C, 40 s at 60 °C and 90 s at 72 °C, and a final extension at 72 °C for 10 min. Amplicons were separated using 1.5% agarose gel, purified, cloned and sequenced. The partial transcripts were elongated through 5′ and 3′ RACE according to the 2nd generation of 5′/ 3′ RACE kit (Roche). Supplementary Table 1 reports the specific primers used for amplicons production and their elongation through 5′- and 3′-RACE and for the *in situ* hybridisation experiments reported in [1].

The sequences from GenBank, reported in Supplementary Tables 2–5, were used for alignments and sequence comparisons with the sequences of BsBax, BsAIF1, BsPARP1, BsIAP7, respectively. The

latter were deposited in GenBank and corresponding to the accession numbers GenBank: KU948200, GenBank: KU948201, GenBank: KU948202 and GenBank: KU948203, respectively (Figs. 1–8).

The predicted amino acid sequences of BsBax, BsAIF1, BsPARP1, BsIAP7 are reported in Figs. 2, 4, 6 and 8, respectively. They were aligned with known orthologous sequences from both vertebrate and invertebrates using the MUSCLE programme [3] and data are reported in Figs. 1, 3, 5, and 7, respectively. Identity analysis of the deduced amino acid sequences were performed using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) and LALIGN (http://www.ch.embnet.org/software/LALIGN\_form.html) [4] and are reported in Supplementary Tables 2–5.



Fig. 1. Alignments of the deduced amino acid sequence of BsBax with known orthologous sequences from both vertebrate and invertebrates.

	TGTTA																			
	TAAGTT	GATI	rggc	GAI	ſŦĠĊ	CTC	FTC.	AGG	CTC	TCC	TTG	ATT	CGA	AGT	ΓGΤ	TAG	TAT.	ATA		
1	GGCAGC	CGCI	FATC	CTAP	ACAI	[AT]	AGA'	TAT.	ACG	GTA	TAC	CAT	TCC.	ATT	ICT.	TGT	GTG	GCA		
1																			M ATG	
2	S G	G	G	D	Q	А	K	K	Е	А	G	K	Ι	D	D	K	Ν	D	Е	
4	TCGGGG	GGC	GGTG	GAT	CAAG	GCG.	AAG	AAG	GAA	GCA	GGG	AAG	ATA	GAT	GAT	AAA	AAT	GAC	GAG	
22	D P	R	Ρ	S	R	Ρ	Т	R	Ρ	R	Т	А	R	Ρ	G	R	R	Т	D	
64	GATCCT	CGA	CAI	CTF	4GG(	CCA.	ACC	CGA	CCG.	AGA	ACC	GCC	AGA	CCA	GGT	AGA	AGA	ACA	GAT	
12	D F	c	c	т	17	0	c	7	V	C	C	F	C	c	v	T	D	u	т	
124	CCTGAA	UTCG	ГСТТ	л ГТGC	Y STC	CAG	GGA	GCT	GTT	GGT	GGG	GAA	GGT	U TCA'	TAC	ACG	AGA	CAT	CTA	
101	001011				510.	0110	0011	001	011	001	000	01111	001	- 011				0111		
62	Т Е	Е	Е	R	А	V	G	Ν	Q	А	R	Н	L	L	Т	М	F	Ι	S	
184	ACTGAG	GAG	GAAA	4GAC	GCT	GTT	GGA.	AAC	CAA	GCA	AGG	CAT	CTT	CTG.	ACA	ATG	TTT	ATC	TCT	
		_	_	_	_	_	_	_			_							_	_	
82	DR	A	E	R	D	P	D	L	S	Q	D	K	K	G	A	V	Q	E	T	
244	GACAGA	IGCAU	JAAC	-GI(	эАтс		GAC	CTG	TCG	CAG	GAI	AAG	AAA	GGA	GCT	GTA	CAG	GAA	ACA	
102	v o	L	С	Ν	К	S	S	Е	S	Ν	А	М	S	D	S	D	А	А	L	
304	GTTCAG	- STTA:	- FGT#	AACI	AAGZ	AGC'	TCG	GAA	AGT	AAT	GCG	ATG	AGC	GAT.	AGC	- GAT	GCT	GCT	TTA	
122	R E	V	S	L	Т	L	R	V	Ι	G	D	Q	L	N	Е	D	Ν	D	L	
364	AGGGAG	GTA:	FCAI	ľΤG	ACC.	ΓTG	CGT	GTT	ATT	GGA	GAT	CAG	TTG.	AAT	GAA	GAC	AAT	GAC	CTT	
142	M C	т	т	D	D	37	т	17	0	D	m	V	D	т	E.	м	V	17	C	
424	AACAGO	ц СТТ7	⊥ ∿∏TC	Ъ ЗАТ7	AGA	ን ዓጥጥ	ц СТС	V GTT	CAA	- CCC	ACC	AAG	GAC	⊥ ATC	r TTC	м Атg	AAA	v GTT	TGT	
101						0.1.1	0 1 0	0	01111				0110					0.1.1		
162	K <u>Q</u>	I	F	А	D	Q	Ν	F	Ν	Ŵ	G	R	V	А	А	V	F	Y	F	
484	AAACAG	GATA:	LLLL	3CCC	GAT	CAA.	AAT	TTC	AAC	TGG	GGC	AGA	GTA	GCC	GCT	GTA	TTC	TAC	TTT	
1.0.0		-	-	-	-				~	-	-			-	~	-			-	
182	A Y	R		I	T	K	A	V CTD A	C	T	T	M	D	E	G	I ama	D	W	I	
544	GCGIAC	AGA	JIIF	7 T 1.77	1CM	ang	GCI	GIA	160.	ACA	ACC	AIG	GAI	GAA	GGC	AIA	GAI	199	MII	
202	K D	I	М	S	W	А	М	D	F	L	Y	Q	Y	V	L	W	W	Ι	V	
604	AAAGAT	ATTA	ATGA	\GTJ	гGG	GCG.	ATG	GAC	TTC	TTA	TAT	CAG	TAT	GTT	CTT	TGG	TGG	ATT	GTC	
												_							_	
222	QR	G	G	W	G	M	I	R	Е	Y	F	G	Т	Р	Т	A	Y	V	V	
664	CAAAGG		-1-1-1	r, G.G.C	-iliAA	$\Delta 111 \div$	A 107 - 1	3.013	~ ~ ~				A A.	$\alpha \alpha \alpha$				0.00		
		GGA	3011				AIC	AGA	.GAA	TAC	1.1.1	GGA	ACA	CCG.	ACT	GCG	TAT	GTG	GTT	
242	S L	S	A	А	т	T	I.	AGA	.GAA V	TAC	Y	'GGA V	W	CCG. W	аст к	GCG	TAT	GTG	GTT	259
242 724	S L TCTCTG	STCT	A	A GCA <i>F</i>	T \CA2	I ATA	L CTT.	AGA T ACT	GAA V GTT	TAC L TTG	Y Y TAT	V GTA	W TGG	CCG. W TGG.	ACT K AAG	Q Q CAG	DAT Q CAA	GTG - TGA	GTT	259 850
242 724	S L TCTCTG	S S STCTO	A	A 3CAF	T ACAA	I ATA	L CTT.	AGA T ACT	GAA V GTT	TAC L TTG	Y Y TAT	U GTA	W TGG	CCG. W TGG.	аст к аад	Q Q CAG	Q CAA	GTG - TGA	GTT TCT	259 850
242 724	<mark>S L</mark> TCTCTG TAAACT	STCTO	A GCCG FCAT	A GCAF TCC	T ACAA CTG:	I ATA FTT	L CTT. TTA	AGA T ACT CCA	GAA V GTT AGA	TAC L TTG TCA	Y Y TAT CTG	GGA V GTA	W TGG TAG	CCG. W TGG. ATA	ACT K AAG TGA	Q Q CAG AAT	TAT Q CAA TTA	GTG - TGA CCC.	GTT TCT AGT	259 850
242 724	<mark>S L</mark> TCTCTG TAAACT GCCTTT	STCTO 'TTAT	A GCCG FCAT AACG	A GCAF TCC GC1	T ACAA CTG: FACA	I ATA ITT ATC	L CTT. TTA TGG	AGA T ACT CCA TTT	GAA V GTT AGA TAT	TAC L TTG TCA TGT	Y TAT CTG AGG	V GTA TCT	W TGG TAG. TTC.	CCG. W TGG. ATA' AGA'	ACT K AAG TGA TTT	Q Q CAG AAT CTG	Q CAA TTA TTG	GTG TGA CCC GAC	TCT AGT ATA	259 850
242 724	S L TCTCTG TAAACT GCCTTT CTCTAG	STCTO TTTAT	A GCCG FCAT AACG AGTA	A GCAF ITCC GCI IATC	T ACAA CTG: FACA JGGG	I ATA ITT ATC CAA	L CTT. TTA TGG TCT.	AGA T ACT CCA TTT ATT	GAA V GTT AGA TAT TTT	L TTG TCA TGT GCT	Y TAT CTG AGG CTG	V GTA TCT AAG	W TGG TAG. TTC. ACT	CCG. W TGG. ATA' AGA' GGT'	K K AAG TGA TTT TCC	Q Q CAG AAT CTG AAA	Q CAA TTA TTG ATG	GTG TGA CCC. GAC. CAA	TCT AGT ATA CAC	259 850
242 724	S L TCTCTG TAAACT GCCTTT CTCTAG ACAAAT CTATAA	SGGA( STCTO TTTAT TTAT SATC TTC	A GCCG FCAT AACG AGTA AATT	A GCAF ITCC GCT VATC	T ACA2 DTG: PAC2 DGG0 AGC0	I ATA ITT ATC CAA GAC	L CTT. TTA TGG TCT. TAA	AGA T ACT CCA TTT ATT CCT	GAA V GTT AGA TAT TTT AAT GGC	L TTG TCA TGT GCT TCT AAG	Y TAT CTG AGG CTG GTT AGT	GGA V GTA TCT AAG CAG TAT	W TGG TAG. TTC. ACT TAA TGA	CCG. W TGG. ATA' AGA' GGT' CTA' A A T	ACT K AAG TGA TTT TCC TCA	Q CAG AAT CTG AAA GTG	Q CAA TTA TTG ATG ATT AGA	GTG TGA CCC. GAC. CAA TGA	TCT AGT ATA CAC CCA	259 850
242 724	S L TCTCTG TAAACT GCCTTT CTCTAG ACAAAT CTATAA CCGGTC	SGGA( STCTO STCTO STCTO SATCA SATCA SATCA STTCA STCCA STCCA	A GCCG ICAI AACG AGTA AATI FAAA	A GCAA ITCC GGCI AATC AATC AACC JCGI	T ACAZ DTG PACZ DGG0 AGC0 DTT0 PGG0	I ATA TTT ATC CAA GAC CAC STT	L CTT. TTA TGG TCT. TAA TTA. GCC	AGA T ACT CCA TTT ATT CCT. ACT GAC	GAA GTT AGA TAT TTT AAT GGC. CAC	TAC L TTG TCA TGT GCT TCT AAG TCC	Y TAT CTG AGG CTG GTT AGT TCC	V GTA GTA CAG TAT ATA	W TGG TAG. TTC. ACT TAA TGA. TTA	CCG. W TGG. ATA' AGA' GGT' CTA' AAT' CAT'	ACT K AAG TGA TTT TCC TCA CCA TCA	Q CAG AAT CTG AAA GTG CGC TTG	Q CAA TTA TTG ATG ATT AGA TGA	GTG TGA CCC GAC CAA TGA CCG AAT	TCT_ AGT ATA CAC CCA TCA CGA	259 850
242 724	S L TCTCTG TAAACT GCCTTT CTCTAG ACAAAT CTATAA CCGGTC TCTAGA	SGGA( STCT( STCT( STCT) SATCA SATCA STTCA STTCA STGCA STGCA	A GCCG ICAI AACG AGTA AATI FAAA AAAC GATG	A GCAJ ITCC GCT AATC IAAJ ACC JCGT JAAT	T ACA2 PACA2 PAC2 CGGC AGCC 2TTC PGGC PATC	I ATA ITT ATC CAA GAC CAC GTT CAA	L CTT. TTA TGG TCT. TAA TTA GCC TGT	AGA T ACT CCA TTT ATT CCT. ACT GAC TAT	GAA V GTT AGA TAT TTT GGC. GCA	TAC L TTG TCA TGT GCT TCT AAG TCC TTG	TTT TAT CTG AGG CTG GTT AGT TCC AAA	V GTA TCT AAG CAG TAT ATA ATA	W TGG TAG. TTC. ACT TAA TGA. TGA. TGT	CCG. W TGG. ATA' AGA' GGT' CTA' CAT' CAT' CCA.	ACT K AAG TGA TTT TCC TCC CCA TCA ACA	Q CAG AAT CTG AAA GTG CGC TTG ATT	TAT Q CAA TTA TTG ATG AGA TGA GCA	GTG TGA CCC GAC CAA TGA CCG AAT GAG	TCT AGT ATA CAC CCA TCA CGA CCA	259 850
242 724	S L TCTCTG GCCTTT CTCTAG ACAAAT CTATAA CCGGTC TCTAGA CTGAAA	STCTO STCTO STCTO STCTO STTATA SATCA SATCA STCCA STCCA STCCA STCCA STCCA STCCA STCCA STCCA STCCA STCCA STCCA STCTO	A GCCG ICAI AACG AGTA AATI IAAA AAAC GATG GTTI	A GCA/ GCCJ GCCJ AATC IAAA AACC ICAC ICAC	T ACAA PACAA PACA CGGC AGCC CTTC CGGC PATC 3GA	I ATA ATA CAA CAA CAC CAC CAC CAC	L CTT. TTA TGG TCT. TAA TTA GCC TGT CTT	AGA T ACT CCA TTT ATT CCT ACT GAC TAT GAA	GAA V GTT AGA TAT TTT AAT GGC. CAC GCA AAT.	TAC L TTG TCA TGT GCT TCT AAG TCC TTG AAA	Y TAT CTG AGG CTG GTT AGT TCC AAA ATG	GGA GTA GTA GTAG CAG TAT ATA ATA GAA	W TGG TAG TTC. ACT TAA TGA TGA TGT CGA	CCG. TGG. TGG. ATA' AGA' GGT' CTA' CTA' CAT' CCA. TGA'	ACT K AAG TGA TTT TCC TCA TCA ACA TTC	Q CAG AAT CTG AAA GTG CGC TTG ATT ACT	TAT Q CAA TTA TTG ATG ATG AGA GCA ATC	GTG TGA CCC. GAC CAA TGA CCG AAT GAG AGT	TCT AGT ATA CAC CCA TCA CGA CCA TTT	259 850
242 724	S L TCTCTG GCCTTT CTCTAG ACAAAT CTATAA CCGGTC TCTAGA GGCATA	SGGA( STCT( STCT( STCT) SATCA SATCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STTCA STCTA	A GCCG ICAI AACG AGTA AATI FAAA AATI FAAA GTTI ACAA	A GCA/ GCT GCT GCT AATO IAA/ CCGT GAAT ICAO	T ACAZ PACZ CGGC AGCC CTTC PGGC SGAT STGZ	I ATA ATA ATC CAA GAC CAA GTT CAA TTT	L CTT TTA TGG TCT TAA TTA GCC TGT CTT TGC	AGA T ACT CCA TTT ATT GAC TAT GAA CAT	GAA GTT AGA TAT TTT AAT GGC. GCA AAT. AGC.	L TTG TTG TCA TGT GCT TCT AAG TCC TTG AAA	Y TAT CTG AGG CTG GTT AGT TCC AAA ATG CTT	V GGTA TCTT AAG CAG CAG CAG TAT ATA ATA TTT GAA TCA	W TGG TAG. TTC. ACT TAA TGA. TGA. TGT CGA TTT	CCG. W TGG. ATA AGA' CTA CCA. CCA. TGA GGA.	ACT K AAAG TGA TTT TCC TCA CCA TCC ACA TTC AAAA	Q CAG AAT CTG AAA GTG CGC TTG ATT ACT TTA	TAT Q CAA TTA TTG ATG ATG AGA TGA GCA ATC TCA	GTG TGA CCCC GAC CAA TGA CCG GAG AGT GGA	TCT AGT ATA CAC CCA TCA CGA CCA TTT TTT	259 850
242 724	S L TCTCTG GCCTTT CTCTAGA CCGATA CCGGTC TCTAGA CTGAAA GGCATA TGTGAG	SGGA( STCT( STTAT) SATC) 'TTC) 'TTC) 'TTC) 'TGC) 'TGCO 'TGCO STGGO	A GCCG FCAT AACG AGTA AATT FAAA AAAC GATG GTTT ACAA GGTG	A GCA/ GCCJ GCCJ VATC ICAA/ VACC CCGJ GAAJ ICAC VACC JTA/	T ACA/ DTGS DGGG DGGC DTTC FGGC FATC SGAS STG/ ATCS	I ATA' TTT' ATC' CAA' GAC' CAC' CAA' CAA' TTT' ACA' ICC'	L CTT TTA TGG TCT. TAA TTA GCC TGT CTT TGC GTC.	AGA T ACT CCA TTT ATT CCT ACT GAC TAT GAA CAT AAT	GAA GTT AGA TAT TTT AAT GGC. CAC GCA AAT. AGC. TGG	L TTG TTG TCA TGT GCT TCT AAG TCC TTG AAA AAA TTG	Y TAT CTG AGG CTG GTT TCC AAA ATG CTT ATG	V GTA TCT AAG CAG TAT ATA ATA ATA CATA GAA	W TGG TAG. TTC TAA TGA TGA TGA TGA TTA TGT CGA TTT AAA	CCG. W TGG. ATA' AGA' CTA' CCA. TGA TGA GGA. GCT'	ACT K AAG TGA TTT TCC TCA ACA TTC ACA ACA GAA	Q CAG AAT CTG AAA GTG CGC TTG ATT TTA TTA	TAT Q CAA TTA TTG ATG ATG ATG AGA TGA GCA ATC TCA CGA	GTG TGA CCCC GAC CCA TGA CCG AAT GAG GAG TAC	TCT AGT ATA CAC CCA TCA CGA CCA TTT TTT TTA	259 850
242 724	S L TCTCTG GCCTTT CTCTAGA CCGGTC TCTAGA CCGGTC TCTAGA GCCATA GCCATA	SGGA( STCT( STCT) SATCI SATCI SATCI STCCI STGCI STGCO SAGTI SCCT SCCT SCCT SCCT SCCT SCCT SCCT SC	A GCCC FCAT FCAT AACG AGTA AAAT SATG STTT ACAA GGTG ATCC	A GCAI ITCC GGCI IAAC ICAC ICAC ICAC ICAC ICAC	T ACA/ CTG? FAC/ CGGC AGCC CTTC FGGC FATC SGA? STG/ SGT?	I ATA ATA TTT CAA GAC CAA GAC CAA CAA FTT- ACA FTC ITG ITG	L CTT TTA TGG TCT TAA TTA GCC TGT CTT TGC GTC.	AGA T ACT CCA TTT ATT GAC GAC CAT AAT AAT	GAA V GTT AGA TAT TTT GGC. CAC GCA AAT AGC. TGG ATA	L TTG TCA TGT GCT TCT AAG TCC TTG AAA AAA TTG ATT	Y TAT CTG CTG GTT AGT TCC AAA ATG CTT ATG GGT	V GTA TCT AAG CAG CAG CAG TAT ATA ATA TTT GAA TCA GAC	W TGG TAG. TTC. ACT TAA TGA TGA TGA TTT AAA CTGG	CCG. W TGG. ATA' AGA' CTA' CCA. TGA GGA. GCT' TCG COC	ACT K AAG TGA TTT TCC TCA CCA TCC AAA ACA TTC GAA	Q CAG AAT CTG AAA GTG CGC TTG ACT TTA TGT GCA	TAT Q CAA TTA TTG ATG ATG ATG ATG ACC CGA TAT	GTG TGA CCCC GAC CAA TGA CCG AAT GAG GAG TAC GGT CTC	TCT AGT ATA CAC CCA TCA CGA CCA TTT TTT TTA GTC CTA	259 850
242 724	S L TCTCTG GCCTTT CTCTAG ACAAAT CTGTAAA CCGGTC TCTAGA CTGAAA GGCATA TGTGAG GTTTCC TTTTCC	STTCI STTTAT TTTAT/ SATC/ TTTC/ TTC/ TTC/ TTC/ TTC/ SATC/ SAGT/ SAGT/ TTC/	A GCCC FCAT AACC AGTA AACT FAAA AAAC GTTT ACAA GGTC ACCAC AAAT	A GCA/ GCCJ GGCJ AATC ICAC CCGJ GAAJ ICAC JCAC JCAC JCAC JCAC JCAC	T ACAJ CTG <sup>C</sup> FACJ CGGC CTTC FGGC CTTC SGA <sup>C</sup> STGJ STGJ STGJ SGC <sup>C</sup> CAJ SCAJ	I ATA ATA ATC CAA GAC CAA GAC CAA CAA FTT ACA FTT ACA FTG ATG AATG	L CTT TTA TGG TCT TAA TTA GCC TGT GTC. GTC. GAT AAT	AGA T ACT CCA TTT ATT GAC TATT GAA CAT AAT AAA TAT GAT	GAA V GTT AGA TAT TTT AAT GGC CAC GCA AAT AGC TGG ATA GCT CCC	TAC L TTG TCA TGT GCT TCT AAG TCC TTG AAA TTG AAA TTG ATT CCA	TTT Y TAT CTG AGG GTT AGT TCC AAA ATG GGT GTT ATG	V GTA TCT AAG CCAG CCAG TAT ATA ATA TTT GAA TCA GAC GTC AGT TCG	W TGG TAG TTC. ACT TGA TGA TGA TGT CGA TTT CGA CTG CAT TCT	CCG. W TGG. ATA AGA GGT CTA CCA. TGA GGA. GGA. GGA. CCA	ACT K AAAG TGA TTT TCC CCA TCCA CCA TTC AACA TTC AAA GAA TTT TTA	Q CAG AAT CTG AAA GTG CGC TTG ACT TTA TGT GCA TAC ACA	TAT Q CAA TTA TTG ATG ATG GCA ATC CGA TCA CGA TAT AAG TGC	GTG TGA CCCC GAC CAA TGA CCG AAT GAG AGT GGT GGT GAT	TCT AGT ATA CCCA CCA CCA CCA TTT TTT TTT GTC CTA CTC	259 850
242 724	S L TCTCTG GCCTTT CTCTAG ACAAAT CTATAA CCGGAC TCTAGA GGCATA TGTGAG GTTTTC TTTTCC CTTTTT	S TTTA: TTTA: TTTA: SATC: TTTC: TTC: TTC: TTC: TTC: TTC: TTC:	A GCCC FCAT AACG AGTA AAATT FAAA GTT GAAC GTT ACAA ACCAC ACAC ACAC	A GCA/ FTCC GGCT AATC FAAA CCGT GAAT FCAC AAGC STAA CCC CCA/ CCA/ GAC GTAC	T ACAA DTG TACA DGGC DTG TGGC TTG TGG TGA TGA TGC TGA TGC TGA TGA TGA TGA TGA TGA TGA TGA TGA TGA	I ATA ATA CAA GAC CAA CAC CAC CAC CAC CAC CAC CA	L CTT TTA TGG TCT. TAA TTA GCC TGT GTC. GTC. GTC. GAT AAT	AGA T ACT CCA TTT ATT CCT ACT GAC CAT AAT AAA TAT GAT TCC	GAA V GTT AGA TAT TTT AAT GGC. CAC GCA AAT. AGC. TGG ATA GCT CCC AAT	TAC L TTG TCA TGT GCT TCT TCC TTG AAA ATTG AAA TTG CCA TGA	TTT Y TAT CTG AGG CTG GTT AGT CTT ATG GTT ATG GTC	V GTA TCT AAAG CAGG TAT ATA ATA ATA CATA GAC GACC GTC AGT TCG	W TGG TAG TTC. ACT TAA TGA TGA TGT CGA TTT CCAT TCT	CCG. W TTGG. ATA AGA GGT CTA TGA GGA. GGA. GGA GGA GGA GGA GTA	ACT K AAAG TGA TTT TCC TCA CCA TTC CCA TTC AAAA GAA TTT TTA TTG TTT	Q CAG AAT CTG AAA GTG CGC TTG ACT TTA TGT GCA TAC TAC ACA	TAT Q CAA TTA TTG ATG ATG ATG ACC ATC CGA TCA CGA TAT AAG GGC	GTG TGA CCCC GAC CCAA TGA CCGG AAT GGAG GGA GGA GGAT TCT	TCT AGT ATA CAC CCA TCA CCA TTT TTT GTC CTA CTC ATA	259 850
242 724	S L TCTCTG GCCTTT GCCTTT CTCTAG ACAAAT CTGTAGA GGCATA TGTGAG GTTTTC CTTTTT GATCAA	S TTTA: TTTA: TTTA: TTTA: TTTC: TTC: TTC:	A GCCC ICA1 AACG AGTA AAT1 IAAA SATG SATG ACAC ACAC ACAC ACAC ACAC ACAC	A GCAJ ITTCC GGCJ LATC ICAC CCGJ JAAC CCC JCAJ ICAC JCAC JCAC JCAC JTAC ICAC	T ACAA CTG TACA CGGG CTTC TGGG TATC SGAT SGAT SGAT SCAA SAC SAC SAC	I ATA ATA TTT ATC CAA GAC CAA GAC CAA TTT ACA ICC ICC ITTG ATG ATG ATG AAT. ITG GAA	L CTT TTA TGG TCT. TAA TTA GCC TGT CTT TGC GTC. GAT AAT TGA	AGA T ACT CCA TTT ACT GAC CAT GAC CAT GAA CAT AAA TAT GAT TCC AGC	GAA V GTT AGA TAT TTT AAT GGC CAC GCA AAT AGC TGG ATA CCC CAC	TAC L TTG TCA TGT GCT TCT TCT TTG AAA ATTG AAA TTG TGA TGG ATG	TTT Y TAT CTG AGG CTG GTT AGT TCC AAA ATG GTT ATG GTT ATG GTC ATT	V GTA TCT AAAG CAAG TATA ATA ATA TTT GAA GAC GAC C GAC TCG TCG	W TGG TAG TTC. TAA TGA TGA TGT CGA TTT AAA CTG CAT TCT TTT TCA	CCG. W TGG. ATA' AGA' GGT CTA' CCA. TGA' GGA. GGA. GGA. GCTA' GAG' CTA' GTA' TAT'	ACT K AAAG TGA TTT TCCA TCA CCA TTCA CCA TTCA GAA TTTC GAA TTTG TTTG	Q CAG CAG CAG CTG CTG CTG CGC TTG ACT TTA TGT GCA TAC ACA TTT TCT	TAT Q CAA TTA TTG ATG ATG ATG ATG ACC ATCA CGA TCA CGA TAT TAG GGC GGT	GTG TGA CCCC GAC CCA TGA CCG GAC GAG GAG TAC GGG GGT GGT GCG	TCT AGT ATA CAC CCA TCA CCA TTT TTA GTC CCA CTA CTA CTA CTA CTA	259 850
242 724	S L TCTCTG GCCTTT CTCTAG ACAAAT CTGTAGA CCGGTC TCTAGA GGCATA TGTGAG GTTTTC CTTTTTC CTTTTTTCC	S GTCT( TTAT, TTAT, AATC/ TTTC/ TTC/ TTC/ GCAA TGGC/ GAGTA TTC/ GGTAC GACAC	A GCCC ICA1 AACG AGTA AAATI IAAA AAAC GTT SGTC VCAC VCAC VCAC VCAC VCAC VCAC STT STG I	A GCAI FTCC GCT AATC FAAA CCCGT GAAT CCAC STAA CCCC GTAA CCCC FTAC TTZ TCT	T ACAA CTGS FACA CGGC CTTC FGGC CTTC FGGC FATC SGAS SGAS SGAS SGAS SGAS SGAS SGAS SGA	I ATA ATA ATC CAA GAC CAA GAC CAA CAA CAA TC AAA TC AAA TC AAA TG GAA TG GAA	L CTT TTA TGG TCT. TAA TTA GCC TGT CTT TGC GAT GAT TGA TGG	AGA T ACT CCA TTT ACT GAC CAT GAA CAT AAT AAA TAT GAT TCC AGC GAG	GAA V GTT AGA TAT TTT AAT GGC CAC GCA AAT CCC AAT CCC AAT CTA AAA	L TTG TTG TGT GCT TCT AAG TCC TTG AAA TTG AAA TTG AAT TGG ATG ATG AAG	TTT Y TAT CTG AGG CTG GTT TCC AAA ATG GTT ATG GTT ATG GTC CTG	V GTA TCT AAAG CAG TAT ATA ATA TTT GAA TCA GAC GAC TCG TCG TAC CGT	W TGG TAG TTC. ACT TAA TGA TGA CGA TTT TCT TCT TCT TCA CCTC	CCG. W TGG. ATA ATA AGA CTA CCA. CCA. TGA GGA. GCT CCTA GGA. GCTA TCC GTA TCC TCC CTA	ACT K AAAG TGA TTT TCC CCA TCC ACA TTC ACA TTT TTT TT	Q CAG CAG CTG CTG CTG CGC TTG ACT TTG TTA CCA TACA TTT TCT GTG	TAT Q CAA TTA TTG ATG ATG ATG GCA CGA TCA CGA TAT CAA GGC GGT ATT	GTG TGA CCCC GAC CAA TGA CCGG AAT GGA GGA TAC GGT GGA TCT GCG ATT	TCT AGT ATA CAC CCA TCA CCA TTT TTT GTC CTA GTC CTA CTC ATA TGG TGA	259
242 724	S L TCTCTG GCCTTT CTCTAG ACAAAT CTGTAGA CTGTAGA GGCATA TGTGAG GTTTTC CTTTTC CTTTTTT GATCAA	S GTCT( TTAT) GTTCT( TTAT) GTTC/ TTC/ TTC/ GCA GCA TTC/ GGTA GAC( GAC( GAC( C) GTA( C) C) C) C) C) C) C) C) C) C) C) C) C)	A GCCC ICA1 AACG AGTA AACG AGTA AAATI ICAA GTT1 ACAA GTT1 ACAA ACCAA ACAA CAAC ACAA CAAC CAAC	A FTCC GCD LATC FAAA LACC CCGJ GAA CCCC STAA STAA CCCC STAA CCCC STAA CCCC STAA CCCC STAA STAA CCCC STAA	T ACAA CTG <sup>12</sup> FAC2 CGGC CTTC FGGC FATC 3GA1 3GA1 3GA1 3GA1 3GA1 3GA1 3GA1 3GA1	I ATA ATA ATC CAA GAC CAA GAC CAA CAA FTC ATA FTG ATG ATG ATG ATG ATG ATG ATG ATG ATG A	L CTT TTA TGG TCT. TAA TTA GCC TGT TGC GTC. GAT TGA TGA TGA AAT TGA AAT	AGA T ACT CCA TTT ACT ACT GAC TAT GAC TAT GAA TAT GAA TAT CCA AGC GAG ATCC	GAA V GTT AGA TAT TTT AAT GGCC CAC GCA AAT CAC AAT CCC CAC AAT CTA AAA CAA	L TTG TTG GCT TCT AAG TCC TTG AAA ATTG AAA TTG ATT CCA TGA ATG ATG ATG	TTT Y TAT CTG AGG CTG AGT TCC AAA ATG GTT ATG GTT ATG GTT ATG ATT CTG ATA	V GTA TCT AAGG CAG TAT ATA ATA ATA GAC GTC GGC TCG GTCG CGTC CGTC	W TGG TAG. TTC. ACT TAA TGA TGA CGA TTT TCA CTC TCA	CCG. W TGG. ATA AGA GGT <sup>1</sup> CCA. TCA GGA GCTA GGAG GTA TCTG GTA TCT. GCG.	ACT K AAG TGA TTT TCC CCA TCA CCA TTC ACA TTT GCA ACT TTG CCA TTT TTG CCA	Q CAG CAG CTG CTG CGC TTG AAA TTT TTA TGT GCA TAC ACA TTT TCT GTAC GTAC	TAT Q CAA TTA TTG ATG ATG GCA TGA TGA TCA CGA TCA CGA TGC GGTT AAG TGC GGTT ATG	GTG TGA CCCC GAC CAA TGA CCGG AAT GGA GGA TAC GGA GGA TCT GCG GCG ATT CAA	TCT AGT ATA CCA CCA CCA CCA TTT TTT TTA GTC CTA CTA CTA CTA TTG TGA TGA	259 850
242 724	S L TCTCTG GCCTTT CTCTAG ACAAAT CTGTAGA CTGTAGA GCCATA TGTGAG GTTTTC CTTTTTC GATCAA ATTGCT	SGGA( STCT( TTA! SATC/ TTC/ TTC/ TTC/ TTC/ TTC/ GCAA TTC/ TACC GTAC GTAC GTAC GTAC GTAC	A GCCC FCAT AACG AGTA AAACG GTTT AAAA GTTT ACAA ACAA	A FTCC GGC5 AATC FAAA CCG7 JAAT CAC CG7 JAAC CC7 JAAC CG7 JAAC CC7 JAAC JAAC	T ACAA CTG <sup>CD</sup> CTG <sup>CD</sup> CTG <sup>CD</sup> AGCC CTTC CGGC CTTC CGGC CTTC CGGC CTTC CGGC ACCAA SCCAAA SCCAA S	I ATA ATA ITT ATC CAA GAC CAA CAA CCAC CAA ICC ITTG ATG AAT ICC ATG AAT ICCA ITG. CAA ICCA	L CTT TTA TGG TCT. TAA TTA GCC TGT TGC GTC GTC GTC GAT GAAT TGA TGA AAT TGA AAA	AGA T ACT CCA TTT ACT GACT GACT GAC GAC TAT GAA TAT TCC GAG GAG GAG GAG AAA	GAA V GTT AGA TAT TTT AAT GGC CAC GCA AAT AGG CCC CAA AAT CCC CAA AAA CAG AAT	TAC L TTG TCA TGT GCT TCT AAG AAG TTG AAA TTG AAT TGG AAT CAT CAT	TTT Y TAT CTG AGG CTG AGT TCC AAA ATG GTT ATG GTT ATG GTT ATG ATG ATG	V GTA TCT AAG CAG TAT ATA ATA ATA GAC GAC TCG GAC TCG TCG TAC CGT TCC TAC	W TGG TAG TTC. ACT TGA TGA TGA TGT CGA TTT TCA CTC TCA TTT	CCG. W TGG. ATA' AGA GGT' CCA. TGA' GGA. GCTA' TCG' GTA' TCT. GCG' TTG'	ACT K AAG TGA TTT TCC CCA TCC ACA TTC ACA TTC ACA TTT TTG TTT TGC TTT TAG	Q CAG AAT CTG AAA GTG CGC TTG ACT TTA TGT GCA ACA TTT TCT GTG GTG TAC ATA	TAT Q CAA TTA TTG ATG ATG ATT AGA TGA GGC GGT TAT GGC GGT ATT GAT ACC	GTG TGA CCCC GACC CAA TGA GAG GAG GAG TAC GGG GGT GCG GCG GCT CAA TAC	TCT AGT ATA CCA CCA CCA TTA GTC CTA CTA CTA CTA TTT TTA GTC CTA CTC ATA TGG TGA TTT AAA	259

Fig. 2. cDNA and deduced amino acid sequence of BsBax. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. In yellow, the Bcl-2 domain and, within this, BH1, BH2 and BH3 domains are underlined. Light blue: transmembrane domain.



Fig. 3. Alignments of the deduced amino acid sequence of BsAIF1 with known orthologous sequences from both vertebrate and invertebrates.

	TTATAAGTATAGATTCAACTATAGAACTCTTCAATC
1	M Y C F S R M H atgtattgttttagtcggatgcac
9	N R V G I L L K N V S K S K E Y V Q T F
25	ANTCGGGTTGGAATACTGCTAAAAAAGCAATAGAATACGTTCAAAACATTT
29	G V R G Y R F T R K D Q R P Q Q L H T H
85	GGGGTCAGAGGATACAGGTTCACTAGAAAAGACCAACGACGACGACGAGGTTCACAGGCAT
49	R S L D L P T H V G S A V C I G T G M M
145	CGGTCACTTGACCTCCCACGCATGTTGGGTCTGCTGTTGCATTGGAACGGGGATGATG
69	F F G I A G V Y Y A M E P K A A E V A E
205	TTCTTTGGAATTGCTGGAGTTTATTATGCCATGGAGCCAAAAGCTGCTGGAGTCGCTGAA
89	V P S E E S G P L L P T V P E S A Q Y L
265	GTACCGTCGGAAGAATCTGGCCCACTGTTACCAACTGTCCCGGAGTCTGCACAATATCTT
109	L I G G G T A S H S A M R A I R K R D P
325	TTGATCGGAGGGAGGGCATCTCATTCGGCATGCGGGCCATCAGGAAAAGAGATCCA
129	T À K I L I I T E E E L R P Y M R P P L
385	ACTGCAAAAATTTTAATCATCGCGAAGAAGAACTTCGACCTTACATGCGACCCCCGCTC
149	S K E L W Y S D D R E L A R Q L I F K Q
445	TCARAGGAACTCTGGTACTCGGAGGTGGCGGGGGGGGGGG
169	W N G K T R S L F F E K E G F Y S D P S
505	TGGAACGGAAAAACGAGGGGTTATTTTTTGAGAAGGAAGG
189	E L S S L E S G G V G L L T G K K V L S
565	GAATGTCGTCGCTAGAAAGCGGGGGGGCGTCGGATACTTACT
209	ISVKDNSVTLDDGSIVKFEK
625	ATATCTGTAAAAGATAATTCCGTTACTCGATGACGGGGCGATTGTGAAATTGAGAAA
229	C L I A T G G K P K S L P V F E N H S N
685	TGTCTCATCGCGACCGGCGGGAAACGAAAAGTTTGCCAGTTTTGAAAATCACAGCAAC
249	K T T L F R T A A D F E T L D S L T D G
745	AAAACAACGCTTTTTCGAACGGCGGGCCGATTTCGAGACTTTGGATTCTCACCGATGGA
269	I E S V T V I G G G F L G S E L A C A L
805	ATAGAATCTGTTACTGTGATAGGAGGAGGATCTTGGGAAGCGAACTGGCTTGCGCGCTT
289	G H K G S K S G L K V T Q I F P E A G N
865	GGCCATAAAGGCTCCAAGTCAGGCTTGAAAGTGACGCCAAATCTTTCCGGAAGCCGGCAAC
309	M G R V L P E Y L T K W T T E K V R K E
925	ATGGGCCGCGTTCTGCCGGAATATCTCACAAAGTGGACGACGGAAAAGTTCGTAAAGAG
329	G V D V I T H S V V K S V S E D N D K V
985	GGTGTGGATGTGATCACTCGTCGTGGAGGTGTGAGAAGACAATGACAAAGTG
349	Q L S L S N G E K I A T D H L V V A V G
1045	CAGCTGTCTCTCAGTAACGGAGAAAGATCGCCACCGATCATCTGGTTGTTGCCGTAGGT
369	L D I D T K L A A S A G I E V D E T Q G
1105	TTAGACATCGACACGAAACTGGCAGGTTCCGCAGGTATCGAAAGGCAAAGGC
389	G Y R V N À E L Q À R N N I W V À G D À
1165	GGGTATCGGGTTATGCCGAGTTGCAAGCGAGAACAATATCTGGGTGGCAGGAGACGCA
409	A C F Y D I K L G R R R V E H H D H A V
1225	GCGTGTTTCTACGACATCAAGCTTGGAAGGAGGAGGAGGAGGAGCACCATGACCATGCGGTG
4	129 V S G R L <mark>A G E N M T G A G K P Y W <b>H Q</b> 85 gtcagcggcagattggccggaaaatatgaccggtgccggtaaaccatactggcaccaa</mark>
4 13	Image: Second system     Image: Se
4 14	A     K     L     P     T     G     V     F     A     A     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     P     R     A     D     T     D     T     D     T     D     T
4 14	A A E A T G E G V R S E Q E A A E P A I       65     GCAGCAGAAGCCATGGCGAAGGCGTACGGTTTGAACAGGAAGCGGCAGAACTGCTATT
15	009 R V T S A N V D S E D F G K G V V F Y V   25 CGCGTGACTTCAGCCAATGTTGACTCGGAAGACTTCGGCAAGGGTGTAGTTTCTACGTG
15	29 R Q K K V V G I L L W N I F G R T G I A 85 CGACAAAAGAAAGTCGTCGGGATACTGCTGTGGAACATTTTCGGGCGCACTGGAATTGCG
16	49 <b>R K I I H D Q Q E H D D Y R E L A K L F</b> 45 AGAAAGATCATTCATGACCAACGAGAACACGACGACTACCGAGAACTCGCAAAACTGTTT
5 17	69 K I H K E D F E D E - 578 1737 AAAATTCACAAAGAAGAATTTTGAAGACGAA ATCCAAACGTTTCGATATCTCAGATGC
	CAAATATCATATUTCTACTATAATATCAGAACTTTCGTAAATAATCCTCTTAAATGCTCTT ACCAACAAGAAGCCGAACCAACGAGAGTGAGAGTGGAGTGGAGTTTCACATTTCCT GGTAGGACGTTAACAACTAACTACATCAAGTTTCCAAGTTTCCAATATAGAGTTCGAATT GTAGGTACAGTTTTTTTAAATTAGACTAGAGGAGGAGGAGATATTATAAATTAGAATCACTG AATTTTAAAAAAAA

**Fig. 4.** cDNA and deduced amino acid sequence of BsAIF1. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. Orange: pyridine nucleotide-disulphide domain (PNDD); yellow: NADH-binding PNDD; light blue: C-terminal domain of dimerisation; light purple: nuclear localisation motif. Bold: the N-terminal transmembrane domain.



Fig. 5. Alignments of the deduced amino acid sequence of BsPARP1 with known orthologous sequences from both vertebrate and invertebrates.

1	Δ	q
	-	-

	GTATTAAGATTCCTTGGGTGTAATTCAAATTGATACTTGATCGAA	506	G S K Y T A T L S L V D M N T G A N S Y	
1	M D G R S ATGGATGGGCGTAGT	1516	GGCTCGAAGTACACCGCAACCCTATCCTTGGTGGACATGAATACGGGAGCAAATTCCTA	.T
6	N D Y D Y V A E Y A K S N K A M C R D C	526	Y K L Q L L E H D S K K K Y Y V F R A W	3
16	AATGACTATGACTATGTCGCAGAATATGCGAAGAGTAACAAGGCTATGTGCCGTGATTGC	1576	TATAAGTTGCAGCTTCTGGAGCATGACTCGAAGAAGAATATTATGTGTTCCGCGCCTG	
26	K A L I A K E S L R I A Y M V Q S R F F	546	G R V G T T I G G D K V E D F Y S V D A	3
76	AAAGCCTTGATAGCCAAGGAAAGTCTGAGAATTGCTTACATGGTACAGGCACGATTTTT	1636	GGTCGAGTTGGGACGACGATAGGGGGGGATAAAGTCGAGGACTTCTATTCCGTTGATGC	
46	D G K Q P P W Y H Y T C F F R K Y R P S	566	A V R N F E E V Y L D K T G N E F G T A	Г
136	GATGGCAAGCAACCACCGTGGTACCATTACACTTGCTTCTCCGAAAATATCGACCTTCC	1696	GCGGTGAGGAATTTTGAAGAAGTCTATTTAGATAAAACTGGCAACGAATTTGGCACCGC	
66	C S D D I K G F N N L R F P D Q E K I R	586	H F M K M P K K F Y P L D I D Y G E N D	г
196	TGTTCCGATGACATAAAGGGCTTCAACAATCTCCGGATTTCCAGACCAAGAAAAGATTCGA	1756	CATTCATGAAGATGCCCAAAAAGTTTTATCCCCTTGACAATCGATACGGTGAAAACGA	
86	T Q I D G F A N S S K S A K G G T K I I	606	E K L K V L Q S A G S N <mark>S K L H Q E V Q</mark>	A
256	ACGCAAATCGATGGCTTTGCAAATAGTTCAAAATCTGCAAAAGGTGGAACAAAGATCATA	1816	GAGAAGCTGAAGGTCTTGCAATCTGCCGGTAGGAATTCGAAGGTCACAAGAAGTGCA	
106	S T T L R D F V V <mark>Q Y A A S N R S K C R</mark>	626	K I I K M I F D V E S M K K A M V E F E	3
316	TCGACGACGCTGCGAGACTTTGTAGTTCAGTATGCGGCGCTCTAACAGGAGCAAGTGCCGC	1876	AAGATCATCAAGATGATATTTGACGTGGAAAGCATGAAGAAGCCGATGGTGGAATTCGA	
126	Q C E D K I E K D E I R L S H K E K H P	646	L D M K K M P L G K L S K K O L Q R A Y	0
376	CAATGTGAGGATAAAATCGAGAAGGACGAAATCCGGCTCTCGCACAAAGAGAAGCATCCC	1936	CTGGATATGAAAAAGATGCCTCTCCGGTAAGTTAAGTAAG	
146 436	E K P Q L G L V D R W H H V G C F L K H GAAAAGCCGCAGCTCGGCCTGTTGATCGTTGGCATCATGTCGGATGTTTCTTGAAACAC	666 1996	G V L S E V T G I I E C E T P S A P K I Geggtacteteccgaagtgaccgcggttattgatggaaacgccaagcgctcccaaat	A
166	K A A R G W E D H F S A D M L S G F Q G	686	L D C S N R F Y T L I P H D F G M K Q P	3
496	AAGGCCGCCCGTGGCTGGGAAGATCACTTCAGGCGCGACATGCTCTCAGGGTTCCAAGGG	2056	CTGGATTGCTCAAACAGATTTTACACTTTGATCCCGCATGACTCGGAATGAAACAGCC	
186	L D A D D K G Q M R K L L <u>K P K</u> G K A K	706	PMLNNIELVKSKVEMIDSLL	3
556	CTGGACGCGGACGACAAAGGGCAGATGAGGAAACTACTGAAGCCTAAAGGAAAGGCAAAA	2116	CCGATGCTGAACAACATTGAGTTGGTCAAGTCCAAGGTCGAGATGATCGACAGCCTGCT	
206	K I A E Q P S T S E I R T E E D P I M K	726	D I E V A Y N L L Q G D Y A D V K E E D	0
616	AAGATTGCGGAACAGCCTTCAACTTCTGAAATTCGAACGGAAGAGGATCCGGATCATGAAA	2176	GATATTGAGGTCGCATACAATCTGCTTCAAGGCGACTATGCCGACGGAAGGAGAAGA	
226	K I K S Q S T L M W R N I D R L K S I E	746	PIDTQYKKKLQCCRIEPLNAND	0
676	AAAATCAAATCGCAGTCGACCTGATGTGGAGGAACATCGACAGACTCAAAAGTATCGAA	2236	CCCATCGACACTCAGTACAAGAAACTGCAATGCCGAATAGAGCCACTCAATGCCAACGA	
246	L R K S T F E D I L E E N H Q K I P I G	766	E M R K V I S D M V Q N T H A K T H S S	3
736	CTGCGTARAAGTACCTTCGAAGACATACTGGAAGAAATCATCAGAAAATACCCATCGGC	2296	GAGATGCGTAAGGTGATCAGCGACATGGTGCAAAACACTCACGCCAAGACCCATTCCTC	
266	Y D A L Y L A I G D G M T F G A L P W C	786	Y N L E V E E I F E L D R K G E K S R Y	0
796	TACGACGCATTGTATCTCGCCATCGGACGCATGACTTTCGGGGCTTTGCCCTGGTGT	2356	TACAATCTCGAAGTCGAAGAAAAAGCCCCTA	
286	P E C K A G P L V V C S D G Y K C T G Q	806	R P F K K L H N R R L L W H G S R V T N	D
856	CCGGAATGCAAAGCGGGACCGCTGGTTGTCTGCAGCGACGGATAAAATGCACCGGACAG	2416	CGGCCGTTCAAGAAGCTGCAAACCGTCGTCTGCTGGCATGGCTGCCGTGGACAA	
306	I S E W T R C T Y R T A V V E R P T W S	826	Y A G I L S Q G L R I A P P E A P V T G	A
916	ATATCGGAATGGACCCGGTGCACGTATAGGACGGCCGTCGTCGAGAGGCCCACTGGTCG	2476	TATGCCGGCATTTTGTCACAGGGCTTGAGGATTGCGCCCCGAAGCGCCTTGACGGG	
326	F P K S V R E S D P F L E K Y K Y K K M	846	Y M F G K G L Y F A D M V S K S A N Y C	0
976	TTCCCGAAATCCGTCCGAGAATCCGACCCGTTTCTTGAGAAAATAAGTATAAAAAATG	2536	TATATGTTCGGCAAGGGGTTGTACTTTGCCGATATGGTTTCGAAAAGTGCGAATTACTG	
346	E R V F P P K A P <b>F K P L K K M K V I V</b>	866	H A S H G N P Y G L L L L A E V A L G N	
1036	GAGCGCGTTTTCCTCCGAAAGCGCCTTTCAAGCCGTTGAAAAAATGAAGGTTATCGTT	2596	CATGCGTCGCATGGGAATCCATACGGACTGTTATTGCTCGGTGAAGTAGCGCTGGGAAAT	
366	L T K T P S P K P D I K A S V E K L G G	886	M R E L K N A M D V R K L P K G K H S V	
1096	CTTACGAAAACACCCTCGCCGAAGCCTGACATCAAGGCGTCTGTCGAAAAACTCGGCGGT	2656	ATGCGCGAGTTGAAGAATGCCATGGACGTACGGAAGCTACCGAAAGGAAAGCACAGCGTC	
386	S I T T S V H L A Y F C I S T K E G V A	906	K G M G S T C P D E S S D Y T M K D G T	
1156	TOGATCACTACGTCCGTGCATCTGGGGGACTTTTGCATCAGGAGGGAG	2716	AAGGGTATGGGGTCCACATGCCCGGACGAATGAGAGGGACCACAGATGAAAGATGGCACC	
406	K M S K K M T F G S A T R C F C R F S R	926	I A H V G S G K E S G C E G T S L L Y N	
1216	ARAATGAGCAAAAAAATGACATTCGGCTCAGCTACAAGATGTTTTTGTCGTTTTCTCAAGA	2776	ATAGCCCATGTAGGGTCCGGGAAGGAGTCTGGCTGGAAGGAA	
426	Y F E R S R G R V G A T A D V A F A Q K	946	E Y I V Y D V A Q V N L R Y L V K T K F	
1276	TATTTGAAAGATCTCGAGGCCGTGTCGGAGGCCACAGCCGATGTTGCCTTTGCTCAAAA	2836	GAATACATAGTGTACGATGTGAGGCGAAGTGAACTTGAGATATTTAGTGAAGACTAAATTC	
446 1336	A Q R V R L G I G A T R I E A W G I K R GCACAGCGTGTGCGATTGGGGATCGGAGCATCGAGGATCGAGCTTGGGGTATAAAGAGA	966 2896		971 2916
466 1396	A L E E V S G N V P E K K M K M V V K A GCTCTCGAGGAGGTTAGTGGTAACGTGCCAGAGAAAAAAAA		TGATTTCATTGCCAAGTTCGCAATAGTAAGCATCTGTTGCACATCATTTCGCTTTGCT CCTTTTTGCCTCATGTTCGCCGCTGTTCTTTTTCCGTGTAAATAAA	3
486 1456	G A A V D Q Q S G L E H K T H V L V K N GGCGCTGCAGTTGATCAGCAGTCTGGGTTGGAGCACAAAACTCATGTGCTGGTCAAGAAC			

**Fig. 6.** cDNA and deduced amino acid sequence of BsPARP1. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. Light blue: N-terminal zinc-finger domains; pink: PADR1 (poly(ADP-ribose)-synthase 1) domain; bold: BRCT (BRCA1 (breast cancer susceptibility protein C-terminus) domain; light purple: WGR (tryptophane-, glycine-, arginine-rich) motif; orange; C-terminal regulatory PARP domain; yellow: C-terminal catalytic PARP domain.



Fig. 7. Alignments of the deduced amino acid sequence of BsIAP7 with known orthologous sequences from both vertebrate and invertebrates.

#### AAGGTATTCTTTGTACCTATTATATAGTGTATATATGGAGTTGGGACGAGTAGA ACGCAGCAAAATACGTGCCAAGTTCTGTATGATAGTTCGCCTTGTGAAACTGGA

1 1	L																		M ATG	I ATT
	G	S	D	S	R	R	S	Q	L	Q	A	V	N	G	S	G	S	V	S	T
	GG2	ATCG	GAC	AGC.	AGG.	AGGʻ	FCA	CAG	CTT	CAA	GCT	GTG.	AAT	GGA	AGC	GGA	AGC	GTG	TCAJ	ACA
23	3 S	G	I	P	T	S	S	Q	G	G	L	S	A	N	H	G	L	F	T	D
67	7 TC:	IGGT	ATA	CCT.	ACC.	AGT'	FCT(	CAA(	GGC	GGA	CTA'	TCA	GCA	AAC	CAT	GGT	CTC	TTC	ACG	GAC
43	3 I	F	N	C	A	N	P	T	S	P	A	Y	T	G	K	S	Q	R	L	S
127	7 ATO	CTTC	AAC	TGT	GCA	AAT(	CCA	ACT'	TCT	CCA	GCA'	TAT.	ACG	GGG	AAG	TCG	CAA	AGG	CTT'	TCA
63	3 <mark>T</mark>	F	K	N	W	N	K	H	N	I	L	S	P	D	K	L	A	D	A	G
187	7 ACC	CTTT	'AAA	AAC	TGG.	AAC	AAG	CAC	AAC.	ATA	CTG.	AGC	CCA	GAT	AAA	TTA	GCC	GAC	GCA	GGA
247	L	F	Y	L	G	E	R	D	R	C	K	C	F	Y	C	N	G	G	L	Q
	7 TTC	GTTT	TAT	TTG	GGC	GAAJ	AGA	GAC	AGA	TGT.	AAA'	TGC	TTT	TAC	TGC.	AAT	GGA	GGA	TTG	CAG
103	3 N	W	V	A	G	D	D	P	M	E	E	H	A	K	Y	F	P	Q	C	E
327	7 AA:	FTGG	GTG	GCA	GGC	GAC	GAC	CCAJ	ATG	GAG	GAG	CAT	GCC	AAA	TAT	TTC	CCA	CAA	IGC(	GAA
123	3 Y	V	L	<mark>s</mark>	R	M	G	P	D	Y	V	A	N	L	N	A	R	Y	P	N
387	7 TAC	CGTT	CTG	tca	AGA	ATG(	GGA(		GAT'	TAC	GTT	GCA	AAC	CTG	AAT	GCA	AGA	TAT	CCGJ	AAT
143	3 L	T	R	P	T	I	P	S	P	P	P	Q	M	Q	Y	D	A	R	P	I
447	7 TT2	AACG	CGA	CCG.	ACG.	ATT(	CCG'	TCA(	CCA	CCC	CCG	CAA	ATG	CAA	TAT	GAC	GCC	CGG	CCAJ	ATT
163	3 Т	L	S	P	P	S	S	A	V	H	R	P	Q	N	V	Q	S	S	P	S
507	7 АС(	GCTG	TCA	CCG	CCC.	AGCJ	AGC	GCT(	GTT	CAC	AGG	CCA	CAA	AAC	GTG	CAA	TCA	TCG	CCA'	TCA
183	3 L	s	R	S	P	A	A	A	S	Q	S	H	I	D	Q	D	L	A	E	L
567	7 СТ(	Cagc	CGG	TCG	CCA	GCA	GCT(	GCA	TCA	CAA	AGT	CAT.	ATT	GAT	CAA	GAC	TTG	GCG	GAA'	TTA
203	3 V	S	Q	A	E	Q	F	G	F	D	R	S	I	I	E	E	I	L	S	D
627	7 GT2	ATCG	CAA	GCA	GAA	CAA'	FTC	GGT	TTT	GAT	CGA'	TCA	ATT	ATT	GAG	GAA	ATA	TTA	TCG(	GAT
223	3 K	y	R	S	T	G	C	Y	H	S	S	L	T	E	L	I	D	D	I	I
687	7 AAG	gtac	AGG	TCA	ACG	GGT	IGT	FAT(	CAC.	AGC:	AGC	CTC.	ACA	GAA	CTT.	ATT	GAT	GAC	ATC:	ATT
243	3 R	K	E	S	G	I	T	F	D	P	Q	P	M	E	D	V	E	V	E	C
747	7 AG2	AAAA	.gaa	AGT	GGG.	ATT	ACC'	FTC	GAT	CCC	CAG	CCA	ATG	GAA	GAC	GTC	GAA	GTC	GAA'	IGT
263	3 Н	E	T	I	R	N	P	P	I	A	C	V	A	E	G	R	S	C	G	N
807	7 СА:	Igaa	.aca	ATA	CGG.	AAT(	CCA	CCG2	ATA	GCA'	TGC	GTG	GCA	GAG	GGT.	AGA	TCC	TGT	GGTJ	AAT
283	3 V	E	S	D	S	P	L	R	D	Q	L	Q	S	A	I	N	V	T	R	C
867	7 GT2	Agaa	.AGC	GAC	TCG	CCG	CTA	CGC(	GAT	CAG	CTG	CAA	TCA	GCA	ATA	AAC	GTA	ACC	CGT'	IGT
303	3 <u>m</u>	S	C	N	R	R	N	R	D	C	L	F	L	D	C	G	H	L	C	C
927	7 AT(	GTCA	.TGT	AAC.	AGG	CGG2	AAT(	CGT(	GAT	TGC'	TTA'	TTC	TTA	GAT	TGC	GGT	CAC	TTG	IGC	IGC
323	3 C	Y	E	C	G	K	A	K	Q	R	C	V	I	C	G	T	R	V	R	E
987	7 TG:	FTAC	GAA	TGC	GGG.	AAA(	GCAJ	AAG(	CAA	AGA'	TGC	GTG.	ATT	TGT	GGG.	ACC	CGG	GTT.	AGG(	GAA
343	V	I	K	I	F	T	T	S	F	A	L	R	. E	נ יז	: :	s	L	W	L	L N
1047	GTG	ATA <i>I</i>	AAAA	ATA:	FTT.	ACG	ACA	TCI		CGC2	ATT	AAG	GTI	ככי:	FTT(	Cac	TGT	GGC	TTC	CTAAAC
363 1107	Y TAC'	C IGC <i>I</i>	N AAT <mark>1</mark>	- FAA																
	CAT	rcg:	TG	i GT:	ATA' FCT(	TAG GAT	TTA AAA	TCI AAA	TC <i>I</i>	AAG	ATC	TTA	TTI.	TCO	GTT	ΓTΑ	СТА	TAA	TGP	AATAAA

**Fig. 8.** cDNA and deduced amino acid sequence of BsIAP7. START and STOP codon are in green and red, respectively; 5' and 3' UTR are in grey. Orange: BIR domain; blue: RING (really interesting new gene) finger domain.

## Acknowledgements

This research was supported by grants from the CARIPARO Foundation (Excellence Project 2009) and the Italian M.I.U.R (P.R.I.N. grant 20109XZEPR to LB).

#### Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi. org/10.1016/j.dib.2016.05.017.

#### **Appendix B. Transparency Document**

Transparency data associated with this article can be found in the online version at http://dx.doi. org/10.1016/j.dib.2016.05.017.

### References

- [1] N. Franchi, F. Ballin, L. Manni, F. Schiavon, G. Basso, L. Ballarin, Recurrent phagocytosis-induced apoptosis in the cyclical generation change of the compound ascidian *Botryllus schlosseri*, Dev. Comp. Immunol. 62 (2016) 8–16.
- [2] D. Campagna, F. Gasparini, N. Franchi, N. Vitulo, F. Ballin, L. Manni, G. Valle, L. Ballarin, Transcriptome dynamics in the asexual cycle of the chordate *Botryllus schlosseri*, BMC Genomics 17 (2016) 275.
- [3] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32 (2004) 1792–1797.
- [4] X. Huang, W. Miller, A time-efficient linear-space local similarity algorithm, Adv. Appl. Math 12 (1991) 337-357.