

Molecular Cloning, Characterization, and Expression Analysis of a Prolyl 4-Hydroxylase from the Marine Sponge *Chondrosia reniformis*

Marina Pozzolini¹ · Sonia Scarfi¹ · Francesca Mussino² · Sara Ferrando¹ · Lorenzo Gallus¹ · Marco Giovine¹

Received: 4 August 2014 / Accepted: 27 January 2015 / Published online: 28 April 2015
© Springer Science+Business Media New York 2015

Abstract Prolyl 4-hydroxylase (P4H) catalyzes the hydroxylation of proline residues in collagen. P4H has two functional subunits, α and β . Here, we report the cDNA cloning, characterization, and expression analysis of the α and β subunits of the P4H derived from the marine sponge *Chondrosia reniformis*. The amino acid sequence of the α subunit is 533 residues long with an M_r of 59.14 kDa, while the β subunit counts 526 residues with an M_r of 58.75 kDa. Phylogenetic analyses showed that α P4H and β P4H are more related to the mammalian sequences than to known invertebrate P4Hs. Western blot analysis of sponge lysate protein cross-linking revealed a band of 240 kDa corresponding to an $\alpha_2\beta_2$ tetramer structure. This result suggests that P4H from marine sponges shares the same quaternary structure with vertebrate homologous enzymes. Gene expression analyses showed that α P4H transcript is higher in the choanosome than in the ectosome, while the study of factors affecting its expression in sponge fragmorphs revealed that soluble silicates had no effect on the α P4H levels, whereas ascorbic acid strongly upregulated the α P4H mRNA. Finally, treatment with two different tumor necrosis factor (TNF)-alpha inhibitors determined a significant downregulation of α P4H gene expression in fragmorphs demonstrating, for the first time in Porifera, a positive

involvement of TNF in sponge matrix biosynthesis. The molecular characterization of P4H genes involved in collagen hydroxylation, including the mechanisms that regulate their expression, is a key step for future recombinant sponge collagen production and may be pivotal to understand pathological mechanisms related to extracellular matrix deposition in higher organisms.

Keywords Collagen · P4H · PDI · TNF · *Chondrosia reniformis*

Introduction

The phylum Porifera is the oldest metazoan group. The marine sponge *Chondrosia reniformis* is a common Mediterranean demosponge, which body consists of a large proportion of collagen. Due to its peculiar physical–chemical characteristics (Wilkie et al. 2006), collagen derived from this sponge has been the object of study for 40 years (Garrone et al. 1975; Heinemann et al. 2007), and many biotechnological applications of *C. reniformis* collagen extracts in drug preparations (Nicklas et al. 2009a, b) as well as in cosmetics (Swatschek et al. 2002) have been described lately. Recently, a non-fibrillar collagen type from this sponge was cloned, and some information on the regulation of its gene expression (i.e., up-regulation by soluble silicates) was also unveiled (Pozzolini et al. 2012). The molecular characterization of sponge collagens as well as of the enzymes involved in their biosynthesis would open the way to the recombinant production of these proteins on a large scale. Although there is some information regarding the molecular cloning of sponge collagen genes (Exposito and Garrone 1990 and Exposito et al. 1990; Pozzolini et al. 2012), there are to date no data regarding the enzymes involved in the post-translational modifications of

Electronic supplementary material The online version of this article (doi:10.1007/s10126-015-9630-3) contains supplementary material, which is available to authorized users.

✉ Marina Pozzolini
marina.pozzolini@unige.it

¹ Department of Territory Environment and Life Sciences, University of Genova, Via Pastore 3, 16132 Genova, Italy

² Department of Experimental Medicine, Section of Biochemistry, University of Genova, Viale Benedetto XV, 1, 16132 Genova, Italy

these molecules in Porifera, which knowledge is essential for the development of recombinant sponge collagen.

Prolyl 4-hydroxylase (P4H, EC 1.14.11.2) is a key enzyme of the collagen synthesis. It catalyzes the formation of 4-hydroxyproline collagen residues that are essential for the folding of the newly synthesized polypeptide chain into triple helical molecules (for review, see Kivirikko et al. 1989; Kivirikko and Myllyharju 1998; Myllyharju 2003). The pivotal role of this enzyme in collagen metabolism makes it a possible target for the pharmacological treatment of fibrotic diseases. Prolyl 4-hydroxylases reside in the lumen of the endoplasmic reticulum (ER), where the folding and the assembling of collagen occur. The vertebrate P4Hs are $\alpha_2\beta_2$ tetramers whose α subunits contain the catalytic active site for proline hydroxylation, whereas the β subunits are identical to protein disulfide isomerase (PDI), a thiol-disulfide oxidoreductase with numerous roles in the cells (Kersteen and Rainers 2003; Gilber 1997). In the P4H tetramers, the PDI acts by maintaining the solubility of the α subunits and retaining the enzyme in the ER lumen (Kivirikko et al. 1989). Two main catalytic α type subunits were identified in vertebrates, α (I) and α (II), both associating with the same PDI/ β subunit to form $[\alpha$ (I) $_2\beta_2]$ or $[\alpha$ (II) $_2\beta_2]$ tetramers. The P4H type I is the most widespread form, whereas the type II is confined to chondrocyte, osteoblast, and endothelial cells (Annunen et al. 1998). More recently, a third α P4H, named α (III), was cloned and described (Kukkola et al. 2003). Its expression was demonstrated in several human tissues but at significantly lower levels than α (I) and α (II) isoforms.

To date, little information is available on invertebrate P4Hs, only described at the molecular level in *Caenorhabditis elegans* (Veijola et al. 1994) and in *Drosophila melanogaster* (Annunen et al. 1999). In *C. elegans*, up to five α P4H isoforms were found, and surprisingly, insect cell expression experiments revealed that this enzyme is an $\alpha\beta$ dimer. In *D. melanogaster* genome, up to 20 P4H-related genes were found (Abrams and Andrew 2002), and the recombinant expression of a *Drosophila* P4H isoform resulted in the formation of an $\alpha_2\beta_2$ tetramer as in vertebrates (Annunen et al. 1999). Co-expression experiments in insect cells of a *C. elegans* α P4H subunit with a human PDI revealed the formation of an active enzyme, while the co-expression of an α subunit from *D. melanogaster* with a human PDI failed to form an active P4H indicating, in the latter case, a species-specific α - β interaction (Veijola et al. 1994; Annunen et al. 1999). Relatively little data are available on conditions and factors affecting P4H expression and activity in vivo. In mammals, P4H expression may be regulated by nitric oxide (Cao et al. 1997), cytokines (Hiramatsu et al. 1982; Kawaguchi et al. 1992; Muguerza et al. 2001), and cigarette smoking (Raveendrana et al. 2004). Recently, the complete structure of the human P4H I promoter was elucidated (Chen et al. 2006), and the molecular pathway leading to

tumor necrosis factor-alpha (TNF-alpha)-driven P4H down-regulation in the mammalian cardiovascular wall system was described (Zhang et al. 2008).

The present work is the first description, at the molecular level, of the P4H derived from the marine sponge *C. reniformis*, a common Mediterranean demosponge. In particular, we report the molecular cloning and characterization of the α and β (PDI) subunits of P4H from *C. reniformis*, as well as the quantitative expression analysis of the α subunit in order to elucidate the sponge transcript tissue localization and to investigate the possible regulatory effects of soluble silicates, ascorbic acid, and cytokines (i.e., TNF-alpha).

Materials and Methods

Experimental Animals

Specimens of *C. reniformis* were collected in the area of Portofino Promontory (Liguria, Italy) at depths of 10–20 m. During sampling and transport, the temperature was maintained at 14–15 °C. Short-term stabulation was performed as described in Pozzolini et al. (2014). In particular, the sponges were stored at 14 °C in 200-L aquaria containing natural sea water (NSW) collected in the same area of Portofino Promontory with a salinity of 37‰ and equipped with an aeration system.

Preparation and Incubation of *C. reniformis* Fragmorphs

Fragmorphs were prepared as previously described (Pozzolini et al. 2012). Briefly, specimens of *C. reniformis* were taken from the aquaria and immediately transferred to plastic containers filled up with filtered natural sea water (FNSW). The sponges were cut into cylindrical fragments of 70 mm Ø (fragmorphs), quickly transferred into 12-well plates, and brought back to aquaria for a 2-month storage approximately, until complete healing and attachment to the plastic plates were reached. For each set of experiments, fragmorphs obtained from a single large sponge were used.

A 12-well plate containing 2-month-old fragmorphs was then taken from the aquarium. The fragmorph samples were rinsed twice with FNSW and transferred to a 14 °C incubator and pre-incubated overnight in artificial sea water (ASW). Specimens were then incubated for 24 h in ASW with or without 120 μ M sodium meta-silicate or 120 μ M ascorbic acid. Alternatively, fragmorphs were incubated for 48 h in FNSW in the presence of two different TNF-alpha inhibitors: 2 mM pentoxifylline (PTX) (Sigma-Aldrich, Milan, Italy) (Poulakis et al. 1999) or 22 μ M SPD304 (Sigma-Aldrich) (He et al. 2005).

RNA Extraction and cDNA Synthesis

Total RNA was extracted either from a whole *C. reniformis* sponge or alternatively from the ectosome, from the choanosome, or from stimulated fragmorphs using Isol-RNA Lysyl reagent (5 Prime GmbH, Hilden, Germany).

The poly-A fraction was then isolated using FastTrack® MAG mRNA isolation kit (Life Technologies, Milan, Italy) according to the manufacturer's instructions.

The sponge cDNA was synthesized using Superscript III RT-PCR System™ (Life Technologies). In a final volume of 20 µl, 20 ng of purified mRNA was incubated in the appropriate buffer containing either 2.5 µM Odt-adapter (Table 1) or 50 ng/µl of random hexamers, 0.5 mM dNTP mix, 5 mM DTT, 40 units of ribonuclease inhibitor (RN_{ASE}OUT, Life Technologies), and 15 units of SuperScriptIII™. The reaction was carried out for 50 min at 60 °C. Finally, to remove the remaining RNA strand complementary to the cDNA, 2 units of *Escherichia coli* RNase H (Life Technologies) was added and incubated for 20 min at 37 °C.

Full-Length Cloning of P4H Alpha Subunit from *C. reniformis*

A first 3'-partial sequence of cDNA coding for a P4H from *C. reniformis* was obtained by PCR approach. The amplification reaction was performed using 1 µl of *C. reniformis* cDNA synthesized as described in the “RNA extraction and cDNA synthesis” section, with an oligonucleotide complementary to the 3'-adapter as reverse primer (Rev1) coupled to an Fw1 sense primer both reported in Table 1.

Fw1 oligonucleotide was designed in position 478–500 within the *Amphimedon queenslandica* EST clone (GW175750) corresponding to one of the regions of higher identity with mammalian P4H cDNA sequences. The specific choice of the Fw1 primer sequence was based on the corresponding region encoding for the most conserved amino acids with the lowest degeneration level.

To complete the 5' end region, the GeneRacer™ Kit (Life Technologies) was used. A first PCR reaction was performed with High Fidelity Platinum Taq Polymerase (Life Technologies), using 1 µl of cDNA synthesized as described in the “RNA extraction and cDNA synthesis” section using random hexamers and 5' GeneRacer forward primer (supplied by the kit) coupled with Rev1 primer (Table 1), as reverse primer, designed within the previously identified 3'-nucleotide sequence. The touchdown thermal profile employed was as follows: 94 °C for 2 min, followed by 5 cycles at 94 °C for 30 s, 65 for 30 s, and 68 °C for 30 s; 5 cycles at 94 °C for 30 s, 62 °C for 30 s, and 68 °C for 30 s; 25 cycles at 94 °C for 30 s, 60 °C for 30 s, and 68 °C for 30 s; and 1 cycle at 68 °C for 10 min.

Successively, a nested PCR reaction was carried out using 1 µl of a tenfold dilution of the first PCR product with 5'-Generacer Nested sense primer coupled with Rev2 reverse primer (Table 1). The thermal profile used was 94 °C for 2 min, followed by 35 cycles at 94 °C for 15 s, 65 °C for 30 s, and 68 °C for 2 min and one cycle at 68 °C for 10 min.

Finally, to confirm the sequence, the α prolyl 4-hydroxylase full-length cDNA was amplified with High Fidelity Platinum Taq Polymerase (Life Technologies), and the PCR product was cloned in the pCR 2.1 vector for sequencing.

Table 1 Oligonucleotide sequences used as primers for PCR reactions

Name	Sequence (5'–3')	Position
Fw1 (forward)	5'-GGTGGACAGTATGAACCACATTA-3'	470–500 ^a
Rev1 (reverse)	5'-CGCAAAGTGGTATGACAATGACAGTGA-3'	1,654–1,680 ^c
Rev2 (reverse)	5'-ATTCTGCCTCCAGATTGAGTAAACACC-3'	1,411–1,437 ^c
FwPHB (forward)	5'-ATTCTATGCCCCATGGTGTG-3'	1,204–1,223 ^b
RevPHB1 (reverse)	5'-CCGCCATGAAGTCAATTGGAGACT-3'	1,332–1,356 ^d
RevPHB2 (reverse)	5'-CCATCAATCTCGTTGGCTGTGCGAGT-3'	1,254–1,278 ^d
Fgapd (forward)	5'-AAGCCACCATCAAGAAGG-3'	882–899 ^e
Rgapd (reverse)	5'-CCACCAGTTTCACAAAGC-3'	1,023–1,040 ^e
FP4H (forward)	5'-AGAGGAGGAACGATTAGGAGAAC-3'	187–209 ^c
RP4H (reverse)	5'-CCGCATAATGACAGCCAAGG-3'	294–313 ^c

^a With respect to GW175750 EST sequence

^b With respect to JQ699291 sequence

^c With respect to XM_003382762 sequence

^d With respect to JQ699292 sequence

^e With respect to KM217385 sequence

Full-Length Cloning of P4H (PDI) Beta Subunit from *C. reniformis*

The same approach described in the previous section was used for the isolation of the full-length cDNA coding for *C. reniformis* protein disulfide isomerase. A first 3'-partial region of the cDNA was amplified using 1 μ l of *C. reniformis* 3'-Odt-adapted cDNA with an oligonucleotide complementary to 3'-adapter as reverse primer, coupled to FwPHB sense primer, both described in Table 1.

In particular, FwPHB primer was designed in position 1204-1223 of the *A. queenslandica* PDI cDNA (XM_003382762) within one of the regions of highest identity with the mammalian PDI sequences.

The 5' end region was obtained using the GeneRacer™ Kit (Life Technologies), as described in the previous section. For the first PCR reaction performed with High Fidelity Platinum Taq Polymerase, 1 μ l of cDNA synthesized with random hexamers was amplified with 5' GeneRacer forward primer (supplied by the kit) coupled with RevPHB1 as reverse primer (Table 1), designed within the previously identified 3'-nucleotide sequence. The thermal profile employed was 94 °C for 2 min, followed by 35 cycles at 94 °C for 30 s, 72 °C for 30 s, and 68 °C for 2 min and one cycle at 68 °C for 10 min.

Then, a nested PCR reaction was carried out using 1 μ l of a tenfold dilution of the first PCR product with 5' Generacer Nested sense primer coupled with RevPHB2 as reverse primer (Table 1). The thermal profile used was as follows: 94 °C for 2 min, followed by 35 cycles at 94 °C for 15 s and 68 °C for 2 min and one cycle at 68 °C for 10 min.

To confirm the sequence, the PDI full-length cDNA was amplified with High Fidelity Platinum Taq Polymerase (Life Technologies), and the PCR product was cloned in the pCR 2.1 vector for sequencing.

Sequence Analyses

The obtained full-length cDNAs of P4H and PDI nucleotide sequences were analyzed using the BLAST algorithm of the National Centre for Biotechnology Information (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The deduced amino acid sequence was analyzed using the simple Modular Architecture Research Tool (SMART) program (<http://smart.embl-heidelberg.de/>). The presence of signal peptides was detected using SignalP 3.0 server (<http://www.cbs.dtu.dk/services/SignalP/>), whereas the consensus sites for the N-linked or O-linked glycosylation sites were checked using NetNGlyc 1.0 Server (<http://www.cbs.dtu.dk/services/NetNGlyc/>) and NetOGlyc 3.1 Server (<http://www.cbs.dtu.dk/services/NetOGlyc/>), respectively.

Multiple alignments and the phylogenetic tree were obtained using ClustalW2 free program (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>).

α P4H Subunit Transcript Analysis

To elucidate the tissue distribution of transcripts, the amounts of α P4H mRNA were analyzed by qPCR in the two main regions of *C. reniformis* body: ectosome and choanosome. The same mRNA was also quantified in 2-month-old fragmorphs treated for 24 h with soluble silicates or ascorbic acid or for 48 h with two different TNF inhibitors (see “Preparation and incubation of *C. reniformis* fragmorphs” section). GAPDH (KM217385) was used as reference gene for sample normalization.

Each PCR reaction was performed in a volume of 20 μ l containing: 1 \times master mix iQ SYBR® Green (Bio-Rad, Milan, Italy), 0.2 μ M of each primer, and 0.8 μ l of synthesized cDNA as previously described using random hexamers. All samples were analyzed in triplicate.

The following conditions were used: initial denaturation for 3 min, followed by 45 cycles with denaturation at 95 °C for 15 s and annealing and elongation at 57.7 °C for 60 s. The fluorescence was measured at the end of each elongation step. The next step was a slow heating (1 °C/s) of the amplified product from 55 to 92 °C in order to generate a melting temperature curve.

All primers, Fgapd and Rgapd for reference gene and FP4H and RP4H for target gene (Table 1), were designed using the Beacon Designer 7.0 software (Premier Biosoft International, Palo Alto CA, USA) and obtained from TibMolBiol (Genova, Italy). Data analyses were performed using the DNA Engine Opticon® 3 Real-Time Detection System Software program (3.03 version). Furthermore, in order to detect the relative gene expression of the α subunit gene as compared to an untreated sample, the comparative threshold Ct method (Aarskog and Vedeler 2000) within the Gene Expression Analysis software for iCycler iQ Real-Time Detection System (Bio-Rad) was used (Vandescompele et al. 2002).

Sponge Protein Cross-Link and Western Blot Analysis

Fresh *C. reniformis* specimens were cut into pieces of about 1 cm³ and further divided into smaller slices (10–15 mm³) using a sterile scalpel. Slices were then transferred to a 50-ml tube adding an equal volume of ipotonic buffer (10 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.8, 25 mM KCl, 1 mM EGTA, 1 % protease inhibitor cocktail, 0.1 % Triton X-100) and ice-cold homogenized with an ULTRA TURRAX T25 basic homogenizer (IKA, Verke Staufen, Germany). The homogenate was then centrifuged at 500 \times g for 10 min at 4 °C to remove bigger fragments, and then, supernatant was further centrifuged at 15,000 \times g for 15 min at 4 °C to remove cell debris. The supernatant, containing the sponge cell lysate, was collected and protein content was quantified by Bradford assay (Bradford 1976). For

protein cross-linking experiments, 500 µg of sponge proteins were incubated in 20 mM HEPES, pH 7.5, in the presence or absence of 2 mM disuccinimidyl suberate (DSS, Pierce Biotechnology, Rockford, IL, USA) for 1 h at room temperature under gentle shaking. The reaction was then stopped by adding Tris–HCl, pH 7.5, to a final concentration of 50 mM. Equal amounts of protein lysate (30 µg) (with or without cross-link) were loaded onto 8 % SDS-PAGE gels (in non-reducing conditions), electrophoretically separated, and transferred onto nitrocellulose membranes (VWR, Milan, Italy). Membranes were blocked and incubated with an anti-P4HAI antibody or an anti-PDI antibody at 0.2 µg/ml final concentration (Santa Cruz Biotechnology, Santa Cruz, CA, USA). Western blots were developed with SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific, Rockford, IL, USA), according to the manufacturer's instructions. Band detection and densitometry were performed using the Chemi-Doc System and the relative software package (Bio-Rad).

Immunohistochemistry and Confocal Microscopy Analyses

Samples of *C. reniformis* were fixed in 4 % paraformaldehyde (Carlo ERBA Reagents srl, Milan, Italy) in 0.1 M phosphate buffer (pH 7.4) at 4 °C, Paraplast embedded (McCormick Scientific, S. Louis, MO, USA), and sectioned at 6-µm thickness. Immunofluorescence was performed using a goat anti-P4HAI polyclonal antibody (1:200 in PBS, Santa Cruz Biotechnology) and a chicken anti-goat Alexa 488 (1:600 in PBS, Life Technologies). Nuclei were counterstained, after a pretreatment with RNAase A (10 µg/ml, Sigma-Aldrich, Milan, Italy), with propidium iodide (10 µg/ml, Life Technologies) in saline phosphate buffer (pH 7.4) and RNAase A (10 µg/ml). Alternatively, nuclei were counterstained with 10 µg/ml 4',6-diamidino-2-phenylindole (DAPI).

Specimen fluorescence was observed through a Leica DMRB light and epifluorescence microscope (Leica Microsystems srl, Milan, Italy), equipped with Nomarski contrast. Images were acquired using a Leica CCD camera DFC420C. Alternatively, images were obtained using a Leica TCS SL confocal microscope (Leica Microsystems) equipped with argon/He-Ne laser sources, a HCX PL APO CS 63.0×1.40 oil, and HC PL FLUOTAR 20.0×0.5 air objectives. During image acquisition, the 488 laser was set at 40 % energy, and the emission ranges were set between 500–550 and 600–700 nm for Alexa 488-conjugated antibodies and for propidium iodide detection, respectively. The photomultiplier voltage gain was set to eliminate specimen autofluorescence. Single plane images or, alternatively, stacks of 50 sections with a Z-step of 122 nm for a total thickness of 5.9–6.0 µm were taken for each image. 3D reconstruction from CLSM optic slices was performed using ImageJ 3D

Viewer (Rasband WS, ImageJ, U.S. National Institutes of Health, Bethesda, MD, USA, <http://imagej.nih.gov/ij/>, 1997–2014.).

Compliance with Ethics Guidelines

All experiments complied with institutional, national, and international ethics guidelines concerning the use of animals in research. *C. reniformis*, used in all experiments, is not included among the endangered species.

Results

C. reniformis αP4H Subunit Cloning

C. reniformis αP4H cDNA (a.n. JQ699291) is 1,789-nucleotide long. It contains a 5' untranslated region (UTR) of 16 nucleotides, an open reading frame of 1,599 nucleotides coding for a 533-amino acid protein, a 3' UTR of 174 nucleotides including a stop codon (TAA), and a putative polyadenylation consensus signal (AATAAA) preceding the poly(A) tail (Fig. 1).

The putative translation product has an estimated M_r of 59.14 kDa. Analysis with SignalP 3.9 Server reveals the presence of a possible signal peptide (1–27) with a cleavage site between Cys₂₇ and Gly₂₈ (Fig. 1).

Identity analysis with known αP4H amino acid sequences (Fig. 2a) reveals that the amino acid residues involved in Fe²⁺ binding supposedly are His₄₂₉, Asp₄₃₁, and His₅₀₀, while the residue involved on the 2-oxoglutarate binding is most likely Lys₅₀₉. Furthermore, also in the *C. reniformis* protein, the additional His residue involved in the 2-oxoglutarate C1 carboxyl group binding and decarboxylation results located in position 517. P4H multiple alignment analyses revealed that five conserved cysteines located in positions 176, 291, 307, 503, and 527, respectively, and involved in intra-chain disulfide bonds are also present in the *C. reniformis* αP4H protein. Furthermore, in position 513, a sixth Cys residue was found, a feature that is shared with the *A. queenslandica* αP4H sequence. Finally, a potential N-glycosylation site is located on the Asn₁₅₇ residue.

The αP4H C-terminals of *C. elegans* and *D. melanogaster* are an average 20 amino acids longer than their vertebrate counterparts; surprisingly, the Porifera αP4H C-terminal size resembles that of vertebrate sequences more than the invertebrate ones. In general, the overall amino acid sequence identity between *C. reniformis* αP4H and human α (I) P4H is higher than the identity with the invertebrate isoforms (Fig. 2b). The phylogenetic tree, obtained by multiple alignment of αP4Hs from representative species (Fig. 2b), shows the derivation from an ancestral progenitor gene which seems

Fig. 1 Nucleotide sequence of full-length cDNA and accordingly the deduced amino acid sequence of *C. reniformis* α P4H (JQ699291). The overlapping cDNA corresponds to a 1,789-nucleotide mRNA, including a 16-nucleotide 5'UTR and a 174-nucleotide 3'UTR. Numbering begins at the start codon. Start and stop codons are in *bold*. The signal peptide is *double underlined*. The Asn residue potentially involved in N-glycosylation is in *light gray*. The amino acid residues involved in the Fe^{2+} or 2-oxoglutarate binding are in *gray*. The six conserved Cys residues are in the *box*. The putative polyadenylation consensus signal (AATAAA) in the 57-nucleotides 3'-UTR is *underlined*

```

tcgacatatccggggtt -16
atgcagttggtattagggatgagttctactacttctgtgtgctgtgtagttaggagtg 60
M O L V L G M S S T T L L L L C S L G V 20
gttattccatctctgtctgtggagaaatgtttacagccctttacatatggaaggactg 120
V I P S L V C G E M F T A L L H M E G L 40
gtagagttagaagaacaattagtgcaacacctaataatcctataaaggaagaggagaa 180
V E L E E Q L V Q H L K S Y I R K E E E 60
cgattaggagaacttaagaagtttttagcttcagcagagaatgctcaaagtttgccaaga 240
R L G E L K K F L A S A E N A Q S L P R 80
aatgaaccaagagagcacctctaccacccgaccaatgccttggtgctcattatgctggtat 300
N E P R E H L Y H P T N A L A V I M R Y 100
cacagtggtgggacaaaactgtcagaatgtctaccaggataactctcatgatctt 360
H S G W D Q K L S E Y V Y Q D N S H D L 120
atgtctgctgtggcttggagaagtatcgcttctcctcaaggaggactactctggtgca 420
M S A V A L E K Y R F P S K E D Y S G A 140
gtgactgctatcattcgactgcaagacgcatacaagattagtcaggaaacctgacacag 480
V T A I I R L Q D A Y K I S P R N L T Q 160
tctatgttggggagagggcagccatataagtgacaaccagaatttgccttggaaataggt 540
S I L G E R Q A I K V T T R I C F E I G 180
agagaggcctattatctagagaactattggcactaaagagtggaatggtgagtgctg 600
R E A Y Y L E N Y W H T K E W M L E C L 200
aggaaaatggatgaggagaatgatataagaatgtcagcctaccagaaatattatgacttt 660
R K M D E E N D Y E D V S L P E I Y D F 220
ctggcattttctgaatacaagttggtaacctgycgcaaagcaatgcaatatacgaaggac 720
L A F S E Y K V G N L R K A M Q Y T K D 240
atgttgcaagtgacccactcatgaaagggccttgagaaatttaaattttttttcggaa 780
I L Q S D P T H E R A L R N L N F F S E 260
caaatgaaagaagatcctgaagcatttaacagagtagttacagagcagacaaggaatgta 840
Q M K E D P E A F N R V V T E Q T R N V 280
tatcctgaaacggttaggctatgaggaactgtgtgtagagaagcaagccaattccaaaggag 900
Y P E T L G Y E E L C R E A K P I P K E 300
aaccatcacaagatggtctgcttctatttcacacataagaacaatcctcgtcttattctc 960
N H H K M V C F Y F T H K N N P R L I L 320
cgcccaatataagtggaagtagctcacttaaaaccaagaattggatcttcaaaaagttt 1020
R P I K V E V A H L K P R I W I F K K F 340
ctttctgaaacagaaatggctctgctgaggaactggtgctaccctcaaaactaaacagca 1080
L S E T E M A R L R E L A V P K L K R A 360
actgctagaaatggaaaacgggagaattgaaacgggctgattataggttagtaagagt 1140
T A R N W K T G E F E P A D Y R I S K S 380
ggttggtcagtgaaatgatgacgaatctactgatattgtacaccggataaataatagg 1200
G W L S E D D D E S T D I V H R I N N R 400
attgacgattccacaggtctctccatggcaacagcagaggacctacaagtagtgattac 1260
I D D S T G L S M A T A E D L Q V V N Y 420
ggaatagtggaattatgaaccgattatgattttgcaagaagaatgaaagtagcttc 1320
G I G G H Y E P H Y D F A R K N E D A F 440
tcaaggttgggtgggttaataggatctactctacttatttatatgagcaatgtaacg 1380
S R L G W G N R I S T L L I Y M S N V T 460
ctgggagtgccacgggttactgaatctggaggcagaattataccttacaatggtgat 1440
L G G A T V F T E S G G R I I P Y N G D 480
gcagtgtactggtggaactgaagcgttcaggtgaggggtgatgagaacaagacatgca 1500
A V Y W W N L K R S G E G D M R T R H A 500
gcctgtccagtgctagtaggaaccaaatgggtctgtaacaagtgattcagcagggctggg 1560
A C P V L V G T K W V C N K W I H E A G 520
caggaatttagaagaccatgctcgttaataaaaaccagtaatgacattttgtgacctt 1620
Q E F R R P C S L N K N Q 533
tcaaaatctttcaaatctactgtcattgtcataccactttgcatatatttatcttaaat 1680
ttgataataatattatgataactttatttaggtatgattcatgctgtaagtgct 1740
atgtgattttcatgtggtggtggtggtgtagaataaaaaaaaaa 1800

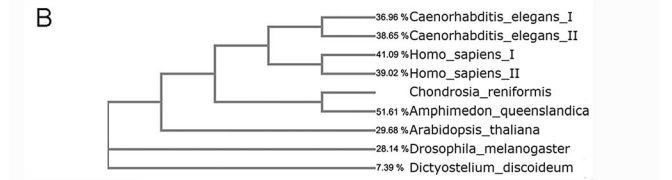
```

A

Caenorhabditis elegans_I	-----MRLALLVLTATIG-----YAVADLFTSIADNQNLLETERNIPKI	38
Caenorhabditis elegans_II	-----MRAVLVCLLAG-----LAHADLFTAIADLQMLGAEKDVDTTI	38
Homo_sapiens_I	-----MTWV-ILIIIGILL-----PQSLAHPGFPTSIGQMTDLIHTPKDLVTS	41
Homo_sapiens_II	-----MKLWSALLMAWFG-----VLSCVQAEFFTSIGHMDDLIAEKKELVOS	43
Chondrosia reniformis	MQLVLGMSSTLTLCSLGVVIFSLVCGEMFTALLHMEGLVLEELQVQS	50
Amphimedon queenslandica	---MRVPATLFSFILLGDTLLSSTRGEMFTALLHMEGLQIDERNLSSQ	46
Drosophila melanogaster	---MLLAKVFLVFLVQLS-----CYGEFFSSTGLAKLFTFTEVVLVAE	41
Arabidopsis thaliana	-----MLLAKVFLVFLVQLS-----CYGEFFSSTGLAKLFTFTEVVLVAE	41
Dictyostelium discoideum	-----MLLAKVFLVFLVQLS-----CYGEFFSSTGLAKLFTFTEVVLVAE	41
Caenorhabditis elegans_I	LDKYIHDEERLVLQKLLSEYKSKNEISIEENGLKDTNPINAFLLIKRK	88
Caenorhabditis elegans_II	IDQYIEAEARLDLRRYAHFYVHRNHAHESVGPFEVFNINAVLLIKRL	88
Homo_sapiens_I	LKDYIKAEEDKLEQIKKWA-KLDRILTSATKDPGFVHPVNAFKMLKRL	90
Homo_sapiens_II	LKEYIIVBEAKLSEKISWANKMEALTSKSAADAEYLHPVNAFKMLKRL	93
Chondrosia reniformis	LKSYIRKEERLQELKFLSAENASQSLPNEPREHLVHPNLAIVIMRY	100
Amphimedon queenslandica	LHNYTAKEREKLVLESFAKREVEGALDTMGDGIAGHLHVPVNAFLTNRF	96
Drosophila melanogaster	LQNYVNEINQHAALQSEIDAI RVEHLNADGIDDDYLNPNVAFRLIKRL	91
Arabidopsis thaliana	-----MDISNLPFHRIQQGLLQSLKQPQ-----	22
Dictyostelium discoideum	-----MDISNLPFHRIQQGLLQSLKQPQ-----	22
Caenorhabditis elegans_I	IFDW-KBIESKMNANKAGNVSSITDSSYG--VRYPTADDLSGAAILGLLR	135
Caenorhabditis elegans_II	TTW-KVVENMLNNAKSTFLNKIITDNRVSEVKFPGPEEDLSGAATALLR	137
Homo_sapiens_I	NTEW-SELENDLVLDKMSDGFISNLTIQRQY--FPNDEQVGAALKALLR	135
Homo_sapiens_II	NTDW-PALBDELVLQDSAGFIANLSVQRQF--FPTDEDEIGAALKALMR	138
Chondrosia reniformis	HSQWQKLSVEYVQDSDHLSMAVALEKRY--FSPKEDYSGAVTAIIR	146
Amphimedon queenslandica	TNGW-KMMEHYIVSDNGQDFMAMISVNRHS--FPTEDYSGAMTALLR	141
Drosophila melanogaster	HSWDTPEQSVTADSSSNLYLTMANLKEN--LSFSPQDFVGSALIAITR	139
Arabidopsis thaliana	-----QNNDESSSSNNKNNLINNE-----KVSNNLID	49
Dictyostelium discoideum	-----QNNDESSSSNNKNNLINNE-----KVSNNLID	49
Caenorhabditis elegans_I	LQDTYRLDTKDLADGKIYADQNYTFSAKDCFETARAAYNEHDFYHTVMW	185
Caenorhabditis elegans_II	LQDTYSLDLDLNSGIIIGGKVSNKLSGHDTFEVRGSAYNQDKYHCLMR	187
Homo_sapiens_I	LQDTYSLDLDLNSGIIIGGKVSNKLSGHDTFEVRGSAYNQDKYHCLMR	187
Homo_sapiens_II	LQDTYSLDLDLNSGIIIGGKVSNKLSGHDTFEVRGSAYNQDKYHCLMR	187
Chondrosia reniformis	LQDTYRLDDEGTSIRGELPGTKYQAMLSDVDCFGMGRSAYNEGDYYHTVLW	188
Amphimedon queenslandica	LQDAYSIPRNLQCS-LLGERQAKIVTTRICFEGREAYLLENVYHTKWEV	195
Drosophila melanogaster	LQDTYRLDDEGTSIRGELPGTKYQAMLSDVDCFGMGRSAYNEGDYYHTVLW	188
Arabidopsis thaliana	LQDTYRLDDEGTSIRGELPGTKYQAMLSDVDCFGMGRSAYNEGDYYHTVLW	188
Dictyostelium discoideum	LQDTYRLDDEGTSIRGELPGTKYQAMLSDVDCFGMGRSAYNEGDYYHTVLW	188
Caenorhabditis elegans_I	MEEAQRLQDDE--VEPTVEVEDILEYLAFAFYKQNNLKHAKLKTLELYK	232
Caenorhabditis elegans_II	MEQALVRIENE--NPTTEWEIILEYLAFLYKQNNLKHAKLKTLELYK	234
Homo_sapiens_I	MEQALRQDDEG--EISTIDKVSVDLYLSYAVVQGGDLKALLTKKLLK	232
Homo_sapiens_II	MEQVQLQDAG--EATTTKQVLDYLSYAVVQGGDLRHLAELTRRLLS	235
Chondrosia reniformis	MLAELRQDDEG--NDYEDVSLPEIYDFLAFSEYKVGKRNKAMQYTKDILQ	243
Amphimedon queenslandica	MLAELRQDDEG--NDYEDVSLPEIYDFLAFSEYKVGKRNKAMQYTKDILQ	243
Drosophila melanogaster	MQSMQLQDDE--SYGSESASLDFMEAVVEYHREMGHESALSVDNHLVRS	237
Arabidopsis thaliana	QLAFINRLDSE--YGTGPPSLRGLRG	49
Dictyostelium discoideum	NKESLNQLEKK-----GYLIIDNPLNDLNKINLIYD	93
Caenorhabditis elegans_I	MNPTHPRAKGNVWYEDLLEQEGVRRSDMRKNLPPIQNRRP--DSVVG	278
Caenorhabditis elegans_II	IAPNHPRAKGNVWYEDLLEQEGVRRSDMRKNLPPIQNRRP--DSVVG	274
Homo_sapiens_I	LDPESHARAGNLRYPFQGLLEBER--EKTLLTQTEAEIATPEGIYERPDY	283
Homo_sapiens_II	LDPESHARAGNLRYPFQGLLEBER--EKTLLTQTEAEIATPEGIYERPDY	283
Chondrosia reniformis	SDPHERALNRLNFPSEKQMDPEAFNRVVTBGT-----RNV	280
Amphimedon queenslandica	HDPHERAISNREYFNRSREPEPKF--VDHEGV-----LDD	274
Drosophila melanogaster	IEPDQRSHLLEARQLEELITDGDKNGLHLHTARRP-----GDHYE	278
Arabidopsis thaliana	ONTRYLRDVS-----EELITDGDKNGLHLHTARRP-----GDHYE	278
Dictyostelium discoideum	ESNYQFKEMK-----EELITDGDKNGLHLHTARRP-----GDHYE	103
Caenorhabditis elegans_I	NTERTMYEALCRNEVP-VSQKDISRLYCYKGRDR--PPLVYAPIKVEIKR	325
Caenorhabditis elegans_II	IVERDAYEALCRGEPVPEPKQWKNLRCYLKRDK--PPLKLAIPKVEIKR	322
Homo_sapiens_I	LPERQYEMLRCSGKIMTFRQKQLFCRYHGNDRNPKLAPAKQEDBW	331
Homo_sapiens_II	LPERQYVESLCSGKIMTFRQKQLFCRYHGNDRNPKLAPAKQEDBW	333
Chondrosia reniformis	YPETLGYEELCR-EAKPIKRNHMMVCFYFTHKNPNLILRPIKVEVAV	329
Amphimedon queenslandica	ESHEAVYKLCR-EPAIPKSHLKKLLCYFHNKRNPLILRPIKVEVAV	323
Drosophila melanogaster	SRFRMYEQVCRGELAPLPSVQKRN-LACRLRKR--LQVAPFKLEBLH	313
Arabidopsis thaliana	-----RWANDKDAELLRINQVNPVVS	81
Dictyostelium discoideum	-----LIEAGMNGKTDK	115
Caenorhabditis elegans_I	FNPLAVLFDKVISDDEVAAIQELAKPKLARATVHDS-VTGKLVATYTRIS	374
Caenorhabditis elegans_II	FDPLAVLFDKVISDDEVAAIQELAKPKLARATVHDS-VTGKLVATYTRIS	371
Homo_sapiens_I	DKPRIIFRFDIISDAEIEIVKDLAKPLSRATVHDP-ETGKLTAAQYRVS	380
Homo_sapiens_II	DSPHIVRYDVMDSBEIERIKELAKPKLARATVHDP-KTGVLTVASYRVS	382
Chondrosia reniformis	LKPRIWIFKFLSETEMARLELAVPKLKRATARNV-KTGFEPEPADYRIS	378
Amphimedon queenslandica	VKPKIYIFDYITDREIERLKELANPKLNARATVHG--ENGLLHATYRIS	371
Drosophila melanogaster	LDPVLVQLHGVIGSKDSDSQKTRAPRIKRSTVYSLGGNGGSGTAAAFRTS	373
Arabidopsis thaliana	WSPRIIVLHDFLSPFEECYLKAIRPRLQVSTVVDV-KTGQVKSQDVRTS	370
Dictyostelium discoideum	WK-----DKSIRGDIQWHRDSSNRITQDKLDSST-----TRININLLD	154

Caenorhabditis elegans_I	KSAWLKEWEG--DVTVENKRIQYMTNLEMETAEELQIANVIGGHYDP	421
Caenorhabditis elegans_II	KSAWLKGDDL--PVIDRNNRIEDFTNLNAGTSEELQVANYGLGGHYDP	418
Homo_sapiens_I	KSAWLSGYEN--PVSIRNNRIQDLTGLDVTAEELQVANYGVGGQYEP	427
Homo_sapiens_II	KSWLEDDDD--PVARVNNRQHIITGLTVKTAELQVANYGVGGQYEP	429
Chondrosia reniformis	KSGWLEDDDESDTIVRHNINRIDSSTGLSMTAEELQVANYGVGGQYEP	428
Amphimedon queenslandica	KSGWLESDDD--PLGVVDRIQRIEDVTGLTMTAEELQVANYGVGGQYEP	420
Drosophila melanogaster	QGASFNYSRN--AATKLSRSHVDFSGLNMDVYAEGLQVANYGVGGQYEP	420
Arabidopsis thaliana	SGMPLTHVER-SYPIIQALEKRIAVFQVPAENGLIQVLRYPQGFYK	179
Dictyostelium discoideum	KLDLIKNEFD-----NVIPFNMSIKTQY--QLAVLYNGRYIK	190
Caenorhabditis elegans_I	HFTHAKKEESKSFESLCTGNRIATVPLFMSQFSGGGTVFTEAKS----	466
Caenorhabditis elegans_II	HFDFARKEEKNAFKTLAGTGNRIATVPLFMSQFSGGGTVFTHLGT----	463
Homo_sapiens_I	HFDFARKEEDPAFKELCTGNRIATVPLFMSQFSGGGTVFTEAKS----	472
Homo_sapiens_II	HFDFSRPRPDSGKTE--GNRLATPLFMSQFSGGGTVFTEAKS----	472
Chondrosia reniformis	HYDFARKN-EDAFSLRQGNRIISTLLYMSVTLGGATVFTESGG-----	472
Amphimedon queenslandica	HYDFARTG-EDTFTSLSGNRIISTLLYMSVTLGGATVFTESGG-----	464
Drosophila melanogaster	HWDSFPEN-HIYQEGDLHGNRMATGIYLSLVEKGGGATVFFGLPL-----	464
Arabidopsis thaliana	HHYDFADT--FNLKRGQRVATMLMYLTDVDEGGYTFPLAGDGDCTC	225
Dictyostelium discoideum	HRDSFYSS-----ESLTIISRIITMIIYVNMKDWKGGGELRLYTN-----	230
Caenorhabditis elegans_I	-----TILPTKNDALFWNLYKQDGNPDRTHAACPVLVGKIVKWSNK	508
Caenorhabditis elegans_II	-----AVFPSKNDALFWNLRDGEDLRTTHAACPVLVGKIVKWSNK	505
Homo_sapiens_I	-----SVWPKKGTAVFWNLFASGEGDYSTRTHAACPVLVGKIVKWSNK	514
Homo_sapiens_II	-----AIWPKKGTAVFWNLLRSSEGDYRTRTHAACPVLVGKIVKWSNK	514
Chondrosia reniformis	-----RIIPFGNDVFWNLLKSGEGDMRTHAACPVLVGKIVKWSNK	514
Amphimedon queenslandica	-----RLVPTKRAVFWNLLKSGEGDYSTRTHAACPVLVGKIVKWSNK	506
Drosophila melanogaster	-----LVTPEPGRSLFWNLLKSGEGDQPRTHAACPVLVGKIVKWSNK	506
Arabidopsis thaliana	GGIKMGIKSVKPTKGDALFWSMGLDQSDQSIHGGCEVLEPESKWSATK	275
Dictyostelium discoideum	-----PNNTNQKELKTEEFIDIEPIADRLLIFLSPFLHE	267
Caenorhabditis elegans_I	WIHEKGFERRPGLKSSDYERFVGDIG-YGPEPRNANPNVSNLAKDVME	557
Caenorhabditis elegans_II	WIHEKGFERRPGLKSSDYERFVGDIG-YGPEPRNANPNVSNLAKDVME	557
Homo_sapiens_I	WLHERGQEFRRPGLKSSDYERFVGDIG-YGPEPRNANPNVSNLAKDVME	539
Homo_sapiens_II	WFERGQEFRRPGLKSSDYERFVGDIG-YGPEPRNANPNVSNLAKDVME	533
Chondrosia reniformis	WIHERGQEFRRPGLKSSDYERFVGDIG-YGPEPRNANPNVSNLAKDVME	533
Amphimedon queenslandica	WIHERGQEFRRPGLKSSDYERFVGDIG-YGPEPRNANPNVSNLAKDVME	525
Drosophila melanogaster	WIERNQDINVRPGLKSSDYERFVGDIG-YGPEPRNANPNVSNLAKDVME	535
Arabidopsis thaliana	WMRQKATS-----FEPRIAITTWIY-----	283
Dictyostelium discoideum	VLQCN--FEPRIAITTWIY-----	284
Caenorhabditis elegans_I	TL 559	
Caenorhabditis elegans_II		
Homo_sapiens_I		
Homo_sapiens_II		
Chondrosia reniformis		
Amphimedon queenslandica		
Drosophila melanogaster		
Arabidopsis thaliana		
Dictyostelium discoideum		

Fig. 2 a Multiple alignment obtained using ClustalW2 free program of *C. reniformis* αP4H amino acid sequence with other αP4Hs, comparing human α(I) (NP_000908), human α(II) (NP_001017973), *C. reniformis* αP4H (JQ699291), *A. queenslandica* αP4H (XM_003382475), *C. elegans* PHY-1 (AJ270999), *C. elegans* PHY-2 (JQ699291), *D. melanogaster* αP4H (AF096284), *A. thaliana* αP4H (AEC10206),



to split at least in three branches: the Protozoa (*Dictyostelium*), the Arthropoda (*Drosophila*), and a third branch with a common progenitor for Plants (*Arabidopsis*), Porifera (*Chondrosia* and *Amphimedon*), Nematoda (*Caenorhabditis*), and Chordata (*Homo*).

C. reniformis βP4H Subunit (PDI) Cloning

C. reniformis βP4H cDNA (a.n. JQ699292) is 1,814-nucleotide long. It contains a 5' UTR of 63 nucleotides, an open

and *D. discoideum* (Q86KR9). The six Cys residues are in *light gray*, and the amino acid residues involved in the Fe²⁺ or 2-oxoglutarate binding are in *gray*. **b** Phylogenetic tree obtained with ClustalW2 free program of *C. reniformis* αP4H amino acid sequence and other αP4Hs. The tree is represented as a cladogram. The identity percentages compared with *C. reniformis* αP4H are indicated

reading frame of 1,578 nucleotides coding for a 526-amino acid protein, and a 3' UTR of 157 nucleotides including a stop codon (TAA), two mRNA instability motifs (ATTTA), and a poly(A) tail (Fig. 3).

The putative translation product has an estimated *M_r* of 58.7 kDa. Analysis with SignalP 3.9 Server revealed the presence of a signal peptide (1–16) with a cleavage site between Gly₁₆ and Ala₁₇.

The domain organization reflects the typical α–β–β'–α' modular organization of mammalian βP4H (Kivirikko et al.

Fig. 3 Nucleotide sequence of full-length cDNA and accordingly the deduced amino acid sequence of *C. reniformis* β P4H (JQ699292). The cDNA corresponds to a 1,814-nucleotide mRNA, including 63-nucleotide 5'UTR and 157-nucleotide 3' UTR. Numbering begins at the start codon. The start and the stop codons are in *bold*. The signal peptide is *double underlined*. The two Cys residues present in the two typical redox active sites (TRX motif: C-X-X-C) are marked with *asterisk*. The two mRNA instability motives (ATTTA) in the 157-nucleotides 3'-UTR are *underlined*. The ER lumen retention signature KIEL is in the *box*

```

ttg -63
tacaactccaagtaagccacagctctgctagactatctcttcagagccaagtggatgacagaa -60
atggttaaaggaaatcattttctctatattggccattctaattcttggagcagacgatatc 60
M L K G I I F S I L A I L I L G A D D I 20
cctgaagaagatggagtggttagtgcgacatcagacaactttaagcaagcagtagggagt 120
P E E D G V L V S T S D N F K Q A V G S 40
atagacaacatactcgtagaattttatgcaacggtggcggtcattgcaagctctggcg 180
I D N I L V E F Y A P W C* G H C* K A L A 60
ccaaatgatgcaagcagcagaattgttgcagaaggaagatctgaaatccgctctcgc 240
P N Y A K A A E L L Q K E G S E I R L A 80
aaagtggatgcaacaatcactcggatctggcacagaagtttggagtgagaggttacccc 300
K V D A T I H S D L A Q K F G V R G Y P 100
accatcaagtctctcagaggagaagacaatgtggggactattcagctggcaggcaagca 360
T I K F F R G E D N V V D Y S A G R Q A 120
gaggacatttgaaactggtaaaagaagaagactggccccctgcccactgacctgaccaca 420
E D I V N W L K K K T G P P A T D L T T 140
gctgaaacagccaaggagttctctgagagtgacagagctgtggtggttggattctttgaa 480
A E Q A K E F S E S A E A V V V G F F E 160
tctgctgaatcagaggagccaaaggcgtacatagcagcagttaccaagcttggagacatg 540
S A E S E E A K A Y I A A V T K L E D M 180
ccggctgggaattgtcaccaataaggaggtgccaatcacttggagccactgttagcagt 600
P A G I V T N K E V A E S L G A T V S S 200
attgtaatgtttcaaaagtgtgacgatggaagggtgtttacagtggggagtagacaacggt 660
I V M F Q K F D D G K V V Y S G E Y N G 220
gtagacatactgtcctttgtgaaatgctgaaacggctcccattgttcattgaaatattgat 720
V D I L S F V N A E R L P L F I E F N D 240
aagactgctaacagatcttcagcggcagcatcaaggcacttcttcttcttattca 780
K T A N S I F S G S I K V H F L V F Y S 260
gatgactcggatgaaagcagccagcatttcagaggctgtcactgctgtagccaaagaattt 840
D D S D E A A S I S E A V T A V A K E F 280
agaggcagaataatccttggtagggtaataacagatgtggagaaatacaacagaattgtt 900
R G R I I F V R V N T D V E N N N R I V 300
gagtttttcaatcttaaggaggaagcatttcccttggtagcggctgattcagctggagcat 960
E F F N L K E D D F P C G R L I Q L D D 320
gacatgagaaaatttaggcccgactttgatcactcaatgaagaaaatattaggactttc 1020
D M R K F R P D F D S L N E E N I R T F 340
ttaaccaactttttgtaggtaaaactgaagccacatctcagctcgaagatattccccgaa 1080
L T N F F D G K L K P H L S S E D I P E 360
gatgggatgccaaaccagtcagattcttctcggacgtaactttgggatggtatcaag 1140
D W D A K P V K I L V G R N F W D V I K 380
agtgggaaagatgcttggtaggtttatgcccctgggtgtggtcactgcaaggccctg 1200
S G K D V F V E F Y A P W C* G H C* K A L 400
gctcccacttgggatgagcttggagagaagtttgcgacagcgaagcattattgcaaaa 1260
A P T W D E L G E K F V D S D V I I A K 420
atagactcgcagccaaagcagattgatggatataagattgaagggtttccaactctaaaa 1320
I D S T A N E I D G I K I E G F P T L K 440
tatttcaagatgatgatccagagtctccaattgacttcaatgggaggcaggacacttgac 1380
Y F K S D D P E S P I D F M G G R T L D 460
gacctgtcaagttttagagtcctcaaggactgatcgcgggtgcaaggagccgctccgaa 1440
D L V K F V E S K G T D R G V K E P S E 480
gacctcggagatgaattcttgacagcctgaagatgaagattttgaagatgaagattat 1500
D L G D E F F D S L E D E D F E D E D Y 500
gaagatgaagatgaagatgaaggaatggatgagcaggaggatgtaggagagcctgctaaa 1560
E D E D E D E G M D E Q E D V G E P A K 520
gataaaaagattgaacttaggtgttaattgatagaaaatatttactccttatcaaat 1620
D K K I E L 526
aatcaaataggagagcagatgtaagcttttgacaatgatcagttctgaatttaattta 1680
ctctgtatgtatggtttgctttccatcgacaataatataatgaatgtttgtcgaaaaa 1740
aaaaaaaaaa

```


1989) (Fig. 3) followed by a highly acidic region located at the C-terminal of the protein. The last four amino acid residues Lys-Ile-Glu-Leu (KIEL) are likely involved in the ER lumen retention. Multiple alignment performed with ClustalW2 of *C. reniformis* P4H β subunit with known PDIs indicates that, although the sponge sequence presents a C-terminal extension similar to the PDIs from the nematode and from the calcareous sponge *Sycon ciliatum* (Fig. 4a), nevertheless, it results strongly related to the mammalian PDI. *C. reniformis* PDI, in fact, shares the highest level of identity with its human counterpart (48.44 %, Fig. 4b).

Western Blot Analysis

C. reniformis sponge cell lysate was analyzed to detect the presence of P4H α and β (PDI) subunits by Western blot analysis by using commercial antibodies with a broad species spectrum raised against conserved regions of human α P4H and PDI proteins. Results displayed in Fig. 5a show that each antibody could detect a band corresponding to the expected molecular weights of α P4H and PDI (58 and 59 kDa); furthermore, both antibodies were able to detect a band of 120 kDa that could correspond to an α P4H/PDI heterodimer (Nietfeld et al. 1981). Finally, cross-linked sponge lysate Western blot using the α P4H antibody mainly shows two bands: one corresponds to the possible heterodimer while the other, with an estimated molecular weight of 240 kDa, indicates the presence of a tetramer. This result suggests that also in the *C. reniformis*, marine sponge the P4H enzyme is active in the $\alpha_2\beta_2$ tetramer form as in mammals and in *D. melanogaster*.

Immunohistochemistry Analyses

P4H localization in the body of *C. reniformis* was studied with standard immunofluorescence techniques by use of epifluorescence and confocal microscopy. P4H was detected both in the sponge choanosome as well as in the ectosome.

Overall, sponge P4H is more expressed in the choanosome region. In particular, in the immunohistochemistry analysis, the enzyme appears mainly surrounding the numerous choanocyte chambers (Fig. 5b (I), c (I–II) and Online resource Movie 1) with similar intensity among the different chambers. Conversely in the ectosome, intense immunoreactivity (ir) was solely detectable around the numerous sand grains usually incorporated by this sponge (Fig. 5b (II), c (III–VI) and Online resource Movie 2, 3, and 4). The nuclear staining allowed recognizing the presence of cells on the grain surfaces. Some, but not all of the nuclei, were surrounded by ir; immunoreactive cell processes were also visible. These observations suggest that, in the ectosome, the collagen synthesis takes place close to the grains. The collagen network, coordinated with the grains, could compensate for the lack of

spicules and endoskeleton. Indeed, the grains could replace the function of the megascleres in sustaining the tissues.

Tissue Localization and Gene Expression Regulation of α Subunit mRNA

The expression patterns of the α P4H gene were studied in *C. reniformis* by quantitative PCR analysis. A first analysis aimed at defining the tissue localization in the two main tissues of the sponge revealed that the mRNA level of α P4H was 13.65 ± 4.81 -fold higher ($p < 0.005$) in the choanosome than in the ectosome (Fig. 6a), somehow confirming the data observed in the immunohistochemistry analyses.

Afterwards, α P4H gene expression was studied in 2-month-old fragmorphs obtained from a single sponge. Specimens were incubated at 14 °C in the presence or absence of 120 μ M Na₂SiO₃ or 120 μ M ascorbic acid for 24 h, and the relative α P4H gene expression was then quantified by qPCR as described in “Materials and methods.” This time point was chosen according to previously published data (Pozzolini et al. 2012) in which a non-fibrillar sponge collagen, a P4H substrate, was found to be upregulated at 24 h by silicates. As shown in Fig. 6b, in silicate-treated fragmorphs, the amount of α P4H mRNA was not significantly different with respect to the untreated control sample (1.26 ± 0.337 -fold increase), whereas a fold increase of 4.96 ± 0.577 was observed in the ascorbic acid-treated samples with respect to control, indicating that the positive action of vitamin C in the extracellular matrix biosynthesis likely takes place already upstream, at the level of gene expression.

Finally, to evaluate the influence of cytokines known to be involved in α P4H gene expression (i.e., TNF- α), 2-month-old fragmorphs were incubated in FNSW for 48 h at 14 °C in the presence of two specific TNF- α inhibitors, namely 2 mM PTX and 22 μ M SPD304. This time point was chosen to ensure both an optimal penetration of the drugs into the fragmorphs and the highest inhibition since the effect one of the two inhibitors (pentoxifylline) needs time to be observed acting at the level of TNF- α protein synthesis. Figure 6c shows that the α P4H gene expression level was significantly diminished with respect to the control by use of both TNF inhibitors (0.36 ± 0.038 - and 0.64 ± 0.012 -fold, respectively). Thus, these results indicate a possible positive involvement of a still unknown TNF- α -like cytokine in the *C. reniformis* α P4H expression and sponge matrix deposition.

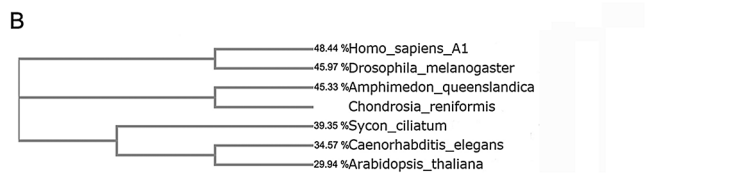
Discussion

Currently, there is an increasing interest in sponge collagen production for biotechnological purposes (Nicklas et al. 2009a, b; Swatschek et al. 2002). Although some information

Fig. 4 a Multiple alignments of *C. reniformis* βP4H amino acid sequence with other PDIs performed with ClustalW2 free program Muscle program, comparing human PDI A1 (NP_000909), *C. reniformis* βP4H (JQ699292), *A. queenslandica* PDI A2 (XP_003382810), *C. elegans* PDI A1 (NP_497746), *S. ciliatum* PDI (CCQ18635), *D. melanogaster* PDI A2 (NP_524079), and *A. thaliana* A1 (NP_173594). **b** Phylogenetic tree obtained using multiple alignment of *C. reniformis* βP4H amino acid sequence and of other PDIs. The tree is represented as a cladogram. The identity percentages compared with *C. reniformis* βP4H are indicated

A

Homo_sapiens_A1_	--MLRRALLCLAVAALVR-----ADAPEE-EDHVLVLRKSNFA 35
Drosophila_melanogaster	MKFLICALFLAASYVAAS-----AAEAVKVEBGLVATVDNFK 38
Amphimedon_queenslandica	--MK--VLAFLCAVLLG-----AVVRADBSLVLVLTQDTFH 33
Chondrosia_reniformis	--MLKGIIFSILAILLIG-----ADDIPED-GVLVSTSDNFK 35
Sycon_ciliatum	--MASSYRLLLSLVLLACFHGSLSEGBAETAPEEISDVLVLTDENFD 48
Caenorhabditis_elegans	--MSLSVSFIFLIVASIG-----AVVADSENVLVLTSDNFT 45
Arabidopsis_thaliana	MAMRGPTLFSILVLSLCS-----SIRSBETETKPVVLTLDHTNP 31
Homo_sapiens_A1_	EALAAHKYLLVEFYAPWCGHCKALAPYAKAAGKLAKEGSEIRLAKVDAT 85
Drosophila_melanogaster	QLIADNEFVLFYAPWCGHCKALAPYAKAQAQLAEKSPKIKLAKVDAT 88
Amphimedon_queenslandica	EAISSNENILVEFYAPWCGHCKALAPYAKAAMIKRGGMDFTLAKVDAT 83
Chondrosia_reniformis	QAVGSIDNLFYAPWCGHCKALAPYAKAALQLKEGSEIRLAKVDAT 85
Sycon_ciliatum	EAIKDNENILVEFYAPWCGHCKALAPYAKAALRBEBSAIRLAKLDGT 98
Caenorhabditis_elegans	ETINQNEFVLFYAPWCGHCKALAPYAKAALDLKEGSEIRLAKVDAT 84
Arabidopsis_thaliana	DTINKHDFIVVEFYAPWCGHCKQLAPYEAKAASLSSNVPVVLAKIDAS 91
Homo_sapiens_A1_	BES--DLAQYGVRYGPTIKFRNGDTASPEYTAGREADDIVNMLKRR 133
Drosophila_melanogaster	VEG--ELAEQYAVRYGPTIKFRFSG--SPVEYSGGQAAIDIAWTKKT 133
Amphimedon_queenslandica	VEK--ELAEQYAVRYGPTIKFRFSG--GVRPRYSAGKANDI IAWLEKST 128
Chondrosia_reniformis	IHS--DLAQYGVRYGPTIKFRFSG--DNVVYDSAGRAEIDIVNMLKKT 131
Sycon_ciliatum	VHK--KAVERFSLKGYPTIKFRFSG--GETEYSGGRTSKS IAWLSKRT 143
Caenorhabditis_elegans	ENQ--ALASKFVRYGPTIKFRFSG--KPTKYTGGRATQIDVDMWKKS 129
Arabidopsis_thaliana	BETNREBFAQYEVGFPTIKFRFSG--KAVQBYNGPREABEIVTYLKKQS 140
Homo_sapiens_A1_	GPAATLDPGAAEELVESSEVAVIGFFKVDSESAKQFLQAEEA-IDD 182
Drosophila_melanogaster	GPPAKDLTSVADAEQFLKDNELTIIIGFFKDLSEBAKTFKRVANA-LDS 182
Amphimedon_queenslandica	GPVVELATAAEIKAFNDKADVSIVGYFSPNETDEAKAYISAADSGIEGL 178
Chondrosia_reniformis	GPPATDLTAEQAKFSESAAVGVYFSESASEBAKAYIAAVTK-LDM 180
Sycon_ciliatum	GPPAVALPDVDAKTFVGDRSVAIIIGFFPDETTQEARGLFSAQADAPDL 193
Caenorhabditis_elegans	GPTVTVESVEQLLEELGKTRVVLVLYGPKDAKSDAATITVNEVADS-VDDA 178
Arabidopsis_thaliana	GPASABIKSADDASEVSDKVVVVGIFPKLGSSEFDSFMAIAEKLSREL 190
Homo_sapiens_A1_	PFGITNSDVFSKYQLDKDG-VLFLPKFDEGRNNFEG---EVTKENLLDF 228
Drosophila_melanogaster	VFGVSSNADVIKYEAKDNG-VLFLPKFPDDKKSVPFEG---ELNEENLKKF 228
Amphimedon_queenslandica	NFALCINPETTKEMAEVNT-VLFLYKFDGKSVFPAAADSNWTTESIVRF 227
Chondrosia_reniformis	PAGIVTNKEVAESLGGATVSS-IVMFQKFDGKVVVSG---EYNGVDILSF 226
Sycon_ciliatum	RYGLTSSSDVAKEYGIESEFQIVAVRTFDEPQVQVQDQ--DNYTAEASIF 241
Caenorhabditis_elegans	FFAVAGSAEVAASALNEDGVALITDGDSDSETSTIAEAETINTIALKQW 228
Arabidopsis_thaliana	DFAHSTDAKLLPRGESSVTGPVVRFLKPFDEQFVDSK---DFDGEALEKF 237
Homo_sapiens_A1_	IKHNQLPLVIEFTEQT---APKIFGGEIKTHLLFLPKSVSDYDGLKSN 274
Drosophila_melanogaster	AQVQSLPLIVDFNHES---ASKIFGGSIKSHLLFPVSRGGHIEKYVDP 274
Amphimedon_queenslandica	ISDERLPVYVTLFSDET---APIIFGGSIKNHLSPFASDDEKYEYTMEN 273
Chondrosia_reniformis	VNAERLPLFIEFNKDT---ANSIFSGIKVHFLVFSDDSDRAASISEA 272
Sycon_ciliatum	ARLSSMPVYIEFSQDN---AAKIFSGVTKRQFLYFGSQQSEDEYSHVAV 287
Caenorhabditis_elegans	LHAYKLSAVTEFTHES---AQEIVGGDLKPKHFLIIRKSDSEFDETIK 274
Arabidopsis_thaliana	VKESSIPLITVFDKDPNHPVYIKFPESNTKAMLFINFTGGAESLSKSK 287
Homo_sapiens_A1_	FKTAESFPGKILFIFIDSDHTDNRILEFFGLKKECAVRLITLLEE-E 323
Drosophila_melanogaster	LKEIAKRYDDILFVTISSDEEDHTRIFEFMGKNEEVPTRLIKLEE-D 323
Amphimedon_queenslandica	LKVIKGEFRGKIVVHIDSKKESERIMEFFGTTKDDLPALRIIHLSE-D 322
Chondrosia_reniformis	VTAVAKEFRGRIIVRVNTDVENNRIIVEFFNLKEDDFPCGRILQLDD-D 321
Sycon_ciliatum	LLEQAKKHGEMLFVFTNTDEENVKILDVFLGNSNTSPAVRIVIELAARD 337
Caenorhabditis_elegans	FTEVAKKFRAKIVFLVLDVBEENARILEPLVGDVAKMTANRIVISLAD-Q 323
Arabidopsis_thaliana	YREVATSNKGGQLSFLLG-DAENSQGAFOYFGLSESVQLIIQTADD-- 334
Homo_sapiens_A1_	MTKYKPESEELTAERITFPCHRFLEGIKPHLMSQELPEDWDKQPVKVLV 373
Drosophila_melanogaster	MAKYKPESDLSAETIEAFKFLDGLKIPHLLSQELPEDWDKQPVKVLV 373
Amphimedon_queenslandica	MKKYRDPQEIETEKLRGFPVQGLDGTITPHLNTTEVPEDDWDKQPVKVLV 372
Chondrosia_reniformis	MRKFRDPDLSNEENIRTLTNFFDGGKLPHLSESDIPEDWDKQPVKVLV 371
Sycon_ciliatum	MKKYQPTGETADVIAPPIEGVNDGSIKQSLKSEIPEDDWDKQPVKVLV 387
Caenorhabditis_elegans	VEKPKPQ---EGEDFEATNSYLEGKSAQDLKADLPEDWNLQPVKVLV 369
Arabidopsis_thaliana	-KKYLKTN--VEVDQIBSWVDKFDKDKIAPHKKSQPIPAENN-EPVKVVV 380
Homo_sapiens_A1_	GKNFEDVAFDEKKNVVEFYAPWCGHCKQLAPITWIKLIGETKYKHENIVIA 423
Drosophila_melanogaster	SNFPEVALDKSKSVLVEFYAPWCGHCKQLAPITYDQLAEKYNKEDNIVIA 423
Amphimedon_queenslandica	GKNFKEVALDETKHAFVEFYAPWCGHCKQLAPITWIKLIGEBHYKNDQIVIA 422
Chondrosia_reniformis	GRNFWDVYKSG-KDVFVEFYAPWCGHCKQLAPITWDLTEGKVFVDS-VIIA 419
Sycon_ciliatum	GKNFNDVIINNEKHALIQFYAPWCGHCKKLMFVWDEVEGYFDGRDDHVVG 437
Caenorhabditis_elegans	ASNFNEIALDETKTVFVFKYAPWCGHCKQLVFPVWDELAEKYESNPNVIA 419
Arabidopsis_thaliana	SDSLDDIVLNSGKNVLLFVYAPWCGHCKQLAPILDEVAVPSQSDSSVVIA 430
Homo_sapiens_A1_	KMDSTANEVEAVKVSFPPTLKFPPASADRTVIDYNGERTLDGFKKFLSE 473
Drosophila_melanogaster	KMDSTANELESIKSSFPPTIKYPRK-EDNKVIDPNLDRITLDDFKFLDAN 472
Amphimedon_queenslandica	KMDSTKNEVDGIQITGFPPTIKFPKSGKES-HDYVGGRTQEDLIQYVEDR 471
Chondrosia_reniformis	KIDSTANEIDGKIBGPTLKYPKSDDPESPIDPMGRTLDDLVKPFVSK 469
Sycon_ciliatum	KMDATANEVGLKIBSFPPTIYFLPNGKSRPKTYKTRLEALLVFAEAG 487
Caenorhabditis_elegans	KLDATLNEADLVKNSFPPTLKLWPAGSTP-VVDYGDNRLEKPEFBNKY 468
Arabidopsis_thaliana	KLDATANDFPKDTPDKGFPPTIYFKSASGNVVVEGDRTEKEDPISFVDKN 480
Homo_sapiens_A1_	GQ-----DGAGDDDDLEDLEAEPEPD---MEEDDDQKAVKDEL 508
Drosophila_melanogaster	G-----EVADESEPEETE---EIEEEAPK---KDEL 496
Amphimedon_queenslandica	L----AGKPLEKGEE--VDEGMDTMEDMNDPAMPPEEEDGATRDEL 514
Chondrosia_reniformis	G----TDRGVKPESEDLDGDFPDSLEDEDEDEDEDEDEDEGMDQED 516
Sycon_ciliatum	GRDGGRPDPHPVPAKP-KPEPADFEADEEDLPLETEEKEKILAEAEQ 534
Caenorhabditis_elegans	AG-----SASESETASQDHEEL 485
Arabidopsis_thaliana	KD-----TVGEPKKEEETTEEVKDEL 501
Homo_sapiens_A1_	-----
Drosophila_melanogaster	-----
Amphimedon_queenslandica	-----
Chondrosia_reniformis	VGEPAKDKKIEL-- 526
Sycon_ciliatum	AAAEAEEMPAKEL 550
Caenorhabditis_elegans	-----
Arabidopsis_thaliana	-----



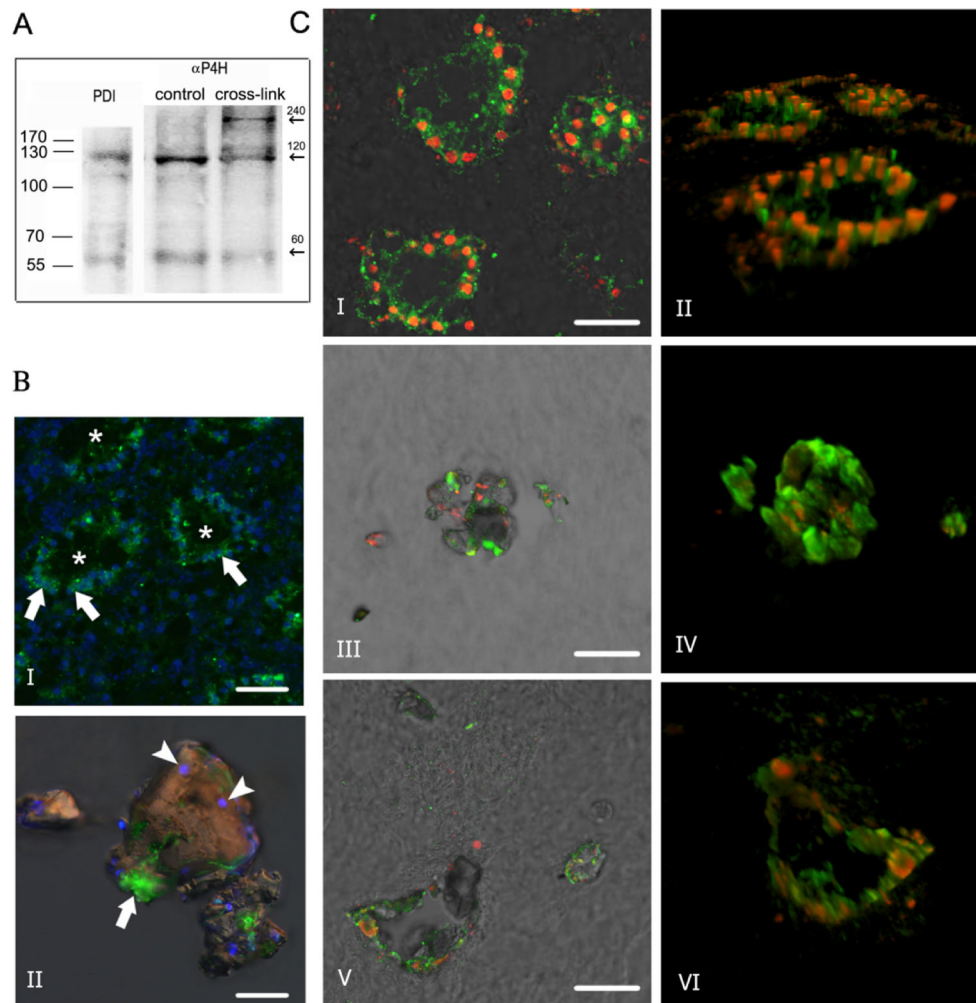


Fig. 5 **a** Western blot analysis of PDI (β P4H) and α P4H expression in *C. reniformis* marine sponge. On the left are standard molecular weight markers (kDa), and on the right are molecular weights (kDa) of major immunoreactive bands (arrows). **b** Epifluorescence of α P4H staining (green) in *C. reniformis*: nuclei counterstained with DAPI (blue) and obtained with a Leica DMRB microscope. **I** Choanosome: immunofluorescence is visible around choanocyte chambers (asterisks), in the cytoplasm of choanocytes (arrows) surrounding nuclei. **II** Ectosome: immunofluorescence is detectable close to the sand grain (arrow), and fluorescence overlapped to Nomarski contrast image. Some cells close to the grain (arrowheads) are not immunoreactive. To obtain this image, six different photographs (three channels and two

different focus levels) were combined. Scale bars, 20 μ m. **c** Confocal microscopy of α P4H immunoreactivity (green) in *C. reniformis*: nuclei counterstained with PI (red) and obtained with a Leica TCS SL microscope. **I** Choanosome: choanocytes (green immunofluorescence) and red nuclei, single plane confocal image overlapped to light Nomarski contrast. **II** 3D reconstruction of α P4H immunoreactivity and nuclei in the choanocyte chambers from stacks of 50 sections with a Z-step of 122 nm for a total thickness of 5.9–6.0 μ m. Scale bar 20 μ m. **III–V** Ectosome: α P4H immunoreactivity (green) around sand grains, single plane confocal image overlapped to Nomarski contrast. Scale bars 40 μ m. **IV–VI** cropped 3D reconstruction of α P4H immunoreactivity and nuclei close to sand grains from stacks of 50 sections

regarding the molecular cloning of sponge collagen genes (Exposito and Garrone 1990 and Exposito et al. 1990) is available, there are to date no data regarding the enzymes involved in the post-translational modifications of these proteins in Porifera that would definitely enable the recombinant production of sponge collagens on a large scale.

The present data report the first molecular cloning and characterization of the full-length cDNAs encoding for the α and β subunits of the prolyl 4-hydroxylase enzyme derived from the marine sponge *C. reniformis* which is directly involved in the post-translational modifications of collagen.

The deduced amino acid sequence of the α subunit consists of 533 residues, with a cleavable signal peptide of 27 additional residues (Fig. 1). The identification of the amino acids involved in the Fe^{2+} and 2-oxoglutarate binding (Fig. 2a), by multiple alignment, likely indicates similar enzymatic kinetics in the species analyzed, as well as the presence of an additional His residue involved in the 2-oxoglutarate C1 carboxyl group binding and decarboxylation. In *Drosophila* α P4H, this residue is replaced with an Arg, and accordingly, its K_m for 2-oxoglutarate is four times higher with respect to the K_m of human α P4H (Annunen et al. 1999); Hence, sponge α P4H

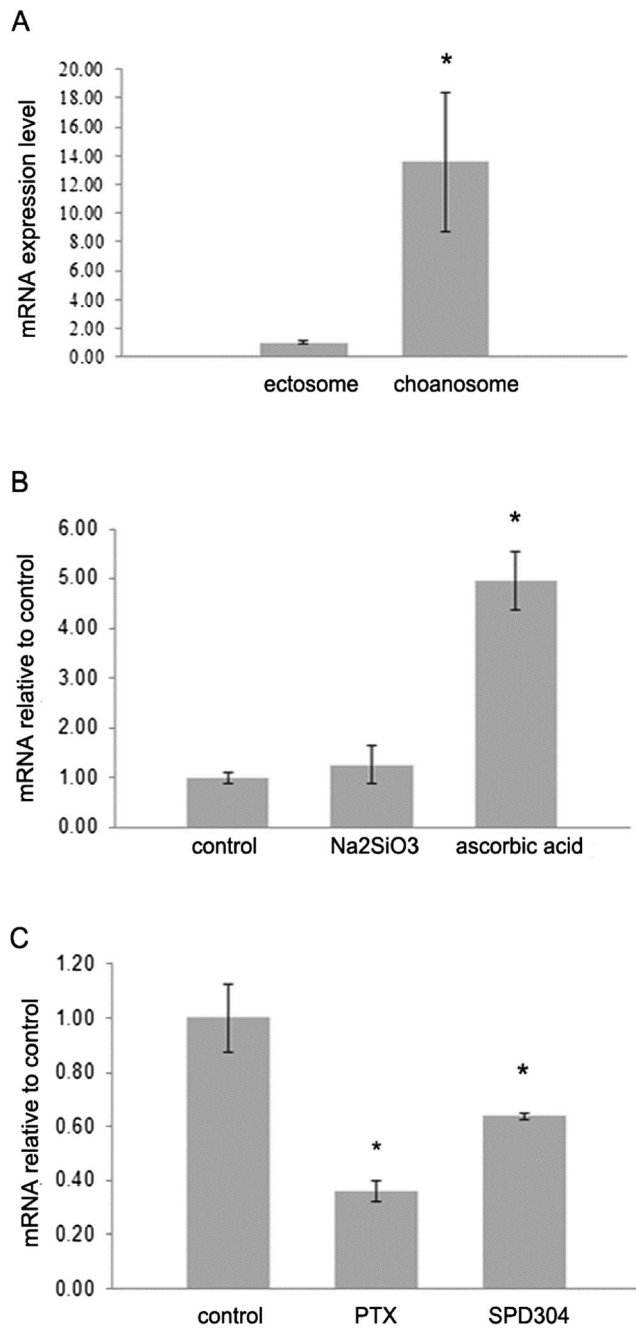


Fig. 6 **a** Relative mRNA expression determined by qPCR in different sponge tissues, from ectosome samples and choanosome samples, respectively. Each bar represents the mean of three replicates ($n=3$; $p<0.005$). **b** α P4H mRNA expression analysis in 2-month-old fragmorphs treated with 120 μ M Na₂SiO₃ or 120 μ M ascorbic acid for 24 h. Fold increase expression is compared to untreated (control) sample. Each bar represents the mean of three replicates ($n=3$; $*p<0.005$ with respect to control). **c** α P4H mRNA expression analysis in 2-month-old fragmorphs treated with 2 mM pentoxifylline (PTX) or with 22 μ M SPD304 for 48 h, two specific TNF- α inhibitors. Fold increase expression is compared to untreated (control) sample. Each bar represents the mean of three replicates ($n=3$; $*p<0.005$ with respect to control)

presumably possesses a higher enzymatic activity similar to its human counterpart. By the comparison of *C. reniformis* α P4H

with other α P4Hs in a hypothetical phylogenetic tree obtained with ClustalW2 (Fig. 2b), an overall higher level of identity with the mammalian form (42 %) with respect to all the other invertebrate proteins is inferable. Thus, the tree suggests that even if taxonomically very distant, during the evolution process, sponge and mammalian P4Hs have evolved in parallel from a common ancestor coming to a similar successful molecular organization. This evidence is consistent with previous studies on protein parallelism between sponges and humans that has also been highlighted in other past occasions. Gamulin et al. in fact clearly showed a significant higher similarity of several sponge proteins to mammals and humans, with respect to homologous polypeptides in *C. elegans* (Gamulin et al. 2000).

The multiple alignment reveals also the presence of the five conserved cysteine residues involved in intra-chain bonds (Kivirikko and Myllyharju 1998), while the presence of an additional cysteine residue (Cys₅₁₂) suggests that the sponge P4H is most related to the human type II α P4H. This isoform expression is mainly restricted to osteoblasts, endothelial cells, and chondrocytes (Annunen et al. 1998). This could indicate that the human type II isoform may have evolved first, and not surprisingly, its localization is associated with cartilage which is considered a more ancient type of tissue than bone. The putative amino acid sequence of *C. reniformis* α P4H has dimensions similar to those of vertebrate α P4Hs and lacks the C-terminal extension which rather characterizes the *C. elegans* α P4Hs. The higher level of similarity of sponge α P4H with the mammalian α P4H compared to the nematode also suggests that the *C. reniformis* protein, in its native state, forms an $\alpha_2\beta_2$ tetramer with its β P4H subunit (PDI) and not a dimer as in *C. elegans* (Veijola et al. 1994).

As for the α P4H, also the deduced amino acid sequence of the β subunit (PDI) has a signal peptide and shows an identity level with *A. queenslandica* PDI A2 of 44.5 % (Fig. 4b). Overall, *C. reniformis* PDI presents the typical structure of higher organism PDIs, the only different feature being an acidic amino acid rich C-terminal region of about 20 residue longer than the other PDIs. This protein region in particular is also present in the calcareous sponge *S. ciliatum*. Western blot analyses on sponge cell lysates, in non-reducing conditions, showed two bands at the expected molecular weights of α P4H and PDI (58 and 59 kDa, Fig. 5a), while in the cross-linked sample, the presence of a band corresponding to 240 kDa strengthens our hypothesis of the tetramer as the active form of this enzyme also in sponges.

The specific tissue expression pattern of *C. reniformis* α P4H was approached in two different ways: by immunohistochemistry (Fig. 5b, c) and by comparing the α P4H mRNA transcript levels between ectosome and choanosome by qPCR (Fig. 6a). These two techniques indicate that the sponge P4H is more expressed in the choanosome region. In particular, in the immunohistochemistry analysis, the enzyme appears

mainly surrounding the numerous choanocyte chambers (Fig. 5b (I), c (I–II) and Online resource Movie 1) with similar intensity among the different chambers. This observation is not surprising as, in vertebrates, P4H is known to be expressed also in endothelial and epithelial cells (Nissi et al. 2001; Steiling et al. 2009), and accordingly, collagen synthesis is known to occur in different cell types (Bader et al. 2013; Edogawa et al. 2014). As *C. reniformis* endosome is extremely dynamic (Bonasoro et al. 2001), it is possible to speculate a main role of choanocytes in the regulation of collagen deposition in the endosome, at least the fibrillar one, in order to maintain a functional aquiferous system both in quiescence as well as in changing shape.

Conversely, in the cortical region (ectosome), the transcript levels are lower (Fig. 6a) and the ir for the enzyme appears only in close proximity of the sediment grains that the animal usually incorporates (Fig. 5b (II), c (III–VI)). This latter localization could be strictly related to a particular phenomenon taking place in these sponges which have the amazing capability to selectively engulf siliceous particles in the ectosome (Bavestrello et al. 1998) and to etch quartz grains in the same district (Bavestrello et al. 1995). The etching action, developed in a sort of fibrotic envelope around the quartz grains, is caused by a peculiar chemical reaction of ascorbic acid, abundant in these sponges (Cerrano et al. 1999), slowly dissolving crystalline silica (Bavestrello et al. 1995, 2003). Thus, the local presence of soluble silicates as well as of ascorbic acid likely stimulates, in paracrine cells, both the production of collagen (either by soluble silicates (Pozzolini et al. 2012) and by ascorbic acid (Lyons and Schwarz 1984)) and of P4H (by ascorbic acid solely, as reported here) strengthening the fibrotic envelope around the engulfed quartz particles. This ancestral response to quartz in *C. reniformis* could be the precursor of the well-known pathological responses of the mammalian lung to silica inhalation where the formation of fibrotic nodules around crystalline quartz is observed as well, possibly explaining one of the reasons behind the onset of pulmonary fibrosis in humans. Differently from sponges, in fact, the mammalian lung tissue is unable to dissolve the crystalline particles leading with time to the development of silicosis (Giovine et al. 2013).

The low transcript levels of this enzyme in the cortical region, compared to the internal region (choanosome), are in contrast with the distribution of a non-fibrillar collagen previously described in the same animal, which in turn results more expressed in the ectosome (Pozzolini et al. 2012). It is possible that the P4H identified in the present work is mainly associated with the maturation of fibrillar collagen types, whose transcripts are, in fact, more abundant in the choanosome as observed in preliminary experiments (data not shown).

External factors possibly affecting P4H gene regulation in sponges were investigated by use of an in vitro model of *C. reniformis* tissue explant (fragmorph). Understanding the

mechanisms influencing the fibrogenesis and consequently the biomass growth of these animals is of extreme interest for their biotechnological potentialities as collagens and pharmacological metabolite producers increasing their performances in mariculture or in ex situ culture systems (Bergman et al. 2011; Sipkema et al. 2005). Thus, we analyzed the effect of two compounds on P4H gene expression: soluble silicates, previously described as able to enhance non-fibrillar collagen expression in the same animal (Pozzolini et al. 2012; Giovine et al. 2013), and ascorbic acid, a fundamental enzymatic co-factor of the proline hydroxylation reaction and a well-known positive regulator of collagen gene expression (Lyons and Schwarz 1984). The results obtained indicate that while silicates do not affect P4H transcript levels in sponges, surprisingly, micromolar ascorbic acid supplementation leads to a fivefold increase of the same transcript (Fig. 6b). To our knowledge, this is the first time that ascorbic acid is reported to stimulate P4H expression at the level of its transcript; hence, this result suggests further studies to confirm this behavior in higher organisms and to elucidate its signal transduction pathway. Evidently in sponges, while silicates act specifically only on collagen gene regulation, ascorbic acid seems to carry out its regulatory role on a wider display of genes involved in the biosynthesis of the extracellular matrix (i.e., collagen and P4H).

Some of the more detailed information on P4H gene regulation in mammals concern the negative effects of *pro*-inflammatory cytokines, such as TNF- α , in human aortic smooth muscle cells (AoSMC) involved in cardiovascular diseases (Zhang et al. 2008). For what concerns TGF- β , although this cytokine is considered one of the master signals of fibrosis and of myofibroblast differentiation (Gharaee-Kermani et al. 2009), to date no data are available on its direct involvement in P4H gene expression. Furthermore, in sponges and in invertebrates (molluscs, nematodes, insects, and echinoderms), no TGF- β genes have been described so far but only BMP family-related proteins (Dzik 2010). Hence, these proteins and their receptors may not be involved in immune responses in invertebrates.

Thus, in order to investigate the possible involvement of cytokines on sponge P4H gene regulation, fragmorphs were treated with two different TNF inhibitors, and transcript levels were quantified accordingly. Interestingly, the sponge P4H gene resulted strongly repressed in presence of both TNF inhibitors (Fig. 6c). These results suggest the existence of some TNF-like agonist with positive effects on P4H gene expression, differently from AoSMC. This positive effect could better find an analogy with the behavior of TNF- α in mammalian pulmonary fibrosis, where this cytokine raises a strong fibrogenic response (Piguet et al. 1990). The molecular characterization of TNF in sponges, with further definition of its physiological role in these primitive animals, is certainly object for future work also in the effort to better elucidate the

peculiar sponge cell reactivity to siliceous grains (Giovine et al. 2013).

In conclusion, the knowledge of the mechanisms underlying the regulation of genes involved in the extracellular matrix biosynthesis in sponges, such as the P4H gene involved in collagen hydroxylation, has a meaning, not only for biotechnological purposes of recombinant sponge collagen production but also to better understand similar physiological as well as pathological mechanisms in higher animals.

Acknowledgments This work was supported by grant from EU (FP7 grant agreement no. 266033 SPongE Enzyme and Cell for Innovative AppLication—SPECIAL) to Marco Giovine and by the University of Genova Funding (PRA 2012) to Sonia Scarfi. The authors thank Carlo Cerrano for the supplying of fresh marine sponges.

Conflict of Interest The authors declare no conflict of interests.

References

- Aarskog NK, Vedeler CA (2000) Real-time quantitative polymerase chain reaction. A new method that detects both the peripheral myelin protein 22 duplication in Charcot-Marie-Tooth type 1A disease and the peripheral myelin protein 22 deletion in hereditary neuropathy with liability to pressure palsies. *Hum Genet* 107:494–498
- Abrams EW, Andrew DJ (2002) Prolyl 4-hydroxylase alpha-related proteins in *Drosophila melanogaster*: tissue-specific embryonic expression of the 99F8-9 cluster. *Mech Dev* 112:165–171
- Annunen P, Autio-Harmainen H, Kivirikko KI (1998) The novel type II prolyl 4-hydroxylase is the main enzyme form in chondrocytes and capillary endothelial cells, whereas the type I enzyme predominates in most cells. *J Biol Chem* 273:5989–5992
- Annunen P, Koivunen P, Kivirikko KI (1999) Cloning of the alpha subunit of prolyl 4-hydroxylase from *Drosophila* and expression and characterization of the corresponding enzyme tetramer with some unique properties. *J Biol Chem* 274(10):6790–6796
- Bader HL, Lambert E, Guiraud A, Malbouyres M, Driever W, Koch M, Ruggiero F (2013) Zebrafish collagen is required for proper function of certain basement membranes. *J Biol Chem* 288(10):6777–6787
- Bavestrello G, Arillo A, Benatti U, Cerrano C, Cattaneo-Vietti R, Cortesogno L, Gaggero L, Giovine M, Tonetti M, Sarà M (1995) Quartz dissolution by the sponge *Chondrosia reniformis* (Porifera, Demospongiae). *Nature* 378:374–376
- Bavestrello G, Benatti U, Calcinai B, Cattaneo-Vietti R, Cerrano C, Favre A, Giovine M, Lanza S, Pronzato R, Sarà M (1998) Body polarity and mineral selectivity in the demosponge *Chondrosia reniformis*. *Biol Bull* 195:120–125
- Bavestrello G, Benatti U, Cattaneo-Vietti R, Cerrano C, Giovine M (2003) Sponge cell reactivity to various forms of silica. *Microsc Res Tech* 62:327–335
- Bergman O, Mayzel B, Anderson MA, Shpigel M, Hill RT, Ilan M (2011) Examination of marine-based cultivation of three demosponges for acquiring bioactive marine natural products. *Mar Drugs* 9(11):2201–2219
- Bonasoro F, Wilkie IC, Bavestrello G, Cerrano C, Candia Carnevali MD (2001) Dynamic structure of the mesohyl in the sponge *Chondrosia reniformis* (Porifera, Demospongiae). *Zoomorphology* 121:109–121
- Bradford MA (1976) Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72:248–254
- Cao M, Westerhausen-Larson A, Niyibizi C, Kavalkovich K, Georgescu HI, Rizzo CF, Hebda PA, Stefanovic-Racic M, Evans CH (1997) Nitric oxide inhibits the synthesis of type-II collagen without altering Col2A1 mRNA abundance: prolyl hydroxylase as a possible target. *Biochem J* 324:305–310
- Cerrano C, Bavestrello G, Arillo A, Benatti U, Bonpadre S, Cattaneo-Vietti R, Gaggero L, Giovine M, Leone L, Lucchetti G, Sarà M (1999) Calcium oxalate production in the marine sponge *Chondrosia reniformis*. *Mar Ecol Prog Ser* 179:297–300
- Chen LI, Shen YH, Wang X, Wang J, Gan Y, Chen N, Wang J, LeMaire SA, Coselli JS, Wang XL (2006) Human prolyl-4-hydroxylase alpha(I) transcription is mediated by upstream stimulatory factors. *J Biol Chem* 281:10849–10855
- Dzik JM (2010) The ancestry and cumulative evolution of immune reactions. *Acta Biochim Pol* 57:443–466
- Edogawa S, Sakai A, Inoue T, Harada S, Takeuchi T, Umegaki E, Hayashi H, Higuchi K (2014) Down-regulation of collagen I biosynthesis in intestinal epithelial cells exposed to indomethacin: a comparative proteome analysis. *J Proteomics* 103:35–46
- Exposito JY, Garrone R (1990) Characterization of a fibrillar collagen gene in sponges reveals the early evolutionary appearance of two collagen gene families. *Proc Natl Acad Sci U S A* 7:6669–6673
- Exposito JY, Ouazana R, Garrone R (1990) Cloning and sequencing of a Porifera partial cDNA coding for a short-chain collagen. *Eur J Biochem* 190:401–406
- Gamulin V, Muller I, Muller WEG (2000) Sponge proteins are more similar to those of *Homo Sapiens* than to *Caenorabditis elegans*. *Biol J Linn Soc* 71:821–828
- Garrone R, Hu A, Junqua S (1975) Fine structure and physicochemical studies on the collagen of the marine sponge *Chondrosia reniformis* Nardo. *J Ultrastruct Res* 52:261–275
- Gharaee-Kermani M, Hu B, Phan SH, Geytko MR (2009) Recent advances in molecular targets and treatment of idiopathic pulmonary fibrosis: focus on TGF-beta signaling and the myofibroblasts. *Curr Med Chem* 16:1400–1417
- Gilber HF (1997) Protein disulfide isomerase and assisted protein folding. *J Biol Chem* 272:29399–29402
- Giovine M, Scarfi S, Pozzolini M, Penna A, Cerrano C (2013) Cell reactivity to different silica. *Prog Mol Subcell Biol* 54:143–174
- He MM, Stroustrup Smith A, Oslob JD, Flanagan WM, Braisted AC, Whitty A, Cancilla MT, Wang J, Lugovskoy AA, Yoburn JC, Fung AD, Farrington G, Eldredge JK, Day ES, Cruz LA, Cachero TG, Miller SK, Friedman JE, Choong IC, Cunningham BC (2005) Small-molecule inhibition of TNF-alpha. *Science* 310:1022–1025
- Heinemann S, Ehrlich H, Douglas T, Heinemann C, Worch H, Schatton W, Hanke T (2007) Ultrastructural studies on the collagen of the marine sponge *Chondrosia reniformis* Nardo. *Biomacromolecules* 8:3452–3457
- Hiramatsu M, Kumegawa M, Hatakeyama K, Yajima T, Minami N, Kodama H (1982) Effect of epidermal growth factor on collagen synthesis in osteoblastic cells derived from newborn mouse calvaria. *Endocrinology* 111:1810–1816
- Kawaguchi Y, Kitani A, Hara M, Harigai M, Hirose T, Suzuki K, Kawakami M, Hidaka T, Ishizuka T, Kawagoe M (1992) Cytokine regulation of prolyl 4-hydroxylase production in skin fibroblast cultures from patients with systemic sclerosis: contribution to collagen synthesis and fibrosis. *J Rheumatol* 9:1195–1201
- Kersteen EA, Rainers RT (2003) Catalysis of protein folding by protein disulfide isomerase and small-molecule mimics. *Antioxid Redox Signal* 5:413–424
- Kivirikko KI, Myllyharju J (1998) Prolyl 4-hydroxylase and their protein disulfide isomerase subunit. *Matrix Biol* 16:357–368
- Kivirikko KI, Myllylä R, Pihlajaniemi T (1989) Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrate and a multi-functional subunit. *FASEB J* 3:1609–1617

- Kukkola L, Hieta R, Kivirikko KI, Myllyharju J (2003) Identification and characterization of a third human, rat, and mouse collagen prolyl 4-hydroxylase isoenzyme. *J Biol Chem* 278:47685–47693
- Lyons BL, Schwarz RI (1984) Ascorbate stimulation of PAT cells causes an increase in transcription rates and a decrease in degradation rates of procollagen mRNA. *Nucleic Acids Res* 12:2569–2579
- Muguerza B, Castilla-Cortázar I, García M, Quiroga J, Santidrián S, Prieto J (2001) Antifibrogenic effect in vivo of low doses of insulin-like growth factor-I in cirrhotic rats. *Biochim Biophys Acta* 1536:185–195
- Myllyharju J (2003) Prolyl 4-hydroxylase, the key enzymes of collagen biosynthesis. *Matrix Biol* 22:15–24
- Nicklas M, Schatton W, Heinemann S, Hanke T, Kreuter J (2009a) Preparation and characterization marine sponge collagen nanoparticles and employment for the transdermal delivery of 17 β -estradiol-hemihydrate. *Drug Dev Ind Pharm* 35:1035–1042
- Nicklas M, Schatton W, Heinemann S, Hanke T, Kreuter J (2009b) Enteric coating derived from marine sponge collagen. *Drug Dev Ind Pharm* 35:1384–1388
- Nietfeld JJ, Van der Kraan I, Kemp A (1981) Dissociation and reassociation of prolyl 4-hydroxylase subunits after cross-linking of monomers. *Biochim Biophys Acta* 661:21–27
- Nissi R, Autio-Harmainen H, Marttila P, Sormunen R, Kivirikko KI (2001) Prolyl 4-hydroxylase isoenzymes I and II have different expression patterns in several human tissues. *J Histochem Cytochem* 49:1143–1153
- Piguet PF, Collart MA, Grau GE, Sappino AP, Vassalli P (1990) Requirement of tumour necrosis factor for development of silica-induced pulmonary fibrosis. *Nature* 344:245–247
- Poulakis N, Androutsos G, Kazi D, Bastas A, Provata A, Bitsakou C, Kontozoglou T, Polyzogopoulou C, Tassiopoulou A (1999) The differential effect of pentoxifylline on cytokine production by alveolar macrophages and its clinical implications. *Respir Med* 93:52–55
- Pozzolini M, Bruzzone F, Berilli V, Mussino F, Cerrano C, Benatti U, Giovine M (2012) Molecular characterization of a nonfibrillar collagen from the marine sponge *Chondrosia reniformis* Nardo 1847 and positive effects of soluble silicates on its expression. *Mar Biotechnol* 14:281–293
- Pozzolini M, Mussino F, Cerrano C, Scarfi S, Giovine M (2014) Sponge cell cultivation: optimization of the model *Petrosia ficiformis* (Poiret 1789). *JEMBE* 454:70–77
- Raveendrana M, Senthila D, Utama B, Shena Y, Dudley D, Wanga J, Zhang Y, Wang XL (2004) Cigarette suppresses the expression of P4H α and vascular collagen production. *Biochem Biophys Res Commun* 323:592–598
- Sipkema D, Osinga R, Schatton W, Mendola D, Tramper J, Wijffels RH (2005) Large-scale production of pharmaceuticals by marine sponges: sea, cell, or synthesis? *Biotechnol Bioeng* 90:201–222
- Steiling K, Kadar AY, Bergerat A, Flanigan J, Sridhar S, Shah V, Ahmad QR, Brody JS, Lenburg ME, Steffen M, Spira A (2009) Comparison of proteomic and transcriptomic profiles in the bronchial airway epithelium of current and never smokers. *PLoS ONE* 4, e5043
- Swatschek D, Schatton W, Kellermann J, Müller WE, Kreuter J (2002) Marine sponge collagen: isolation, characterization and effects on the skin parameters surface-pH, moisture and sebum. *Eur J Pharm Biopharm* 53:107–113
- Vandescompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3(7):RESEARCH0034
- Veijola J, Koivunen P, Annunen P, Pihlajaniemi T, Kivirikko KI (1994) Cloning, baculovirus expression, and characterization of the alpha subunit of prolyl 4-hydroxylase from the nematode *Caenorhabditis elegans*. This alpha subunit forms an active alpha beta dimer with the human protein disulfide isomerase/beta subunit. *J Biol Chem* 269(43):26746–26753
- Wilkie IC, Parma L, Bonasoro F, Bavestrello G, Cerrano C, Carnevali MD (2006) Mechanical adaptability of a sponge extracellular matrix: evidence for cellular control of mesohyl stiffness in *Chondrosia reniformis* Nardo. *J Exp Biol* 209:4436–4443
- Zhang C, Zhang MX, Shen YH, Burks JK, Zhang Y, Wang J, LeMaire SA, Yoshimura K, Aoki H, Coselli JS, Wang XL (2008) TNF-alpha suppresses prolyl-4-hydroxylase alpha1 expression via the ASK1-JNK-NonO pathway. *Arterioscler Thromb Vasc Biol* 27(8):1760–1767