

Contents lists available at ScienceDirect

# Achievements in the Life Sciences



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

# Molecular Cloning and Homology Modeling of Novel Tyrosylprotein Sulfotransferase of Marine Mollusk

Maria S. Avtushenko<sup>a,\*</sup>, Svetlana N. Kovalchuk<sup>a</sup>, Galina N. Likhatskaya<sup>a</sup>, Tatyana N. Zvyagintseva<sup>a</sup>, Thomas Haertlé<sup>b</sup>

<sup>a</sup> G.B. Elyakov Pacific Institute of Bioorganic Chemistry, Far Eastern Branch, Russian Academy of Sciences, Vladivostok 690022, Russia
<sup>b</sup> UR1268, Biopolymeres, Interactions, Associations, Institut National de la Recherche Agronomique, B.P. 71627, 44316 Nantes, Cedex 3, France

## ARTICLE INFO

Article history: Received 14 February 2016 Accepted 1 June 2016 Available online 16 June 2016

Keywords: Tyrosylprotein sulfotransferase TPST Post-translational modification Marine mollusk Littorina sitkana Homology modeling

## ABSTRACT

The gene of tyrosylprotein sulfotransferase, which was discovered in mammals, has been widely found in marine mollusk *Littorina sitkana*. High conservation of this gene indicates the functional importance of TPST in the metabolism of the living world. The cDNA encoding TPST in the mollusk was cloned and sequenced, and the enzyme was assigned on the basis of amino acid sequence similarity as tyrosylprotein sulfotransferase-2 (TPST-2). The putative homology model for the catalytic domain of TPST from *L. sitkana* was constructed according to crystal structure of the catalytic domain of the human TPST-2. The putative model of dimer structure showed that the active site involved two monomers and the dimer contains two active centers.

© 2016 Far Eastern Federal University. Hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

## Introduction

Tyrosylprotein sulfotransferases (TPSTs<sup>1</sup>, EC 2.8.2.20) are the Golgi-localized type II transmembrane proteins, which transfer a sulfuryl group  $(SO_3^{--})$  from the universal sulfate donor 3'-phosphoadenosine-5'-phosphosulfate (PAPS) to the hydroxyl of tyrosine side chain (Lee and Huttner, 1983). Tyrosine sulfation is a common post-translational modification of peptides and proteins. TPST activity has been described in many species of animals and plants (Niehrs and Huttner, 1990; Sane and Baker, 1993; Kasinathan et al., 2005; Nishimura and Naito, 2007; Hanai et al., 2000). In 2012, this enzyme was first discovered in a Gramnegative bacterium (Han et al., 2012), although tyrosine sulfation has not previously been reported in prokaryotes. In spite of the functional importance of these enzymes, almost all known TPSTs are of mammalian origins. Only few enzymes from invertebrates have been described (from nematodes *Caenorhabditis elegans* and *Brugia malayi*, insects *Culex quinquefasciatus*, *Drosophila* spp., and *Anopheles gambiae*, sea squirts *Ciona intestinalis* and *Halocynthia roretzi*, trematode *Schistosoma japonicum*, marine mollusk *Crassostrea gigas*) mainly via cDNA or genome sequencing.

The sulfation of biomolecules plays an important role in the metabolism of pro- and eukaryotes. The growing scientific interest in the sulfation of different natural compounds stems from the high biological activities of sulfated derivatives and their role in organisms. The known roles of tyrosine sulfation in mammals are various and multiple, for example maintenance of hemostasis,

*E-mail address*: mpesentseva@gmail.com (M.S. Avtushenko). Peer review under responsibility of Far Eastern Federal University.

http://dx.doi.org/10.1016/j.als.2016.06.001

*Abbreviations*: TPST, tyrosylprotein sulfotransferases; PAPS, 3'-phosphoadenosine-5'-phosphosulfate; LsTPST, tyrosylprotein sulfotransferases from *Littorina sitkana.* \* Corresponding author at: 159, Pr-t 100-letiya Vladivostoka, Vladivostok 690022, Russia. Tel./fax: +7 423 231 07 05.

<sup>2078-1520/© 2016</sup> Far Eastern Federal University. Hosting by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

triggering inflammatory responses, strengthening leucocyte adhesion, defining specificity of chemokine receptors and enhancement of potency of bioactive peptides (Coughtrie et al., 1998; Kehoe and Bertozzi, 2000; Moore, 2003; Seibert and Sakmar, 2008).

Marine hydrobionts are extremely rich in sulfated metabolites with diverse biological activities (Kornprobst et al., 1998). Unfortunately, sulfation/desulfation processes are studied considerably less than, for example, phosphorylation. The knowledge about the proteins undergoing sulfation in marine organisms is particularly poor. Frequently enzymes from marine habitats have unique properties; some of them are very useful for biotechnological applications. The enzymes with interesting specificities and high level of activities were found in the digestive glands of *Littorina sitkana*. Previously we isolated and characterized from this mollusk the enzymes, which catalyze hydrolysis and transformation of carbohydrate-containing natural products. There are alginate lyase (Favorov et al., 1979), two forms of fucoidanases (Kusaykin et al., 2003; Bilan et al., 2005), endo-1,3- $\beta$ -D-glucanase (Pesentseva et al., 2012),  $\beta$ -D-glucosidase (Pesentseva et al., 2008) and sulfatase (Kusaykin et al., 2006). Herbivorous marine gastropod *L. sitkana* is widely spread on the coasts of the Pacific and the Atlantic Oceans, and it was interesting to search the TPST gene in this animal. The data of the crystal structure of human tyrosylprotein sulfotransferase-2 (Teramoto et al., 2013) and the homology model for human TPST-1 (Nedumpully-Govindan et al., 2014) will give more information about this enzyme.

Thus, the present study was devoted to the search for the tyrosylprotein sulfotransferase gene in the marine gastropod, cloning and sequencing of the cDNA encoding this protein, and 3D-structure homology modeling.

#### **Materials and Methods**

### **Biological Material**

Marine mollusks *L. sitkana* were collected in the Posieta Bay (northwestern part of the Sea of Japan) in August 2014 near the Marine Experimental Station of the Pacific Institute of Bioorganic Chemistry.

## RNA Isolation and cDNA Synthesis

Total RNA was isolated from the liver of *L. sitkana* by the TRIzol Reagent (Invitrogen, USA) and cDNA was synthesized with Mint cDNA Synthesis Kit (Eurogen, Russia) according to the provided protocols.

#### Amplification of TPST cDNA

The fragments of cDNA of TPST from *L. sitkana* were obtained by PCR using cDNA from *L. sitkana*. The SU-F1, SU-F2, SU-R2, SU-R3, and SU-R4 primers (Table 1) synthesized on the basis of the conserved peptides were used for the amplification, which was carried out for 35 cycles (10 s at 95 °C, 15 s at 55 °C, 40 s at 72 °C). The terminal cDNA regions were obtained by rapid amplification of the cDNA fragments (RACE). The amplification was performed with SU-5race and SU-3race primers for 38 cycles (10 s at 95 °C, 20 s at 63 °C, 60 s at 72 °C).

## Cloning and Sequencing of cDNA of TPST From L. sitkana

The PCR products were cloned with InsTAclone PCR Cloning Kit from Fermentas (Lithuania) according to the manufacturer's recommendations. Bacterial colonies containing plasmids with the desired insertion were screened by PCR using M13 universal primers. Nucleotide sequences were determined with ABI Prism Big Dye Terminator 3.1 Cycle Sequencing Kit from Applied Biosystems (USA) on ABI Prism 310 Genetic Analyzer.

## Analysis of Nucleotide and Amino Acid Sequences

The nucleotide and amino acid sequences were analyzed using the programs CHROMAS 2.01 (http://www.technelysium.com. au/chromas\_lite.html) and GENERUNNER 3.05. Amino acid sequence was established on the basis of the nucleotide sequence of the cDNA encoding TPST with EXPASY (http://expasy.org/tools/dna.html). Search of TPST homologous was performed using the BLAST2 (http://www.ebi.ac.uk/blastall). Domain architectures were identified by SMART tool (http://smart.embl-heidelberg.de)

Table 1	
Primers for amplification of TPST cDNA of L. sitkar	ıa.

No.	Primer	Nucleotide sequence
1	SU-F1	5'-ATTTTTATTGGIGGIGTICCICG-3'
2	SU-F2	5'-ATGCTTGATGCTCATCCTGATGT-3'
3	SU-R2	5'-ATTAACAGGTTTAATAACTTGATCAGT-3'
4	SU-R3	5'-ATCAGGTTTACCATAATTAGGAGGATT-3'
5	SU-R4	5'-ATCAGGTTTACCATAATTAGGAGGATT-3'
6	SU-5race	5'-TTCCGGTAGCTCTTGAGGT-3'
7	SU-3race	5'-GACGGAGAAGTCGACGGAC-3'

and InterProScan (http://www.ebi.ac.uk/Tools/pfa/iprscan/). The multiple sequence alignment was performed with the ClustalW2 (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

## Molecular Modeling

The sequence of LsTPST (GenBank AHJ26006.1 or Uniprot W6FFX7) was used for fold recognition and homology modeling. The LsTPST sequence analysis and fold recognition were carried out with servers PHYRE (http://www.sbg.bio.ic.ac.uk/phyre2/), I-Tasser (http://zhanglab.ccmb.med.umich.edu/I-TASSER/) and FUGUE (http://tardis.nibio.go.jp/fugue/). The LsTPST monomer and homodimer model construction and analysis were carried out with program MOE 2013.08 (Molecular Operating Environment, 2011.10, Chemical Computing Group Inc., 1010 Sherbrooke St. West, Suite #910, Montreal, QC, Canada H3A 2R7, 2013).

The model of LsTPST monomer (residues 47–346) was built with the homology module in MOE 2013.08 using as template the crystal structure of the catalytic domain (residues 52–353) of human tyrosylprotein sulfotransferase-2 monomer (TPST-2; PDB code: 3AP1; resolution 1.9 Å) (Teramoto et al., 2013). The model of LsTPST homodimer was generated using optimized structure of LsTPST monomer and human TPST-2 homodimer as template.

## Results

## cDNA Cloning and TPST Primary Structure

An amino acid sequence of the *L. sitkana* TPST was determined by the methods of molecular biology. Total RNA isolated from *L. sitkana* liver was used for the synthesis of the first strand of cDNA. For amplification of cDNA fragments encoding the TPST degenerated oligonucleotide primers were used (Table 1). These primers were synthesized on the basis of the conservative sites of the amino acid sequences of human TPSTs and the partial sequences of TPST from marine mollusks, which were found in GenBank. The cDNA fragments, obtained through PCR, were cloned in plasmid vector and sequenced. Analysis of the nucleotide sequence of this cDNA fragments by BLAST2 had shown that a fragment ~750 bp in length, obtained with primers SU-F2 and SU-R4, was homologous to cDNA encoding the TPSTs of other animals. Gene-specific primers for the amplification of the terminal cDNA sites of the mollusk *L. sitkana* TPST were constructed on the basis of the nucleotide sequence of the obtained cDNA fragments of 810 and 1500 bp in length, which were cloned and sequenced. As a result, we determined the nucleotide sequence of three overlapping fragments of cDNA encoding the LSTPST. Their mutual correlation allowed reconstruction of the complete nucleotide sequence of TPST cDNA. The amino acid sequence of the LSTPST was deduced from the nucleotide sequence. Analysis of the cDNA sequence indicates one prolonged open reading frame of 1224 bp in length, which encodes a polypeptide of 407 amino acid residues. Its calculated molecular mass is 46.63 kDa. The isoelectric point of this protein is 8.94.

#### Molecular Modeling

The LsTPST amino acid sequence analysis using I-Tasser, FUGUE and PHYRE2 servers showed that the protein central part has a fold with high confidence similar to the hTPST catalytic domain. A model of the full amino acid sequence of LsTPST, which was constructed by server I-Tasser, included cytoplasmic and transmembrane parts, and additional 4 helixes at the C-terminus (data not shown). The I-Tasser model with high C-score had TM-score 0.47, which showed the structural similarity between model of LsTPST fragment and catalytic domain of hTPST (3AP1) structures. The LsTPST sequence analysis by FUGUE and PHYRE2 servers showed with high confidence that the LsTPST fragment had folding similar to the catalytic domain of hTPST (PDB ID: 3AP1).

The alignment of amino acid sequences LsTPST and human TPST-2 showed that LsTPST had 64% of identity and 77% of similarity with the catalytic domain of TPST-2 (Fig. 1). The theoretical model of the spatial structure of the LsTPST catalytic domain (residues 47–346) was obtained by MOE 2013.08 (Fig. 2A). The crystal structure of the catalytic domain of the human TPST-2 (PDB code 3AP1\_A; Teramoto et al., 2013) was used as a template for modeling LsTPST structure. Analysis of the obtained model by MOE program showed that the monomer structure of the LsTPST catalytic domain had the following proportions of secondary structures  $\beta$ -strands (7.3%),  $\alpha$ -helix (48%), loop and disordered structure (44.7%). The structure of LsTPST was stabilized by 86 hydrogen and 19 ionic bonds, 131 hydrophobic contacts and two disulfide bonds. Monomers of LsTPST and human TPST-2 contained free SH groups Cys204 and Cys210 in the catalytic domain. Superposition of LsTPST and human TPST-2 structures showed that the value RMSD for C $\alpha$  atoms of the theoretical model and template crystal structure was 0.55 Å. Superposition of the model and template structures showed the identity of the active sites and substrate binding site structures (Fig. 2B). The LsTPST dimer was modeled (Fig. 2C) by analogy with the human TPST-2 dimer, which was functionally active.

The monomers, included in the dimer, were both involved in the formation of the active center of the LsTPST, wherein the binding of the substrate and two active centers was present in the LsTPST dimer structure. Thus, the models of the LsTPST catalytic domain in the form of monomer and dimer were generated and the residues Arg72 and Glu93 were determined as the catalytic residues (Fig. 3).



Fig. 1. Sequence alignment of the LsTPST (residues 47–346) and hTPST (residues 52–353) catalytic domains. Secondary structural elements are shown as cylinders ( $\alpha$  helix) and arrows ( $\beta$  strand). Disulfide-bonds (C96–C156 and C225–C233 for hTPST; C90–C150 and C219–C227 for LsTPST) are indicated at the bottom. The alignment was conducted using Maestro 9.3 (http://www.schrodinger.com/Maestro/).

## Discussion

The primary structure of TPST from marine mollusk *L. sitkana* was determined by cDNA sequencing. The comparison of the LsTPST amino acid sequence with two human TPSTs and the TPSTs of other invertebrate species revealed 70–85% of structural homology. It should be noted that the maximal degree of homology was 85% with the protein from *C. gigas* (GenBank EKC35269). The degree of amino acid homology with TPSTs of vertebrates is also high: 74–81%. Analysis of the sequence of LsTPST using the InterProScan server indicates that it contains a catalytic domain of TPST-2 from Arg5 to Gly355.

The multiple alignment of the amino acid sequences of TPST-1 (GenBank O60507) and TPST-2 (GenBank O60704) from human and some invertebrates — insects *Drosophila melanogaster* (GenBank AAM94031) and *C. quinquefasciatus* (GenBank EDS40555), clam *C. gigas* (GenBank EKC35269), nematodes *C. elegans* (GenBank O77081) and sea squirts *H. roretzi* (GenBank AAM09087) is presented in Fig. 4.

Their comparison demonstrated the conservative fragment PRSGTTLM. Amino acids of this fragment are involved in the binding of PAPS (Teramoto et al., 2013). The transmembrane region from Leu12 to Ser30 of LsTPST (Fig. 4) was defined with InterProScan server (http://www.ebi.ac.uk/Tools/pfa/iprscan/). It coincides with the transmembrane regions of other animal TPSTs (Stone et al., 2009).

The multiple alignment of TPST sequences revealed conserved residues, which allowed to speculate on their potential functions of LsTPST. Five of the eight cysteine residues (Cys90, Cys150, Cys204, Cys219, Cys227) are conserved and four of them formed two disulfide bonds, suggesting that these bonds are required for the stabilization of the monomeric molecule of LsTPST. Based on the available literature data (Teramoto et al., 2013) the Cys31 is a dimer-forming residue of the enzyme. Mutational analysis data obtained for human TPST-2 showed that conserved residues Arg72 and Glu93 are directly involved in catalysis, and Lys152 and Ser277, being conserved also, are needed to stabilize the transition state of the enzyme during catalysis (Teramoto et al., 2013). According to the data on the directed mutagenesis of human TPST-2 the positively charged conserved residues Arg95, Arg99, Arg116, Lys158, Lys109 and Arg276 as well as Thr192 are included in the substrate-binding site, where they contribute to the recognition of acidic amino acid residues of the acceptor molecule. The region from transmembrane site to the beginning of the catalytic domain has remarkably low sequence identity, and it is often named as unstructured "stem". There are suggestions that this region in addition to some conserved residues is responsible for the specificity of the enzyme (Stone et al., 2009).

Taking into account that data on the tyrosylprotein sulfotransferases in marine mollusks are absent we compared the amino acid sequence of LsTPST with partial sequences from marine mollusks, which was found in GenBank. This comparison has



**Fig. 2.** (A) Structural model of the LsTPST catalytic domain (residues 47–346). The structure of the protein is shown in ribbon diagram form:  $\alpha$ -helixes (red),  $\beta$ -strands (yellow), coils (white) and turns (blue). (B) Superimposition of the structural model of LsTPST (residues 47–346; yellow) with template crystal structure of hTPST (residues 52–353; blue) in complex with PAP (magenta) and C4 peptide (green). LsTPST catalytic residues Arg72 and Glu93 are shown as spacefill. (C) The model of the complex of C4 peptide (P and P') with LsTPST homodimer from M1 (magenta) and M2 (orange) monomers. C4 peptides in the active sites of M1 and M2 monomers are shown as sticks, and catalytic residues Arg72 and Glu93 are shown as spacefill.



Fig. 3. Superposition of the catalytic sites of TPST2 (white) and LsTPST (dark gray). TPST2 is complexed with PAP (black) and C4 peptide (white) which is shown as ball and stick. Atom's colors are shown in red (O), blue (N) and magenta (P).

TPST2 HUMAN (060704)	MRLSVRRVLLAAGCALVLVLAVOLGOOVLECRAVLAGLRSPRGA	44
TDOTE HUMAN (OCOFOR)	MUCKI KONI LI ACI UT COUTUPVI COUAMECUUDI EED CODUKI	4.4
IPSII_HOMAN (060507)		44
D.melanogaster (AAM94031)	MRLPYRNKKVTLWVLFGIIVITMFLFKFTELRPTCLFKVDAANELSS	47
C.quinquefasciatus(EDS40555)	MRNRKIVLGVAGAAILLLMVFFKGTELGARCLANGGGSVG	40
C.elegans (077081)	MRKNRELLLVLFLVVFILFYFITARTADDPYYSNHREKFNG	41
L.sitkana (KFC35269)	MALCRQSTKRKLLLCGLAVTVLYFFYISAPCDQKSQY	37
C.gigas (EK914395)		
H. roretzi (AAM09087)	MGLGTAGTGGVKRTWWISLWLLTIVETVLVSSGVPKEGKMGTSFINRDIM	50
Interest (Harlester)		50
	1	
TDOTO ULIMAN	MODEOFFLUMUCTNI UVEVOVCKAMDI, I EVCCUDOSCTTI MDAMI DA	00
TPST2_HOMAN		01
TPSTI_HOMAN	ESTRITVRIGLDLKANKIFAYHKDMPLIFIGGVPRSGITLMRAMIDA	91
D.melanogaster	QMVRVEKYLTDDNQRVYSYNREMPLIFIGGV <b>PRSCTTLMRAMLDA</b>	92
C.quinquefasciatus	EMVRGETGSGGEDDGHGLGGSYKYHRNMPLIFIGGV <mark>PRSGTTLMRAMLDA</mark>	90
C.elegans	AAADDGDESLPFHQLTSVRSDDGYNRTSPFIFIGGVPRSGTTLMRAMLDA	91
L.sitkana	IMVPKEKNHFIYDSKNRAIPYGEDMEIIFVGGMPRSGTTLMRVMLDA	84
C.gigas	MPRSGTTLMRVMMDA	15
H.roretzi	MVKNSQEYQYNSVDQKPTLVKNDQDMPLIFIGGMPRSGTTLMRTMLDA	98
	*******	
	40 0 0 0 0	
TPST2 HUMAN	HPEVRCGEETRIIPRVLAMRQAWSKSGREKLRLDEACVTDEVLDAAMOAF	140
TPST1 HUMAN	HPDIRCCEPTRVIPRILALKOMWSRSSKEKTRUDFACVTDEWLDSAMOAE	141
D.melanogaster	HEDVRCCORTRVIPRIDOLRSHWIKSRKISLEDORACTTKRWMISATAOR	142
C guinguofoggiotug	UDDUDCCOPTIDUT DETLOT DE CUMMUSEURSUBLUEN OF TREET NONTACE	140
C. quinquelasciacus		140
c.eregans	NEE VICCEBETRVI PRIDNERSQWARSERBWNREDQQAGVIGEVINNAISSE	141
L.sitkana	HPDVRCCEETRVLPRILGLRTQWEKSAMEKKRLEAACVTGEVLDSAVRAF	134
C.gigas	HPEVRCGEETRVIPRILGMRNHWQKSEIERKRLVEAGINDAVIDSAVSAF	65
H.roretzi	HEDVRCCCETRVIPRLIFMRNNWIRSVKERRRLEEACILPKVLDQAVSQF	148
TPST2_HUMAN	ILEVIAKHGEPARVLCNKOPFTLKSSVYLSRLFPNSKFLLMVROGRASVH	190
TPST1_HUMAN	LLEIIVKHGEPAPYLCNKDPFALKSLTYLSRLFPNAKFLLMVRDGRASVH	191
D.melanogaster	CLEIIAKHGEPAPRLCNKDPLTLKMGSYVIELFPNAKFLFMVRDGRATVH	192
C.quinquefasciatus	CLEIIAKHGDPAPRLCNKDPLTLKMGSYVIELFPEAKFLFMVRDGRATVH	190
C.elegans	IMEIMVGHGDRAPRLCNKDPFTMKSAVYLKELFPNAKYLLMIRDGRATVN	191
L.sitkana	ILEVVAKHGEAAPRICNKDPFTIKSAIYLSHOFPRSRFIFMIRDGRAVVH	184
Caiass	TUETTAKHODAAPPLONKDEFTLKSSOVISSLEPNAKETIMTEDORAVTH	115
H roretzi	TLETTINHCSAAKELONKDEETIKATTYLHELEDNAKETIMIRDGRATAH	198
		200
	٥ ٥	
TPST2 HUMAN	SMITRKVTIAGFDLSSYRDCLTKWNKAIEVMYAQCMEVGKEKCLPVYYEQ	240
TPST1 HUMAN	SMISRKVTIAGEDLNSYRDCLTKWNRAIETMYNOCMEVGYKKOMLWHYPO	241
D.melanogaster	STISRKVTTTGEDLSSYROCMOKWNHAIEVMHEOCRDICKDROMMVYVDO	242
C. quinquefasciatus	STISRKVTTTGEDLSNYROCMTKWNON LOTMHDOCKETCKDROMMWYWRO	240
C elecane	STISPKVTTTCEDLNDEROCMTKANAALOTMUDOCESVCEKNOLKWYYPO	241
L aitkana	STITUKUTTI SCRDI VSVENCI OKUSSAMENINVSOCI PUCDSECMENIVERO	231
D.SICKANA		234
C.gigas	SVISRKVTISGFLIKNPKMCLEKWNTAMEVMYSQCLKVGPMRCMPVYYEQ	165
H. roretzi	SIISKKVIIAGIDIISYKDVIIKMNRALEVIHNOONEVGSKYOLSVHYEI	248
	^ <b>_</b>	
TRANS INIMAN		000
TPSTZ HUMAN	INTERARS EXECTED OF TAMED AVENING BUT OK POOV SUSK I BRSTDOVI	290
TPSTI_HUMAN	LVDHPERWMRTLLKFLQIPWNHSVLHHEEMIGKAGGVSLSKVPRSTDQVI	291
D.melanogaster	LVIHPEEWMRKILKELDVPWNDAVLHHEEFINKPNGVPLSKVERSSDQVI	292
C.quinquefasciatus	LVIHPEEWMRKILQELDIPWNDSVLHHEEFINKENGVALSKVERSSDQVI	290
C.elegans	LVLHPEAQMRRITEFLDIPWDDKVLHHEQLIGKDISLSNVERSSDQVV	289
L.sitkana	LAIHPREWMMKIVKFLDLPWNETVLHHEDFIGDKISLSRTEKSTDQVI	282
C.gigas	LALHPEVWMHRILEFLDIPWNNSVLHHEDFIGKPGGASLSKTEKSTDOVI	215
H.roretzi	LVLRPVQETKRIFTELDIPWSDHVLHHEKYLPDIQLSKLEKSTDQVQ	295
TOTO INNAN		240
TPST2_HUMAN	N-VNLBALSKWICHIPGDVVRDMAQIAPMLAQLGYDPYANPPNYGNPDPF	340
TPST1_HUMAN	KPVNVGALSKWVGKIPPDVLQDMAVIAPMLAKLGYDPYANPPNYGKPDPK	341
D.melanogaster	KPVNLEAMSKWVGQIPGDVVRDMADIAPMLSVLGYDPYANPEDYGKPDAW	342
C.quinquefasciatus	KPVNLEAMSKWVGHIPEDVVRDMADVAPMLSVLGYDPYANPPDYGKPDTY	340
C.elegans	KPVNLDALIKWVGTIPEDVVADMDSVAPMLRRLGYDPNANPPNYGKPDEL	339
L.sitkana	KPVNVEALSKWVGNIPDDVVKDMHTIAPMLMTLGYDPHANPPNYGKPDAQ	332
C.gigas	KPVNIEALSKWVGFYSDEIIADMAKIAPMLQTLGYDPDKNPPDYGKPDPK	265
H.roretzi	RELYLDALTSWECHI PEDVERDMAKIAPMLTKLCYDPSIORPSYCEPDEF	345

Fig. 4. Sequence alignment of human TPSTs with several species of invertebrates. Conserved residues are highlighted in black (strictly conserved) or in gray (homologous replacement). Residues in the putative transmembrane region are underlined. Symbols indicate PAPS-binding residues (\*), catalytic residues (↓), substrate-binding residues (\$), and residues which are involved in stabilization of enzyme molecule during catalysis (**V**). All residues are marked for the human TPST-2.

shown 66% of identity with uncharacterized protein of gastropod Lottia gigantea (GenBank B4BVQ4). As a result of this study a new source of TPST was found. We have determined the amino acid sequence of the TPST from marine gastropod L. sitkana. The studied enzyme has been identified as TPST-2, on the basis of the analysis of its primary structure. A high degree of homology between TPSTs from terrestrial vertebrates, insects and marine invertebrates indicates the importance of sulfation of tyrosine residues for the metabolism of invertebrates and vertebrates. The investigation of TPST from marine mollusks is interesting for understanding the functions of these enzymes and of their substrates - proteins and peptides. A recent survey of SwissProt data finds only 561 experimentally detected sulfated proteins (Khoury et al., 2011). Hence the studies of the metabolic importance of sulfation of tyrosines are at their very beginning. We hope our results could help to identify novel tyrosine-sulfated proteins and to understand the extent and the importance of the sulfation processes in marine organisms between others. Also, due to the high identity some structural features of LsTPST may be crucial for TPSTs of other types of shellfish.

## Acknowledgments

This work was supported by the grant of the Russian Foundation for Basic Research grant # 15-04-01004\_a. The authors express their gratitude to the Service of Science and Technology of the French Embassy in Moscow for the PhD in co-tutelle fellowship.

#### References

- Bilan, M.I., Kusaykin, M.I., Grachev, A.A., Tsvetkova, E.A., Zvyagintseva, T.N., Nifantiev, N.E., Usov, A.I., 2005. Effect of enzyme preparation from the marine mollusk Littorina kurila on fucoidan from the brown alga Fucus distichus. Biochemistry (Mosc) 70, 1321–1326.
- Coughtrie, M.W.H., Sharp, S., Maxwell, K., Innes, N.P., 1998. Biology and function of the reversible sulfation pathway catalyzed by human sulfotransferases and sulfatases. Chem. Biol. Interact. 109, 3–27.
- Favorov, V.V., Vozhova, E.I., Denisenko, V.A., Elyakova, L.A., 1979. A study of the reaction catalysed by alginate lyase VI from the sea mollusk, *Littorina* sp. Biochem. Biophys. Acta 569, 259–266.
- Han, S.W., Lee, S.W., Bahar, O., Schwessinger, B., Robinson, M.R., Shaw, J.B., Madsen, J.A., Brodbelt, J.S., Ronald, P.C., 2012. Tyrosine sulfation in a Gram-negative bacterium. Nat. Commun. 3, 1153.
- Hanai, H., Nakayama, D., Yang, H., Matsubayashi, Y., Hirota, Y., Sakagamin, Y., 2000. Existence of a plant tyrosylprotein sulfotransferase: novel plant enzyme catalyzing tyrosine O-sulfation of preprophytosulfokine variants in vitro. FEBS Lett. 470, 97–101.
- Kasinathan, C., Ramaprasad, P., Sundaram, P., 2005. Identification and characterisation of tyrosylprotein sulfotransferase from human saliva. Int. J. Biol. Sci. 1, 141–145. Kehoe, J.W., Bertozzi, C.R., 2000. Tyrosine sulfation: a modulator of extracellular protein–protein interactions. Chem. Biol. 7, 57–61.
- Khoury, G.A., Baliban, R.C., Floudas, C.A., 2011. Proteome-wide post-translational modification statistics: frequency analysis and curation of the SwissProt database. Sci. Rep. 1. http://dx.doi.org/10.1038/srep00090.

Kornprobst, J.-M., Sallenave, C., Barnathan, G., 1998. Sulfated compounds from marine organisms. Comp. Biochem. Physiol. B 119, 1–51.

- Kusaykin, M.I., Burtseva, Y.V., Svetasheva, T.G., Sova, V.V., Zvyagintseva, T.N., 2003. Distribution of O-glycosylhydrolases in marine invertebrates. Enzymes of the marine mollusk Littorina kurila that catalyze fucoidan transformation. Biochem. Mosc. 68, 317–324.
- Kusaykin, M.I., Pesentseva, M.S., Sil'chenko, A.S., Avilov, S.A., Sova, V.V., Zvyagintseva, T.N., Stonik, V.A., 2006. Aryl sulfatase of unusual specificity from the liver of marine mollusk *Littorina kurila*. Bioorg. Khim. 32, 63–70.
- Lee, R.W., Huttner, W.B., 1983. Tyrosine-O-sulfated proteins of PC12 pheochromocytoma cells and their sulfation by a tyrosylprotein sulfotransferase. J. Biol. Chem. 258, 11326–11334.
- Matz, M.V., Shagin, D.A., Bogdanova, O.V., Lukyanov, S.A., Diatchenko, L., Chenchik, A., 1999. Amplification of cDNA ends based on template-switching effect and stepout PCR. Nucleic Acids Res. 2, 1558–1560.

Moore, K.L., 2003. The biology and enzymology of protein tyrosine O-sulfation. J. Biol. Chem. 278, 24243-24246.

- Nedumpully-Govindan, P., Li, L., Alexov, E.G., Blenner, M.A., Ding, F., 2014. Structural and energetic determinants of tyrosylprotein sulfotransferase sulfation specificity. Bioinformatics 30, 2302–2309.
- Niehrs, C., Huttner, W.B., 1990. Purification and characterization of tyrosylprotein sulfotransferases. EMBO J. 9, 35-42.
- Nishimura, M., Naito, S., 2007. Tissue-specific mRNA expression profiles of human carbohydrates sulfotransferase and tyrosylprotein sulfotransferases. Biol. Pharm. Bull. 30, 821–825.
- Pesentseva, M.S., Kusaykin, M.I., Anastyuk, S.D., Sova, V.V., Zvyagintseva, T.N., 2008. Catalytic properties and mode of action of endo-(1→3)-β-D-glucanase and β-D-glucosidase I from marine mollusk *Littorina kurila*. Carbohydr. Res. 343, 2393–2400.
- Pesentseva, M.S., Kovalchuk, S.N., Anastyuk, S.D., Kusaykin, M.I., Sova, V.V., Rasskazov, V.A., Zvyagintseva, T.N., 2012. Endo-(1→3)-β-D-glucanase GI from marine mollusk Littorina sitkana: amino acid sequence and ESIMS/MS-estimated features of transglycosylation and hydrolysis reactions in comparison to analogous enzyme LIV from Pseudocardium sachalinensis. J. Mol. Catal. B Enzym. 75, 73–79.
- Sane, D.C., Baker, M.S., 1993. Human platelets possess tyrosylprotein sulfotransferase (TPST) activity. Thromb. Haemost. 69, 272–275.
- Seibert, C., Sakmar, T., 2008. Toward a framework for sulfoproteomics: synthesis and characterization of sulfotyrosine-containing peptides. Biopolymers 90, 459–477. Stone, M.J., Chuang, S., Hou, X., Shoham, M., Zhu, J.Z., 2009. Tyrosine sulfation: an increasingly recognized post-translational modification of secreted proteins. New Biotechnol. 25, 299–317.
- Teramoto, T., Fujikawa, Y., Kawaguchi, Y., Kurogi, K., Soejima, M., Adachi, R., Nakanishi, Y., Mishiro-Sato, E., Liu, M.C., Sakakibara, Y., Suiko, M., Kimura, M., Kakuta, Y., 2013. Crystal structure of human tyrosylprotein sulfotransferase-2 reveals the mechanism of protein tyrosine sulfation reaction. Nat. Commun. 4, 1–9.