



Molecular cloning and characterization of a tyrosine phosphatase from *Monosiga brevicollis*



Benjamin F. Zhao^{a,b}, Zhizhuang Joe Zhao^{a,*}

^a Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK, USA

^b Oklahoma School of Science and Mathematics, Oklahoma City, OK 73104, USA

ARTICLE INFO

Article history:

Received 2 October 2014

Available online 13 October 2014

Keywords:

Tyrosine phosphatase

SHP

SH2 domain

RNA splicing

Enzyme kinetics

Monosiga brevicollis

ABSTRACT

Protein tyrosine phosphorylation is thought to be a unique feature of multicellular animals. Interestingly, the genome of the unicellular protist *Monosiga brevicollis* reveals a surprisingly high number and diversity of protein tyrosine kinases, protein tyrosine phosphatases (PTPs), and phosphotyrosine-binding domains. Our study focuses on a hypothetical SH2 domain-containing PTP (SHP), which interestingly has a predicted structure that is distinct from SHPs found in animals. In this study, we isolated cDNA of the enzyme and discovered that its actual sequence was different from the predicted sequence as a result of non-consensus RNA splicing. Contrary to the predicted structure with one SH2 domain and a disrupted phosphatase domain, *Monosiga brevicollis* SHP (MbSHP) contains two SH2 domains and an intact PTP domain, closely resembling SHP enzymes found in animals. We further expressed the full-length and SH2 domain-truncated forms of the enzyme in *Escherichia coli* cells and characterized their enzymatic activities. The double-SH2 domain-truncated form of the enzyme effectively dephosphorylated a common PTP substrate with a specific activity among the highest in characterized PTPs, while the full-length and the N-terminal SH2 domain-truncated forms of the enzyme showed much lower activity with altered pH dependency and responses to ionic strength and common PTP inhibitors. This indicates that SH2 domains suppress the catalytic activity. SHP represents a highly conserved ancient PTP, and studying MbSHP should provide a better understanding about the evolution of tyrosine phosphorylation.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Protein tyrosine phosphorylation has a crucial role in cell proliferation, differentiation, and transformation [1]. This process is controlled by the coordinate action of protein tyrosine kinases and protein tyrosine phosphatases (PTPs). PTPs have diverse structural features and physiological functions [2,3]. They contain one or two conserved catalytic domains linked to various domains or segments. One subfamily of these enzymes contains two SH2 domains in the N-terminus followed by a single catalytic domain [4]. SH2 domains interact with specific tyrosine-phosphorylated structure motifs thereby playing a targeting or regulatory role. In the human genome there are two such enzymes, namely, SHP-1 (PTPN6) and SHP-2 (PTPN11). Both enzymes have been extensively studied, and they are known to play a crucial role in signal transduction and have major pathological implications [5–7].

Signaling through tyrosine phosphorylation has long been considered a hallmark of intercellular communication, unique to multicellular animals. However, genomic analyses of the unicellular choanoflagellate *Monosiga brevicollis* has revealed that the organism contains a remarkable count of 128 tyrosine kinases, 39 PTPs, and 123 phosphotyrosine-binding SH2 proteins, all higher than those seen in any metazoan [8–10]. Choanoflagellates are a group of small unicellular protists which consists of over one hundred species. They are phagotrophs and are thought to be the closest living relatives of animals as they resemble the feeding cells of sponges. The recent identification of intricate tyrosine phosphorylation signaling components in the genome of *M. brevicollis* provides further supports to this notion. *M. brevicollis* thus provide a unique system to study the evolution of tyrosine phosphorylation and related signaling circuitries. While many *M. brevicollis* PTPs should be isolated and characterized, one particular member of these PTPs, namely, MbSHP, attracted our attention. Unlike other SHP members, the predicted sequence of MbSHP contains a single SH2 domain followed by PTP domain with a 12 amino acid insert disrupting the so-called PTP signature motif (Genbank accession XM_001747931). In this study, we isolated MbSHP cDNA and

* Corresponding author at: Department of Pathology, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73104, USA.

E-mail address: joe-zhao@ouhsc.edu (Z.J. Zhao).

found the actual sequence differ from the prediction and revealed that there is non-consensus RNA splicing in the *MbSHP* gene. We further expressed the enzyme and characterized its PTP activity.

2. Materials and methods

2.1. Materials

M. brevicollis (stain Mx 1, ATCC® PRA-258™) was obtained from American Type Culture Collection (ATCC). Para-nitrophenol phosphate (pNPP), sodium vanadate, and hydrogen peroxide were from Sigma–Aldrich. A stock solution of pervanadate was made by incubating 0.1 M sodium vanadate and 0.2 M hydrogen peroxide at room temperature for 20 min before use [11].

2.2. Cell culture and isolation of DNA and RNA

M. brevicollis cells were obtained from ATCC with feeder bacterium *Flavobacterium* sp. included. The cells were cultured at 25 °C with Wards cereal grass medium made in filtered natural sea water (collected from the Gulf of Mexico). The culture was maintained at a density of 1–5 million cells/ml. The cells were collected by centrifugation at 2000g for 20 min at 4 °C after one hour sitting of culture suspensions on ice. Genomic DNAs were purified from the cell pellets by using the phenol/chloroform method after proteinase K digestion of whole cell lysates. Total RNAs were isolated by using the Trizol reagent (Invitrogen), and single-strand cDNAs were synthesized with random primers by using the QuantiTect reverse transcription kit from Qiagen.

2.3. PCR amplification and DNA sequencing

Initially, two PCR primers, namely, MbSHP5 (5'-GGTC AAATGTCCGGCCAGACAG) and MbSHP3C (5'-TGAAACGACACAG GCGAAAATCCAC), were synthesized to amplify the coding region of MbSHP based on the predicted cDNA sequence (Genbank accession XM_001747931). Following sequencing analyses of initial PCR products and *M. brevicollis* genomic sequence in the database, a new primer named mbSHP5f (5'-CGTGGAAATGACCTTAGTACTGACAC) derived from further 5'-upstream was synthesized and used to amplify the full-length form of MbSHP cDNA and the genomic MbSHP DNA together with primer MbSHP3. PCR was performed with high fidelity Phusion DNA polymerase with genomic DNA and single-strand cDNA as templates. PCR products were gel-purified and subjected to DNA sequencing analyses by using an ABI 3730 capillary sequencer at the core facility of University of Oklahoma Health Sciences Center. DNA sequencings were performed from both forward and reserve directions.

2.4. Molecular cloning and expression of MbSHP

For expression of the full-length and two SH2 domain-truncated forms of MbSHP in *Escherichia coli* cells, cDNA fragments corresponding to specific regions of MbSHP were amplified by PCR with PCR primers containing convenient restriction enzyme sites and cloned into the pGex-2T vectors. The sequences of inserts in the plasmids were verified by DNA sequencing. For expression of recombinant proteins, *E. coli* cells were induced by 50 µM isopropyl β-D-1-thiogalactopyranoside (IPTG) overnight at 25 °C. Cells were extracted by sonication, and cell extracts were purified by using glutathione-Sepharose 4B beads (GE Healthcare) according to standard protocols.

2.5. PTP activity assays

PTP activities were measured by using pNPP as a substrate at room temperature as previously described [12–14]. The assays were performed in a 50 mM acetate, MES, or Tris buffer (pH 4–9) supplemented with 1.0 mM EDTA and 2.0 mM dithiothreitol. The reactions usually ran for 10 min before termination by 0.2 M NaOH. Absorbance at 405 nm was measured with a spectrophotometer, and a molar extinction coefficient of 18,600 M⁻¹cm⁻¹ was used for calculation of concentrations. One unit of activity is defined as 1 nmol of phosphate released per minute.

3. Results

3.1. Molecular cloning of MbSHP and identification of non-consensus RNA splicing in the MbSHP gene

As described above, the genome of *M. brevicollis* contains a peculiar hypothetical PTP which has a single SH2 domain and a catalytic domain with an unusual insert disrupting the PTP signature motif; this is not seen in any other organisms. We thought it is important to isolate the enzyme and to characterize its enzymatic activity. For this purpose, we amplified the single-strand cDNA of *M. brevicollis* cells by using PCR primers covering the predicted coding sequencing of MbSHP cDNA. Interestingly, instead of the anticipated 1216 bp product encoding the predicted MbSHP protein, we obtained a single 1243 bp PCR product which codes for a polypeptide with a 21 amino acid insertion in the end of the SH2 domain and a deletion of the 12 amino acid insert in the PTP signature motif. This essentially disapproved the predicted MbSHP sequence as the experimental data revealed a translated protein product more similar to SHPs found in animals except that it has only one SH2 domain. We thus carefully reanalyzed the genomic DNA sequence in the database and found there were non-consensus RNA splicing sites (see below) and DNA segments encoding an additional SH2 domain in the *MbSHP* gene. We thus derived a new 5' primer from upstream sequence and used it to amplify the genomic DNA and single-strand cDNA together with the original 3' PCR primer. The complete sequences of resultant genomic and cDNA products are shown in Fig. 1. The PCR product amplified from genomic DNA was 3298 bp long, and its sequence matched that in the Genbank database except for a G-to-T substitution in the middle of intron 3. The PCR product from cDNA was 1610 bp with a sequence encoding a 513 amino acid protein containing two SH2 domains and a PTP domain. Alignment of the genomic and cDNA sequences revealed 16 exons in the *MbSHP* gene. All the exon/intron junctions follow the GT-AG rule except for two non-consensus GC splice donor sites which were not considered in the original prediction of MbSHP cDNA. One of the non-consensus splicing sites alters the sequence encoding the PTP signature motif, and other affects the C-terminal SH2 domain. We thus obtained a complete cDNA sequence of MbSHP in *M. brevicollis*. The sequenced has been submitted to the GenBank and was assigned an accession number of KM609288.

3.2. Sequence comparison of MbSHP with other SHPs

A BLAST search of protein databases demonstrated that MbSHP resembles the structure of SHPs found in all animals. It share 31%, 51%, and 52% overall sequence identities with SHPs found in *Caenorhabditis elegans*, *Drosophila melanogaster*, and zebra fish, respectively. When compared with human SHP-1 and SHP-2, it display 50% and 52% sequence identities, respectively. A detailed structural comparison MbSHP and human SHP-2 is shown in Fig. 2. The N-terminal SH2 domain appears to be mostly con-

CGTGGAAATGACCTTTAGGTACTGACACA**ATGGCTCGCAA**GTGAGTAGACCACGTTCTAGACGCACGGTATCGGCTCTCCGATGCCCCCTGCACCACCACC
 M A R K
 CGTAAACCTGACAGAGCCTTGTGCCATGCCCACGCGATAGATGGTTTTCATCCCAACATCTCGGGCGAGGAGGCCGAGCGAATGCTGCTACCCAGGGCTT
 W F H P N I S G E E A E R M L L T Q G F
 TGACGGCTCGTTCTCGTGCCGACAAAGCCGAGCAACAAGGAGACTATGCTCTCTGTCGGTGTGTCGGTGTCTTCTGTCCCTGAGTCGCCTGCCA
 D G S F L V R Q S R S N K G D Y A L S V R
 CTGTGCATCGTCACAACAGCTCAGCTCATCATGTATCTGGTGATCACCCGCGGCTCGGAATGGACCTCACCAAAATACCGTTCTCCTCCACACAGCCG
 R
 CGGTGCCTCGGTCACTCACATCAAGATTCAAAACAGTGGTGACTTTCTCGACTTGTATGGCGCGGAAAAGTTTGCCAACTTGACCGAACTGATTCAGTTT
 G A S V T H I K I Q N S G D F L D L Y G G E K F A N L T E L I Q F
 TACACGACCCACCCGGGCGAGGTGACACTTGAACCCACATCTCTGACTCCACCCACTTGGACTCGTCCCCCGCCCCCTCCTCTTTTTTTTTTTTTTTT
 Y T T H P G E
 TTTTTTG**GGCCTGCCACCATTGGCCTATTATCCTTTTGGCCTCTCAAATGCACTCTCATGCCGCGCCCTTTCCAACATGTTACGCTCAAGGAAGTGG**
 L K E V
 ATGGTGATGTTATTGATCTGCGACACCCGCTCCTCAGCGAGGACCCACCTCTGAGCGCTGGTTCATGGTCAAATGTCCGGCCAGACAGCGAGGCGGC
 D G D V I D L R H P L L S E D P T S E R W F H G Q M S G P D S E A A
 CCTGCGTGAGCGCGGAGCAGCGGAGCTTTCTCGTCCGCGCTCCAGTCCCAACCTGGTGGCTTTGCCCTCTCAGTCCGGTAGGACCCCTGCGCCTTG
 L R E R G E H G S F L V R A S Q S Q P G R F A L S V R
 AGCCCAACGTATCGCGAGCCCTCGGCTCTCTGCCCTGGATCTGGCCAGACTCATAGTCTCTGCGCTGCAACAGCTGTGGAGATGCCATCTCTCACA
 C G D A I S H
 TTATCATTCGCAACAACGGTGGCTCTTTGATGTGCGCGCGGTGCTCTGTTTACCGATCTCACCTCCCTCGGTATGCCCTGTGCGCTTTCTTTCTCTT
 I I I R N N G G S F D V G G G A S F T D L T S L
 TGACCCCGTCCGTCGCAAGCGCCTGCTAACGGCGGTCTCACGCTCTCTTTCAGTGGAAATATTACAAACGCAAGCCCTTGGTGAAGCAATGGCAACGT
 V E Y Y K R K P L V E A N G N V
 TGTGCACTTGCTCAAGGTGCGACTTCTTGACAGCCCCCGCTGCCTTCCACCAGGCTTTCCAACACTCACCGCCACACGCACTGTGCTTCTCCACCCCA
 V H L L K
 ATTTCAAACGACTAGCCCTTTAATGCCACCCGCTCAACCTGGCCAGCATTGATGGCGCTTTTATTGAACCTCAGCAAG**GCA**CAGCAAGCGTCAATCCCTCCC
 P F N A T R L N L A S I D G R F I E L S K
 CTGTCTTGGCTTGATATTGAAATCCCTGCGCTCATTCTGCCCTTTGTGCCCGCTGACTAACTCACGTTGTAATGCATCTGCAG**GAA**CCGAGGAGGAGT
 E T E E E
 TTGGCAAGGCTGGTTTCAATGAGGAGTTTGAGCAACTCCAGACGATGGAGCACCAACACAAGTTTGATCGCATGGAAGCGCTCGCATCGAGAACAAGTC
 F G K A G F N E E F E Q L Q T M E H Q H K F D R M E G A R I E N K S
 AAAGAATCGCTACAAGAACAATCTTGGCTTAGGTCCACCTTGCCCGCGCAACACAACCTGTATCTCTTCCATTGCTGCTTGTCTTGTCTTTTC
 K N R Y K N I L P
 TTATCAATTGTGCACTGCGCGTGCACAACCCGAGTTGAAGCGTTGTGGTTCCGTGAGCTCACGTGTGTACCTCCATTGTGCAG**AGC**ATCACACTCGCGT
 Y D H T R V
 CAAGCTCCGCAACGTTCTCTGATGGCGTGATTGGCGCGGACTACATCAACGCCAATTTTATCAACGGCGAAGCGCCAGGCACTGACAAGGCCTACATTGGC
 K L R N V P D G V I G A D Y I N A N F I N G E A P G T D K A Y I A
 TCGCAAGGCATATGCCCCACACGGTGGCTGACTTTTGGCAAATGATTGTCGAACAAACTGTATGCTTCTCAACAGCCTCTGCTCACCTTCTAACGCA
 S Q G T M P D T V A D F W Q M I V E Q N
 GACATGCAAGACTTGCTTGGCGCATTCTCTTGTCTACGCTCTATGTCTCGCTTCTATTTTGGCTTGTGGCGTGGACTGTGCACGGGAACCTCAGGCC
 C
 GTCTGGTTGTCATGGCCACTAATCCGGTTGAGAAGGGCAAGGTGTGTGGTTTCATGTCTTATGAGTGAAGGATGGCTTTTAGGGCCTCGTACA
 R L V V M A T N P V E K G K
 ACTCATCATAATTTACCTGCTTTGCAGCACAGTGCACGCTTATTGGCCTAGCAGCGAAGCGCGACCGGCGCTTTTGGCAATTATGTTGTGAGCTTC
 H K C T P Y W P S S E R R P A R F G N Y V V E L
 TTACGGAAGACAACCTTGTATGCGTACACGCTCCGTGCGATGACGCTTGGCTTGTAGCTGTACAGCTACACCGCGTGAGACAACACCAAGCGGCGA
 L T E D N F D A Y T V R R M S V A V
 CAATGAACCATGCGTTTGTCTTTCTCTCTTACCCGTTTATCGCTTCTTTTCTCTGTGCTCCTCATTGATACCTGCATTGACAACAGCGCAAGGC
 G E G
 GATCCTCGCGAGATTACCAATACCACTACACGTCCTGGCCTGACCACGGTACGACGCTAGGTGCCGCTCCTGTGATGATGGGGAGGCGATGCTCATGGG
 D P R E I Y Q Y H Y T S W P D H
 CTTTCCATCTGATATGAACAACGTTGGTAGCGCTGCCACCCGAGCGTGAATTTTTCGAGCTTATCATGAAAATTCGCCAGCATCGCGACCGTCTGCGTAAG
 G V P P P E R E F L Q L I M K I R Q H R D R L R K
 AGCATCCCCGGCTCGGCCCCAACGCTGGTTCAGTGCAG**GCA**AGTACATGCTAATCCGCGAGACCAATTTCAAAGGTGCAGGGAGCATGCTTAACAGTCG
 S I P G L G P T L V H C S
 TTGGTGCACCTCGGGCCTTGCTGGTATCGGTGCGACCGGTACTTTCTTGGTCATTGACATTGCTGCTGGACCTCATCCGCCACTACGGTCAGTTGT
 A G I G R T G T F L V I D I V L D L I R H Y
 CGCACCATCGTGAAGCGAGAAAAATCAAGAAGGCTCAACGCTTGAAGTCTATCTATTGTGCGGTGCTACGCTTAGTCTCTTTGCCCTCCATCATGTGGT
 CATTGACAGGCATTGACTGCGCATTGACATCCAAAGCATCATTACAAGTGGCGTGACAGCGCTCAGGCATGATCCAGACAGCGGTATGTTACTCTT
 G I D C D I D I Q S I I H K L R G Q R S G M I Q T A
 GACAAAATAATAGTAGAACAGAAAAAACAACAAAAACAATAACATGCTTCTTACCATGTGTGTGCACGGGTGTGTGCTAGGCACAAATA
 A Q Y
 TCGCTTCTATCTACAAGCCGTGATGGACTACATCCAATCCCAAAAAGCGCCCC**TAA**AGACGAGGGCCCAACGTGGATTTTCGCTGTGCTCGTTCA
 R F I Y K A V M D Y I Q S H K S R P *

Fig. 1. Genomic DNA sequence, exons, introns, and translated product of the *MbSHP* gene. DNA sequences were determined by Sanger sequencing of PCR products. The genomic sequence matches that in the Genbank database except for a G-to-T substitution indicated in a square. Exons (highlighted in gray) are assigned based on the cDNA sequence. All the introns follow the GT-AG rule except for the two non-consensus GC splice donor sites which are double-underlined and bold-faced. The translation and termination sites and the PTP signature motif are bold-faced. The DNA sequences corresponding to the 5' and 3' PCR primers are underlined.

served with a 67% identity, while the C-terminal SH2 domains are 57% identical. The PTP domains share a 52% sequence identity with the amino acid residues in the PTP signature motif region essentially unchanged. Note that the major sequence diversity is

found in the C-terminus. MbSHP has a short C-terminal segment missing the two conserved tyrosine phosphorylation sites which bind GRB2 in SHP-1 and SHP-2 [5]. Overall, SHP appears to be highly conserved during evolution.

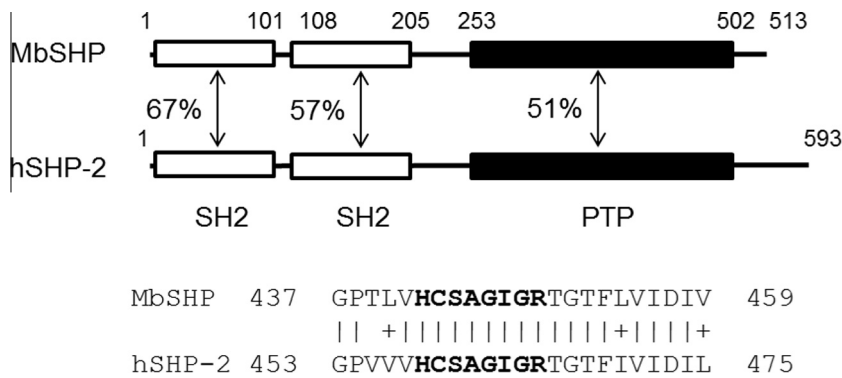


Fig. 2. Schematic structure comparison of MbSHP with human SHP-2. Two SH2 domains and the catalytic PTP domain are shown in blocks. Amino acid sequence identities and positions of domains are indicated. The amino acid sequences of MbSHP and human SHP-2 in the conserved PTP signature motif (bold-faced) are also compared.

3.3. Expression and PTP activity assays of full-length and SH2 domain-truncated forms of MbSHP

To analyze the PTP activity of MbSHP, we expressed the full-length and two SH2 domain-truncated enzymes as GST fusion proteins in *E. coli* cells. A schematic presentation of these recombinant enzymes is shown in Fig. 3A. The recombinant proteins were highly expressed in *E. coli* cells and were found in the soluble cell extracts. By passing the cell extracts through glutathione-Sepharose columns, the recombinant proteins were purified to near homogeneity with over 95% purity (Fig. 3B). PTP activities of the purified enzymes were determined by using a routinely used PTP substrate, pNPP. As shown in Fig. 3C, Δ2SH2-MbSHP exhibited robust pNPP dephosphorylation ability while ΔNSH2-MbSHP and full-length MbSHP showed a lower but clear activity. This indicates that MbSHP acts as a typical PTP.

3.4. Characterization of full-length and SH2 domain-truncated forms of MbSHP

As observed with other PTPs, the activity of MbSHP is affected by changes in pH (Fig. 4A). Full-length MbSHP has a clear preference for acidic pH with an optimum around 5.5. This was also found for both full-length and SH2 domain-truncated forms of SHP-1 and SHP-2 [12–14]. In contrast, the two SH2 domain-truncated forms of MbSHP displayed an optimal pH at 6.5. We further performed activity assays with different concentrations of pNPP under optimal pH. As shown in Fig. 4B, all three forms of MbSHP obey the Michaelis–Menten kinetics. Δ2SH2-MbSHP showed $k_m = 1.8$ mM and $V_{max} = 60$ nmole/min/μg. This V_{max} value corresponds to a turnover number of 60 times per second which is among the highest in characterized PTPs [12–15]. On the other hand, ΔNSH2-MbSHP and full-length MbSHP displayed much

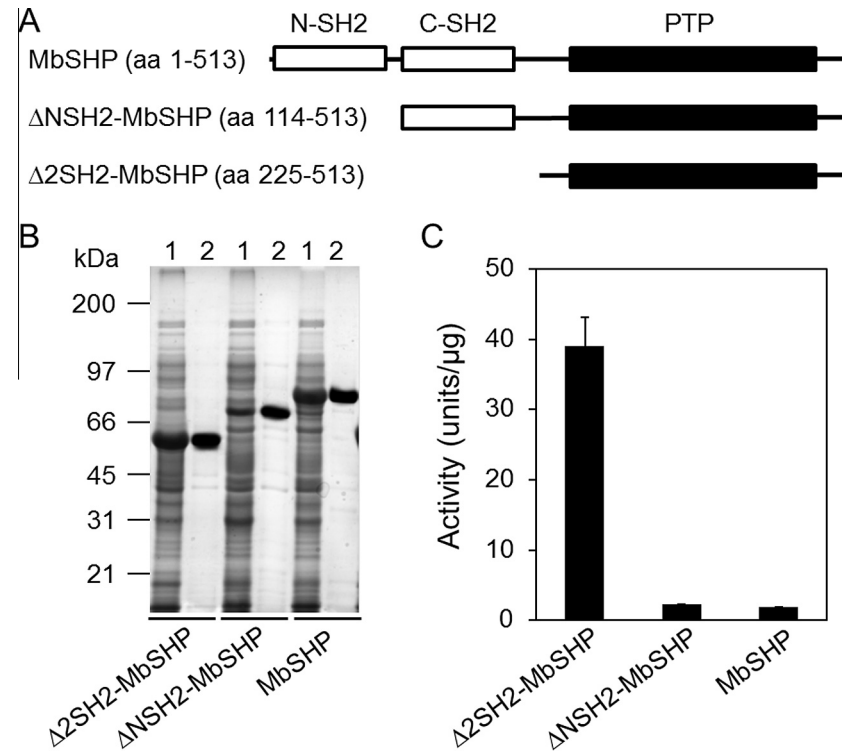


Fig. 3. Expression, purification, and PTP activity assays of truncated and full-length forms of MbSHP. (A) Schematic presentation of full-length and SH2 domain-truncated forms of MbSHP expressed as GST fusion proteins in *E. coli* cells. GST is attached to N-terminus, and amino acid residue numbers of MbSHP are indicated. (B) Separation of crude *E. coli* cell extracts (lane 1) and purified MbSHP proteins (lane 2) on a 10% SDS-polyacrylamide gel stained with Coomassie blue. (C) Comparison of the PTP activity of full-length and SH2 domain-truncated forms of MbSHP. Activity assays were performed with 10 mM pNPP as substrate in a buffer containing 50 mM MES-NaOH (pH 6.0), 1 mM EDTA, and 2 mM dithiothreitol. Error bars denote standard deviation ($n \geq 3$).

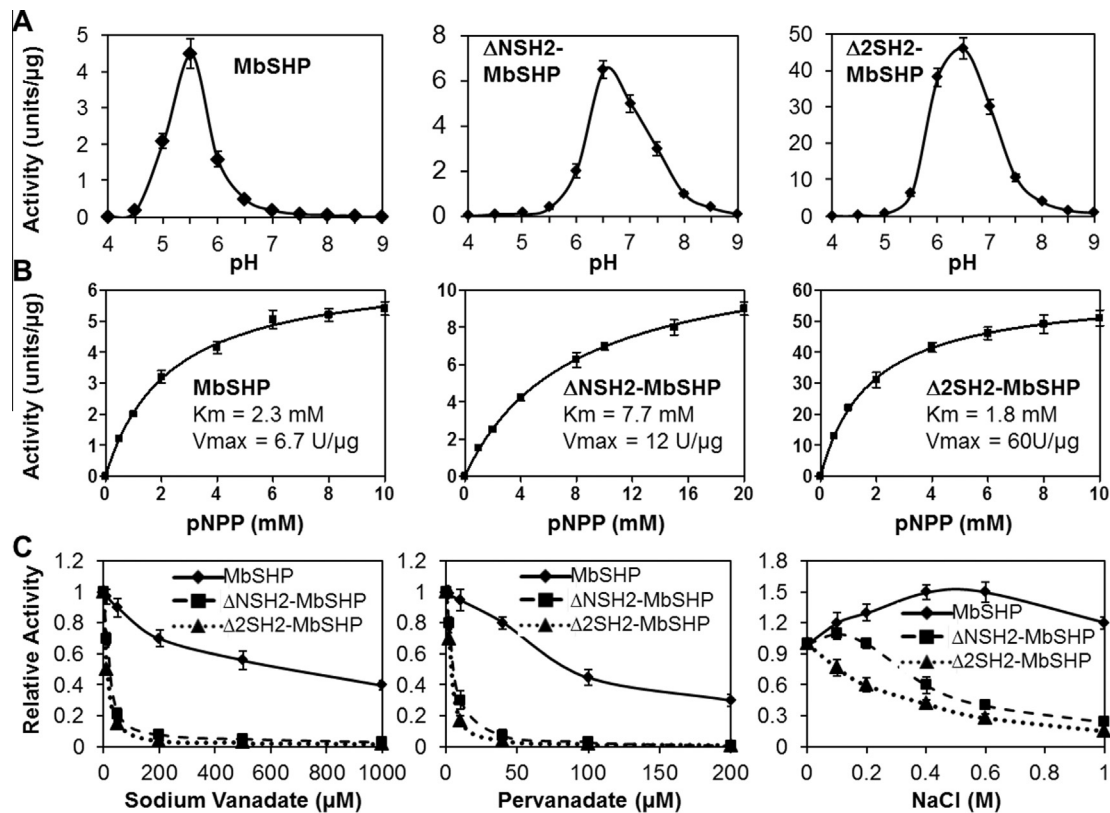


Fig. 4. Enzymatic characterization of full-length and SH2 domain-truncated forms of MbSHP. (A) pH dependency. PTP activities were measured by using 10 mM pNPP in a 50 mM Tris–acetate buffer at indicated pH. (B) Substrate saturation curves. PTP activity assays were performed with various concentrations of pNPP in 50 mM MES–NaOH buffer at optimal pH (pH 5.5 for MbSHP, pH 6.5 for ΔNSH2-MbSHP and Δ2SH2-MbSHP). Non-linear regression analyses with the Michaelis–Menten equation produced the curves and yielded the indicated k_m and V_{max} values which were consistent with the data obtained by linear regression analyses of Lineweaver–Burk plots. (C) Effectors of MbSHP. PTP activity was performed with 10 mM pNPP at optimal pH as described above. Error bars denote standard deviation ($n = 3$).

lower specific activity with k_m values of 7.7 and 2.3 mM and V_{max} values of 12 and 6.7 nmole/min/μg, respectively. The lower activity observed with full-length MbSHP in comparison with Δ2SH2-MbSHP is somewhat expected as the same phenomenon has been demonstrated with both human SHP-1 and SHP-2 [12–14]. In fact, crystal structures of these two human enzymes revealed that the N-terminal SH2 domain interacts with the catalytic domain and partly blocks the substrate entrance [16,17]. In this study, we also characterized N-terminal SH2 domain-truncated ΔNSH2-MbSHP which remained poorly active with a low V_{max} value and a high k_m value. How the C-terminal SH2 domain suppresses the catalytic activity is to be defined. Since it affected the pH dependency of the enzyme (see Fig. 4A), the C-terminal SH2 domain may suppress the catalytic activity in a different way from the N-terminal SH2 domain.

Like other PTPs, the phosphatase activity of MbSHP can be inhibited by common PTP inhibitor sodium vanadate and pervanadate (Fig. 4C). In fact, the two SH2 domain-truncated forms of MbSHP nearly lost all PTP activities in the presence of 200 μM vanadate or 40 μM pervanadate. Interestingly, however, the full-length enzyme was relatively resistant, with more than 30% remaining activity at 1 mM vanadate or 200 μM pervanadate. This may be attributed to the blockage of the catalytic center by the N-terminal SH2 domain. Similar phenomenon has been seen with human SHP1 and SHP2 [13,14]. The effects of ionic strength on MbSHP activity are also noteworthy. While human SHP-1 and SHP-2 (both full-length and SH2 domain-truncated forms) are inhibited by increases in ionic strength [13,14], MbSHP enzymes are resistant. Interestingly, the activity of full-length MbSHP initially increased with increases in ionic strength and remained fully

active at 1 M NaCl. This may be physiological relevant considering that *M. brevicollis* is a unicellular marine protist.

4. Discussion

By using experimental approaches, we identified the actual coding sequence of MbSHP. MbSHP resembles the structure of other SHPs without unusual structure features. Whole genome sequencing has become a common tool to study biological organism and systems. Computer-based analyses serve as essential tools to decipher the gene sequences but with limitations. The incorrect prediction of MbSHP coding sequence is largely due to the presence of two non-consensus RNA splicing within the MbSHP gene. While nearly 99% of introns have the canonical GT–AG junctions, there are exemptions with GC–AG junctions being the most common non-consensus splicing sites [18,19]. In the MbSHP gene, we found two such sites out of total of 15 introns. It is to be determined if this non-consensus splicing is more common in *M. brevicollis*.

Our study indicates that MbSHP is highly homologous to human SHP-1 and SHP-2. PTPs are highly diverse in their structures, and each species of animals contains a large number of them. However, only a handful of PTPs, such as SHP, have orthologs across the animal kingdom. In fact, SHP is one of the only two PTP orthologs that are shared by *C. elegans*, *Drosophila*, *Xenopus*, fish, and mammals. Apparently, SHP is an ancient PTP which is evolutionarily indispensable. Therefore, studying SHP may be particularly informative for understanding the evolution of protein tyrosine phosphorylation as a regulatory mechanism. Studies have demonstrated that SHPs play a crucial role in development. In *Drosophila* and *Xenopus*,

disruption of SHP function caused impairment of embryonic development [20,21]. In mice, knockout of the SHP-2 gene caused death of mouse embryos at mid-gestation [22]. SHP-2 appears to play a positive role in controlling the MAP kinase signaling pathway [23–25]. It is to be determined if MbSHP has a similar role in a unicellular protist like *M. brevicollis*.

Previous studies have demonstrated that the full-length form of SHP-1 and SHP-2 display very low specific activity *in vitro* compared with other PTPs due to an internal suppression [13,14]. In fact, X-ray crystal structures revealed an autoinhibition mechanism due to an interaction of the catalytic PTP domains with the N-terminal SH2 domain [16,17]. The present study demonstrates that MbSHP also displays that property. It is generally believed that SHPs are activated by binding of their SH2 domains to tyrosine-phosphorylated receptors and other proteins concomitant with translocation of the enzymes to their targets [5]. While activation of SHP-1 and SHP-2 is achieved by tyrosine phosphorylated signaling motifs including the so-called immunoreceptor tyrosine-based inhibition motifs [26], it is to be determined if such motifs exist in *M. brevicollis*. Studying the regulation of MbSHP should provide a better understanding about the origin of tyrosine phosphorylation signaling system.

Acknowledgments

This work was supported by grant HL079441 from the National Institutes of Health and a grant from Oklahoma Center for the Advancement of Science & Technology.

References

- [1] T. Hunter, The genesis of tyrosine phosphorylation, *Cold Spring Harb. Perspect. Biol.* 6 (2014) a020644.
- [2] A. Alonso, J. Sasin, N. Bottini, et al., Protein tyrosine phosphatases in the human genome, *Cell* 117 (2004) 699–711.
- [3] J.N. Andersen, P.G. Jansen, S.M. Echwald, et al., A genomic perspective on protein tyrosine phosphatases: gene structure, pseudogenes, and genetic disease linkage, *FASEB J.* 18 (2004) 8–30.
- [4] M. Adachi, E.H. Fischer, J. Ihle, et al., Mammalian SH2-containing protein tyrosine phosphatases, *Cell* 85 (1996) 15.
- [5] B.G. Neel, H. Gu, L. Pao, The “Shp”ing news: SH2 domain-containing tyrosine phosphatases in cell signaling, *Trends Biochem. Sci.* 28 (2003) 284–293.
- [6] A. Ostman, C. Hellberg, F.D. Bohmer, Protein-tyrosine phosphatases and cancer, *Nat. Rev. Cancer* 6 (2006) 307–320.
- [7] G. Chan, D. Kalaitzidis, B.G. Neel, The tyrosine phosphatase Shp2 (PTPN11) in cancer, *Cancer Metastasis Rev.* 27 (2008) 179–192.
- [8] N. King, M.J. Westbrook, S.L. Young, et al., The genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans, *Nature* 451 (2008) 783–788.
- [9] G. Manning, S.L. Young, W.T. Miller, et al., The protist, *Monosiga brevicollis*, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan, *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 9674–9679.
- [10] D. Pincus, I. Letunic, P. Bork, et al., Evolution of the phospho-tyrosine signaling machinery in premetazoan lineages, *Proc. Natl. Acad. Sci. U.S.A.* 105 (2008) 9680–9684.
- [11] Z. Zhao, Z. Tan, C.D. Diltz, et al., Activation of mitogen-activated protein (MAP) kinase pathway by pervanadate, a potent inhibitor of tyrosine phosphatases, *J. Biol. Chem.* 271 (1996) 22251–22255.
- [12] D. Pei, B.G. Neel, C.T. Walsh, Overexpression, purification, and characterization of SHPTP1, a Src homology 2-containing protein-tyrosine-phosphatase, *Proc. Natl. Acad. Sci. U.S.A.* 90 (1993) 1092–1096.
- [13] Z. Zhao, P. Bouchard, C.D. Diltz, et al., Purification and characterization of a protein tyrosine phosphatase containing SH2 domains, *J. Biol. Chem.* 268 (1993) 2816–2820.
- [14] Z. Zhao, R. Larocque, W.T. Ho, et al., Purification and characterization of PTP2C, a widely distributed protein tyrosine phosphatase containing two SH2 domains, *J. Biol. Chem.* 269 (1994) 8780–8785.
- [15] Z.Y. Zhang, Protein tyrosine phosphatases: structure and function, substrate specificity, and inhibitor development, *Annu. Rev. Pharmacol. Toxicol.* 42 (2002) 209–234.
- [16] P. Hof, S. Pluskey, S. Dhe-Paganon, et al., Crystal structure of the tyrosine phosphatase SHP-2, *Cell* 92 (1998) 441–450.
- [17] J. Yang, L. Liu, D. He, et al., Crystal structure of human protein-tyrosine phosphatase SHP-1, *J. Biol. Chem.* 278 (2003) 6516–6520.
- [18] M. Burset, I.A. Seledtsov, V.V. Solovyev, SpliceDB: database of canonical and non-canonical mammalian splice sites, *Nucleic Acids Res.* 29 (2001) 255–259.
- [19] T.A. Thanaraj, F. Clark, Human GC-AG alternative intron isoforms with weak donor sites show enhanced consensus at acceptor exon positions, *Nucleic Acids Res.* 29 (2001) 2581–2593.
- [20] L.A. Perkins, I. Larsen, N. Perrimon, Corkscrew encodes a putative protein tyrosine phosphatase that functions to transduce the terminal signal from the receptor tyrosine kinase torso, *Cell* 70 (1992) 225–236.
- [21] T.L. Tang, R.M. Freeman Jr., A.M. O'Reilly, et al., The SH2-containing protein-tyrosine phosphatase SH-PTP2 is required upstream of MAP kinase for early *Xenopus* development, *Cell* 80 (1995) 473–483.
- [22] T.M. Saxton, M. Henkemeyer, S. Gasca, et al., Abnormal mesoderm patterning in mouse embryos mutant for the SH2 tyrosine phosphatase Shp-2, *EMBO J.* 16 (1997) 2352–2364.
- [23] K. Yamauchi, K.L. Milarski, A.R. Saltiel, et al., Protein-tyrosine-phosphatase SHPTP2 is a required positive effector for insulin downstream signaling, *Proc. Natl. Acad. Sci. U.S.A.* 92 (1995) 664–668.
- [24] Z. Zhao, T. Tan, J.H. Wright, et al., Altered expression of protein tyrosine phosphatase 2C in 293 cells affects protein tyrosine phosphorylation and map kinase activation, *J. Biol. Chem.* 270 (1995) 11765–11769.
- [25] A.M. Bennett, S.F. Hausdorff, A.M. O'Reilly, et al., Multiple requirements for SHPTP2 in epidermal growth factor-mediated cell cycle progression, *Mol. Cell. Biol.* 16 (1996) 1189–1202.
- [26] E.O. Long, Regulation of immune responses through inhibitory receptors, *Annu. Rev. Immunol.* 17 (1997) 875–904.