

Supporting Information

Deng et al. 10.1073/pnas.1007883107

Fig. S1. Nucleotide sequence alignment of the C-terminal domain coding exon (exon6) of *Ld* sialic acid synthase (SAS)-A and *Ld*SAS-B and exon2 of three typical antifreeze protein (AFP)III genes, *Ld ψ* AFPIII, *Ld*AFPIII-1, and *Ld*AFPIII-2, from the *Lycodichthys dearborni* AFPIII locus. Among AFPIII genes, *Ld ψ* AFPIII shares the greatest sequence similarity with the *Ld*SAS genes, including having six additional nucleotides (shown in red) not present in the other two AFPIII genes, suggesting that *Ld ψ* AFPIII is the earliest member of the AFPIII gene family. Asterisks indicate identical nucleotides between the six sequences, and the asterisks in red indicate identical nucleotides shared between *Ld ψ* AFPIII and the *Ld*SAS genes. The stop codon (TGA for SAS and TAG for AFPIII) is underlined.

[Fig. S1 \(PDF\)](#)

Fig. S2. Sequence alignments of *Ld*SAS-A, *Ld*SAS-B, and *Ld ψ* AFPIII, the putative primitive AFPIII gene members from *L. dearborni*. (A) Nucleotide sequence alignment and deduced amino acid sequence of *Ld*SAS-A and *Ld*SAS-B. The proximal 5' flanking sequences (~600 nt) of the two genes share no sequence similarity and thus, are not homologous. The locations of the six exons are labeled, and exon sequences and translated amino acids are bold. Amino acids and codons of *Ld*SAS-B that differ from *Ld*SAS-A are shown in blue. (B) Nucleotide sequence alignment of the homologous regions between *Ld*SAS-B and *Ld ψ* AFPIII. The proximal 5' flanking regions of the two genes are homologous. The color highlighted *Ld*SAS-B sequence regions (in the ancestral *Ld*SAS-B) gave rise to the AFPIII gene. The signal peptide of AFPIII (exon1) was derived from partial 5' flanking sequences and the first six codons of exon1 of *Ld*SAS-B (yellow), AFPIII intron 1 from *Ld*SAS-B intron 5 (turquoise), and the mature AFPIII (exon2) from *Ld*SAS-B C-terminal domain (exon6; yellow). (C) Nucleotide sequence alignment of the homologous regions between *Ld*SAS-A and *Ld ψ* AFPIII. *Ld*SAS-A and *Ld ψ* AFPIII share lower nucleotide identities (66% id) than *Ld*SAS-B and *Ld ψ* AFPIII (70% identity) (Fig. 2). The proximal 5' flanking regions of the two genes are not homologous, and without this homologous region, *Ld*SAS-A could not have given rise to *Ld ψ* AFPIII. In all sequences, the stop codon is in red, nucleotides not shown are indicated by number of nucleotides, asterisks indicate identical nucleotides, and dashes indicate gaps. Sequences have been deposited in GenBank under accession numbers GQ368894 (*Ld ψ* AFPIII) and GQ368892 (*Ld*SAS-A and *Ld*SAS-B).

[Fig. S2 \(PDF\)](#)

Fig. S3. The synthetic E6 domain of *Ld*SAS-A as well as the full-length *Ld*SAS-A showed similar ice-binding activities as those of *Ld*SAS-B. Single ice-crystal growth morphology in (A) water and solutions of (B) *Ld*SAS-A E6 (12 mg/mL), (C) *Ld*SAS-B E6 (12 mg/mL), and (D) AFPIII (0.12 mg/mL) at their respective equilibrium freezing points. Single ice-crystal growth morphology in (E) water and solutions of (F) *Ld*SAS-A E6 (12 mg/mL) and (G) *Ld*SAS-B (12 mg/mL) at temperatures below the freezing points. (H) In contrast, in native AFPIII (0.12 mg/mL), the seed ice crystal did not grow until the temperature was lowered to -0.104°C . (I) Initial seed ice crystal at the equilibrium freezing point (fp) of the recombinantly expressed full-length *Ld*SAS-A solution. Ice growth is inhibited during 20 min of holding the ice crystal at the equilibrium fp (J) or small cooling ($\sim 0.004^{\circ}\text{C}$ to 0.074°C) below the equilibrium fp (K). When the nonequilibrium fp is reached, (L) ice expands rapidly as a hexagonal disk. The detectable ice-inhibition activity of *Ld*SAS-A is slightly greater than that of *Ld*SAS-B.

[Fig. S3 \(PDF\)](#)

Fig. S4. The expression and sialic acid synthetic activity of *Ld*SAS-A, *Ld*SAS-B, and *Ld*SAS-B mutant (*Ld*SAS-Bm4). (A) SDS/PAGE of the expressed enzymes after purification and dialysis. (B) Sialic acid synthetic activity of *Ld*SAS-A, *Ld*SAS-B, and *Ld*SAS-Bm4. The *Ld*SAS-Bm4 is a mutant form of the *Ld*SAS-B gene, of which four amino acids (V299, G304, V306, and T334) in the C-terminal domain were substituted with their homologs (Q9, N14, A16, and Q44) in the AFPIII ice-binding surface, respectively. Together with the T305, T308 preexisted in the *Ld*SAS-B, the mutant mimic of the evolutionary conversion of the SAS-B C-terminal to AFPIII in the active site. One unit (U) of sialic acid synthetic activity was defined as the amount of the enzyme that synthesizes 1 μmol sialic acid per 1 min. Sialic acid synthetic activity was assayed using a classical method (1). Briefly, 160 μL substrate solution (containing required amounts of ManNAc-6-P, 8.3 mM phosphoenolpyruvate (PEP), and 12.5 mM MgCl_2 in 50 mM bicine, pH 7.5) were added to 90 μL enzyme solution (1 mg/mL). The reactions were carried out at 37°C for 1 h and stopped by boiling for 3 min. After centrifugation for 3 min at $13,000 \times g$, the samples were analyzed using the thiobarbituric acid (TBA) assay (2). The enzyme activity was calculated using a molar absorption coefficient for sialic acid of $57,000 \text{ M}^{-1}\cdot\text{cm}^{-1}$ at 549 nm. Sialic acid synthetic activity was not detected for the mutant *Ld*SAS-Bm4 protein, even when enzyme concentration was increased to 3 mg/mL under the same reaction conditions.

[Fig. S4 \(PDF\)](#)

1. Reaves ML, Lopez LC, Daskalova SM (2008) Replacement of the antifreeze-like domain of human N-acetylneuraminic acid phosphate synthase with the mouse antifreeze-like domain impacts both N-acetylneuraminic acid 9-phosphate synthase and 2-keto-3-deoxy-D-glycero-D-galacto-nonulosonic acid 9-phosphate synthase activities. *BMB Rep* 41:72–78.
2. Warren L (1959) The thiobarbituric acid assay of sialic acids. *J Biol Chem* 234:1971–1975.

Fig. S8. Expression analysis showing that the precursor sequence for the AFPIII signal peptide found in the *SAS-B* gene possesses the same potential as the native AFPIII signal peptide in secreting AFPIII. (A) The expression construct containing *L. dearborni* AFPIII second exon without any signal peptide sequence but with an initiation codon [non-signal peptide (SP) construct]. (B) The expression construct containing *L. dearborni* AFPIII cDNA including the native signal peptide and the AFPIII second exon (pre-AFPIII construct). (C) The expression construct containing the AFPIII signal peptide precursor from the corresponding *LdSAS-B* gene and the AFPIII second exon (SP precursor construct). Each construct was linked to a flag tag at the C terminal. (D) Western blot analysis to detect the flag tag of the expressed protein exported to culture media from transfected HEK293T cells. Lane 1: HEK293T transfected with the empty pCS2-flag4 vector. Lane 2: HEK293T transfected with the non-SP construct (AFPIII second exon without any leading sequence). Lane 3: HEK293T transfected with the pre-AFPIII construct (the native AFPIII signal peptide and the AFPIII second exon). Lane 4: HEK293T transfected with the SP precursor construct (the SP precursor from *LdSAS-B* gene and AFPIII second exon). Confluent HEK293T cells cultured on six-well plates were transiently transfected with one of the constructs and incubated at 37 °C for 3 d. Culture media (about 1 mL) from each well were collected and centrifuged at 106 × *g* to remove any suspended cells; 12 μL spun culture medium from each transfection experiment were electrophoresed on a 15% SDS-PAGE gel and transblotted to a PVDF transfer membrane. Rabbit anti-flag antibody (Beyotime) and HRP-labeled goat anti-rabbit IgG secondary antibody were used to detect the presence or absence of secreted AFPIII (*D Top*). As controls, the same amount of culture media (12 μL) and the total proteins extracted from the transfected cells were electrophoresized and transblotted under the same conditions as above and probed with rabbit anti-human β-actin antibody (Beyotime) and HRP-labeled goat anti-rabbit IgG secondary antibody. It shows that the cytoplasmic β-actin was not detectable in any culture media and excludes the presence of intracellular protein in culture media because of cell lysis (*D Middle*). In contrast, similar amounts of cytoplasmic β-actin were detectable in the intracellular components of the four transfected cells, indicating that the antibodies used are valid (*D Bottom*). All of the photos were produced by exposing the blots to X-ray films (Kodak) for the same amount of time (1 min).

[Fig. S8 \(PDF\)](#)

Fig. S1

<i>LdSAS-A</i>	CTGGGTAAAGTCCTTGGTGGCCACGGTCAAGATCCCCAAAGGCACCGTCTGACTCAGGACATGTTGACGGTGAAGGTGGCCGAGCCGATGGGCGTCGCGGCCGAGGACATCTTCCAATG
<i>LdSAS-B</i>	CTGGGTAAATCTGTGGTGGCCAAGGTCAAGATCCCCAAAGGAACCGTCTGACTCAGGACATGTTGGCGGTGAAGGCCGCCGAGCCGATGGGTATCGCGGCCGAGGACTTGTGCAAAATG
<i>LdPAFP111</i>	-----GCGTCCGTGGTGGCCAACCAAGCTGATCCCATAAACTAGCCCTGACTCTGATAATGATGAAGGCGAAAAGTGGTACCCCCAATGGGCATCCCCGCCGAGGACATTCCCCGAATA
<i>LdAFP111-1</i>	-----GCGTCCGTGGTGGCCAACCAAGCTGATCCCATAAACTAGCCCTGACTCTGATAATGATGAAGGCGGAGGTGGTACCCCCAATGGGCATCCCCGCCGAGGACATTCCCCGAATA
<i>LdAFP111-2</i>	-----GCGTCCGTGGTGGCCAACCAAGCTGATCCCATAAACTAGCCCTGACTCTGATAATGATGAAGGCGGAGGTGGTACCCCCAATGGGCATCCCCGCCGAGGACATTCCCCGAATA

** *****
 ***** * ** * ***** * *** ** * * * * * ** ***** * * ***

<i>L</i> αSAS-A	GTGGGAAAGACCGTGACGAAGGACGTGGAGGAGGACGCGAGCCTCTTGCCAGAGGTGGTGACGGGTACT <u>TGCAAG</u> AAGAGGAAGTGCTGA
<i>L</i> αSAS-B	GTGGGAAAGACCGTGACGGAGGACGTGGAGGAGGACGACAGCGTCATGCCAGAGATGGTGAAAGGGTACT <u>TGCAAG</u> AACAAGAAGTGCTGA
<i>L</i> αPAFPIII	ATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCTCATGCCAGACATGGTGAAAACGTAC <u>CGCGTG</u> GCGAAGT <u>AGT</u> TCTGA
<i>L</i> αAFPIII-1	ATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCTCATGCCAGACATGGTGAAAACTAT-----GAGAAGT <u>AGT</u> TCTGA
<i>L</i> αAFPIII-2	ATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCTCATGCCAGACATGGTGAAAACTAT-----GAGAAGT <u>AGT</u> TCTGA

* * * * *

Fig. S2

A

```

LdSAS-A  -----gaacaaggaaccaggaagctcattagagtgggtgggtagctcttaaaagagccgttgtgtgagcgacacctcaactctgccagaccatgaaagcatccaagcaccgccgtagct
LdSAS-B  -----tgtcattgtgctatttgatagaggagaccatgccgatctagacagctgatattatgaaacagcccaacaataacagcaagcaaatatgaatgtgcatgctgaggagtaa

LdSAS-A  ctggcgaattttgcagaaatgttgtgtgtgatgtggacgggcccgtatcaaagagccaaaacatgaaattgaaatcgatgccatcagccagcttctccggcgggagaggggtgaggcga
LdSAS-B  ttggaagatgtatcagaaataaactacaggcatgcaagagtatagtcttctgactggactcttattttgtcttttcttatatcattttgggatgggcgatgggagagttttccacctg

LdSAS-A  cggcggatgccttcttgatactttcatgggtccagtttagcttgaagtgtctctcgaaggggatgaatcttgtatcgctccgtcgctgactagatagtttcttatgctaagtactgt
LdSAS-B  tctcagcttgtttttacccaaaatattgtatactattagaaccgttgcacaggggttcaaatgaatgttttcgttttagtttgatcgtgtaagtgttgtgttcgacagcaagaagca

LdSAS-A  gctgtcgctcgacaggaattggctgacggcagggggcggttctaacagtacatgagaaccgattttatcgtctagtttcataagtcaataaataaaccgtctatatgaaaaactacga
LdSAS-B  gaggatctgtgtgcaggaagtcagctgctgcatggatcacaagttatacgtgcaagcttgtgatagtttgcaaaaaacaagttataaaacagtttacttgaagaatataaaatttcc

LdSAS-A  gttaaagcctatttttacacttctgtgacttaataactaaagtactaaatattttctcatttaattaagaacttttgtgtctgtaaaggaaatccggatattacgtcatactacaagcgg
LdSAS-B  actacaattggcataaggaggtggtggcagtggtgacggtgtcacagtgctctaccactttcagaccaaacgggcccgttgaacaggatatactggttgaacaagagtggctgctcctgtt

LdSAS-A  aagtcctccacatcc--tggtatgtgcagcacattaattgggtccggtcagattcaagtcctcagctacggcttttcgcttcgt----gctaaataattc---gtgctaaataattc--ta
LdSAS-B  aagtcctaccgcataatattagatgtgcagcacatgaactcatcctttcagattcaagtcctcagctacggcttttcgcttccttcgccgtaataattccggtattttcaccagtttgttt
***** ** *** * ***** ** * *** ***** ** * ***** ** * ***** ** * ***** ** * ***** ** * ***** ** *

      └─►Exon 1
      M P L T F E L C P G R M I G G D H P C F I I A E I G Q N H Q G D
LdSAS-A  aatagtttgttttctgtggacaaaATGCCGTTAACGTTTCGAGCTGTGTCCCGGTAGAATGATCGGAGGAGATCATCCGTGCTTCATCATCGCTGAAATTGGACAAAATCACCAGGGAGAC
LdSAS-B  tccgttttcttttctgcagacaaaATGCCGTTAACGTTTCGAGCTGTGTCCCGGTAGAATGATCGGAGGAGATCATCCGTGCTTCATCATTGCTGAAATTGGACAAAATCACCAGGGAGAC
      *** ***** *****

      E I A K K M I K M A K
LdSAS-A  ATTGAGATTGCCAAGAAAATGATCAAAATGGCAAAGgtataaaatcgacacgtatcgctttgtattgaatgtgaatacgtatatatttctgat-----
LdSAS-B  ATTGAGATTGCCAAGAAAATGATCAAAATGGCAAAGgtataaaatcgacac-----tttatattgaatgtgaatacgtatatatttctgatgttcttatcttaacagccaagacaca
***** *****

LdSAS-A  ---acgagcactgacgtaatcagaactaaattaaaaatgtaccattgaatatgtgtgcgcttatcttcc---tctt-----ctgaatcattactt-----
LdSAS-B  cagacaataactaacttattcagaactcagtttaaaatgtgcagattgaatatgtgtgcgctcatcgctcgaatcattactttaataatactataaactaaactagtgtgtgtatgt
      ** * *** ** * ***** * ** ***** * ***** ***** ** * ** * ** * ** * ** * ** * ** * ** * ** *

LdSAS-A  -----
LdSAS-B  agacctttatgctttaattaagtattataatgtgcaatgttgcataagagcaagtggttccgaaggggtgtttgtataaagacgtgcactaggatgttaaatgaccggatagtga
      **

LdSAS-A  aatgaagtttggcttgtgttagttcgtggccatgttgggtaact--atctgtactgaagccaaagcacacgcccgttttacgtgtattcagtttaattaaatacctcatagacataaaga
LdSAS-B  aatgaagtttggcttgtgtttgttcgtggccatgttgggtaactgtatctgtactgaagccaaagcacacgcccgttttatgtgtattcagttt---aaatacctcatagacataaaga
***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** *****

LdSAS-A  agcagcaatgggtgagaaactttaaatgcacttaacatataaaactaaaaagacaactt-----cctttaaacctt---tagagccttgttttgcag
LdSAS-B  agcagcaatgggtgagaaactttaaatgcactttacaaa-aaaaataaaaacacaacttttagtgaaacacttcagtggtgctcctttaaaccttcagtagagccttgttttgcag
***** ***** ** * ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** ***** *****

LdSAS-A  gtgtatatgttaattatttctgaaaagaactgtct--tttgggtccagttacaggg-aacatagacaactttt-gagttatttaagttataagcttttgttaatttctagaggaattt
LdSAS-B  gtgtatatgttaattatttctgaaaataagtggtatttttgggtccagttaccagtaaaaatagacaactttttgagttatttcagtgatacatttttgttaatttctagaataattt

```

***** ** ** * ***** * ** ***** ***** ** *

LdSAS-A ccttcagaagaactacatcaaaccacagtattaaacatatcttctatctatttattaatgggaacttcattccccatatatgttgcttattatgccgcttttagatcccattaattcatg
LdSAS-B cctttagaataactaaattaaaccacagta-aaaaaaatcttgatctatttattaatgtgaacttcattccccatatatgttgcttattatgccgcttttagatcccattaattcatg
*** **

LdSAS-A atcagctcacctgggtgaaaaaac--tatttgaactactagaaatagatttgaatagggtataat-gaaataattaatggtaagtgtagtaattgaa-----cacgggtctcctctt--aa
LdSAS-B atcagctcacctgggtgaaaaaacaactatttgaactactaggaatatatttgaatagggtataatggaaaccgacttggataggggcattgaaggagacgactcaggggatgcttttcag
***** **

LdSAS-A tagta-----tgaatatagtcctttctaggaac-cttactgtgttttttag-cacaatctgtag---gagcccttaa-atcaatcatcactacttagtatgttttccaa...43nt..
LdSAS-B catttatttgctgcagaagacaataacaacacacacatatattaacttctgggtctcctctgcactcctgctccctcaacgatacaaaacaccatttgcatacgaacaa..233nt..
* * * * *

→ Exon 2

LdSAS-A D Q/CG A D C A K F Q K S E L E H/C K F N K K A L E R E Y N S V H S W G K T Y G D H
LdSAS-B GACCAGGGTGGGATTGTGCCAAATTCAGAAGAGTGAATTGGAGCATAAATTCACAAGAAAGCTTTGGAGCGGAGTACAACCTCTGTTCACTCCTGGGGAAAAACCTATGGTGACCAC
GACTGCGGTGCTGATTGCGCAAATTCAGAAGAGTGAATTAGAGTGTAAATTCACAAGAAAGCCTTGGAGCGGAGTACAACCTCTGTTCACTCCTGGGGAAAAACCTATGGTGACCAC
*** **

LdSAS-A K R H L E F S H D Q Y R E L K/QK Y A E/AE V G I C/FF T A S G M D E
LdSAS-B AAGCGCCATCTGGAGTTGAGCCATGATCAGTACAGGGAAGCTCAAGAAATATGCAGAGGAGGTGGGGATCTGCTTCACTGCCTCTGGGATGGATGAGgttaagaaagagacgtctgaagaag
AAGCGCCATCTGGAGTTGAGCCATGATCAGTACAGGGAAGCTCAGAAATATGCTGCGGAGGTGGGGATCTTCTTCACTGCCTCTGGGATGGATGAGgttaagaaagagacgtctgaagaag

→ Exon 3

LdSAS-A gctttctgtgggttttaatacaggttgaagtttaaacgtgt--ttctttctgtcttttagATGGCAGTGGAGTTCCCTCCATGAGCTTAATGTGCCCTTCTTGAAAGTGGGCTCAGGAGACA
LdSAS-B gctttctgtgggttttaatacaggttgaagtttaaacatagtttctttctgtcttttagATGGCAGTGGAGTTCCCTCCATGAGCTTAATGTGCCCTTCTTTAAAGTGGCCTCAGTAGACG

LdSAS-A /AN N F P/SY L/ME K T A K K G
LdSAS-B CCAACAACCTCCCTATCTGGAGAAGACCGCCAAGAAAGgtgagagtgaactgcagcagatgtattttccgttgcatatttaagtgtctaaaatcatcagtgactttgttgggacttcac
CCAACAACCTTCTCCTATATGGAGAAGACCGCCAAGAAAGgtgagagtgaactgcagcagatgtattttatgttgcatatttaaatgctaaaatcatcagtgactt-----cac

LdSAS-A catgtggttgaacaaaggaggttgaagttgggcctacatggttcaacagccccaccagacgttaagttctcacaacttaacactgggggtctcattggacgagggccacaggcgccaca
LdSAS-B cattacat--aaataaatactatattaagta-----aaacaataaaaacagaaatagaagcagaagaagtgcacaaatgtgtccgtagaa---attagttaagag--acgaaaaataaa
** * **

LdSAS-A gggggctccctgtgggtgcgactgagactataaaaatctgcattatctccagctcggagccatcccacgcaattcgattcgcatgaaagg-----1009nt-----
LdSAS-B tgaagtttcgcttgatt--gttga-actacaggag-c---tcattaagaggagagagaagggttgcattttcccttttgccc--at-g-----45nt-----
* * * * *

→ Exon 4

LdSAS-A R P M V V S S G M Q S M E T M R R V Y K T/IV K K/EH N P N F T I L Q C T S A/SY P L
LdSAS-B GACGGCCCATGGTGGTGTCCAGCGGGATGCAGAGCATGGAGACGATGCGTCGGGTCTACAAAACGGTGAAGAAGCACAACCCCAACTTCACCATCCTGCAGTGCACCAAGCCTACCTC
GACGGCCCATGGTGGTGTCCAGCGGGATGCAGAGCATGGAGACGATGCGTCGGGTCTACAAAATAGTGAAGGAGCACAACCCCAACTTCACCATCCTGCAGTGCACCAAGCTCCTACCTG

LdSAS-A /VE/DA/PE/KD/NV N L R/KV I T/S
LdSAS-B TGGAAAGCTGAAGATGTCAACCTCAGAGTGATAACGgtgagaagaacagagcattactcctttatcaaaagggttggtacagaacaagatctcaacaagttgagttgggaaattgtctca
TGGATCCTAAAAATGTCAACCTCAAGTGATATCGgtgagaaaagtaggattttaaactcctgagcgtctcact--taaagtaatagtttgacatttcag--gagttggtggattttgtta

ctccattgaaagtagacgttttcagcagatgcatattgagttttaaaagccggttaagatatcacatgtcttattgtcaatttgtgttttagcatttattccacaaattaaacttacaat
ct--gttgggtgagccaggct-agctgttccccctggtttcca--gtcattatgctaagctaagctaaccggcagctggc---tctagcttg-----aaatttaactcacaaa
** ** *

Exon 5

aattaagatgaatcttcctgttttgatggccgtgtagaattgacact---488nt---GAATACCAGAAGGAATTTCCCGATATTTCCCATCGGGTACTCCGGCCATGAGTCTGGGATC
tgtgagtgagg-tgtcaatcgctcgagag-att-tagagtaacact---20nt---GAATACCAGAAGGAATTTCCCGATATTTCCCATCGGGTACTCCGGCCATGAGACTGGGATC
* * * * *

S I S V/IA A V/AA/VM G A K V I E R H V T L D K T W K G N D H E/SA S L E P S/GE L A/KE
AGTATTTTCAGTTGCGGCCGTGGCTATGGGTGCAAAGGTCATCGAGCGCCACGTAACCTGGACAAGACGTGGAAAGGAAATGACCACGAAGCCTCTCTGGAGCCCGAGTACGAGTACCCGAG
AGTATTTTCAGTTGCGGCCGCGAGTTATGGGTGCAAAGGTCATCGAGCGCCACGTAACCTGGACAAGACGTGGAAAGGAAATGACCACAGCGCCTCTCTGGAGCCCGAGTACGAGTACCCGAG

L V R S/CI R L/EV E R/MA Q G T/SG/CI K Q M L P C E K P/AC H D/AK
CTGGTTCGTTCCATCCGACTGGTGGAGAGGGCGCAGGGCACCGGCATCAAGCAGATGTTACCTTGTGAGAAGCCGTGCCATGATAAGgtgagatattctgtctcactagggaaacttaac
CTGGTTCGTTGTCATCCGAGAGGTGGAGATGCGCAGGGGAGCTGCATCAAGCAGATGTTACCTTGTGAGAAGGCGATGCCATGCTAAGgtgagatattctgtctccacaaaaacgtttttt

tt---gtactgttcagaaagacagaggaccttt-----taagaaaaca-----ttttgaaattgttttcaactcccaaccacgagaaacaataataatttcagacctttttattttct
tttaaaatgcgagccacagttaaattacatttttactgtttagaagacagataaccttaagaaaacatttctactgt--gccatgagaaacaataataatttcagacctttttattttct
** * * * * *

Exon 6

attctccaagCTGGGTAAGTCTTGGTGGCCACGGTCAAGATCCCCAAAGGCACCGTCTGACTCAGGACATGTTGACGGTGAAGGTGGCCGAGCCGATGGGCGTCGCGGCCGAGGACAT
attctccaagCTGGGTAAGTCTGTTGGTGGCCAGGTCAAGATCCCCAAAGGCACCGTCTGACTCAGGACATGTTGCGGGTGAAGGCAGCCGATGGGTATCGCGGCCGAGGACAT

F/CQ/KM V G K T V T K/ED V E E D G/DS L/VL/MP E V/MV D/KG Y C K K/NR/KK C Ter
CTTCCAAATGGTGGGAAAGACCGTGACGAAGGACGTGGAGGAGGACGGCAGCCTCTTGCCAGAGGTGGTGGACGGGTACTGCAAGAAGAGGAAGTGCAGgggtcgcaaggaggttcatca
GTGCAAAATGGTGGGAAAGACCGTGACGGAGGACGTGGAGGAGGACGACAGCGTCATGCCAGAGATGGTGAAAGGGTACTGCAAGAACAGGAAGTGCAGgagatggcaaggaggttataa
* * * * *

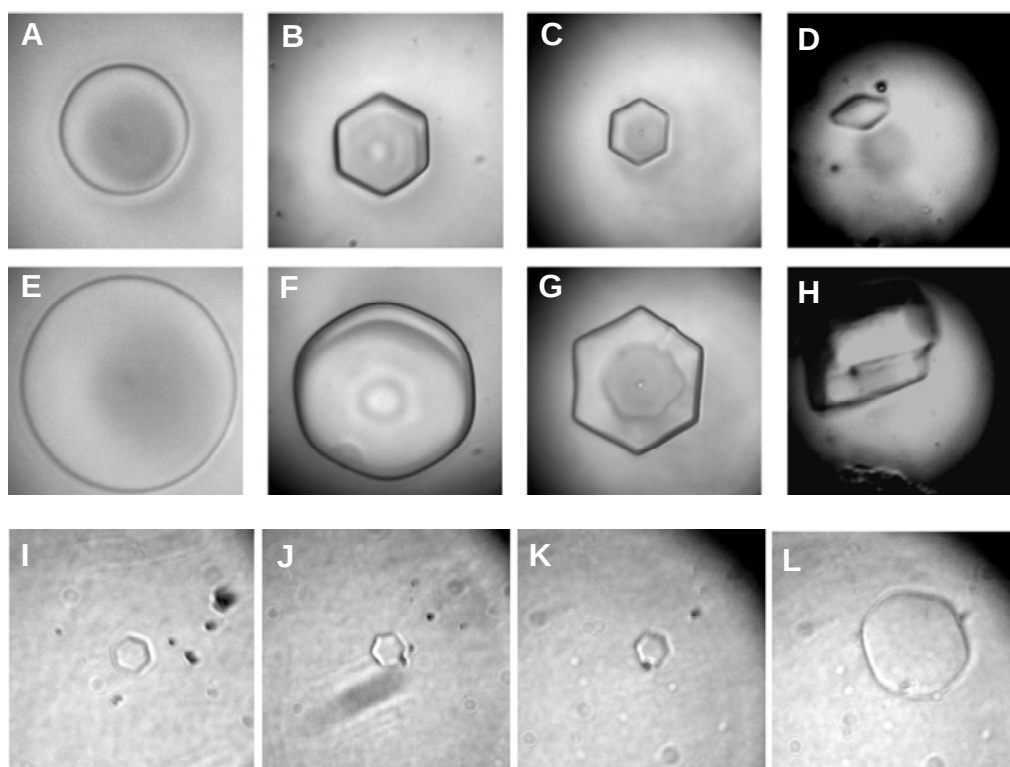
atgggtataaaaggcacatcatttgagaggaaacgtgatcgattcctctggtcattttgtttggtggtgatgggattgtcctggacgaggagtaattggttcgtgcattgctctgatagt
atgggtagaaaaggcacatcatttgagaggaaatgt-----t-----ctggcgat-a-----ttg-----

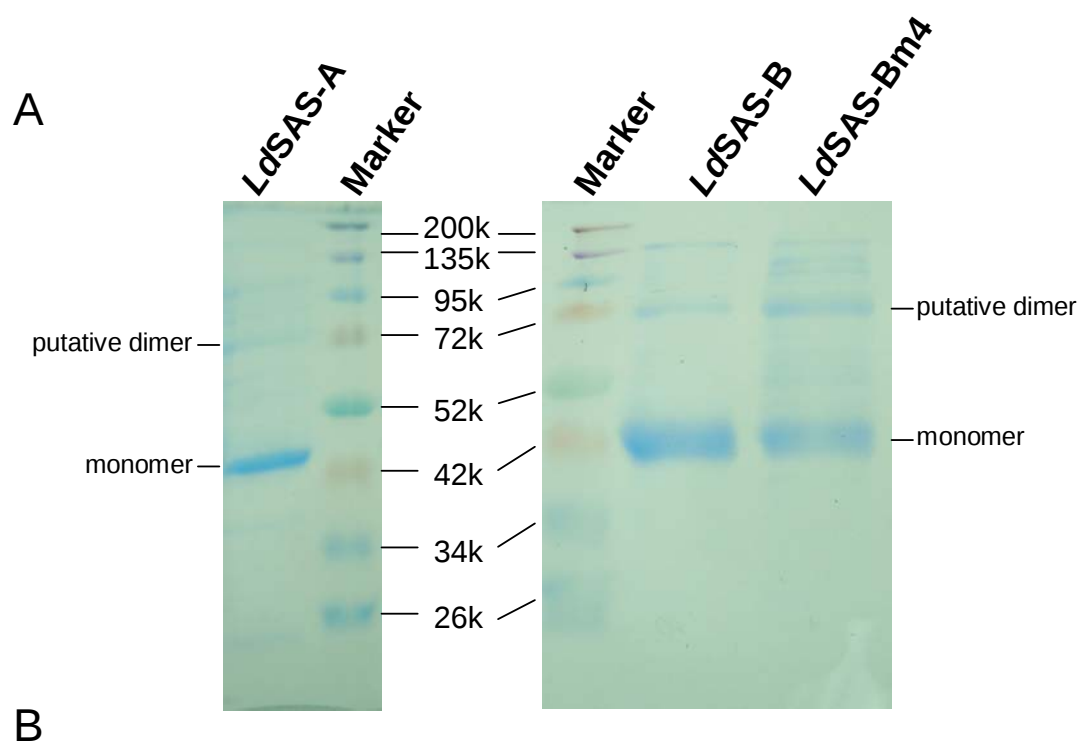
ttcaatatatcacagctgggggttcaaacggttactttgtcgaagtataaaccaccataatcttacaagaggctcagtagtagtctcagtcctgtggtcagtagtagtgcagtgctttc
-----g-tcattttg-tgt-----c-----tcttacaagaggctcagtagtagtctcagtcctgtggtcagtagtagtgcagtgctttc
* * * * *

tgagggtggcacaaaagccagttgttgttactgtgaacttaacatactttgcactatagttttgcctgctctgtagagattaaaggaacattgaacacacgaagcggtcactggtcatct
tgagggtggcacaaaagccagttgttgttactgtgaacttaacatactttgcactatagttttgcctgctctgtagagattaaaggaacattgaacacacgaagcggtcactggtcatct

		→ <i>LdPAFPIII</i> Exon 2
<i>LdPAFPIII</i>	ttatatttccattcttcaag	GCG-----TCCGTGGTGGCCAACCAAGCTGATCCCCATAAATACTGCCCTGACTCTGATAATGATGAAGGCGAAAGTGGTCACCCCAATGGGCATCCCCGC
<i>LdSAS-A</i>	ttatatttctattctccaag	CTGGGTAAGTCCTTGGTGCCACGGTCAAGATCCCCAAAGGCACCGTCCTGACTCAGGACATGTTGACGGTGAAGGTGGCCGAGCCGATGGGCGTCGCGGC
	*****	* * * * *
		→ <i>LdSAS-A</i> Exon 6
<i>LdPAFPIII</i>	CGAGGACATCCCCGAATAATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCCTCATGCCAGACATGGTGAAAACGTACCGCGTGCCGAAG	AGttctgagcgtgccaagg
<i>LdSAS-A</i>	CGAGGACATCTTCCAAATGGTGGGAAAGACCGTGACGAAGGACGTGGAGGAGGACGGCAGCCTCTTGCCAGAGGTGGTGGACGGGTACTGCAAGAAGAGGAAGTGC	TAggggtcgcaagg
	*****	* * * * *
<i>LdPAFPIII</i>	agttt	cttcccaaaaccaaagaagaatgcctcctctcacaattaaccttggtttttgtcacaacccaagtcctgctccggatgttaactgaacatgtcaaaacctgtggcgactatgttg
<i>LdSAS-A</i>	agttc..518nt..	cttcccaaaaccaaagaagaatgccccctctcacaattaaccttggtttttgtcatgaaccaaaccgtccggatg-----aaacctgtggagcct-----
	****	* * * * *
<i>LdPAFPIII</i>	agatgtgatgttctgaaaagataaagcctataaaatgttgccaaatttcctgcctgaggtttttccttgctctttgctacatcgctttgctgctcgatcgactcactctgtgtatgcca	
<i>LdSAS-A</i>	-----ccaaa-----	taaaatgttaccgaatttcctccctgaggttttccttgctctttgctacatcgctttgatgctcgatcgactccttcctt-----ca
	* *	* * * * *
<i>LdPAFPIII</i>	cattcactttgtactctccttctcacggtaggtttattatattt	
<i>LdSAS-A</i>	cattcactttgtactctccttctcatggtaggtttattatctt	
	*****	* * * * *

Fig. S3





	<i>LdSAS-A</i>	<i>LdSAS-B</i>	<i>LdSAS-Bm4</i>
Sialic acid synthetic activity	$1.4 \pm 0.2 \text{ mU/mg}$	$2.7 \pm 0.2 \text{ mU/mg}$	NO activity detected

Fig. S5

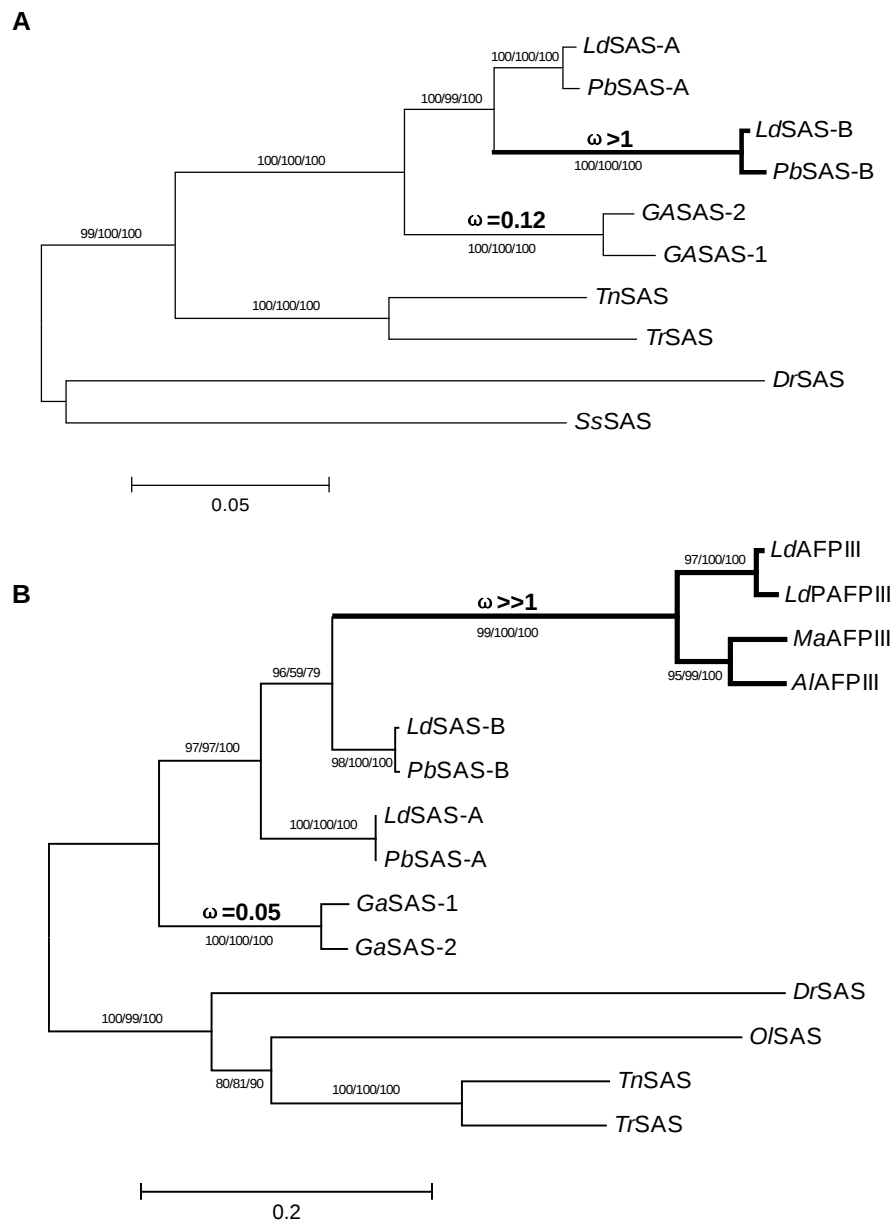
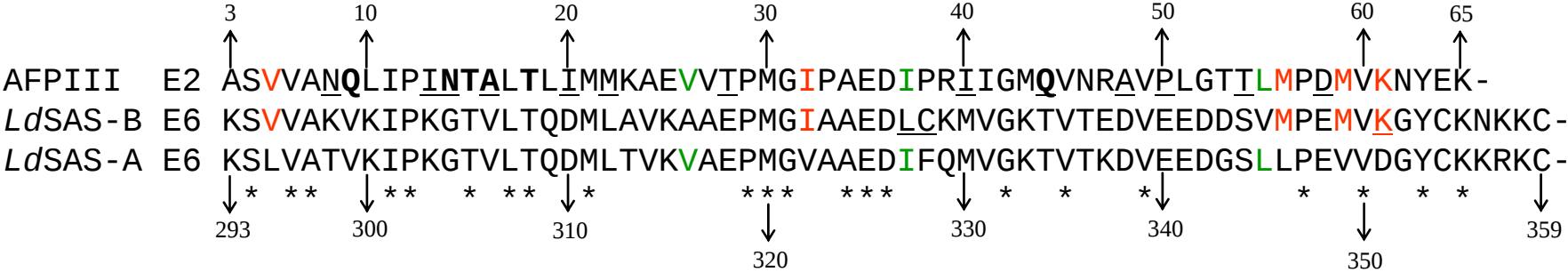


Fig. S6



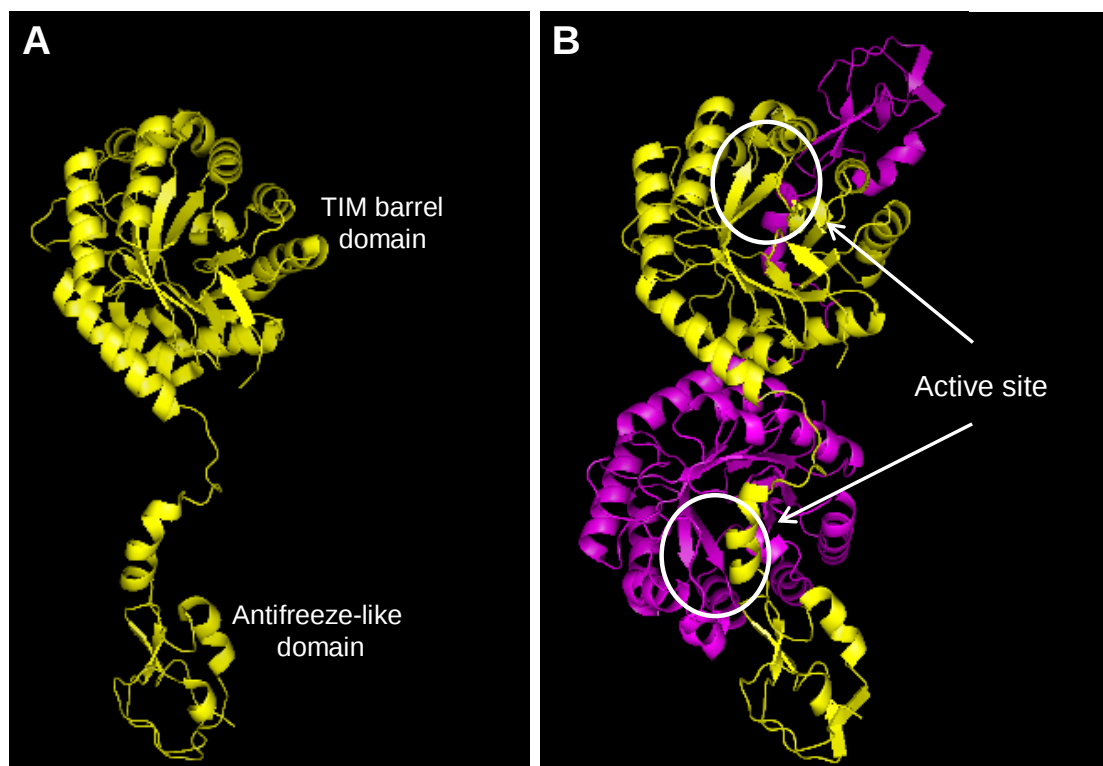


Fig. S8

