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 *LdSAS-B* and exon2 of three typical antifreeze protein (*AFP*)III genes, *Ldψ AFPIII*, *LdAFPIII-1*, and *LdAFPIII-2*, from the *Lycodichthys dearborni AFPIII* locus. Among *AFPIII* genes, *Ldψ AFPIII* shares the greatest sequence similarity with the *LdSAS* genes, including having six additional nucleotides (shown in red) not present in the other two *AFPIII* genes, suggesting that *Ldψ 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 *Ldψ AFPIII* and the *LdSAS* genes. The stop codon (TGA for *SAS* and TAG for *AFPIII*) is underlined.

Fig. S1 (PDF)

Fig. S2. Sequence alignments of *LdSAS-A*, *LdSAS-B*, and *Ldψ AFPIII*, the putative primitive *AFPIII* gene members from *L. dearborni*. (*A*) Nucleotide sequence alignment and deduced amino acid sequence of *LdSAS-A* and *LdSAS-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 *LdSAS-B* that differ from *LdSAS-A* are shown in blue. (*B*) Nucleotide sequence alignment of the homologous regions between *LdSAS-B* and *Ldψ AFPIII*. The proximal 5′ flanking regions of the two genes are homologous. The color highlighted *LdSAS-B* sequence regions (in the ancestral *LdSAS-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 *LdSAS-B* (yellow), *AFPIII* intron 1 from *LdSAS-B* intron 5 (turquoise), and the mature *AFPIII* (exon2) from *LdSAS-B* C-terminal domain (exon6; yellow). (*C*) Nucleotide sequence alignment of the homologous regions between *LdSAS-A* and *Ldψ AFPIII*. *LdSAS-A* and *Ldψ AFPIII* share lower nucleotide identities (66% id) than *LdSAS-B* and *Ldψ AFPIII* (70% identity) (Fig. 2). The proximal 5′ flanking regions of the two genes are not homologous, and without this homologous region, *LdSAS-A* could not have given rise to *Ldψ 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 (*Ldψ 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 (~0.004 °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 × *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 $M^{-1}\cdot cm^{-1}$ 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. S5. The phylogenetic trees used in the nonsynonymous substitution per nosynonymous site (Ka)/synonymous substitution per synonymous site (Ks) ratio tests. (*A*) The phylogenetic relationship of *SAS* genes from two AFPIII-bearing Antarctic eelpouts, *L. dearborni* and *Pachycara brachycephalum*, and other five teleost species inferred by neighbor-joining, maximum likelihood, and Bayesian algorithms. The nonsynonymous substitution per nosynonymous site (dN)/synonymous substitution per synonymous site (dS) (ω) value was calculated using the branch model of codeml in the Phylogenetic Analysis by Maximum Likelihood (PAML) program for each branch, and the values of the *SAS-B* branch of *L. dearborni* and *P. brachycephalum* and the two *SAS* genes of *Gasterosteus aculeatus* were shown. The ω values of all other branches, like those of the *G. aculeatus* branch, are smaller than 1 and therefore, are not shown. The eelpout *SAS-B* branch was then assigned as the foreground branch in the branch-site model and tested for selection against other branches of the tree as the background (Table 1 has detailed results). (*B*) The phylogenetic tree used to test positive selection in the *AFPIII* branch. The tree was constructed using the three homologous regions between the *AFPIII* and *SAS* genes, including the signal peptide, the intron, and the C-terminal/AFP regions. Sequences from each species were aligned with Clustalw, and evolutionary relationship was inferred by the same three algorithms as in *A*. A ω value greatly larger than 1 was detected in the branch leading to the *AFPIII* genes, whereas all other branches showed ω values less than 1. The *AFPIII* branch was then tested for selection using the branch-site model against the background of other branches (results shown in Table 1). *Ld*, *L. dearborni*; *Pb*, *P. brachycephalum*; *Ga*, *G. aculeatus*; *Dr*, *Danio rerio*; *Ss*, *Salmo salar*; *Tn*, *Tetraodon nigroviridis*; *Tr*, *Takifugu rubripes*; *Ma*, *Macrozoarces americanus.* The teleost *SAS* sequences were retrieved from GenBank or Ensembl databases. The GenBank accession numbers are *Ld*SAS-A and *Ld*SAS-B, GQ368892; *Dr*SAS, BC050955; *Ga*SAS-1, BT028736; *Ga*SAS-2, groupIX, 13638390:13682490; *Ss*SAS, BT045230; *Tr*SAS, AJ705104; *Ld*AFPIII, U20439; *Ma*AFPIII, J03924; *Al*AFPIII, M22125; and *Ld*ψAFPIII, GQ368894. The Ensembl Gene IDs are *Tn*SAS, ENSTNIT00000021471; and *Ol*SAS, ENSORLG00000020815.

Fig. S5 (PDF)

Fig. S6. Amino acid alignment of the SAS C-terminal (exon6) domain and mature AFPIII (exon2). The position numbers of amino acids for AFPIII (first line) and SAS (second and third lines) are based on the length of the mature proteins. Underlined residues are detected to experience positive Darwinian selection with posterior probability greater than 95% (Table 1). Bold residues constitute a putative flat ice-binding surface (1). Residues in red are identical between AFPIII and *Ld*SAS-B but different from *Ld*SAS-A, whereas those in green are identical between AFPIII and *Ld*SAS-A but different from *Ld*SAS-B. Asterisks indicate identical residues in AFPIII and *Ld*SAS-A and -B .

Fig. S6 (PDF)

1. Jia Z, DeLuca CI, Chao H, Davies PL (1996) Structural basis for the binding of a globular antifreeze protein to ice. *Nature* 384:285–288.

Fig. S7. The monomer and homodimer structures of sialic acid synthase reported by Gunawan et al. (1). (*A*) View of the SAS monomer showing the triose-phosphate isomerase (TIM) barrel (N-terminal) and the antifreeze-like (C-terminal) domain. (*B*) Arrangement of the SAS holoenzyme with two monomers (one in yellow and the other in purple) in a domain-swapped manner. Each subunit has the surface area of about 12,900 $Å^2$, with the interface area of about 3,880 $Å^2$. The majority of the residues within the active site come from loop regions and S2 and S4 of the C-terminal end of the TIM barrel, with additional contacts made by the helix and coiled region of the linker domain and an extended 11-residue loop of the antifreeze-like domain (residues 285–314) (1). The structure was viewed with the PyMOL program (2) based on deposit (Protein Data Bank (PDB) ID = 1XUZ).

Fig. S7 (PDF)

1. Gunawan J, et al. (2005) Structural and mechanistic analysis of sialic acid synthase NeuB from *Neisseria* meningitidis in complex with Mn2+, phosphoenolpyruvate, and N-acetylmannosaminitol. *J Biol Chem* 280:3555–3563.
2. DeLano WL (2002) *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, CA).

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 *Ld*SAS-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 *Ld*SAS-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 antiflag 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

```
LdSAS-A        CTGGGTAAGTCCTTGGTGGCCACGGTCAAGATCCCCAAAGGCACCGTCCTGACTCAGGACATGTTGACGGTGAAGGTGGCCGAGCCGATGGGCGTCGCGGCCGAGGACATCTTCCAAATG
LdSAS-B        CTGGGTAAATCTGTGGTGGCCAAGGTCAAGATCCCCAAAGGAACCGTCCTGACTCAGGACATGTTGGCGGTGAAGGCGGCCGAGCCGATGGGTATCGCGGCCGAGGACTTGTGCAAAATG
LdPAFPIII      ------GCGTCCGTGGTGGCCAACCAGCTGATCCCCATAAATACTGCCCTGACTCTGATAATGATGAAGGCGAAAGTGGTCACCCCAATGGGCATCCCCGCCGAGGACATTCCCCGAATA
LdAFPIII-1     ------GCGTCCGTGGTGGCCAACCAGCTGATCCCCATAAATACTGCCCTGACTCTGATAATGATGAAGGCGGAGGTGGTCACCCCAATGGGCATCCCCGCCGAGGACATTCCCCGAATA
LdAFPIII-2     ------GCGTCCGTGGTGGCCAACCAGCTGATCCCCATAAATACTGCCCTGACTCTGATAATGATGAAGGCGGAGGTGGTCACCCCAATGGGCATCCCCGCCGAGGACATTCCCCGAATA
                       **  *********      ******** *   ** * ******** *   *** **  ** * * * ** *   ** *****  ** * ********* *   *  ***

LdSAS-A        GTGGGAAAGACCGTGACGAAGGACGTGGAGGAGGACGGCAGCCTCTTGCCAGAGGTGGTGGACGGGTACTGCAAGAAGAGGAAGTGCTGA
LdSAS-B        GTGGGAAAGACCGTGACGGAGGACGTGGAGGAGGACGACAGCGTCATGCCAGAGATGGTGAAAGGGTACTGCAAGAACAAGAAGTGCTGA
LdPAFPIII      ATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCCTCATGCCAGACATGGTGAAAACGTACCGCGTGGCGAAGTAGTTCTGA
LdAFPIII-1     ATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCCTCATGCCAGACATGGTGAAAAACTAT------GAGAAGTAGTTCTGA
LdAFPIII-2     ATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCCTCATGCCAGACATGGTGAAAAACTAT------GAGAAGTAGTTCTGA
                * **** *   ****    **  ***  *  ** *   * * ** *******  ***** *    **  **  *   * * *** ****
```

A

```
LdSAS-A     ------gaacaaggaaccaggaagctcattagagtgggtgggtagctcttaaaagagccgttgtgtgagcgacacctcaactctgccagaccatgaaagcatccaagcaccccgctagct
LdSAS-B     ------tgtcattgtgctattggatagaggagaccatgccgatctagacagctgatattatgaaaacagccccaaacaataacagcaagcaaataatgaatgtgcatgcgtgaggagtaa

LdSAS-A     ctggcgaattttgcagaaatgttgttgttgatgtggacgggccgctatcaaatgagccaaaacatgaaattgaaatcgatgccatcagccagcttctccggcgggagaggggtgaggcga
LdSAS-B     ttggaagatgtatcagaaataaactacaggcatgcaagagtatagtcttcctgactggactcttattttgtcttttcttatatcattttgggatgggcgcatgggagagttttccacctg

LdSAS-A     cggcggatgccttcttggatactttcatgggtccagttagcttgaagtgtctctcgaaggggatgaatcttgtatcgctccgtcgctgactagatagtttccttatgctaatgagactgt
LdSAS-B     tctcagcttgtttttacccaaaatattgtatacctattagaaccgttgtcacagggttcaaatgaatgttttcgtttagttttgatcgtgtaagtgcttgttgttcgacagcaagaagca

LdSAS-A     gctgtcgctcgcacaggattggctgacggcagggggcgggtcttaacagtacatgagaaccgatttatacgtctagtttcataagtcaataaataaaccgtctatatgaaaaaactacga
LdSAS-B     gaggatctgtgtgcaggaagtcagctgctgcatggatcacaagttatacgatgcaagcttgtgatagtttgcacaaaaacaagttataaaacagtttacttgaaagaatataaaatttcc

LdSAS-A     gttaaagcctatttttacacttctgtgacttaaatactaaagtactaaatattttctcatttaattaagaacttttgtgtctgtaaaggaaatccggatattacgtcatactacaagcgg
LdSAS-B     actacaattggcataaggaggtggtggcagtgtgacggtgtcacagtgtcctaccactttcagaccaaacgggccggttgaacaggatatactggttgaacaagagtggctgctcctgtt


LdSAS-A     aagtcctcccacatcc--tggatgtgcagcacattaattggtccggtcagattcaagtctcagctacggctttcgcttcgt-----gctaaataattc----gtgctaaataattc--ta
LdSAS-B     aagtcctaccgcatatattagatgtgcagcacatgaactcatcctttcagattcaagtctcagctacggctttccgttccttcgccgctaaataattccgttattttcaccagtttgttt
            ******* ** ***    * ************** ** *  ***  ****************************  *** *     ************     *  * *  * **   *
```

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     aatagtttgttttctgtggacaaaATGCCGTTAACGTTCGAGCTGTGTCCCGGTAGAATGATCGGAGGAGATCATCCGTGCTTCATCATCGCTGAAATTGGACAAAATCACCAGGGAGAC
LdSAS-B     tccgttttcttttctgcagacaaaATGCCGTTAACGTTCGAGCTGTGTCCCGGTAGAATGATCGGAGGAGATCATCCGTGCTTCATCATTGCTGAAATTGGACAAAATCACCAGGGAGAC
                 *** *******  ************************************************************************ ******************************


              E  I  A  K  K  M  I  K  M  A  K
LdSAS-A     ATTGAGATTGCCAAGAAAATGATCAAAATGGCAAAGgtaataaaatcgacacgtatcgctttgtattgaatgtgaatacgtatattttctgat---------------------------
LdSAS-B     ATTGAGATTGCCAAGAAAATGATCAAAATGGCAAAGgtaataaaatcgacac-------tttatattgaatgtgaatacgtatattttctgatgttcttatctttaacagccaagacaca
            ****************************************************       *** ******************************


LdSAS-A     ---acgagcactgacgtaatcagaactaaattaaaaaatgtacccattgaatatttgtgtcgccttatcttcc---tctt--------------------ctgaatcattactt------
LdSAS-B     cagacaataactaacttattcagaactcagtttaaaaatgtgcagattgaatatttttgtcgcctcatcgtctgaatcattactttaaatatactataaactaaactagtgcttgtatgt
               ** *  *** ** ** ******** * ** ******** *  *********** ******** *** **    ** *                   ** **  * * ***

LdSAS-A     ----------------------------------------------------------------------------------------------------------------------ta
LdSAS-B     agacctttatgctttaattaagttattaagattataatgtgcaatgttgcatagagcaagtgggtccgaagggtgtttgtataaagacgtgcactaggatgttaaatgaccggatatgta
                                                                                                                                **

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

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

LdSAS-A     gtgtatattgttaattatttcctgaaaagaactgtct--tttggtcccagttacaggg-aacatagacaactttt-gagttatttaagttataagcttttgttaatttctagaggaattt
LdSAS-B     gtgtatattgttaattatttcctgaaaataagtgtgtattttggtcccagttacccagtaaaatagacaactttttgagttatttcagtgatacatttttgttaatttctagaataattt
```

```
                ***************************** ** *** *  ***************   * ** ************* ********* *** ***   *****************  *****

LdSAS-A         ccttcagaagaactacatcaaaccacagtattaaacatatttcctatctatttattaatgggaacttcattccccatatatgttgcttattatgccgcttttagatcccattaattcatg
LdSAS-B         cctttagaataactaaattaaaccacagta--aaaaaaatcttgtatctatttattaatgtgaacttcattccccatatatgttgcttattatgccgcttttagatcccattaattcatg
                **** **** ***** ** ***********  *** * ** *  *************** ************************************************************

LdSAS-A         atcagctcacctggtgaaaaaac---tatttgtaacactagaaatagatttgaataggtataat-gaaataattaatggtaagtgtagtaattgaa------cacggtctcctctt--aa
LdSAS-B         atcagctcacctggtgaaaaaacaactatttgtaacactaggaatatattttaataggtataattggaaccgacttggataggggcattgaaggagacgactcaggggatgcttttccag
                ************************   *************** **** **** ************ * **        * ** * * * * *  **       ** **  * ** **  *

LdSAS-A         tagta------tgaaatagtcctttctaggaac-cttactgtgtttttag-cacaatctgtag----gagcccttaa-atcaatcatcactacttagtatgttttccaa...43nt...
LdSAS-B         catttatttgctgcagaagacaatacaacacacacatatattaacttctggtctccctctgcactcctgctccctcaacgatacaaaacaccattttgcatacggaacaa..2331nt..
                 * *      ** *  ** *  * * *   ** * **   *   ** * * * *  **** *     *  **** **    *   * *** * ** * **      ***
                 ┌─▶ Exon 2
                 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-A         GACCAGGGTGCGGATTGTGCCAAATTCCAGAAGAGTGAATTGGAGCATAAATTCAACAAGAAAGCTTTGGAGCGCGAGTACAACTCTGTTCACTCCTGGGGAAAAACCTATGGTGACCAC
LdSAS-B         GACTGCGGTGCTGATTGCGCCAAATTCCAGAAGAGTGAATTAGAGTGTAAATTCAACAAGAAAGCCTTGGAGCGCGAGTACAACTCTGTTCACTCCTGGGGGAAAACCTATGGTGACCAC
                ***   ***** ***** ********************** ***  ****************** ********************************** ******************

                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-A         AAGCGCCATCTGGAGTTCAGCCATGATCAGTACAGGGAACTCAAGAAATATGCAGAGGAGGTGGGGATCTGCTTCACTGCCTCTGGGATGGATGAGgtaagaaagagacgtctgaagaag
LdSAS-B         AAGCGCCATCTGGAGTTCAGCCATGATCAGTACAGGGAACTGCAGAAATATGCTGCGGAGGTGGGGATCTTCTTCACTGCCTCTGGGATGGATGAGgtaagaaagagacgtctgaagaag
                ****************************************  ********** * ************** ************************************************
                                                                      ┌─▶Exon 3
                                                                      M  A  V  E  F  L  H  E/GL  N  V  P  F  L/FK  V  G/AS  G/LD  T
LdSAS-A         gctttctgtggttttaatcaggttgaagtttaaacgtgt---ttctttctgtcttttagATGGCAGTGGAGTTCCTCCATGAGCTTAATGTGCCCTTCTTGAAAGTGGGCTCAGGAGACA
LdSAS-B         gctttctgtggttttaatcaggttgaagtttaaccatatgttttctttctgtcttttagATGGCAGTGGAGTTCCTCCATGGGCTTAATGTGCCCTTCTTTAAAGTGGCCTCACTAGACG
                ********************************* * * *   ************************************** ****************** ******* ****  ****


                /AN  N  F  P/SY  L/ME  K  T  A  K  K  G
LdSAS-A         CCAACAACTTCCCCTATCTGGAGAAGACCGCCAAGAAAGgtgagagtgaactgcagcagatgtattttccgttgcatatttaagtgctaaaatcatcagtgcactttgttgggacttcac
LdSAS-B         CCAACAACTTCTCCTATATGGAGAAGACCGCCAAGAAAGgtgagagtgaactgcagcagatgtattttatgttgcatatttaaatgctaaaatcatcagtgcactt-----------cac
                *********** ***** ************************************************  ************* ********************          ***

LdSAS-A         catgtggttggaacaaagggagttggaagttgggcctacatggttcaacagccccaccagacgttaagtctcacaaacttaacactggggtctcattggacgagggccacaggcgccaca
LdSAS-B         cattacat--aaataaatactatattaagta------aaacaataaaacagaaatagaagcagaaagaatgcacaaatgtgtccgtagaa----attagttaagag--acgaaaaataaa
                ***    *   ** ***     *   ****       * *   *  *****    *  **  *  *     ******  *  *  * *      *** *   ** *  **       * *

LdSAS-A         gggggctccctgtggtgcgactgagactataaaaatctgcattatctccagctcggagccatcccacgcaattcgattcgcatgaaagg-----------1009nt--------------
LdSAS-B         tgaagtttcgcttgatt--gttga-actacaggag-c----tcattaagaggagagagaagggttgcattttcccctttgccc--at-g------------45nt--------------
                 *  * * *   ** *     *** **** *  *  *    * **    **    ***        *    * *  ** **    *  *
                 ┌─▶ Exon 4
                  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-A         GACGGCCCATGGTGGTGTCCAGCGGGATGCAGAGCATGGAGACGATGCGTCGGGTCTACAAAACGGTGAAGAAGCACAACCCCAACTTCACCATCCTGCAGTGCACCAGCGCCTACCCTC
LdSAS-B         GACGGCCCATGGTGGTGTCCAGCGGGATGCAGAGCATGGAGACGATGCGTCGGGTCTACAAAATAGTGAAGGAGCACAACCCCAACTTCACCATCCTGCAGTGCACCAGCTCCTACCCTG
                ************************************************************  ****** ************************************** ********

                /VE/DA/PE/KD/NV  N  L  R/KV  I  T/S
LdSAS-A         TGGAAGCTGAAGATGTCAACCTCAGAGTGATAACGgtgagaagaacagagcattactcctttatcaaaaggttggtacagaaacaagatctcaaacaagttgagttgggaaattgtctca
LdSAS-B         TGGATCCTAAAAATGTCAACCTCAAAGTGATATCGgtgagaaaagtaggattttaaatcctgagcgtctcact--taaagtaatagtttgacatttcag--gagttggtggatttgtta
```

```
                 ****  ** ** ************ ******* ********* *  **    ***   * * * *      *  ** ** ** *   *  **    **  *******   *** * * *

LdSAS-A          ctccattgaaagtagacgttttcagcagatgcatattgagttttaaaagccgttaagatatcacatgtcttattgcaatttgtgtttttagcatttatttccacaaattaaacttacaat
LdSAS-B          ct--gttgggctgagccaggct-agctgttcccccctggtttcca--gtcattatgctaagctaagctaaccggcagctggc---tctagcttg---------aaatttaactcacaaa
                 **   ***     ** *    * *** * * *       ****  *  * * *** * **    * *       ***  * *    * **** *          ***** **** ****

                                                                                  Exon 5
                                                                                  E  Y  Q  K  E  F  P  D  I  P  I  G  Y  S  G  H  E  S/TG  I
LdSAS-A          aattaagatgaatcttcctgttttgatggccgtgtagaattgacact----488nt----GAATACCAGAAGGAATTTCCCGATATTCCCATCGGGTACTCCGGCCATGAGTCTGGGATC
LdSAS-B          tgtgagtgtgg-tgtcaatcgtctgagag—att-taggagtaacact-----20nt----GAATACCAGAAGGAATTTCCCGATATTCCCATCGGGTACTCCGGCCATGAGACTGGGATC
                   * *   **  * *   *  * ***  *   * *** * * *****             ************************************************* ********

                 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
LdSAS-A          AGTATTTCAGTTGCGGCCGTGGCTATGGGTGCAAAGGTCATCGAGCGCCACGTAACCCTGGACAAGACGTGGAAAGGAAATGACCACGAAGCCTCTCTGGAGCCCAGTGAGCTAGCCGAG
LdSAS-B          AGTATTTCAATTGCGGCCGCAGTTATGGGTGCAAAGGTCATCGAGCGCCACGTAACCTTGGACAAGACTTGGAAAGGAAATGACCACAGCGCCTCTCTGGAGCCCGGTGAGCTAAAAGAG
                 ********* *********  * ********************************** ********** ******************   *************** ********   ***

                 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
LdSAS-A          CTGGTTCGTTCCATCCGACTGGTGGAGAGGGCGCAGGGCACCGGCATCAAGCAGATGTTACCTTGTGAGAAGCCGTGCCATGATAAGgtgagatattcttgctccactaggaaacttaac
LdSAS-B          CTGGTTCGTTGCATCCGAGAGGTGGAGATGGCGCAGGGGAGCTGCATCAAGCAGATGTTACCTTGTGAGAAGGCATGCCATGCTAAGgtgagatattctggctccacaaaaacgtttttt
                 ********** *******  ******** ********* * * ***************************** * ******* *************** ******* *  *   **

LdSAS-A          tt----gtactgttcagaaagacagaggaccttt------taagaaaaca------tttttgaaattgttttcaactcccaaccacgagaacaataataatttcagaccttttattttct
LdSAS-B          tttaaaatgcgagccacagttaaattacatttttactgtttagaaagacagataacctttaagaaaacatttctactgt--gccatgagaacaataataattgcagaccttttattttct
                 **     * *    ** *   * *    *  ***      **  ** ***       *** **    **** ***       *** *************** *****************

                          Exon 6
                          L  G  K  S  L/VV  A  T/KV  K  I  P  K  G  T  V  L  T  Q  D  M  L  T/AV  K  V/AA  E  P  M  G  V/IA  A  E  D  I/L
LdSAS-A          attctccaagCTGGGTAAGTCCTTGGTGGCCACGGTCAAGATCCCCAAAGGCACCGTCCTGACTCAGGACATGTTGACGGTGAAGGTGGCCGAGCCGATGGGCGTCGCGGCCGAGGACAT
LdSAS-B          attctccaagCTGGGTAAATCTGTGGTGGCCAAGGTCAAGATCCCCAAAGGAACCGTCCTGACTCAGGACATGTTGGCGGTGAAGGCGGCCGAGCCGATGGGTATCGCGGCCGAGGACTT
                 ****************** **  ********* ****************** ************************ ********* ***************  ************** *

                  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
LdSAS-A          CTTCCAAATGGTGGGAAAGACCGTGACGAAGGACGTGGAGGAGGACGGCAGCCTCTTGCCAGAGGTGGTGGACGGGTACTGCAAGAAGAGGAAGTGCTGAgggtcgcaaggagttcatca
LdSAS-B          GTGCAAAATGGTGGGAAAGACCGTGACGGAGGACGTGGAGGAGGACGACAGCGTCATGCCAGAGATGGTGAAAGGGTACTGCAAGAACAAGAAGTGCTGAggatggcaaggagtttataa
                  * * ********************** ****************** **** ** ******** ***** * ************** * ************ * ********** ** *

LdSAS-A          atgggtataaaaggcacatcatttgagaggaaacgtgatcgattcctctggtcattttgtttggtggtgatgggattgtcctggacgaggagtaattggttcgtgcattgctctgatagt
LdSAS-B          atgggtagaaaaggcacatcatttgagaggaaatgt-------t----------------ctggcgat-a----------ttg-------------------------------------
                 ******* ************************ **       *                 *** * * *           **

LdSAS-A          ttcaatatatcacagctggggttcaaacggttactttgctcaagtgataaacccaccataatcttacaagaggctcagtagtagtctcagtcctgtggtcagtactagtgcagtgctttc
LdSAS-B          ----------------------------g-tcattttg-tgt-----------c-------tcttacaagaggctcagtagtagtctcagtcctgtggtcagtactagtgcagtgctttc
                                             * * * **** *             *       ************************************************************

LdSAS-A          tgaggtggcacaaaagccagttgttgttactgtgaacttaacatactttgcactatagtttttgcctgctctgtagagattaaaggaacattgaacacacgaagcggctactggtcatct
LdSAS-B          tgaggtggcacaaaagccagttgttgttactgtgaacttaacatactttgcactatagtttttgcctgctctgtagagattaaaggaacattgaacacacgaagcggctactggtcatct
                 ******************************************************************************************************************
```

```
LdSAS-A     cctcacacaacggtgagatcacagagcccctttgccaaacactcattgtattcttcttgtacctgctcattcatgtaagttcatgatatcctgcatggtttgtgtttgtgctgcttcttt
LdSAS-B     cctcacacaacggtgagatcacagagcccctatgccaaacactcattgtattcttcttgtacctgctcattaatgtaagttcatgatatcctgcatggtttgtgtttgtgctgcttcttt
            ******************************** ************************************** ***********************************************

LdSAS-A     gagagaagcacaataagactactggagccagacgcttcccaaaaccaaaagaagaaatgcccccctctcacaattaaccttgtttttgtcatgaacccaaacccgtccggatgaaacctgt
LdSAS-B     gagagaagcacaataagactactagagccagacgcttcccaaaaccaaaagaagaaatgcccccctctctcaattaaccttgtttttgtcatgaacccaaacccgtccggatgaaacctgt
            ************************ ******************************************** ***********************************************

LdSAS-A     ggagcctccaaataaaatgttacccgaatttcctccctgaggttttcctttgtctttgctacatcgctttgatgctcggatcgactccttccttcacattcactttgtactctccttctc
LdSAS-B     ggagcctccaaataaaatattacccgaatttcctccctgaggttttcctttgtctttgctacatcgctttgatgctcggatcgactccttccttcacattcactttgtactctccttctc
            ****************** ***********************************************************************************************

LdSAS-A     atggtaggtttattatctttagatgtgcagttagtttctgtgaaataacataccacacactgatattgtctgtgcattgacttggttatgca
LdSAS-B     atggtaggtttattatctttagatgtgcagttagtttctgtgaaataacataccacacactgatattgtctgtgcattgacttggttatgca
            ******************************************************************************************
```

B

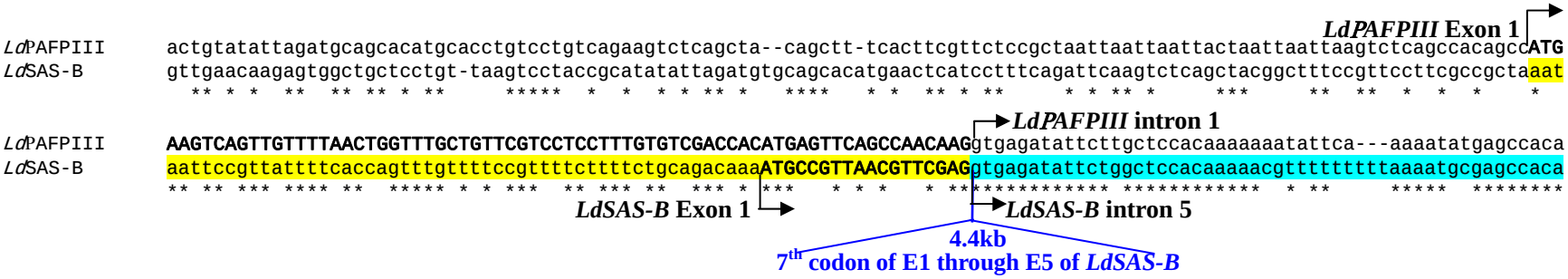
```
LdPAFPIII   tgacattgtgctattggaaagaagagacca-gctgatctagacagttgatatcatgatcaacagccccaaacaa--------------------gtgtgcatgcgtgaggagtgattggc
LdSAS-B     tgtcattgtgctattggatagaggagaccatgccgatctagacagctgatattatgaa-aacagccccaaacaataacagcaagcaaataatgaatgtgcatgcgtgaggagtaattgga
            ** *************** *** ******* ** *********** ****** ****  ***************                     ****************** *****

LdPAFPIII   agatgtatgagaactgaaccatgggcttaggcatgcaagacta--gtcttcttgactgaacttgcgctagagcttcattggaggcatctattttgtcttttctcgtatgattttgggatg
LdSAS-B     agatgtatcagaaataaactac------aggcatgcaagagtatagtcttcctgactggactc-----------------------ttattttgtcttttcttatatcattttgggatg
            ******** **** * *** *       ************ **  ****** ****** ***                       *****  *** ********************

LdPAFPIII   gacgcatgggagagtttttc—-ccgtctcagcttgctttttaccctaaatattgtatacctattagaaccgttgtcacagggttcaaattaacattttcgtttagttttgatcatg-- 45nt--taag
LdSAS-B     ggcgcatgggagagttttccacctgtctcagcttg-tttttacccaaaatattgtatacctattagaaccgttgtcacagggttcaaatgaatgttttcgtttagttttgatcgtg--505nt--taag
            * ***************** *  * *********** ********* ****************************************** **  ******************* **        ****

LdPAFPIII    ggcttgttatacgacagaaagaag---aggatatgcgtgcaggctgtcagctaataaatggatcacaagttatagaatgcaagcttgtgatagtttgcacaaaaacaagttatac-----
LdSAS-B      tgcttgttgttcgacagcaagaagcagaggatctgtgtgcaggaagtcagctgctgcatggatcacaagttatacgatgcaagcttgtgatagtttgcacaaaaacaagttataaaacag
             ******* * ****** ******   ***** ** *******  *******  *  *****************  ***************************************

LdPAFPIII   tttacttataagaatataaaatttccattacaattggcataaggaggtg-------tgatacagtgacctactttcaggccaataggaaacgggatatgccggttaagtcctcc-cacat
LdSAS-B     tttacttgaaagaatataaaatttccactacaattggcataaggaggtggtggcagtgtgacggtgtcacagtgtcctaccactttcagaccaaacgggccggttgaacaggatatactg
            *******  ****************** ********************       **  ** *** *  * * **   *** *   * **   *   ******* *         **
```




```
LdPAFPIII   actgtatattagatgcagcacatgcacctgtcctgtcagaagtctcagcta--cagctt-tcacttcgttctccgctaattaattaattactaattaattaagtctcagccacagccATG
LdSAS-B     gttgaacaagagtggctgctcctgt-taagtcctaccgcatatattagatgtgcagcacatgaactcatcctttcagattcaagtctcagctacggctttccgttccttcgccgctaaat
              ** * *  **  ** ** * **     *****  *  *  * * ** *   ****   * *  ** * **     * * ** *     ***     **  **  *  *  *    *

LdPAFPIII   AAGTCAGTTGTTTTAACTGGTTTGCTGTTCGTCCTCCTTTGTGTCGACCACATGAGTTCAGCCAACAAGgtgagatattcttgctccacaaaaaaatattca---aaaatatgagccaca
LdSAS-B     aattccgttattttcaccagtttgttttccgttttcttttctgcagacaaaATGCCGTTAACGTTCGAGgtgagatattctggctccacaaaaacgtttttttttaaaatgcgagccaca
            ** ** *** **** **  ***** * * ***  ** *** **  *** * ***   * * *   * ************** ************  * **     *****  ********
```

```
LdPAFPIII    gtaaaattcaattgttttctgttcagaaagacagagaacctttaagaaaacatttttgaaatagtttttcaactgtgccatgagaacattaataatgtctgacctttttattttccattc
LdSAS-B      gttaaattacatt-tttactgtttagaaagacagataacctttaagaaaacatttct---------------actgtgccatgagaacaataataattgcagacctttttattttctattc
             ** *****  *** *** ***** *********** ******************* *               ***************** *******  * ************** ****

                   LdPAFPIII Exon 2
LdPAFPIII    ttcaagGCG------TCCGTGGTGGCCAACCAGCTGATCCCCATAAATACTGCCCTGACTCTGATAATGATGAAGGCGAAAGTGGTCACCCCAATGGGCATCCCCGCCGAGGACATTCCC
LdSAS-B      tccaagCTGGGTAAATCTGTGGTGGCCAAGGTCAAGATCCCCAAAGGAACCGTCCTGACTCAGGACATGTTGGCGGTGAAGGCGGCCGAGCCGATGGGTATCGCGGCCGAGGACTTGTGC
             * ****  *      ** ***********      ******** *   ** * ******** *   *** **  ** *** * ** *   ** ***** *** * ********* *   *
                   LdSAS-B Exon 6
LDPAFPIII    CGAATAATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCCTCATGCCAGACATGGTGAAAACGTACCGCGTGGCGAAGTAGttctgagcgtgccaaggagttt       cttcccaa
LDSAS-B      AAAATGGTGGGAAAGACCGTGACGGAGGACGTGGAGGAGGACGACAGCGTCATGCCAGAGATGGTGAAAGGGTACTGCAAGAACAAGAAGTGCTGAggatggcaaggagttt.399nt.cttcccaa
               ***  * **** *   ****    **  ***  *  ** *   * * ********** *********  **** **  *   *** *** *****  ** **********       ********

LdPAFPIII    aaccaaaagaagaaatgcctcctctcacaattaaccttgtttttgtcacaaacccaagtctgtccggatgttaactgaacatgtcaaaacctgtggcgactatgttgagatgtgatgttc
LdSAS-B      aaccaaaagaagaaatgcccccctctctcaattaaccttgtttttgtcatgaacccaaacccgtccggatg----------------aaacctgtggagcct------------------c
             ******************* ****** *********************  *******  * *********                ********** * **                  *

LdPAFPIII    tgaaaagataaagcctataaaatgttgcccaaatttcctgcctgaggtttttccttgtctttgctacatcgctttgctgctcggatcgactcactctgtgtatgccacattcactttgta
LdSAS-B      caa-------------ataaaatattacccgaatttcctccctgaggtttttcctttgtctttgctacatcgctttgatgctcggatcgactccttcctt------cacattcactttgta
               *             ******* ** *** ******** *********** * ********************** ***************  **  *      ***************

LdPAFPIII    ctctccttctcacggtaggtttattatttt
LdSAS-B      ctctccttctcatggtaggtttattatctt
             ************ ************** **
```

C

```
LdPAFPIII    ----------------------------------------------------5'flanking region-----------------------------------------------
LdSAS-A      ----------------------------------------------------5'flanking region-----------------------------------------------
                                                          No sequence similarity

LdPAFPIII    aagtcctcccacatactgtatattagatgcagcacatgcacctgtcctgtcaga----agtctcagctacagctttcacttcgttctccgctaattaattaattactaattaattaagtc
LdSAS-A      aagtcctcccacatcctggatg-----tgcagcacattaattggtccggtcagattcaagtctcagctacggctttcgcttcgt-----gctaaataattcgt-gctaaataatt-----
             ************** *** **      **********  *   **** ******    ************ ****** ******     ***** *****  *  **** *****

                   LdPAFPIII Exon 1
LdPAFPIII    tcagccacagccATGAAGTCAGTTGTTTTAACTGGTTTGCTGTTCGTCCTCCTTTGTGTCGACCACATGAGTTCAGCCAACAAGgtgagatattcttgctccacaaaaaaatattcaaaa
LdSAS-A      ---------------------------ctaaatagtttgtt------------ttctgtggacaaaATGCCGTTAACGTTCGAGgtgagatattcttgctccactaggaaac--------
                                          *** * ***** *            ** *** *** * ***   * * *   * ********************** *  ***
                                                                                 LdSAS-A Exon 1
                                                                                 3.8 kb
                                                                                 7th codon of E1 through E5 of LdSAS-A

LdPAFPIII    atatgagccacagtaaaattcaattgttttctgttcagaaagacagagaaccttt-aagaaaacatttttgaaatagtttttcaactgtgccatgagaacattaataatgtctgacctt
LdSAS-A      -------------------ttaact-tgtactgttcagaaagacagaggacctttaagaaaacatttttgaaattgttttcaactcccaaccacgagaacaataataatttcagacctt
                                 * ** * * * ***************** ****** ****************** *****     *    *** ******* ******* ** ******
```

```
                                  ►LdPAFPIII Exon 2
LdPAFPIII     ttattttccattcttcaagGCG------TCCGTGGTGGCCAACCAGCTGATCCCCATAAATACTGCCCTGACTCTGATAATGATGAAGGCGAAAGTGGTCACCCCAATGGGCATCCCCGC
LdSAS-A       ttattttctattctccaagCTGGGTAAGTCCTTGGTGGCCACGGTCAAGATCCCCAAAGGCACCGTCCTGACTCAGGACATGTTGACGGTGAAGGTGGCCGAGCCGATGGGCGTCGCGGC
              ******** ***** ****  *      *** *********       ******** *   ** * ******** *   *** *** ** *** **** *   ** ****** ** * **
                                  ►LdSAS-A Exon 6

LdPAFPIII     CGAGGACATTCCCCGAATAATCGGAATGCAAGTGAACAGGGCAGTGCCGTTGGGCACAACCCTCATGCCAGACATGGTGAAAACGTACCGCGTGGCGAAGTAGttctgagcgtgccaagg
LdSAS-A       CGAGGACATCTTCCAAATGGTGGGAAAGACCGTGACGAAGGACGTGGAGGAGGACGGCAGCCTCTTGCCAGAGGTGGTGGACGGGTACTGCAAGAAGAGGAAGTGCTGAgggtcgcaagg
              *********   ** ***  * **** *   ****  * **  ***  *  ** *   * **** *******  ***** *   **** **  *  ** * *** ***** **  *****

LdPAFPIII     agttt         cttcccaaaaccaaaagaagaaatgcctcctctcacaattaaccttgtttttgtcacaaacccaagtctgtccggatgttaactgaacatgtcaaaacctgtggcgactatgttg
LdSAS-A       agttc..518nt..cttcccaaaaccaaaagaagaaatgccccctctcacaattaaccttgtttttgtcatgaacccaaacccgtccggatg----------------aaacctgtggagcct------
              ****          ************************** ***************************  *******  * *********                *********** * **

LdPAFPIII     agatgtgatgttctgaaaagataaagcctataaaatgttgcccaaatttcctgcctgaggtttttccttgtctttgctacatcgctttgctgctcggatcgactcactctgtgtatgcca
LdSAS-A       ------------ccaaa-------------taaaatgttacccgaatttcctccctgaggttttcctttgtctttgctacatcgctttgatgctcggatcgactccttcctt------ca
                           *  **             ********* *** ******** *********** * ********************** ***************  **  *      **

LdPAFPIII     cattcactttgtactctccttctcacggtaggtttattatttt
LdSAS-A       cattcactttgtactctccttctcatggtaggtttattatctt
              ************************* ************** **
```

Fig. S3

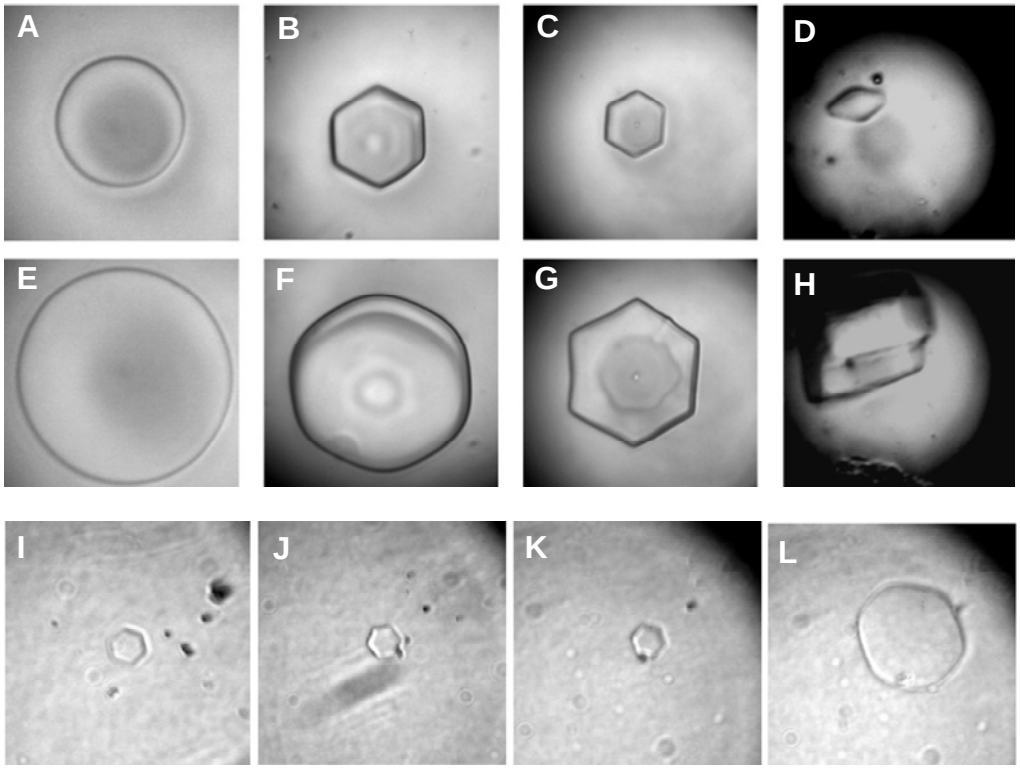
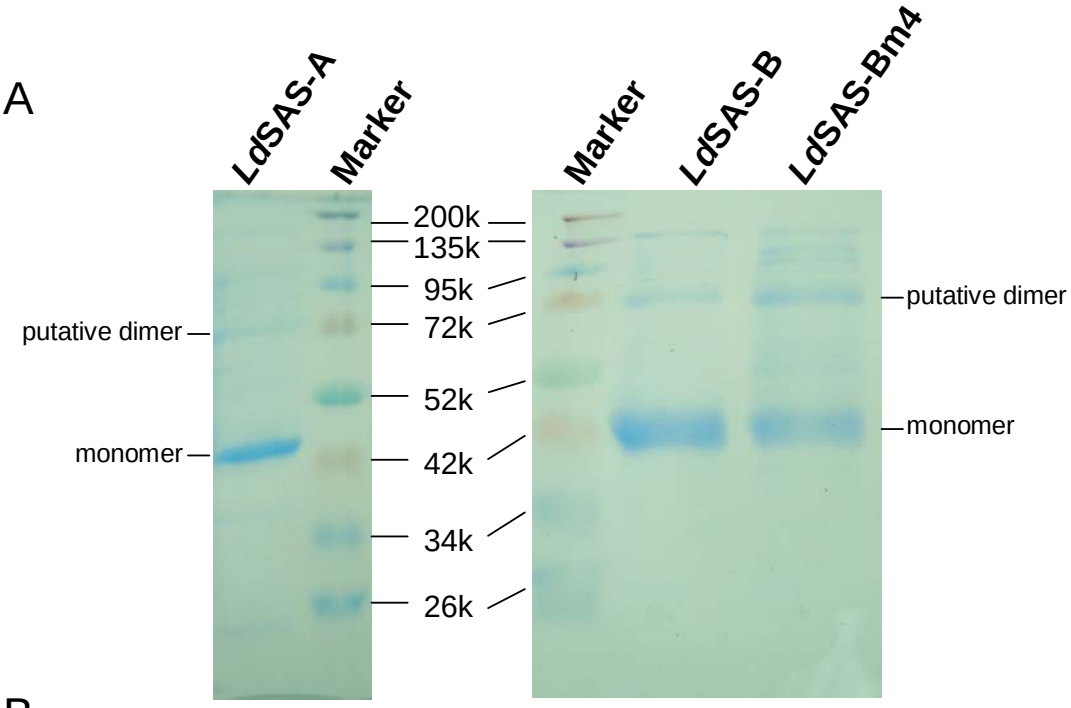


Fig. S4

A



B

	*Ld*SAS-A	*Ld*SAS-B	*Ld*SAS-Bm4
Sialic acid synthetic activity	**1.4±0.2mU/mg**	**2.7±0.2mU/mg**	**NO activity detected**

Fig. S5

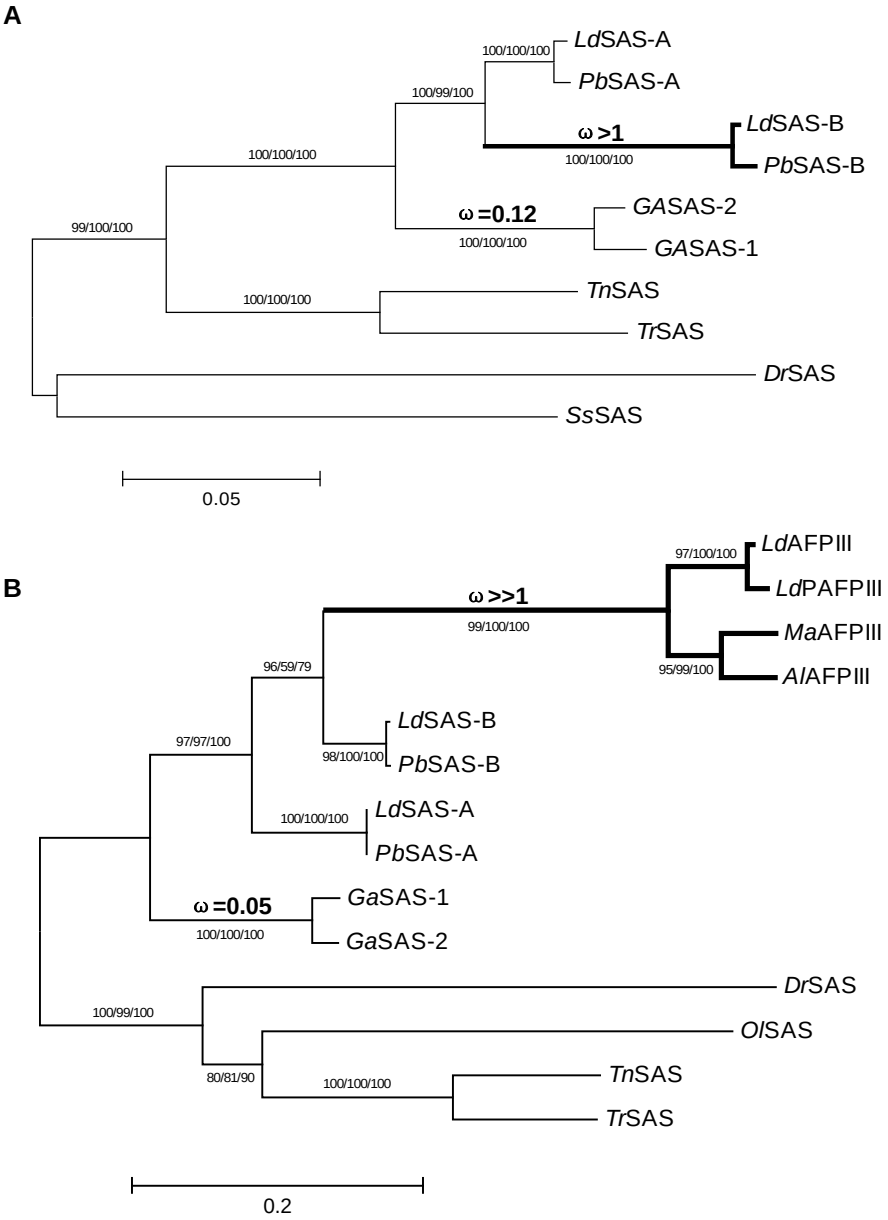

A
100/100/100
LdSAS-A
PbSAS-A
100/99/100
ω >1
100/100/100
LdSAS-B
PbSAS-B
100/100/100
ω =0.12
100/100/100
GASAS-2
GASAS-1
99/100/100
100/100/100
TnSAS
TrSAS
DrSAS
SsSAS
0.05
B
97/100/100
LdAFPIII
LdPAFPIII
ω >>1
99/100/100
MaAFPIII
95/99/100
AlAFPIII
96/59/79
LdSAS-B
97/97/100
98/100/100
PbSAS-B
100/100/100
LdSAS-A
PbSAS-A
ω =0.05
GaSAS-1
100/100/100
GaSAS-2
DrSAS
100/99/100
OlSAS
80/81/90
100/100/100
TnSAS
TrSAS
0.2

Fig. S6

```
                3         10            20            30            40            50            60      65
AFPIII   E2 ASVVANQLIPINTALTLIMMKAEVVTPMGIPAEDIPRIIGMQVNRAVPLGTTLMPDMVKNYEK-
LdSAS-B  E6 KSVVAKVKIPKGTVLTQDMLAVKAAEPMGIAAEDLCKMVGKTVTEDVEEDDSVMPEMVKGYCKNKKC-
LdSAS-A  E6 KSLVATVKIPKGTVLTQDMLTVKVAEPMGVAAEDIFQMVGKTVTKDVEEDGSLLPEVVDGYCKKRKC-
            * **   ** *  ** *        ***  ***   * *  *   *        *  *  * *
            293      300          310           320           330           340           350           359
```

Fig. S7

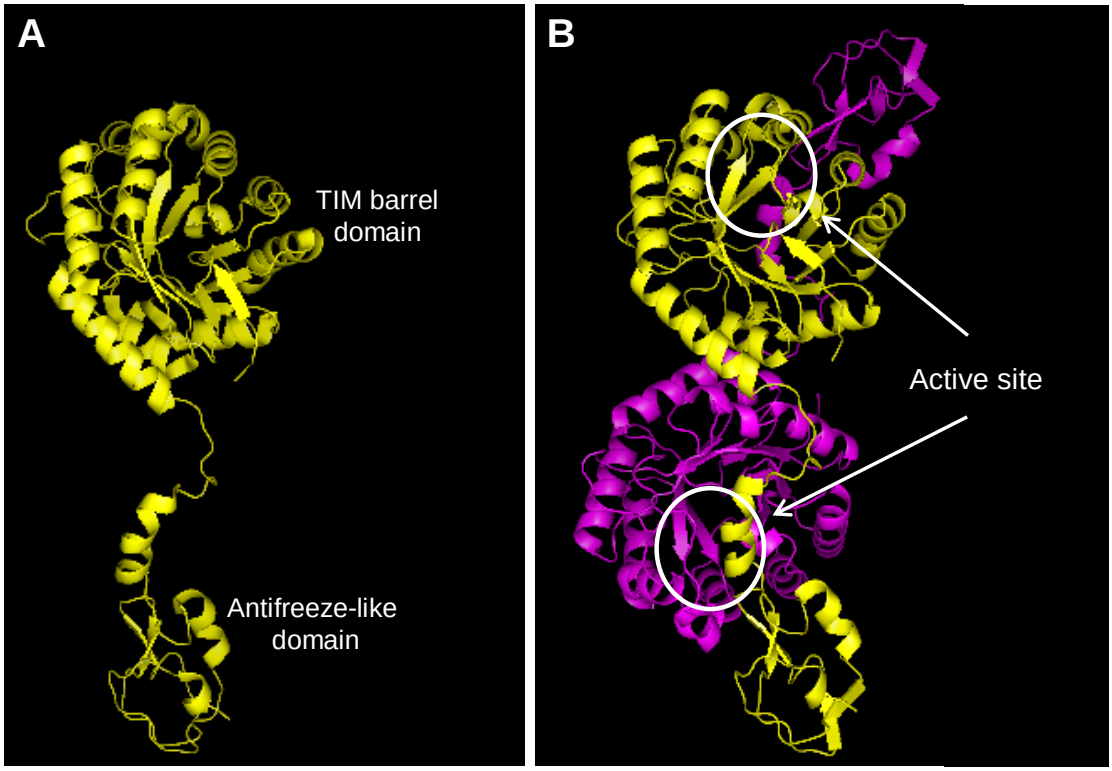
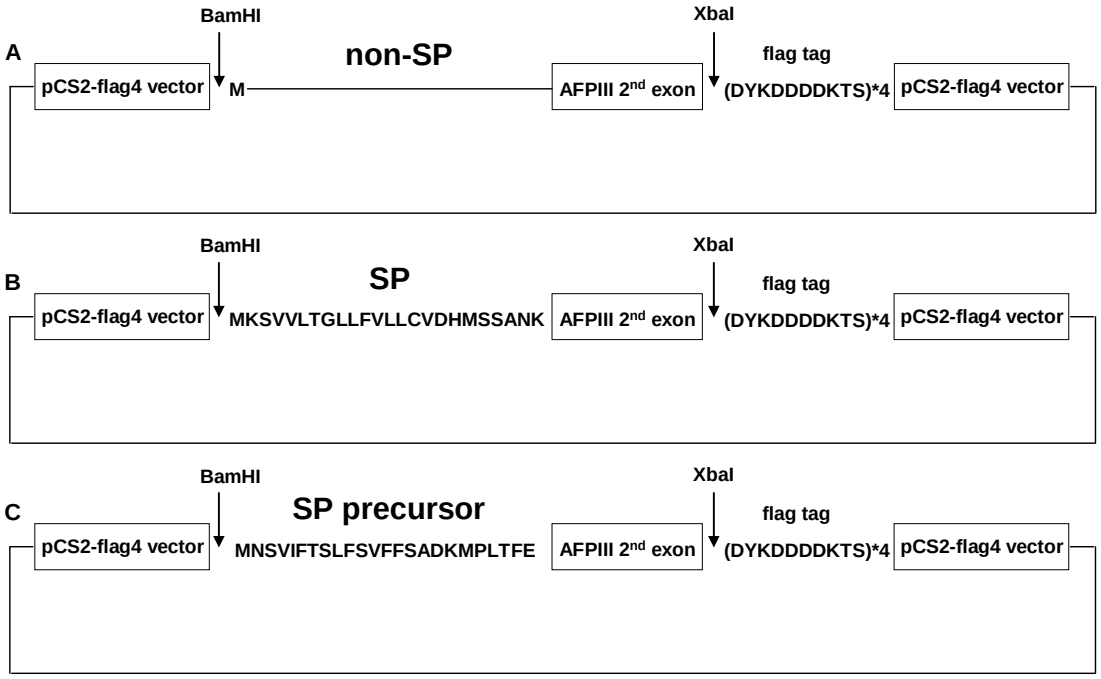
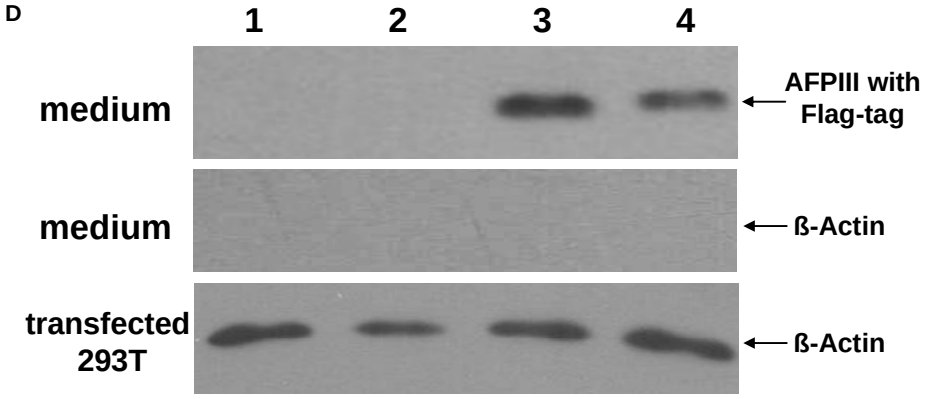


Fig. S8


A
BamHI
XbaI
non-SP
flag tag
pCS2-flag4 vector
M
AFPIII 2nd exon
(DYKDDDDKTS)*4
pCS2-flag4 vector
B
BamHI
XbaI
SP
flag tag
pCS2-flag4 vector
MKSVVLTGLLFVLLCVDHMSSANK
AFPIII 2nd exon
(DYKDDDDKTS)*4
pCS2-flag4 vector
C
BamHI
XbaI
SP precursor
flag tag
pCS2-flag4 vector
MNSVIFTSLFSVFFSADKMPLTFE
AFPIII 2nd exon
(DYKDDDDKTS)*4
pCS2-flag4 vector


D
1
2
3
4
medium
AFPIII with Flag-tag
medium
ß-Actin
transfected 293T
ß-Actin