Supporting Information

Synthesis of bioengineered heparin by recombinant yeast *Pichia pastoris*

Yonglin Zhang,[a,c] Yang Wang,[a,c] Zhengxiong Zhou,[a,c] Peilin Wang,[a,c] Xintong Xi,[a,c] Shan Hu,[a,c] RuiRui Xu,[a,c] Guocheng Du,[a,c] Jianghua Li,[a,c*] Jian Chen,[b,c] Zhen Kang[a,b,c*]

[a]The Key Laboratory of Carbohydrate Chemistry and Biotechnology, Ministry of Education, Jiangnan University, Wuxi, China.

[b]The Science Center for Future Foods, Jiangnan University, Wuxi, 214122, China.

[c]The key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi, China.

*Corresponding authors

Zhen Kang or Jianghua Li

Email: zkang@jiangnan.edu.cn; lijianghua@jiangnan.edu.cn

T: (86)-510-85918307, F: (86)-510-85918309

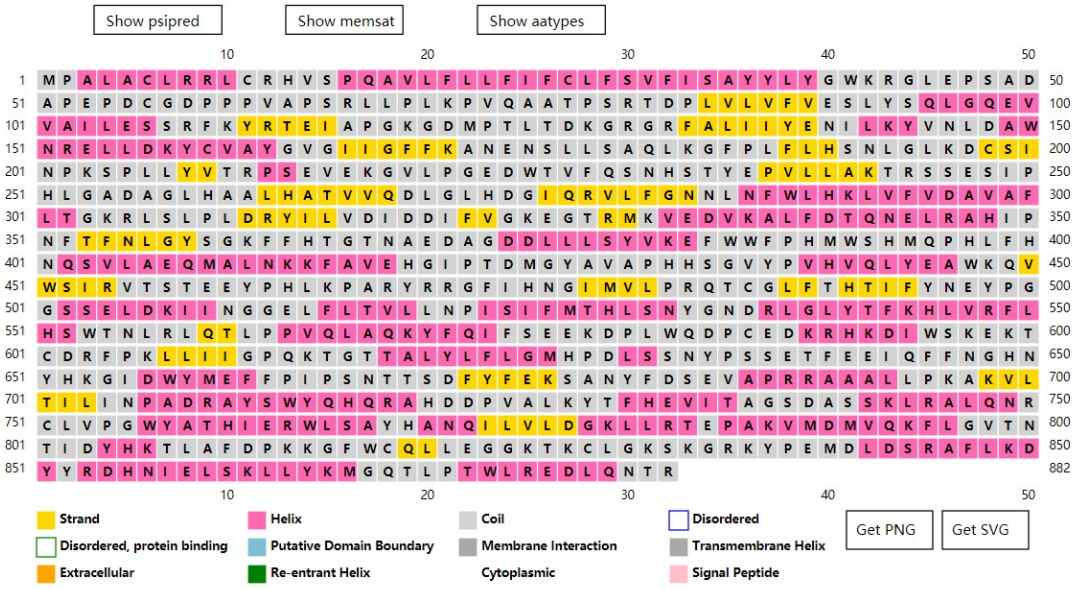

Fig. S1.

Secondary structure prediction of NDST1 using the Protein Structure Prediction Server (PSIPRED)[1]. N-terminal signal peptide region was indicated with green line.

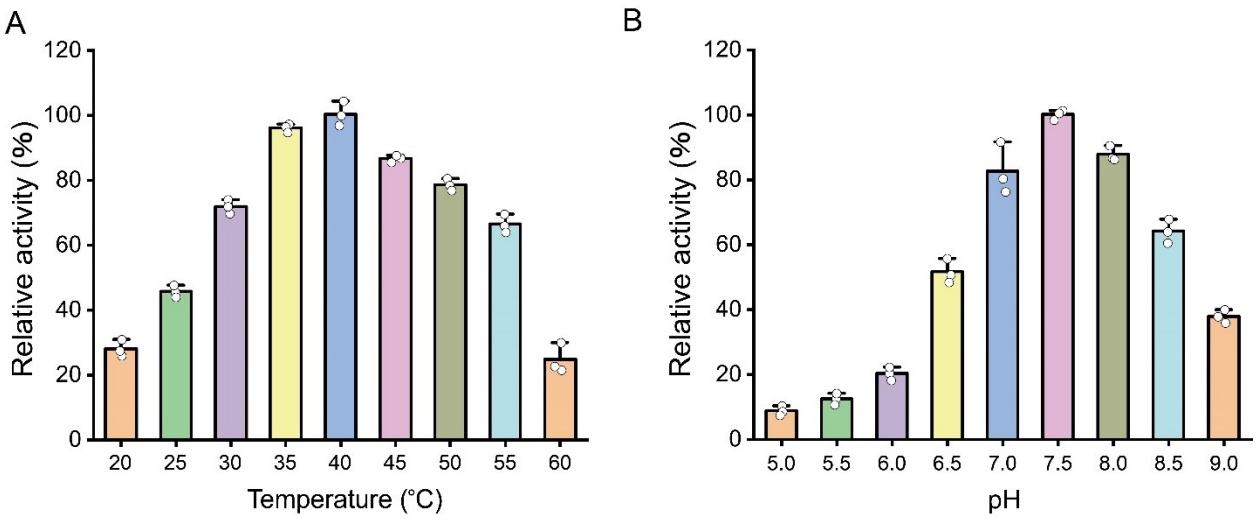

Fig. S2.

Effect of temperature and pH on NDST activity. (**A**) The relative activity of NDST at different temperatures. (**B**) The relative activity of NDST at different pH. All the data are expressed as the mean ± S.D. from three (n=3) biologically independent replicates.

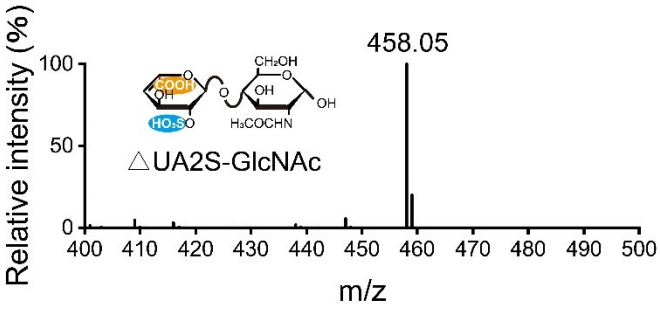

Fig. S3.

HPLC-MS analysis of the ΔUA2S-GlcNAc (*m/z* 458.05 $[M-H]^-$) structure.

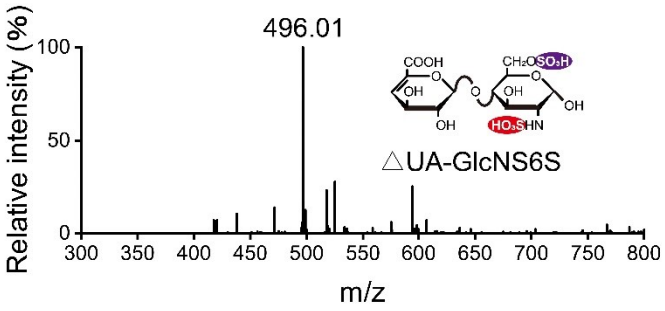

Fig. S4.

HPLC-MS analysis of the ΔUA-GlcNS6S (*m/z* 496.01 $[M-H]^-$) structure.

Fig. S5.

Standard curves of eight different disaccharide calibrants. Linear equation of eight disaccharide standards are listed in the table at the right of the figure.

Table S1. Molecular weights of bioengineered heparin in this work.

Bioengineered heparin	M_W (kDa)
Cell-free system	68.4±10.5
Engineered *P. pastoris*	349.1±11.7

Note: M_W is average molecular weight. All the data are expressed as the mean ± S.D. from three (n=3) biologically independent replicates.

References

1. L. J. McGuffin, K. Bryson and D. T. Jones, *Bioinformatics*, 2000, **16**, 404-405.

Table S2.

Strains used in this study.

Strains	Feature	Source
E. coli 109	*Escherichia coliI* K12 JM109	NEB
E. coli BL21	*E. coli* BL21(DE3)	Invitrogen
S. cerevisiae	*S. cerevisiae* CEN.PKI-IC	Invitrogen
P.pastoris	*P.pastoris* GS115	Invitrogen
Ec01	*E. coli BL21* derivative*, E. coli* BL21(DE3) harboring plasmid pET28a-*hNDST1*$^{\Delta N43}$, Ampicillin resistance	This study
Ec02	*E. coli BL21* derivative, *E. coli* BL21(DE3) harboring plasmid pET28a-*mNDST1*$^{\Delta N43}$, Ampicillin resistance	This study
Sc01	*S. cerevisiae* derivative, *S. cerevisiae* CEN.PKI-IC-IVM harboring plasmid pY26-TEF/GPD-*mNDST1*$^{\Delta N43}$-6His	This study
Sc02	*S. cerevisiae* derivative, *S. cerevisiae* CEN.PKI-IC-IVM harboring plasmid pY26-TEF/GPD-*hNDST1*$^{\Delta N43}$-6His	This study
Pp01	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*hNDST1*$^{\Delta N43}$ (SalI)	This study
Pp02	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*mNDST1*$^{\Delta N43}$ (SalI)	This study
Pp03	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*hNDST1*$^{\Delta N83}$ (SalI)	This study
Pp04	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*hNDST1*$^{\Delta N91}$ (SalI)	This study
Pp05	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*hNDST1*$^{\Delta N95}$ (SalI)	This study
Pp06	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*MBP-hNDST1*$^{\Delta N83}$ (SalI)	This study
Pp07	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*SUMO-hNDST1*$^{\Delta N83}$ (SalI)	This study
Pp08	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*TrxA-hNDST1*$^{\Delta N83}$ (SalI)	This study
Pp09	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*AP2-hNDST1*$^{\Delta N83}$ (SalI)	This study

Pp10	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*PDI-hNDST1*$^{\Delta N83}$ (SalI)	This study
Pp11	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*C5 epi* (SalI)	This study
Pp12	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*2-OST* (SalI)	This study
Pp13	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*6-OST1* (SalI)	This study
Pp14	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pPIC3.5K-*3-OST1* (SalI)	This study
Pp15	*P. pastoris* derivative*, P. pastoris* GS115 harboring plasmid pGAP815-*kfiA'-T2A1-kfiC'-T2A2-Ppcugd* (SalI)	This study
Pp16	*P. pastoris* derivative, *P. pastoris GS115 P. pastoris* GS115 harboring plasmid pAO815-*kfiA'-T2A1-kfiC'-T2A2-Ppcugd* (SalI)	This study
Pp17	*P. pastoris* GS115 harboring plasmid pGAP815-*kfiA'-T2A1-kfiC'-T2A2-Ecugd* (SalI)	This study
Pp18	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pAO815-*kfiA'-T2A1-kfiC'-T2A2-Ecugd* (SalI)	This study
Pp19	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pGAP815-*kfiA'-T2A1-kfiC'-T2A2-BsytdA* (SalI)	This study
Pp20	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pAO815-*kfiA'-T2A1-kfiC'-T2A2-BsytdA* (SalI)	This study
Pp21	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pGAP815-*kfiA'-T2A1-kfiC'-T2A2-CgugdA2* (SalI)	This study
Pp22	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pAO815-*kfiA'-T2A1-kfiC'-T2A2-CgugdA2* (SalI)	This study
Pp23	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pGAP815-*kfiA'-T2A1-kfiC'-T2A2-BstuaD* (SalI)	This study
Pp24	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pAO815-*kfiA'-T2A1-kfiC'-T2A2-BstuaD* (SalI)	This study
Pp25	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pGAP815-*kfiA'-T2A1-kfiC'-T2A2-BstuaD'* (SalI)	This study
Pp26	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmid pAO815-*kfiA'-T2A1-kfiC'-T2A2-BstuaD'* (SalI)	This study

Pp27	*P. pastoris* derivative, *P. pastoris* GS115 harboring plasmids pAO815-*kfiA'-T2A1-kfiC'-T2A2-BstuaD'-T2A5-MBPhNDST1*$^{\Delta N83}$ (SalI) and pPIC3.5k-*C5 epi-T2A5-2-OST-T2A1-6-OST1-T2A2-3-OST1* (SacI)	This study

Table S3.

Plasmids used in this study.

Plasmid	Feature	Source
pET-28a (+)	*E. coli* expression vector, Kan^R	Invitrogen
pY26-TEF/GPD	*S. cerevisiae* expression vector, Amp^R	Invitrogen
pAO815	*P. pastoris* expression vector, AOX promoter	Lab stock
pGAP815	*P. pastoris* expression vector, GAP promoter	This study
pPIC3.5K	*P. pastoris* expression vector, AOX promoter	Lab stock
pPIC3.5KHy	*P. pastoris* expression vector, AOX promoter, inserted Hygr cassette	Lab stock
pET28a-*hNDST1$^{\Delta N43}$*	pET28a plasmid containing gene *hNDST1$^{\Delta N43}$*	This study
pET28a-*mNDST1$^{\Delta N43}$*	pET28a plasmid containing gene *mNDST1$^{\Delta N43}$*	This study
pY26-TEF/GPD-*hNDST1$^{\Delta N43}$*	pY26 plasmid containing gene *hNDST1$^{\Delta N43}$*, inserted between GDP promoter and CYC1 terminator	This study
pY26-TEF/GPD-*mNDST1$^{\Delta N43}$*	pY26 plasmid containing gene *mNDST1$^{\Delta N43}$*, inserted between GDP promoter and CYC1 terminator	This study
pPIC3.5K-*hNDST1$^{\Delta N43}$*	pPIC3.5K plasmid containing gene *hNDST1$^{\Delta N43}$*	This study
pPIC3.5K-*mNDST1$^{\Delta N43}$*	pPIC3.5K plasmid containing gene *mNDST1$^{\Delta N43}$*	This study
pY26-TEF/GPD-*hNDST1$^{\Delta N43}$*-6His	pY26 plasmid containing gene *hNDST1$^{\Delta N43}$* and 6x histidine-tag codons, inserted between GDP promoter and CYC1 terminator	This study
pY26-TEF/GPD-*mNDST1$^{\Delta N43}$*-6His	pY26 plasmid containing gene *mNDST1$^{\Delta N43}$* and 6x histidine-tag codons, inserted between GDP promoter and CYC1 terminator	This study

pPIC3.5K-*hNDST1$^{\Delta N43}$*-6His	pPIC3.5K plasmid containing gene *hNDST1$^{\Delta N43}$* and 6x histidine-tag codons	This study
pPIC3.5K-*mNDST1$^{\Delta N43}$*-6His	pPIC3.5K plasmid containing gene *mNDST1$^{\Delta N43}$* and 6x histidine-tag codons	This study
pPIC3.5K-*hNDST1$^{\Delta N83}$*-6His	pPIC3.5K plasmid containing gene *hNDST1$^{\Delta N83}$* and 6x histidine-tag codons	This study
pPIC3.5K-*hNDST1$^{\Delta N91}$*-6His	pPIC3.5K plasmid containing gene *mNDST1$^{\Delta N91}$* and 6x histidine-tag codons	This study
pPIC3.5K-*hNDST1$^{\Delta N95}$*-6His	pPIC3.5K plasmid containing gene *hNDST1$^{\Delta N95}$* and 6x histidine-tag codons	This study
pPIC3.5K-*MBP-hNDST1$^{\Delta N83}$*-6His	pPIC3.5K plasmid containing gene *MBP-hNDST1$^{\Delta N83}$* and 6x histidine-tag codons	This study
pPIC3.5K-*SUMO-hNDST1$^{\Delta N83}$*-6His	pPIC3.5K plasmid containing gene *SUMO-hNDST1$^{\Delta N83}$* and 6x histidine-tag codons	This study
pPIC3.5K-*TrxA-hNDST1$^{\Delta N83}$*-6His	pPIC3.5K plasmid containing gene *TrxA-hNDST1$^{\Delta N83}$* and 6x histidine-tag codons	This study
pPIC3.5K-*AP2-hNDST1$^{\Delta N83}$*-6His	pPIC3.5K plasmid containing gene *AP2-hNDST1$^{\Delta N83}$* and 6x histidine-tag codons	This study
pPIC3.5K-*PDI-hNDST1$^{\Delta N83}$*-6His	pPIC3.5K plasmid containing gene *PDI-hNDST1$^{\Delta N83}$* and 6x histidine-tag codons	This study
pPIC3.5k-*C5*	pPIC3.5K containing gene *C5-epi* and 6x histidine-tag codons	This study

epi-6His		
pPIC3.5k-*2-OST*-6His	pPIC3.5K containing gene *2-OST* and 6x histidine-tag codons	This study
pPIC3.5k-*6-OST1*-6His	pPIC3.5K containing gene *6-OST1* and 6x histidine-tag codons	This study
pPIC3.5k-*3-OST1*-6His	pPIC3.5K containing gene 3-OST1 and 6x histidine-tag codons	This study
pAO815-*kfiA'-T2A1-kfiC'-T2A2-BstuaD*	pAO815 containing gene *kfiA*, *kfiC and BstuaD* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pGAP815-*kfiA'-T2A1-kfiC'-T2A2-BstuaD*	pGAP815 containing gene *kfiA*, *kfiC and BstuaD* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pAO815-*kfiA'-T2A1-kfiC'-T2A2-BsytdA*	pAO815 containing gene *kfiA*, *kfiC and BsytdA* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pAO815-*kfiA'-T2A1-kfiC'-T2A2-BsytdA*	pGAP815 containing gene *kfiA*, *kfiC and BsytdA* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pAO815-*kfiA'-T2A1-kfiC'-T2A2-CgugdA2*	pAO815 containing gene *kfiA*, *kfiC and CgugdA2* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pGAP815-*kfiA'-T2A1-kfiC'-T2A2-CgugdA2*	pGAP815 containing gene *kfiA*, *kfiC and CgugdA2* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pAO815-*kfiA'-T2A1-kfiC'-T2A2-Ecugd*	pAO815 containing gene *kfiA*, *kfiC and Ecugd* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pGAP815-*kfiA'-T2A1-kfiC'-T2A2-Ecugd*	pGAP815 containing gene *kfiA*, *kfiC and Ecugd* expression cassette which assembled by T2A1 and T2A2 peptide	This study

pAO815-*kfiA′-T2A1-kfiC′-T2A2-Ppcugd*	pAO815 containing gene *kfiA*, *kfiC and Ppcugd* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pGAP815-*kfiA′-T2A1-kfiC′-T2A2-Ppcugd*	pGAP815 containing gene *kfiA*, *kfiC and Ppcugd* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pGAP815-*kfiA′-T2A1-kfiC′-T2A2-BstuaD′*	pGAP815 containing gene *kfiA*, *kfiC and BstuaD′* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pAO815-*kfiA′-T2A1-kfiC′-T2A2-BstuaD′*	pAO815 containing gene *kfiA*, *kfiC and BstuaD′* expression cassette which assembled by T2A1 and T2A2 peptide	This study
pAO815-*kfiA′-T2A1-kfiC′-T2A2-BstuaD′-T2A5-MBPhNDST1$^{\triangle N83}$*	pAO815 containing gene *kfiA*, *kfiC*, *tuaD* and *MBPhNDST1*$^{\triangle N83}$ expression cassette which assembled by T2A3, T2A2 and T2A5 peptide	This study
pPIC3.5k-*C5 epi-T2A5-2-OST-T2A1-6-OST1-T2A2-3-OST1*	pPIC3.5K containing gene *C5-epi*, *2-OST*, *6-OST1 and 3-OST1* expression cassette which assembled by T2A5, T2A1 and T2A2 peptide	This study

Table S4.

Primers used in this study.

Primers	Sequence (5' to 3')
pET28a-hNDST1$^{\Delta N43}$-F	TTAACTTTAAGAAGGAGATATACCATGGGTTTGGAACCTTCTGCTGATGCTC
pET28a-hNDST1$^{\Delta N43}$-R	CTCAGCTTCCTTTCGGGCTTTGTTATCTAGTATTTTGCAAATCCTCTCTCAACCAAGTTGGC
pET28a-mNDST1$^{\Delta N43}$-F	AACTTTAAGAAGGAGATATACCATGGGTTTGGAACCTTCTGCTGATGCTTC
pET28a-mNDST1$^{\Delta N43}$-R	TCAGCTTCCTTTCGGGCTTTGTTATCTAGTATTTTGCAAATCCTCTCTCAACCAAGTTG
pY26-hNDST1$^{\Delta N43}$-F	GACGGATTCTAGAACTAGTGGATCATGGGTTTGGAACCTTCTGCTGATGCTC
pY26-hNDST1$^{\Delta N43}$-R	ATTACATGACTCGAGGTCGACGGTATCGTTATCTAGTATTTTGCAAATCCTCTCTCAACCAAGTTGGC
pY26-mNDST1$^{\Delta N43}$-F	GACGGATTCTAGAACTAGTGGATCATGGGTTTGGAACCTTCTGCTGATGCTTC
pY26-mNDST1$^{\Delta N43}$-R	ATTACATGACTCGAGGTCGACGGTATCGTTATCTAGTATTTTGCAAATCCTCTCTCAACCAAGTTG
pPIC3.5k-hNDST1$^{\Delta N43}$-F	TTATTAGCTTACTTTCATAATTGCATGGGTTTGGAACCTTCTGCTGATGCTC
pPIC3.5k-hNDST1$^{\Delta N43}$-R	CGAATTAATTCGCGGCCGCTTATTATCTAGTATTTTGCAAATCCTCTCTCAACCAAGTTGGC

pPIC3.5k-mNDST1$^{\Delta N43}$-F	TTATTAGCTTACTTTCATAATTGCATGGGTTTGGAACCTTCTGCTGATGCTTC
pPIC3.5k-mNDST1$^{\Delta N43}$-R	CGAATTAATTCGCGGCCGCTTATTATCTAGTATTTTGCAAATCCTCTCTCAAC CAAGTTG
pPIC3.5k-hNDST1-R	CATGGTGGCCGTTTGGATCCTTCGAATAATTAGTTG
pPIC3.5k-hNDST1$^{\Delta N83}$-F	CGAAGGATCCAAACGGCCACCATGCCATTGGTTTTGGTTTTCGTTGAATCTTT GTAC
pPIC3.5k-hNDST1$^{\Delta N91}$-F	CGAAGGATCCAAACGGCCACCATGTCTTTGTACTCTCAATTGGGTCAAGAAG TTG
pPIC3.5k-hNDST1$^{\Delta N95}$-F	CGAAGGATCCAAACGGCCACCATGCAATTGGGTCAAGAAGTTGTTGCTATTT TGG
MBP-hNDST1$^{\Delta N83}$-F	TTATTAGCTTACTTTCATAATTGCATGAAAATCGAAGAAGGTAAACTGGTAAT CTGG
MBP-hNDST1$^{\Delta N83}$-R	AGAAGGTTCCAAACCTCTCATGGATCCACCTCCAGTCTGCGCGTCTTTCAGGG CTTCATC
SUMO-hNDST1$^{\Delta N83}$-F	TTATTAGCTTACTTTCATAATTGCATGTCGGACTCAGAAGTCAATCAAGAAG
SUMO-hNDST1$^{\Delta N83}$-R	AGAAGGTTCCAAACCTCTCATGGATCCACCTCCACCACCAATCTGTTCTCTGT GAGCCTC
TrxA-hNDST1$^{\Delta N83}$-F	TTATTAGCTTACTTTCATAATTGCATGAGCGATAAAATTATTCACCTGACTGA CG
TrxA-	AGAAGGTTCCAAACCTCTCATGGATCCACCTCCGGCCAGGTTAGCGTCGAGG

hNDST1$^{\Delta N}$ ^{83}-R	AACTCTTTC
AP2-hNDST1$^{\Delta N}$ ^{83}-F	AAGCTAAAGCTAAAGCTGAAGCTGAAGCTAAAGCTAAAATGAGAGGTTTGG AACCTTCTGCTGATGC
AP2-hNDST1$^{\Delta N}$ ^{83}-R	CTTCAGCTTCAGCTTTAGCTTTAGCTTCAGCTTCAGCGGTGGCCGTTTGGATC CTTCGAATAATTAGTTG
PDI-hNDST1$^{\Delta N}$ ^{83}-F	ATTATTCGAAGGATCCAAACGGCCACCATGCAATTCAACTGGAATATTAAAA CTGTGG
PDI-hNDST1$^{\Delta N}$ ^{83}-R	AGAAGGTTCCAAACCTCTCATGGATCCACCTCCAAGCTCGTCGTGAGCGTCT GCCTCAC
pPIC3.5k-C5epi-F	GTCGCGGATCCGCCACCATGAGATGTTTGGCTGCTAGAGTTA
pPIC3.5k-C5epi-R	TGGCGGCCGCTTAATTATGTTTAGCTCTAGAACCTTTCAAATAAGATTTC
pPIC3.5k-2OST-F	GTCGCGGATCCGCCACCATGGATGGTCCTAGACAAGAAGTTG
pPIC3.5k-2OST-R	TGGCGGCCGCTTAATTAGATTTAGGATAAATTTTTTCATAGAAGAAA
pPIC3.5k-6OST1-F	GTCGCGGATCCGCCACCATGGCTTTCGATATGAAAGGTGAAGATG
pPIC3.5k-6OST1-R	GTGGCGGCCGCTTAGTCCACTTTTCAATAATATGAGACATGTAATCTTCAG
pPIC3.5k-3OST1-F	GTCGCGGATCCGCCACCATGAAGGGTGGTACTAGAGCTTTGTTGG
pPIC3.5k-3OST1-R	GTGGCGGCCGCTTAATGCCAATCAAAAGTTCTACCAACC
kfiA'-F	CTAATTATTCGAAACGAGGAATTCGCCACCATGATTGTTGCAAATATGTCATC
kfiA'-R	GTCACCGCAAGTAAGCAAAGAACCTCTACCCTCTCCGGATCCCCCTTCCACAT TATACACTAATTCGAG

kfiC'-F	GTTCTTTGCTTACTTGCGGTGACGTTGAGGAAAACCCAGGTCCAATGAACGC AGAATATATAAATTTAGTTGAAC
kfiC'-R	CGTCACCGCAAGTAAGTAGGGATCCACGACCTTCTCCGGATCCTTGTTCAATT ATTCCTGATACATCTTTAAAC
BstuaD-F	TCCCTACTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCGATGAAGAAAA TTGCTGTTATTGGTACTGG
BstuaD-R	GTCTCCACATGTAAGAAGTGATCCCCTACCCTCGGCACGTCCGGATCCCAAA TTAACAGAACCCAAATCTTTTGCCAA
BstuaD'-F	TCCCTACTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCGATGATGAAGA AAATTGCTGTTATTGGTACTGG
BstuaD'-R	GTCTCCACATGTAAGAAGTGATCCCCTACCCTCGGCACGTCCGGATCCTTACA AATTAACAGAACCCAAATCTTTTGCCAATTC
BsytdA-F	TCCCTACTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCGATGATAAAAA AAGCGATCATTCCTGCAGG
BsytdA-R	GTCTCCACATGTAAGAAGTGATCCCCTACCCTCGGCACGTCCGGATCCTTGGG GGTTCTTTTCCGTTTTCAGCATATCG
CgugdA2-F	TCCCTACTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCGATGCGGATGA CAGTGATTGGTACGGGTTACCTTGGCGCTACG
CgugdA2-R	GTCTCCACATGTAAGAAGTGATCCCCTACCCTCGGCACGTCCGGATCCCTAA AGGTTGCGGCCGAGCGCTTCCATTTCCCAACCG
Ecugd-F	TCCCTACTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCGATGAAAATCA CCATTTCCGGTACTGG
Ecugd-R	GTCTCCACATGTAAGAAGTGATCCCCTACCCTCGGCACGTCCGGATCCGTCGC TGCCAAAGAGATCGCGGGTG
Ppcugd-F	TCCCTACTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCGATGAAGGTCA CGGTTTTCGGAACCGG
Ppcugd-R	GTCTCCACATGTAAGAAGTGATCCCCTACCCTCGGCACGTCCGGATCCAGCT GGCGCAATCTTGCCACGTC
MBPhNDS T1$^{\triangle N83}$-F	ATCACTTCTTACATGTGGAGACGTCGAGGAGAACCCTGGTCCAATGAGAGGT TTGGAACCTTCTGCTGATG
MBPhNDS T1$^{\triangle N83}$-R	AGTCATGTCTAAGGCGAATTCTTATCTAGTATTTTGCAAATCCTCTCTCAACC AAG
C5epi-F	AACTAATTATTCGAAGGATCCGCCACCATGAGATGTTTGGCTGCTGGTGTTCA

C5epi-R	ACATGTAAGAAGTGATCCCCTACCCTCGGCACGTCCGGATCCATTGTGCTTAG CTCTACCACCTTTCAAG
2OST-F	AGGGGATCACTTCTTACATGTGGAGACGTCGAGGAGAACCCTGGTCCAATGG ATGGTCCTAGACAAGAAGTTGC
2OST-R	CTTCAACATCACCACAAGTCAGAAGAGAACCACGACCTTCTCCGGATCCATT AGATTTAGGATAAATTTTTTCATAGAAGAAATTTTG
6OST1-F	TGACTTGTGGTGATGTTGAAGAAAACCCAGGTCCAATGGCTTTCGATATGAA AGGTGAAG
6OST1-R	CTCTACGTCACCGCAAGTAAGTAGGGATCCACGACCTTCTCCGGATCCCCACT TTTCAATAATATGAGACATGTAATC
3OST1-F	CTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCGATGAAGGGTGGTACTA GAGCTTTG
3OST1-R	TTAATGCCAATCAAAAGTTCTACCAACCAATTTAAAAAAAGGCGAATTAATT CGCGGCCGC

Table S5.

DNA sequences used in this study.

T2A sequence	Coding sequence 5' to 3'
T2A1	GAGGGTAGAGGTTCTTTGCTTACTTGCGGTGACGTTGAGGAAAACCCAGGTCCA
T2A2	GAAGGTCGTGGATCCCTACTTACTTGCGGTGACGTAGAGGAAAACCCTGGTCCG
T2A5	GAGGGTAGGGGATCACTTCTTACATGTGGAGACGTCGAGGAGAACCCTGGTCCA