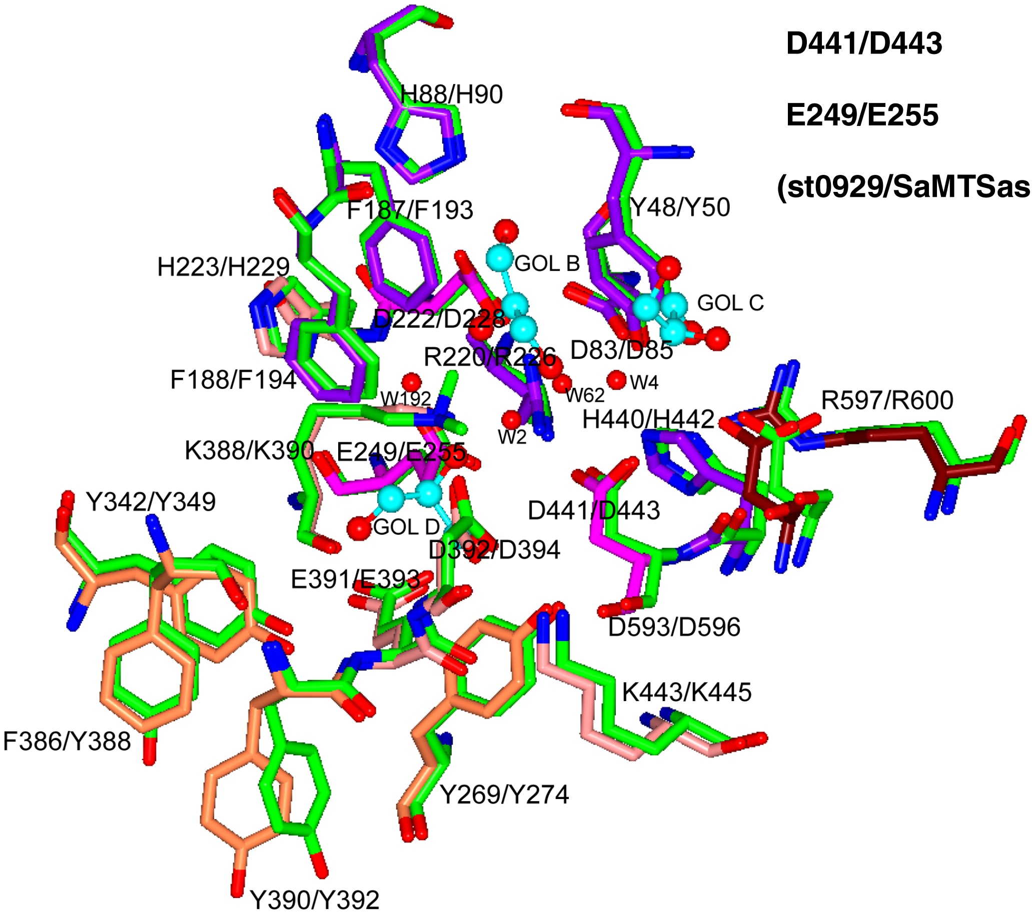
Supplementary Figure 1 Superimposed st0929 and SaMTSase (green) active site residues. The st0929 putative catalytic triad, subsites +1, -1, -2 residues are depicted in magenta, coral, violet and brown, respectively.



Catalytic triad:

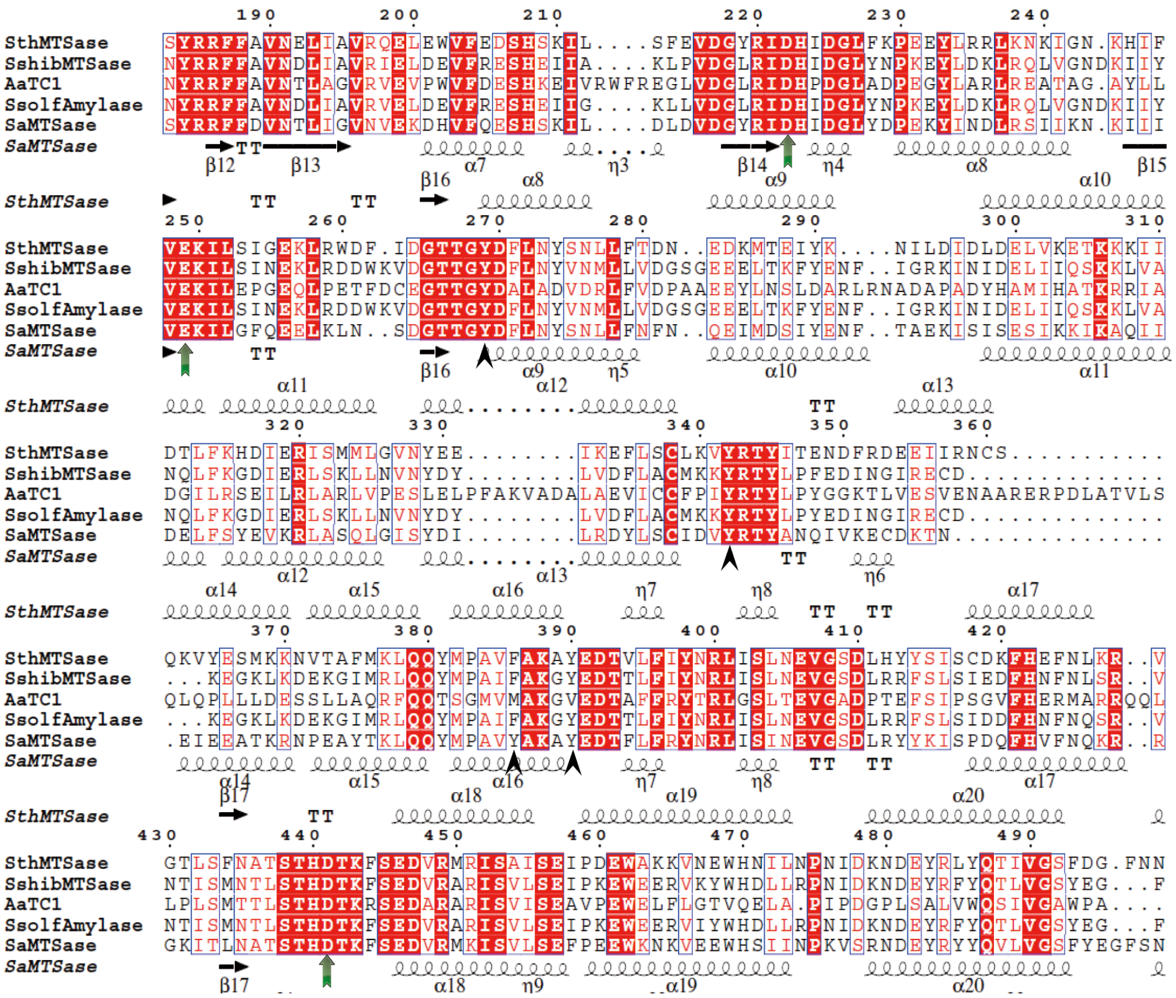
D222/D228

D441/D443

E249/E255

(st0929/SaMTSase)

Supplementary Figure 2 Sequence Alignment of known MTSases. The abbreviations SthMTSase, SshibMTSase, AaTC1, SsolfAmylase, SaMTSase represent st0929 protein, *Sulfolobus shibatae* MTSase, *Arthrobacter aurescens* strain TC1 MTSase, *Sulfolobus solfataricus* glycosyl-trehalose producing enzyme, and *Sulfolobus acidocaldarius* MTSase (PDB code: 1IV8) respectively. The black arrows indicate the residues comprising the tyrosine cluster (Y269, Y342, F386, Y390 in st0929) in subsite +1. The green arrows indicate the catalytic triad residues (D222, E249, D441 in st0929). The sequence alignment was carried out using *Blast* (Altschul, S.F., Madden T.L., Schaffer A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997). *Nucleic Acids Res.* 25, 3389-3402; Kabsch, W. & and Sander, C. (1983). *Biopolymers.* **22**, 2577-2637). This figure was produced using CLUSTAL W.



```
                    190       200       210           220       230       240
SthMTSase     SYRRFFAVNELIAVRQELEWVFEDSHSKIL....SFEVDGYRIDHIDGLFKPEEYLRRLKNKIGN.KHIF
SshibMTSase   NYRRFFAVNDLIAVRIELDEVFRESHEIIA....KLPVDGLRIDHIDGLYNPKEYLDKLRQLVGNDKIIY
AaTC1         NYRRFFAVNTLAGVRVEVPWVFDESHKEIVRWFREGLVDGLRIDHPDGLADPEGYLARLREATAG.AYLL
SsolfAmylase  NYRRFFAVNDLIAVRVELDEVFRESHEIIG....KLLVDGLRIDHIDGLYNPKEYLDKLRQLVGNDKIIY
SaMTSase      SYRRFFDVNTLIGVNVEKDHVFQESHSKIL....DLDVDGYRIDHIDGLYDPEKYINDLRSIIKN.KIII

              250       260       270       280       290       300       310
SthMTSase     VEKILSIGEKLRWDF.IDGTTGYDFLNYSNLLFTDN..EDKMTEIYK....NILDIDLDELVKETKKKII
SshibMTSase   VEKILSINEKLRDDWKVDGTTGYDFLNYVNMLLVDGSGEEELTKFYENF..IGRKINIDELIIQSKKLVA
AaTC1         VEKILEPGEQLPETFDCEGTTGYDALADVDRLFVDPAAEEYLNSLDARLRNADAPADYHAMIHATKRRIA
SsolfAmylase  VEKILSINEKLRDDWKVDGTTGYDFLNYVNMLLVDGSGEEELTKFYENF..IGRKINIDELIIQSKKLVA
SaMTSase      VEKILGFQEELKLN..SDGTTGYDFLNYSNLLFNFN..QEIMDSIYENF..TAEKISISESIKKIKAQII

                     320       330                 340       350       360
SthMTSase     DTLFKHDIERISMMLGVNYEE.........IKEFLSCLKVYRTYITENDFRDEEIIRNCS..........
SshibMTSase   NQLFKGDIERLSKLLNVNYDY.........LVDFLACMKKYRTYLPFEDINGIRECD.............
AaTC1         DGILRSEILRLARLVPESLELPFAKVADALAEVICCFPIYRTYLPYGGKTLVESVENAARERPDLATVLS
SsolfAmylase  NQLFKGDIERLSKLLNVNYDY.........LVDFLACMKKYRTYLPYEDINGIRECD.............
SaMTSase      DELFSYEVKRLASQLGISYDI.........LRDYLSCIDVYRTYANQIVKECDKTN..............

                     370       380       390       400       410       420
SthMTSase     QKVYESMKKNVTAFMKLQQYMPAVFAKAYEDTVLFIYNRLISLNEVGSDLHYYSISCDKFHEFNLKR..V
SshibMTSase   ...KEGKLKDEKGIMRLQQYMPAIFAKGYEDTTLFIYNRLISLNEVGSDLRRFSLSIEDFHNFNLSR..V
AaTC1         QLQPLLLDESSLLAQRFQQTSGMVMAKGVEDTAFFRYTRLGSLTEVGADPTEFSIPSGVFHERMARRQQL
SsolfAmylase  ...KEGKLKDEKGIMRLQQYMPAIFAKGYEDTTLFIYNRLISLNEVGSDLRRFSLSIDDFHNFNQSR..V
SaMTSase      .EIEEATKRNPEAYTKLQQYMPAVYAKAYEDTFLFRYNRLISINEVGSDLRYYKISPDQFHVFNQKR..R

          430       440       450       460       470       480       490
SthMTSase     GTLSFNATSTHDTKFSEDVRMRISAISEIPDEWAKKVNEWHNILNPNIDKNDEYRLYQTIVGSFDG.FNN
SshibMTSase   NTISMNTLSTHDTKFSEDVRARISVLSEIPKEWEERVKYWHDLLRPNIDKNDEYRFYQTLVGSYEG...F
AaTC1         LPLSMTTLSTHDTKRSEDARARISVISEAVPEWELFLGTVQELA.PIPDGPLSALVWQSIVGAWPA....
SsolfAmylase  NTISMNTLSTHDTKFSEDVRARISVLSEIPKEWEERVIYWHDLLRPNIDKNDEYRFYQTLVGSYEG...F
SaMTSase      GKITLNATSTHDTKFSEDVRMKISVLSEFPEEWKNKVEEWHSIINPKVSRNDEYRYYQVLVGSFYEGFSN
```