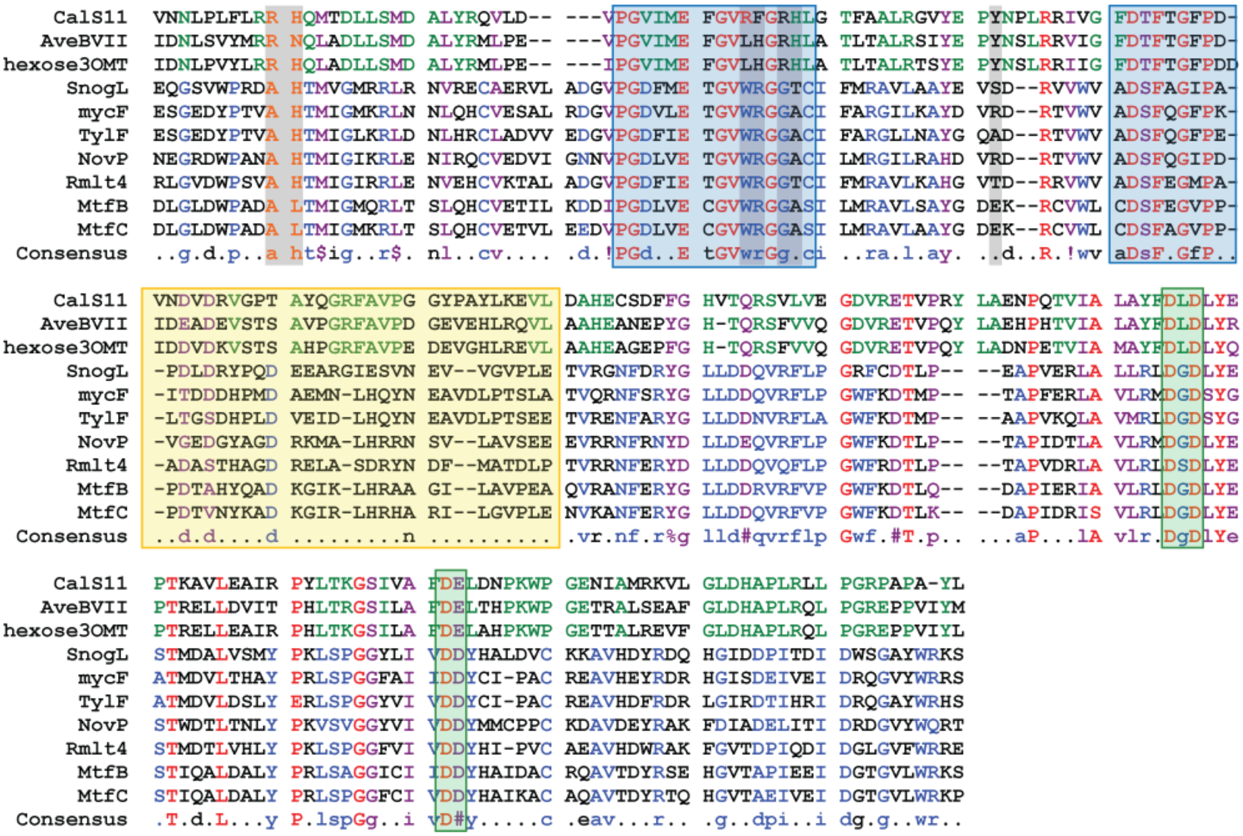
Supporting Information for:

Structural and functional characterization of Cals S11, a TDP-rhamnose 3'-O-methyltransferase involved in calicheamicin biosynthesis.

Shanteri Singh[1], Aram Chang[2], Kate E. Helmich[2], Craig A. Bingman[2], Russel L. Wrobel[2], Emily T. Beebe[2], Shin-Ichi Makino[2], David J. Aceti[2], Kevin Dyer[3], Greg L. Hura[3], Manjula Sunkara[4], Andrew J. Morris[4], George N. Phillips Jr.[2,5*], and Jon S. Thorson[1*]

[1]Center for Pharmaceutical Research and Innovation, University of Kentucky College of Pharmacy, 789 South Limestone Street, Lexington, KY 40536-0596; [2]Department of Biochemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA**;** [3]Physical Bioscience Division, Lawrence Berkeley National Laboratory, Berkeley CA 94720, USA; [4]Division of Cardiovascular Medicine, University of Kentucky, Lexington, KY 40536; [5]Department of Biochemistry and Cell Biology, Rice University, Houston, Texas 77005, USA**;** [*]Correspondence to Jon S Thorson, (e-mail: jsthorson@uky.edu) or George N. Phillips, Jr. (e-mail: georgep@rice.edu)

Table of Contents



```
     CalS11  VNNLPLFLRR HQMTDLLSMD ALYRQVLD-- ---VPGVIME FGVRFGRHLG TFAALRGVYE PYNPLRRIVG FDTFTGFPD-
    AveBVII  IDNLSVYMRR NQLADLLSMD ALYRMLPE-- ---VPGVIME FGVLHGRHLA TLTALRSIYE PYNSLRRVIG FDTFTGFPD-
 hexose3OMT  IDNLPVYLRR HQLADLLSMD ALYRMLPE-- ---IPGVIME FGVLHGRHLA TLTALRTSYE PYNSLRRIIG FDTFTGFPDD
      SnogL  EQGSVWPRDA HTMVGMRRLR NVRECAERVL ADGVPGDFME TGVWRGGTCI FMRAVLAAYE VSD--RVVWV ADSFAGIPA-
       mycF  ESGEDYPTVA HTMIGMKRLN NLQHCVESAL RDGVPGDVLE TGVWRGGACI FARGILKAYD VRD--RTVWV ADSFQGFPK-
       TylF  ESGEDYPTVA HTMIGLKRLD NLHRCLADVV EDGVPGDFIE TGVWRGGACI FARGLLNAYG QAD--RTVWV ADSFQGFPE-
       NovP  NEGRDWPANA HTMIGIKRLE NIRQCVEDVI GNNVPGDLVE TGVWRGGACI LMRGILRAHD VRD--RTVWV ADSFQGIPD-
      Rmlt4  RLGVDWPSVA HTMIGIRRLE NVEHCVKTAL ADGVPGDFIE TGVWRGGTCI FMRAVLKAHG VTD--RRVWV ADSFEGMPA-
       MtfB  DLGLDWPADA LTMIGMQRLT SLQHCVETIL KDDIPGDLVE CGVWRGGASI LMRAVLSAYG DEK--RCVWL CDSFEGVPP-
       MtfC  DLGLDWPADA LTMIGMKRLT SLQHCVETVL EEDVPGDLVE CGVWRGGASI LMRAVLAAYG DEK--RCVWL CDSFAGVPP-
  Consensus  ..g.d.p..a ht$ig..r$. nl..cv.... .d.!PGd..E tGVwrGg.ci ..ra.l.ay. ..d..R.!wv aDsF.GfP..

     CalS11  VNDVDRVGPT AYQGRFAVPG GYPAYLKEVL DAHECSDFFG HVTQRSVLVE GDVRETVPRY LAENPQTVIA LAYFDLDLYE
    AveBVII  IDEADEVSTS AVPGRFAVPD GEVEHLRQVL AAHEANEPYG H-TQRSFVVQ GDVRETVPQY LAEHPHTVIA LAYFDLDLYR
 hexose3OMT  IDDVDKVSTS AHPGRFAVPE DEVGHLREVL AAHEAGEPFG H-TQRSFVVQ GDVRETVPQY LADNPETVIA MAYFDLDLYQ
      SnogL  -PDLDRYPQD EEARGIESVN EV--VGVPLE TVRGNFDRYG LLDDQVRFLP GRFCDTLP-- --EAPVERLA LLRLDGDLYE
       mycF  -ITDDDHPMD AEMN-LHQYN EAVDLPTSLA TVQRNFSRYG LLDDQVRFLP GWFKDTMP-- --TAPFERLA VLRMDGDSYG
       TylF  -LTGSDHPLD VEID-LHQYN EAVDLPTSEE TVRENFARYG LLDDNVRFLA GWFKDTMP-- --AAPVKQLA VMRLDGDSYG
       NovP  -VGEDGYAGD RKMA-LHRRN SV--LAVSEE EVRRNFRNYD LLDEQVRFLP GWFKDTLP-- --TAPIDTLA VLRMDGDLYE
      Rmlt4  -ADASTHAGD RELA-SDRYN DF--MATDLP TVRRNFERYD LLDDQVQFLP GWFRDTLP-- --TAPVDRLA VLRLDSDLYE
       MtfB  -PDTAHYQAD KGIK-LHRAA GI--LAVPEA QVRANFERYG LLDDRVRFVP GWFKDTLQ-- --DAPIERIA VLRLDGDLYE
       MtfC  -PDTVNYKAD KGIR-LHRHA RI--LGVPLE NVKANFERYG LLDDQVRFVP GWFKDTLK-- --DAPIDRIS VLRLDGDLYE
  Consensus  ..d.d....d .........n .......... .vr.nf.r%g lld#qvrflp Gwf.#T.p.. ...aP...lA vlr.DgDlYe

     CalS11  PTKAVLEAIR PYLTKGSIVA FDELDNPKWP GENIAMRKVL GLDHAPLRLL PGRPAPA-YL
    AveBVII  PTRELLDVIT PHLTRGSILA FDELTHPKWP GETRALSEAF GLDHAPLRQL PGREPPVIYM
 hexose3OMT  PTRELLEAIR PHLTKGSILA FDELAHPKWP GETTALREVF GLDHAPLRQL PGREPPVIYL
      SnogL  STMDALVSMY PKLSPGGYLI VDDYHALDVC KKAVHDYRDQ HGIDDPITDI DWSGAYWRKS
       mycF  ATMDVLTHAY PRLSPGGFAI IDDYCI-PAC REAVHEYRDR HGISDEIVEI DRQGVYWRRS
       TylF  ATMDVLDSLY ERLSPGGYVI VDDYCI-PAC REAVHDFRDR LGIRDTIHRI DRQGAYWRHS
       NovP  STWDTLTNLY PKVSVGGYVI VDDYMMCPPC KDAVDEYRAK FDIADELITI DRDGVYWQRT
      Rmlt4  STMDTLVHLY PKLSPGGFVI VDDYHI-PVC AEAVHDWRAK FGVTDPIQDI DGLGVFWRRE
       MtfB  STIQALDALY PRLSAGGICI IDDYHAIDAC RQAVTDYRSE HGVTAPIEEI DGTGVLWRKS
       MtfC  STIQALDALY PRLSPGGFCI VDDYHAIKAC AQAVTDYRTQ HGVTAEIVEI DGTGVLWRKP
  Consensus  .T.d.L...y P.lspGg..i vD#y.....c .eav...r.. .g..dpi..i dg.g..wr..
```

Figure S1. Multiple sequence alignment of CalS11 using DELTA-BLAST. AveBVII is a TDP-6-deoxy-L-hexose 3-*O*-MT from *Streptomyces avermitilis* (avermectin); hexose 3OMT is a NDP-hexose 3-*O*-methyltransferase from *Streptomyces rochei*; SnogL is a nogalose *O*-MT from *Streptomyces nogalater* (nogalamycin); MycF is a mycinose 3-*O*-MT from *Micromonospora griseorubida* (mycinamycin II); TylF is a a mycinose 3-*O*-MT from *Streptomyces fradiae* (tylosin); NovP is a noviose 3-*O*-MT from *Streptomyces spheroids* (novobiocin); RmlT4, MtfB and MtfC are *O*-MTs from *Mycobacteria*. Residues in blue boxes are conserved motifs of MTs, which include a glycine-rich motif and a signature motif containing a conserved aspartate, both of which are key to cofactor binding; green boxes indicate conserved acidic residues involved in metal binding; the yellow box highlights the region within CalS11 and NovP which correlates to the flexible lid.

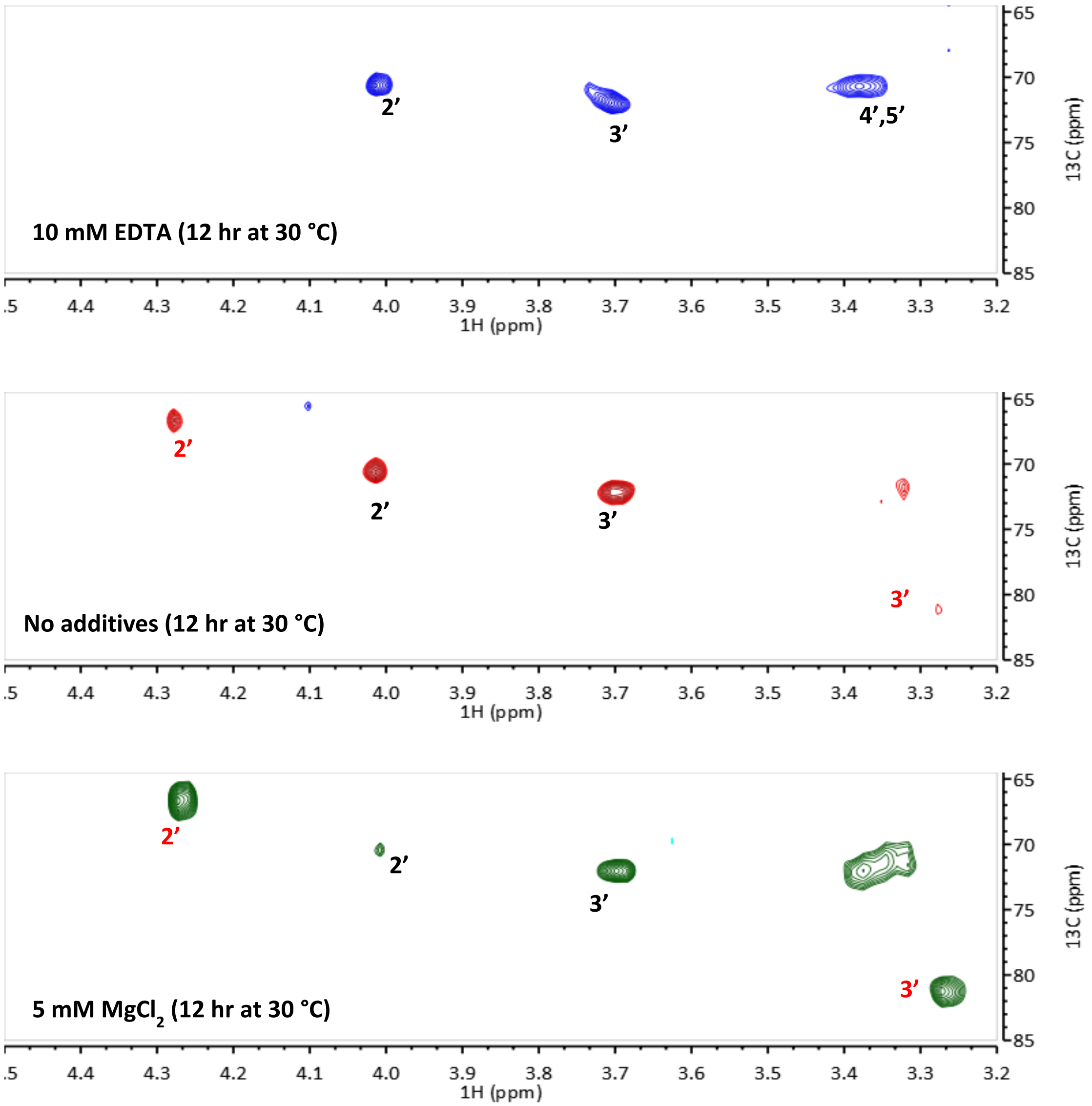


Figure S2. Metal dependency of CalS11. Region of ^{13}C-^{1}H HSQC spectrum where colors blue, red and green represent TDP-L-[*U*-^{13}C]rhamnose and the corresponding ^{13}C-methylated product, respectively within a standard reaction (middle panel), a standard reaction containing 10 mM EDTA (top panel) and a standard reaction containing 5 mM MgCl$_2$ (bottom panel).

Table S1. Proteins structurally-related to CalS11.

PDB	Enzyme	Organism	P-Score	Rmsd	L1[a]	L2[a]	%ID
2WK1	NOVP Noviose-4'-*O*-MT	*Streptomyces caeruleus*	8.97E-11	2.54	206	242	19
3SSO	MycE Mycinamycin-2'-*O*-MT	*Micromonospora griseorubida*	3.9E-5	3.01	206	221	9
3BXO	DesVI TDP-Desosamine-3'-*N*,*N*-diMT	*Streptomyces venezuelae*	6.13E-5	3.03	206	152	7
3GXO	MmcR Mitomycin-7-*O*-MT	*Streptomyces lavendulae*	1.18E-4	3.08	206	183	6
3BUS	RebM Rebeccamycin-4'-*O*-MT	*Lechevalieria aerocolonigenes*	2.85E-4	3.04	206	251	8
4E2Z	TcaB9 TDP-tetronitrose-3'-*C*-MT	*Micromonospora chalcea*	4.22E-4	3.12	206	290	9
1QZZ	RdmB Aclacinomycin-10-hydroxylase	*Streptomyces purpurascens*	0.0011	3.24	206	250	6
3LST	CalO1 Orsellinic acid-*O*-MT	*Micromonospora echinospora*	0.00157	3.29	206	180	10
3PFG	TylM1 TDP-Mycaminose-3'-*N*,*N*-diMT	*Streptomyces fradiae*	0.00158	3.18	206	241	8

[a] L1 and L2 indicate the number of amino acid residues of CalS11 and enzyme under question used for structural alignment.

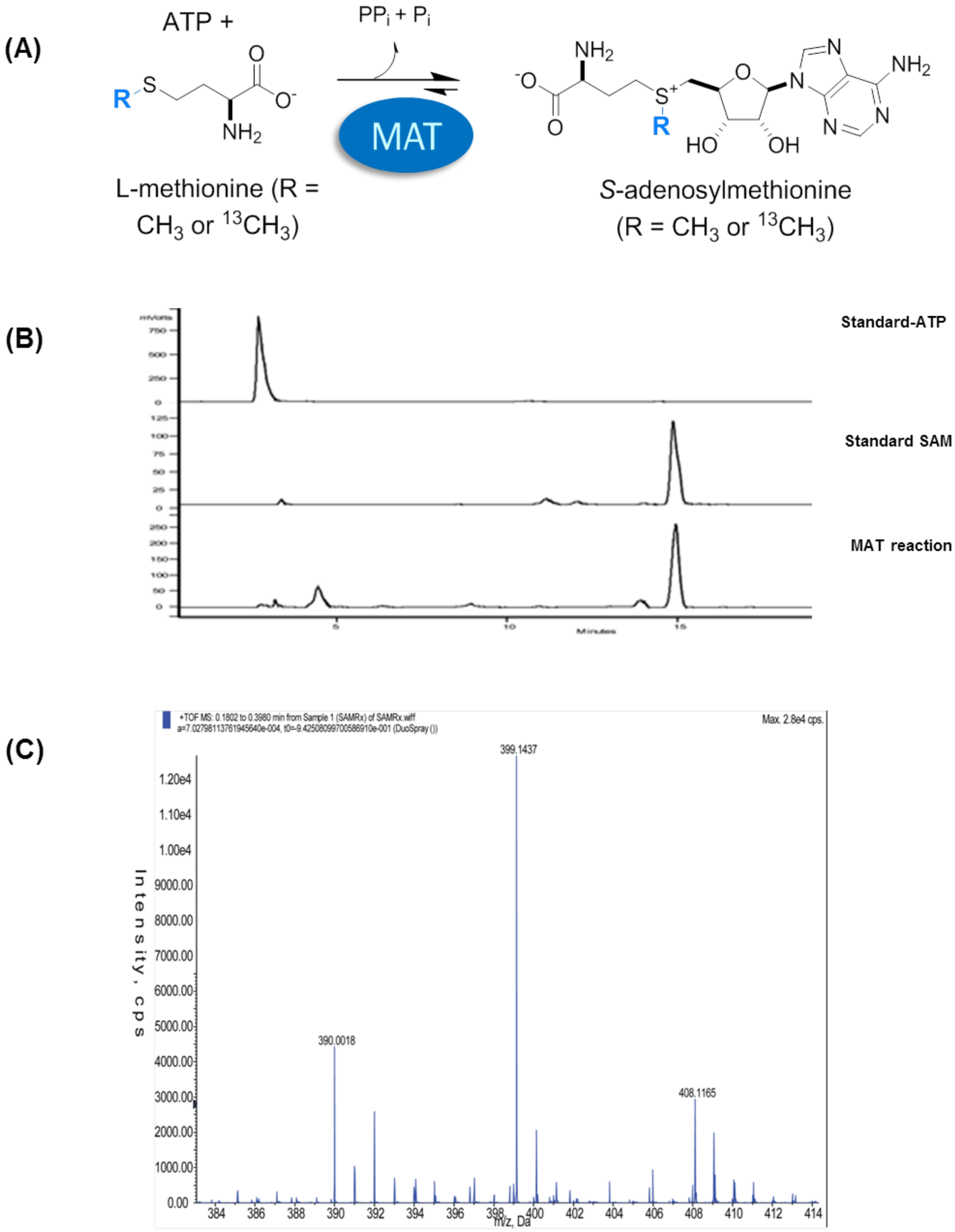


Figure S3. Methionine adenosyltransferase (MAT) reaction. **(A)** Methionine adenosyltransferase reaction. **(B)** HPLC chromatogram [Gemini-NX C-18 5 µm, 250 x 4.6 mm column (Phenomenex, Torrance, California, USA) with a gradient of 5% to 50% CH_3CN (solvent B) over 20 min (A = 20 mM sodium phosphate buffer pH 3.5 and 1 g/L 1-octanesulfonic acid (Sigma-Aldrich, St. Louis, MO); flow rate = 1 mL min^{-1}; A_{254} nm] of standards (ATP and SAM) and a corresponding MAT reaction (1.1 mM L-methionine, 1.2 equivalent of ATP and 50 µg *Sulfolobus solfataricus* methionine adenosyltransferase in 20 mM NaH_2PO_4 buffer, 10 mM $MgCl_2$, 100 mM KCl, pH 8.0 at 37 ^{o}C for 1 hr). **(C)** HRMS mass of MAT reaction product; *m/z* 399.1437, calcd 399.1446.

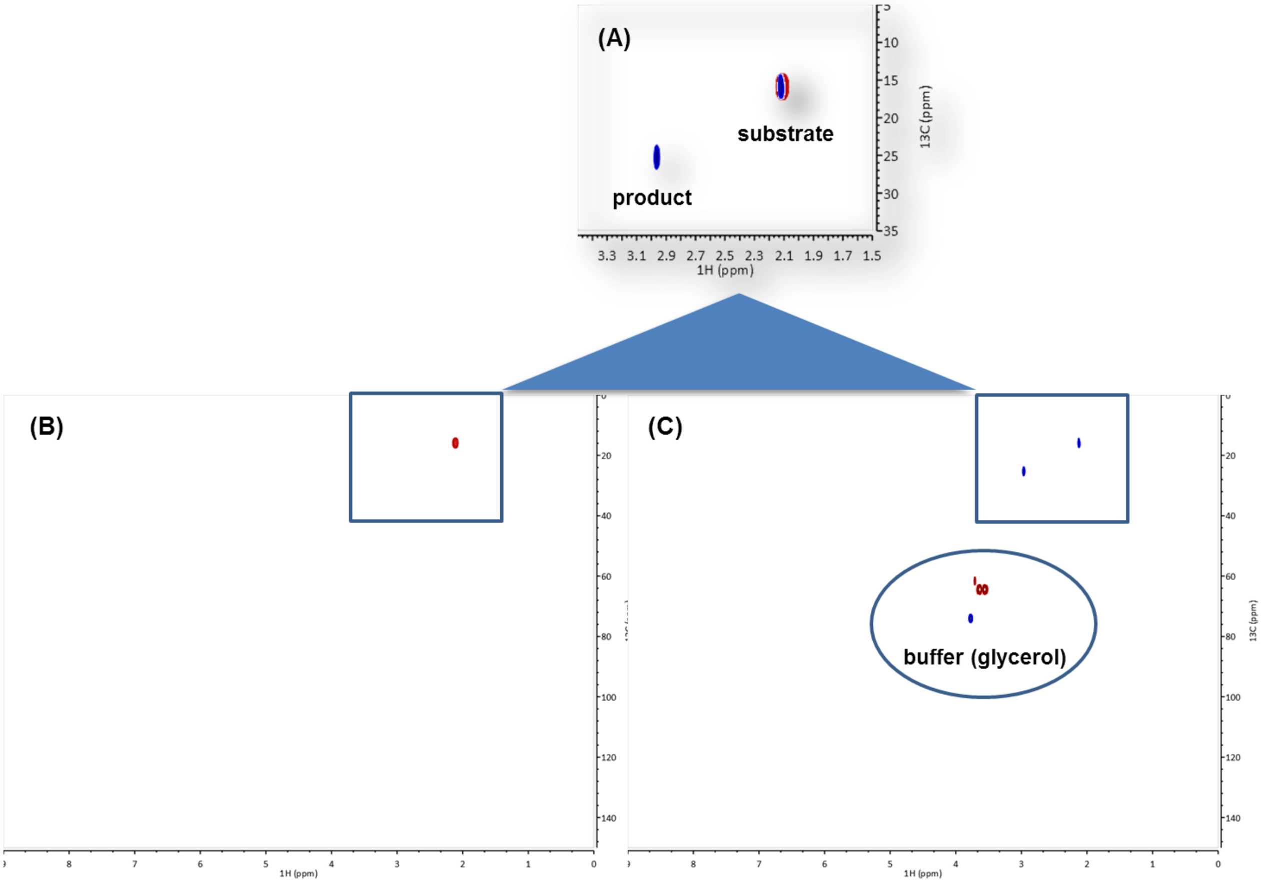


Figure S4. NMR of Methionine adenosyltransferase (MAT) reaction. **(A)** Overlay of region of ^{13}C-^{1}H HSQC spectrum consisting of *S*-CH_3 groups of [*S*-^{13}C-methyl]-L-methionine and the corresponding ^{13}C-methylated product ($^{13}CH_3$-SAM). Spectrum colored in red corresponds to reference [*S*-^{13}C-methyl]-L-methionine spectrum and that in blue is that from a partial MAT reaction containing *S*-CH_3 groups of substrate and product. The NMR chemical shifts of *S*-CH_3 groups of [*S*-^{13}C-methyl]-L-methionine and $^{13}CH_3$-SAM are ^{1}H-2.12 ppm/^{13}C-15.9 ppm and ^{1}H-2.96 ppm/^{13}C-25.4 ppm, respectively. For comparison, the full ^{13}C-^{1}H HSQC spectrum of [*S*-^{13}C-methyl]-L-methionine **(B)** and the corresponding ^{13}C-methylated product (^{13}C-SAM) **(C)** are also provided. The circled signals arise from residual glycerol within the buffer.