Science Immunology
AAAS

Supplementary Materials for

Sulfated bile acid is a host-derived ligand for MAIT cells

Emi Ito *et al.*

Corresponding author: Sho Yamasaki, yamasaki@biken.osaka-u.ac.jp

Sci. Immunol. **9**, eade6924 (2024)
DOI: 10.1126/sciimmunol.ade6924

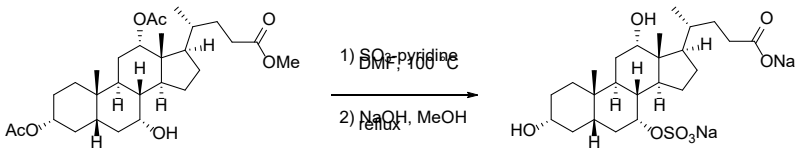
The PDF file includes:

Other Supplementary Material for this manuscript includes the following:

SUPPLEMENTARY MATERIALS AND METHODS

General methods for the synthesis of compounds

^{1}H NMR spectra were recorded using a JEOL ECA-500 or JEOL ECZ600R spectrometer. Chemical shifts are reported in δ (ppm) relative to Me_4Si (in $CDCl_3$ and CD_3OD) and MeCN signal (in D_2O) as internal standards. ^{13}C NMR spectra were recorded using a JEOL ECA-500 or JEOL ECZ600R spectrometer and referenced to the residual $CHCl_3$ signal (in $CDCl_3$), CD_3OD signal (in CD_3OD) and MeCN signal (in D_2O). IR spectra were obtained on a JASCO FT/IR-4100 spectrometer. High-resolution mass spectrometry (HRMS) spectra were recorded on a Shimadzu LC-ESI-IT-TOF-MS equipment (ESI). Optical rotations were measured with a JASCO P-1020 polarimeter. Column chromatography was performed using flash chromatography on a Wakogel C-300E (Wako). For thin-layer chromatography, a silica gel 70 F_{254} plate was employed. Compounds CA7S (Compound 1) (*85*), CA3S (Compound 2) (*85, 86*), CA12S (Compound 3) (*85*), TCA7S (Compound 4) (*85*), S7 (*85*), and S9 (*85*) are known and were synthesized by modifying previously published methods. Compounds TCA3S (Compound 7) (*85, 86*), S1 (*87*), S2 (*86*), S3 (*88*), S8 (*85*) and RL-7-Me (*24*) were also previously described and were prepared according to the previously published methods.

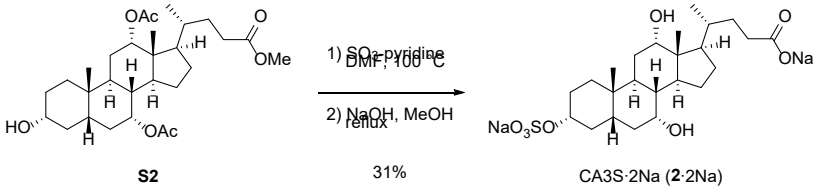


S1 64% CA7S·2Na (**1**·2Na)

Disodium cholate 7-sulfate (CA7S·2Na)

CA7S·2Na was synthesized by modifying of a previously published method (*89, 90*). Sulfur trioxide-pyridine complex (565 mg, 3.54 mmol) was added to a stirred solution of Compound S1 (179 mg, 0.352 mmol) in dry DMF (2 ml) under an argon atmosphere. After 24 hours of

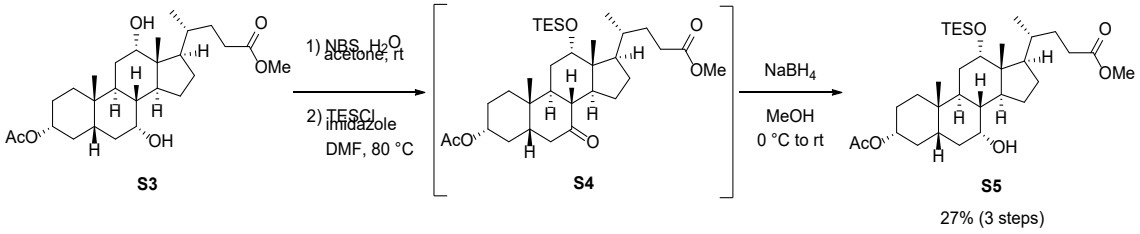
constant stirring at 100°C, the mixture was concentrated in vacuo. The residue was dissolved in a solution of NaOH (0.493 g, 12.3 mmol) in MeOH (10 ml). After 18 hours of constant stirring at 90°C, the mixture was adjusted to pH 8 with 10% aqueous HCl and concentrated in vacuo. MeOH was added to the residue and the resulting precipitate was filtered. After the filtrate was concentrated in vacuo, the residue was dissolved in water (5 ml), which was purified by Sep-Pak® C18 cartridge. The cartridge was activated by MeOH and equilibrated by water before loading. After the residue was loaded and desalinated by water (10 ml), the product was eluted by 20% MeCN aq. (20 ml, four times). The eluent was lyophilized to yield CA7S·2Na (Compound 1·2Na) as a white solid (119 mg, 64%). NMR measurements were performed after dissolving the product in 20% MeCN aq. containing 0.1% TFA followed by lyophilization: mp 173−174°C; $[\alpha]^{25}_{D}$ +10.6 (*c* 0.700, MeOH); IR (neat cm^{-1}): 3336 (OH), 1559 (C=O), 1412 (S=O); ^{1}H NMR (600 MHz, D_2O) δ 0.73 (s, 3H), 0.93 (s, 3H), 0.97 (d, J = 6.9 Hz, 3H), 1.00−1.07 (m, 1H), 1.08−1.16 (m, 1H), 1.25−1.41 (m, 3H), 1.42−1.53 (m, 2H), 1.56−1.61 (m, 1H), 1.61−1.71 (m, 3H), 1.71−1.98 (m, 8H), 1.92−1.98 (m, 1H), 1.99−2.10 (m, 2H), 2.32 (ddd, J = 15.1, 9.0, 7.5 Hz, 1H), 2.45 (ddd, J = 15.1, 9.8, 5.5 Hz, 1H), 3.49−3.55 (m, 1H), 4.05−4.08 (m, 1H), 4.49−4.52 (m, 1H); $^{13}C\{^{1}H\}$ NMR (150 MHz, D_2O) δ 12.5, 17.1, 22.4, 23.1, 27.59, 27.62, 28.1, 29.8, 31.0, 31.3, 31.7, 34.5, 35.0, 35.4, 38.4, 39.3, 41.4, 42.1, 46.7, 47.1, 72.2, 73.8, 80.1, 180.8; HRMS (ESI-TOF) *m/z*: $[M + H - 2Na]^{-}$ calculated for $C_{24}H_{39}O_8S$, 487.2371; found, 487.2379.



S2 31% CA3S·2Na (**2**·2Na)

Disodium cholate 3-sulfate (CA3S·2Na)

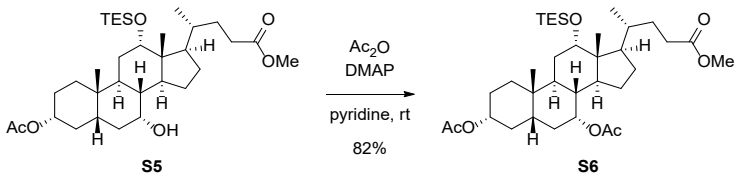
Using a procedure identical with that described for synthesis of Compound 1·2Na from Compound S1, cholic acid derivative Compound S2 (100 mg, 0.197 mmol) was converted into CA3S·2Na (Compound 2·2Na) as a white solid (32.5 mg, 31%): mp 193−195°C; $[\alpha]^{25}_{D}$ +33.8 (*c* 0.405, MeOH); IR (neat cm^{-1}): 3369 (OH), 1565 (C=O), 1409 (S=O); ^{1}H NMR (500 MHz, D_2O) δ 0.72 (s, 3H), 0.93 (s, 3H), 0.97 (d, J = 6.9 Hz, 3H), 1.02−1.10 (m, 1H), 1.11−1.21 (m, 1H), 1.25−1.36 (m, 2H), 1.38−1.46 (m, 1H), 1.47−1.58 (m, 3H), 1.58−1.62 (m, 1H), 1.62−1.71 (m, 4H), 1.71−1.81 (m, 2H), 1.81−1.94 (m, 4H), 1.99−2.04 (m, 1H), 2.07−2.14 (m, 2H), 2.17−2.27 (m, 2H), 3.89−3.91 (m, 1H), 4.06−4.08 (m, 1H), 4.18−4.24 (m, 1H); $^{13}C\{^{1}H\}$ NMR (125 MHz, D_2O) δ 12.5, 17.3, 22.3, 23.3, 26.9, 27.7, 28.0, 28.1, 32.9, 34.1, 34.8, 35.0, 35.3, 35.8, 36.7, 39.5, 41.6, 42.2, 46.7, 47.4, 69.1, 74.0, 82.1, 185.3; HRMS (ESI-TOF) *m/z*: $[M + H - 2Na]^{-}$ calculated for $C_{24}H_{39}O_8S$, 487.2371; found, 487.2376.



Compound S5

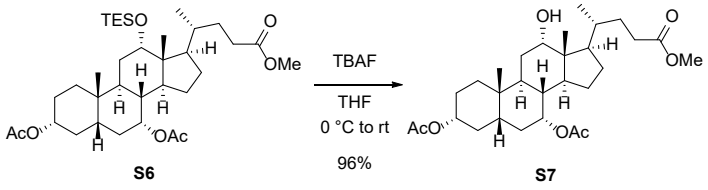
NBS (1.40 g, 7.85 mmol) was added to a stirred solution of Compound S3 (2.92 g, 6.28 mmol) in water (39 ml) and acetone (56 ml) at room temperature. After 16 hours of constant stirring at this temperature, the mixture was concentrated in vacuo, and the residue was dissolved in CH_2Cl_2. The organic layer was washed with brine, dried over Na_2SO_4, and concentrated in vacuo to yield a crude product (1.73 g). Imidazole (1.54 g, 22.7 mmol) and chlorotriethylsilane (2.52 ml, 15.0 mmol) were then added to a solution of the crude product (1.73 g) in dry DMF

(9.3 ml) at room temperature. After constant stirring for 3 hours at 80°C under an argon atmosphere, the mixture was diluted with MeOH and concentrated in vacuo. The residue was dissolved in EtOAc, washed with brine, and dried over Na_2SO_4. After concentration in vacuo, the residue was purified by flash chromatography over silica gel with *n*-hexane–EtOAc (7:1 to 5:1) to yield a ketone, Compound S4, which was used without further purification. $NaBH_4$ (192 mg, 5.08 mmol) was added to a solution of the ketone Compound S4 in MeOH (16 ml) at 0°C. After constant stirring for 2 hours at room temperature, the mixture was diluted with saturated aqueous NH_4Cl. The whole mixture was extracted with Et_2O and dried over Na_2SO_4. After concentration in vacuo, the residue was purified by flash chromatography over silica gel with *n*-hexane–EtOAc (6:1 to 4:1) to yield Compound S5 as a white amorphous solid (957 mg, 27% over three steps): mp 43−44°C; $[\alpha]^{25}_D$ +46.7 (*c* 0.473, $CHCl_3$); IR (neat cm^{-1}): 3543 (OH) 1728 (C=O); ^1H NMR (500 MHz, $CDCl_3$) δ 0.63 (q, J = 8.0 Hz, 6H), 0.66 (s, 3H), 0.90 (s, 3H), 0.96 (d, J = 6.3 Hz, 3H), 0.99−1.15 (m, 2H), 1.00 (t, J = 8.0 Hz, 9H), 1.20−1.40 (m, 4H), 1.42−1.55 (m, 6H), 1.57−1.65 (m, 1H), 1.65−1.75 (m, 2H), 1.76−1.86 (m, 2H), 1.87−1.96 (m, 3H), 1.97−2.05 (m, 1H), 2.01 (s, 3H), 2.18−2.34 (m, 3H), 2.37 (ddd, J = 15.5, 10.3, 5.2, 1H), 3.67 (s, 3H), 3.82−3.85 (m, 1H), 4.03−4.06 (m, 1H), 4.52−4.60 (m, 1H); $^{13}C\{^1H\}$ NMR (125 MHz, $CDCl_3$) δ 5.8 (3C), 7.3 (3C), 12.4, 17.6, 21.5, 22.7, 23.4, 26.6, 26.8, 27.8, 28.9, 31.0, 31.1, 34.1, 34.87, 34.93, 35.4, 35.7 39.8, 41.2, 41.3, 46.1, 47.2, 51.5, 68.4, 73.8, 74.2, 170.9, 174.7; HRMS (ESI-TOF) *m/z*: $[M + Na]^+$ calculated for $C_{33}H_{58}NaO_6Si$, 601.3895; found, 601.3894.



Compound S6

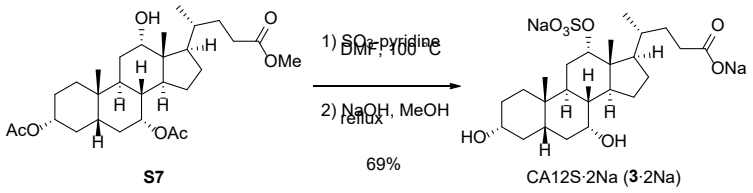
DMAP (21.4 mg, 0.175 mmol) and Ac_2O (465 μl, 4.92 mmol) were added to a solution of Compound S5 (949 mg, 1.64 mmol) in pyridine (5.3 ml) at 0°C. After being stirred for 23 hours at room temperature, the mixture was diluted with CH_2Cl_2, washed with saturated aqueous $CuSO_4$ and brine, and dried over Na_2SO_4. After concentration in vacuo, the residue was purified by flash chromatography over silica gel with *n*-hexane–EtOAc (10:1 to 7:1) to yield Compound S6 as a colorless oil (831 mg, 82%): $[\alpha]^{25}_D$ +38.3 (*c* 0.630, $CHCl_3$); IR (neat cm^{-1}): 1737 (C=O); ^1H NMR (500 MHz, $CDCl_3$) δ 0.64 (q, J = 8.0 Hz, 6H), 0.64 (s, 3H), 0.92 (s, 3H), 0.95 (d, J = 6.3 Hz, 3H), 0.99−1.12 (m, 2H), 1.03 (t, J = 8.0 Hz, 9H), 1.19−1.24 (m, 1H), 1.29−1.53 (m, 7H), 1.55−1.63 (m, 3H), 1.69−1.75 (m, 1H), 1.76−1.90 (m, 4H), 1.90−1.96 (m, 1H), 1.96−2.08 (m, 2H), 2.03 (s, 3H), 2.03 (s, 3H), 2.16−2.29 (m, 2H), 2.34−2.40 (m, 1H), 3.67 (s, 3H), 4.05−4.07 (m, 1H), 4.53−4.60 (m, 1H), 4.87−4.90 (m, 1H); $^{13}C\{^1H\}$ NMR (125 MHz, $CDCl_3$) δ 5.8 (3C), 7.1 (3C), 12.2, 17.6, 21.4, 21.5, 22.6, 23.4, 26.8, 27.7, 27.8, 28.8, 31.01, 31.04, 31.5, 34.7, 34.78, 34.82, 35.7, 38.4, 41.0, 41.1, 46.1, 47.1, 51.5, 71.2, 73.7, 74.0, 170.4, 170.8, 174.7; HRMS (ESI-TOF) *m/z*: $[M + Na]^+$ calculated for $C_{35}H_{60}NaO_7Si$, 643.4001; found, 643.4002.



Compound S7

To a solution of Compound S6 (831 mg, 1.34 mmol) in THF (22 ml) was added a solution of TBAF (1M in THF 3.35 ml, 3.35 mmol) at 0°C. After constant stirring for 15 hours at room temperature, the mixture was diluted with EtOAc. The whole mixture was washed with saturated aqueous NH_4Cl and brine, and dried over Na_2SO_4. After concentration in vacuo, the residue was purified by flash chromatography over silica gel with *n*-hexane–EtOAc (3:1 to

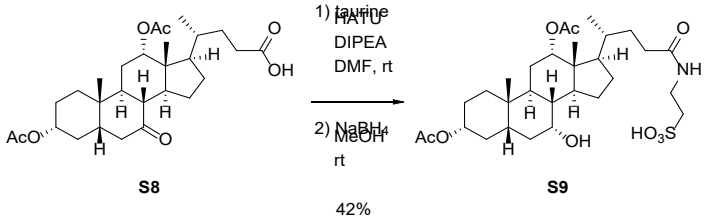
0:100) to yield Compound S7 as a white solid (653 mg, 96%): mp 181−183°C; $[\alpha]^{26}_{D}$ +23.5 (*c* 0.668, $CHCl_3$); IR (neat cm^{-1}): 3498 (OH), 1725 (C=O); ^1H NMR (500 MHz, $CDCl_3$) δ 0.69 (s, 3H), 0.92 (s, 3H), 0.98 (d, J = 6.3 Hz, 3H), 1.04−1.13 (m, 2H), 1.24−1.54 (m, 8H), 1.55−1.65 (m, 4H), 1.66−1.75 (m, 2H), 1.77−1.90 (m, 4H), 1.95 (ddd, J = 15.7, 5.7, 3.4 Hz, 1H), 2.03 (s, 3H), 2.07 (s, 3H), 2.04−2.13 (m, 1H), 2.17−2.27 (m, 2H), 2.38 (ddd, J = 15.4, 10.3, 5.1 Hz, 1H), 3.67 (s, 3H), 3.99−4.02 (m, 1H), 4.54−4.62 (m, 1H), 4.87−4.91 (m, 1H); $^{13}C\{^1H\}$ NMR (125 MHz, $CDCl_3$) δ 12.5, 17.3, 21.5, 21.6, 22.5, 22.9, 26.6, 27.2, 28.1, 28.5, 30.8, 31.0, 31.2, 34.3, 34.5, 34.7, 34.9, 38.0, 40.9, 42.1, 46.5, 47.1, 51.5, 70.8, 72.7, 74.1, 170.5, 170.7, 174.6; HRMS (ESI-TOF) *m/z*: $[M + Na]^+$ calculated for $C_{29}H_{46}NaO_7$, 529.3136; found, 529.3136.



Disodium cholate 12-sulfate (CA12S·2Na)

By a procedure identical with that described for synthesis of Compound 1·2Na from Compound S1, cholic acid derivative Compound S7 (102 mg, 0.201 mmol) was converted into CA12S·2Na (Compound 3·2Na) as a white amorphous solid (74.3 mg, 69%): mp 195−198°C; $[\alpha]^{25}_{D}$ +40.3 (*c* 0.740, MeOH); IR (neat cm^{-1}): 3383 (OH), 1560 (C=O), 1410 (S=O); ^1H NMR (500 MHz, D_2O) δ 0.76 (s, 3H), 0.92 (s, 3H), 0.96−1.05 (m, 1H), 0.98 (d, J = 6.3 Hz, 3H), 1.08−1.19 (m, 1H), 1.24−1.36 (m, 2H), 1.36−1.56 (m, 5H), 1.56−1.73 (m, 5H), 1.74−1.87 (m, 3H), 1.88−1.96 (m, 1H), 1.96−2.06 (m, 2H), 2.07−2.20 (m, 3H), 2.20−2.27 (m, 1H), 3.42−3.50 (m, 1H), 3.87−3.91 (m, 1H), 4.69−4.73 (m, 1H); $^{13}C\{^1H\}$ NMR (125MHz, D_2O) δ 12.3, 17.7, 22.6, 23.3, 25.1, 27.3, 27.7, 29.9, 32.9, 34.4, 34.9, 35.2, 35.4, 35.9, 39.1, 39.4, 41.6, 42.9, 46.3,

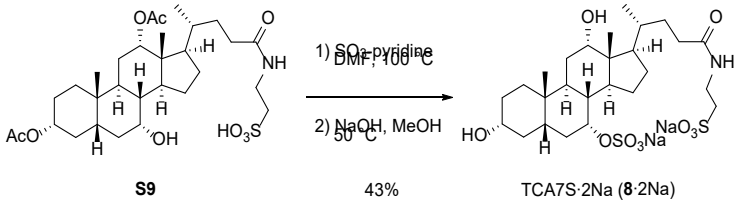
47.1, 69.0, 72.4, 84.1, 185.3; HRMS (ESI-TOF) *m/z*: $[M + H - 2Na]^-$ calculated for $C_{24}H_{39}O_8S$, 487.2371; found, 487.2378.



Compound S9

Diisopropylethylamine (421 µl, 2.41 mmol) and HATU (461 mg, 1.21 mmol) were added to a stirred solution of Compound S8 (395 mg, 0.805 mmol) in dry DMF (6.0 ml) at room temperature under an argon atmosphere. After constant stirring for 30 min, taurine (153 mg, 1.22 mmol) was added. The mixture was stirred for 17.5 hours at room temperature. After concentration in vacuo, the residue was diluted with Et_2O and saturated aqueous $NaHCO_3$. After the aqueous layer was concentrated in vacuo, MeOH was added to the residue and the resulting precipitate was filtered. After the filtrate was concentrated in vacuo, the residue was dissolved in water (3 ml). Using a procedure identical to that described for synthesis of Compound 1 from Compound S1, the aqueous solution of sulfonic acid was purified by Sep-Pak® C18 cartridge and lyophilized to yield the taurine-conjugated product as a pale yellow amorphous solid (283 mg), which was used without further purification. MeOH (4.0 ml) and $NaBH_4$ (34.6 mg, 0.915 mmol) were added to a stirred solution of the taurine-conjugated product (283 mg) at room temperature. After constant stirring for 17.5 hours at this temperature, the mixture was diluted with saturated aqueous NH_4Cl and concentrated in vacuo. The residue was dissolved in $CHCl_3$. The organic layer was washed with 10% aqueous HCl three times and dried over Na_2SO_4. After concentration in vacuo, the residue was purified by flash

chromatography over silica gel with $CHCl_3$-MeOH (5:1 to 4:1) to yield Compound S9 as a white solid (203 mg, 42% over two steps). All spectral data were in agreement with previously reported findings (*85*): ^{1}H NMR (500 MHz, CD_3OD) δ 0.78 (s, 3H), 0.85 (d, J = 6.3 Hz, 3H), 0.94 (s, 3H), 1.00−1.08 (m, 1H), 1.11−1.20 (m, 1H), 1.24−1.48 (m, 7H), 1.49−1.66 (m, 4H), 1.66−1.81 (m, 6H), 1.90−2.00 (m, 2H), 1.99 (s, 3H), 2.05−2.14 (m, 3H), 2.10 (s, 3H), 2.24 (ddd, J = 13.7, 10.3, 5.1 Hz, 1H), 2.35−2.41 (m, 1H), 2.96 (t, J = 6.9 Hz, 2H), 3.58 (t, J = 6.9 Hz, 2H), 3.80−3.83 (m, 1H), 4.47−4.55 (m, 1H), 5.06−5.08 (m, 1H).



S9 43% TCA7S·2Na (**8**·2Na)

Disodium taurocholate 7-sulfate (TCA7S·2Na)

TCA7S·2Na was synthesized by modifying a previously published method (*89, 90*). To a stirred solution of Compound S9 (80.1mg, 0.134 mmol) in dry DMF (1.6 ml) was added sulfur trioxide-pyridine complex (214 mg, 1.35 mmol) under argon atmosphere. After being stirred for 50 hours at 100°C, the mixture was concentrated in vacuo. A solution of NaOH (0.792 g, 19.8 mmol) in MeOH (16 ml) was added to the mixture. After 45 hours of stirring at 50°C, the mixture was adjusted to pH 8 with 10% aqueous HCl and concentrated in vacuo. MeOH was added to the residue, and the resulting precipitate was filtered. After the filtrate was concentration in vacuo, the residue was dissolved in water (3 ml), which was purified by Sep-Pak® C18 cartridge and lyophilized to yield TCA7S·2Na (Compound 8·2Na) as a pale yellow solid (36.8 mg, 43%): mp 186°C (decomp.); $[\alpha]^{24}_D$ +7.6 (*c* 0.520, MeOH); IR (neat cm^{-1}): 3336 (OH), 1654 (C=O), 1449 (S=O); ^{1}H NMR (500 MHz, D_2O) δ 0.71 (s, 3H), 0.93 (s, 3H), 0.97 (d, J = 6.3 Hz, 3H), 0.99−1.17 (m, 2H), 1.20−1.44 (m, 4H), 1.46−1.53 (m, 1H), 1.55-1.97 (m,

13H), 1.99−2.10 (m, 2H), 2.20 (ddd, *J* = 14.1, 7.4, 7.4 Hz, 1H), 2.31 (ddd, J = 14.1, 9.7, 5.1 Hz, 1H), 3.07 (t, *J* = 6.9 Hz, 2H), 3.47−3.53 (m, 1H), 3.56 (t, *J* = 6.9 Hz, 2H), 4.03−4.07 (m, 1H), 4.47−4.51 (m, 1H); $^{13}C\{^{1}H\}$ NMR (125 MHz, D_2O) δ 12.6, 17.3, 22.5, 23.2, 27.6, 27.7, 28.2, 29.8, 31.0, 32.2, 33.4, 34.6, 35.1, 35.4, 35.6, 38.4, 39.4, 41.5, 42.1, 46.7, 47.2, 50.4, 72.2, 73.8, 79.9, 178.16; HRMS (ESI-TOF) *m/z*: $[M + H - 2Na]^-$ calculated for $C_{26}H_{44}O_{10}NS_2$, 594.2412; found, 594.2410.

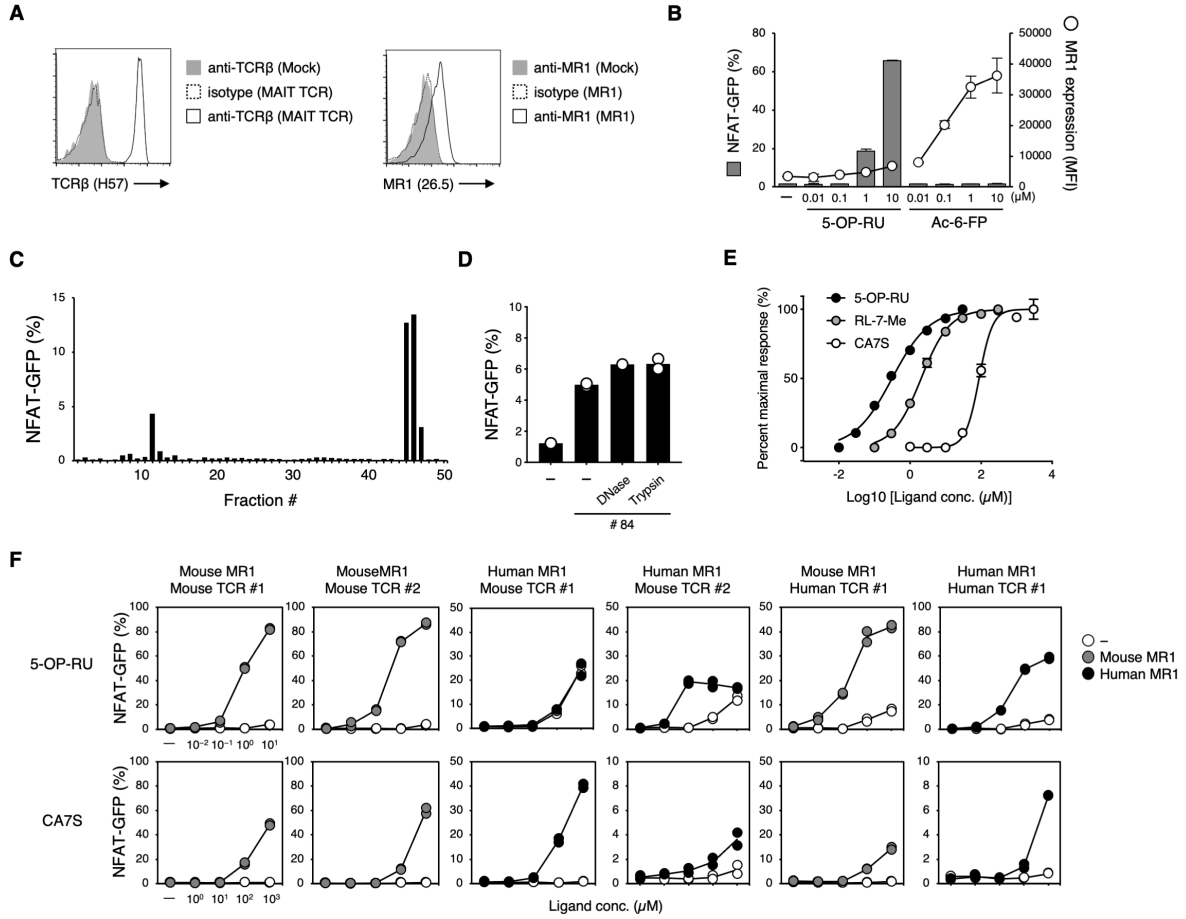


Fig. S1. Purification and characterization of CA7S. (**A**) A MAIT TCR and MR1 were co-expressed in TCR-deficient T cell hybridoma (*81*). Surface expression levels of MAIT TCR and MR1 on reporter cells were assessed by surface staining with anti-mouse TCRβ (H57-597) (left) or anti-human/mouse/rat MR1 (26.5) (right). HTK888 and MOPC-173 were used as isotype control Abs, respectively (dotted line). (**B**) NFAT-GFP reporter cells were stimulated with vehicle control, 5-OP-RU or Ac-6-FP. GFP expression and MR1 surface expression were evaluated at 6 hours after stimulation. Closed columns indicate GFP expression. Open circles indicate MR1 surface expression. (**C**) Percentages of GFP^+ cells upon stimulation with the 50 fractions from the secondary purified intestine of SPF mice. (**D**) Percentages of GFP^+ cells upon stimulation with DNase or trypsin treated fraction #84. (**E**) Dose-response curves for 5-OP-RU, RL-7-Me, and CA7S. (**F**) Combinations of human or mouse MR1 with three clones of MAIT TCR were coexpressed in NFAT-GFP reporter cells. These cells were stimulated with vehicle control, 5-OP-RU, or CA7S and analyzed by flow cytometry after 20 hours. Data are presented as the means ± SD of triplicates (B and E) or individual values of duplicate assays (D and F). All data are representative of more than two independent experiments.

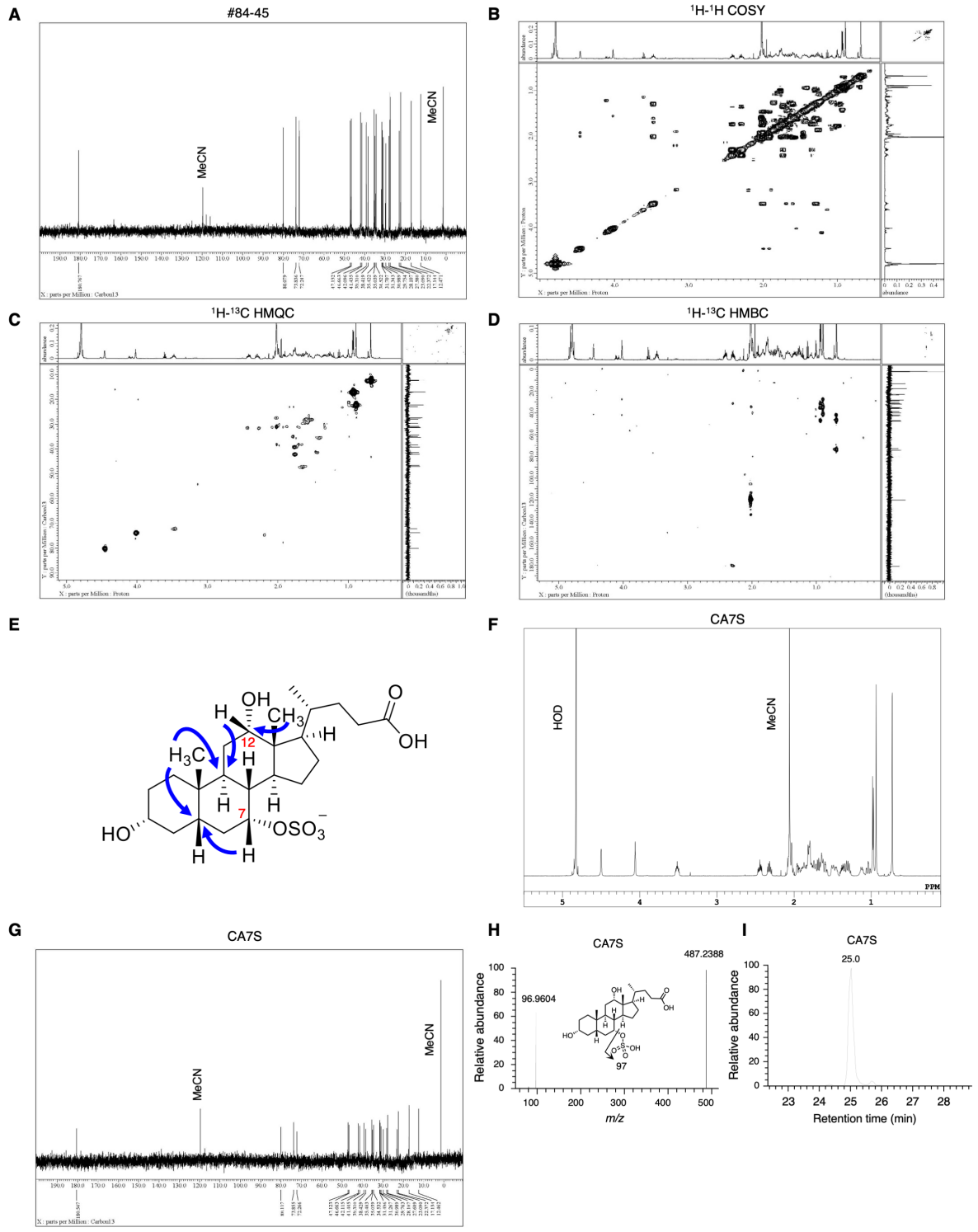


Fig. S2. Structural determination of CA7S. (**A**) ^{13}C NMR spectra (150 MHz, D_2O) of fraction #84-45. (**B** to **D**) 2D NMR analyses, including ^{1}H-^{1}H correlation spectroscopy (COSY) (B), ^{1}H-^{13}C heteronuclear multiple quantum correlation spectroscopy (HMQC) (C) and ^{1}H-^{13}C heteronuclear multiple-bond correlation spectroscopy (HMBC) (D) of fraction #84-45. (**E**) Key ^{1}H-^{13}C HMBC correlations of fraction #84-45. (**F** to **I**) ^{1}H NMR (600 MHz, D_2O) spectra (F), ^{13}C NMR (150 MHz, D_2O) spectra (G), MS/MS spectra (H), and extracted ion chromatograms (I) of synthetic CA7S.

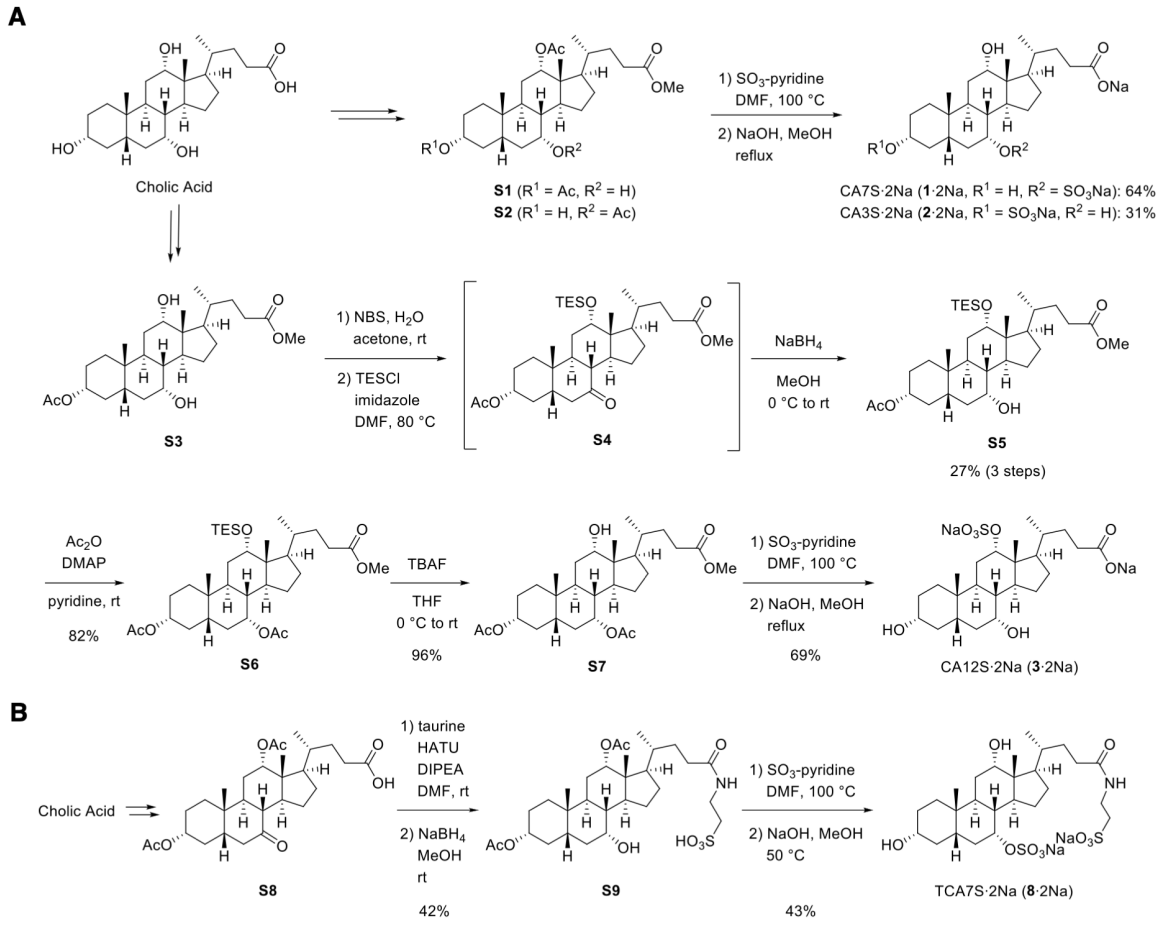


Fig. S3. Chemical synthesis of bile acid metabolites. (**A**) Synthetic schemes of CA7S (Compound 1), CA3S (Compound 2), and CA12S (Compound 3). (**B**) Synthetic scheme of TCA7S (Compound 8).

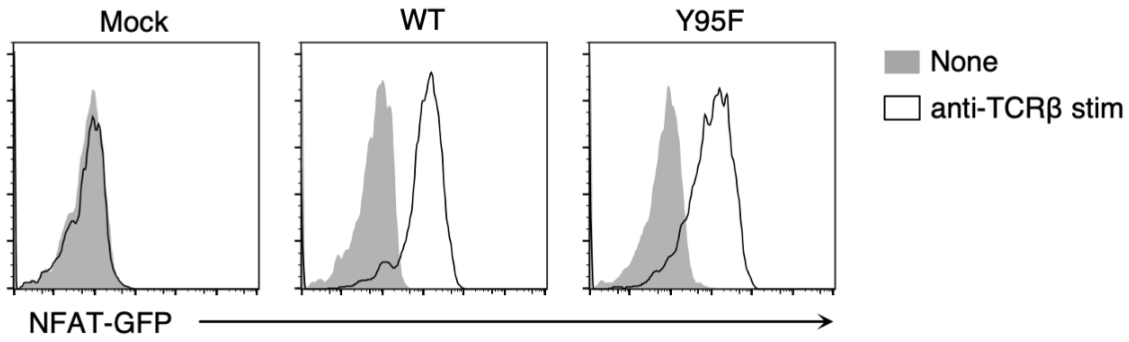


Fig. S4. Stimulation of reporter cells expressing mutant TCRα. Reporter cells were transfected with vector alone (Mock), MAIT TCRβ together with WT MAIT TCRα (WT), or mutant TCRα (Y95F). The reporter cells were stimulated with 10 μg/ml of plate-coated anti-mouse TCRβ (H57-597). GFP expression was analyzed by flow cytometry after 20 hours.

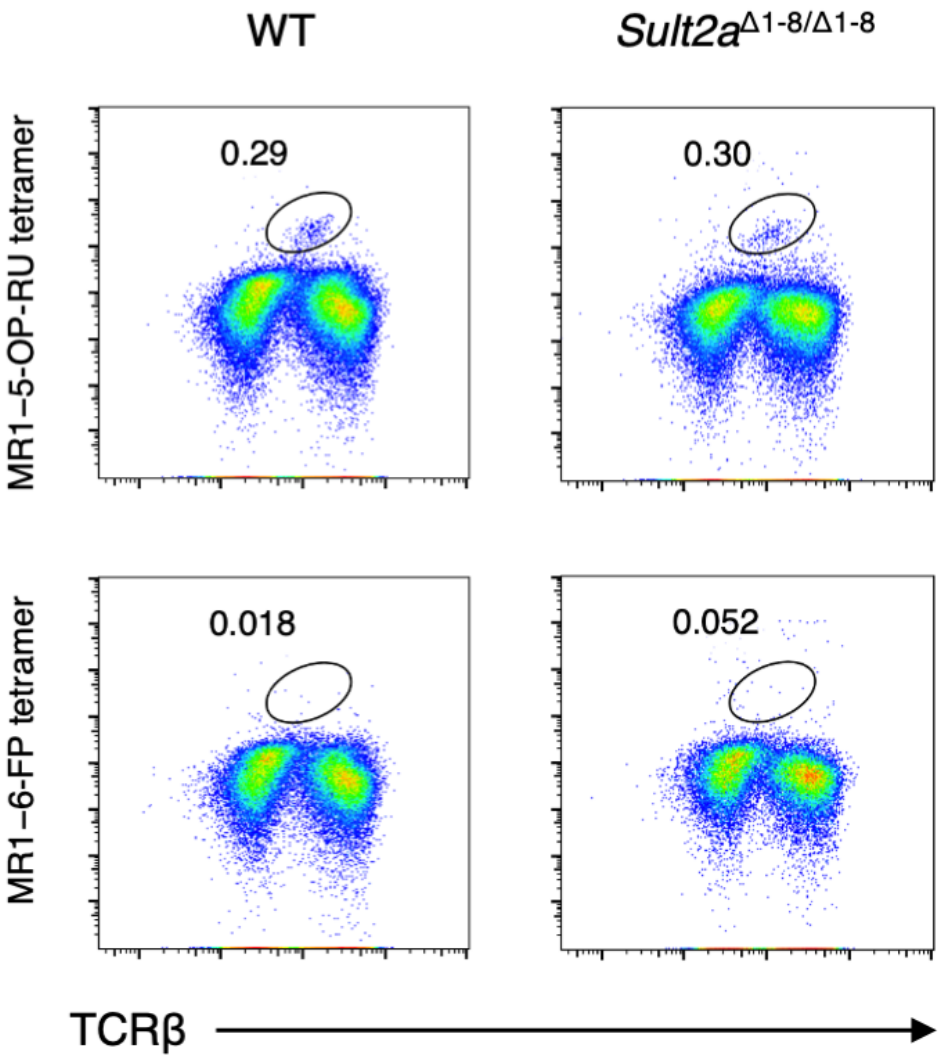


Fig. S5. Liver MAIT cells in *Sult2a*-deficient mice. Representative flow cytometry dot plots of liver $CD45^{+}CD19^{-}TCR\beta^{+}$MR1−5-OP-RU tetramer^{+} MAIT cells from WT or *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. MR1−6-FP tetramer staining was used as a negative control for the MR1−5-OP-RU tetramer.

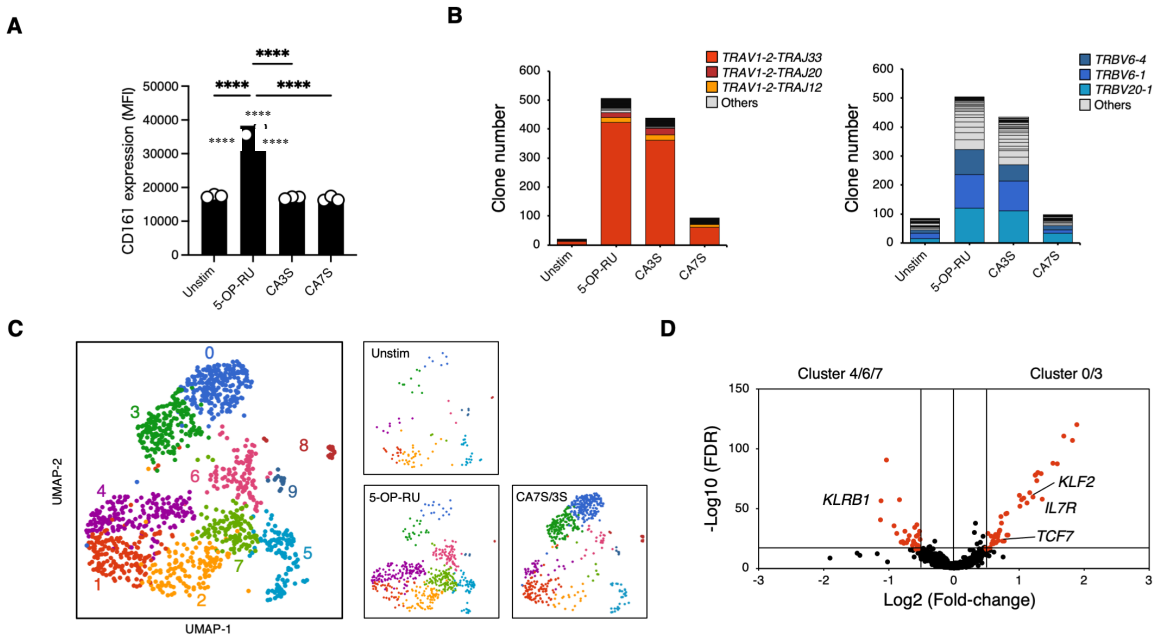


Fig. S6. Responses of human MAIT cells upon stimulation with CAS or 5-OP-RU. (A) CD161 expression (MFI) on $CD3^{+}CD161^{+}MR1-5\text{-OP-RU-tetramer}^{+}$ MAIT cells in PBMCs 6 days after stimulation with or without 5-OP-RU, CA3S, or CA7S. (**B** to **D**) Human PBMCs were stimulated with vehicle control (Unstim), 5-OP-RU (10 μM), CA3S (500 μM), or CA7S (500 μM) for 17 days. (B) TCRαβ usage of $CD3^{+}CD161^{+}MR1{-}5\text{-OP-RU tetramer}^{+}$ cells for each stimulus group. Vα-Jα (left) and Vβ (right) are shown. (C) UMAP projection based on mRNA expression of all (left) or each stimulus group (Unstim, 5-OP-RU, or CA7S/3S, right) (D) Volcano plot of mRNA expression comparing the characteristic clusters of (A) 5-OP-RU stimulation (cluster 4/6/7) and CA7S/3S stimulation (cluster 0/3). Data are presented as individual values and the means of triplicate assays. ****P<0.001, by two-tailed, one-way ANOVA followed by Tukey's multiple comparison test.

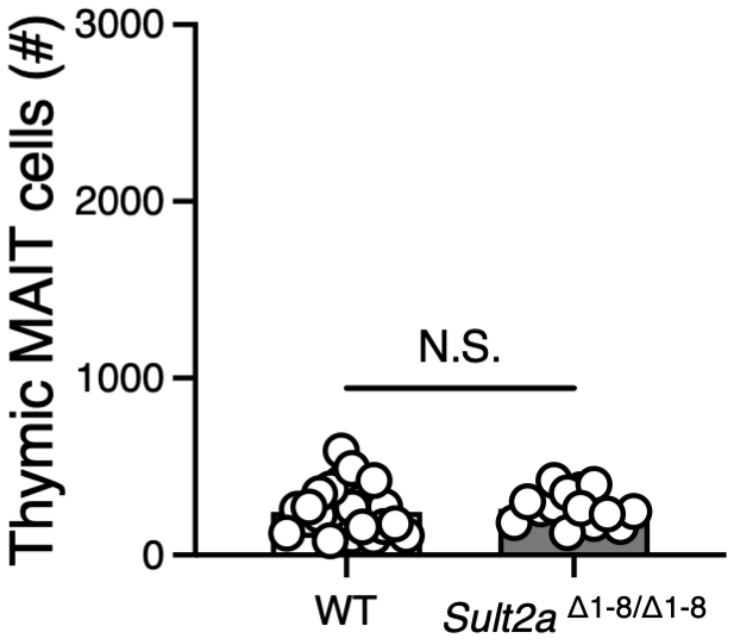


Fig. S7. Thymic MAIT cells in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice under germ-free conditions. Germ-free *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice were rederived by in vitro fertilization. Thymic MAIT cells from 8-week-old WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice under germ-free conditions were analyzed by tetramer staining as described in Fig. 6A. Data are from experiments with 14 or more mice per group. Each symbol indicates individual mice. N.S., not significant.

A

1st purify conditions			
Column	COSMOSIL PBr 4.6mm I.D. x 250mm		
Eluent A	A: water (0.1%TFA)		
Eluent B	B: AcCN: water = 80:20 (0.1%TFA)		
Flow Rate	1.0mL/min gradient		
Column Temperature	R.T.		
Fractionation	0.5mL/Fr		
	Time (min)	A (%)	B (%)
Gradient Condition	0	100	0
	7	100	0
	55	0	100
	60	0	100
	62	100	0
	65	stop	

B

2nd purify conditions			
Column	TSKgel Amide-80 5μm 4.6mm I.D. x 250mm		
Eluent A	A: water (0.1%TFA)		
Eluent B	B: AcCN (0.1%TFA)		
Flow Rate	1.0mL/min gradient		
Column Temperature	R.T.		
Fractionation	0.5mL/Fr		
	Time (min)	A (%)	B (%)
Gradient Condition	0	0	95
	7	0	95
	55	100	0
	60	100	0
	62	0	100
	65	stop	

Table S1. HPLC conditions for purification. (**A** and **B**) HPLC conditions for (A) the primary purification on reversed-phase column chromatography and (B) the secondary purification on hydrophilic interaction chromatography (HILIC).

	CA3S	CA12S
Tissue	Average ± SD (pmol/mg)	
Thymus	N.D.	N.D.
Liver	N.D.	N.D.
Gallbladder	N.D.	N.D.
Duodenum	N.D.	N.D.
Jejunum	N.D.	N.D.
Ileum	N.D.	N.D.
Cecum	N.D.	N.D.
Colon	N.D.	N.D.

Table S2. Quantitative determination of CA3S and CA12S in mouse tissues. Tissue distribution of CA3S and CA12S in various tissues of SPF mice. Data are presented as the means ± SD from experiments with three mice per group and are representative of two independent experiments. N.D., not detected.

Cluster 1			
Name	ES	p-value	FDR
Innate immune response	-0.67865	0.001	0.04724
Defense response to other organism	-0.67095	0.001	0.04724
Response to cytokine	-0.62764	0.001	0.04724
Defense response	-0.60893	0.001	0.04724
Positive regulation of immune system process	-0.53741	0.001	0.04724
Regulation of immune response	-0.51989	0.001	0.04724
Organic acid metabolic process	-0.50649	0.001	0.04724
Regulation of immune system process	-0.50648	0.001	0.04724
Transmembrane transport	-0.50306	0.001	0.04724
Cell cell signaling	-0.4987	0.001	0.04724
Protein localization to organelle	-0.49436	0.001	0.04724
Proteolysis	-0.48758	0.001	0.04724
Mrna metabolic process	-0.4786	0.001	0.04724
Exocytosis	-0.47674	0.001	0.04724
Regulation of protein localization	-0.47035	0.001	0.04724
Cellular protein catabolic process	-0.47031	0.001	0.04724
Secretion	-0.4668	0.001	0.04724
Protein modification by small protein conjugation	-0.46519	0.001	0.04724
Regulation of cell death	-0.46224	0.001	0.04724
Protein catabolic process	-0.46132	0.001	0.04724
Cellular macromolecule catabolic process	-0.44836	0.001	0.04724
Intracellular protein transport	-0.4478	0.001	0.04724
Organonitrogen compound catabolic process	-0.44531	0.001	0.04724
Positive regulation of molecular function	-0.43894	0.001	0.04724
Intracellular transport	-0.43225	0.001	0.04724

Cluster 0			
Name	ES	p-value	FDR
Positive regulation of vascular wound healing	0.99986	0.00202	0.06026
Negative regulation of toll like receptor 4 signaling pathway	0.97634	0.00225	0.0605
Negative regulation of t cell mediated cytotoxicity	0.96164	0.00244	0.0605
Acylglycerol homeostasis	0.93427	0.00244	0.0605
Adenylate cyclase activating adrenergic receptor signaling pathway involved in heart process	0.94884	0.00244	0.0605
Negative regulation of interleukin 8 secretion	0.93358	0.00244	0.0605
G protein coupled receptor signaling pathway involved in heart process	0.94897	0.00252	0.06105
Histone h3 k14 acetylation	0.84244	0.00261	0.06156
Actin filament severing	0.86	0.00261	0.06156
Regulation of aspartic type endopeptidase activity involved in amyloid precursor protein catabolic process	0.8468	0.00261	0.06156
Negative regulation of p38mapk cascade	0.85534	0.00261	0.06156
Negative regulation of microglial cell activation	0.90483	0.00261	0.06156
Negative regulation of macrophage activation	0.86433	0.00269	0.06246
Detection of mechanical stimulus	0.85941	0.00269	0.06246
Angiogenesis involved in wound healing	0.89993	0.00269	0.06246
Neutrophil extravasation	0.85242	0.00269	0.06246
Regulation of cell cell adhesion mediated by integrin	0.82478	0.00272	0.06268
Regulation of toll like receptor 4 signaling pathway	0.89975	0.00272	0.06268
Negative regulation of smoothened signaling pathway	0.81159	0.00277	0.06326
Negative regulation of camp mediated signaling	0.85076	0.00292	0.06519
V d j recombination	0.86972	0.00296	0.06534
Forebrain cell migration	0.84898	0.00337	0.06795
Positive regulation of wound healing	0.75979	0.00337	0.06795
Negative regulation of g protein coupled receptor signaling pathway	0.71663	0.00342	0.06825
Positive regulation of response to wounding	0.73368	0.00342	0.06825

Table S3. Gene ontology upon stimulation with CA7S or 5-OP-RU. Top 25 gene ontology terms of biological processes based on differentially expressed genes comparing the characteristic clusters of 5-OP-RU stimulation (cluster 1, left) and CA7S stimulation (cluster 0, right), related to Fig. 7, D and E. Inflammatory (cluster 1) and tissue repair/wound healing-related (cluster 0) gene ontology terms are both shown in bold.

A

Mammals	MR1	TRAV1	Sult2a
Human	+	+	+
Mouse	+	+	+
Cow	+	+	+
Cat	–	–	–
Dog	–	–	–
Panda	–	–	–
Ferret	–	–	–
Microbat	+	+	+

B

```
Human Sult2A1          --------------------------MSDDFLWFEGIAFPTMGFRSETLR KVRDEFVIRDEDVIILTYPKSGTNWLAEILCLMHSKGDAKWIQSVPIWER
Mouse Sult2A2          --------------------------MMSDYNWFEGIPFPAISYQREILE DIRNKFVVKEEDLLILTYPKSGTNWLNEIVCLIQTKGDPKWIQTVPIWDR
Cow Sult2A1            -------------------------------------------MEFPSEHLR EVQEDFIIKDEDVLLLTYPQSGTNWLIETVCLIYSKGDPKWVQSEPIWDR
Microbat Sult2A1       --------------------------MSDGYVWFEGIPFPVVDFTPEHMR EVQENFVFKDEDVLILSFPKSGTNWLIEILSLIYSKGDPKWIQSVPIWER
Human Sult2B1          MDGPAEPQIPGLWDTYEDDISEISQKLPGEYFRYKGVPFPVGLYSLESIS LAENTQDVRDDDIFIITYPKSGTTWMIEIICLILKEGDPSWIRSVPIWER
Panda Sult2B1          ---------------------------------------------------- MVEN-ADVQDDDIFIITYPKSGTNWMIEILSLILKDGDPSWIHSVPIWKR
Ferret Sult2B1         MDGPSEPRIPALWDTHENNISELSQKLSGEYFRYKGVPFPVGIYSPESIR MVEN-ADVKDDDIFIITYPKSGTNWMIEILSLILKDGDPSWIRSVPIWKR
Dog Sult2B1            MDGPAEPRIPALWDTHENNISEISQKLSGEYFRYKGIPFPVGIYSPESIR MVEN-ADVQDDDIFIITYPKSGTNWMIEILSLILKDGDPSWIRSVPIWKR
Cat Sult2B1            MDGPAEPRIPALWDGSED-ISDISQKLSGEYFRYKGIPFPVGIYSPESIR MVEN-ADVQDDDIFIITYPKSGTNWMIEILSLILKDGDPSWIRSVPIWKR
                                                                              .:   .:::*::::::*:***.*: * :.*: ..**..*::: ***.*

Human Sult2A1          SPWVESEIGYTALSETESPRLFSSHLPIQLFPKSFFSSKAKVIYLMRNPR DVLVSGYFFWKNMKFIKKPKSWEEYFEWFCQGTVLYGSWFDHIHGWMPMR
Mouse Sult2A2          SPWIETEIGYSAIINKEGPRLITSHLPIHLFSKSFFSSKAKAIYLMRNPR DILVSGYFFWGNTNLVKNPGSLGTYFEWFLQGNVLFGSWFEHVRGWLSMR
Cow Sult2A1            SPWVETKHGYELLKEKEGPRLISSHLPIQLFPKSFFKSKAKMIYLIRNPR DVLVSGYFFWRSAKFVKRPQSLEQYFEWFVEGNVVFGSWFDHARGWMSMR
Microbat Sult2A1       SPWLETMYGYKTLKDKEGPRFMTSHVPIHLFPKSLFKSKAKVIYIIRNPR DILVSGYFFWITSNFVKKPESLEQYFEWFIQGKVVFGSWFDHVRGWISMQ
Human Sult2B1          APWCETIVGAFSLPDQYSPRLMSSHLPIQIFTKAFFSSKAKVIYMGRNPR DVVVSLYHYSKIAGQLKDPGTPDQFLRDFLKGEVQFGSWFDHIKGWLRMK
Panda Sult2B1          APWCETIMGALSLSDQTSPRLMSSHLPIQLFTKAFFNSKAKVIYVGRNPR DVVVSFYHYSKIAGQLKDPGTPDQFLQNFLKGEVQFGSWFDHIKGWIRMQ
Ferret Sult2B1         APWCETIMGAFSLSDQPSPRLMSSHLPIQLFTKAFFNSKAKVIYMGRNPR DVVVSLYHYSKIAGQLKDPGTPDQFLQNFLKGEVQFGSWFDHIKGWIRMQ
Dog Sult2B1            APWCETIMGAFSLSDQPRPRLMSSHLPIQLFTKAFFNSKAKVIYMGRNPR DVVVSLYHYSKIAGQLKDPGTPDQFLQNFLKGEVQFGSWFDHIKGWIRMQ
Cat Sult2B1            APWCETILGAFSLSDQPRPRLMSSHLPIQLFAKAFFTSKAKVIYIGRNPR DVVVSLYHYSKIAGQLKDPGTPDQFLQNFLKGEVQFGSWFDHIKGWIRMK
                       :** *:  *   : :   **:::**:**::*.*::*.**** **: **** *::** *.:      :* * :   ::. * :* * :****:* :**: *:

Human Sult2A1          EEKNFLLLSYEELKQDTGRTIEKICQFLGKTLEPEELNLILKNSSFQSMK ENKMSNYSLLSVDYVVD-KAQLLRKGVSGDWKNHFTVAQAEDFDKLFQEK
Mouse Sult2A2          EWDNFLVLYYEDMKKDTKGTIKKICDFLGKNLGPDELDLVLKYSSFQAMK ENNMSNYSLIKEDRVTN-GLKLMRKGTTGDWKNHFTVAQAEAFDKVFQEK
Cow Sult2A1            DKENFLILSYEEMKWDMRSTVEKICQFLGKKLEPEELNSVLKNSSFQAMK ENSMSNYSLLKGQYFEE-NGQLLRKGVTGDWRNYFTVAQAETFDKLFQEK
Microbat Sult2A1       GKENFLILSYEELKRDTRSAVQKLCQFLGKKLEPEELNSVLQNSSFQAMK ENKMSNYSMLRGWLLDQKNGSFMRKGISGDWKNYFTVAQAEVFDKIFQEK
Human Sult2B1          GKDNFLFITYEELQQDLQGSVERICGFLGRPLGKEALGSVVAHSTFSAMK ANTMSNYTLLPPSLLDHRRGAFLRKGVCGDWKNHFTVAQSEAFDRAYRKQ
Panda Sult2B1          GKENFLFITYEELQQDLHGSVKRICQFLGRPLGEEALGSVVAHSAFGAMK ANAMSNYTLLPPSLLDQRHGAFLRKGVCGDWKNHFTVAQSEAFDRVYREQ
Ferret Sult2B1         GKENFLFITYEELQQDLHGSVQRICQFLGRPLGEEALGSVVAHSAFGAMK ANAMSNFTLLPPSLLDQRHGAFLRKGVCGDWKNHFTVAQSEAFDRVYREQ
Dog Sult2B1            GKENFLFITYEELHQDLHGSVQRVCEFLGRPLGEEALGSVVAHSAFGAMK ANAMSNFTLLPPTLLDQRRGAFLRKGVCGDWKNHFTVAQSEAFDRVYREQ
Cat Sult2B1            GKENFLFITYEELQQDLHSCLQDICQFLGWPLGKEALDSVVAHSAFGAMK ANAMSNFTLLPPSLLDQRHGAFLRKGVCGDWKNHFTVAQSEAFDRVYREQ
                         .***.: **::: *    :: :* ***  *  : *. ::  *:* :**  * *******    . .    ::***  ***:*:*****:* **: ::::

Human Sult2A1          MADLPRELFPWE---------------------------------------- ----------------
Mouse Sult2A2          MAGFPPGMFPWE---------------------------------------- ----------------
Cow Sult2A1            MADLPQDLFLWKQ--------------------------------------- ----------------
Microbat Sult2A1       TADLPRELFPWEEWSNTSRSQR----------------------------- ----------------
Human Sult2B1          MRGMP--TFPWDEDPEEDGSPDPEPSPEPEPKPSLEPNTSLEREPRPNSS PSPSPGQASETPHPRPS
Panda Sult2B1          MRGLP--TFPWDVSEDASPDPDPSPSPPEASEPPNP--------------- ----------------
Ferret Sult2B1         MRGLP--TFPWDVPDDASPDPDPSPSTPEASEPPNP--------------- ----------------
Dog Sult2B1            MRGMP--TFPWDVPEDTSPDPDPSPSPPQAGSEPLHP-------------- ----------------
Cat Sult2B1            MRGLP--TFPWDDPEDASPDPDPSPTPAQASEHPPDL-------------- ----------------
                         .:*   * *.
```

Table S4. Alignment of *SULT2A1* homologues in mammalian species. (**A**) Correlation among the presence of functional genes encoding *MR1*, *TRAV1*, and sulfotransferase 2 (*SULT2*) in various mammalian species. The amino acid sequences of genes that show the highest similarities with human *SULT2A1* were obtained with TBLASTN searches. We found no *SULT2A1*-like genes in cats, dogs, pandas, and ferrets,. Rather, the obtained sequences were similar to human *SULT2B1* with respect to amino acid sequence identities, number of exons, and domain architectures. (**B**) Multiple sequence alignments of *SULT2*-related genes from the indicated mammalian species. The gene accession codes are as follows: human *SULT2A1* (Q06520), human *SULT2B1* (O00204), mouse *Sult2a2* (L27121), cow *SULT2A1* (XM_005219462), dog *SULT2B1* (XM_038656443), cat *SULT2B1* (XM_023245236), panda *SULT2B1* (XM_019798249), ferret *SULT2B1* (XM_004767429), and microbat *SULT2A1* (XM_019658129).

Data S1. Chemical information of targeted bile acids.