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

Emi Ito[1,2], Shinsuke Inuki[3], Yoshihiro Izumi[4], Masatomo Takahashi[4], Yuki Dambayashi[3], Lisa Ciacchi[5], Wael Awad[5], Ami Takeyama[1,2], Kensuke Shibata[6], Shotaro Mori[1,2], Jeffrey Y. W. Mak[7], David P. Fairlie[7], Takeshi Bamba[4], Eri Ishikawa[1,2], Masamichi Nagae[1,2], Jamie Rossjohn[5,8] and Sho Yamasaki[1,2,9*]

[1]Department of Molecular Immunology, Research Institute for Microbial Diseases, Osaka University, Suita, Osaka, 565-0871, Japan

[2]Laboratory of Molecular Immunology, Immunology Frontier Research Center, Osaka University, Suita, Osaka, 565-0871, Japan

[3]Graduate School of Pharmaceutical Sciences, Kyoto University, Kyoto, Kyoto, 606-8501, Japan

[4]Division of Metabolomics, Medical Institute of Bioregulation, Kyushu University, Fukuoka, Fukuoka, 812-8582, Japan

[5]Infection and Immunity Program and Department of Biochemistry and Molecular Biology, Biomedicine Discovery Institute, Monash University, Clayton, Victoria, Australia

[6]Department of Microbiology and Immunology, Graduate School of Medicine, Yamaguchi University, Ube, Yamaguchi, 755-8505, Japan

[7]Institute for Molecular Bioscience, University of Queensland, Brisbane, Qld 4072, Australia.

[8]Institute of Infection and Immunity, Cardiff University, School of Medicine, Heath Park, Cardiff, UK

[9]Center for Infectious Disease Education and Research (CiDER), Osaka University, Suita, Osaka, 565-0871, Japan

*Corresponding author. Email: yamasaki@biken.osaka-u.ac.jp

Abstract: Mucosal-associated invariant T (MAIT) cells are innate-like T cells that recognize bacterial riboflavin–based metabolites as activating antigens. Although MAIT cells are found in tissues, it is unknown if any host tissue-derived antigen(s) exist. Here we report that a sulfated bile acid, cholic acid 7-sulfate (CA7S), binds the non-classical MHC class I protein MR1 and is recognized by MAIT cells. CA7S is a host-derived metabolite whose levels were reduced by more than 98% in germ-free mice. Deletion of the sulfotransferase 2a family of enzymes (*Sult2a1-8*) responsible for CA7S synthesis reduced the number of thymic MAIT cells in mice. Moreover, recognition of CA7S induced MAIT cell survival and the expression of a homeostatic gene signature. By contrast, recognition of a previously described foreign antigen, 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU), drove MAIT cell proliferation and the expression of inflammatory genes. Thus, CA7S is an endogenous antigen for MAIT cells, which promotes their development and function.

One-sentence summary: Cholic acid sulfate binds to MR1 and is recognized by MAIT cells.

INTRODUCTION

Mucosal-associated invariant T (MAIT) cells are innate-like T cells that recognize riboflavin-based metabolites, such as 5-OP-RU, presented by MHC class I-related protein (MR1) as foreign antigens through semi-invariant T cell receptors (TCR) (*1–5*). Although conventional T cells recognize self-derived weak antigens for selection and maintenance, microbiome-derived metabolites mediate the development of MAIT cells, as evidenced by the fact that MAIT cell numbers are decreased in germ-free mice (*6*, *7*). As the microbiome has a strong effect on endogenous metabolite levels (*8*), it is possible that host-derived antigen(s) may also contribute to MAIT cell development.

In many mammalian species, MAIT cells are most abundant in the liver (*9*, *10*) and are mainly localized in the hepatic sinusoid around bile ducts (*11–13*). In contrast to circulating MAIT cells, liver MAIT cells exhibit an activated phenotype despite being in a non-proliferating state (*11*, *14*, *15*). This suggests that weak MR1 ligand(s) may be expressed in these specific regions. MAIT cell ligands of this type have not yet been identified however, although previous studies have suggested the presence of self-derived ligands for MR1 (*16–18*).

In this study, we report that sulfated bile acid is presented by MR1 and recognized by MAIT TCRs to promote MAIT cell homeostasis. Deletion of enzyme required for its synthesis in mice impaired MAIT cell development.

RESULTS

Cholic acid 7-sulfate (CA7S) is a MAIT cell ligand presented by MR1

Not all MR1 ligands act as agonistic antigens to MAIT cells (*4*, *19*, *20*). To simultaneously evaluate MR1 binding and antigen activity, we established an NFAT-GFP reporter T cell line expressing both MR1 and a MAIT TCR (fig. S1A). A known MAIT cell antigen, 5-OP-RU, induced the dose-dependent activation of these reporter cells (fig. S1B) and marginal upregulation of MR1 on the cell surface (fig. S1B). By contrast, a non-stimulatory MR1 ligand, acetyl-6-formylpterin (Ac-6-FP), potently enhanced MR1 surface expression (fig. S1B) as previously reported (*21*, *22*) without inducing GFP expression. Thus, the reporter cells can function as both antigen-presenting cells and responder T cells.

To search for host-derived ligand(s) for MAIT cells, we used this cell line to screen fractionated intestinal extracts from specific-pathogen-free (SPF) mice. Among the 100 fractions separated by reversed-phase column chromatography (table S1A), three major peaks were detected around fractions #11, 42, and 84 (Fig. 1A). To determine which peak corresponded to a known microbe-derived antigen, we subjected 5-OP-RU to the same column separation process (Fig. 1B). The activities of fractions #11 and #42 were again observed (Fig. 1B) corresponding to the expected retention of 5-OP-RU and its related lumazine-derivatives, respectively (*5*).

To determine the chemical structure of the activating molecule in fraction #84, we further purified this fraction by hydrophilic interaction chromatography (HILIC) and collected a single fraction (designated as #84-45) (table S1B and fig. S1C). The activity was nuclease- and protease-resistant, suggesting a metabolite(s) other than polypeptides or oligonucleic acids (fig. S1D). High-resolution mass spectrometry (HRMS) analysis revealed a major peak at *m/z* 487.2368 in negative ion mode, suggesting a small compound (Fig. 1C). We next performed 1D- and 2D-NMR) spectroscopic analysis (Fig. 1D and fig. S2, A to D) showed that the compound is a cholic acid analogue with a functional group attached to the hydroxy group at

the C7 position (fig. S2E). The molecular formula estimated from the main peak at *m/z* 487.2368 was $C_{24}H_{40}O_8S$ $[M - H]^-$, suggesting that the major component of the fraction was cholic acid 7-sulfate (CA7S) (Fig. 1E). Indeed, HRMS/MS analysis confirmed the presence of a sulfate group (Fig. 1F). We next synthesized CA7S to authenticate the assignment (fig. S3A). The ^1H and ^{13}C NMR spectra of fraction #84-45 agreed with those of the synthetic CA7S (Fig. 1D and fig. S2, A, F, and G), and liquid chromatography HRMS (LC/HRMS) analysis showed identical retention times for fraction #84-45 and the CA7S standard (Fig. 1, F and G, and fig. S2, H and I). Synthetic CA7S activated reporter cells expressing a MAIT TCR, but the efficacy and potency were weaker than known antigens (Fig. 1H). As absolute concentrations of unstable 5-OP-RU converted from 5-amino-6-D-ribitylaminouracil (5-A-RU) cannot be estimated precisely (*23*), more stable synthetic ribityllumazine, 7-methyl-8-D-ribityllumazine (RL-7-Me), was used for comparison (*24*). The activity of CA7S was approximately 5% of RL-7-Me, indicating that CA7S is a weak agonist to MAIT cells (fig. S1E). Synthetic CA7S also elevated cell-surface MR1 expression (Fig. 1, I and J). CA7S was recognized by several different MAIT TCRs derived from mouse (m) and human (h) on m/h MR1 (fig. S1F). Thus, CA7S appears to be a ligand for MAIT cells presented by MR1.

CA7S is present in various SPF mouse tissues

CA7S is a bile acid metabolite biosynthesized by host sulfotransferase 2a (Fig. 2, A and B) (*25*). As bile acid metabolism is largely influenced by symbiotic bacteria, we next examined the effect of microbiota on intestinal bile acid metabolites including CA7S. To assess this, intestines from SPF and germ-free (GF) mice were analyzed using targeted metabolomics (data S1). The concentration of many bile acid derivatives were affected by symbiotic bacteria (Fig. 2A) that mediate the deconjugation/dehydroxylation of bile acids (Fig. 2B) (*26–28*). Among these, CA7S were decreased by more than 98% under GF conditions (Fig. 2A), presumably due to the lack of deconjugation of tauro CA7S (TCA7S) by symbiotic bacteria (Fig. 2A) (*29*), as TCA7S was increased in GF mice (Fig. 2A).

In addition to the intestines, CA7S was detected in the gallbladders, livers, and thymi of SPF mice. CA7S levels were detectable, but lower, in all GF tissues examined (Fig. 2C). LC-MS/MS analyses confirmed that other regioisomers of CA7S, such as CA3S and CA12S could not be detected in mouse tissue (table S2) as previously reported (*25*). Thus, CA7S is a host-derived bile acid metabolite, but its abundance is largely dependent on symbiotic bacteria.

Cholic acid sulfate is a unique agonistic bile acid metabolite

CA7S is biosynthesized by sulfate-conjugation of cholic acid (CA), which is a major component of bile acid released into the intestine from the gallbladder and which aids absorption of dietary lipids through micellization (*30*). However, CA itself did not activate reporter cells expressing MAIT TCR (Fig. 3, A and C), suggesting that the sulfate group was required for agonistic function.

In addition to the hydroxy group at position 7 (7-OH) of CA, 3-OH and 12-OH groups are also potential sulfation sites. We therefore also synthesized these isomers (fig. S3A). CA3S, which is present in some mammalian species (*25*), but not CA12S, activated MAIT reporter cells in the presence of MR1 (Fig. 3, A and C). By contrast, all three CA sulfate forms (CA3S, CA7S, and CA12S) upregulated the cell-surface expression of MR1 to a similar extent (Fig. 3, B and C), suggesting that CA sulfates may have the ability to bind to MR1.

As primary mouse bile acids are often conjugated with taurine, we next synthesized taurine-conjugated forms of CA3S (TCA3S) and CA7S (TCA7S) (fig. S3B) and examined the effect

of C24 modification on activity. Neither TCA, TCA3S, nor TCA7S showed any substantial activity either as an antigen or an MR1 ligand (Fig. 3, A to C). Thus, this amino acid conjugation of CA, which incorporates the addition of a hydrophilic group at position 24, abolishes any MR1-binding capability.

Primary bile acids are further dehydroxylated by symbiotic bacteria, producing secondary bile acids. However, the sulfated forms of secondary bile acids, such as deoxy CA3S (DCA3S), litho CA3S (LCA3S), and taurolitho CA3S (TLCA3S) did not have activity in reporter cells (Fig. 3, A and C). Thus, the sulfation of the 3- or 7-OH of cholic acid is the key structural feature for MAIT cell activation.

CA7S is presented by MR1 in a manner similar to bacterial ligands

We next investigated the mode of CA7S presentation by MR1. MAIT TCR recognition of CA7S was inhibited by blocking anti-MR1 antibodies and Ac-6-FP (*21*) in a dose-dependent manner (Fig. 4, A and B), implying that CA7S is presented by MR1 in a conventional manner (*4*, *5*) despite lacking a carbonyl group that could form Schiff base with MR1. We therefore assessed the possible MR1-binding mode by introducing mutations in MR1. Most mutations affecting 5-OP-RU binding to MR1 also affected CA7S binding, consistent with CA7S and 5-OP-RU potentially binding to a similar region of MR1 (Fig. 4C). We next assessed the competitive ligand-binding to MR1 of CA7S in the presence of the fluorescent ligand, JYM20, which has been reported to bind in the A′ pocket of MR1 (*31*). CA7S outcompeted JYM20 with a similar potency to that of the reported ligand, diclofenac (DCF) (*19*) (Fig. 4D). TCRα Y95 was also essential for the recognition of CA7S, as has also been reported for 5-OP-RU (Fig. 4E) (*5*), although this mutant TCR could still respond normally to anti-TCRβ monoclonal antibody stimulation (fig. S4). Thus, although CA7S differs in both size and structure from other reported MR1 ligands (*4*, *5*, *18*, *19*), it is presented by the MR1 A′ pocket and is recognized by MAIT TCRs.

Thymic MAIT cell numbers are reduced in *Sult2a*-deficient mice

CA7S is biosynthesized by the SULT2A family of enzymes, which incorporate a sulfate group into CA (*32*). As eight isoforms of SULT2A (*Sult2a1-8*) exist in a cluster on mouse chromosome 7A1, we generated mice lacking all eight isoforms using the CRISPR/Cas9 system (*Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice) (Fig. 5, A and B). Targeted metabolomics analyses revealed that CA7S was significantly decreased in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice, whereas other bile acid metabolites were unchanged (Fig. 5C). Nevertheless, *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice were viable and showed no macroscopic abnormalities, suggesting that CA7S is not a vital metabolite in mice. As SULT2A is abundantly expressed in the liver (*33*, *34*), we confirmed the lack of expression of all *Sult2a* isoforms in hepatocytes from *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice via single-cell RNA sequencing (scRNA-seq) (Fig. 5, D and E). However, hepatocyte and resident immune cell clusters were similar in WT and *Sult2a*-deficient mice (Fig. 5F). Thus, CA7S deficiency does not grossly affect the development of hepatocytes and adjacent immune cells, allowing us to use *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice to investigate the contribution of CA7S to MAIT cell development and function.

A significant decrease in the number of MR1−5-OP-RU tetramer^{+} (MR1−5-OP-RU tet^{+}) thymic MAIT cells was observed in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice compared with littermate controls (Fig. 6, A and B). Thymic MAIT cells can be divided into three developmental stages. The most mature stage (stage 3) was dominant in WT mice (*35*), whereas this stage was significantly less frequent in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice (Fig. 6, C and D). Mature thymic MAIT cells skew to a MAIT17

subset, an IL-17–producing MAIT cells, which requires more signaling for development (*36*), whereas MAIT17 was less dominant in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice, as assessed by surface marker expression (Fig. 6, E and F). scTCR-RNA-seq analyses of WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ thymocytes expressing MAIT clonotypes (*Trav1-Traj9/12/33*) also supported the reduction of MAIT17 (Fig. 6G). Thus, thymic MAIT cell development is impaired under CA7S-deficient conditions.

Because MAIT cells are highly abundant in the livers of many mammalian species (*9*, *10*, *37*), we also examined MR1−5-OP-RU tet^{+} mature MAIT cells in this organ. The frequency of liver MAIT cells was not significantly different in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice (fig. S5) and most MAIT cell clones in the livers of both WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice expressed previously reported MAIT1 signature genes (*38*) (Fig. 6H). However, genes significantly downregulated in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice compared to WT mice included those involved in TCR signaling and costimulation (e.g., *Lcp2* encoding SLP-76, *Cd84* encoding Slamf5, *Nfatc1, Nfkbia*, and *Tab2*) (*38–40*) and cell adhesion (*Icam1* and *Cd164*) (*38*). The common gamma chain (γ_c) (*Il2rg*) (a component of IL-7R–mediated licensing of MAIT cells (*14*)), Bcl-2 family proteins (*Mcl1*) (*41*), and other anti-apoptotic proteins (*Bag1, Birc6* and *Rb1cc1*) (*42–45*) were also downregulated, suggesting that liver MAIT cells in the absence of SULT2A may not receive sufficient signals to support their homeostasis (Fig. 6I). By contrast, these genes were not downregulated in iNKT and conventional T cells in the livers of *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice (Fig. 6I). Thus, CA7S specifically contributes to MAIT cell homeostasis.

CA7S contributes to the maintenance of canonical human MAIT cells

To confirm whether CA7S acts on primary human MAIT cells, peripheral blood mononuclear cells (PBMCs) were stimulated with CA7S. To compare the functional differences for each agonist, we stimulated cells with minimum effective concentrations of 5-OP-RU (10 μM) and CA7S (1000 μM) based on their reporter cell activity. As assessed by Cell Trace Violet (CTV) dilution, 5-OP-RU induced cell division of MR1−5-OP-RU tet^{+} MAIT cells, whereas CA7S did not (Fig. 7, A and B). However, the number of MAIT cells cultured in the presence of CA7S was four times greater than that of control cultures without CA7S at day 6 (Fig. 7C). 5-OP-RU, but not CA7S, increased the surface expression of CD161 (*KLRB1*) during culture (fig. S6A), suggesting that these ligands induced different responses via the MAIT TCR. To test this hypothesis, we compared gene expression signatures using scTCR-RNA-seq of MAIT cells that were either unstimulated (Unstim) or stimulated with 5-OP-RU or CA7S for 1 day. Uniform manifold approximation and projection (UMAP) clustering analysis of MAIT cells (*TRAV1-2*$^{+}$) demonstrated that gene expression signatures were distinct depending on the ligand. 5-OP-RU mainly induced cluster 1, whereas cluster 0 was enriched upon stimulation with CA7S. Both were distinct from the Unstim signature (clusters 3 and 7) (Fig. 7D). Comparison of clusters 0 and 1 revealed that distinct genes were upregulated upon stimulation with 5-OP-RU versus CA7S (Fig. 7E). *CD69, GZMB*, *IFNG*, and *TNF*, for example, were induced by 5-OP-RU (*46*, *47*), whereas *CXCR4* and *IL7R* were selectively induced upon CA7S stimulation (Fig. 7E). Indeed, CA7S induced surface expression of CXCR4, but not CD69, in the peripheral MAIT cells (Fig. 7F). Gene ontology enrichment analysis also supported this differential gene expression, as 5-OP-RU upregulated immune activation-related genes as previously reported (table S3) (*46–48*), whereas CA7S induced genes involved in wound healing (Top enrichment score, P=0.002) and negative regulation of immune activation (second and third enrichment score, P=0.002) (table S3). This differential regulation was also observed within the most frequent clonotypes expressing the canonical MAIT TCR pair (*TRAV1-2*-CAVRDSNYQLIW-*TRAJ33 – TRBV28-TRBJ2-5*) (Fig. 7G). 5-OP-RU preferentially induced genes in area 3 rather than 1, whereas CA7S-induced genes were included in area 2 rather than

4 (Fig. 7G). Thus, CA7S appears to trigger different signals than those induced by 5-OP-RU through the same MAIT TCR. We also performed scTCR-RNA-seq after long-term stimulation with CA7S and CA3S, which are both present in human. MAIT cells that survived in the presence of CA7S/3S expressed canonical MAIT TCRs, such as *TRAV1-2-TRAJ33/12/20* paired with *TRBV6-4/6-1/20-1*, which were also enriched upon 5-OP-RU stimulation (fig. S6B), suggesting that CA7S/3S is a ligand for canonical MAIT cells. However, differential gene expression signatures were also retained, as CA7S/3S induced homeostatic genes, such as *IL7R* (*14*, *49*), *KLF2* (*50*), and *TCF7* (*36*, *51*, *52*), which were not induced by 5-OP-RU (fig. S6, C and D). Thus, sulfated bile acid is a host-derived ligand distinct from the microbe-derived antigen 5-OP-RU, which contributes to the development and survival of canonical MAIT cells.

DISCUSSION

In the bile acid metabolic pathway, sulfation contributes to the excretion of excess bile acids through the gastrointestinal tract for bile acid homeostasis (*25*). However, the "active" physiological roles of CA7S have not been fully understood (*29*, *53*). The current study suggests that an abundant byproduct of one physiological process can be utilized for the regulation of other physiological processes. CA sulfates are mainly produced in the liver and small intestine, both of which express SULT2A (*33*, *34*, *54*) and where MAIT cells are enriched in humans (*55*), suggesting that this metabolite may play an important role in MAIT cell homeostasis in these tissues. Although there are some limitations to exploring the effect of CA7S deficiency using laboratory mice in which MAIT cells are rare, further analyses across multiple species or use of MAIT-enriched mice (*56–58*) may help clarify this issue. As common genetic polymorphisms for human SULT2A1 result in the reduction of its enzymatic activity (*59*, *60*), future work should investigate whether there is any association between MAIT cell development/function and the presence of these polymorphisms.

Impairment of thymic MAIT cells, particularly stage 3, in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice suggests that CA7S plays a preferential role in MAIT cell maturation. The reduction of CA7S in GF thymi may explain impaired thymic MAIT cell development in addition to the lack of 5-OP-RU (*6*, *7*). Although CA7S was detected in the thymus, it is currently unknown how this metabolite comes to be located in this lymphoid organ. In GF mice (*3*, *6*, *7*, *35*), SPF *Sult2a*-deficient mice (Fig. 6B) and even GF *Sult2a*-deficient mice (fig. S7), a few thymic MAIT cells were detected. It is therefore possible that other unknown endogenous ligand(s) may also contribute to the early development of thymic MAIT cells. An alternative but not mutually exclusive hypothesis is that an intrinsic binding affinity of MR1 to MAIT TCRs (*21*) may also play a role in thymic selection in a ligand-independent manner.

Conventional T cells are positively selected within the thymus by moderate interaction with weak-affinity peptides (*61*). In the periphery, these weak peptides cannot induce activation, but contribute to the maintenance or homeostasis of mature T cells (*52*, *62*, *63*). CA7S also induced homeostatic gene expression in the peripheral MAIT cells, although it did not trigger inflammatory responses. The role of CA7S in liver MAIT cell function is not clearly understood, but we observed a difference in gene expression in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. CA7S may be involved in tissue residency, local survival or functional priming/licensing within the tissue. Indeed, one of the CA7S-induced molecules, CXCR4, contributes to the residency of MAIT cells in the liver (*64*), although further investigation is warranted. Liver MAIT cells respond to biliary epithelial cells (BEC) in an MR1-dependent manner (*11*). Thus, it is possible that hepatic cells, such as BEC, liver sinusoidal endothelial cells (LSECs), or hepatocytes may act as a niche for MAIT cells by presenting bile acid metabolites on MR1, as MAIT cells are

localized in the liver sinusoid around bile ducts (*9*). Alternatively, as BEC are normally separated from MAIT cells in the sinusoid under steady-state conditions, MAIT cells may recognize CAS presented by BEC when the LSEC barrier is breached by trauma or infection, therefore serving as a possible damage sensor leading to tissue repair (*7*, *46–48*). Considering their unique tissue distribution, MAIT cells may need to be resistant to bile acid stresses and indeed MAIT cells express high levels of the ATP-binding cassette subfamily (ABCB1) (*37*), also called multidrug resistance 1 polypeptide (MDR1), which is known to protect T cells in the ileum from bile acid-driven stresses (*65*).

CA7S may accumulate in disease settings such as cholestasis-related inflammation, as bile sulfation is required for excretion of excessive bile (*25*, *66*). Indeed, correlation between MAIT cell activation and primary sclerosing cholangitis (PSC) or primary biliary cholangitis (PBC) have been reported (*67*, *68*). Whether excessive CA7S is involved in the development of such inflammatory diseases through the dysregulation of MAIT cells is an important area for future research.

Mammalian species lacking MR1 lose TRAV1, suggesting the coevolution of the MR1–MAIT cell system (*69*). Intriguingly, some such species also lack *Sult2a* homologues (table S4, A and B). Both MR1 and MAIT cells are absent in fish and amphibians (*70*). These animals utilize larger C27 bile alcohols as major bile salts (*71*) and these are unlikely to bind the MR1 pocket even if sulfated. Further phylogenetic studies will be needed to validate a possible link between sulfated bile acids and the MR1–MAIT system.

In conclusion, this study extends previous work underscoring the structural diversity of ligands presented by MR1 as compared with other MHC class I-related molecules (*72*). In order to monitor a wide variety of exogenous and endogenous metabolites, mammals may have acquired a “promiscuity” in presenting ligands with various structures and affinities (*4*, *18–20*, *73*) on MR1, which has a certain intrinsic affinity to MAIT TCRs (*21*). Further identification of unexpected structures of natural ligands therefore is likely possible and may contribute to further understanding of MAIT cell physiology.

MATERIALS AND METHODS

Study design

The goal of this study is to identify and characterize host-derived MAIT cell ligand(s). To enable this end, we established a platform utilizing column chromatography–based separation and mass spectrometry together with a sensitive reporter assay. To determine the chemical structure of the ligand, we used NMR spectroscopy and chemical synthesis. We established mice lacking a particular bile acid metabolite using mouse genetics and metabolomics. Sample sizes were determined on the basis of previous studies and pilot experiments in our laboratory and are indicated in the figure legend. All experiments were performed three times unless otherwise indicated in the figure legends. Mice of the same sex and similar ages were randomly assigned to each experimental group. For in vivo experiments, the investigators were blinded to the genotype of the mouse samples. None of the data were excluded from our analysis.

Mice

Sult2a$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice were generated using the CRISPR/Cas9 system, as previously described (*74*). In brief, Cas9 mRNA and sgRNAs were microinjected into fertilized embryos of BDF1 mice. Homozygous KO mice were born from a heterozygous intercross. All mice were genotyped 2-4 weeks after birth, using PCR with specific primers (P1, 5′-ACCTGGAAAGACTAATACTTGCC-3′, P2, 5′-CCCCACAGAGACAGACCAAT-3′ and P3, 5′-CAAATGATCTCTCAATGAGTTCAC-3′). *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice were used for phenotypic analyses in 3-to-8-week-old with wild-type littermate controls (*Sult2a*$^{+/+}$ or *Sult2a*$^{+/\Delta 1\text{-}8}$). Germ-free mice (C57BL/6NJcl [Gf]) were purchased from CLEA Japan Inc and analyzed in the same weeks of age compared with SPF mice. Germ-free *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice are generated by in vitro fertilization maintained in vinyl isolators within the facility in CLEA Japan. All animal protocols were approved by the committee of Ethics on Animal Experiment, Research Institute for Microbial Diseases, Osaka University (Biken-AP-R03-17-0).

Reagents

5-OP-RU for the cellular assay was prepared with 5-A-RU (Toronto Research Chemical) and methylglyoxal (Sigma-Aldrich) as previously described (*5*). Ac-6-FP (Cat. No. 11.418) were purchased from Schircks laboratories. 2-hydroxy-5-methoxybenzaldehyde (HMB) (Cat. No. 146862), DCF (Cat. No. D6899) and epigallocatechin gallate (EGCG) (CAS No. 93894) were supplied by Sigma-Aldrich. For FP assays, 5-OP-RU (*75*), JYM20 (*76*), and RL-6-Me-7-OH (*4*) were synthesized as previously described. The NLV peptide (Cat. No. 181329) was synthesized by GL Biochem. TCA3S, LCA3S and TLCA3S were purchased from Cayman. DCA3S and THBA were purchased from Avanti. CA and TCA were purchased from Nacalai Tesque. DCA were purchased from FUJIFILM Wako. Synthetic CA7S, CA3S, CA12S, TCA3S, TCA7S and RL-7-Me were prepared as described in supplementary methods. Anti-human/mouse/rat MR1 (26.5), anti-mouse CD3 (2C11), anti-mouse CD19 (6D5), anti-mouse TCRβ (H57-597), anti-mouse CD45 (30-F11), anti-mouse CD319 (4G2), anti-mouse CD138 (281-2), anti-mouse CD24 (M1/69), Mouse IgG2a, κ isotype control (MOPC-173), Armenian Hamster IgG isotype control (HTK888), anti-human CD3 (HIT3a), anti-human CD161 (HP-3G10), anti-human CD69 (FN50), anti-human CXCR4 (12G5), anti-mouse CD16/32 (93) and anti-human TotalSeq-C Hashtags (LNH-94; 2M2) antibodies were purchased from BioLegend. Anti-mouse CD44 (IM7) were from BD Biosciences. Dead cells were stained with Propidium iodide (PI) (Sigma-Aldrich) or 7-amino-actinomycin D (7AAD) (Biolegend). Mouse or human MR1 tetramers loaded with 5-OP-RU or 6-FP conjugated to PE or APC were provided by the NIH tetramer core facility (Emory University) (*5*).

Cells

TCRα- and TCRβ-chain cDNA sequences were cloned into retroviral vector pMX-IRES-rat CD2. TCRαβ sequences for mouse MAIT TCR #1 clonotype (*TRAV1*-CAVRDSNYQLIW-*TRAJ33 – TRBV13-3-TRBJ2-1*) and mouse MAIT TCR #2 clonotype (*TRAV1*-CAVRDSNYQLIW-*TRAJ33 – TRBV13-3-TRBJ2-7*) were from previous reports (*3*, *77*). αβTCR sequences for human MAIT TCR #1 clonotype (*TRAV1-2*-CAVKDSNYQLIW-*TRAJ33 – TRBV6-2-TRBJ2-2*) were from Protein Data Bank (4L4T). Human and mouse MR1 were cloned into retroviral vector pMX-IRES-human CD8 (*78*). These plasmids were co-transfected into Phoenix packaging cells using PEI MAX (Polysciences). Supernatant containing retroviruses was used for infection (*79*) into TCR-deficient mouse T cell hybridoma with an NFAT-GFP reporter gene (*80*). MR1 and TCRα mutants were constructed by site-directed mutagenesis using KOD-FX (TOYOBO) following the manufacturer's instruction.

Ligand extraction and purification

Mouse large intestines were frozen in liquid nitrogen. They were then crushed and extracted with 6 ml of double-distilled water. The supernatant was ultrafiltered (<100 kDa) and fractionated by HPLC (JASCO LC-NetII/ADC) using COSMOSIL PBr 4.6 mm I.D.×250 mm (Nacalai Tesque) and TSK gel Amide-80 5 µm 4.6 mm I.D.×250 mm (TOSOH Bioscience) as described in fig. S1, C and D. Purified samples were dried using a freeze-dryer (EYELA FDU-1200).

Reporter cell assay

For the stimulation of reporter cells, tissue-extracted samples and synthetic compounds were dissolved in water and added to each well of the 96-well plate in the absence of methylglyoxal. Some water-insoluble samples were dissolved in ethanol or chloroform/methanol (2:1, vol/vol) and coated on plates as described (*81*). NFAT-GFP reporter cells (1-3×10^4 cells/100 µl/well) were cultured for 6−20 hours at 37°C and analyzed for GFP and MR1 expression by flow cytometry. Tissue extracted fractions were treated with DNase I (Roche) or Trypsin (Promega) overnight at 37°C before 5-OP-RU and CA7S stimulation in fig. S1F. For the inhibition assay, anti-MR1 antibody (0.001-1 µg/ml) and Ac-6-FP (0.1-100 µM) were added to reporter cells for 1 hour before the stimulation with 5-OP-RU and CA7S.

Human studies

The institutional review boards of Osaka University (approval number 29-4-10) approved blood draw protocols for healthy individuals. The research was performed in accordance with all relevant guidelines and regulations.

Stimulation of peripheral blood mononuclear cells (PBMCs)

To prepare PBMCs, whole blood was collected in heparin-coated tubes and centrifuged to separate the cellular fraction and plasma using Lymphocyte Separation Solution (d=1.077) (Nacalai Tesque). PBMCs were labeled with Cell Trace Violet (Invitrogen), according to the manufacturer's instructions. PBMCs were cultured with 5-OP-RU, CA3S, and CA7S in the absence of cytokines in RPMI 1640 medium (Sigma-Aldrich) supplemented with 10% fetal bovine serum, 100 U/ml of penicillin (Sigma-Aldrich), 100 µg/ml of streptomycin (MP Biomedicals), and 50 µM 2-mercaptoethanol (Nacalai Tesque). At the indicated days after stimulation, PBMCs were analyzed by flow cytometry (Attune NxT Flow Cytometer, Thermo Fisher Scientific) and $CD3^+CD161^+MR1^-$5-OP-RU tetramer$^+$ cells were sorted by SH800S Cell Sorter (Sony Biotechnology) for sc-TCR-RNA-seq analyses. For survival assays, 1×10^6 PBMCs were left untreated or treated with 1 mM CA7S for 6 days in 150-µl cultures using 96-

well round-bottom plates. At day 0, 150 µl of culture suspension was analyzed by flow cytometry and the number of living (PI-negative) $CD3^+CD161^+MR1-5\text{-}OP\text{-}RU\text{-}tetramer^+$ MAIT cells was defined as 100%. Every 2 days, the same staining was performed and the survival ratio was calculated as a percentage against day 0.

Preparation of MAIT cells from mouse tissues

Thymi were dissociated by homogenizer to make single-cell suspensions. Single-cell suspensions of thymocyte were stained with APC-mouse MR1−5-OP-RU tetramer (NIH Tetramer Core Facility, Emory University Vaccine Center, Atlanta, USA). These were enriched by anti-APC microbeads according to the manufacturer's instructions (Miltenyi Biotec). Livers were dissociated over 40-µm nylon mesh. Liver mononuclear cells were isolated by density gradient centrifugation on a discontinuous 40%–80% Percoll (Cytiva) gradient.

Single-cell-based transcriptome and TCR repertoire analysis

Single-cell transcriptome and TCR repertoire analysis were performed using Chromium Controller (10x Genomics) according to the manufacturer's instructions as previously described (*82*). Gene expression-based clustering was performed using the Seurat R package (v3.1; (*83*)). Cells with a mitochondrial content >10% and cells with <200 or >4000 genes detected were considered outliers (dying cells and empty droplets and doublets, respectively) and filtered out. The cluster analysis based on UMAP plot, volcano plot and differential expression analysis were performed by BBrowser (BioTuring).

NMR spectroscopy

Fraction #84-45 was dissolved in 750 µl of D_2O containing approximately 0.1% (v/v) MeCN as an internal standard. NMR spectra were acquired on a JEOL ECZ600R spectrometer and were referenced internally according to residual solvent signals (HOD, MeCN). Signals characteristic of cholic acid analogues were observed (three methyl protons at δ 0.73, 0.93, 0.97 ppm, two methylene protons at δ 2.32, 2.45 ppm, three methine protons at δ 3.52, 4.07, 4.51 ppm in the ^{1}H NMR spectrum and a carbonyl carbon at δ 180.8 ppm in the ^{13}C NMR spectrum, fig. 1D and S2A). Through analyses including 2D-NMR (fig S2, B to E), two peaks at δ 4.07 and 4.51 ppm in the ^{1}H NMR spectrum were assigned to methine protons H-12 and H-7, respectively. The H-7 signal appeared more deshielded than predicted, indicating that a functional group is linked to the hydroxy group located at the C7 position of the cholic acid skeleton.

LC-MS/MS quantitative analysis

Bile acids were extracted from each of the mouse tissue samples by 1 ml of extraction solvent (ice-cold methanol) after supplementing of deuterium-labeled internal standard (IS) compounds (d_4-CA; d_4-CDCA; d_4-TCA; d_4-TCDCA; d_4-LCA; d_4-DCA; d_4-GDCA; and d_5-TLCA). Samples were mixed vigorously by vortex for 1 min followed by 5 min of sonication. The samples were centrifuged at 16,000*g* for 5 min at 4°C. The supernatant (800 µl) was then collected into clean tubes. For the targeted quantitative analysis of bile acids (data S1), the supernatant was diluted with methanol according to the abundance of bile acid contained in each tissue to prepare an analytical sample. A triple-quadrupole mass spectrometer equipped with an electrospray ionization (ESI) ion source (LCMS-8060; Shimadzu Corporation) was used in MRM mode. The conditions for the LC-MS/MS analysis were as follows: column, metal free peek-coated InertSustain C18 (2.1×150 mm; particle size, 3 µm; GL Sciences Inc.); column temperature, 50°C; flow rate, 0.3 ml/min; mobile phase, water/acetonitrile/acetic acid (3:1:0.05, v/v/v) (A) and methanol/isopropanol/acetic acid (1:19:0.05, v/v/v) (B); gradient curve, 0% B at 0 min, 38% B at 17 min, 100% B at 25 min, 100% B at 35 min, 0% B at 35.1

min, and 0% B at 50 min; injection volume, 5 µl; mass analysis mode, negative ion mode, electrospray voltage: –3.0 kV; nebulizer gas flow rate: 2.0 liters per min, drying gas flow rate: 10.0 liters per min, desolvation temperature: 250°C, heat block temperature: 400°C, and detector voltage: 2.16 kV. The MRM mode and a dwell time of 10 ms per channel were used. Other optimized MRM parameters for bile acids are shown in data S1. The absolute contents of bile acids were calculated based on the calibration curves. The MRM calibration curves were generated from the triplicate analyses of these standard solutions using the chromatographic peak area of each analyte to that of the IS.

MR1 Protein expression and purification

Recombinant human MR1 was produced in vitro as previously described (*19*, *84*). Briefly, DNA encoding the extracellular domains of the MR1 heavy chain and β2m were transformed into *Escherichia coli* BL21(DE3). The proteins were subsequently overexpressed and purified as inclusion bodies. Refolding of MR1-HMB was performed via addition of 80 µM HMB, 120 mg of denatured inclusion body MR1 heavy chain, and 60 mg of β2m protein in 1 liter of refold buffer consisting of 0.1 M Tris pH 8.5, 2.5 M of urea, 0.4 M of L-arginine, 2 mM Na-EDTA, 0.5 mM oxidized glutathione, and 5 mM reduced glutathione for 16 hours at 4°C. The refolding solution was dialyzed against 10 mM Tris pH 8.0. The MR1-HMB complex was purified via size exclusion (S200 Superdex 16/600, GE Healthcare and then dialyzed against 10 mM Bis-Tris propane pH 6.5, 150 mM NaCl at 4°C for 16 hours. The resultant MR1-empty was buffer exchanged to 10 mM Tris pH 8.0 and purified using anion exchange (HiTrap Q HP, GE Healthcare) chromatography as previously described (*31*).

Fluorescence polarization (FP) inhibition assay

Bile acids (CA7S, CA3S, and THBA) were each solubilized in ultra-pure water. The remainder of ligands were dissolved in dimethyl sulfoxide (Sigma-Aldrich). FP assays were conducted with serial dilutions of each ligand starting at a maximum concentration of 50 mM for CA7S, CA3S, and THBA, 20 mM DCF, 5 µM Ac-6-FP, 500 nM 5-OP-RU, 500 µM RL-6-Me-7-O, 5 mM EGCG, and 1 mM NLV peptide. Each compound was incubated with the previously reported synthetic fluorescent MR1-specific ligand, JYM20 (10 nM) and human MR1-empty (100 nM) in assay buffer (25 mM HEPES pH 7.5, 150 mM NaCl, and 5 mM Na-EDTA) (*31*). FP was measured after 24 hours at 25°C using the PHERAstar microplate reader (BMG LabTech) as previously reported (*31*). Ligand binding curves were graphed as a sigmoidal concentration response curve using Prism Version 9.3.0 (GraphPad Software Inc.). IC_{50} values for binding affinity were calculated at the ligand concentration required for 50% of inhibition for JYM20 binding to MR1-empty.

Non-targeted LC-HRMS/MS analysis

Non-targeted liquid chromatography high-resolution tandem mass spectrometry (LC-HRMS/MS) analyses were performed using a Nexera X2 UHPLC system (Shimadzu Co.) coupled with a Q Exactive high-performance benchtop quadrupole Orbitrap mass spectrometer (Thermo Fisher Scientific Inc.). LC analytical conditions were identical to those used in the targeted bile acid analysis method using LC-MS/MS. The full scanning HRMS analysis conditions were as follows: polarity, positive and negative ionization; sheath gas flow rate, 50 arb; auxiliary (Aux) gas flow rate, 10 arb; spray voltage for positive ion mode, 4.0 kV; spray voltage for negative ion mode; −3.0 kV; capillary temperature, 250°C; S-lens level, 60; heater temperature, 400°C; mass resolution, 70000; automatic gain control (AGC) target, 3×10^6; maximum injection time, 200 ms; and scan range, 150−1500 (*m/z*). The conditions for data dependent MS^2 (dd-MS^2) were as follows: mass resolution, 17500; AGC target, 5×10^4; trap fill time, 80 ms; isolation width, ± 1.2 Da; fixed first mass, *m/z* 50; stepped normalized collision

energy, 10, 30, and 45 eV; intensity threshold of precursor ions for dd-MS^2 analysis, 1×10^4; apex trigger, 2−4 s; and dynamic exclusion, 2 s. The Compound Discoverer ver. 3.0 (Thermo Fisher Scientific Inc.) was used for HRMS data processing. Infusion analysis was performed using an ion trap/time-of-flight high-resolution mass spectrometry (HRMS) fitted with an electrospray ionization ion source (Shimadzu). The HRMS operating conditions were as follows: polarity, negative; electrospray voltage, −3.5 kV; curve desolvation line (CDL) temperature 200°C; heat block temperature, 200°C; nebulizing gas (N_2) flow, 1.5 liters per min; drying gas (N_2) pressure, 0.1 MPa; scan range, *m/z* 100–1500.

Statistical analysis

An unpaired two-tailed Student's *t* test or one-way ANOVA with the Tukey's multiple comparison test were performed for the statistical analyses using GraphPadPrism (Version 9.1.0, GraphPad Software Inc.). Asterisks denote the level of statistical significance (*$P<0.05$, **$P<0.01$, ***$P<0.005$, ****$P<0.001$). *P*-values were adjusted for multiple comparisons using the Benjamini–Hochberg method (FDR<0.05) using GraphPad Prism version 9.1.0.

REFERENCES AND NOTES

1. D. I. Godfrey, H. F. Koay, J. McCluskey, N. A. Gherardin, The biology and functional importance of MAIT cells. *Nat. Immunol.* **20**, 1110–1128 (2019).

2. F. Legoux, M. Salou, O. Lantz, MAIT cell development and functions: the microbial connection. *Immunity*. **53**, 710–723 (2020).

3. E. Treiner, L. Duban, S. Bahram, M. Radosavljevic, V. Wanner, F. Tilloy, P. Affaticati, S. Gilfillan, O. Lantz, Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature*. **423**, 1018 (2003).

4. L. Kjer-Nielsen, O. Patel, A. J. Corbett, J. Le Nours, B. Meehan, L. Liu, M. Bhati, Z. Chen, L. Kostenko, R. Reantragoon, N. A. Williamson, A. W. Purcell, N. L. Dudek, M. J. McConville, R. A. J. O'Hair, G. N. Khairallah, D. I. Godfrey, D. P. Fairlie, J. Rossjohn, J. McCluskey, MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature*. **491**, 717–723 (2012).

5. A. J. Corbett, S. B. G. Eckle, R. W. Birkinshaw, L. Liu, O. Patel, J. Mahony, Z. Chen, R. Reantragoon, B. Meehan, H. Cao, N. A. Williamson, R. A. Strugnell, D. Van Sinderen, J. Y. W. Mak, D. P. Fairlie, L. Kjer-Nielsen, J. Rossjohn, J. McCluskey, T-cell activation by transitory neo-antigens derived from distinct microbial pathways. *Nature*. **509**, 361–365 (2014).

6. F. Legoux, D. Bellet, C. Daviaud, Y. El Morr, A. Darbois, K. Niort, E. Procopio, M. Salou, J. Gilet, B. Ryffel, A. Balvay, A. Foussier, M. Sarkis, A. El Marjou, F. Schmidt, S. Rabot, O. Lantz, Microbial metabolites control the thymic development of mucosal-associated invariant T cells. *Science (80-.).* **366**, 494–499 (2019).

7. M. G. Constantinides, V. M. Link, S. Tamoutounour, A. C. Wong, P. J. Perez-Chaparro, S. J. Han, Y. E. Chen, K. Li, S. Farhat, A. Weckel, S. R. Krishnamurthy, I. Vujkovic-Cvijin, J. L. Linehan, N. Bouladoux, E. D. Merrill, S. Roy, D. J. Cua, E. J. Adams, A. Bhandoola, T. C. Scharschmidt, J. Aubé, M. A. Fischbach, Y. Belkaid, MAIT cells are

imprinted by the microbiota in early life and promote tissue repair. *Science (80-.).* **366**, 445 (2019).

8. E. Blacher, M. Levy, E. Tatirovsky, E. Elinav, Microbiome-modulated metabolites at the interface of host immunity. *J. Immunol.* **198**, 572–580 (2017).

9. A. Kurioka, L. J. Walker, P. Klenerman, C. B. Willberg, MAIT cells: new guardians of the liver. *Clin. Transl. Immunol.* **5**, e98 (2016).

10. M. D. Edmans, T. K. Connelley, S. Jayaraman, C. Vrettou, M. Vordermeier, J. Y. W. Mak, L. Liu, D. P. Fairlie, E. A. Maze, T. Chrun, P. Klenerman, S. B. G. Eckle, E. Tchilian, L. Benedictus, Identification and phenotype of MAIT cells in cattle and their response to bacterial infections. *Front. Immunol.* **12**, 1–17 (2021).

11. H. C. Jeffery, B. Van Wilgenburg, A. Kurioka, K. Parekh, K. Stirling, S. Roberts, E. E. Dutton, S. Hunter, D. Geh, M. K. Braitch, J. Rajanayagam, T. Iqbal, T. Pinkney, R. Brown, D. R. Withers, D. H. Adams, P. Klenerman, Y. H. Oo, Biliary epithelium and liver B cells exposed to bacteria activate intrahepatic MAIT cells through MR1. *J. Hepatol.* **64**, 1118–1127 (2016).

12. E. von Seth, C. L. Zimmer, M. Reuterwall-Hansson, A. Barakat, U. Arnelo, A. Bergquist, M. A. Ivarsson, N. K. Björkström, Primary sclerosing cholangitis leads to dysfunction and loss of MAIT cells. *Eur. J. Immunol.* **48**, 1997–2004 (2018).

13. M. J. Lett, H. Mehta, A. Keogh, T. Jaeger, M. Jacquet, K. Powell, M. A. Meier, I. Fofana, H. Melhem, J. Vosbeck, G. Cathomas, A. Heigl, M. H. Heim, E. Burri, K. D. Mertz, J. H. Niess, O. Kollmar, C. J. Zech, R. Ivanek, U. Duthaler, P. Klenerman, D. Stroka, M. F. Sinnreich, Stimulatory MAIT cell antigens reach the circulation and are efficiently metabolised and presented by human liver cells. *Gut*, 2526–2538 (2022).

14. X.-Z. Tang, J. Jo, A. T. Tan, E. Sandalova, A. Chia, K. C. Tan, K. H. Lee, A. J. Gehring, G. De Libero, A. Bertoletti, IL-7 licenses activation of human liver intrasinusoidal mucosal-associated invariant T cells. *J. Immunol.* **190**, 3142–3152 (2013).

15. R. Lamichhane, F. Munro, T. W. R. Harrop, S. M. de la Harpe, P. K. Dearden, A. J. Vernall, J. L. McCall, J. E. Ussher, Human liver-derived MAIT cells differ from blood MAIT cells in their metabolism and response to TCR-independent activation. *Eur. J. Immunol.* **51**, 879–892 (2021).

16. M. Lepore, A. Kalinichenko, S. Calogero, P. Kumar, B. Paleja, M. Schmaler, V. Narang, F. Zolezzi, M. Poidinger, L. Mori, G. de Libero, Functionally diverse human T cells recognize non-microbial antigens presented by MR1. *Elife*. **6**, 1–22 (2017).

17. M. D. Crowther, G. Dolton, M. Legut, M. E. Caillaud, A. Lloyd, M. Attaf, S. A. E. Galloway, C. Rius, C. P. Farrell, B. Szomolay, A. Ager, A. L. Parker, A. Fuller, M. Donia, J. McCluskey, J. Rossjohn, I. M. Svane, J. D. Phillips, A. K. Sewell, Genome-wide CRISPR–Cas9 screening reveals ubiquitous T cell cancer targeting via the monomorphic MHC class I-related protein MR1. *Nat. Immunol.* **21**, 178–185 (2020).

18. M. H. Young, L. U'Ren, S. Huang, T. Mallevaey, J. Scott-Browne, F. Crawford, O. Lantz, T. H. Hansen, J. Kappler, P. Marrack, L. Gapin, MAIT cell recognition of MR1 on bacterially infected and uninfected Cells. *PLoS One*. **8**, 6–11 (2013).

19. A. N. Keller, S. B. G. Eckle, W. Xu, L. Liu, V. A. Hughes, J. Y. W. Mak, B. S. Meehan, T. Pediongco, R. W. Birkinshaw, Z. Chen, H. Wang, C. D'Souza, L. Kjer-Nielsen, N. A. Gherardin, D. I. Godfrey, L. Kostenko, A. J. Corbett, A. W. Purcell, D. P. Fairlie, J. McCluskey, J. Rossjohn, Drugs and drug-like molecules can modulate the function of mucosal-associated invariant T cells. *Nat. Immunol.* **18**, 402–411 (2017).

20. M. J. Harriff, C. McMurtrey, C. A. Froyd, H. Jin, M. Cansler, M. Null, A. Worley, E. W. Meermeier, G. Swarbrick, A. Nilsen, D. A. Lewinsohn, W. Hildebrand, E. J. Adams, D. M. Lewinsohn, MR1 displays the microbial metabolome driving selective MR1-restricted T cell receptor usage. *Sci. Immunol.* **3**, eaao2556 (2018).

21. S. B. G. Eckle, R. W. Birkinshaw, L. Kostenko, A. J. Corbett, H. E. G. McWilliam, R. Reantragoon, Z. Chen, N. A. Gherardin, T. Beddoe, L. Liu, O. Patel, B. Meehan, D. P. Fairlie, J. A. Villadangos, D. I. Godfrey, L. Kjer-Nielsen, J. McCluskey, J. Rossjohn, A molecular basis underpinning the T cell receptor heterogeneity of mucosal-associated invariant T cells. *J. Exp. Med.* **211**, 1585–1600 (2014).

22. H. E. G. McWilliam, S. B. G. Eckle, A. Theodossis, L. Liu, Z. Chen, J. M. Wubben, D. P. Fairlie, R. A. Strugnell, J. D. Mintern, J. McCluskey, J. Rossjohn, J. A. Villadangos, The intracellular pathway for the presentation of Vitamin B-related antigens by the antigen-presenting molecule MR1. *Nat. Immunol.* **17**, 531–537 (2016).

23. J. Y. W. Mak, W. Xu, R. C. Reid, A. J. Corbett, B. S. Meehan, H. Wang, Z. Chen, J. Rossjohn, J. McCluskey, L. Liu, D. P. Fairlie, Stabilizing short-lived Schiff base derivatives of 5-aminouracils that activate mucosal-associated invariant T cells. *Nat. Commun.* **8**, 14599 (2017).

24. T. Matsuoka, C. Motozono, A. Hattori, H. Kakeya, S. Yamasaki, S. Oishi, H. Ohno, S. Inuki, The effects of 5-OP-RU stereochemistry on its stability and MAIT-MR1 axis. *ChemBioChem*. **22**, 672–678 (2021).

25. Y. Alnouti, Bile acid sulfation: A pathway of bile acid elimination and detoxification. *Toxicol. Sci.* **108**, 225–246 (2009).

26. M. Funabashi, T. L. Grove, M. Wang, Y. Varma, M. E. McFadden, L. C. Brown, C. Guo, S. Higginbottom, S. C. Almo, M. A. Fischbach, A metabolic pathway for bile acid dehydroxylation by the gut microbiome. *Nature*. **582**, 566–570 (2020).

27. B. V. Jones, M. Begley, C. Hill, C. G. M. Gahan, J. R. Marchesi, Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 13580–13585 (2008).

28. I. Choucair, I. Nemet, L. Li, M. A. Cole, S. M. Skye, J. D. Kirsop, M. A. Fischbach, V. Gogonea, J. Mark Brown, W. H. Wilson Tang, S. L. Hazen, Quantification of bile acids: A mass spectrometry platform for studying gut microbe connection to metabolic diseases. *J. Lipid Res.* **61**, 159–177 (2020).

29. S. N. Chaudhari, J. N. Luo, D. A. Harris, H. Aliakbarian, L. Yao, D. Paik, R. Subramaniam, A. A. Adhikari, A. H. Vernon, A. Kiliç, S. T. Weiss, J. R. Huh, E. G. Sheu, A. S. Devlin, A microbial metabolite remodels the gut-liver axis following bariatric surgery. *Cell Host Microbe*. **29**, 408–424 (2021).

30. D. W. Russell, Fifty years of advances in bile acid synthesis and metabolism. *J. Lipid Res.* **50**, S120–S125 (2009).

31. C. J. H. Wang, W. Awad, L. Liu, J. Y. W. Mak, N. Veerapen, P. T. Illing, A. W. Purcell, S. B. G. Eckle, J. Mccluskey, G. S. Besra, D. P. Fairlie, Quantitative af fi nity measurement of small molecule ligand binding to major histocompatibility complex class-I – related. *J Biol Chem.* **298**, 1–12 (2022).

32. J. Huang, S. P. Bathena, J. Tong, M. Roth, B. Hagenbuch, Y. Alnouti, Kinetic analysis of bile acid sulfation by stably expressed human sulfotransferase 2A1 (SULT2A1). *Xenobiotica*. **40**, 184–194 (2010).

33. Y. Alnouti, C. D. Klaassen, Tissue distribution and ontogeny of sulfotransferase enzymes in mice. *Toxicol. Sci.* **93**, 242–255 (2006).

34. L. Feng, Y. L. Yuen, J. Xu, X. Liu, M. Y. C. Chan, K. Wang, W. P. Fong, W. T. Cheung, S. S. T. Lee, Identification and characterization of a novel PPARα-regulated and 7α-hydroxyl bile acid-preferring cytosolic sulfotransferase mL-STL (Sult2a8). *J. Lipid Res.* **58**, 1114–1131 (2017).

35. H. F. Koay, N. A. Gherardin, A. Enders, L. Loh, L. K. Mackay, C. F. Almeida, B. E. Russ, C. A. Nold-Petry, M. F. Nold, S. Bedoui, Z. Chen, A. J. Corbett, S. B. G. Eckle, B. Meehan, Y. D'Udekem, I. E. Konstantinov, M. Lappas, L. Liu, C. C. Goodnow, D. P. Fairlie, J. Rossjohn, M. M. Chong, K. Kedzierska, S. P. Berzins, G. T. Belz, J. McCluskey, A. P. Uldrich, D. I. Godfrey, D. G. Pellicci, A three-stage intrathymic development pathway for the mucosal-associated invariant T cell lineage. *Nat. Immunol.* **17**, 1300–1311 (2016).

36. H. F. Koay, S. Su, D. Amann-Zalcenstein, S. R. Daley, I. Comerford, L. Miosge, C. E. Whyte, I. E. Konstantinov, Y. D'Udekem, T. Baldwin, P. F. Hickey, S. P. Berzins, J. Y. W. Mak, Y. Sontani, C. M. Roots, T. Sidwell, A. Kallies, Z. Chen, S. Nüssing, K. Kedzierska, L. K. Mackay, S. R. McColl, E. K. Deenick, D. P. Fairlie, J. McCluskey, C. C. Goodnow, M. E. Ritchie, G. T. Belz, S. H. Naik, D. G. Pellicci, D. I. Godfrey, A divergent transcriptional landscape underpins the development and functional branching of MAIT cells. *Sci. Immunol.* **4**, eaay6039 (2019).

37. M. Dusseaux, E. Martin, N. Serriari, I. Péguillet, V. Premel, D. Louis, M. Milder, L. Le Bourhis, C. Soudais, E. Treiner, O. Lantz, Human MAIT cells are xenobiotic-resistant, tissue-targeted, CD161 hi IL-17-secreting T cells. *Blood*. **117**, 1250–1259 (2011).

38. M. Salou, F. Legoux, J. Gilet, A. Darbois, A. Du Halgouet, R. Alonso, W. Richer, A. G. Goubet, C. Daviaud, L. Menger, E. Procopio, V. Premel, O. Lantz, A common transcriptomic program acquired in the thymus defines tissue residency of MAIT and NKT subsets. *J. Exp. Med.* **216**, 133–151 (2019).

39. J. Dias, E. Leeansyah, J. K. Sandberg, Multiple layers of heterogeneity and subset diversity in human MAIT cell responses to distinct microorganisms and to innate cytokines. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E5434–E5443 (2017).

40. H. Flament, M. Rouland, L. Beaudoin, A. Toubal, L. Bertrand, S. Lebourgeois, C. Rousseau, P. Soulard, Z. Gouda, L. Cagninacci, A. C. Monteiro, M. Hurtado-Nedelec, S. Luce, K. Bailly, M. Andrieu, B. Saintpierre, F. Letourneur, Y. Jouan, M. Si-Tahar, T. Baranek, C. Paget, C. Boitard, A. Vallet-Pichard, J. F. Gautier, N. Ajzenberg, B. Terrier, F. Pène, J. Ghosn, X. Lescure, Y. Yazdanpanah, B. Visseaux, D. Descamps, J. F. Timsit, R. C. Monteiro, A. Lehuen, Outcome of SARS-CoV-2 infection is linked to MAIT cell activation and cytotoxicity. *Nat. Immunol.* **22**, 322–335 (2021).

41. I. Dzhagalov, A. Dunkle, Y.-W. He, The anti-apoptotic Bcl-2 family member Mcl-1 promotes T lymphocyte survival at multiple stages. *J. Immunol.* **181**, 521–528 (2008).

42. S. Takayama, T. Sato, S. Krajewski, K. Kochel, S. Irie, J. A. Milian, J. C. Reed, Cloning and functional analysis of BAG-1: A novel Bcl-2-binding protein with anti-cell death activity. *Cell*. **80**, 279–284 (1995).

43. C. G. Low, I. S. U. Luk, D. Lin, L. Fazli, K. Yang, Y. Xu, M. Gleave, P. W. Gout, Y. Wang, BIRC6 Protein, an Inhibitor of Apoptosis: Role in Survival of Human Prostate Cancer Cells. *PLoS One*. **8** (2013), doi:10.1371/journal.pone.0055837.

44. H. Xia, W. Wang, J. Crespo, I. Kryczek, W. Li, S. Wei, Z. Bian, T. Maj, M. He, R. J. Liu, Y. He, R. Rattan, A. Munkarah, J. L. Guan, W. Zou, Suppression of FIP200 and autophagy by tumor-derived lactate promotes naïve T cell apoptosis and affects tumor immunity. *Sci. Immunol.* **2** (2017), doi:10.1126/sciimmunol.aan4631.

45. T. M. Young, C. Reyes, E. Pasnikowski, C. Castanaro, C. Wong, C. E. Decker, J. Chiu, H. Song, Y. Wei, Y. Bai, B. Zambrowicz, G. Thurston, C. Daly, Autophagy protects tumors from T cell-mediated cytotoxicity via inhibition of TNFα-induced apoptosis. *Sci. Immunol.* **5**, 1–18 (2020).

46. T. S. C. Hinks, E. Marchi, M. Jabeen, M. Olshansky, A. Kurioka, T. J. Pediongco, B. S. Meehan, L. Kostenko, S. J. Turner, A. J. Corbett, Z. Chen, P. Klenerman, J. McCluskey, Activation and in vivo evolution of the MAIT cell transcriptome in mice and humans reveals tissue repair functionality. *Cell Rep.* **28**, 3249-3262.e5 (2019).

47. T. Leng, H. D. Akther, C. P. Hackstein, K. Powell, T. King, M. Friedrich, Z. Christoforidou, S. McCuaig, M. Neyazi, C. V. Arancibia-Cárcamo, J. Hagel, F. Powrie, R. S. Peres, V. Millar, D. Ebner, R. Lamichhane, J. Ussher, T. S. C. Hinks, E. Marchi, C. Willberg, P. Klenerman, TCR and inflammatory signals tune human MAIT cells to exert specific tissue repair and effector functions. *Cell Rep*. **28**, 3077-3091.e5 (2019).

48. R. Lamichhane, M. Schneider, S. M. de la Harpe, T. W. R. Harrop, R. F. Hannaway, P. K. Dearden, J. R. Kirman, J. D. A. Tyndall, A. J. Vernall, J. E. Ussher, TCR- or cytokine-activated CD8+ mucosal-associated invariant T cells are rapid polyfunctional effectors that can coordinate immune responses. *Cell Rep.* **28**, 3061-3076.e5 (2019).

49. A. Willing, J. Jäger, S. Reinhardt, N. Kursawe, M. A. Friese, Production of IL-17 by MAIT cells is increased in multiple sclerosis and is associated with IL-7 receptor expression. *J. Immunol.* **200**, 974–982 (2018).

50. O. A. Odumade, M. A. Weinreich, S. C. Jameson, K. A. Hogquist, Krüppel-like factor 2 regulates trafficking and homeostasis of γδ T cells. *J. Immunol.* **184**, 6060–6066 (2010).

51. L. A. Mielke, Y. Liao, E. B. Clemens, M. A. Firth, B. Duckworth, Q. Huang, F. F. Almeida, M. Chopin, H. F. Koay, C. A. Bell, S. Hediyeh-Zadeh, S. L. Park, D. Raghu, J. Choi, T. L. Putoczki, P. D. Hodgkin, A. E. Franks, L. K. Mackay, D. I. Godfrey, M. J. Davis, H. H. Xue, V. L. Bryant, K. Kedzierska, W. Shi, G. T. Belz, TCF-1 limits the formation of Tc17 cells via repression of the MAF–RORγt axis. *J. Exp. Med.* **216**, 1682–1699 (2019).

52. D. Pais Ferreira, J. G. Silva, T. Wyss, S. A. Fuertes Marraco, L. Scarpellino, M. Charmoy, R. Maas, I. Siddiqui, L. Tang, J. A. Joyce, M. Delorenzi, S. A. Luther, D. E. Speiser, W. Held, Central memory CD8+ T cells derive from stem-like Tcf7hi effector cells in the absence of cytotoxic differentiation. *Immunity*. **53**, 985–1000 (2020).

53. S. N. Chaudhari, D. A. Harris, H. Aliakbarian, J. N. Luo, M. T. Henke, R. Subramaniam, A. H. Vernon, A. Tavakkoli, E. G. Sheu, A. S. Devlin, Bariatric surgery reveals a gut-restricted TGR5 agonist with anti-diabetic effects. *Nat. Chem. Biol.* **17**, 20–29 (2021).

54. W. Teubner, W. Meinl, S. Florian, M. Kretzschmar, H. Glatt, Identification and localization of soluble sulfotransferases in the human gastrointestinal tract. *Biochem. J.* **404**, 207–215 (2007).

55. N. M. Provine, P. Klenerman, MAIT cells in health and disease. *Annu. Rev. Immunol.* **38**, 203–228 (2020).

56. E. Martin, E. Treiner, L. Duban, L. Guerri, H. Laude, C. Toly, V. Premel, A. Devys, I. C. Moura, F. Tilloy, S. Cherif, G. Vera, S. Latour, C. Soudais, O. Lantz, Stepwise development of MAIT cells in mouse and human. *PLoS Biol.* **7**, 0525–0536 (2009).

57. Y. Cui, K. Franciszkiewicz, Y. K. Mburu, S. Mondot, L. Le Bourhis, V. Premel, E. Martin, A. Kachaner, L. Duban, M. A. Ingersoll, S. Rabot, J. Jaubert, J. P. De Villartay, C. Soudais, O. Lantz, Mucosal-associated invariant T cell-rich congenic mouse strain allows functional evaluation. *J. Clin. Invest.* **125**, 4171–4185 (2015).

58. H. Wakao, K. Yoshikiyo, U. Koshimizu, T. Furukawa, K. Enomoto, T. Matsunaga, T. Tanaka, Y. Yasutomi, T. Yamada, H. Minakami, J. Tanaka, A. Oda, T. Sasaki, R. Wakao, O. Lantz, T. Udagawa, Y. Sekiya, K. Higuchi, N. Harada, K. Nishimura, M. Ohtaka, M. Nakanishi, H. Fujita, Expansion of functional human mucosal-associated invariant T cells via reprogramming to pluripotency and redifferentiation. *Cell Stem Cell*. **12**, 546–558 (2013).

59. B. A. Thomae, B. W. Eckloff, R. R. Freimuth, E. D. Wieben, R. M. Weinshilboum, Human sulfotransferase SULT2A1 pharmacogenetics: Genotype-to-phenotype studies. *Pharmacogenomics J.* **2**, 48–56 (2002).

60. S. A. Gohal, M. I. Rasool, A. F. Bairam, E. S. Alatwi, F. A. Alherz, M. S. Abunnaja, A. A. El Daibani, K. Kurogi, M. C. Liu, Effects of genetic polymorphisms on the sulfation of dehydroepiandrosterone and pregnenolone by human cytosolic sulfotransferase SULT2A1. *J. Biochem.* **170**, 419–426 (2021).

61. K. A. Hogquist, S. C. Jameson, W. R. Heath, J. L. Howard, M. J. Bevan, F. R. Carbone, T cell receptor antagonist peptides induce positive selection. *Cell*. **76**, 17–27 (1994).

62. C. Tanchot, F. A. Lemonnier, B. Pérarnau, A. A. Freitas, B. Rocha, Differential requirements for survival and proliferation of CD8 naive or memory T cells. *Science (80-.).* **276**, 2057–2062 (1997).

63. C. Viret, F. S. Wong, C. A. Janeway, Designing and Maintaining the Mature TCR Repertoire. *Immunity*. **10**, 559–568 (1999).

64. Z. Chen, S. Liu, C. He, J. Sun, L. Wang, H. Chen, F. Zhang, CXCL12-CXCR4-mediated chemotaxis supports accumulation of mucosal-associated invariant T cells into the liver of patients with PBC. *Front. Immunol.* **12**, 1–10 (2021).

65. W. Cao, H. Kayama, M. L. Chen, A. Delmas, A. Sun, S. Y. Kim, E. S. Rangarajan, K. McKevitt, A. P. Beck, C. B. Jackson, G. Crynen, A. Oikonomopoulos, P. N. Lacey, G. J. Martinez, T. Izard, R. G. Lorenz, A. Rodriguez-Palacios, F. Cominelli, M. T. Abreu, D. W. Hommes, S. B. Koralov, K. Takeda, M. S. Sundrud, The xenobiotic transporter Mdr1 enforces T cell homeostasis in the presence of intestinal bile acids. *Immunity*. **47**, 1182–1196 (2017).

66. C. L. Corbett, T. C. Bartholomew, B. H. Billing, J. A. Summerfield, Urinary excretion of bile acids in cholestasis: evidence for renal tubular secretion in man. *Clin. Sci. Mol. Med. Suppl.* **61**, 773–780 (1981).

67. L. Valestrand, F. Zheng, S. H. Hansen, J. Øgaard, J. R. Hov, N. K. Björkström, T. H. Karlsen, X. Jiang, E. Melum, Bile from patients with primary sclerosing cholangitis contains mucosal-associated invariant T-cell antigens. *Am. J. Pathol.* **192**, 629–641 (2022).

68. X. Jiang, M. Lian, Y. Li, W. Zhang, Q. Wang, Y. Wei, J. Zhang, W. Chen, X. Xiao, Q. Miao, Z. Bian, D. Qiu, J. Fang, A. A. Ansari, P. S. C. Leung, R. L. Coppel, R. Tang, M. E. Gershwin, X. Ma, The immunobiology of mucosal-associated invariant T cell (MAIT) function in primary biliary cholangitis: regulation by cholic acid-induced interleukin-7. *J. Autoimmun.* **90**, 64–75 (2018).

69. P. Boudinot, S. Mondot, L. Jouneau, L. Teyton, M. P. Lefranc, O. Lantz, Restricting nonclassical MHC genes coevolve with TRAV genes used by innate-like T cells in mammals. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E2983–E2992 (2016).

70. K. Tsukamoto, J. E. Deakin, J. A. M. Graves, K. Hashimoto, Exceptionally high conservation of the MHC class I-related gene, MR1, among mammals. *Immunogenetics*. **65**, 115–124 (2013).

71. A. F. Hofmann, L. R. Hagey, M. D. Krasowski, Bile salts of vertebrates: Structural variation and possible evolutionary signifi cance. *J. Lipid Res.* **51**, 226–246 (2010).

72. W. Awad, J. Le Nours, L. Kjer-Nielsen, J. McCluskey, J. Rossjohn, Mucosal-associated invariant T cell receptor recognition of small molecules presented by MR1. *Immunol. Cell Biol.* **96**, 588–597 (2018).

73. M. Salio, W. Awad, N. Veerapen, C. Gonzalez-Lopez, C. Kulicke, D. Waithe, A. W. J. Martens, D. M. Lewinsohn, J. V. Hobrath, L. R. Cox, J. Rossjohn, G. S. Besra, V. Cerundolo, Ligand-dependent downregulation of MR1 cell surface expression. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 10465–10475 (2020).

74. W. Xue, S. Chen, H. Yin, T. Tammela, T. Papagiannakopoulos, N. S. Joshi, W. Cai, G. Yang, R. Bronson, D. G. Crowley, F. Zhang, D. G. Anderson, P. A. Sharp, T. Jacks, CRISPR-mediated direct mutation of cancer genes in the mouse liver. *Nature*. **514**, 380–384 (2014).

75. J. Y. W. Mak, L. Liu, D. P. Fairlie, Chemical modulators of mucosal associated invariant T cells. *Acc. Chem. Res.* **54**, 3462–3475 (2021).

76. H. E. G. McWilliam, J. Y. W. Mak, W. Awad, M. Zorkau, S. Cruz-Gomez, H. J. Lim, Y. Yan, S. Wormald, L. F. Dagley, S. B. G. Eckle, A. J. Corbett, H. Liu, S. Li, S. J. J. Reddiex, J. D. Mintern, L. Liu, J. McCluskey, J. Rossjohn, D. P. Fairlie, J. A. Villadangos, Endoplasmic reticulum chaperones stabilize ligand-receptive MR1 molecules for efficient presentation of metabolite antigens. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 24974–24985 (2020).

77. F. Tilloy, E. Treiner, S. H. Park, C. Garcia, F. Lemonnier, H. De La Salle, A. Bendelac, M. Bonneville, O. Lantz, An invariant T cell receptor α chain defines a novel TAP-independent major histocompatibility complex class Ib-restricted α/β T cell subpopulation in mammals. *J. Exp. Med.* **189**, 1907–1921 (1999).

78. S. Yamasaki, E. Ishikawa, M. Sakuma, K. Ogata, K. Sakata-Sogawa, M. Hiroshima, D. L. Wiest, M. Tokunaga, T. Saito, Mechanistic basis of pre-T cell receptor-mediated autonomous signaling critical for thymocyte development. *Nat. Immunol.* **7**, 67–75 (2006).

79. S. Yamasaki, E. Ishikawa, M. Sakuma, H. Hara, K. Ogata, T. Saito, Mincle is an ITAM-coupled activating receptor that senses damaged cells. *Nat. Immunol.* **9**, 1179–1188 (2008).

80. Y. Matsumoto, K. Kishida, M. Matsumoto, S. Matsuoka, M. Kohyama, T. Suenaga, H. Arase, A TCR-like antibody against a proinsulin-containing fusion peptide ameliorates type 1 diabetes in NOD mice. *Biochem. Biophys. Res. Commun.* **534**, 680–686 (2021).

81. E. Ishikawa, T. Ishikawa, Y. S. Morita, K. Toyonaga, H. Yamada, O. Takeuchi, T. Kinoshita, S. Akira, Y. Yoshikai, S. Yamasaki, Direct recognition of the mycobacterial glycolipid, trehalose dimycolate, by C-type lectin Mincle. *J. Exp. Med.* **206**, 2879–2888 (2009).

82. X. Lu, Y. Hosono, M. Nagae, S. Ishizuka, E. Ishikawa, D. Motooka, Y. Ozaki, N. Sax, Y. Maeda, Y. Kato, T. Morita, R. Shinnakasu, T. Inoue, T. Onodera, T. Matsumura, M. Shinkai, T. Sato, S. Nakamura, S. Mori, T. Kanda, E. E. Nakayama, T. Shioda, T. Kurosaki, K. Takeda, A. Kumanogoh, H. Arase, H. Nakagami, K. Yamashita, Y. Takahashi, S.

Yamasaki, Identification of conserved sars-cov-2 spike epitopes that expand public ctfh clonotypes in mild covid-19 patients. *J. Exp. Med.* **218**, 1–18 (2021).

83. C. Hafemeister, R. Satija, Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol.* **20**, 1–15 (2019).

84. W. Awad, G. J. M. Ler, W. Xu, A. N. Keller, J. Y. W. Mak, X. Y. Lim, L. Liu, S. B. G. Eckle, J. Le Nours, J. McCluskey, A. J. Corbett, D. P. Fairlie, J. Rossjohn, The molecular basis underpinning the potency and specificity of MAIT cell antigens. *Nat. Immunol.* **21**, 400–411 (2020).

85. J. Goto, H. Kato, F. Hasegawa, T. Nambara, Synthesis of monosulfates of unconjugated and conjugated bile acids. *Chem. Pharm. Bull.* **27**, 1402–1411 (1979).

86. R. Thakare, H. Gao, R. E. Kosa, Y. A. Bi, M. V. S. Varma, M. A. Cerny, R. Sharma, M. Kuhn, B. Huang, Y. Liu, A. Yu, G. S. Walker, M. Niosi, L. Tremaine, Y. Alnouti, A. D. Rodrigues, Leveraging of rifampicin-dosed cynomolgus monkeys to identify bile acid 3-O-sulfate conjugates as potential novel biomarkers for organic anion-transporting polypeptidess. *Drug Metab. Dispos.* **45**, 721–733 (2017).

87. Nonappa, U. Maitra, First chemical synthesis, aggregation behavior and cholesterol solubilization properties of pythocholic acid and 16α-hydroxycholic acid. *European J. Org. Chem.*, 3331–3336 (2007).

88. V. Del Amo, L. Siracusa, T. Markidis, B. Baragaña, K. M. Bhattarai, M. Galobardes, G. Naredo, M. N. Pérez-Payán, A. P. Davis, Differentially-protected steroidal triamines; scaffolds with potential for medicinal, supramolecular, and combinatorial chemistry. *Org. Biomol. Chem.* **2**, 3320–3328 (2004).

89. G. Kakiyama, A. Muto, M. Shimada, N. Mano, J. Goto, A. F. Hofmann, T. Iida, Chemical synthesis of 3β-sulfooxy-7β-hydroxy-24-nor-5-cholenoic acid: An internal standard for mass spectrometric analysis of the abnormal Δ5-bile acids occurring in Niemann-Pick disease. *Steroids*. **74**, 766–772 (2009).

90. C. Festa, B. Renga, C. D'Amore, V. Sepe, C. Finamore, S. De Marino, A. Carino, S. Cipriani, M. C. Monti, A. Zampella, S. Fiorucci, Exploitation of cholane scaffold for the discovery of potent and selective farnesoid X receptor (FXR) and G-protein coupled bile acid receptor 1 (GP-BAR1) ligands. *J. Med. Chem.* **57**, 8477–8495 (2014).

Acknowledgements: We thank D. Motooka, Y-C. Liu, D. Okuzaki, X. Lu, A. Hidaka, C. Motozono, M. Yasuda, A. Tanaka, Y. Harima, T. Kondo, M. Goto and M. Matsuda for experimental support; and C. Schutt for discussion. The MR1 tetramer technology was developed jointly by Prof. J. McCluskey, J.R., and D.P.F., and the material was produced by the National Institutes of Health Tetramer Core Facility as permitted to be distributed by the University of Melbourne.

Funding: This research was supported by AMED (21gm0910010 and 21ak0101070 (S.Y.)), JSPS KAKENHI (JP20H00505, JP22H05183 (S.Y.), JP20K06938, JP20H04773 (S.I.), JP22H05185 (Y.I.) and 21J22863 (E.It.)), JST-Mirai Program (JPMJMI20G1 (Y.I.)) and Chugai Pharmaceuticals. This research was also supported by Platform Project for Supporting

Drug Discovery and Life Science Research (Basis for Supporting Innovative Drug Discovery and Life Science Research (BINDS)) from AMED under Grant Number JP21am0101119 (support number 2774). This work was supported by the Australian Research Council (ARC; DP220102401, CE200100012), the National Institutes of Health (NIH) RO1 AI148407-01A1. W.A. was supported by Australian ARC Discovery Early Career Researcher Award (DECRA) fellowships (DE220101491). J.R. is supported by an NHMRC investigator grant (2008981). D.F. acknowledges ARC Centre of Excellence grant (CE200100012), NHMRC investigator grant (2009551) and US National Institutes of Health (RO1 AI14807-01A1) grant.

Author contributions: Conceptualization: S.Y. Investigation: E.It., M.T, A.T., E.Is., M.N., L.C., W.A., and K.S. Resources: Y.D., J.Y.W.M., and D.F.. Data curation: E.It., S.I., Y.I., M.N., L.C., W.A., T.B., and J.R. Writing—original draft: E.It. and S.Y.. Writing—review and editing: E.It., E.Is, S.Y., and J.R. Supervision: S.Y. and J.R. Funding Acquisition: S.I., Y.I., S.Y., and J.R.

Competing interests: J.R., J.Y.W.M., and D.F. are coinventors on patents describing MR1 tetramers and MR1–ligand complexes. The other authors declare no financial conflicts of interest.

Data and materials availability: Single-cell-based transcriptome data and bulk TCR-sequencing data have been deposited in Gene Expression Omnibus datasets under accession numbers GSE203394, GSE204917, GSE202930, and GSE228498. All other data needed to support the study conclusions are available in the main text or the supplementary materials.

Supplementary Materials

Supplementary Materials and Methods

Figs. S1 to S7

Tables S1 to S4

References (*85–90*)

Data S1

Fig. 1. Purification and structural determination of a ligand for MAIT cells. (**A** and **B**) Screening of MAIT cell agonists from mouse intestine. NFAT-GFP reporter cells expressing MAIT TCR and MR1 were stimulated with HPLC-separated fractions from SPF mouse intestine (A) and freshly prepared 5-OP-RU as a control (B) for 16−20 hours and analyzed by flow cytometry. (**C** and **D**) High-resolution mass spectra (HRMS) (C) and ^{1}H NMR spectra (600 MHz, D_2O) (D) of fraction # 84-45 from SPF mice intestine. In (D), impurities are denoted by asterisks. (**E**) Chemical structure of cholic acid 7-sulfate (CA7S). (**F** and **G**) HRMS/MS spectra (F) and extracted ion chromatogram (G) of fraction # 84-45. (**H** and **I**) NFAT-GFP reporter cells were stimulated with 5-OP-RU, RL-7-Me and CA7S. Percentages of GFP^{+} cells (H) and MR1 expression (I) were analyzed at 20 and 6 hours after stimulation, respectively. MR1 surface expression is presented as MFI values of stimulated cells subtracted with those of vehicle-treated unstimulated cells (ΔMFI). (**J**) MR1 surface expression at 0, 2, 4, 8 and 16 hours after stimulation with vehicle control (None), 5-OP-RU or CA7S. (H to J) Data are presented as the means±SD of triplicate assays and are representative of more than two independent experiments.

Fig. 2. Decrease in the level of CA7S in GF mice. (**A**) Quantification of bile acid metabolites in the cecum and cecal contents from SPF and GF mice. Blue and yellow arrows indicate host- and bacteria-mediated enzymatic responses, respectively. N.D., not detected (< 0.010 nmol/tissue). (**B**) Metabolic pathway map of bile acids in mice, related to (A). Sult, sulfotransferase; HDSH, hydroxysteroid dehydrogenase; BAAT, bile acid-CoA:amino acid N-acyltransferase; BACS, bile acid-CoA synthetase; Cyp, cytochrome P450; BSHs, bile salt hydrolases. (**C**) Tissue distribution of CA7S in various tissues of SPF and GF mice. N.D., not detected (<0.010 pmol/mg). (A and C) Data are presented as the means ± SD from experiments with three or more mice per group and are representative of two independent experiments. $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.005$, $^{****}P<0.001$, by two-tailed, unpaired Student's *t* tests with Benjamini–Hochberg correction.

Fig. 3. Structure–activity relationship of bile acid metabolites for MAIT cell activation. (**A** and **B**) MAIT TCR activation assays using NFAT-GFP reporter cells expressing mouse MAIT TCR and MR1. Reporter cells were stimulated with vehicle control, 5-OP-RU and analogues of CA7S (1); cholic acid 3-sulfate (CA3S: 2), cholic acid 12-sulfate (CA12S: 3), cholic acid (CA: 4), taurocholic acid (TCA: 5), deoxycholic acid (DCA: 6), taurocholic acid 3-sulfate (TCA3S: 7), taurocholic acid 7-sulfate (TCA7S: 8), deoxycholic acid 3-sulfate (DCA3S: 9), lithocholic acid 3-sulfate (LCA3S: 10) and taurolithocholic acid 3-sulfate (TLCA3S: 11). Due to their inherent cell toxicity as surfactants, several bile acids were added at lower concentrations (4, 5, 6, 10, and 11). Some water-insoluble bile acids (4, 5, 6, and 10) were coated on plates as described (*81*). 11 was added as DMSO solution due to its insolubility in both water and organic solvents. NFAT-GFP (A) and MR1 (B) expressions were evaluated at 20 hours and 6 hours after stimulation, respectively. (**C**) Structural formula of bile acid analogues related to (A and B). (A and B) Data are presented as individual values of duplicate assays and are representative of more than two independent experiments.

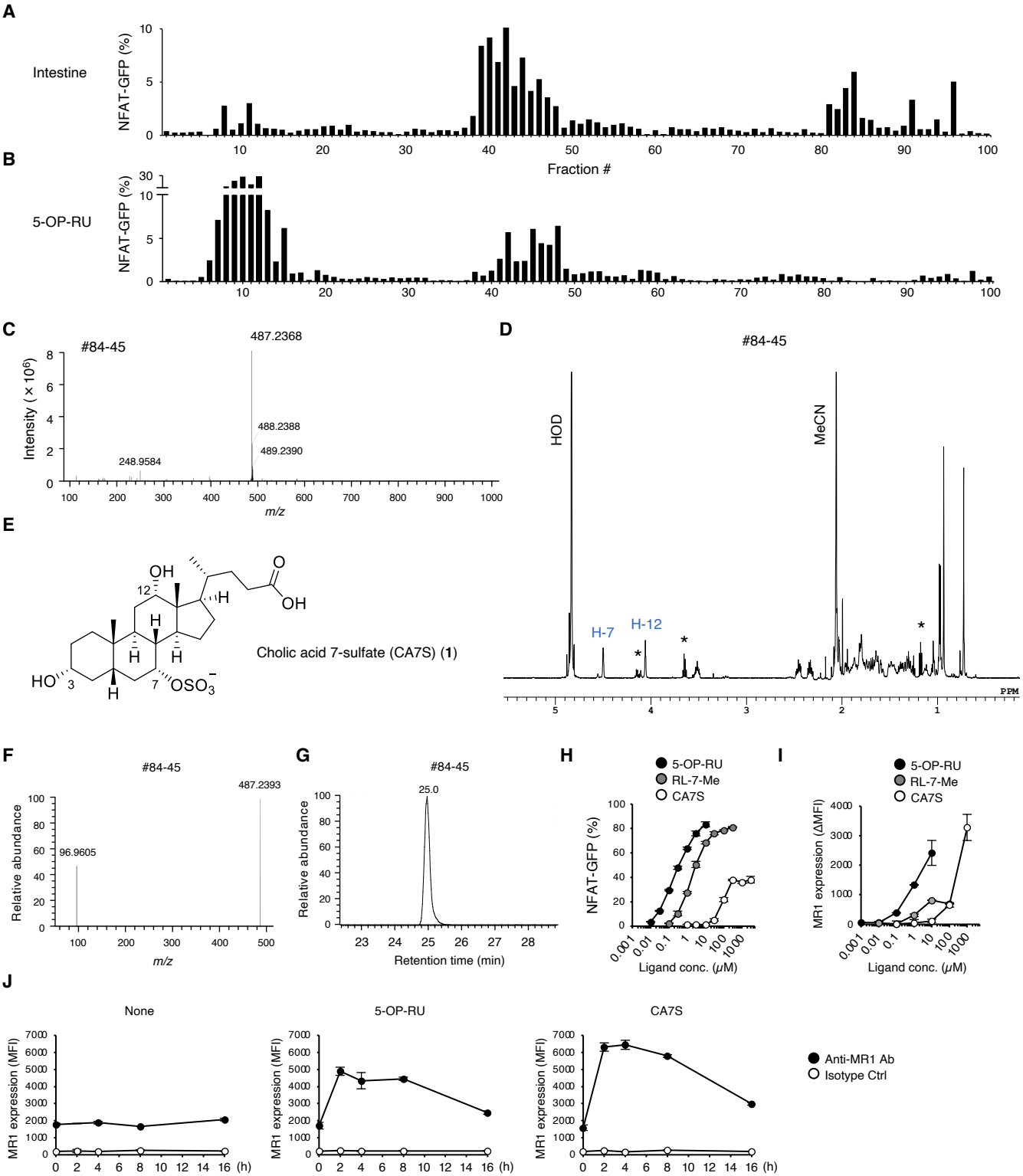
Fig. 4. Binding mode of CA7S to MR1 and MAIT TCR. (**A** and **B**) Inhibition assay of 5-OP-RU (0.5 μM) and CA7S (500 μM) by anti-MR1 Ab (26.5) (A) or Ac-6-FP (B). Percentages of GFP^+ cells were shown. (**C**) Effect of hMR1 mutations (Y7A, R9A, K43A, Y62A, L66A, W69A, M72A, R79A, R94A, W156A and W164A) on recognition of ligands. Cells expressing hMR1 mutants and a MAIT TCR were stimulated with vehicle control, 5-OP-RU or CA7S and analyzed by flowcytometry after 6 hours. MR1 expression was shown as MFI of anti-MR1 staining subtracted by isotype control. (**D**) MR1-restricted ligands affinities (IC_{50}) determined by FP assay (*31*). Left panel: Titration curves of strong MR1 binders (5-OP-RU and Ac-6-FP), moderate MR1 binder (RL-6-Me-7-OH), weak MR1 binders (CA7S, CA3S and diclofenac (DCF)) and MR1 non-binding substances (epigallocatechin gallate (EGCG) and NLV peptide) are displayed. Tetrahydroxy bile acid (THBA) was used as a bile acid control with similar hydrophilicity to CA7S and CA3S. The table represents a summary of IC_{50} for all investigated compounds. (**E**) Effect of MAIT TCRα mutation on ligand recognition. Reporter cells transfected with vector alone (Mock), MAIT TCRβ together with WT MAIT TCRα (WT) or mutant TCRα (Y95F) were stimulated with vehicle control, 5-OP-RU or CA7S and analyzed after 20 hours. (C and D) Data are presented as the means±SD of triplicate assays. (A, B, and E) Data are presented as individual values of duplicate assays. All data are representative of at least two independent experiments. NB, No binding.

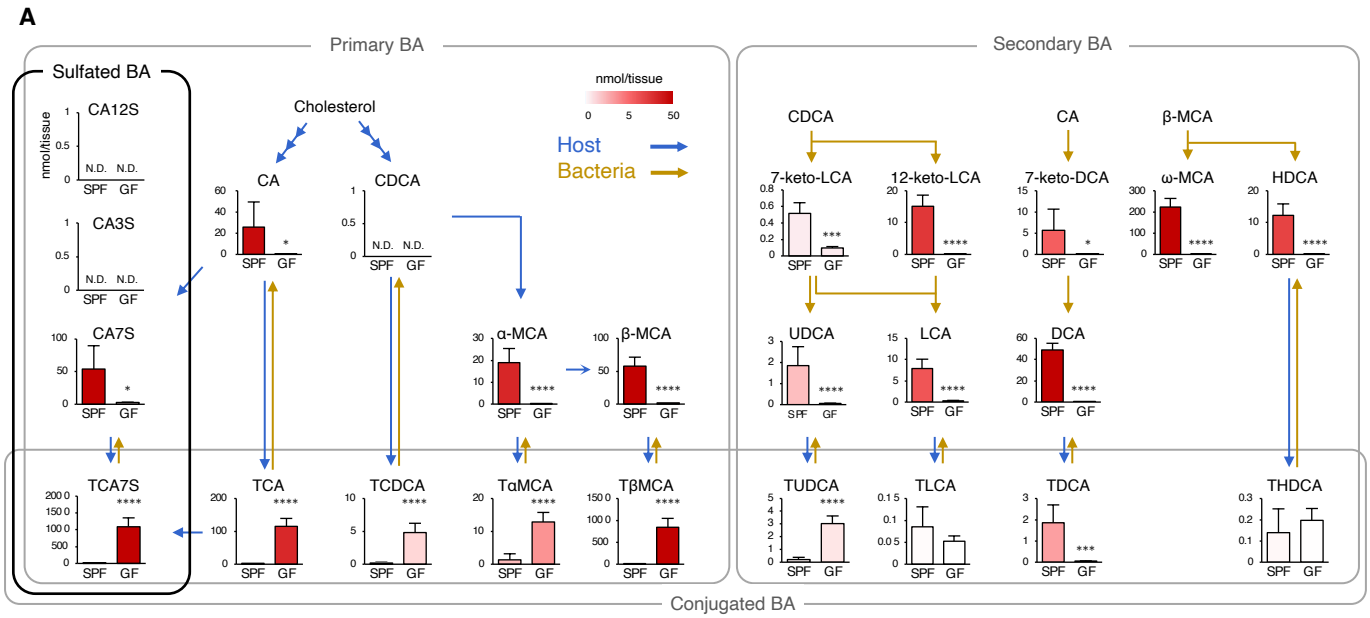
Fig. 5. Generation of *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (A) Gene-targeting strategy for *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. P1, P2, and P3 indicate primers used for genomic PCR. (**B**) Genotyping PCR of *Sult2a*$^{+/+}$, *Sult2a*$^{+/\Delta 1\text{-}8}$, and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (**C**) Quantification of bile acid metabolites in feces from WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. Blue and yellow arrows indicate host- and bacteria-mediated enzymatic responses, respectively. (**D**) UMAP projection based on sc-RNA-seq of density-fractionated liver cell suspensions from WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. Erythrocytes expressing *Hbb-bs* above 1.8 were excluded in the process of analysis. (**E**) Expression of *Sult2a1-8* (left) and *Alb* (right) in the hepatocyte clusters derived from WT (top) and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ (bottom) mice. (**F**) UMAP projection of WT (top) and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ (bottom) mice. (C) Data are presented as the means ± SD from experiment with six or more mice per group. (D to F) Data are from experiments with three mice per group. (C) A direct product of SULT2A, CA7S, was analyzed by two-tailed, unpaired Student's *t* tests (*$P<0.05$). For all other bile acid metabolites, *P*-values were adjusted with Benjamini–Hochberg correction (FDR<0.05).

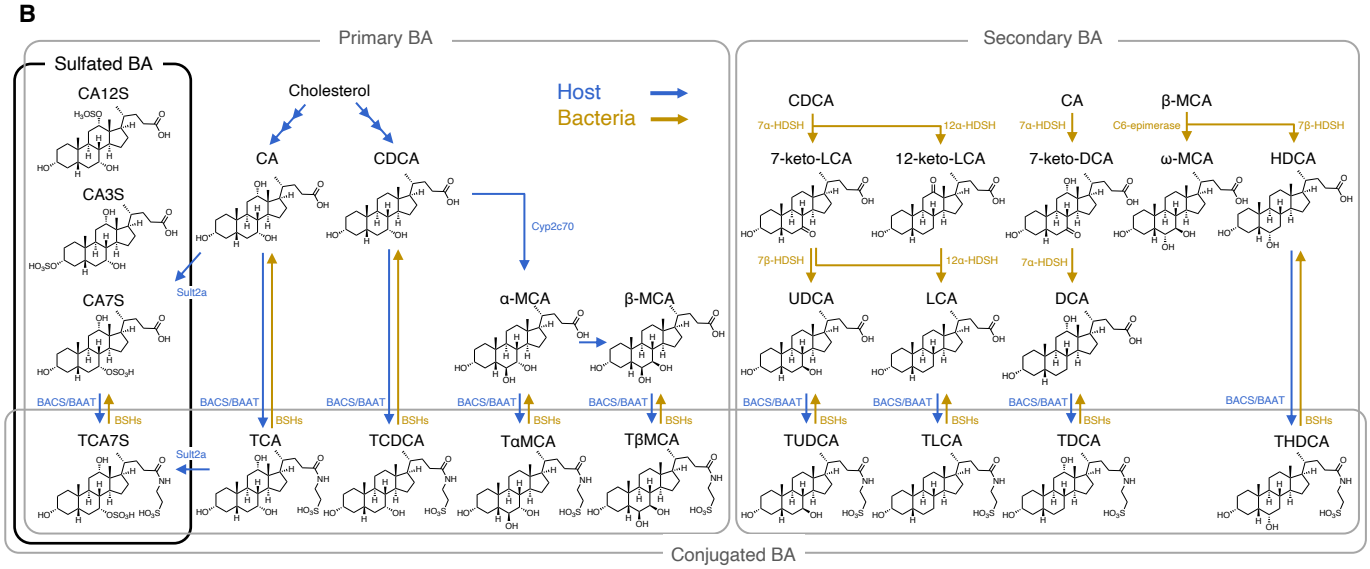
Fig. 6. Impaired MAIT cell development in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (**A** to **F**) Flow cytometry analysis of MAIT cells among MR1−5-OP-RU tetramer-enriched thymocytes from 3-4-week-old WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (A) Representative flow cytometry dot plots of MR1−5-OP-RU $tet^+TCR\beta^+$ MAIT cells in thymocytes. (B) Absolute number of thymic MAIT cells in WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (C) Representative CD44 and CD24 expression in MAIT cells from WT (left) and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ (right) mice. (D) Frequency of stage 1 ($CD44^-CD24^+$), stage 2 ($CD44^-CD24^-$), and stage 3 ($CD44^+CD24^-$) MAIT cells in WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (E) Representative CD319 and CD138 expression in $CD44^+CD24^-$ MAIT cells in WT (left) and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ (right) mice. (F) Frequency of MAIT17 cells ($CD138^+$) and MAIT1 cells ($CD319^+$) in WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (**G** and **H**) Heat map of mRNA expression of MAIT1 and 17 signature genes in canonical MAIT cells (*Trav1-Traj9/12/33*) in the (G) thymi and (H) livers of WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. (**I**) Heat map of mRNA expression in canonical MAIT cells (*Trav1-Traj9/12/33*), canonical iNKT cells (*Trav11-Traj18*), and conventional T cells in the liver of WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. The genes significantly downregulated in *Trav1*-expressing cells from *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice ($P<0.05$) compared to those from WT mice are shown. The gene expression levels are shown as average values. Data are from experiments with three

or more mice per group. *P<0.05, ** P<0.01, by two-tailed, unpaired Student's t tests or one-way ANOVA followed by Tukey's multiple comparison test.

Fig. 7. Responses of human MAIT cell to CA7S. (**A** and **B**) Human PBMCs were labeled with Cell Trace Violet (CTV) and stimulated with vehicle control (Unstim), 5-OP-RU (10 μM) or CA7S (1000 μM) in the absence of cytokines on day 6. Representative flow cytometry dot plots of proliferating $CD3^+$MR1−5-OP-RU tet^+CTV^{lo} MAIT cells (A). Proportion of MR1−5-OP-RU tet^+CTV^{lo} MAIT cells in $CD3^+$-gated cells (B). (**C**) Survival ratio of MAIT cells stimulated with 1000 μM CA7S (CA7S) or left unstimulated (Unstim). (**D** and **E**) Sc-RNA-seq of $CD3^+CD161^+$MR1−5-OP-RU tet^+ MAIT cells stimulated with vehicle control (Unstim), 5-OP-RU (10 μM) or CA7S (1000 μM) for 24 hours. UMAP projection based on mRNA expression was performed on all isolated cells (top left). UMAP of *Trav1-2*$^+$ cells (top right) were separately shown for different stimuli (Unstim, 5-OP-RU, and CA7S) (middle). Proportion of the clusters in *Trav1-2*$^+$ cells of each group are shown as bar graphs (bottom) (D). Volcano plot of mRNA expression comparing the characteristic clusters of 5-OP-RU stimulation (cluster 1) and CA7S stimulation (cluster 0) (E). (**F**) Surface expression of CD69 and CXCR4 after CA7S stimulation. PBMCs were stimulated with indicated concentrations of 5-OP-RU (red) or CA7S (blue) and surface expression of CD69 and CXCR4 within $CD3^+CD161^+$MR1 − 5-OP-RU tet^+ MAIT cells was determined at day 9. (**G**) Heatmap of mRNA expression in the most frequent MAIT clonotypes (*TRAV1-2*-CAVRDSNYQLIW-*TRAJ33 – TRBV28-TRBJ2-5*). The genes with differential expression between cluster 1 and 0 in (E) are shown. (B and C) Data are presented as the means ± SD of triplicate assays and (A, B, C and E) are representative of more than three independent experiments. * P<0.05, **** P<0.001, by two-tailed, unpaired Student's t tests with Benjamini–Hochberg correction or one-way ANOVA followed by Tukey's multiple comparison test.

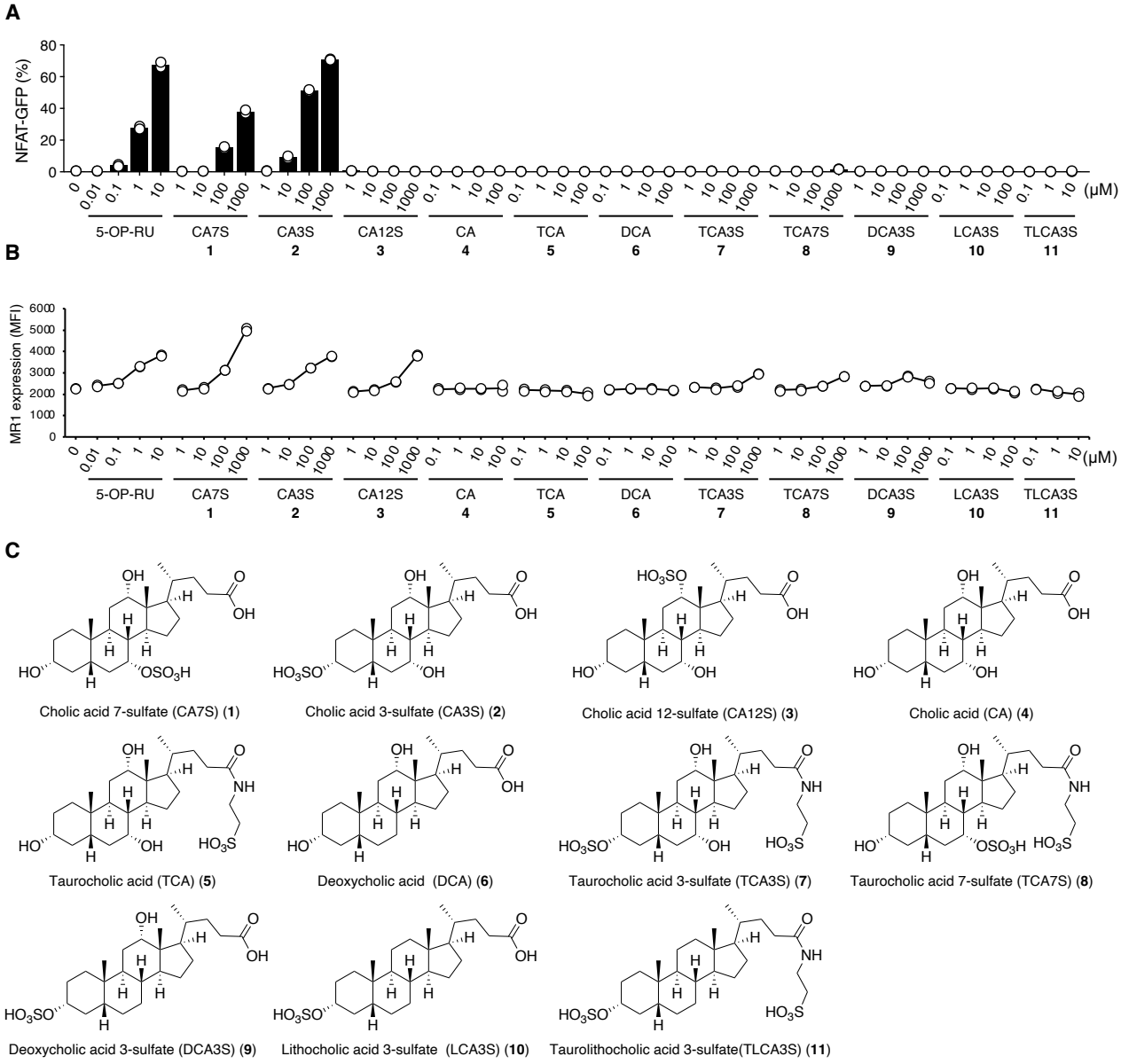

A
Intestine
NFAT-GFP (%)
Fraction #
B
5-OP-RU
NFAT-GFP (%)
C
#84-45
487.2368
488.2388
489.2390
248.9584
Intensity (× 10^6)
m/z
D
#84-45
HOD
MeCN
H-7
H-12
PPM
E
Cholic acid 7-sulfate (CA7S) (1)
F
#84-45
487.2393
96.9605
Relative abundance
m/z
G
#84-45
25.0
Relative abundance
Retention time (min)
H
5-OP-RU
RL-7-Me
CA7S
NFAT-GFP (%)
Ligand conc. (μM)
I
5-OP-RU
RL-7-Me
CA7S
MR1 expression (ΔMFI)
Ligand conc. (μM)
J
None
5-OP-RU
CA7S
MR1 expression (MFI)
(h)
Anti-MR1 Ab
Isotype Ctrl

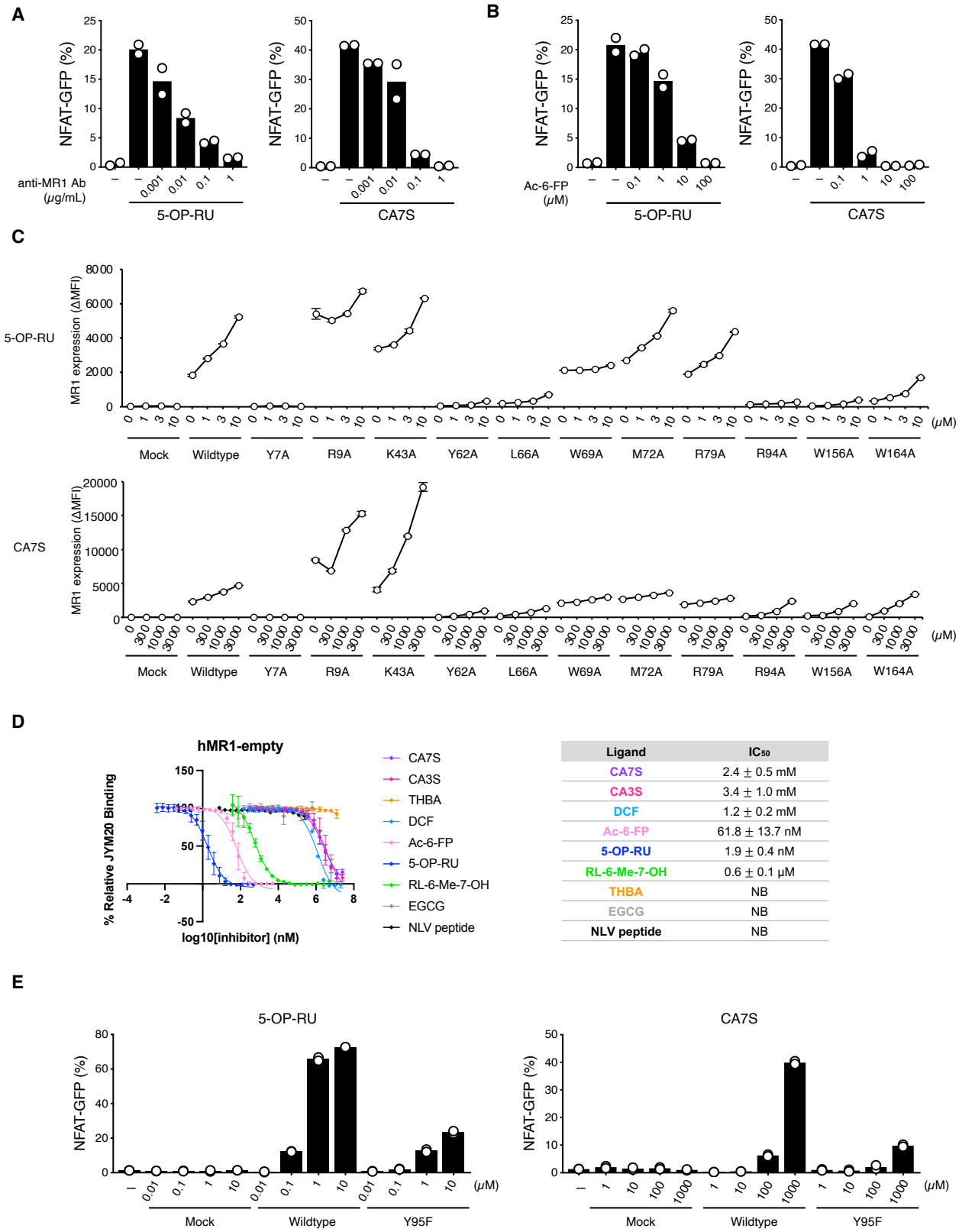




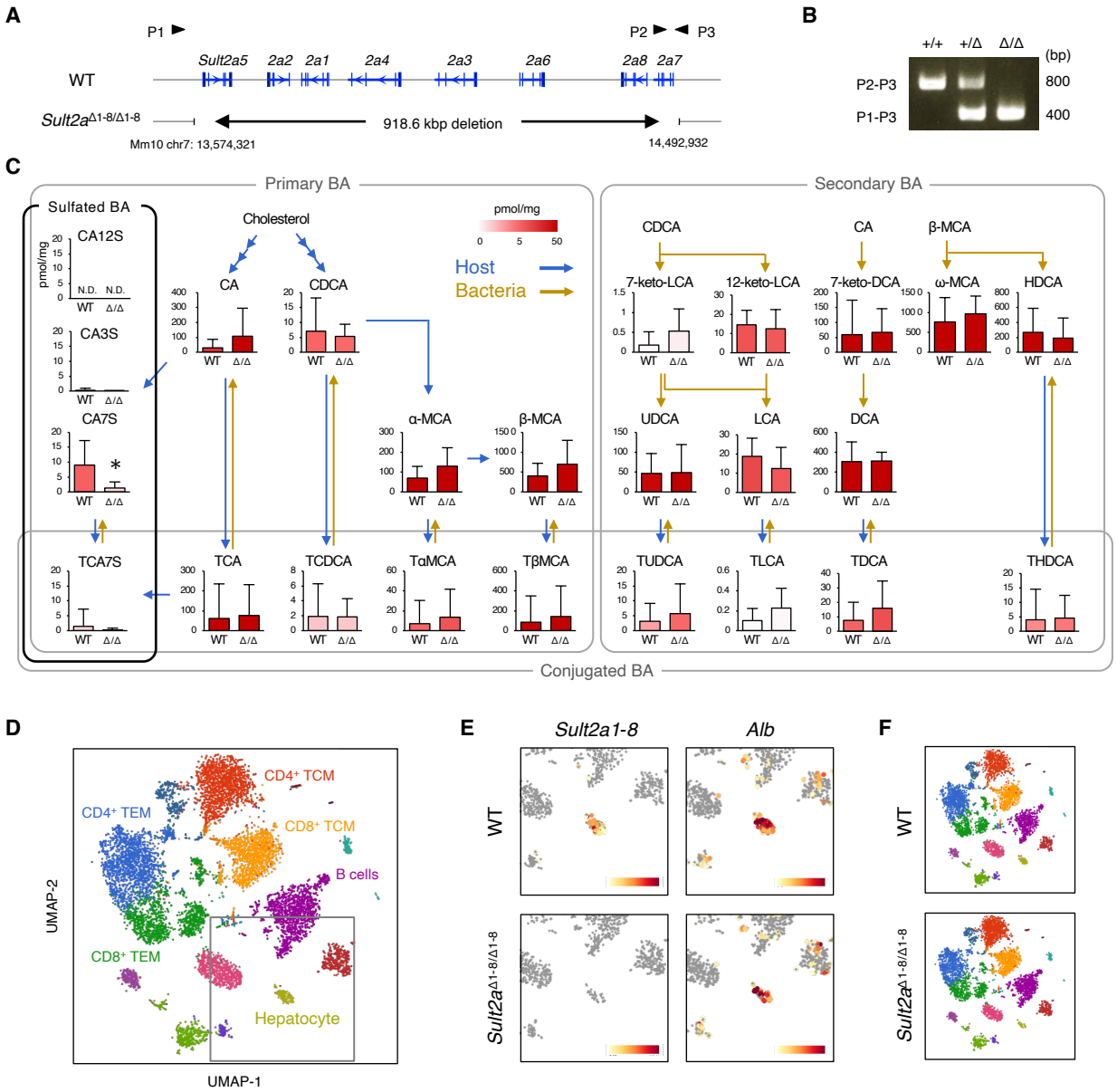
C

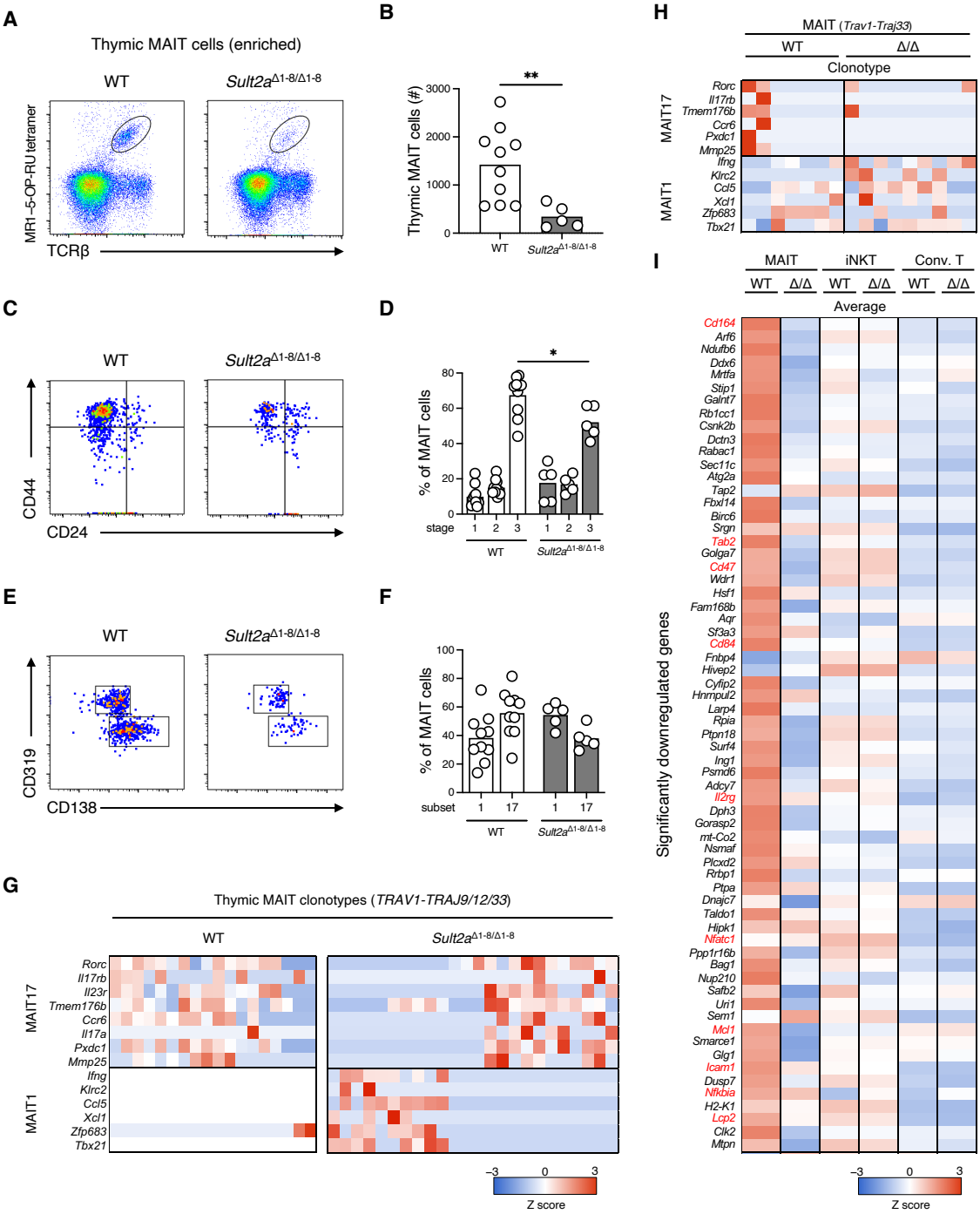
	CA7S	
	SPF	GF
Tissue	Average ± SD (pmol/mg)	
Thymus	0.14 ± 0.02	N.D.
Liver	1.1 ± 0.6	0.022 ± 0.007
Gallbladder	71 ± 27	0.71 ± 0.28
Duodenum	22 ± 8	0.39 ± 0.17
Jejunum	31 ± 17	0.45 ± 0.08
Ileum	36 ± 22	0.51 ± 0.13
Cecum	94 ± 59	0.80 ± 0.15
Colon	33 ± 28	0.39 ± 0.15

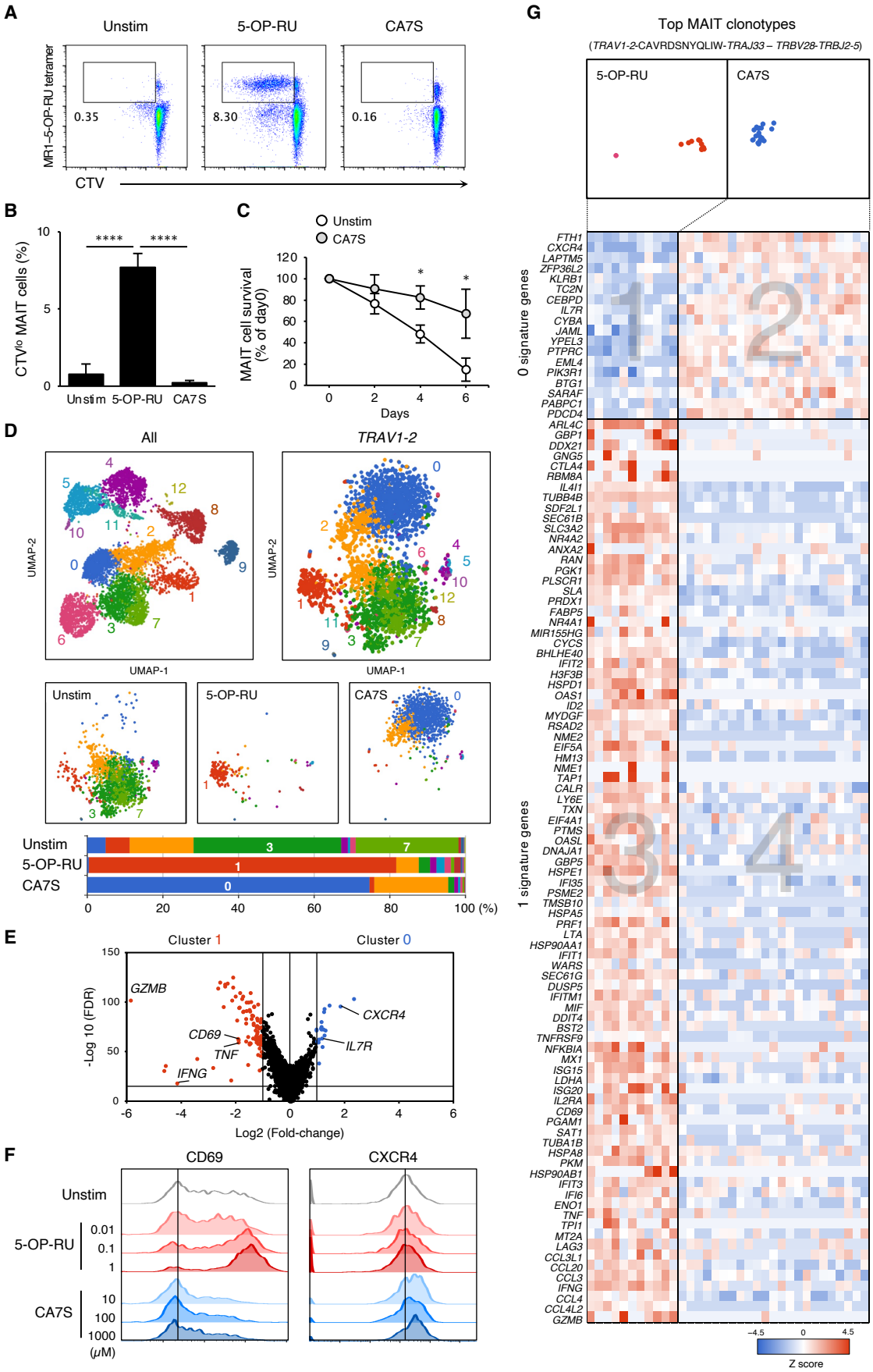

A
NFAT-GFP (%)
80
60
40
20
0
0 0.01 0.1 1 10 1 10 100 1000 1 10 100 1000 1 10 100 1000 0.1 1 10 100 0.1 1 10 100 0.1 1 10 100 1 10 100 1000 1 10 100 1000 1 10 100 1000 0.1 1 10 100 0.1 1 10 100 (μM)
5-OP-RU
CA7S 1
CA3S 2
CA12S 3
CA 4
TCA 5
DCA 6
TCA3S 7
TCA7S 8
DCA3S 9
LCA3S 10
TLCA3S 11
B
MR1 expression (MFI)
6000
5000
4000
3000
2000
1000
0
0 0.01 0.1 1 10 1 10 100 1000 1 10 100 1000 1 10 100 1000 0.1 1 10 100 0.1 1 10 100 0.1 1 10 100 1 10 100 1000 1 10 100 1000 1 10 100 1000 0.1 1 10 100 0.1 1 10 100 (μM)
5-OP-RU
CA7S 1
CA3S 2
CA12S 3
CA 4
TCA 5
DCA 6
TCA3S 7
TCA7S 8
DCA3S 9
LCA3S 10
TLCA3S 11
C
Cholic acid 7-sulfate (CA7S) (1)
Cholic acid 3-sulfate (CA3S) (2)
Cholic acid 12-sulfate (CA12S) (3)
Cholic acid (CA) (4)
Taurocholic acid (TCA) (5)
Deoxycholic acid (DCA) (6)
Taurocholic acid 3-sulfate (TCA3S) (7)
Taurocholic acid 7-sulfate (TCA7S) (8)
Deoxycholic acid 3-sulfate (DCA3S) (9)
Lithocholic acid 3-sulfate (LCA3S) (10)
Taurolithocholic acid 3-sulfate(TLCA3S) (11)



Ligand	IC_{50}
CA7S	2.4 ± 0.5 mM
CA3S	3.4 ± 1.0 mM
DCF	1.2 ± 0.2 mM
Ac-6-FP	61.8 ± 13.7 nM
5-OP-RU	1.9 ± 0.4 nM
RL-6-Me-7-OH	0.6 ± 0.1 μM
THBA	NB
EGCG	NB
NLV peptide	NB


A
P1
P2
P3
WT
Sult2a5
2a2
2a1
2a4
2a3
2a6
2a8
2a7
Sult2a Δ1-8/Δ1-8
918.6 kbp deletion
Mm10 chr7: 13,574,321
14,492,932
B
+/+
+/Δ
Δ/Δ
(bp)
P2-P3
P1-P3
800
400
C
Primary BA
Secondary BA
Sulfated BA
Cholesterol
pmol/mg
Host
Bacteria
CA12S
CA3S
CA7S
CA
CDCA
α-MCA
β-MCA
TCA7S
TCA
TCDCA
TαMCA
TβMCA
CDCA
CA
β-MCA
7-keto-LCA
12-keto-LCA
7-keto-DCA
ω-MCA
HDCA
UDCA
LCA
DCA
TUDCA
TLCA
TDCA
THDCA
WT
Δ/Δ
N.D.
Conjugated BA
D
CD4+ TCM
CD4+ TEM
CD8+ TCM
B cells
CD8+ TEM
Hepatocyte
UMAP-2
UMAP-1
E
Sult2a1-8
Alb
WT
Sult2a Δ1-8/Δ1-8
F
WT
Sult2a Δ1-8/Δ1-8


A
Thymic MAIT cells (enriched)
WT
Sult2a Δ1-8/Δ1-8
MR1–5-OP-RU tetramer
TCRβ
B
Thymic MAIT cells (#)
3000
2000
1000
0
**
WT
Sult2a Δ1-8/Δ1-8
C
WT
Sult2a Δ1-8/Δ1-8
CD44
CD24
D
% of MAIT cells
80
60
40
20
0
*
stage
1 2 3 1 2 3
WT
Sult2a Δ1-8/Δ1-8
E
WT
Sult2a Δ1-8/Δ1-8
CD319
CD138
F
% of MAIT cells
100
80
60
40
20
0
subset
1 17 1 17
WT
Sult2a Δ1-8/Δ1-8
G
Thymic MAIT clonotypes (TRAV1-TRAJ9/12/33)
WT
Sult2a Δ1-8/Δ1-8
MAIT17
Rorc
Il17rb
Il23r
Tmem176b
Ccr6
Il17a
Pxdc1
Mmp25
MAIT1
Ifng
Klrc2
Ccl5
Xcl1
Zfp683
Tbx21
–3 0 3
Z score
H
MAIT (Trav1-Traj33)
WT
Δ/Δ
Clonotype
MAIT17
Rorc
Il17rb
Tmem176b
Ccr6
Pxdc1
Mmp25
MAIT1
Ifng
Klrc2
Ccl5
Xcl1
Zfp683
Tbx21
I
MAIT
iNKT
Conv. T
WT
Δ/Δ
WT
Δ/Δ
WT
Δ/Δ
Average
Significantly downregulated genes
Cd164
Arf6
Ndufb6
Ddx6
Mrtfa
Stip1
Galnt7
Rb1cc1
Csnk2b
Dctn3
Rabac1
Sec11c
Atg2a
Tap2
Fbxl14
Birc6
Srgn
Tab2
Golga7
Cd47
Wdr1
Hsf1
Fam168b
Aqr
Sf3a3
Cd84
Fnbp4
Hivep2
Cyfip2
Hnrnpul2
Larp4
Rpia
Ptpn18
Surf4
Ing1
Psmd6
Adcy7
Il2rg
Dph3
Gorasp2
mt-Co2
Nsmaf
Plcxd2
Rrbp1
Ptpa
Dnajc7
Taldo1
Hipk1
Nfatc1
Ppp1r16b
Bag1
Nup210
Safb2
Uri1
Sem1
Mcl1
Smarce1
Glg1
Icam1
Dusp7
Nfkbia
H2-K1
Lcp2
Clk2
Mtpn
–3 0 3
Z score


A
Unstim
5-OP-RU
CA7S
MR1–5-OP-RU tetramer
0.35
8.30
0.16
CTV
B
CTV^lo MAIT cells (%)
****
****
Unstim
5-OP-RU
CA7S
C
Unstim
CA7S
MAIT cell survival (% of day0)
Days
D
All
TRAV1-2
UMAP-2
UMAP-1
Unstim
5-OP-RU
CA7S
100 (%)
E
Cluster 1
Cluster 0
-Log 10 (FDR)
Log2 (Fold-change)
GZMB
CD69
TNF
IFNG
CXCR4
IL7R
F
CD69
CXCR4
Unstim
5-OP-RU
CA7S
(μM)
G
Top MAIT clonotypes
(TRAV1-2-CAVRDSNYQLIW-TRAJ33 – TRBV28-TRBJ2-5)
5-OP-RU
CA7S
0 signature genes
1 signature genes
Z score

Science Immunology

Supplementary Materials for

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

Emi Ito, Shinsuke Inuki, Yoshihiro Izumi, Masatomo Takahashi, Yuki Dambayashi, Lisa Ciacchi, Wael Awad, Ami Takeyama, Kensuke Shibata, Shotaro Mori, Jeffrey Y. W. Mak, David P. Fairlie, Takeshi Bamba, Eri Ishikawa, Masamichi Nagae, Jamie Rossjohn, and Sho Yamasaki

Correspondence to: yamasaki@biken.osaka-u.ac.jp

This PDF file includes:

Supplementary Materials and Methods
Figs. S1 to S7
Tables S1 to S4
Caption for Data S1

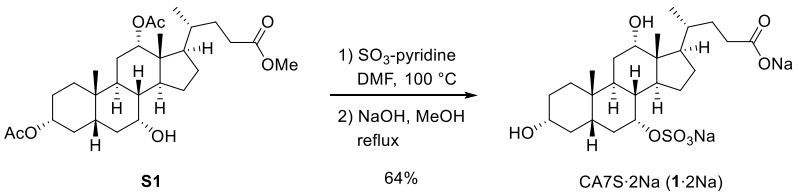
Other Supplementary Material for this manuscript includes the following:

Data S1

SUPPLEMENTARY MATERIALS AND METHODS

General methods for the synthesis of compounds

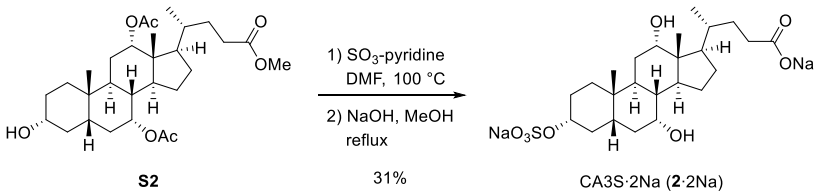
^{1}H NMR spectra were recorded using a JEOL ECA-500 or JEOL ECZ600R spectrometer. Chemical shifts are reported in δ (ppm) relative to Me_4Si (in $CDCl_3$ and CD_3OD) and MeCN signal (in D_2O) as internal standards. ^{13}C NMR spectra were recorded using a JEOL ECA-500 or JEOL ECZ600R spectrometer and 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.



Disodium cholate 7-sulfate (CA7S·2Na)

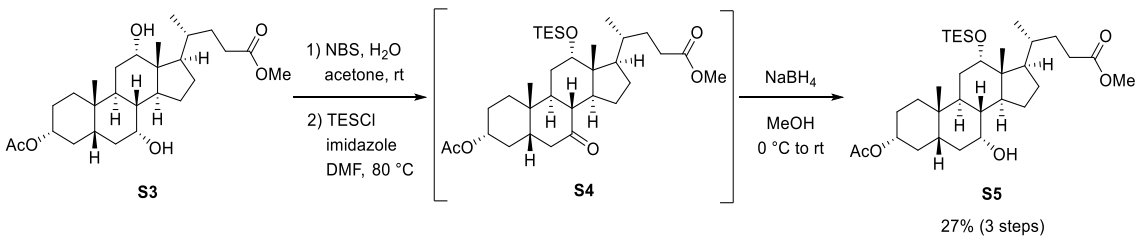
CA7S·2Na was synthesized by modifying of a previously published method (*89, 90*). Sulfur trioxide-pyridine complex (565 mg, 3.54 mmol) was added to a stirred solution of Compound S1 (179 mg, 0.352 mmol) in dry DMF (2 ml) under an argon atmosphere. After 24 hours of

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.



Disodium cholate 3-sulfate (CA3S·2Na)

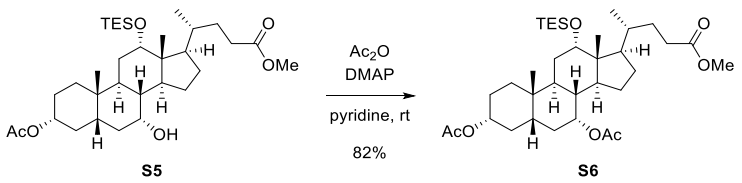
Using a procedure identical with that described for synthesis of Compound 1·2Na from Compound S1, cholic acid derivative Compound S2 (100 mg, 0.197 mmol) was converted into CA3S·2Na (Compound 2·2Na) as a white solid (32.5 mg, 31%): mp 193−195°C; $[\alpha]^{25}_{D}$ +33.8 (*c* 0.405, MeOH); IR (neat cm^{-1}): 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

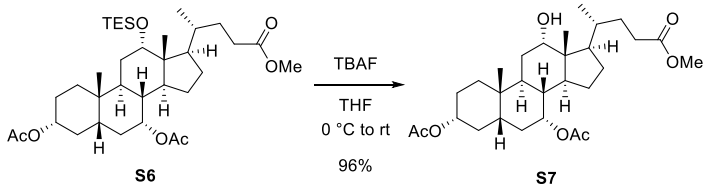
NBS (1.40 g, 7.85 mmol) was added to a stirred solution of Compound S3 (2.92 g, 6.28 mmol) in water (39 ml) and acetone (56 ml) at room temperature. After 16 hours of constant stirring at this temperature, the mixture was concentrated in vacuo, and the residue was dissolved in CH_2Cl_2. The organic layer 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

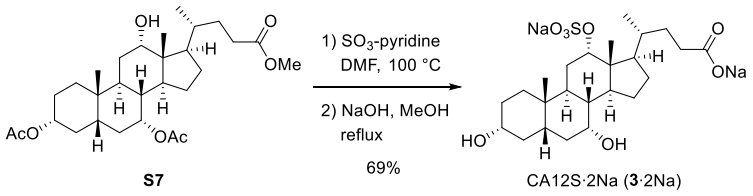
DMAP (21.4 mg, 0.175 mmol) and Ac_2O (465 μl, 4.92 mmol) were added to a solution of Compound S5 (949 mg, 1.64 mmol) in pyridine (5.3 ml) at 0°C. After being stirred for 23 hours at room temperature, the mixture was diluted with CH_2Cl_2, washed with saturated aqueous $CuSO_4$ and brine, and dried over 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

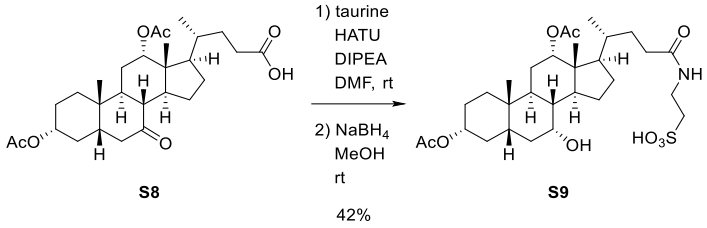
0:100) to yield Compound S7 as a white solid (653 mg, 96%): mp 181−183°C; $[\alpha]^{26}_{D}$ +23.5 (*c* 0.668, $CHCl_3$); IR (neat cm^{-1}): 3498 (OH), 1725 (C=O); ^{1}H NMR (500 MHz, $CDCl_3$) δ 0.69 (s, 3H), 0.92 (s, 3H), 0.98 (d, *J* = 6.3 Hz, 3H), 1.04−1.13 (m, 2H), 1.24−1.54 (m, 8H), 1.55−1.65 (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{^{1}H} 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); ^{1}H 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{^{1}H} 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,

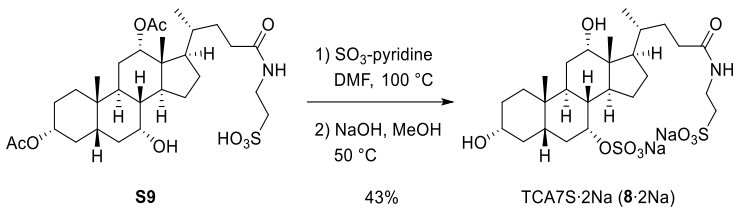
47.1, 69.0, 72.4, 84.1, 185.3; HRMS (ESI-TOF) *m/z*: $[M + H - 2Na]^-$ calculated for $C_{24}H_{39}O_8S$, 487.2371; found, 487.2378.



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).



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.

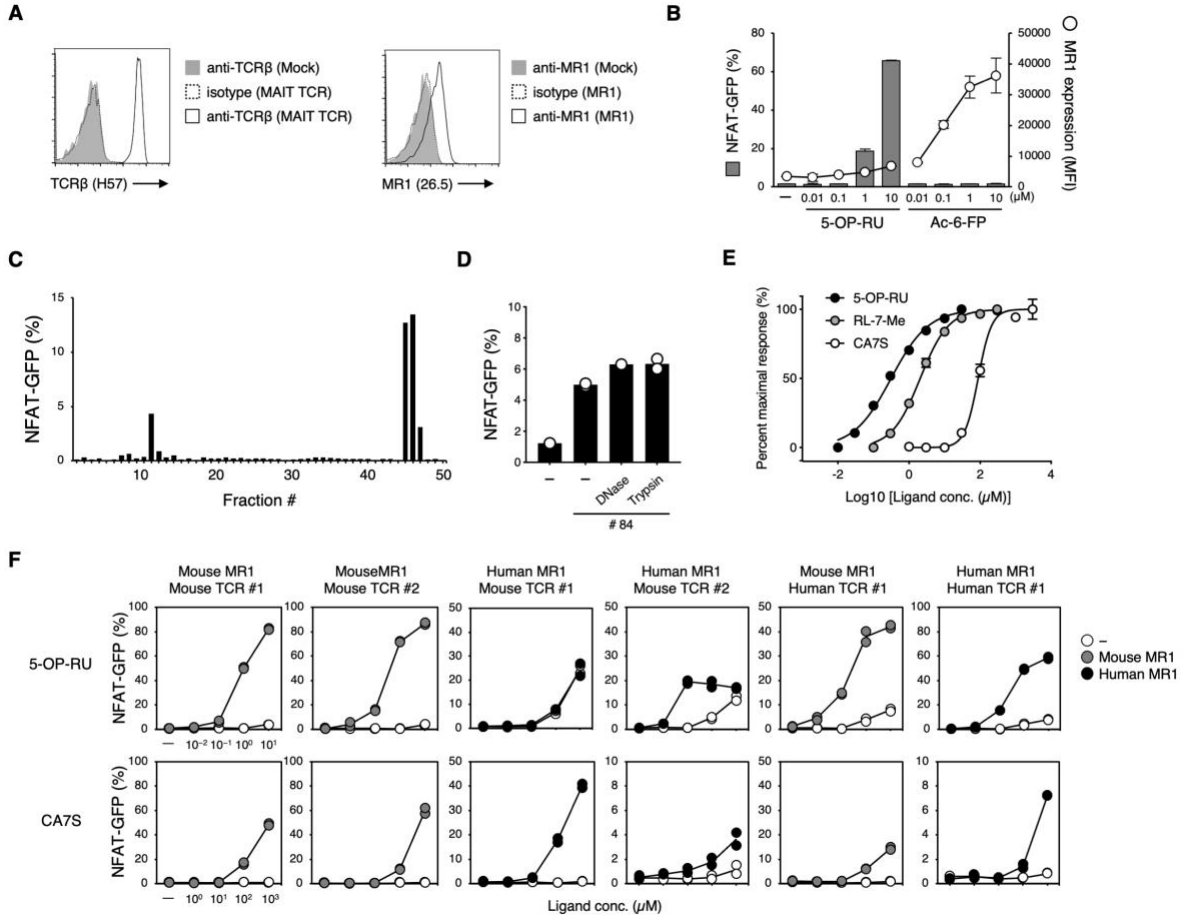


Fig. S1. Purification and characterization of CA7S. (**A**) A MAIT TCR and MR1 were co-expressed in TCR-deficient T cell hybridoma (*81*). Surface expression levels of MAIT TCR and MR1 on reporter cells were assessed by surface staining with anti-mouse TCRβ (H57-597) (left) or anti-human/mouse/rat MR1 (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.

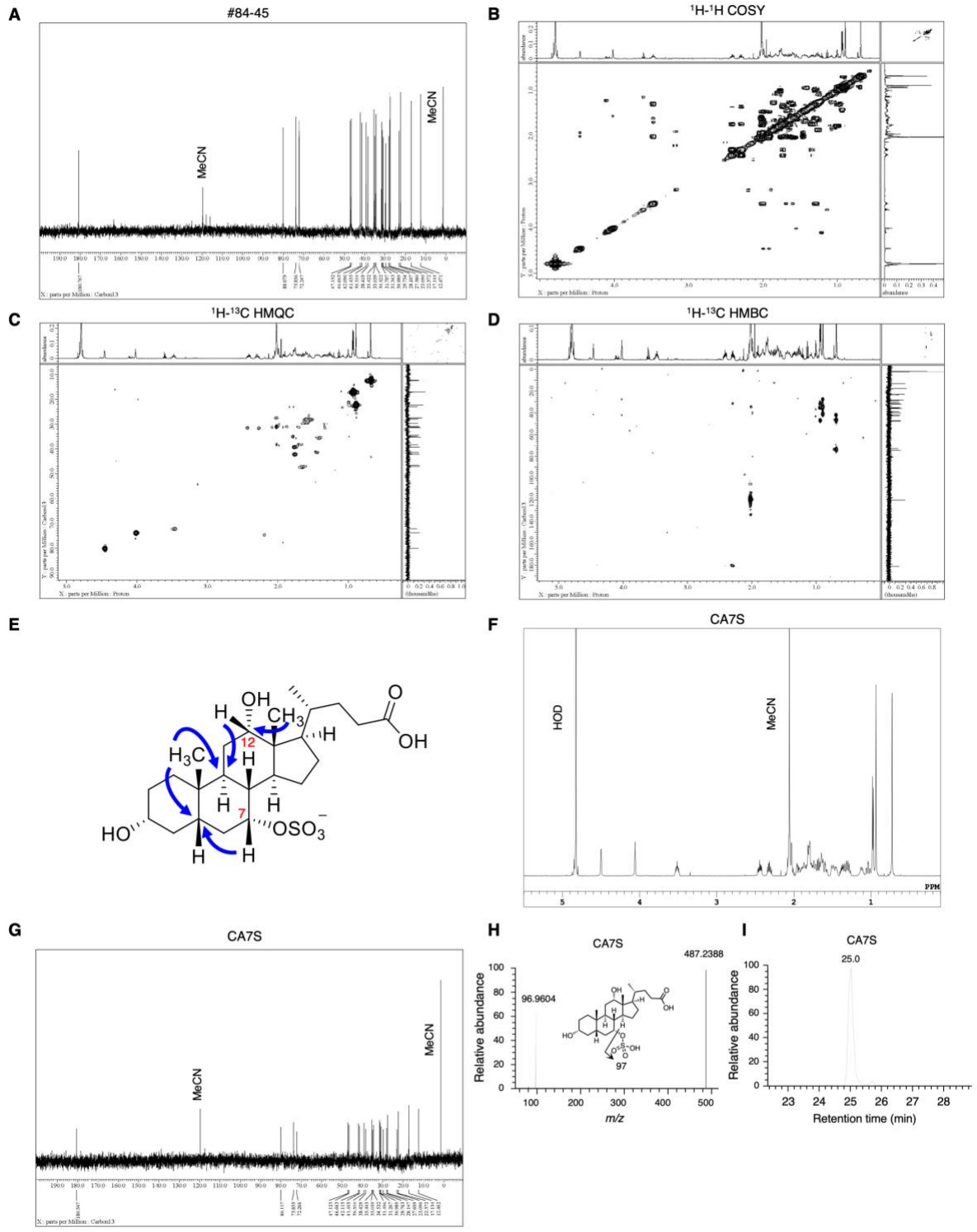


Fig. S2. Structural determination of CA7S. (**A**) ^{13}C NMR spectra (150 MHz, D_2O) of fraction #84-45. (**B** to **D**) 2D NMR analyses, including ^{1}H-^{1}H correlation spectroscopy (COSY) (B), ^{1}H-^{13}C heteronuclear multiple quantum correlation spectroscopy (HMQC) (C) and ^{1}H-^{13}C heteronuclear 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.

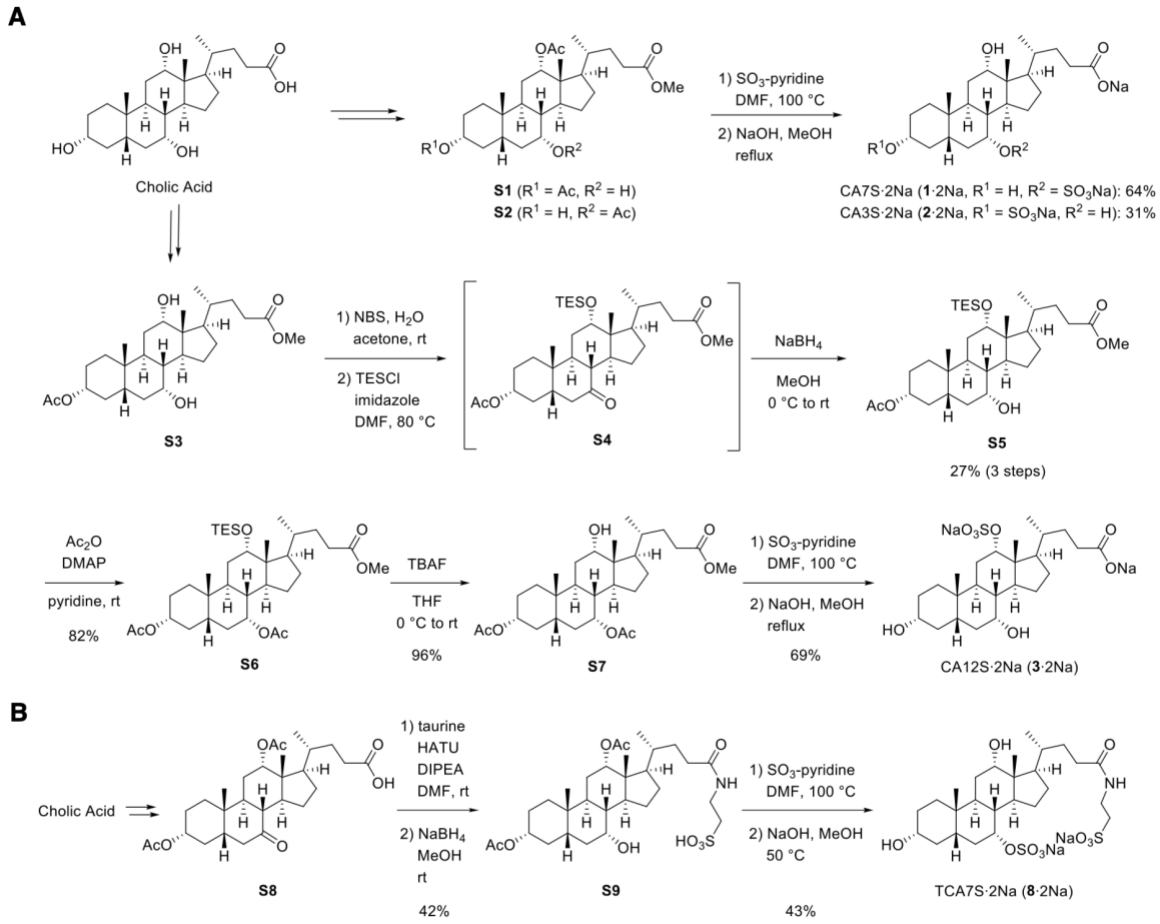


Fig. S3. Chemical synthesis of bile acid metabolites. (**A**) Synthetic schemes of CA7S (Compound 1), CA3S (Compound 2), and CA12S (Compound 3). (**B**) Synthetic scheme of TCA7S (Compound 8).

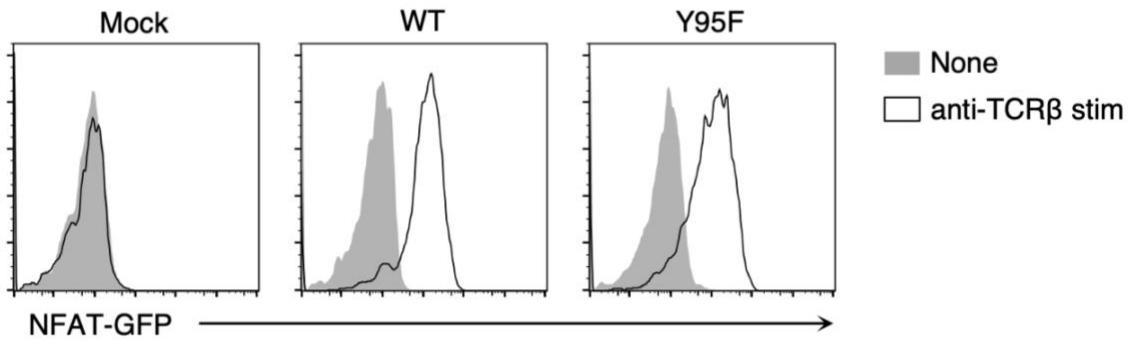


Fig. S4. Stimulation of reporter cells expressing mutant TCRα. Reporter cells were transfected with vector alone (Mock), MAIT TCRβ together with WT MAIT TCRα (WT), or mutant TCRα (Y95F). The reporter cells were stimulated with 10 μg/ml of plate-coated anti-mouse TCRβ (H57-597). GFP expression was analyzed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

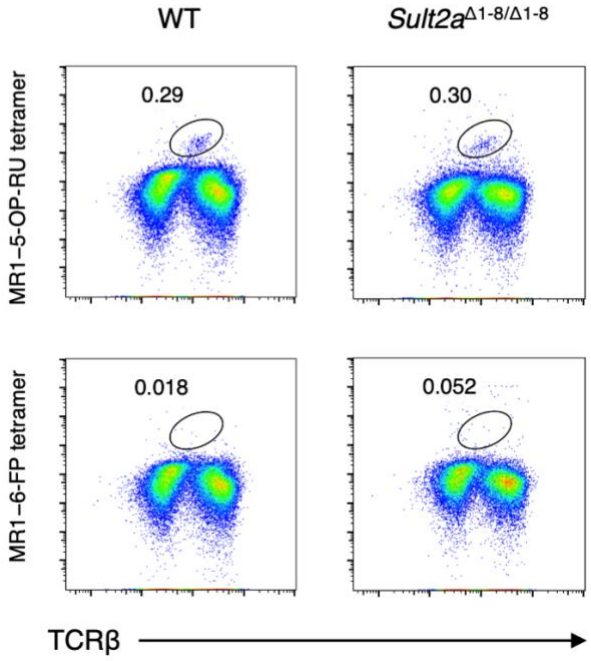


Fig. S5. Liver MAIT cells in *Sult2a*-deficient mice. Representative flow cytometry dot plots of liver CD45^{+}CD19^{-}TCRβ^{+}MR1−5-OP-RU tetramer^{+} MAIT cells from WT or *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice. MR1−6-FP tetramer staining was used as a negative control for the MR1−5-OP-RU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

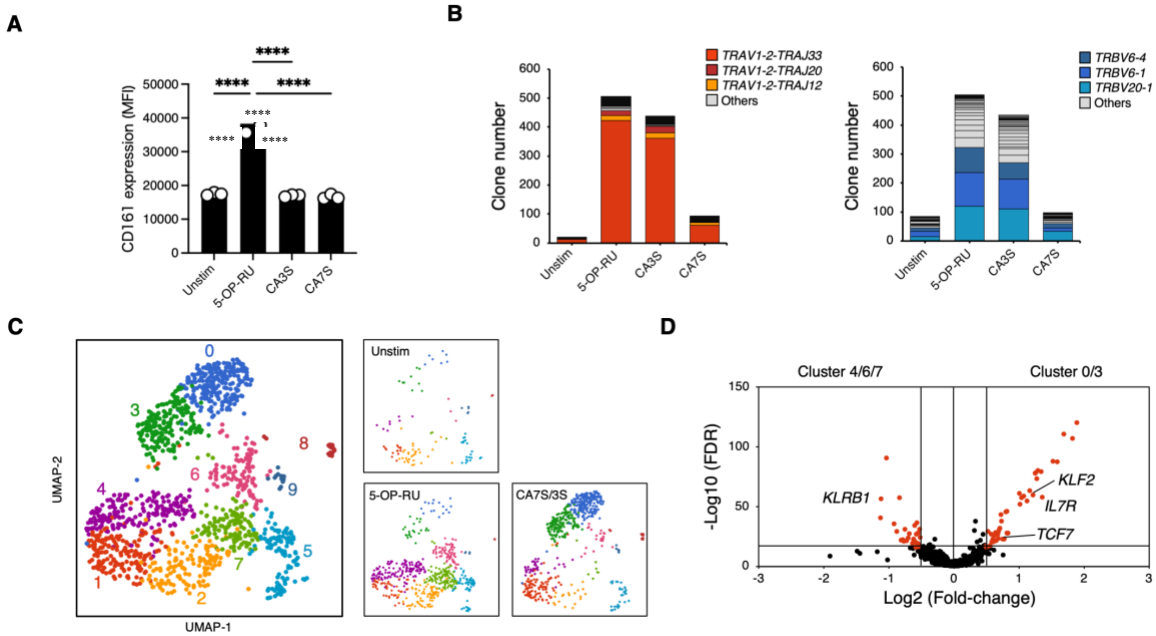


Fig. S6. Responses of human MAIT cells upon stimulation with CAS or 5-OP-RU. (A) CD161 expression (MFI) on $CD3^{+}CD161^{+}MR1-5\text{-}OP\text{-}RU\text{-}tetramer^{+}$ MAIT cells in PBMCs 6 days after stimulation with or without 5-OP-RU, CA3S, or CA7S. (**B** to **D**) Human PBMCs were stimulated with vehicle control (Unstim), 5-OP-RU (10 μM), CA3S (500 μM), or CA7S (500 μM) for 17 days. (B) TCRαβ usage of $CD3^{+}CD161^{+}MR1-5\text{-}OP\text{-}RU$ 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.

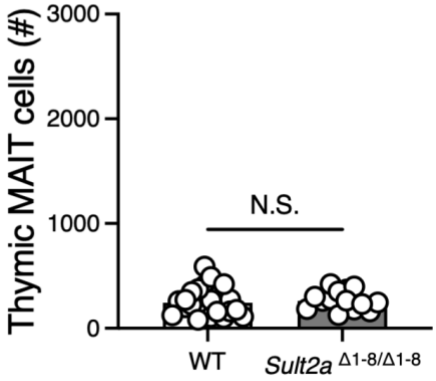


Fig. S7. Thymic MAIT cells in *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice under germ-free conditions. Germ-free *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ mice were rederived by in vitro fertilization. Thymic MAIT cells from 8-week-old WT and *Sult2a*$^{\Delta 1\text{-}8/\Delta 1\text{-}8}$ 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 shown in bold, respectively.

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.