SCISPACE
formerly Typeset

Open access • Journal Article • DOI:10.1038/NATURE11605

MR1 presents microbial vitamin B metabolites to MAIT cells — Source link

Lars Kjer-Nielsen, Onisha Patel, Alexandra J. Corbett, Jérôme Le Nours ...+21 more authors

Institutions: University of Melbourne, Monash University, Clayton campus, Australian Research Council, University of Queensland ...+1 more institutions

Published on: 29 Nov 2012 - Nature (Nature Publishing Group)

Topics: Antigen presentation, CD1, MHC class I, Antigen and Major histocompatibility complex

Related papers:

- Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1
- Human MAIT cells are xenobiotic-resistant, tissue-targeted, CD161hi IL-17–secreting T cells
- Antimicrobial activity of mucosal-associated invariant T cells
- T-cell activation by transitory neo-antigens derived from distinct microbial pathways
- Human Mucosal Associated Invariant T Cells Detect Bacterially Infected Cells

Share this paper:

View more about this paper here: https://typeset.io/papers/mr1-presents-microbial-vitamin-b-metabolites-to-mait-cells-26t8ouvrx0

T cell activation by transitory neo-antigens derived from distinct microbial pathways

Alexandra J. Corbett[1*], Sidonia B.G. Eckle[1*], Richard W. Birkinshaw[2*], Ligong Liu[3*], Onisha Patel[2], Jennifer Mahony[4], Zhenjun Chen[1], Rangsima Reantragoon[1], Bronwyn Meehan[1], Hanwei Cao[1], Nicholas A. Williamson[5], Richard A. Strugnell[1], Douwe Van Sinderen[4,6], Jeffrey Y. W. Mak[3], David P. Fairlie[3,7#], Lars Kjer-Nielsen[1#], Jamie Rossjohn[2,7,8#] & James McCluskey[1#]

[1]Department of Microbiology & Immunology, Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, Victoria 3010, Australia

[2]Department of Biochemistry and Molecular Biology, School of Biomedical Sciences, Monash University, Clayton, Victoria 3800, Australia

[3]Division of Chemistry & Structural Biology, Institute for Molecular Bioscience, The University of Queensland, Brisbane, Qld 4072, Australia

[4]Department of Microbiology, University College Cork, Cork, Ireland.

[5]The Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria, 3010, Australia

[6]Alimentary Pharmabiotic Centre, University College Cork, Cork, Ireland.

[7] Australian Research Council Centre of Excellence in Advanced Molecular Imaging

[8]Institute of Infection and Immunity, Cardiff University, School of Medicine, Heath Park, Cardiff CF14 4XN, UK.

[*]AJC, SBGE, RWB and LL are joint first authors

[#]DPF, LKN, JR and JMc are joint senior authors

e-mail: d.fairlie@imb.uq.edu.au, lkn@unimelb.edu.au, jamie.rossjohn@monash.edu or jamesm1@unimelb.edu.au for correspondence

Running title: MAIT cell antigen formation from distinct chemical entities

T cells discriminate between foreign and host molecules by recognizing distinct microbial molecules, predominantly peptides and lipids [1-4]. Riboflavin precursors found in many bacteria and yeast also selectively activate mucosal-associated invariant T (MAIT) cells [5,6], an abundant population of innate-like T cells in humans [7-9]. However, the genesis of these small organic molecules and their mode of presentation to MAIT cells by the Major Histocompatibility Complex related protein, MR1, are not well understood. We show here that MAIT cell activation requires key genes encoding enzymes that form 5-amino-6-D-ribitylaminouracil (5-A-RU), an early intermediate in bacterial riboflavin synthesis. While 5-A-RU does not bind MR1 or activate MAIT cells directly, it does form potent MAIT-activating antigens *via* non-enzymatic reactions with small molecules, such as glyoxal and methylglyoxal, which are derived from other metabolic pathways. The MAIT antigens formed by the reactions between 5-A-RU and glyoxal/methylglyoxal were simple adducts, 5-(2-oxoethylideneamino)-6-D-ribitylaminouracil (5-OE-RU) and 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU) respectively, which bound to MR1 as shown by crystal structures of MAIT TCR ternary complexes. Although 5-OP-RU and 5-OE-RU are unstable intermediates, they became trapped by MR1 as reversible covalent Schiff base complexes. Mass spectra supported the capture by MR1 of 5-OP-RU and 5-OE-RU from bacterial cultures that activate MAIT cells, but not from non-activating bacteria, indicating that these MAIT Ags are present in a range of microbes. Thus, MR1 is able to capture, stabilize and present chemically unstable pyrimidine intermediates, which otherwise convert to lumazines, as potent antigens to MAIT cells. These pyrimidine adducts are microbial signatures for MAIT cell immunosurveillance.

MAIT cell antigens were previously identified from *Salmonella typhimurium* (strain SL1344) supernatant [5]. Negative mode electrospray ionization–time-of-flight mass spectrometry (ESI–TOF-MS) analysis of MR1-bound ligands from *S. typhimurium* revealed a ligand with a mass to charge (m/z) ratio of 329.11, matching a potent MAIT activating ligand identified during the chemical synthesis of reduced 6-hydroxymethyl-8-D-ribityllumazine (rRL-6-CH_2OH) [5]. While this ligand was identified biochemically, its origin was puzzling, as it is not described in the riboflavin synthesis pathway. We therefore took a genetic approach to evaluate if the riboflavin pathway supplied the MAIT cell ligands.

We examined the capacity of bacterial mutants of the riboflavin pathway to activate MAIT cells (**Figure 1a**). In some bacterial species, including *Lactococcus lactis*, the genes necessary for riboflavin synthesis are grouped together in a single 4-gene operon (RibGBAH), and are regulated by transcriptional repression of a "riboswitch" *via* flavin mononucleotide and riboflavin [10]. Using *L.*

lactis, we tested the ability of bacterial culture supernatant to activate Jurkat cells transduced with a MAIT TCR (Jurkat.MAIT) **(Figure 1b)**. Supernatant from wild type *L. lactis* strain NZ9000 incubated with Ag presenting cells expressing MR1 caused CD69 upregulation in Jurkat.MAIT cells (**Figure 1b**). Addition of riboflavin during culture of NZ9000 inhibited MAIT cell activation consistent with negative regulation of the riboswitch and impaired production of the activating MAIT ligand (**Figure 1b**). Next, three mutant strains of *L. lactis* were employed: two riboflavin overproducers, CB013 and CB021, which produce riboflavin even in the presence of high riboflavin concentrations, and a RibA- strain, which contains a deletion in *ribA*, early in the riboflavin pathway [11]. The riboflavin overproducers activated Jurkat.MAIT cells when grown with or without exogenous riboflavin, whilst there was no MAIT cell activation by supernatant from the RibA- strain (**Figure 1b**). Accordingly, the riboflavin pathway is necessary and sufficient to produce natural MAIT cell antigens.

Next we generated individual mutations in the four genes of the riboflavin operon in *L. lactis*. These were produced using the constitutive riboflavin overproducer strain CB013. Culture supernatants from bacteria with mutant riboflavin pathways were tested for activation of Jurkat.MAIT cells **(Figure 1c)**. The parental CB013 supernatant activated Jurkat.MAIT cells whereas bacteria containing mutations in *RibA* or *RibG* did not activate the reporter cells under similar conditions (**Figure 1c**). Neither *RibB* nor *RibH* mutations, which affect the pathway downstream of 5-amino-6-D-ribitylaminouracil (5-A-RU), had any impact on Jurkat.MAIT activation (**Figure 1c**). Moreover, while the m/z 329.11 species was undetectable in MR1 refolded with supernatant from the RibA and RibG mutants, it was captured by MR1 from the supernatants of the RibB and RibH mutants of *L. lactis* (**Extended Data Figure 1a**). Moreover, culture supernatants from *S. typhimurium* SL1344 with mutated RibD+H did not furnish a detectable m/z 329.11 species that bound MR1, and could not activate Jurkat.MAIT cells (**Extended Data Figures 1b-c**). However, Ag was detectable in the complemented RibD+H SL1344 mutant (**Extended Data Figure 1c**). MR1-restricted ligands were not detected from the supernatant of *Enterococcus faecalis*, which neither possesses the riboflavin pathway nor activates MAIT cells (**Extended Data Figure 1d** and not shown). Analysis of MR1-bound ligands from another MAIT activating strain, *Escherichia coli* (DH5α strain) also revealed a ligand with a mass to charge (m/z) ratio of 329.11 (**Extended Data Figure 1e**). These data are consistent with MAIT activating ligands, from a number of bacterial sources, being derived *via* an unknown mechanism from 5-A-RU.

A key precursor step (**Figure 2a**) in riboflavin biosynthesis is the condensation of 5-A-RU (**1**) with 3,4-dihydroxy-2-butanone-4-phosphate (**2a**) to generate an intermediate 5-(1-methyl-2-

oxopropylideneamino)-6-D-ribitylaminouracil (5-MOP-RU, **3a**), which ring closes readily with dehydration to form RL-6,7-DiMe (**4a**)[12,13] (**Figure 2a**), a biosynthesis that is catalysed by lumazine synthase (RibH). However, RL-6,7-DiMe can also be generated in the absence of lumazine synthase[13,14], suggesting that MAIT antigens might be formed through spontaneous reactions of 5-A-RU with other small molecules *via* non-enzymatic mechanisms (**Figure 2a**). For example, butane-2,3-dione (**2b**), glyoxal (**2c**) and methylglyoxal (pyruvaldehyde, **2d**) can represent by-products arising from a number of metabolic pathways, including glycolysis [15]. Their condensations with 5-A-RU (**1**) would respectively produce pyrimidine adducts 5-MOP-RU (**3b = 3a**), 5-OE-RU (**3c**) and 5-OP-RU (**3d**) *en route* to ribityllumazines RL-6,7-DiMe (**4b = 4a**), RL (**4c**) and RL-7-Me (**4d**) respectively. We found that the initial adducts **3b-d** were formed almost immediately, but readily undergo dehydration upon ring closure to very stable, isolatable compounds **4a-d**, without the need for enzyme catalysis. Adducts **3a-d** (**Figure 2a**) were especially unstable under acidic aqueous conditions (pH < 6), but we could detect them in solution under physiological conditions. We were able to synthesise **3d** in DMSO-d_6, isolate it, unambiguously assign its solution structure by NMR spectroscopy (**Figure 2b & Extended Data Figure 2**), and examine its stability in aqueous media using LCMS. At 37 ºC and pH 6.8, adduct **3d** was clearly formed and had a half-life of around 2 h at 65 μM. It was more stable at lower temperature (e.g. $t_{1/2}$ 14-15h, 15 ºC, pH 6.8–8.0, 65–250 μM) (**Extended Data Figure 3**).

We hypothesized that butane-2,3-dione (**2b**), glyoxal (**2c**), and methylglyoxal (pyruvaldehyde, **2d**), in the presence of 5-A-RU (**1**) might react spontaneously to facilitate MR1 refolding. MR1-5-A-RU was undetectable when MR1 was refolded with 5-A-RU alone (not shown), and the presence of 5-A-RU and butane-2,3-dione failed to yield any MR1-Ag complexes (not shown). However, refolding of MR1 in the presence of 5-A-RU and either glyoxal or methylglyoxal led to a correctly folded MR1-Ag complex. To understand the basis for ligand selectivity by MR1, we determined the structures of the MAIT TCR in complex with MR1 and Ags formed from the condensation of 5-A-RU and either methylglyoxal or glyoxal (**Figures 3a-h, Extended data Table 1 & Extended data Figure 4**)[16]. Surprisingly, the chemically unstable adducts 5-(2- oxopropylideneamino)-6-D-ribitylaminouracil (5-OP-RU, compound **3d**, **Figure 2a**) and 5-(2- oxoethyllideneamino)-6-D-ribitylaminouracil (5-OE-RU, compound **3c**, **Figure 2a**) were observed bound to MR1 (**Extended data Figure 4, Figures 3d & 3e**). Both of these one-ring (pyrimidine) compounds were thus captured by MR1, despite being relatively unstable in the absence of MR1 and readily undergoing dehydrative cyclization to compounds **4d** and **4c**, respectively. The aromatic pyrimidine ring systems of 5-OP-RU and 5-OE-RU superposed on the corresponding ring from the bicyclic lumazine RL-6-Me-7-OH[16] (**Figure 3c**). The creation of 5-OP-RU or 5-OE-RU generated an aliphatic moiety that burrowed into the MR1 cleft, within which the residual carbonyl group formed a Schiff base with Lys43 of MR1 (**Figures 3d**

& 3e). This aliphatic moiety was also stabilised in the cleft by interactions with Tyr7 and Tyr62 (**Figures 3d & 3e**). In contrast, RL-6-Me-7-OH was non-covalently bound within MR1 (**Figure 3f**). Moreover, RL-6-Me-7-OH does not have the propensity to tautomerise into a single ring pyrimidine system due to its ability to form a very stable amide-tautomer. Nevertheless, the ribityl moieties of 5-OP-RU, 5-OE-RU and RL-6-Me-7-OH were all located in essentially identical positions within their respective complexes, each forming a hydrogen bond to Tyr95α [16] (**Figures 3f, 3g & 3h**). Notably, 5-OP-RU and 5-OE-RU are relatively unstable in aqueous media and thus MR1 can capture and stabilise pyrimidine intermediates in the synthesis of lumazines.

Next we undertook ESI-TOF-MS to independently identify the chemical composition of the ligands captured within these refolded MR1-Ag complexes. For MR1 refolded with 5-A-RU and methylglyoxal, a single peak with retention time of 8.9 minutes and m/z 329.11 matched a species that was captured by MR1 from *Salmonella* supernatant [5] and from the reaction mixture during synthesis of rRL-6-CH_2OH [5] (**Extended data Figure 5, upper set of panels, Extended data Figure 5, upper set of panels**). This finding is consistent with the identification within the crystal structure with MR1 of 5-OP-RU, independently assembled from 5-A-RU and methylglyoxal (**Figure 3d**), and supported by the NMR and kinetic characterisation of 5-OP-RU in solution (**Figure 2b, Extended Data Figures 2,3**). Similarly EIC, MS, and MS/MS analysis of MR1 refolded with the mixture of 5-A-RU and glyoxal revealed precursor and product m/z values (**Extended Data Figure 5, middle set of panels; Extended Data Figure 5, lower set of panels**) consistent with identification of 5-OE-RU within the crystal structure of MR1 refolded with 5-A-RU and glyoxal (**Figures 3e, 3h**). Furthermore, mass spectrometric analysis of MR1 refolded with 5-A-RU and ^{13}C-labeled glycolaldehyde yielded expected m/z 317.10 precursor and 179.04 product ions, in agreement with the m/z 315.09 precursor and 177.04 product ions identified in MR1 refolded with 5-A-RU and glyoxal (**Extended Data Figure 5, lower set of panels; Extended Data Figure 5, lower set of panels**).

We asked whether the activity with synthetic rRL-6-CH_2OH [5], might reflect capture by MR1 of a synthetic intermediate. The ligand captured by the mutant K43A-MR1 exposed to the reaction mixture generating rRL-6-CH_2OH [5] was identical by LCMS and MS/MS analysis (m/z 329.11) to the MR1-bound Ag from either *Salmonella* supernatant or derived from 5-A-RU/methylglyoxal condensation (data not shown, **Extended Data Figure 5b**). We therefore evaluated whether the respective MR1-tetramers formed from these distinct synthetic Ags were similar functionally. MR1-5-OP-RU and MR1-5-OE-RU tetramers efficiently stained all human MAIT cells present in PBMCs similarly to the mutant K43A-MR1 tetramers (**Figures 4a & 4b**). We solved the crystal structure of the MAIT TCR-K43A-MR1-Ag complex, which revealed 5-OP-RU as the ligand bound to K43A-

MR1, indicating that MR1 captures an intermediate from the synthesis of rRL-6-CH_2OH (**Extended data Figure 4, Extended Data Table 1**). Thus active MAIT cell ligands are intermediary, open-ring precursors to ribityllumazines that arise from condensing 5-A-RU with small molecule metabolites.

Recombinant MR1 refolded in the presence of folate-deficient culture supernatant from *Salmonella typhimurium* (strain SL1344) captured a dominant species of m/z 329.11 [5]. Mass spectrometric analysis of MR1 refolded with supernatant from *E. coli* (DH5α) also revealed a distinct and abundant m/z 315.09 species with matching liquid chromatography retention time, MS and MS/MS properties to those observed with MR1-5-OE-RU (**Extended Data Figure 6** and data not shown). Closer analysis of the MR1-eluate from *S. typhimurium* also revealed the presence of a m/z 315.09 species, albeit this ligand was much less prevalent (data not shown). Accordingly bacteria with an active riboflavin pathway can produce distinct MAIT activating ligands, the relative abundance of which is dependent upon the bacterial source.

Next, to establish if bacteria also produced free 5-A-RU, ^{13}C-labelled glycolaldehyde was added to *E. coli* supernatant, which was subsequently refolded with MR1. We detected a species with a m/z of 317.1 from these MR1 eluates (**Figure 4c**), consistent with the m/z 317.1 species observed previously (**Extended data Figure 5, lower set of panels**). This indicates that there is sufficient free 5-A-RU released by bacteria to conjugate with exogenously added metabolites. Potent MAIT cell Ags could also potentially be generated by host-derived metabolites forming adducts with 5-A-RU, in a manner somewhat analogous to the genesis of a CD1b-restricted antigen [17]. To test this, we added 5-A-RU to C1R cells transduced with MR1 (C1R.MR1), which led to MR1 cell surface upregulation and activation of Jurkat.MAIT cells (**Figures 4d & 4e**). When exogenous glyoxal or methylglyoxal were added with 5-A-RU to C1R.MR1 cells, we observed a further increase in MR1 upregulation and an increase in Jurkat.MAIT activation, when compared to 5-A-RU added by itself (**Figures 4d & 4e**). Notably, MR1 surface expression was not enhanced, nor was there an increase in Jurkat.MAIT activation, upon co-addition of butane-2,3-dione with 5-A-RU (**Figures 4d & 4e**). These observations suggest that MR1-Ag complexes created from 5-A-RU and glyoxal or methylglyoxal are natively conformed. To test this, the mutation of Tyr95α to either Ala95 or Phe95 ablated recognition of C1R cells to which 5-A-RU had been added, in a manner similar to that observed when synthetic rRL-6-CH_2OH was added to C1R cells [16], consistent with the notion that 5-A-RU is converted to 5-OE-RU or 5-OP-RU within C1R Ag presenting cells (**Figure 4f**). Accordingly, the bacterial riboflavin metabolite, 5-A-RU, can interact with host-derived metabolites analogous functionally to the creation of MR1 ligands found in bacterial supernatant.

We show a unique mechanism for creating T cell ligands from disparate metabolite building blocks. The potent MAIT activating ligands arise from a "core precursor element" of the microbial riboflavin pathway that forms simple adducts with distinct chemical metabolites, via a mechanism that does not require enzymatic catalysis. Thus, MR1 captures, stabilizes and presents otherwise transitory chemical intermediates for MAIT cell recognition. This represents a sophisticated discriminatory mechanism for targeting microbial antigens and protecting the host, whereby distinct metabolic pathways converge to produce T cell antigens.

Methods Summary

Bacterial strains and mutants

Lactococcus lactis strains NZ9000 (wild-type), the NZ9000 RibA- deletion mutant, and the CB013 and CB021 roseoflavin resistant mutants have been previously described [11]. *Salmonella typhimurium* strains SL1344 and BRD509 have been previously described [18]. See Methods for full details of CB013-derivatives and Salmonella Δ*rib*DH mutants.

Chemical syntheses, characterisation and stability

See Methods, Extended data and supplementary information for full details.

MAIT activation and MR1 upregulation assays

Conducted as described previously [5]. See Methods for full details.

MR1 refolding

Conducted as described previously [5]

Mass spectrometry analysis

Conducted as described previously [5]

MR1 tetramer generation

K43A-MR1 tetramers generated as described [6]. See methods for MR1-5-OP-RU and MR1-5-OE-RU tetramers.

MR1 tetramer staining

Conducted as described previously [6].

Crystallization and structure determination

The MAIT TCR in complex with MR1-5-OP-RU, MR1-5-OE-RU and K43A-MR1-5-OP-RU were crystallized and structure determined as described in methods.

References

1 Brigl, M. & Brenner, M. B. CD1: Antigen Presentation and T Cell Function. *Annu Rev Immunol* **22**, 817-890 (2004).

2 Yewdell, J. W. & Haeryfar, S. M. Understanding presentation of viral antigens to CD8+ T cells in vivo: the key to rational vaccine design. *Annu Rev Immunol* **23**, 651-682 (2005).

3 Eckle, S. B. G., Turner, S. J., Rossjohn, J. & McCluskey, J. Predisposed αβ T cell antigen receptor recognition of MHC and MHC-I like molecules? *Current Opinion in Immunology* **25**, 653-659 (2013).

4 Rossjohn, J., Pellicci, D. G., Patel, O., Gapin, L. & Godfrey, D. I. Recognition of CD1d-restricted antigens by natural killer T cells. *Nat Rev Immunol* **12**, 845-857 (2012).

5 Kjer-Nielsen, L. *et al*. MR1 presents microbial vitamin B metabolites to MAIT cells. *Nature* **491**, 717-723 (2012).

6 Reantragoon, R. *et al*. Antigen-loaded MR1 tetramers define T cell receptor heterogeneity in mucosal-associated invariant T cells. *J Exp Med* **210**, 2305-2320 (2013).

7 Gold, M. C. & Lewinsohn, D. M. Co-dependents: MR1-restricted MAIT cells and their antimicrobial function. *Nat Rev Micro* **11**, 14-19 (2013).

8 Treiner, E. *et al*. Selection of evolutionarily conserved mucosal-associated invariant T cells by MR1. *Nature* **422**, 164-169 (2003).

9 Birkinshaw, R. W., Kjer-Nielsen, L., Eckle, S. B. G., McCluskey, J. & Rossjohn, J. MAITs, MR1 and vitamin B metabolites. *Current Opinion in Immunology* **26**, 7-13 (2014).

10 Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends in genetics : TIG* **20**, 44-50 (2004).

11 Burgess, C., O'Connell-Motherway, M., Sybesma, W., Hugenholtz, J. & van Sinderen, D. Riboflavin Production in Lactococcus lactis: Potential for In Situ Production of Vitamin-Enriched Foods. *Applied and Environmental Microbiology* **70**, 5769-5777 (2004).

12 Cushman, M. *et al*. Design, Synthesis, and Evaluation of 6-Carboxyalkyl and 6-Phosphonoxyalkyl Derivatives of 7-Oxo-8-ribitylaminolumazines as Inhibitors of Riboflavin Synthase and Lumazine Synthase. *The Journal of Organic Chemistry* **67**, 5807-5816 (2002).

13 Bacher, A., Eberhardt, S., Fischer, M., Kis, K. & Richter, G. BIOSYNTHESIS OF VITAMIN B2 (RIBOFLAVIN). *Annual Review of Nutrition* **20**, 153-167 (2000).

14 Kis, K., Kugelbrey, K. & Bacher, A. Biosynthesis of Riboflavin. The Reaction Catalyzed by 6,7-Dimethyl-8-ribityllumazine Synthase Can Proceed without Enzymatic Catalysis under Physiological Conditions. *The Journal of Organic Chemistry* **66**, 2555-2559 (2001).

15 Wang, Y. & Ho, C.-T. Flavour chemistry of methylglyoxal and glyoxal. *Chemical Society Reviews* **41**, 4140-4149 (2012).

16 Patel, O. *et al*. Recognition of vitamin B metabolites by mucosal-associated invariant T cells. *Nat Commun* **4**, 2142 (2013).

17 Moody, D. B. *et al*. Cd1b-Mediated T Cell Recognition of a Glycolipid Antigen Generated from Mycobacterial Lipid and Host Carbohydrate during Infection. *The Journal of Experimental Medicine* **192**, 965-976, doi:10.1084/jem.192.7.965 (2000).

18 Hoiseth, S. K. & Stocker, B. A. D. Aromatic-dependent Salmonella typhimurium are non-virulent and effective as live vaccines. *Nature* **291**, 238-239 (1981).

19 Strugnell, R. *et al*. Characterization of a Salmonella typhimurium aro vaccine strain expressing the P.69 antigen of Bordetella pertussis. *Infection and Immunity* **60**, 3994-4002 (1992).

20 Plaut, G. W. E. & Harvey, R. A. in *Methods in Enzymology* Vol. Volume 18, Part B (eds B. McCormick Donald & D. Wright Lemuel) 515-538 (Academic Press, 1971).

21 Huang, S. *et al*. Evidence for MR1 Antigen Presentation to Mucosal-associated Invariant T Cells. *Journal of Biological Chemistry* **280**, 21183-21193 (2005).
22 Reantragoon, R. *et al*. Structural insight into MR1-mediated recognition of the mucosal associated invariant T cell receptor. *The Journal of Experimental Medicine* **209**, 761-774 (2012).

Acknowledgments

We thank the staff of the Monash crystallization facility and the Australian Synchrotron for assistance with crystallization and data collection respectively, and Malcolm J. McConville, Paul O'Donnell and Ching-Seng Ang from the University of Melbourne Bio21 Institute mass spectrometry platform for assistance with mass spectrometry experiments. The Australian Research Council (ARC), a Program Grant and Project Grant of the National Health and Medical Research Council of Australia (NHMRC) supported this research; D. van Sinderen is a recipient of a Science Foundation Ireland (SFI) Principal Investigator award (Ref. No. 08/IN.1/B1909); OP was supported by an ARC Future Fellowship; DPF was supported by NHMRC Senior Principal Research Fellowship (1027369); JR was supported by an NHMRC Australia Fellowship.

Author Contributions

A.J.C, S.B.G.E, R.W.B. L.L are joint first authors. O.P, J.M, Z.C, R.R, B.M, D.V.S, J.Y.W.M. either performed experiments, provided key reagents, and/or analysed data and/or provided intellectual input or helped write the manuscript. D.P.F., L.K.N, JMcC. and J.R. co-led the investigation and contributed to design and interpretation of data, project management, and writing of the manuscript. D.P.F, L.K.N, JMcC and JR are the joint senior and corresponding authors.

Author Information

The atomic coordinates and structure factors for the TCR-MR1-Ag complexes were deposited in the Protein Data Bank (PDB codes: 4NQC, 4NQD and 4NQE). Reprints and permission information is available at www.nature.com/reprints. The authors declare that they have no competing financial interests. Correspondence and requests for materials should be addressed to DPF (email: d.fairlie@imb.uq.edu.au), LKN (e-mail: lkn@unimelb.edu.au), JMcC (e-mail: jamesm1@unimelb.edu.au) or JR (e-mail: jamie.rossjohn@monash.edu).

Figure Legends

Figure 1.The riboflavin pathway furnishes ligands that activate MAIT cells

(**A**) Riboflavin biosynthesis pathway. *RibH*; lumazine synthase. X; hypothetical phosphatase.

(**B**) Cells were incubated overnight with filtered S/N from *L.lactis* NZ9000 (wt), *ribA*- (ribA deletion mutation), CB013 and CB021 (riboflavin overproducers) overnight cultures +/- 3 μg/ml riboflavin then stained for CD3-PE and anti-CD69-APC. MFI CD69-APC for gated Jurkat.MAIT cells, mean +/- SEM. (**C**) Cells were incubated overnight with 10 μl filtered, culture S/N from *Lactococcus* CBO13 (deregulated riboflavin operon), CB013ΔRibA, CB013ΔRibB, CB013ΔRibG and CB013ΔRibH or *S.typhimurium*, then stained for CD3-PE and CD69-APC. MFI CD69-APC for gated Jurkat.MAIT cells, mean +/- SEM. Experiments performed at least 3 times.

Figure 2. Chemical formation of pyrimidines and lumazines from condensation of small metabolites with 5-A-RU

(A) Series a, 5-A-RU (**1**) and 3,4-dihydroxy-2-butanone-4-phosphate (**2a**) forms 5-MOP-RU, **3a**) and RL-6,7-DiMe (**4a**). Series b-d, 5-A-RU (**1**) with butane-2,3-dione (**2b**), glyoxal (**2c**), and methylglyoxal (**2d**) forms 5-MOP-RU (**3b** = **3a**), 5-OE-RU (**3c**) 5-OP-RU (**3d**) respectively, and RL-6,7-DiMe (**4b** = **4a**), ribityllumazine RL (**4c**) RL-7-Me (**4d**). (**B**) 2D NMR spectrum (HMBC) of isolated 5-OP-RU (**3d**) in DMSO-d_6 showing key ^{1}H-^{13}C long range correlations that unambiguously characterize the imine adduct (**3d**), also identified in aqueous media (pH>6).

Figure 3. Structural basis of MR1-binding and recognition of transitory MAIT cell antigens

(**A**) MAIT TCR-MR1-Ag (**B**) docking, (**C**) footprint (**D**) 5-OP-RU and RL-6-Me-7-OH overlay; MR1 contacting (**E**) 5-OP-RU and (**F**) 5-OE-RU; MAIT TCR contacting (**G**) RL-6-Me-7-OH, (**H**) 5-OP-RU or (**I**) 5-OE-RU. MR1 (grey), α-chain (purple), β-chain (cyan). 5-OP-RU (green), 5-OE-RU (yellow) and RL-6-Me-7-OH (magenta), slate (CDR1α), pink (CDR2α) yellow (CDR3α), teal (CDR1β), red (CDR2β) and orange (CDR3β).

Figure 4. MR1-Ag tetramers and MAIT activation.

A) Gating strategy (left), tetramers of MR1-6-FP, MR1-5-OP-RU, MR1-5-OE-RU, or anti-TRAV1-2 (**B)** PBMC co-staining with MR1-5-OP-RU and K43A-MR1-5-OP-RU tetramers. **(C)** EICs of m/z 315.09, or m/z 317.10. (**D**) (i) Activation assay and (ii) MR1 upregulation with 5-A-RU, methylglyoxal (MG), butane-2,3-dione (BD), glyoxal (G), or rRL-6-CH_2OH. MFI of 26.5-PE antibody staining (**E**) (i) CD69 upregulation and (ii) MR1 upregulation with 5-A-RU, MG, BD, or G.

MFI of 26.5-PE. (**F**) Activation of wild-type, Y95A, or Y95F mutant SKW.MAIT cells by 5-A-RU. Mean of triplicates with SEM. Experiments were performed at least twice (b, f) or 3 times (a, c, d, e).

Methods

Bacterial strains and mutants

Lactococcus lactis strains NZ9000 (wild-type), the NZ9000 RibA- deletion mutant, and the CB013 and CB021 roseoflavin resistant mutants have been previously described [11]. CB013-derivatives: CB013ΔRibA, CB013ΔRibB, CB013ΔRibG and CB013ΔRibH were generated by insertion of *Eco*RI or *Eco*RV restriction sites incorporating either one or two stop codons into the individual genes using standard techniques. Inserted sequences:

ΔRibG attaacgtttccccctccttttcgagccagtGAATTCaggattgctaaattcataaaatgctcatcattttccat

ΔRibB gatgctggaagctcgaatgattaatttagacgGAATTCATTAacttatctctcttttgaatttgagttacctctcctat

ΔRibA tatcttttctggactaatcatttcggctgcacGAATTCaaatcagatctccttcattttctctattctcatcatc

ΔRibH ggcccctgctaagagttttgcgttgatgaaGATATCTTATTAttcgttaaaacgtgcaactacaattccatatttgg

The genotype of each mutant was verified by sequencing and multiple rounds of PCRs based on the mutated region to verify the purity of the genotype. The phenotype of each mutant was also checked using growth/absence of growth in Riboflavin assay medium overnight at 30 °C and comparing to a control of CB013. All mutants were unable to grow in the media while CB013 was capable of growth in Riboflavin-limiting conditions as it is an over-producer. *L. lactis* strains were grown at 30 °C without shaking, in M17 medium (Difco) containing 1% glucose and the addition of 3 μg/ml riboflavin where indicated. *Salmonella typhimurium* SL1344 was grown at 37 °C without shaking, in M17 medium (Difco) containing 1% glucose.

Salmonella typhimurium strains SL1344 and BRD509 have been previously described [18]. The Salmonella Δ*rib*DH mutants were constructed on an SL1344 background by lamda red- recombinase mediated allelic replacement followed by general transduction using phage P22 as previously described [19] resulting in strain SL1344ΔRibDH.

Primers:

B2(Sec)F: 5'-TAG GGA TAA CAG GGT AAT-GGT TCG ATA GCG TAA TGG

B2(Sec)R: 5'-TAG GGA TAA CAG GGT AAT-TAT CTT TCC GGC CTG TGA

B2(Kan)F: 5'-CTA AGG AGG ATA TTC ATA TG-GAC CGC GCT TGA AAT GAT

B2(Kan)R: 5'-GAA GCA GCT CCA GCC TAC ACA-ATT GTT AAC AAT GAC ACA

The complement of mutants was performed by transformation of *rib*DH genes resulting in strain SL1344ΔRibDH:RibDH. Mutation and reconstitution were verified by lack of growth or growth on Luria Agar, and by PCR. Mutants were grown on Luria agar containing 20 μg/ml riboflavin.

For MR1 refolds *Salmonella* wt and mutant strains were grown in M9 minimal media supplemented with histidine (77.6 μg/ml) and streptomycin (25 μg/ml) and 3 μg/ml riboflavin. *Enterococcus faecalis* was grown in Folic Acid Assay Medium (Difco) at 37 ˚C without agitation. *E.coli* DH5α was grown in M9 media. *L.lactis* CB013 and CB013 rib mutants were grown in Folic Acid Assay Medium (Difco) supplemented with Xanthine (6 μg/ml) and Yeast Nitrogen Base (6.8 mg/ml) at 30 ˚C without agitation.

Compounds

Glyoxal, methylglyoxal, 1,3-dihydroxyacetone dimer, DL-glyceraldehyde and butane-2,3-dione were purchased from Sigma. [1,2-$^{13}C_2$]glycolaldehyde (glycolaldehyde can be readily air-oxidized to form glyoxal) was purchased from Omicron Biochemicals. A synthesis of rRL-6-CH_2OH has been previously described [5]. 5-amino-6-D-ribitylaminouracil was freshly prepared from 5-nitroso-6-D-ribitylaminouracil following a literature procedure [20]. In brief, 5-nitroso-6-D-ribitylaminouracil (40.0 mg, 0.138 mmol, 1 eq) was dissolved in MilliQ water (3 mL) at 80 ºC under argon. To the red solution was added sodium dithionite powder (1.2–3.3 eq). The colour changed instantly to pale-yellow. After stirring at 80 ºC for 5 min, the solution was cooled under argon in an ice-water bath. For biological studies, the chilled solution was diluted with MilliQ water to make 50 mM stock solutions and stored in 1.5 mL aliquots at –20 ºC for later use.

For NMR characterisation of pyrimidine intermediate 5-OP-RU (**3d**), a freshly prepared solution of 5-amino-6-D-ribitylaminouracil (5-A-RU, **1**) was adjusted to pH 7.0 with 1M sodium hydroxide solution, lyophilized, dissolved in DMSO-d_6 and then filtered to removed salts. The solution was transferred to an NMR tube, filled with argon, and the concentration of 5-A-RU determined by NMR spectroscopy. Methylglyoxal (2 equivalents) was added, and the reaction monitored by NMR. Upon completion, 5-OP-RU was further purified using a Shimadzu preparative HPLC system equipped with a Phenomenex Luna 10 micron C18 250 × 21.20 mm column (P/No 00G-4253-PO-AX) and a SPD-M20A diode array detector. Flow rate was 20 mL/min with linear gradient: 100% solvent A to 100% solvent B over 30 min where solvent A was 20 mM ammonium acetate in H_2O and solvent B was 20 mM ammonium acetate in MeCN–H_2O (80:20, v/v). Compound **3d** was fully characterised by ESI-HRMS (Calculated for $C_{12}H_{17}N_4O_7^-$ m/z 329.1103, measured m/z 329.1116) and 1D and 2D NMR spectroscopy (**Figure 2b & Extended Data Figure 2**).

5-OP-RU (3d) ^{1}H NMR (600 MHz, DMSO-d_6), δ 2.28 (3H, s), 3.38–3.43 (2H, m), 3.47–3.51 (1H, m), 3.52–3.55 (1H, m), 3.56–3.59 (2H, m), 3.73 (1H, m), 7.43 (1H, br s), 8.80 (1H, s); ^{13}C NMR (150 MHz, DMSO-d_6) δ 23.5, 44.1, 63.1, 70.7, 72.8, 72.9, 98.5, 142.0, 152.1, 157.6, 159.1, 200.2. ESI-HRMS calcd for $C_{12}H_{17}N_4O_7$ [M - H]: 329.1103, found: 329.1116.

Stability of 5-OP-RU in aqueous media

The purified 5-OP-RU was dissolved in aqueous TBS buffer (10 mM Tris, 150 mM NaCl, pH 8.0), MilliQ water (pH 6.8), or aqueous ammonium acetate buffer (20 mM, pH 5.4). The consumption of 5-OP-RU was immediately monitored by LCMS. The initial concentration was quantified by comparing to standard solution with known concentration. At 15 ºC, the half-life was determined as 14.5–15 h at pH 8.0 independent of the starting concentrations (65–250 μM), 14.2 h at pH 6.8 (65 μM), and 49 min at pH 5.4 (65 μM). At 37 ºC, pH 6.8, the half-life was 135 min (**Extended Data Figure 3**).

Activation of Jurkat.MAIT and SKW.MAIT cells and detection of MR1 expression on C1R.MR1 cells

Jurkat cells transduced with genes encoding a MAIT TCR comprising the TRAV1-2-TRAJ33 invariant α chain, and a TRBV6-1 β chain, or SKW cells transduced with genes encoding the TRAV1-2-TRAJ33 invariant α chain with either w.t. Tyr95 or mutated Tyr95Ala or Tyr95Phe residues, paired with a TRBV6-1 β chain, were tested for activation by co-incubation with bacterial culture supernatant or compounds and C1R antigen presenting cells expressing MR1 (CIR.MR1, with Jurkat.MAIT cells), or C1R cells (SKW.MAIT cells) for 16 hr. Cells were subsequently stained with PE-Cy7-conjugated anti-CD3 (eBioscience), and APC-conjugated anti-CD69 (BD) antibodies as well as biotinylated anti-MR1 mAb 26.5 [21], followed by Streptavidin-PE (BD), before analysis by flow cytometry. Activation of Jurkat.MAIT or SKW.MAIT cells was measured by an increase in surface CD69 expression. MR1 expression was detected on gated C1R.MR1 cells in the same assay.

Preparation of denatured inclusion body MR1 and β2m

Genes encoding soluble human MR1 or human β2m were expressed for 4 hr in BL21 *E. coli* following induction with 1 mM isopropyl β-D-1-thiogalactopyranoside. *E. coli* were pelleted and resuspended in a buffer containing 50 mM Tris, 25% (w/v) sucrose, 1 mM EDTA, 10 mM DTT pH 8.0. Inclusion body protein was then extracted by lysis of bacteria in a buffer containing 50 mM Tris pH 8.0, 1% (w/v) Triton X-100, 1% (w/v) sodium deoxycholate, 100 mM NaCl, 10 mM DTT, 5 mM $MgCl_2$, and 1 mg DNaseI per litre of starting culture; and subsequent steps involved homogenization with a

polytron homogenizer, centrifugation, and washing inclusion body protein sequentially with firstly a buffer containing 50 mM Tris pH 8.0, 0.5% Triton X-100, 100 mM NaCl, 1 mM EDTA, 1 mM DTT, and secondly a buffer containing 50 mM Tris pH 8.0, 1 mM EDTA, 1 mM DTT. Inclusion body protein was then resuspended in a buffer containing 20 mM Tris pH 8.0, 8 M urea, 0.5 mM EDTA, 1 mM DTT, and following centrifugation the supernatant containing solubilized, denatured inclusion body protein was collected and stored at -80 ºC.

Refolding of MR1-ligand and MAIT TCR

MR1 and β2m were refolded with ligand essentially as described [5]. Briefly, in order to generate MR1-5-OP-RU and MR1-5-OE-RU, 56 mg of MR1 and 28 mg of β2m inclusion body proteins, together with with 2.9 mg of 5-A-RU and 254, 204, or 204 mg of methylglyoxal, glyoxal (Sigma), or ^{13}C-glycolaldehyde (Omicron) respectively were added to a 400 ml refold solution containing 0.1MTris, pH 8.5, 2mMEDTA, 0.4M arginine, 0.5mM oxidized glutathione and 5mM reduced glutathione. Refolded MR1-Ag was then purified by sequential DEAE (GE Healthcare) anion exchange, S75 16/60 (GE Healthcare) gel filtration, and MonoQ (GE Healthcare) anion exchange chromatography. Alternatively, 56 mg of MR1 and 28 mg of β2m inclusion body proteins were refolded in the presence of 400 ml of 0.45 μM-filtered bacterial supernatants or control media, in the absence or presence of 204 mg ^{13}C-glycolaldehyde. The TRBV6-1 MAIT TCR was expressed, refolded and purified essentially as previously described [22].

Analysis of MR1-5-OP-RU and MR1-5-OE-RU by mass spectrometry.

MR1-5-OP-RU or MR1-5-OE-RU (4 μg) were loaded onto an XBridge C18 reversed phase column (Waters) in 20 mM ammonium acetate, pH 5.4, buffer, and detected in an Agilent ESI–TOF mass spectrometer after elution in an acetonitrile gradient. Data was collected in negative ion mode. Different instrumentation resulted in slight variations in retention times of the m/z 329.11/315.09/317.10 species to that reported previously[3].

Generation of MR1-K43A, MR1-5-OP-RU and MR1-5-OE-RU tetramers

The generation of K43A-MR1 tetramers, loaded with synthetic rRL-6-CH_2OH, has been previously described [6]. Briefly, refolded and purified empty C-terminal cysteine-tagged-MR1-K43A was loaded with a 136 molar excess of synthetic rRL-6-CH_2OH for 4 hr at room temperature in the dark. C-terminal cysteine-tagged wild type MR1-5-OP-RU and MR1-5-OE-RU were generated as described above.

Cysteine-tagged-K43A-MR1-5-OP-RU, or cysteine-tagged w.t. MR1-5-OP-RU or cysteine-tagged w.t. MR1-5-OE-RU were then reduced with 5 mM DTT for 20 min prior to buffer exchange into PBS using a PD-10 column (GE Healthcare), and biotinylated with Maleimide-PEG2 biotin (Thermoscientific) with a 30:1 molar ratio of biotin:protein at 4°C for 16 hr in the dark. Biotinylated MR1 was subjected to S200 10/300 GL (GE Healthcare) chromatography to remove excess biotin. Biotinylated, loaded K43A-MR1, or w.t. MR1-5-OP-RU or w.t. MR1-5-OE-RU monomers were tetramerized with streptavidin conjugated to either PE (SA-PE) or Brilliant Violet 421 (SA-BV) (BD Pharmingen).

Isolation of PBMCs

Whole blood from healthy donors was collected (Australian Red Blood Cross Service) and peripheral blood mononuclear cells were separated using Ficoll-Paque Premium (GE Healthcare). PBMCs were harvested and resuspended in fresh RPMI medium. Cells were then washed twice prior to resuspension in 10% DMSO in FCS. Prior to use, PBMCS were stored in liquid nitrogen.

Tetramer staining of human PBMCs.

For co-staining with wild-type and K43A-MR1 tetramers, approximately 5 x 10^5 human PBMCs were stained with K43A-MR1-5-OP-RU-PE tetramer at 20 μg/ml for 40 minutes at room temperature in the dark. Cells were then washed and stained with w.t. MR1-5-OP-RU-BV tetramer at 1.4 μg/ml, CD3-AlexaFluor700 (EBioscience), CD161-PE-Cy7 (Biolegend), CD4-APC-Cy7 (Biolegend) and CD8α-PerCp-Cy5.5 (BD) for 30 minutes at 4°C. Cells were then washed once with 2 ml of FACS wash (2% fetal bovine serum in PBS) and resuspended in 150 μl of FACS fix (2% glucose and 1% paraformaldehyde in PBS) prior to acquisition of data on a BD LSR-Fortessa. Data were analyzed using FlowJo analysis software (Tree Star, Ashland, OR).

For single staining with either MR1-5-OP-RU or MR1-5-OE-RU tetramers, human PBMCs were stained as above with w.t. MR1-5-OP-RU-PE or w.t. MR1-5-OE-RU-PE tetramers at 1.4 ug/ml, and CD3-AlexaFluor700 (EBioscience), CD161-PE-Cy7 (Biolegend), CD4-APC-Cy7 (Biolegend) and CD8α-PerCp-Cy5.5 (BD) for 30 minutes at 4°C, prior to acquisition of data on a BD LSR-Fortessa.

Crystallization and structure determination

Crystals of the soluble MAIT TCR-MR1-Ag complexes were obtained using the hanging drop vapour diffusion method. The MR1-β_2M-5-OP-RU, MR1-β_2M-5-OE-RU, MR1, MR1-K43A-β_2M-5-OP-RU and MAIT TCR were concentrated to 4 mg/ml, mixed in a 1:1 molar ratio, then 1 0.5 μl added to 0.5 μl of a precipitant solution consisting 0.1 M bis-tris propane pH 6.3, 0.2 M sodium acetate and varying concentrations of PEG 3350 between 8-14% w/v. Crystals were observed after incubation at 20°C

for 24 hours in dark conditions and cryoprotected prior to diffraction experiments by soaking in the crystallisation condition modified with between 10-15% v/v glycerol before cooling to 100K. Diffraction images were collected at the Australian Synchrotron MX2 beamline diffracting in a *C2* spacegroup to 2.50 Å, 2.10 Å and 2.20 Å for the MR1-β_2M-5-OP-RU, MR1-β_2M-5-OE-RU, and MR1-K43A-β_2M-5-OP-RU MAIT TCR complexes respectively. The data were processed using Mosflm version 7.0.9 and scaled using AIMLESS or SCALA (MR1-K43A-β_2M-5-OP-RU only) from the CCP4 Suite [23]. The phase problem was solved by molecular replacement using PHASER [24], using MR1 ternary complex (PDB code 4L4T)[16] with CDR loops and ligands removed and using the R_{free} reflection set from the model. The initial solution was refined in Phenix using simulated annealing refinement, with all subsequent refinement steps performed using BUSTER 2.10 [25]. Restraints for 5-OP-RU and 5-OE-RU were generated using the Grade Web Server, with model building performed in COOT using MolProbity for validation [26]. All molecular graphics were made with PyMOL. The buried surface area was calculated with Areaimol and RMSD values calculated using Superpose, both within the CCP4 suite.

Methods References

23 CCP4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-763 (1994).

24 McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* **63**, 32-41 (2007).

25 Zwart, P. H. *et al*. Automated structure solution with the PHENIX suite. *Methods Mol Biol* **426**, 419-435 (2008).

26 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).

Extended data Table 1

	MAIT-MR1-5-OP-RU	MAIT-MR1-5-OE-RU	MAIT-K43A-MR1-5-OP-RU
Data collection			
Temperature	100K	100K	100K
Space group	C2	C2	C2
Cell dimensions□□			
a, *b*, *c* (Å)	218.76, 71.11, 144.28	218.11, 70.60, 143.86	215.58, 68.87, 142.98
□□□□□□□□□□□□□□ (°)	90, 104.87, 90	90, 104.63, 90	90, 104.86, 90
Resolution (Å)	33.42-2.50 (2.55-2.50)	75.41-2.10 (2.21-2.10)	50.00-2.20 (2.3-2.20)
R_{pim}[1]	9.4 (38.8)	6.1 (35.2)	5.9 (36.7)
I/σ_1	7.8 (2.3)	8.1 (2.1)	9.7 (2.3)
Completeness (%)	100 (100)	98.6 (97.1)	97.9 (97.4)
Total N°. observations	307877 (19059)	462978(62837)	509054 (74702)
N°. unique observations	74555 (4584)	122109(17496)	101222 (14608)
Multiplicity	4.1 (4.2)	3.8 (3.6)	5.0 (5.1)
Refinement statistics			
R_{factor}[2] (%)	16.5	18.4	20.8
R_{free}[3] (%)	21.6	22.2	24.5
No. atoms			
• Protein	12424	12396	12514
• Ligand	45	42	46
• Water	1044	900	488
Ramachandran plot (%)			
• Most favoured	97.4	91.4	91.5
• Allowed region	2.5	8.6	8.5
B-factors (Å^2)			
• Protein	29.8	37.9	37.4
• ligand	14.8	23.8	26
rmsd bonds (Å)	0.010	0.010	0.010
rmsd angles (°)	1.16	1.05	1.08

Extended Data Information. Corbett et al

"T cell activation by transitory neo-antigens derived from distinct microbial pathways"

Guide

The extended data comprises one Table and eight figures and associated figure legends

Extended data Figure 1. MR1 ligand identification from different bacterial strains

(A) Detection of m/z 329.11 species in MR1 refolded with 5-A-RU and methylglyoxal (Control), and supernatants from wild-type (CB013), and CB013-derivatives (i.e. ΔRibA, ΔRibB, ΔRibG, or ΔRibH) *L.lactis* bacteria. Shown are counts on the Y-axis versus retention time on the X-axis. (**B**) Lack of activation of Jurkat.MAIT cells by supernatant from mutant ΔRibD/H *S.typhimurium* (strain SL1344) but not. wild-type (wt), or ΔRibD/H + RibD/H bacteria. Shown is MFI of CD69.APC on the Y-axis. (**C**) Detection of m/z 329.11 species in MR1 refolded with supernatants from wild type, ΔRibD/H, or ΔRibD/H + RibD/H *S.typhimurium* bacteria, or control media. Shown are counts on the Y-axis versus retention time on the X-axis. (**D**) Detection of m/z 329.11 species in MR1 refolded with 5-A-RU and methylglyoxal (Control), or bacterial supernatants from *L. lactis* (CBO13) or *E. faecalis* bacteria, or control media. Shown are counts on the Y-axis versus retention time on the X-axis. (**E**) Detection of matching m/z 329.11 species in MR1 refolded with supernatants from *S. typhimurium* (SL1344) or *E. coli* bacteria, or control media. Shown are counts on the Y-axis versus retention time on the X-axis. Experiments A-E were performed three, three, three, two and three times respectively.

Extended Data Figure 2. NMR characterization of 5-OP-RU (**3d**) in DMSO-d_6 with internal solvent peak at 2.50 ppm and 39.52 ppm for ^{1}H and ^{13}C, respectively. (A) ^{1}H NMR (600 MHz); (B) HSQC. The compound 5-OP-RU (**3d**) was synthesised from the reaction of 5-A-RU and methylglyoxal in DMSO-d_6, and then isolated from aqueous media by rpHPLC. Although it was less stable in water, it could still be identified and characterised at pH > 6.

Extended Data Figure 3. (A) Reaction between 5-A-RU (0.5 mM) and methylglyoxal (3 eq) at pH 6.8, 37 ºC in MilliQ water. (B) Stability of purified 5-OP-RU (65 μM) at pH 6.8 and 37 ºC. The half-life was 135 min. (C) Stability of purified 5-OP-RU (65 μM) at variable pH in aqueous TBS buffer (10 mM Tris, 150 mM NaCl, pH 8.0), MilliQ water (pH 6.8), or ammonium acetate buffer (20 mM, pH 5.4) at 15 ºC. The half-lives were 15 h at pH 8.0, 14.2 h at pH 6.8, 49 min at pH 5.4.

Extended data Figure 4. Electron density for ligands, and associated contacts

Electron density of 5-OP-RU in MR1, 5-OE-RU in MR1 and 5-OP-RU in K43A-MR1. Final $2F_o$-F_c map, contoured at 1σ for (a) 5-OP-RU ,(b) 5-OE-RU in the MAIT TCR–MR1–Ag complex and (c) 5-OP-RU in the MAIT TCR–MR1-K43A–Ag complex. Simulated annealing omit maps showing unbiased F_o-F_c electron density, contoured at 3σ, for (d) 5-OP-RU and (e) 5-OE-RU in MR1 and (f) 5-OP-RU in MR1-K43A. MR1-K43A-5-OP-RU MAIT TCR complex showing contacts between MR1-K43A and (G) 5-OP-RU and contacts between MAIT TCR and (H) 5-OP-RU. MR1 is shown in grey and MAIT TCR CDR3α in yellow in and CDR3β in orange with ribbon representation and 5-OP-RU in cyan with stick representation. Hydrogen bonds are indicated with black dashed lines with a water molecule mediating hydrogen bonding between the CDR3β 5-OP-RU shown in dark blue sphere representation.

Extended data Figure 5. Chromatographic and mass spectrometry properties of MR1 ligands.

(A) Ligand eluted from MR1 complexed with product of (i) 5-A-RU and methylglyoxal condensation reaction (upper panels); (ii) 5-A-RU and glyoxal condensation reaction (middle panels) or 5-A-RU and ^{13}C-glycolaldehyde condensation reaction (bottom panels). Shown are extracted ion chromatograms (left); mass-to-charge (m/z) spectrum (centre); and product ions from targeted fragmentation (right). Black diamond: precursor ion. This experiment was performed three times, shown is a representative experiment. (**B**) Mass spectrometry characterisation of 5-OP-RU (upper) and 5-OE-RU and ^{13}C-5-OE-RU (lower).

Extended data Figure 6. Mass spectrometry of the 315.09 species

Extracted ion chromatograms of m/z 315.09 species in MR1 refolded with 5-A-RU and glyoxal (Control), or *E.coli* supernatant, or media. Shown are counts on the Y-axis versus retention time on the X-axis. This experiment was performed three times.

Extended Data Figure 7. NMR characterization of 8-D-ribityllumazine (**RL**) using various 1D and 2D NMR spectroscopic techniques. Spectra were recorded as a solution in D_2O–CD_3OD (9:1) with internal solvent peak at 3.31 ppm and 49.0 ppm for ^{1}H and ^{13}C, respectively. (A) ^{1}H NMR (600 MHz); (B) ^{13}C NMR (150 MHz); (C) HMBC.

Extended Data Figure 8. NMR characterization of 7-methyl-8-D-ribityllumazine (**RL-7-Me**) using various 1D and 2D NMR techniques. Spectra were recorded as a solution in D_2O–CD_3OD (9:1) with internal solvent peak at 3.31 ppm and 49.0 ppm for ^{1}H and ^{13}C, respectively. (A) ^{1}H NMR (600 MHz) and mechanism for deuterium exchange of CH_3 at position-7. Identical exchange was also observed

in pure D_2O at slower rate (not shown); (B) ^{13}C NMR (150 MHz) showing characteristic heptet from 7-CD_3 after complete deuterium exchange; (C) HMBC.

Extended data Table 1 Data collection and refinement statistics

[1] $R_{p.i.m} = \Sigma_{hkl} [1/(N-1)]^{1/2} \Sigma_i | I_{hkl,i} - \langle I_{hkl} \rangle | / \Sigma_{hkl} \langle I_{hkl} \rangle$

[2] $R_{factor} = (\Sigma | |F_o| - |F_c| |) / (\Sigma |F_o|)$ - for all data except as indicated in footnote 3.

[3] 5% of data was used for the R_{free} calculation

Values in parentheses refer to the highest resolution bin.

Extended data Table 1

	MAIT-MR1-5-OP-RU	MAIT-MR1-5-OE-RU	MAIT-K43A-MR1-5-OP-RU
Data collection			
Temperature	100K	100K	100K
Space group	C2	C2	C2
Cell dimensions□□			
a, *b*, *c* (Å)	218.76, 71.11, 144.28	218.11, 70.60, 143.86	215.58, 68.87, 142.98
□□□□□□□□□□□□□□ (°)	90, 104.87, 90	90, 104.63, 90	90, 104.86, 90
Resolution (Å)	33.42-2.50 (2.55-2.50)	75.41-2.10 (2.21-2.10)	50.00-2.20 (2.3-2.20)
R_{pim}[1]	9.4 (38.8)	6.1 (35.2)	5.9 (36.7)
I/σ_1	7.8 (2.3)	8.1 (2.1)	9.7 (2.3)
Completeness (%)	100 (100)	98.6 (97.1)	97.9 (97.4)
Total N°. observations	307877 (19059)	462978(62837)	509054 (74702)
N°. unique observations	74555 (4584)	122109(17496)	101222 (14608)
Multiplicity	4.1 (4.2)	3.8 (3.6)	5.0 (5.1)
Refinement statistics			
R_{factor}[2] (%)	16.5	18.4	20.8
R_{free}[3] (%)	21.6	22.2	24.5
No. atoms			
• Protein	12424	12396	12514
• Ligand	45	42	46
• Water	1044	900	488
Ramachandran plot (%)			
• Most favoured	97.4	91.4	91.5
• Allowed region	2.5	8.6	8.5
B-factors (Å^2)			
• Protein	29.8	37.9	37.4
• ligand	14.8	23.8	26
rmsd bonds (Å)	0.010	0.010	0.010
rmsd angles (°)	1.16	1.05	1.08

Supplementary information

Contains:

1) List of abbreviations used
2) NMR characterisation of the chemical compounds (Supp figures 1-5)
3) Supplementary chemical methods

Abbreviations used:

5-A-RU: 5-amino-6-D-ribitylaminouracil
5-MOP-RU: 5-(1-methyl-2-oxopropylideneamino)-6-D-ribitylaminouracil
5-OE-RU: 5-(2-oxoethylideneamino)-6-D-ribitylaminouracil
5-OP-RU: 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil
Ag: antigen
APC: allophycocyanin
□2M: beta-2-microglobulin
CD3: cluster of differentiation 3, a T-cell marker
CD4: cluster of differentiation 4, a T-cell subset marker
CD8: cluster of differentiation 8, a T-cell subset marker
CD69: cluster of differentiation 69, a marker of T cell activation
CD161: cluster of differentiation 161, a MAIT cell co-receptor marker
DN: double-negative ($CD4^-/CD8^-$), a T-cell subset
FACS: fluorescence-activated cell sorting
mAb: monoclonal antibody
MAIT cell: Mucosal Associated Invariant T cell
MR1: Major histocompatibility complex class I-related protein
m/z: mass to charge
PBMC: peripheral blood mononuclear cell
PBS: phosphate buffered saline
PE: phycoerythrin
RL: ribityllumazine
RL-6,7-DiMe: 6,7-dimethyl-8-D-ribityllumazine
RL-6-Me: 6-methyl-8-D-ribityllumazine
RL-6-Me-7-OH: 7-hydroxy-6-methyl-8-D-ribityllumazine
RL-7-Me: 7-methyl-8-D-ribityllumazine
rRL-6-CH_2OH: reduced 6-hydroxymethyl-8-D-ribityllumazine
TCR: T cell receptor
TRAV1-2: T cell receptor alpha variable 1-2

Supplementary Method 1

General. Analytical HPLC was performed using an Agilent 1200 Series with a diode-array detector on a Phenomenex Luna 5 μm, C18(2) 250□× 4.60 mm column. Flow rate was 1 mL/min with a linear gradient: 100% solvent A over 20 min, followed by 0% to 5% solvent B over 10 min, 5% solvent B over 20 min, 5% to 100% solvent B over 5 min and a further 5 min at 100% solvent B. Solvent A was 20 mM ammonium acetate in MilliQ H_2O and solvent B was 20 mM ammonium acetate in MeCN–H_2O (80:20, v/v). Preparative reversed phase HPLC (rpHPLC) was performed using a Shimadzu preparative HPLC system equipped with a Phenomenex Luna 10 micron C18(2) 250 × 21.20 mm column (P/No 00G-4253-PO-AX) and a SPD-M20A diode array detector. Flow rate was 20 mL/min with linear gradient: 100% solvent A over 20 min, followed by 0% to 5% solvent B over 10 min, 5% to 100% solvent B over 20 min and a further 10 min at 100% solvent B. Electrospray ionization high-resolution mass spectra (ESI-HRMS) measurements were obtained on a Bruker micrOTOF mass spectrometer equipped with an Agilent 1100 Series LC/MSD mass detector in positive ion mode by direct infusion in water at 100 μL/h using sodium formate clusters as an internal calibrant. ^{1}H and ^{13}C NMR spectra were recorded on either a Varian 400 or Bruker Avance 600 spectrometers at 298 K in the solvents indicated and referenced to residual signals in the deuterated solvents (^{1}H: δ 4.79 for D_2O, $^{1}H/^{13}C$: δ 2.50/39.5 for DMSO-d_6 and δ 3.31/49.0 for CD_3OD). Proton and carbon assignments were determined by various one- and two-dimensional NMR experiments (JMOD, gCOSY, gHSQC, gHMBC, NOESY).

Reaction of 5-A-RU with 2b-c: reaction monitoring and isolation of lumazine products

The solutions of 5-A-RU (50 mM, 30 μL, 1.5 μmol) and dicarbonyl compound glyoxal (**2c**), methylglyoxal (**2d**), or butane-2,3-dione (**2b**) (137 mM solution, 30 μL, 2.7 eq) was added into a mixed solvent containing D_2O (60 μL) and H_2O (480 μL) in that order under argon. The final volume ratio of D_2O–H_2O was 1:9. The mixture was vortexed and stored at 4 °C or rt in the dark and the progress of the reaction monitored for 1–7 days by NMR spectroscopy until completion. The optimal conditions for lumazine products were rt, 7 days for RL-6,7-DiMe (**3b**), 4 ºC, 7 days for RL (**4c**), and 4 ºC, 1 days for RL-7-Me (**4d**). The preparative production was carried out at 0.2 mmol scale with 3 equivalent of dicarbonyl compound under optimized conditions. For DL-glyceraldehyde (3 eq) or 1,3-dihydroxyacetone dimer (8 eq), the reaction was conducted at 80 °C for 30 min under argon in the dark. The crude product was diluted with aqueous 20 mM ammonium acetate buffer (pH 5.4) and purified using preparative rpHPLC. The pure product fractions were combined and lyophilized. The initial powder was re-dissolved in MilliQ water and re-lyophilized with complete removal of residual ammonium acetate to give the product as a yellow or green powder. Compounds were fully

characterised by HRMS, 1D and 2D NMR spectroscopy (**Extended Data Figures 7-8 for NMR**).

RL (4c, Extended Data Figure 7 for NMR). ^{1}H NMR (600 MHz, D_2O–CD_3OD 9:1), δ 3.69 (1H, dd, J = 11.9, 6.8 Hz, 1 × H of ribityl-5′-CH_2), 3.79 (1H, dd, J = 6.6, 5.2 Hz, ribityl-3′-CH), 3.82 (1H, dd, J = 11.9, 3.1 Hz, 1 × H of ribityl-5′-CH_2), 3.91 (1H, td, J = 6.8, 3.1 Hz, ribityl-4′-CH), 4.24–4.31 (2H, m, ribityl-2′-CH and 1 × H of ribityl-1′-CH_2), 5.10 (1H, d, J = 11.3 Hz, 1 × H of ribityl-1′-CH_2), 8.36 (1H, d, J = 3.8 Hz, H6), 8.43 (1H, d, J = 3.8 Hz, H7); ^{13}C NMR (150 MHz, D_2O–CD_3OD 9:1) δ 57.2 (CH_2, ribityl-1′-CH_2), 63.5 (CH_2, ribityl-5′-CH_2), 69.3 (CH, ribityl-2′-CH), 72.9 (CH, ribityl-4′-CH), 74.1 (CH, ribityl-3′-CH), 133.9 (CH, 6), 138.0 (C, 4a), 140.5 (CH, 7), 152.5 (C, 8a), 159.0 (C, 2), 163.3 (C, 4). ESI-HRMS calcd for $C_{11}H_{15}N_4O_6^+$ [M + H]$^+$: 299.0986, found: 299.0983. HPLC t_R = 7.8 min.

RL-7-Me (4d, Extended Data Figure 8 for NMR). ^{1}H NMR (600 MHz, D_2O–CD_3OD 9:1), δ 2.90 (3H, s, 7-CH_3), 3.73 (1H, dd, J = 11.9, 6.7 Hz, 1 × H of ribityl-5′-CH_2), 3.87 (1H, dd, J = 11.9, 2.9 Hz, 1 × H of ribityl-5′-CH_2), 3.89 (1H, dd, J = 7.1, 5.0 Hz, ribityl-3′-CH), 3.95 (1H, td, J = 6.9, 3.0 Hz, ribityl-4′-CH), 4.46 (1H, ddd, J = 9.5, 4.8, 2.5 Hz, ribityl-2′-CH), 4.62 (dd, 1H, J = 13.2, 10.3, 1 × H of ribityl-1′-CH_2), 4.97 (1H, dd, J = 13.2, 1.9 Hz, 1 × H of ribityl-1′-CH_2), 8.36 (1H, s, H6); ^{13}C NMR (150 MHz, D_2O–CD_3OD 9:1) δ 24.2 (CH_3, 7-CH_3), 52.2 (CH_2, ribityl-1′-CH_2), 63.6 (CH_2, ribityl-5′-CH_2), 69.7 (CH, ribityl-2′-CH), 73.0 (CH, ribityl-4′-CH), 74.4 (CH, ribityl-3′-CH), 134.3 (C, 4a), 136.6 (CH, 6), 152.7 (C, 8a), 153.5 (C, 7), 159.0 (C, 2), 163.8 (C, 4). ESI-HRMS calcd for $C_{12}H_{17}N_4O_6^+$ [M + H]$^+$: 313.1143, found: 313.1140. HPLC t_R = 15.0 min.

Supplementary Method 2

Kinetic studies of Lego reaction between 5-A-RU and methylglyoxal

The solution of freshly formed 5-A-RU was chilled, adjusted to pH 7.0 with 1M sodium hydroxide solution, diluted with MilliQ water to make 26 mM stock solutions and stored in 1.5 mL aliquots at –20 ºC for later use. This stock was further diluted to the required concentrations with MilliQ water for lower starting concentrations. A chilled (0 °C) 2 mL glass HPLC vial was purged with argon, to which was added 211 μL of aqueous buffer (10 mM Tris, 150 mM NaCl, pH 8.0 or 20 mM ammonium acetate), and chilled solutions of 5-A-RU (26 mM, 25 μL, 0.65 μmol) and methylglyoxal (136mM, 14.4 μL, 1.96 μmol). Total volume was 250 μL and final concentration of 5-A-RU 2.6 μM. To make other starting concentrations, for example 0.5 mM or 0.1 mM final concentration of 5-A-RU but maintaining the same mole ratio to methylglyoxal, stock solutions of 5-A-RU (5 mM or 1 mM) and methylglyoxal (26.1 mM or 5.22 mM) were prepared and used instead. The solution was vortexed for 5s and incubated at the designated temperatures (0, 15 or 37 ºC). At each time point, 5 μL of the reaction mixture was analyzed using LCMS. LCMS analysis was carried out on a Shimadzu LCMS-2020 system equipped with a Phenomenex Luna 3 micron C18(2) 150 × 2.00 mm column (P/No 00F-4251-BO) and a SPD-M20A diode array detector. Flow rate was 0.4 mL/min with linear gradient: 100% Solvent A over 5 min, followed by 0% to 100% Solvent B over 10 min, 100%B over 5 min and regeneration phase including 100% to 0% Solvent B over 1min and a further 0% over 4 min. Solvent A was 20 mM ammonium acetate in H_2O and solvent B was 20 mM ammonium acetate in MeCN–H_2O (80:20, v/v). Quantification was calculated based on the corresponding peak area against a standard solution, the concentration of which was calibrated using the standard PULCON NMR technique (**Extended Data Figure 3a** for condition at 500 μM, 37 ºC, pH 6.8).