Deciphering CD4+ T cell specificity using novel MHC-TCR chimeric receptors

Journal Article

Author(s):
Kisielow, Jan; Obermair, Franz-Josef; Kopf, Manfred

Deciphering CD4+ T cell specificity using novel MHC–TCR chimeric receptors

Jan Kisielow [ID]*, Franz-Josef Obermair and Manfred Kopf [ID]*

αβ T cell antigen receptors (TCRs) bind complexes of peptide and major histocompatibility complex (pMHC) with low affinity, which poses a considerable challenge for the direct identification of αβ T cell cognate peptides. Here we describe a platform for the discovery of MHC class II epitopes based on the screening of engineered reporter cells expressing novel pMHC–TCR (MCR) hybrid molecules carrying cDNA-derived peptides. This technology identifies natural epitopes of CD4+ T cells in an unbiased and efficient manner and allows detailed analysis of TCR cross-reactivity that provides recognition patterns beyond discrete peptides. We determine the cognate peptides of virus- and tumor-specific T cells in mouse disease models and present a proof of concept for human T cells. Furthermore, we use MCR to identify immunogenic tumor neo-antigens and show that vaccination with a peptide naturally recognized by tumor-infiltrating lymphocytes efficiently protects mice from tumor challenge. Thus, the MCR technology holds promise for basic research and clinical applications, allowing the personalized identification of T cell–specific neo-antigens in patients.

As medicine in affluent societies enters a new era of personalization, most therapies and medical products are being progressively tailored to the individual patient. In the case of immunotherapies, antigenic specificities of T cells come into focus and their unbiased, efficient identification, as well as the assessment of cross-reactivity, becomes of central importance[1].

Current methods to identify T cell epitopes[2–9] follow two major strategies. The first is based on detecting pMHC–TCR interactions by staining T cells with recombinant major histocompatibility complex- (MHC)-tetramers or by selecting phage particles, yeast or insect cells displaying peptide–MHC complexes with recombinant TCRs[4–6,10]. These methods are severely constrained by the binding affinities of the recombinant reagents to their ligands, which can differ substantially (for example, specific staining of $OVA_{323-339}$-reactive OT-II T cells requires 400 times more tetramer than the specific staining of lymphocytic choriomeningitis virus (LCMV) gp_{66-77}reactive SMARTA T cells[11]). Low affinity interactions are therefore easily missed. Furthermore, MHC molecules have to be mutagenized to allow efficient surface expression on phages or yeast cells[10,12]. The second strategy is based on measuring T cell activation in co-cultures with antigen presenting cells (APCs) displaying peptide pools[13] or positional scanning combinatorial peptide libraries[2,3]. These cumbersome methods define peptide motifs that lead to T cell activation[3]. However, many T cells do not react with such peptide pools[14]. Furthermore, the predictive power of these co-culture-based methods is variable[14] because random peptide sequences are interrogated and the identified peptide motifs are often ambiguous[15,16] or do not correspond to natural peptides[17]. Finally, because the diversity of MHC class II presented peptides is several orders of magnitude higher than that of MHC class I, it is particularly difficult to find specific peptides for CD4+ T cells.

To address these problems, we designed a method enabling the discovery of epitopes recognized by CD4+ T cells on the basis of screening of engineered reporter cells expressing libraries of novel pMHC–TCR (MCR) hybrid molecules carrying cDNA-derived peptides. This approach overcomes constraints of currently used methods by providing a large collection of APCs, each presenting a unique peptide, from which cells presenting cognate peptides can be enriched by iterative co-culture with T cells and flow sorting. The MCR libraries are enriched for natural peptides, containing peptides derived from differentially spliced messenger RNAs, but also those encoded by alternative-reading frames, allowing screening of common neo-epitopes such as alternative-reading-frame-derived peptides (ARFPs) stemming from mutated DNA or defective ribosomal products (DRiPs)[18]. This enables unbiased identification of the specificity of individual T cells, or T cell subsets enriched for a particular reactivity. Furthermore, MCR random peptide libraries allow mapping of TCR cross-reactivities, while individual reporter clones with MCRs carrying specific neo-epitopes provide efficient means for finding reactivities in oligo- or polyclonal tumor-infiltrating lymphocyte (TIL) populations. We provide results validating the MCR platform and show its efficiency in unbiased discovery of immunogenic epitopes recognized by anti-viral and tumor-infiltrating CD4+ T cells.

Results

MCR is a novel MHC–TCR hybrid receptor tailored for T cell epitope screening. To efficiently identify cells presenting T cell epitopes we constructed novel chimeric receptors (termed MCRs) by grafting the extracellular domains of the MHC carrying directly tethered peptides[4–6] on top of the TCR truncated at the membrane (see Methods). In contrast to chimeric antigen receptors (CARs)[19], MCRs expressed in TCR− T cell lines assemble with the native CD3 complex and can use the whole signaling machinery of the TCR (Fig. 1a,b). To test their functionality, we designed and cloned MCRs for the screening of cognate peptides of MHC class II–restricted T cells (MCR2). MHC class II has an open peptide-binding groove and two transmembrane chains, allowing easy peptide linkage and grafting on the two TCR transmembrane domains. MCR2-gp61 was generated by linking the extracellular domains of the I-Ab MHC class II α- and β-chains to truncated TCRα and TCRβ chains, respectively, with the LCMVgp_{61-77} peptide tethered to the amino-terminal end of the I-Ab β-chain (Fig. 1a). Similarly, we cloned MCR2-OVA carrying the $OVA_{323-339}$ peptide. After transduction into BEKO

Institute of Molecular Health Sciences, ETH Zürich, Zürich, Switzerland. *e-mail: jan.kisielow@biol.ethz.ch; manfred.kopf@ethz.ch
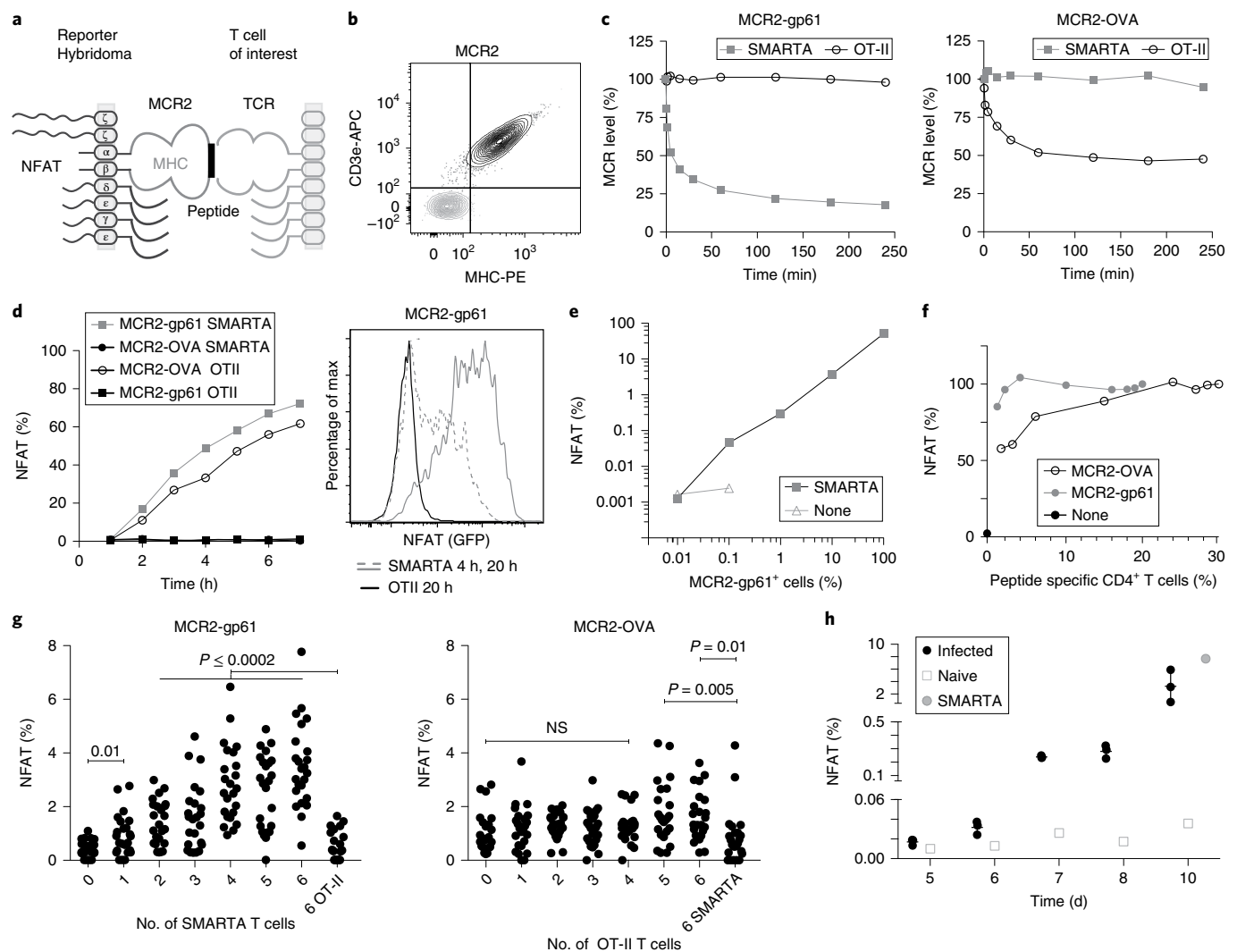
Fig. 1 | Structure and peptide-specific reactivity of the MCR2 sensor. a, Schematic representation of the MCR2 and its interaction with the peptide-specific TCR. **b**, Surface expression of the MCR2. TCR-negative cells were transduced with MCR2 and stained for MHC and CD3e. **c**, Time course of MCR2-gp61(left panel) or MCR2-OVA (right panel) down-regulation in BEKO cells co-cultured with gp61-specific SMARTA or OVA-specific OT-II TCR transgenic CD4+ T cells. Values show MCR2 levels depicted as percentages of mean fluorescence intensity at the start of co-culture. **d**, Time course of peptide-specific NFAT activation (GFP expression) in H18.3.13 cells transduced with MCR2-gp61 or MCR2-OVA co-cultured with SMARTA or OT-II CD4+ T cells. The histogram shows examples of NFAT activation measurements in MCR2-gp61+ cells at indicated time points. **e**, Sensitivity of detection of peptide-specific reporter cells. MCR2-gp61+ and MCR2-OVA+ H18.3.13 cells were mixed at different ratios and NFAT activation was measured in untreated cells (triangles) or after co-cultured with SMARTA CD4+ T cells (squares). The graph shows a linear ($R > 0.99$) correlation between the percentage of H18.3.13 cells with activated NFAT and the percentage of cells carrying the MCR2-gp61. **f**, Minimal frequency of peptide-specific T cells able to trigger robust NFAT activation in MCR2+ reporter cells. Splenocytes from SMARTA and OT-II transgenic mice were mixed at different ratios and used to stimulate MCR2-gp61+ or MCR2-OVA+ H18.3.13 cells. The graph shows percentage of cells with activated NFAT among MCR2-gp61+ H18.3.13 or MCR2-OVA+ H18.3.13 cells as a function of the percentage of peptide-specific CD4+ T cells (determined by FACS), after 8 h of co-culture. **g**, The minimal number of peptide-specific T cells able to trigger a response in reporter cells. NFAT activation in MCR2-gp61+ (left) and MCR2-OVA+ (right) H18.3.13 cells following overnight co-culture with different numbers of SMARTA or OT-II CD4+ T cells ($n = 24$ co-cultures). **h**, Kinetics of LCMV-specific CD4+ T cell expansion in the blood of infected mice ($n = 3$) detected by the activation of MCR2-gp61+ H18.3.13 cells. Percentage of reporter cells with activated NFAT after overnight co-culture with CD4 T cells from blood taken at different days post infection is shown, mean ± s.d. was used. Blood from a naive SMARTA mouse was used as a positive control. Data are representative of one (**e,g,h**), two (**d,f**) or three (**b,c**) independent experiments. Student's *t*-test (two-tailed, unpaired) was used to compare the groups, *P* values are indicated: NS, not significant.

thymoma cells lacking MHC class II and TCR, we found MCR2-gp61 and MCR2-OVA expressed efficiently on the cell surface, indicating correct assembly with the CD3 components of the TCR complex. To verify their specificity, we co-cultured MCR2-gp61+ or MCR2-OVA+ cells with purified gp61-specific SMARTA[20] or OVA-specific OT-II CD4+ T cells. We observed a very fast, peptide-specific down-regulation of MCR2 from the cell surface with kinetics identical to

conventional TCRs[21]. MCR2-gp61 was downregulated in co-cultures with SMARTA T cells, but not in the presence of OT-II T cells (Fig. 1c). Likewise, MCR2-OVA was downregulated in co-cultures with OT-II T cells, but not with SMARTA T cells (Fig. 1c). We further assessed the ability of the MCR2 to trigger nuclear factor of activated T cells (NFAT) activation by transducing it into a TCR-deficient T cell hybridoma carrying the NFAT–green fluorescent

protein (GFP) reporter (H18.3.13). Again, NFAT was only triggered when MCR2-carrying hybridomas were co-cultured with peptide-specific T cells (Fig. 1d and Supplementary Fig. 1a). The response was robust and easily measurable already after 2 h (Fig. 1d). Proof-of-concept experiments for I-A^d and I-E^k gave similar results (data not shown).

We tested the sensitivity of the MCR system by mixing MCR2-gp61$^+$ and MCR2-OVA$^+$ reporter cells at different ratios and measuring NFAT activation after co-culture with SMARTA (Fig. 1e) or OT-II (Supplementary Fig. 1b) CD4$^+$ T cells. Due to rare, spontaneous activation of the NFAT-GFP in reporter cells, direct detection of cells carrying MCR2-gp61 was only possible when their frequency was above 1/10,000 (Fig. 1e). However, the linear correlation between the percentage of GFP-expressing cells and the percentage of MCR2-gp61$^+$ cells indicates that MCR2-gp61$^+$ cells present at frequencies lower than 1/10,000 should be enriched among the GFP-expressing cells. Purifying such rare cells requires iterative rounds of co-culture and sorting of GFP-expressing cells. Because tissue biopsies usually include many different cell types, we also determined the lowest frequency of peptide-specific T cells in a heterogeneous population that was able to trigger robust NFAT activation in MCR2$^+$ cells. Splenocytes from SMARTA and OT-II mice were mixed at different ratios and used to stimulate MCR2-gp61$^+$ or MCR2-OVA$^+$ cells (Fig. 1f). The presence of only 1% peptide-specific CD4$^+$ T cells in the mixture triggered 50% of the maximal NFAT activation in MCR2$^+$ cells. A single SMARTA CD4$^+$ T cell was sufficient to trigger significant NFAT activation in MCR2-gp61$^+$ cells, while five OT-II cells were required for a measurable NFAT-GFP signal in MCR2-OVA$^+$ cells (Fig. 1g), probably due to a lower interaction affinity of the latter.

These results demonstrate a great sensitivity and specificity of the MCR technology and indicate that it could be used as a sensitive diagnostic tool for monitoring T cell reactivities in the blood taken from patients. Indeed, using MCR2-gp61$^+$ reporter cells, we were able to efficiently track gp61-specific CD4$^+$ T cell expansion in the blood of LCMV-infected mice (Fig. 1h).

However, the sensitivity of the MCR system could potentially be a problem, as every TCR of a T cell leaving the thymus has a basal affinity toward autologous pMHC complexes, selected for during positive selection[22] and important for peripheral survival[23]—referred to as 'TCR tickling'. To test whether such, very low affinity, non-activating, interactions would induce undesired MCR signals, we generated T cell hybridomas (ex.H12) from single-peptide mice[24]. All the T cells in these mice are positively selected by the same peptide (Ep) and therefore recognize the Ep peptide with very low affinity, but co-culture with MCR2-Ep$^+$ cells did not induce MCR downregulation nor NFAT activation (Supplementary Fig. 1d,e). We also generated hybridomas from autoreactive 2D2 T cells that recognize the neurofilament medium chain (NFM$_{18-30}$)[25] with an activating, but still relatively low affinity that is 80 times lower than that of the SMARTA TCR–I-A^b–gp61 interaction[26]. The 2D2 TCR–I-A^b–NFM$_{18-30}$ interaction was easily detected (Supplementary Fig. 1f). Thus, the MCR system detects activating pMHC–TCR interactions of varying affinities irrespective of whether a foreign antigen (gp61, OVA) or an autoantigen (NFM) is recognized and is not picking up noise caused by very low affinity interactions with positively selecting or 'tickling' ligands.

Screening for gp61 mimotopes with different functional properties. To identify rare, specific peptides in a complex library, multiple iterative cycles of co-culture and sorting of activated reporter cells are necessary (Fig. 2). Because disappearance of NFAT-reporter signals in the sorted cells is required for efficient detection of activation in subsequent rounds of stimulation, we replaced the very stable GFP with a slow fluorescent timer (FT)[27,28]. This mCherry mutant changes 'color' with time, enabling us to distinguish recent

NFAT activation (blue-FT; referred to as 'NFAT activation' throughout this manuscript) from past NFAT activation (red-FT) and therefore allowing better background discrimination and much shorter intervals between subsequent rounds of stimulation. The NFAT-FT reporter was as sensitive as NFAT-GFP (Supplementary Fig. 1c).

As the first test of the MCR platform, we searched for mimotopes of gp61. The MCR2-gp61-RSS library composed of unique mutants of gp61 was generated by randomizing amino-acid residues through RAG-mediated rearrangement (Fig. 3a,b and Methods). After transducing this library into cells carrying the NFAT-FT reporter (16.2c11), we sorted MCR2-gp61-RSS$^+$ cells and co-cultured them with gp61-specific or OVA-specific T cell hybridomas derived from SMARTA or OT-II CD4$^+$ T cells, respectively (Fig. 3c). Around 10% of the cells showed NFAT activation when co-cultured with the SMARTA hybridoma. These cells were sorted as single cells, expanded and re-screened (Fig. 3d and data not shown). Sequencing of the peptides carried by their MCRs revealed two new mimotopes in addition to the original gp61 peptide (Fig. 3e). The lysine at position 9 of gp61 was mutated to serine (gp61S) or to asparagine (gp61N), indicating that it is not absolutely required for SMARTA T cell activation. However, the level of NFAT activation in reporter cells indicated that SMARTA TCR binds the new mimotopes (in particular gp61S) with lower affinity (Fig. 3d). Indeed, in co-cultures of SMARTA T cells with dendritic cells, the mimotopes induced quantitatively and qualitatively different effector CD4$^+$ T cell responses. Whereas gp61N, like the original peptide, induced strong proliferation and predominantly interferon-γ (IFN-γ)-producing (T$_H$1) cells, gp61S was less efficient in driving proliferation and induced a strong IL-4 (T$_H$2) response (Fig. 3f), consistent with earlier reports showing that low avidity signals induced predominantly T$_H$2 polarization[29]. Thus, the MCR platform allows efficient searches for mimotopes with different functional properties.

Screening for viral epitopes recognized by mouse T cells. To validate the MCR technology outlined in Fig. 2 in a setting of infectious disease, we screened for influenza virus epitopes recognized by CD4$^+$ T cells of infected C57BL/6 mice. We generated MCR$^-$flu$^+$ 16.2c11 reporter cells carrying a complex MCR2 library made directly from viral RNA and containing all possible overlapping peptides of the influenza virus. This cDNA-derived peptide (CDP) library (Fig. 2 and Supplementary Fig. 2c,d) has an important advantage over random or combinatorial peptide libraries, because a high frequency (around 1/6) of the cloned cDNA fragments encode peptides that exist in proteins in vivo (see Methods). MCR2-flu$^+$ reporter cells were then co-cultured with freshly generated hybridomas made from influenza virus–reactive T cells isolated from the lungs of infected mice. To scale-up the process, in the first round we used a mixture of eight hybridomas. In subsequent rounds of co-culture, we used the hybridomas individually and eventually determined the peptides carried by the MCRs of responding reporter cells. The peptides varied in length substantially (from 16 to 33 amino acids), but all incorporated known epitopes from three different influenza virus genes (Fig. 4a, boxed). The NP peptide 2 (NP2) recognized by the hybridoma C2 incorporated two known epitopes: NP264 and NP282. We therefore co-cultured this hybridoma with 16.2c11 reporters carrying the NP2 or NP1 peptide, and compared its reactivity to the NP1-reactive hybridoma H1 (Fig. 4b). Both hybridomas recognized both peptides equally well, indicating that they recognize the epitope NP264 and that very long loops connecting the epitope to the MCR (29 amino acids in the case of NP2) do not disturb MCR2–TCR interactions. Finally, in a similar screening for LCMV-derived T cell antigens, we found a new CD4$^+$ T cell epitope (Supplementary Figs. 2 and 3). These data validate the MCR system for studying T cell specificities in mouse disease models.
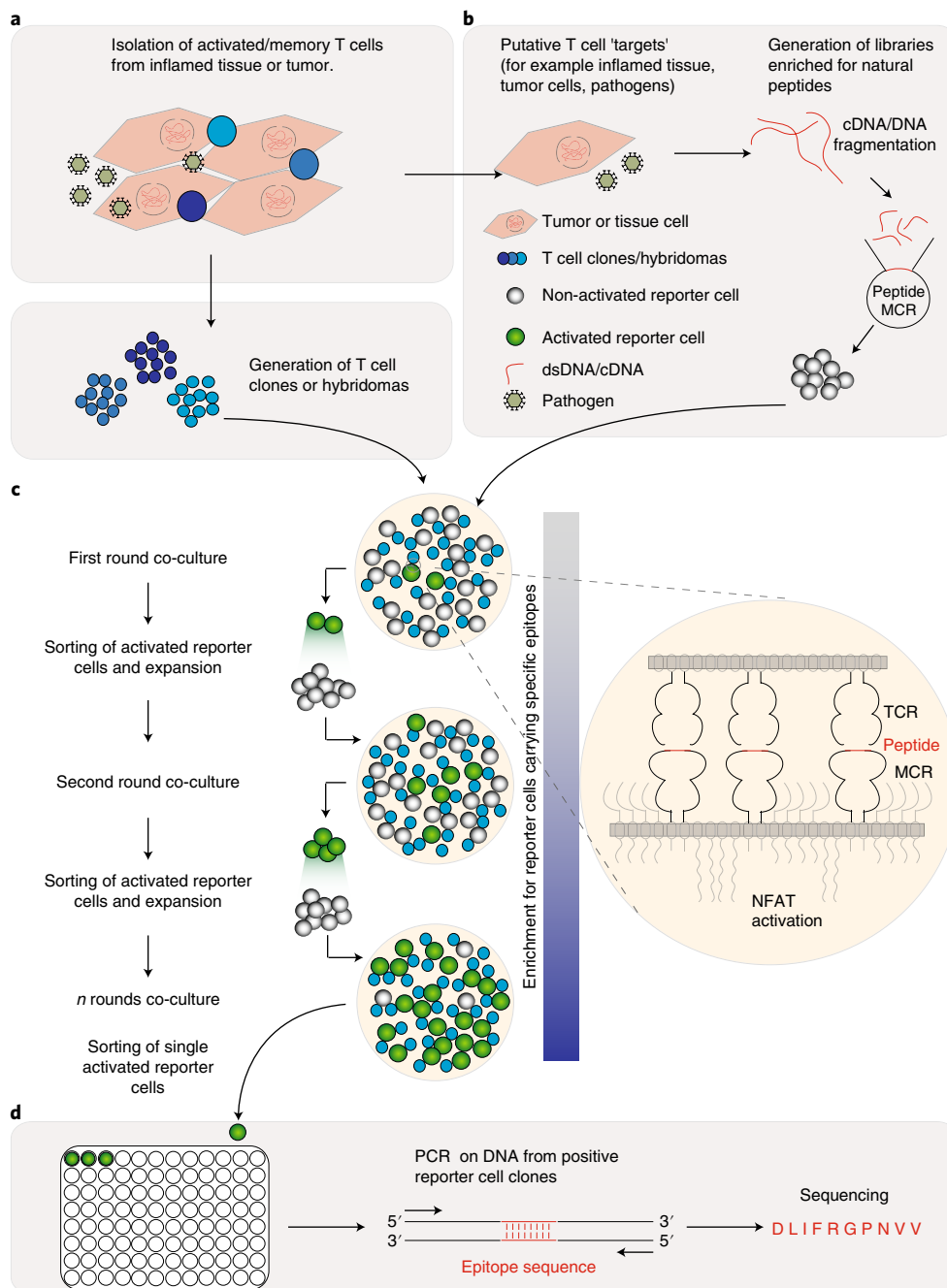
Fig. 2 | The MCR method in a nutshell. a,b, Step 1. Activated and memory T cells are isolated from an affected tissue (infected, autoimmune or tumor-bearing) and used to generate clones or hybridomas (**a**). **b**, In parallel, cDNA or DNA is prepared from the target tissue, tumor or the infecting pathogen, fragmented and cloned into the peptide-encoding part of the MCR vector. In the resulting library, every sixth peptide-MCR molecule presents a natural peptide from the organ, cell or pathogen of interest. This MCR CDP library is retrovirally transduced into an NFAT-reporter cell line. dsDNA, double-stranded DNA. **c**, Step 2. MCR-CDP-expressing reporter cells are co-cultured with the T cell clone of interest and activated cells are sorted by FACS, expanded and co-cultured again with the same T cell clone. Due to the low frequency of target peptides in the library, initially the number of specifically activated cells is not big enough to be distinguished from the spontaneously activated cells (background), they are nevertheless enriched in the sorted population. Iterative co-cultures are continued until a clear population of activated cells becomes apparent and can be distinguished from the background. **d**, Step 3. Single activated reporter cells are sorted, expanded and DNA is prepared. Finally, the peptide-encoding part of the integrated MCR construct is polymerase chain reaction (PCR) amplified from the isolated DNA and sequenced, revealing the T cell–specific epitope.

T cells reactive to the same epitope show diverse antigenic cross-reactivity. We then used the NP264reactive hybridomas H1 and C2 to screen an MCR2 library carrying over 5×10^6 peptides with randomized TCR-contact residues (MCR2–8aaSR, see Methods). We isolated many reactive reporter clones for each of the hybridomas and tested them criss-cross (Fig. 4c). Even though both hybridomas

were monoclonal, as confirmed by TCR β-chain variable region (V$_\beta$ chain) flow cytometry analysis and sequencing (Supplementary Fig. 5a,b) and shared the same peptide specificity (that is, NP264), out of 166 H1-reactive reporters only three (carrying the G1 peptide sequence) were also recognized by the C2 hybridoma and none of the 145 C2-reactive reporters were recognized by H1 (Fig. 4d).
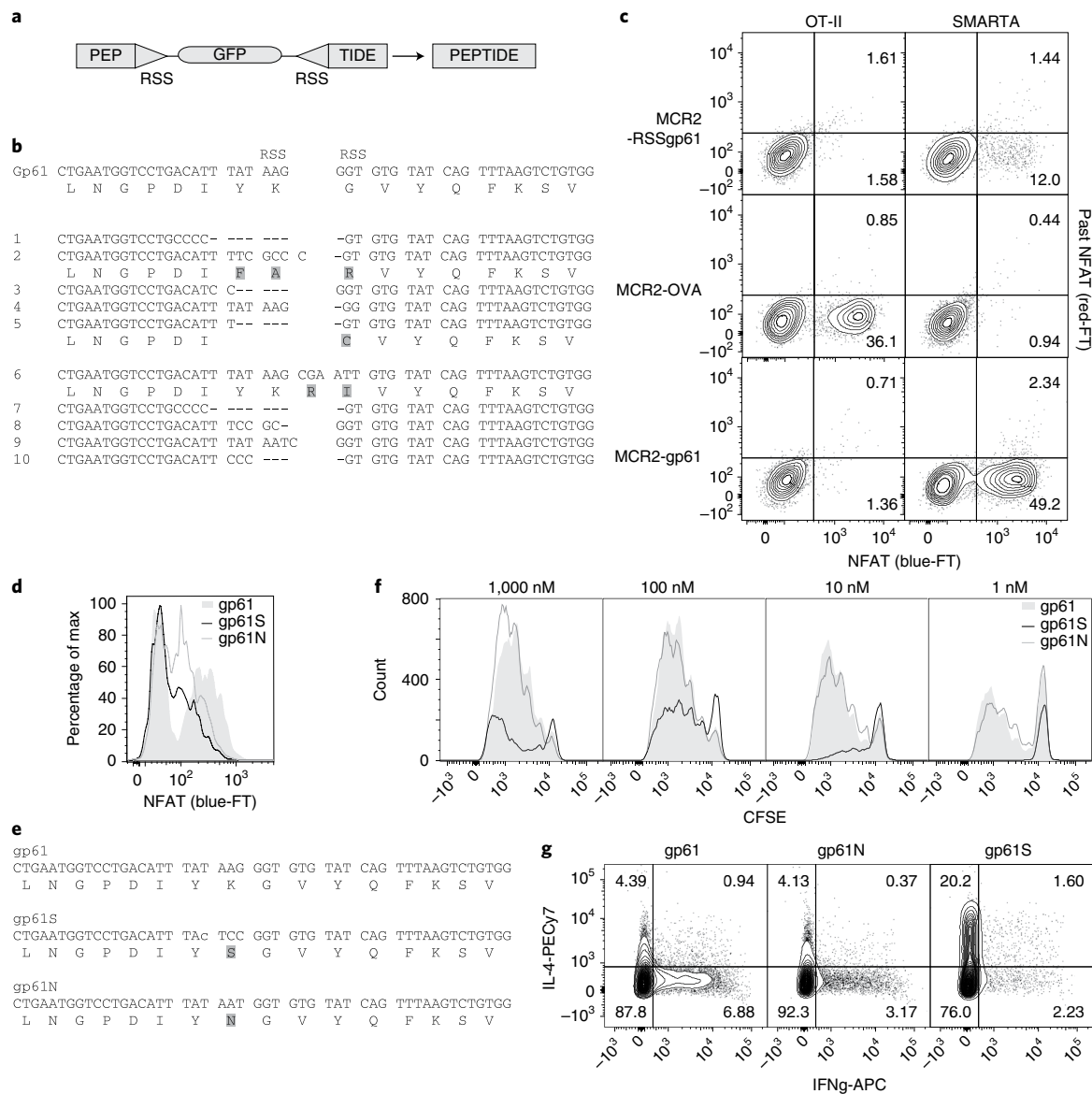
Fig. 3 | Screening for gp61 mimotopes. a, A scheme of RAG-mediated peptide randomization, on the basis of which the MCR2-gp61-RSS mimotope library was generated (see Methods). RSS, recombination signal sequence. **b**, Examples of altered gp61 sequences found in the library (new amino acids are shown in gray). **c**, The 16.2c11 cells carrying the NFAT-FT reporter were transduced with the MCR2-gp61-RSS library or with MCR2-OVA or MCR2-gp61 as controls and co-cultured with SMARTA or OT-II CD4+ T cells hybridomas (ratio 1/1, all cells shown). Single MCR2-gp61-RSS+ cells showing NFAT activation (blue-FT) after 9 h co-culture with the gp61-specific SMARTA hybridoma (top right panel) were sorted and expanded. **d**, Activation of MCR2-gp61+, MCR2-gp61S+ and MCR2-gp61N+ 16.2c11 cells in co-cultures with the SMARTA hybridoma (ratio 1/1). **e**, Sequences of the original gp61 and two new gp61 mimotopes found in the MCR2-gp61-RSS library. **f,g** Dilution of the division-tracking dye CFSE (carboxyfluorescein succinimidyl ester) in SMARTA CD4+ T cells after a 3-day co-culture with dendritic cells pulsed with gp61 mimotopes at different concentrations (**f**) and cytokine production at a concentration of 100 nM (**g**). Numbers in plots indicate percent cells in each quadrant or gated area. Data are representative of one (**c**) or two (**d**,**f**,**g**) independent experiments.

Sequencing of the MCRs from several candidate clones revealed that the peptides they carried were very diverse and shared little homology with NP264 (Fig. 4d). Common residues were present (H7 or K8), but motifs derived from the consensus sequences were very different and neither would allow unambiguous identification of influenza virus NP264 as the antigen of the H1 and C2 hybridomas. Furthermore, homology-based analysis done with Clustal (Supplementary Fig. 4) showed that while most H1- and C2-reactive peptides formed separate clusters, they were distant from the NP264 peptide, indicating that random peptide library screens coupled with bioinformatics approaches would probably fail to correctly predict the native T cell reactivity. Given the diversity of

the discovered epitopes, it was important to verify that they are also recognized as peptides presented by bone marrow derived dendritic cells (BMDCs). Co-culture of H1 and C2 hybridomas with BMDCs pulsed with synthetic peptides confirmed that (Supplementary Fig. 5c), but at high concentrations, some of the peptides showing no reactivity in the MCR screen, stimulated a hybridoma anyhow (for example, H1-A1 stimulated the C2 hybridoma). This indicates possible false-negative results in the MCR assay. However, compared to peptides that scored positive in the MCR assay, 10- to 1,000-fold higher concentrations were required for the same response (Supplementary Fig. 5d); therefore, small impurities in the synthetic peptides (90% pure) could also be the explanation. All peptides in
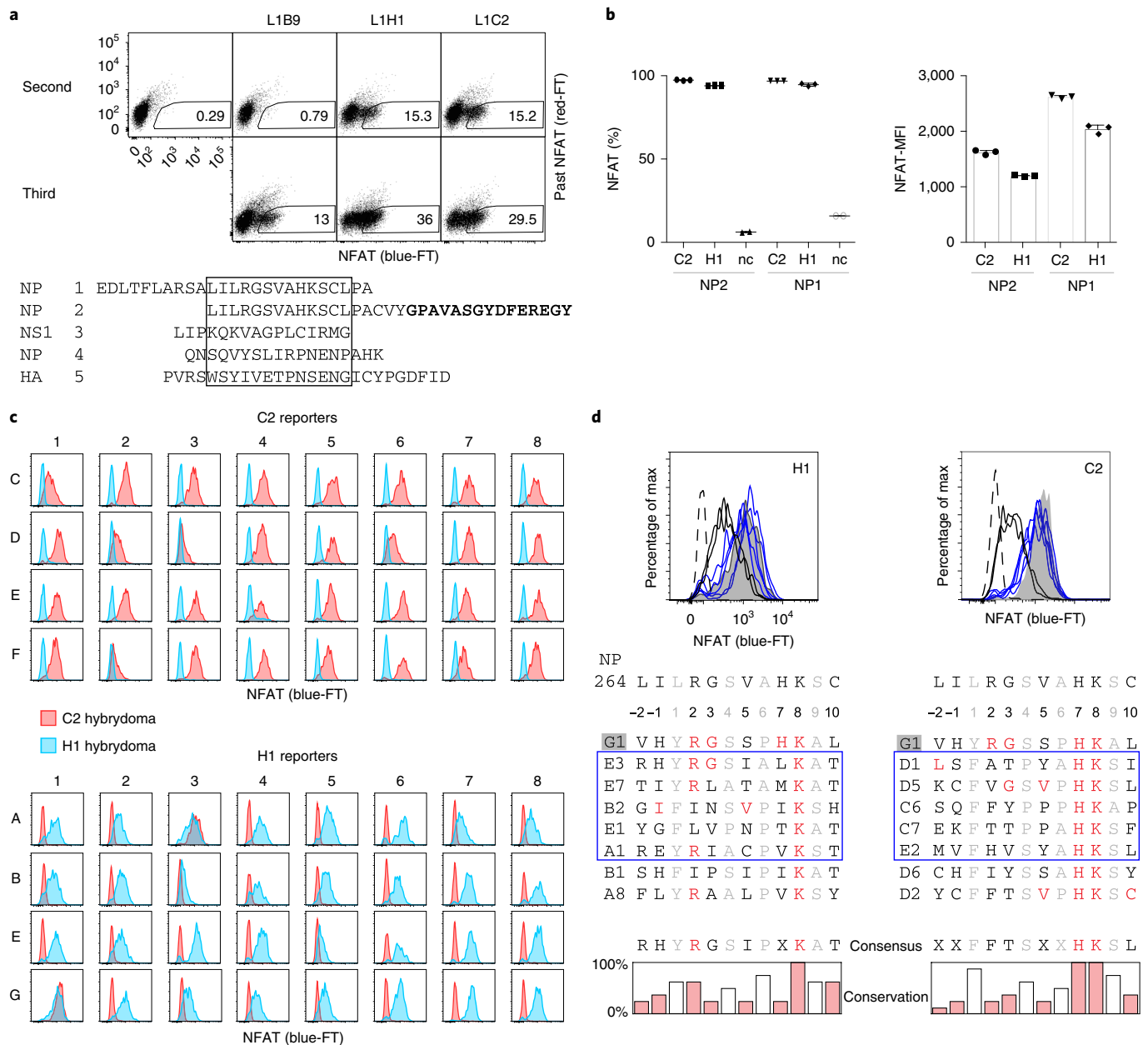
```
NP  1   EDLTFLARSALILRGSVAHKSCLPA
NP  2          LILRGSVAHKSCLPACVYGPAVASGYDFEREGY
NS1 3      LIPKQKVAGPLCIRMG
NP  4          QNSQVYSLIRPNENPAHK
HA  5      PVRSWSYIVETPNSENGICYPGDFID
```


Fig. 4 | Search for viral epitopes. a, MCR2-flu⁺ 16.2c11 reporter cells were co-cultured with eight flu-reactive CD4⁺ hybridomas of unknown peptide specificity. Dot plots show example NFAT activation after two or three rounds of co-culture. Shown below are the different peptide sequences found in this screen, which incorporate five known influenza virus epitopes (boxed and bold). **b**, The fraction responding and MFI of reporters carrying NP1 or NP2 co-cultured with hybridomas H1 and C2 is shown ($n = 3$ or $n = 2$ co-cultures), mean ± s.d. was used. nc, no co-culture. **c**, Histograms show NFAT response of 32 (of 145) individual C2- and 32 (of 166) H1-reactive reporter clones from the fourth round of screening of the MCR2–8aaSR library co-cultured with the C1 or H1 hybridomas. **d**, Comparison of the NFAT activation in example reporter clones (gray histogram represents the response of the G1 reporter clone, dashed histogram negative control cultures). Peptide sequences recovered from these reporters are shown; those with reactivity matching the G1 reporter are boxed. TCR facing amino acids homologous to NP264 are shown in red, anchor residues in gray. Numbers indicate position in the MHC peptide-binding groove. Data are representative of one (**a,c,d**) or two (**b**) independent experiments.

this experiment were selected for recognition by related TCRs (H1 and C2 recognize the same NP264 epitope), therefore amino-acid substitutions or length variations could turn any of these peptides into one being recognized by both hybridomas. Still, direct linking of peptides to the MHC might cause conformational constrains leading to false-negative results in the MCR assay. Notably, in CDP libraries, random sheering of cDNA allows presentation of epitopes in several different registers partially circumventing this problem.

Together, these results highlight the big advantage of CDP libraries over random peptide libraries and bioinformatics for the

discovery of native T cell epitopes. Nevertheless, screening random or semi-random MCR2 libraries allows mapping of cross-reactivity and delivers many unrelated yet highly T cell-reactive peptides, which could be interesting for vaccine development.

Proof of concept for human T cells. To test the MCR technology for the identification of human T cell specificities, we screened an influenza virus full-genome library generated in a vector containing the HLA-DRB1*4 (hMCR2-DRB1*4-flu) with the human CD4 T cell clone hTCC11-46 specific for the influenza virus matrix
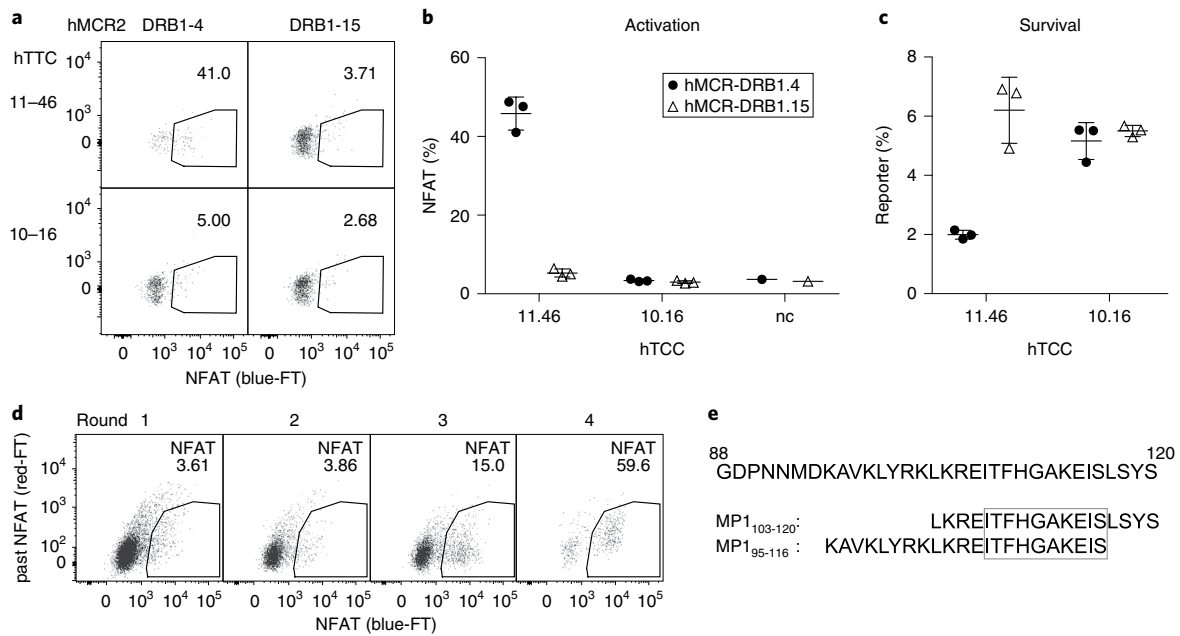
Fig. 5 | Proof of concept for human T cells. a–c, hMCR2-DRB1.4-MP1$_{103-120}$$^+$ and hMCR2-DRB1.15-MP1$_{103-120}$$^+$ 16.2c11 reporter cells were co-cultured ($n=3$ co-cultures) with two human CD4$^+$ T cell clones (hTCC11-46 and hTCC10-16) at a ratio of 1/20. NFAT activation (blue-FT) (**b**) and the proportion of the reporter cells (**c**) were measured, mean ± s.d. is shown. **d**, The 16.2X reporter cells were transduced with the hMCR2-DRB1.4-flu library and used for screening with hTCC11-46. Dot plots depict the NFAT activation analysis after the subsequent rounds of enrichment. **e**, The sequences of influenza virus matrix protein-1 amino acids 88 to 120, MP1$_{103-120}$ positive control peptide and the MP1$_{95-116}$ peptide found in the hTCC11-46 screen, are shown. Data are representative of one (**d,e**) or three (**a–c**) independent experiments.

protein-1–derived core epitope (MP1$_{107-117}$)[13]. The hTCC10-16 T cell clone with a different specificity was used as a negative control. We generated positive and negative controls (that is, hMCR2-DRB1*4-MP1$_{103-120}$$^+$ and hMCR2-DRB1*15-MP1$_{103-120}$$^+$ 16.2c11 cells) and showed that hTCC11-46 indeed recognized MP1$_{103-120}$ in the context of HLA-DRB1*4, but not HLA-DRB1*15 (Fig. 5a,b). However, at the same time it killed the reporter cells in an antigen-specific manner, as observed by the decrease of the fraction of viable reporter cells after co-culture (Fig. 5c). Killing of the reporter cells could be inhibited by adding concanamycin-A (CMA)[30], but this inhibitor of the vacuolar ATPase H$^+$ Transporting V0 Subunit C was toxic for reporter cells and caused their death within 48 h (data not shown). To solve this problem, we generated a library of mutant ATPases by PCR-mediated mutagenesis, over-expressed it in 16.2c11 cells and selected a CMA-resistant 16.2c11 reporter line (termed here '16.2X', see Methods). We used it for screening of the hMCR2-DRB1*4-flu library (Fig. 5d) and found that all the selected clones carried an MP1$_{95-116}$ peptide comprising the hTCC11-46-specific core epitope (Fig. 5e). These results validate the MCR technology for use with human CD4 T cells. In a clinical setting all patients HLA alleles need to be tested, we therefore started building a collection of human MCRs, which at this point includes over 20 functional DRB, DQB and DPB alleles (data not shown).

Probing the reactivity of tumor-infiltrating T cells to identify immunogenic neo-antigens. CD4$^+$ TILs can recognize neo-antigens arising from somatic mutations present in tumors[1,31]. We thought that the MCR platform might be well suited for the identification of such immunogenic epitopes, especially as in this case a simple co-culture with no iterative enrichment would suffice. To test this, we chose a metastasizing mouse breast cancer cell line (E0771), often used in studies searching for effective cancer therapies[32]. Whole-exome sequencing of E0771 cells revealed 627 somatic mutations, from which 300 non-synonymous single-nucleotide polymorphisms (SNPs) were cloned as 23-amino-acid-long

peptides in an MCR2-E0771SNP library. Eighty-one single reporter clones from this library were then co-cultured with polyclonal CD4$^+$ TILs recovered directly from E0771-derived cancer tissue or CD4$^+$ splenocytes. Several reporter clones were specifically recognized by TILs (Fig. 6a,b). Sequencing of the MCRs carried by these clones revealed SNP-carrying peptides derived from several different genes (Fig. 6c), indicating that these SNPs are frequently recognized by TILs in the E0771-derived tumors. The MCR technology allowed direct detection of neo-antigen recognition in freshly isolated TILs, whereas other studies relied on extensive enrichment of specific TILs through in vitro culture[1]. Finally, to test the limits of detection in the human system we mixed hTCC11-46 with hTCC10-16 at different ratios and measured NFAT activation in hMCR2-DRB1*4-MP1$_{103-120}$$^+$ reporter cells after 8 h of co-culture (Supplementary Fig. 6). MP1$_{103-120}$-specific hTCC11-46 cells present at frequencies above 1/2,000 triggered significant responses over background, indicating that most neo-antigen-reactive TILs present in tumor biopsies from patients should be efficiently detected, as they are often found at frequencies higher than 0.05%[33].

Unbiased discovery of tumor-specific antigens recognized by T cells. To identify immunogenic tumor-specific antigens in an unbiased manner, we isolated CD4$^+$ TILs from E0771-derived solid tumors by cell sorting. TIL hybridomas (TILomas) were made and used to screen a CDP library generated from fragmented E0771 cDNA (the 'MCR2-E0771' library). This CDP library comprised over 10^7 individual clones of an insert size of less than 150 base pairs, encoding peptides of 50 amino acids or less that should cover most of the mRNA-derived peptides (for calculation, see Methods). Expression levels of the antigen defines the frequency of peptides in the MCR2 library and consequently how many rounds of co-culture need to be performed; for example, detection of a peptide present at a frequency of 10^{-6} required six rounds of enrichment (Supplementary Fig. 7). Thus, after 7–8 rounds, even antigens expressed as low as one mRNA copy per cell should be detected.
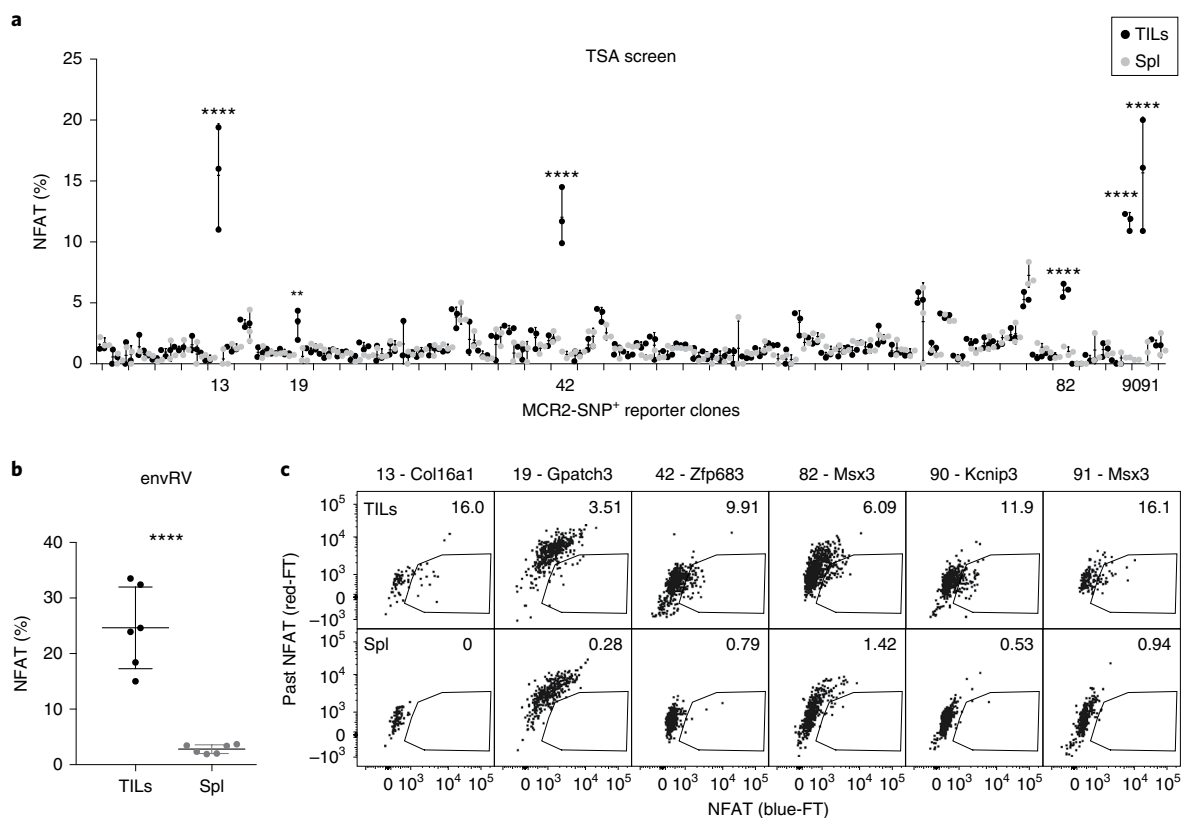
Fig. 6 | Probing TIL reactivity to identify immunogenic neo-antigens. NFAT activation of MCR2-SNP[+] reporter clones carrying individual neo-antigenic peptides derived from SNPs discovered in the E0771 cell line, co-cultured with CD4[+] TILs or CD4[+] splenocytes ($n = 3$ co-cultures). **a–c**, The graph shows the percentage of NFAT[+] cells (**a,b**) and dot plots (**c**) from the clones specifically recognized by the TILs and mean ± s.d. is shown. Clone number and protein names from which the mutated peptides are derived are shown above the dot plots. Similar results were obtained in two independent experiments. Two-way analysis of variance (ANOVA) with multiple pair-wise comparisons was done. **P = 0.004, ****P < 0.0001.

We transduced the MCR2-E0771 library into TCR[−] 16.2c11 reporter cells and co-cultured these with the freshly generated TILomas. We screened 16 TILomas and purified several reactive reporters for 13 of them. After testing each reporter against every TILoma we discovered that TILoma-1.4 reporter responded specifically, but the other reporters cross-reacted with several TILomas (Fig. 7a). Sequencing of the epitopes revealed that the cross-reactive reporters carried peptides mostly encoded by antisense mRNA, untranslated regions, alternative-reading-frames or other non-coding fragments present in our CDP library (for examples, see Supplementary Table 1). Considering that these epitopes are not derived from natural mouse proteins and some represent ARFPs or DRiPs, they are probably recognized like foreign antigens. Polyclonal CD4[+] T cells from naive C57BL/6 mice also recognized these artificial peptides (Fig. 7b), representing an intriguing example of cross-reactivity present in a relatively high fraction of all T cells. Still, screening of MCR2-E0771 library with TILoma-1.4 yielded specific reporter clones, despite the fact that TILoma-1.4 cross-reacted with some of the artificial peptides (Fig. 7a). Sequencing revealed that all of these reporter clones carried a new tumor-specific antigen: a peptide derived from the murine leukemia virus (MLV) envelope protein (env$_{112–134}$). One copy of MLV is integrated in the C57BL/6J mouse genome and is usually silenced[34]. However, in the E0771 cell line (and E0771-derived tumors), MLVenv carries mutations at residues 126 (S126R) and 127 (D127V) and it is actively transcribed (Fig. 7c). Therefore, the mutated env^{S126R,D127V} peptide, referred to as envRV, represents a novel immunogenic tumor-specific antigen, clearly different from the WT env peptide present in the C57BL/6 mouse genome.

Even though the reactivity with artificial peptides did not prevent the discovery of envRV, to remove them, we purged the MCR2-E0771 library of cross-reactive peptides by co-culture with naive polyclonal T cells from C57BL/6 mice (see Methods) and repeated the screen with eight TILomas. This procedure drastically reduced artificial peptides present in the library, but no additional tumor-specific peptides were found; probably due to the fact that many TILs are bystanders and do not recognize tumor-specific antigens. The envRV peptide, however, was found again and confirmed as the specific antigen for TILoma-1.4.

TILoma-1.4 was much less cross-reactive than the H1 and C2 hybridomas. When we used it to screen the random peptide library, we discovered only reporter clones carrying a peptide whose sequence (GVYRMATPGCSR) resembled very much envRV, the cognate antigen. This indicates that T cell cross-reactivity may vary considerably.

Finally, to verify that envRV is the true antigen of TILoma-1.4, we pulsed BMDCs with the envRV peptide and co-cultured them with TILoma-1.4 carrying an NFAT-GFP reporter. The env peptide served as a control. We confirmed that TILoma-1.4 specifically recognized the envRV peptide (Fig. 7d). However, the peptide had to be added at high concentrations, indicating that envRV does not bind the MHC with high affinity and therefore might not be very efficiently presented. To avoid this problem, we transduced BMDCs with constructs encoding envRV or gp61 peptides tethered to the I-A[b] beta chain and dsRed as a marker. dsRed[+] BMDCs were sorted and co-cultured with the TILoma-1.4. TILoma-1.4 reacted with envRV[+] BMDCs but not with gp61[+] BMDCs (Fig. 7e). We then immunized mice with the transduced BMDCs or with envRV and env peptides in
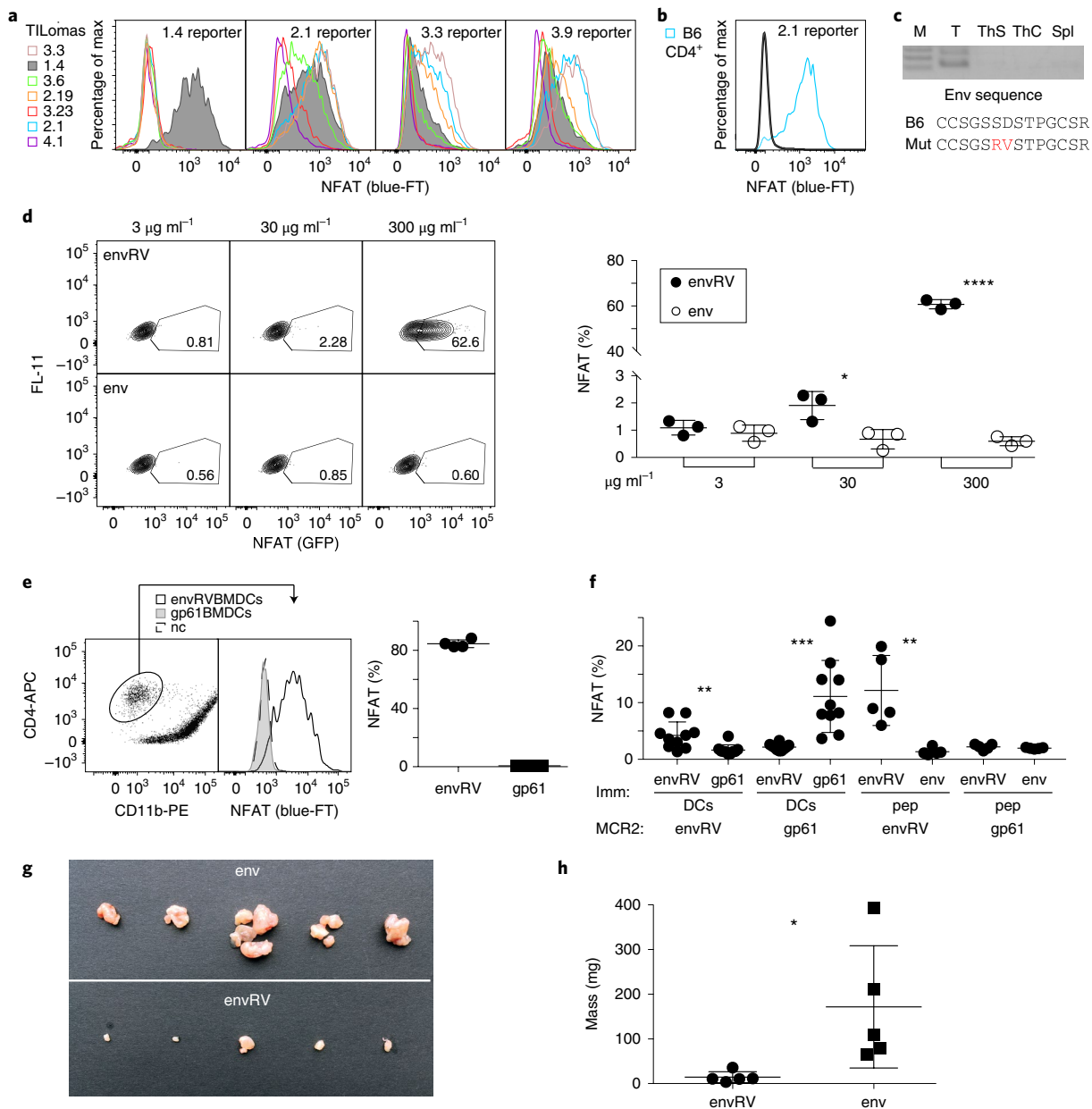
Fig. 7 | Unbiased search for tumor-specific antigens recognized by T cells. MCR2-E0771⁺ reporter cells carrying a CDP library comprising mRNA-derived peptides from the tumor cell line E0771 were screened by co-cultured with freshly generated TILomas. **a,b**, Example NFAT activation (blue-FT) in reporter cells derived from different screenings co-cultured with several different TILomas or B6 CD4⁺ T cells. **c**, Expression of MLVenv mRNA in the tumor cell line E0771 (T), thymic stroma (ThS), thymocytes (ThC) or splenocytes (Spl) tested by RT-PCR. M, DNA size marker. Mutated sequence of the env peptide recovered from the TILoma-1.4 reporters compared to wild type C57BL/6. **d**, TILoma-1.4 carrying the NFAT-GFP reporter was co-cultured (n = 3 co-cultures) with BMDCs pulsed with different amounts of envRV and env peptide. Dot plots and graph show NFAT activation (GFP) in TILomas cells (gated for CD4⁺). **e**, NFAT-GFP TILoma-1.4 was co-cultured with dsRed⁺ BMDCs transduced with a construct encoding envRV or gp61 peptides tethered to the I-Aᵇ β chain and dsRed (envRV-dendritic cells, gp61-dendritic cells). Histograms and graph show NFAT-reporter activation (GFP) in TILoma cells (gated for dsRed⁻). **f**, envRV- or gp61-specific CD4⁺ T cells in the blood of mice immunized with envRV- or gp61-BMDCs (n = 10 mice) or envRV or env peptides in complete Freund's adjuvant (CFA) (n = 5 mice), detected by NFAT (blue-FT) activation in MCR2-envRV⁺ or MCR2-gp61⁺ reporter cells. Percentage of reporter cells with activated NFAT after overnight co-culture with CD4⁺ T cell from blood taken at day 8 post immunization is shown. **g,h**, Mice (n = 5) were challenged with E0771 tumor cell line 9 weeks after a single immunization with peptide/CFA. Shown are the excised tumor tissues (**g**) and the analysis of their weight (**h**). Results are representative of two independent experiments, mean ± s.d. are shown. Student's t-test (two-tailed, unpaired) was used; *P < 0.05, **P ≤ 0.005, ***P = 0.0003, ****P ≤ 0.0001.

CFA. After 8 days we measured the reactivity in the blood of immunized mice with MCR2-envRV⁺ reporter cells and detected envRV-specific T cells (Fig. 7f). Finally, we tested whether vaccinated mice were protected from challenge with the E0771 tumor cells. Indeed, tumor growth was greatly inhibited in mice immunized with envRV

(Fig. 7g,h). These results demonstrate that envRV is the cognate antigen specificity of TILoma-1.4 and that a single immunization with it provides effective protection from tumor challenge.

Taken together, these data provide evidence for the great sensitivity and versatility of the MCR platform, allowing the identification

of T cell–specific peptides even in very complex peptide libraries covering the whole tumor peptidome.

Discussion

In this study, we describe a platform enabling efficient discovery of the peptide specificities of MHC class II–restricted T cells, with the help of the MCR, a novel pMHC–TCR hybrid receptor. CARs composed of various molecules directly linked to the intracellular domain of CD247 (CD3ζ chain) have been successfully used to redirect T cell cytotoxicity towards different targets[19,35–37]. However, CARs containing MHC molecules were shown to be inefficiently phosphorylated and triggered week calcium fluxes[38]. While their function could be improved by addition of signaling domains from molecules such as CD28 or 4–1BB to the ζ-chain[39] or direct linking of Lck[40], an attempt to use them for CD8+ T cell antigen screening was not successful[41]. We reasoned that connecting MHC complexes with directly tethered peptides to the whole TCR signaling machinery should provide more sensitive, physiological signals and minimize false-positives. We therefore built the MCR2 by fusing the peptide and MHC to truncated TCRα and TCRβ chains, enabling assembly with the native CD3 complex. Such molecules turn TCR-deficient T cell hybridomas into artificial APCs where peptide–MHC engagement by the TCRs of specific T cells can be efficiently monitored using an NFAT-reporter system. Co-culture of these cells transduced with a library of MCR2 molecules carrying different peptides with antigen-specific T cells, allows direct and efficient identification of cognate peptide specificities of CD4+ T cells by massively parallel, functional screening in mammalian cells. Optimally, a T cell clone, hybridoma or a cell line expressing the cloned TCR of interest should be used. To increase throughput, several clones can be combined in initial rounds of screening, but the number is constrained by combinatorial limits. While generation of hybridomas is efficient in mice, making human T cell clones is tedious and might be a limitation. Thus, single-cell sequencing of human TILs, TCR cloning and expression in a suitable cell line might be necessary for high-throughput screens.

Several key features of the MCR epitope-screening platform allow efficient, unbiased T cell epitope identification in entire transcriptomes. First, cDNA-derived peptides covalently linked to the MHC part of the MCR2 warrant high density of functional receptors on the cell surface and allow multivalent interaction between the MCRs and TCRs, enabling detection of low affinity ligands. Second, peptides encoded by differentially spliced mRNAs and ARFPs arising from DNA mutations, alternative translation initiation, ribosomal frame shifting or translated untranslated regions, which can generate tumor or autoimmune neo-antigens[42] are efficiently screened. Third, MCR2 uses the complete TCR signaling machinery providing physiological sensitivity and specificity important for finding natural antigens. This feature effectively discriminates pMHC-engaging TCR binding from background TCR binding to cell surface, substantially decreasing false-positive hits compared to the common MHC display techniques. Fourth, we do not rely on functional responses of T cells, so TCRs from anergic T cells can be screened. Fifth, the timer mCherry-variant used in the NFAT reporter enables discrimination of specific signals from the background and efficient iterative screening.

MCR2-CDP libraries made from target tissue, allow screening of all possible peptides encoded by the tissue and the pathogen(s) present in it, independently from endogenous antigen processing, at the same time focusing on real, natural peptides. If the tissue is largely destroyed (for example, pancreatic islets in patients with type 1 diabetes) one can use the RNA or DNA from another source (pancreatic tissue from other patients or post mortem) or, as a last resort, screen random peptide libraries. Nevertheless, antigens derived from peptide splicing, hybrid peptides or some post-translational modifications not performed by our reporter cell line

(for example, citrullination) might not be picked up by the MCR assay. Further genetic modifications of the reporter cell line might circumvent these constrains in the future.

Our results indicate that the MCR technology works very efficiently and provides an advantageous alternative to the currently used methods. Indeed, even though many LCMV T cell epitopes have been described[43], we were able to identify a novel one (that is, NP547). Furthermore, peptides binding MHC with low affinity, such as the identified tumor antigen envRV or epitopes derived from ARFPs, were efficiently detected while they would probably be missed by methods using soluble peptides. The power of the MCR platform also stems from the fact that peptides are screened in the context of native TCR and MHC molecules expressed on the surface of mammalian cells and production of individual recombinant TCRs and MHC tetramers, or mutagenesis of the MHC, are not required. In addition, the large dynamic range of MCR-NFAT signaling allows the design of screening strategies tailored for discovery of high and/or low affinity epitopes. As differences in affinity regulate the quality and quantity of effector and memory T cell responses, the MCR platform has the unique potential to deliver unbiased screening and identification of T cell epitopes together with immune- or disease-relevant information. Furthermore, individual MCR reporter clones, are well suited for testing oligo- or polyclonal T cell populations to find immunogenic SNP-derived neo-epitopes in tumors.

Besides validating the MCR platform, our data provide more evidence for the degeneracy of TCR antigen recognition, by showing that T cells reactive against the same influenza virus epitope can have different antigenic cross-reactivity. These results highlight the ability of T cells to keep potential reactivity toward a plethora of antigens even on epitope-driven expansion.

Taken together, we believe that the MCR platform should be useful for basic research and be adapted for diverse clinical applications. Defining specificities of regulatory and effector tumor-infiltrating T cells might lead to discovery of novel immunogenic tumor antigens. Defining the specificities of autoreactive tissue-infiltrating T cells might help the development of antigen-specific therapies for autoimmune diseases. Furthermore, screening of mimotope libraries should enable the discovery of high affinity peptide variants and the development of sensitive flow cytometry-based tests for antigenic reactivity of CD4+ T cells circulating in the blood of patients.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, statements of data availability and associated accession codes are available at https://doi.org/10.1038/s41590-019-0335-z.

References

1. Linnemann, C. et al. High-throughput epitope discovery reveals frequent recognition of neo-antigens by CD4+ T cells in human melanoma. *Nat. Med.* **21**, 81–85 (2015).
2. Gundlach, B. R. et al. Determination of T cell epitopes with random peptide libraries. *J. Immunol. Methods* **192**, 149–155 (1996).
3. Borras, E. et al. Findings on T cell specificity revealed by synthetic combinatorial libraries. *J. Immunol. Methods* **267**, 79–97 (2002).
4. Dogan, I. et al. Phage-displayed libraries of peptide/major histocompatibility complexes. *Eur. J. Immunol.* **34**, 598–607 (2004).
5. Crawford, F., Huseby, E., White, J., Marrack, P. & Kappler, J. W. Mimotopes for alloreactive and conventional T cells in a peptide–MHC display library. *PLoS Biol.* **2**, e90 (2004).
6. Wen, F., Esteban, O. & Zhao, H. Rapid identification of CD4+ T-cell epitopes using yeast displaying pathogen-derived peptide library. *J. Immunol. Methods* **336**, 37–44 (2008).
7. Siewert, K. et al. Unbiased identification of target antigens of CD8+ T cells with combinatorial libraries coding for short peptides. *Nat. Med.* **18**, 824–828 (2012).

8. Newell, E. W. et al. Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. *Nat. Biotechnol.* **31**, 623–629 (2013).
9. Birnbaum, M. E., Dong, S. & Garcia, K. C. Diversity-oriented approaches for interrogating T-cell receptor repertoire, ligand recognition, and function. *Immunol. Rev.* **250**, 82–101 (2012).
10. Birnbaum, M. E. et al. Deconstructing the peptide-MHC specificity of T cell recognition. *Cell* **157**, 1073–1087 (2014).
11. Mandl, J. N., Monteiro, J. P., Vrisekoop, N. & Germain, R. N. T cell-positive selection uses self-ligand binding strength to optimize repertoire recognition of foreign antigens. *Immunity* **38**, 263–274 (2013).
12. Adams, J. J. et al. Structural interplay between germline interactions and adaptive recognition determines the bandwidth of TCR-peptide-MHC cross-reactivity. *Nat. Immunol.* **17**, 87–94 (2015).
13. Schmid, D., Pypaert, M. & Münz, C. Antigen-loading compartments for major histocompatibility complex class II molecules continuously receive input from autophagosomes. *Immunity* **26**, 79–92 (2007).
14. Sospedra, M. et al. Combining positional scanning peptide libraries, HLA-DR transfectants and bioinformatics to dissect the epitope spectrum of HLA class II cross-restricted CD4+T cell clones. *J. Immunol. Methods* **353**, 93–101 (2010).
15. Judkowski, V. et al. GM-CSF production allows the identification of immunoprevalent antigens recognized by human CD4+ T cells following smallpox vaccination. *PLoS ONE* **6**, e24091 (2011).
16. Wooldridge, L. et al. A single autoimmune T cell receptor recognizes more than a million different peptides. *J. Biol. Chem.* **287**, 1168–1177 (2012).
17. Anderson, B., Park, B. J., Verdaguer, J., Amrani, A. & Santamaria, P. Prevalent CD8+ T cell response against one peptide/MHC complex in autoimmune diabetes. *Proc. Natl Acad. Sci. USA* **96**, 9311–9316 (1999).
18. Wei, J. & Yewdell, J. W. News and views. *Nat. Med.* **23**, 409–410 (2017).
19. Sadelain, M., Rivière, I. & Brentjens, R. J. Targeting tumours with genetically enhanced T lymphocytes. *Nat. Rev. Cancer* **3**, 35–45 (2003).
20. Oxenius, A., Bachmann, M. F., Zinkernagel, R. M. & Hengartner, H. Virus-specific MHC-class II-restricted TCR-transgenic mice: effects on humoral and cellular immune responses after viral infection. *Eur. J. Immunol.* **28**, 390–400 (1998).
21. Valitutti, S., Müller, S., Cella, M., Padovan, E. & Lanzavecchia, A. Serial triggering of many T-cell receptors by a few peptide-MHC complexes. *Nature* **375**, 148–151 (1995).
22. Kisielow, P. & Miazek, A. Positive selection of T cells: rescue from programmed cell death and differentiation require continual engagement of the T cell receptor. *J. Exp. Med.* **181**, 1975–1984 (1995).
23. KirbergJ., Berns, A. & von Boehmer, H. Peripheral T cell survival requires continual ligation of the T cell receptor to major histocompatibility complex–encoded molecules. *J. Exp. Med.* **186**, 1269 (1997).
24. Ignatowicz, L., Kappler, J. & Marrack, P. The repertoire of T cells shaped by a single MHC/peptide ligand. *Cell* **84**, 521–529 (1996).
25. Krishnamoorthy, G. et al. Myelin-specific T cells also recognize neuronal autoantigen in a transgenic mouse model of multiple sclerosis. *Nat. Med.* **15**, 626–632 (2009).
26. Rosenthal, K. M. et al. Low 2-dimensional CD4 T cell receptor affinity for myelin sets in motion delayed response kinetics. *PLoS ONE* **7**, e32562-11 (2012).
27. Subach, F. V. et al. Monomeric fluorescent timers that change color from blue to red report on cellular trafficking. *Nat. Chem. Biol.* **5**, 118–126 (2009).
28. Bending, D. et al. A timer for analyzing temporally dynamic changes in transcription during differentiation in vivo. *J. Cell Biol.* **217**, 2931–2950 (2018).
29. Ruedl, C., Bachmann, M. F. & Kopf, M. The antigen dose determines T helper subset development by regulation of CD40 ligand. *Eur. J. Immunol.* **30**, 2056–2064 (2000).
30. Kataoka, T. et al. Concanamycin A, a powerful tool for characterization and estimation of contribution of perforin- and Fas-based lytic pathways in cell-mediated cytotoxicity. *J. Immunol.* **156**, 3678–3686 (1996).
31. Sahin, U. et al. Personalized RNA mutanome vaccines mobilize poly-specific therapeutic immunity against cancer. *Nature* **547**, 222–226 (2017).
32. Stagg, J. et al. Anti-CD73 antibody therapy inhibits breast tumor growth and metastasis. *Proc. Natl Acad. Sci. USA* **107**, 1547–1552 (2010).
33. Tran, E. et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science* **350**, 1387–1390 (2015).
34. Hayashi, H. et al. Molecular cloning and characterization of the gene encoding mouse melanoma antigen by cDNA library transfection. *J. Immunol.* **149**, 1223–1229 (1992).
35. Brocker, T., Peter, A., Traunecker, A. & Karjalainen, K. New simplified molecular design for functional T cell receptor. *Eur. J. Immunol.* **23**, 1435–1439 (1993).
36. Maher, J., Brentjens, R. J., Gunset, G., Rivière, I. & Sadelain, M. Human T-lymphocyte cytotoxicity and proliferation directed by a single chimeric TCRζ/CD28 receptor. *Nat. Biotechnol.* **20**, 70–75 (2002).
37. Jyothi, M. D., Flavell, R. A. & Geiger, T. L. Targeting autoantigen-specific T cells and suppression of autoimmune encephalomyelitis with receptor-modified T lymphocytes. *Nat. Biotechnol.* **20**, 1215–1220 (2002).
38. Geiger, T. L., Leitenberg, D. & Flavell, R. A. The TCR ζ-chain immunoreceptor tyrosine-based activation motifs are sufficient for the activation and differentiation of primary T lymphocytes. *J. Immunol.* **162**, 5931–5939 (1999).
39. Dotti, G., Gottschalk, S., Savoldo, B. & Brenner, M. K. Design and development of therapies using chimeric antigen receptor-expressing T cells. *Immunol. Rev.* **257**, 107–126 (2013).
40. Geiger, T. L., Nguyen, P., Leitenberg, D. & Flavell, R. A. Integrated src kinase and costimulatory activity enhances signal transduction through single-chain chimeric receptors in T lymphocytes. *Blood* **98**, 2364–2371 (2001).
41. Zhang, T., He, X., Tsang, T. C. & Harris, D. T. SING: a novel strategy for identifying tumor-specific, CTL-recognized tumor antigens. *FEBS. J.* **18**, 600–602 (2004).
42. Kracht, M. J. L. et al. Autoimmunity against a defective ribosomal insulin gene product in type 1 diabetes. *Nat. Med.* **23**, 501–507 (2017).
43. Dow, C. et al. Lymphocytic choriomeningitis virus infection yields overlapping CD4+ and CD8+ T-cell responses. *J. Virol.* **82**, 11734–11741 (2008).

Author contributions

J.K. conceived the study, designed and performed experiments, analyzed data and wrote the manuscript. F.J.O. designed and performed experiments, analyzed and discussed data and reviewed the manuscript. M.K. analyzed and discussed data, corrected and provided critical feedback on the manuscript.

Competing interests

On the basis of this study, a patent application was filed WO2016097334A1. The authors are founders of Tepthera.

Additional information

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Methods

Mice. C57BL/6J, SMARTA and OT-II mice were bred at the ETH Zürich mouse facility (EPIC) under specific pathogen-free conditions. All animal experiments were approved by the local animal ethics committee (Kantonales Veterinärsamt Zürich, licenses ZH061/17), and performed according to local guidelines (TschV) and the Swiss animal protection law (TschG).

hTCC cell culture. hTCCs were expanded in Immunocult-XF T cell expansion medium (STEMCELL) supplemented with hIL-2 (Roche) and Immunocult-XF human CD3/CD28/CD2 T cell activator (STEMCELL).

Immunization with BMDCs or peptides in CFA. Freshly isolated bone marrow cells cultured with GM-CSF for 2 days, were transduced with retroviral supernatants and cultured further for 8 days. Then 10^5 transduced cells were injected intraperitoneally on day 0. For peptide immunization, 100 µg of peptide (Genosphere, purity >90%) in PBS was mixed 1:1 with commercially available CFA (Difco Laboratories) and was injected subcutaneously.

Cell lines. BEKO is a spontaneous thymoma cell line derived from TCRβ-deficient mice. The H18.3.13 NFAT-reporter cell line was generated by retrovirally transducing the NFAT-GFP reporter (carrying four copies of the minimal human IL-2 promoter, each containing three NFAT binding sites ACGCCTTCTGTATGAAACAGTTTTTCCTCC inserted upstream of the enhanced GFP coding sequence) into a TCR− B6 T cell hybridoma. The 16.2c11 reporter cell line was generated by transfecting the 16.2 T cell hybridoma (BALB/c) with the NFAT–slow FT reporter construct and transducing with a vector encoding the murine ecotropic retrovirus receptor Slc7a1. Cells were not tested for Mycoplasma.

Generation of CMA-resistant 16.2c11 cells (16.2X). Mouse ATP6v0c was amplified from cDNA by error-prone PCR (MnCl$_2$-containing buffer and unbalanced dNTPs) and cloned into the pMY vector generating a library of mutants. After transduction into 16.2c11, cell selection with 10 nM CMA was performed. ATP6v0c over-expressed in the surviving cells was cloned and sequenced revealing the ATP6v0c (S45L/M53I) mutant.

Flow cytometry. The following antibodies were used: Fc block (anti-CD16/CD32; 2.4G2; home-made); CD4-APC (GK1.5), CD4-PE (GK1.5), TCRβ-APC (H57–597), MHC class II I-A-PE (M5/114.15.2), CD11b (M1/70), IL-4-PE-Cy7(11B11) and IFNg-APC(XMG1.2). TCR V$_β$ analysis was done with the BD Pharmingen mouse Vβ TCR Screening Panel. Cells were analyzed on FACSCanto II or LSRFortessa (BD Bioscience) and data were analyzed in FlowJo software (Tree Star).

Hybridoma generation. Sorted T cells were activated with plastic-bound anti-CD3ε and anti-CD28 antibodies in the presence of mouse IL-2 for 2–3 days. Equal numbers of activated T cells and the TCRα⁻β⁻ BW5147 fusion partner were fused using PEG-1500, and plated at limiting dilution in the presence of 100 mM hypoxanthine, 400 nM aminopterin and 16 mM thymidine (HAT). In some cases, hybridomas were generated with TCRα⁻β⁻ BW5147 carrying the NFAT-GFP reporter.

T cell hybridoma–BMDC co-culture. Hybridoma cells carrying the NFAT-GFP reporter were co-cultured 8 h with BMDCs pulsed with different amounts of the indicated peptides (Genosphere, purity >90%). GFP expression was measured by FACS.

Cloning of the MCR2. MCR2α and MCR2β were cloned into the pMYsiresGFP retroviral vector, so that MCR2β replaced GFP. Throughout the study, we used this MCR2 vector (pMY-MCR2αiresMCR2β) to generate MCR2s containing various peptides and referred to them as MCR2-peptides. The MCR2 α and β chains contained the following parts: MCR2 α chain: the MHC-II I-A^b α chain residues 1–208 linked to the TCRα chain constant region residues 87–137 (consisting of the hinge region, transmembrane (TM) and intracellular (IC) domains) by the GGSGGSAQ linker. MCR2 β chain: the MHC-II I-A^b β chain residues 1–217 linked to the TCRβ chain constant region (C1) residues 123–173 by the AQSGGSGGSAQ linker. In the MCR2-gp61 residues DS at positions 29 and 30 of the MHC-II part were replaced by the amino-acid sequence SGLNGPDIYKSVGSGGSGGSGDS (containing the gp61 peptide). In MCR2-OVA the same residues were replaced by the amino-acid sequence SISQAVHAAHAEINEAGRGSGGSGGSGDS (containing the OVA peptide). Other candidate peptide sequences and the linker GSGGSGGSGDS were cloned between the GD residues at positions 28 and 29 of the MHC-II I-A^b β chain and inserted into the MCR2 vector.

Human MCR2s were built by replacing the mouse MHC sequences with human HLA DRA and DRB1.4 or DRB1.15.

Retroviral transduction of reporter cell lines and sorted thymocytes. Retrovirus containing supernatants were produced in the ecotropic Phoenix packaging cell line and used to infect reporter cell lines and sorted cells.

RAG-mediated generation of mimotope libraries. To generate the gp61 mimotope library, the MCR2-gp61-RSS-GFP-RSS construct was built by inserting a stuffer fragment containing GFP and the RAG RSSs into the middle of the gp61 peptide in the MCR2-gp61 construct (Fig. 3a). This construct was transduced into sorted CD4−CD8− double-negative thymocytes cultured on Tst-4/DLL1[44]. After a week of culture (during which double-negative cells develop into CD4+CD8+ double-positive cells and recombine their TCR genes, as well as the RSS-GFP-RSS stuffer, by RAG-mediated rearrangement) cells were sorted and cDNA was made. The peptide-encoding part of the recombined MCR2-gp61-RSS-GFP-RSS construct was PCR amplified from cDNA and cloned into the empty MCR2 vector, generating the MCR2(gp61-RSS) mimotope library.

CDP and semi-random peptide library generation. To generate the MCR2-LCMV, MCR2-flu, hMCR2-flu and MCR2-EO771 CDP libraries, cDNA encoding the LCMV proteins GP and NP or influenza virus– or EO771-derived double-stranded cDNA was digested for a limited amount of time (Takara DNA fragmentation Kit). The fragments were ligated with 200 bp linkers homologous to vector sequences flanking a BamHI cloning site (present at the amino acid 28, between the sequences encoding the leader and the GS linker) and PCR amplified. Fragments encoding 50 amino acids or less were purified by gel electrophoresis, cloned into the pMY-MCR2-I-A^b (mouse) or pMY-hMCR2-DRB1.4(human) vector by Gibson assembly and transfected into bacteria generating over 2×10^6 clones for MCR2-LCMV, 10^6 clones for MCR2-flu and 3×10^7 for MCR2-EO771. Assuming 3×10^5 mRNA molecules per cell and an average mRNA length of 1,000 bp, the transcriptome encoded 3×10^8 possible peptides. However, as MHC class II accommodates peptides of varying length, it was not necessary to screen all of them. The 3×10^7 clones of 150 bp, generated by random cDNA degradation, covered the transcriptome 15 times and therefore included most peptides. TCR− 16.2c11 cells were transduced with these libraries and cells were sorted by FACS (for example, 7.6×10^6 MCR2+ cells for the MCR2-EO771 library). Random cDNA pieces were cloned in both the sense and the antisense directions; therefore, around one in sixth peptides was encoded by the natural reading frame. The other peptides may represent ARFPs—e.g., DRiPs.

To remove artificial peptides cross-reacting with many T cells, we incubated reporter cells carrying the MCR2-EO771 library with a 10-fold excess of naive polyclonal C57BL/6 CD4+ T cells for 8 h and sorted the NFAT-negative cells. These cells were expanded to constitute the 'purged' library that was used in the second TILoma screening.

The MCR2 semi-random peptide library MCR2–8aaSR was made by cloning an oligonucleotide (GGTNNNNNNNTWCNNNNNNBCCNNNSCCNNNNNNKC CNNNGGA) into the MCR2 vector using Gibson assembly. This oligonucleotide encoded random amino acids at positions facing the TCR, while putative MHC anchor residues were partially fixed to ensure good presentation. The complexity was 5.5×10^6 bacterial clones and after transduction 11.5×10^6 individual MCR2+ cells were sorted on a BD FACSAria III cell sorter.

Stimulation of MCR2+ H18.3.13 or 16.2c11 cells. If not stated otherwise, MCR2+ cells were co-cultured with a >five-fold excess of sorted CD4+ T cells or CD4+ T cell hybridomas from indicated donor mice for 8–12 h. In some experiments, to decrease background, reporter cells were sorted GFP- or sFT-negative before co-culture. To estimate the minimal number of peptide-specific T cells able to trigger a response in reporter cells, different numbers of T cells were sorted into wells containing 1,000 MCR2+ reporter cells. NFAT activation in reporter cells following overnight co-culture was measured in 24 replicate cultures for each T cell number. A two-tailed t-test was performed.

Exome sequencing, SNP peptide library generation and screening. In the E0771 cell line we identified 627 non-synonymous mutations by whole exome sequencing (done at the Functional Genomic Center Zurich). Sure Select Mouse All Exon kit was used (Agilent) and sequencing was done in pair-end mode. Data analysis consisted of the following steps: raw reads were first cleaned by removing adapter sequences, trimming low quality ends and filtering reads with low quality phred (quality <20) using Trimmomatic (v.0.33) (Trimmomatic). Sequence alignment of the resulting high-quality reads to the Mouse reference genome (Ensembl build GRCm38) was performed using Bowtie2 (Bowtie2). For variant calling, we used GATK's HaplotypeCaller (v.3.5) (GATK). Variant annotation was done with SnpEff v.4.1l (SnpEff). Variants detected in the control and the tumor were substracted using substract command in bedtools. Only variants with an allele frequency of >0.8 and moderate or high impact were kept.

Following this, 23aa long peptides incorporating 300 of these mutations were cloned into the MCR2 vector. Then 16.2c11 cells were transduced with this MCR2-SNP library, cloned by single-cell sorting and expanded. For each analyzed clone, 1,000 cells were seeded in 96-well plates and were co-cultured with 10,000 TILs isolated directly from tumor biopsies by sorting. After 12 h, NFAT-GFP signals were analyzed by FACS on BD LSRFortessa.

Recovery of epitope sequences from selected reporter clones and verification of authenticity. DNA was prepared from selected reporter clones. The peptide-encoding part of the integrated MCR construct was PCR amplified from the

isolated DNA with the oligonucleotides (5' GGGACGTGGTTTTCCTTTGA and 5' GCCTTTCGGAGTCCCCGGAG), cloned into the MCR2 vector by Gibson assembly and sequenced, revealing the T cell–specific epitope. The vectors with newly cloned epitopes were subsequently used to verify their authenticity by transduction of 16.2c11 cells and co-culture with the corresponding hybridoma.

MCR2 downregulation assay. MCR2+ BEKO cells were co-cultured with a five-fold excess of sorted CD4+ T cells from indicated donor mice and MCR2 expression was measured on BD LSRFortessa by MHC class II staining.

LCMV and influenza virus infection. C57BL/6 mice were infected with 200 plaque forming units of LCMV WE strain or 50 plaque forming units of influenza virus PR8 strain. CD4+ T cells were isolated using anti-CD4+ magnetic beads (Miltenyi Biotech) from spleen or lung on day 5 and/or 8 post infection, and used for hybridoma formation.

Isolation of CD4+ T cells from tumor biopsies. E0771 medullary adenocarcinoma cells were implanted subcutaneously in syngeneic female C57BL/6 mice; 2 weeks later, tumor tissue was surgically removed and weighted. For TIL preparation, tumor biopsies were pooled from multiple mice, cut into small fragments and digested with collagenase and DNAse, CD4+ TILs were isolated by sorting on BD FACSAria III cell sorter and used directly in co-culture experiments or for the generation of hybridomas (TILomas).

Statistical analysis. Two-group comparisons were assessed with a Student's *t*-test (two-tailed, unpaired). C57BL/6J mice ($n = 5$ or $n = 10$ per group) were assigned to experimental groups randomly and the investigators were blinded to the group allocation and peptide identity during the vaccination and tumor challenge experiments. Student's *t*-test, ANOVA and non-linear regression analysis was done with Prism 7 (graphpad).

Clustal analysis. Clustal analysis and Phylogenic tree build was done at http://www.ebi.ac.uk/Tools/msa/clustalo/

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Sequence data generated in this study were deposited as Fastq files at the European Nucleotide Archive (https://www.ebi.ac.uk/ena) under accession codes ERS3019426, ERS3019427, ERS3019428 and ERS3019429 (B6 R1, B6 R2, E0771 R1 and E0771 R2, respectively) and study accession code PRJEB30681.

References

44. Masuda, K. et al. T cell lineage determination precedes the initiation of TCRβ gene rearrangement. *J. Immunol.* **179**, 3699–3706 (2007).

nature research

Corresponding author(s):	Jan Kisielow
Last updated by author(s):	Feb 13, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	BDFACS DIVA v(.0.1
Data analysis	FlowJo v10.4.2, Prism v6 and v7, Clustal at www.ebi.ac.uk/Tools/msa/clustalo/, Exom sequencing analysis: Trimmomatic v0.33, Bowtie2, GATK's HaplotypeCaller v3.5, SnpEff v4.1l, Bedtools v2.27

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The reported exom sequencing dataset was uploaded to ENA (Study:PRJEB30681) and can be downloaded from www.ebi.ac.uk/ena/data/view/PRJEB30681

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Most of our experiments are in vitro experiments using cell lines. Sample size was determined during experimental design, based on the measurements made (flow cytomerty, tumor mass measurments), our experience with data variation of each technique, literature and pilot experiments. A power calculation was not done.
Data exclusions	No data was excluded from the analysis.
Replication	If not stated otherwise, experiments were repeated at least 2 times. All attempts at replication were successfull.
Randomization	Mice were allocated to the experimental gorups at a random fasion.
Blinding	The analysis reported in figure 7fgh was performed in a double-blinded fashion. For other experiments, given appropriate handling and equal sample treatment and because data collection and analysis were performed by computer-based methods (e.g. flow cytometric analysis), we believe blinding was not necessary.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Fc block (CD16/CD32; 2.4G2; home-made). BioLegend: CD4-APC(GK1.5)#100412, CD4-PE(GK1.5)#100408, TCRβ-APC (H57-597)#109212, I-A-PE(M5/114.15.2)#107608, CD11b-PE(M1/70)#101208, IL-4-PeCy7(11B11)#504118, IFNg-APC(XMG1.2)#505810. eBioscience: CD3e-APC(2C11)#17-0031-83. BD Pharmingen Vbeta TCR Screening Panel #557004.
Validation	Antibodies were tested using appropriate FMO controls for flow cytometry. All antibodies had validation statement provided on the website of the manufacturer.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Modifications of the BALB/c-derived 16.2 T cell hybridoma, as well as other T cell lines and hybridomas were generated in our laboratory.
Authentication	Cells were not authenticated other than by flow cytometry staining for CD3e, TCR, CD4 and I-A.
Mycoplasma contamination	Cells were not tested for mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	C57Bl/6J Mice, females, 6-8 weeks old were used.

Wild animals	none
Field-collected samples	none
Ethics oversight	All animal experiments were approved by the local animal ethics committee (Kantonales Veterinärsamt Zürich, licenses ZH061/17), and performed according to local guidelines (TschV, Zurich) and the Swiss animal protection law (TschG).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Single cell suspenssions were done by standard tissue culture procedures.
Instrument	BD FACSAria III, BD LSRFortessa
Software	FlowJo v10.4.2
Cell population abundance	Post-sort cell populations were >95% pure, as determined by re-analysis.
Gating strategy	Gates were set according to FMO or unstained/untreated controls.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.