OPEN ACCESS

University of BRISTOL

University of Bristol - Explore Bristol Research

General rights

npg

ORIGINAL ARTICLE

Naive CD8+ T-cell precursors display structured TCR repertoires and composite antigen-driven selection dynamics

Michelle A Neller[1,7], Kristin Ladell[2,7], James E McLaren[2], Katherine K Matthews[1], Emma Gostick[2], Johanne M Pentier[2], Garry Dolton[2], Andrea JA Schauenburg[2], Dan Koning[3], Ana Isabel CA Fontaine Costa[3], Thomas S Watkins[1], Vanessa Venturi[4], Corey Smith[1], Rajiv Khanna[1], Kelly Miners[2], Mathew Clement[2], Linda Wooldridge[2], David K Cole[2], Debbie van Baarle[3], Andrew K Sewell[2], Scott R Burrows[1,5], David A Price[2,6,7] and John J Miles[1,2,5,7]

Basic parameters of the naive antigen (Ag)-specific T-cell repertoire in humans remain poorly defined. Systematic characterization of this 'ground state' immunity in comparison with memory will allow a better understanding of clonal selection during immune challenge. Here, we used high-definition cell isolation from umbilical cord blood samples to establish the baseline frequency, phenotype and T-cell antigen receptor (TCR) repertoire of CD8+ T-cell precursor populations specific for a range of viral and self-derived Ags. Across the board, these precursor populations were phenotypically naive and occurred with hierarchical frequencies clustered by Ag specificity. The corresponding patterns of TCR architecture were highly ordered and displayed partial overlap with adult memory, indicating biased structuring of the T-cell repertoire during Ag-driven selection. Collectively, these results provide new insights into the complex nature and dynamics of the naive T-cell compartment.
Immunology and Cell Biology (2015) **93**, 625–633; doi:10.1038/icb.2015.17; published online 24 March 2015

Reactive T cells from the extrathymic naive pool are expanded and mobilized into the memory repertoire by specific and productive interactions with cognate antigen (Ag). Exactly which clonotypes are recruited during this process, however, remains a source of recurrent immunological inquiry. At present, we understand that extrathymic αβ T-cell selection is dependent on a number of variables, including Ag abundance, priming location, T-cell antigen receptor (TCR) ligand binding parameters and precursor frequency (reviewed in Allen *et al.*[1]).

Recent technical advances in the detection and isolation of T-cell precursors directly *ex vivo* have allowed, for the first time, the unambiguous enumeration and characterization of unmanipulated naive Ag-specific populations.[2] Initial experiments in mice revealed that Ag-specific precursors are present at frequencies of 0.08–890 cells per 100 000 CD4+/CD8+ T cells (reviewed in Jenkins and Moon[3]). In addition, T-cell precursor frequencies were found to cluster numerically by Ag specificity between different mice. Interestingly, early precursor enumerations appeared to correlate positively with immunodominance hierarchies after pathogen challenge.[3] However, recent evidence in other murine systems[4,5] demonstrated that this association can sometimes be inverse, indicating that memory formation is complex and involves the proliferative capabilities of individual T-cell precursors. In humans, initial calculations from adult peripheral blood place Ag-specific precursor frequencies between 2 and 600 cells per 100 000 CD4+/CD8+ T cells.[6–10]

In this study, we aimed to determine several baseline parameters of Ag-specific precursors in humans through the use of umbilical cord blood (UCB). We combined a modified multimer-based magnetic enrichment protocol with high-definition multiparametric flow cytometry to ensure the high-purity isolation, accurate enumeration and detailed phenotypic characterization of Ag-specific precursors directly *ex vivo*. Furthermore, we ascertained TCR usage in these Ag-specific precursor populations by using a quantitative molecular approach and conducted a systematic comparison with our large bank

of memory TCR repertoires from adults to explore the effects of Ag-driven selection. This study represents the first enumeration and analysis of Ag-specific CD8[+] T-cell precursors in human newborns.

RESULTS

Sample selection and high-definition T-cell isolation
We used UCB to characterize the baseline Ag-specific T-cell precursor repertoire in order to bypass certain technical limitations seen when using adult material. First, in adults, identification of authentic naive cells is sometimes difficult due to subsets of memory T cells that exhibit 'naive-like' phenotypes.[11] Second, naive T cells steadily proliferate in adults over time, which can distort precursor calculations.[12] Third, adult humans have lengthy and multifaceted immune histories, such that the naive/memory T-cell ratio can vary widely between individuals.[12] This variable can drastically alter precursor estimates when calculating frequencies across total lymphocytes or lymphocyte subsets. Finally, over the long duration of a human life, it is difficult to rule out contact with any given infectious agent or immunogenic compound. This is particularly applicable to ubiquitous influenza viruses and human herpesviruses, which can produce false negatives or discrepant results during serotesting.[13,14] In contrast, the large bulk of the UCB compartment is naive and, in healthy mothers, vertical transmission of the pathogens investigated here is rare.

To ensure high-definition isolation of low-frequency T cells, we first compared tetramers versus dextramers, which comprise long dextran polymer backbones individually decorated with ~24 peptide-major histocompatibility complex and ~12 fluorochrome molecules,[15] for their ability to identify naive precursors in UCB. At the optimal median fluorescence intensity titration for each reagent, dextramers proved superior to tetramers in parallel multiparametric flow cytometry trials (Figure 1). Specifically, dextramer-labeled populations were more readily identifiable and >12-fold brighter than their tetramer-labeled counterparts. These findings concur with a recent study in which dextramers consistently outperformed tetramers, most strikingly for the identification of low-avidity T-cell populations.[16] To further facilitate precursor isolation, we combined dextramers with dasatinib pre-treatment, which can enhance peptide-major histocompatibility complex multimer fluorescence without concomitant increases in background staining by blocking Ag-induced TCR downregulation.[17] As predicted, we observed brighter staining of Ag-specific precursors in the presence of dasatinib (data not shown). It is noteworthy that this approach does not alter T-cell phenotype.[16] Accordingly, all multimer magnetic enrichments in this study were performed by using a combination of dasatinib pre-treatment and dextramer staining.

Analysis of naive Ag-specific T-cell precursor frequencies
Six different dextramers were synthesized bearing a series of viral and self-derived Ags as follows: HLA-A*0201-ELAGIGILTV (A2-ELA) from human Melan-A/MART-1 (residues 26–35), HLA-A*0201-GILGFVFTL (A2-GIL) from influenza virus matrix (residues 58–66), HLA-A*0201-GLCTLVAML (A2-GLC) from Epstein–Barr virus (EBV) BMLF1 (residues 280–288), HLA-A*0201-NLVPMVATV (A2-NLV) from cytomegalovirus (CMV) pp65 (residues 495–503), HLA-B*0801-FLRGRAYGL (B8-FLR) from EBV EBNA3A (residues 325–333) and HLA-B*0702-TPRVTGGGAM (B7-TPR) from CMV pp65 (residues 417–426). Using these reagents in conjunction with a modified magnetic enrichment protocol (Figures 2a and b) and a strict flow cytometric gating strategy (Figure 2c), we observed clustering of T-cell precursor frequencies based around epitope specificity (Figure 3a and Supplementary Table 1). A2-ELA-specific precursors were by far the most frequent in the UCB repertoire, averaging 229 cells per 100 000 CD8[+] cells (range 120–352). Virus-specific precursors were rarer by more than an order of magnitude in all cases, with a hierarchy per 100 000 CD8[+] cells as follows: A2-GIL-specific precursors, mean 15.8 cells (range 1.2–66.9) > A2-GLC-specific precursors, mean 10.4 cells (range 0.3–34.7) > A2-NLV-specific precursors, mean 3.7 cells (range 1.9–6.9) > B8-FLR-specific precursors, mean 2.3 cells (range 0.4–4.7) > B7-TPR-specific precursors, mean 1 cell (range <1–1.6). In contrast, the immunodominance hierarchy in virus-exposed adults was more or less inverted, with memory T-cells specific for the herpesviruses EBV and CMV dominating the CD8[+] compartment (Figure 3b). A2-GLC-specific precursors were detected at a mean of 0.15 cells per 100 000 CD8[+] cells (range 0.1–0.2) in EBV-seronegative adults and A2-NLV-specific precursors were detected at a mean of 1.3 cells per 100 000 CD8[+] cells (range 0.8–2.1) in CMV-seronegative adults (Figure 3a and Supplementary Table 1).
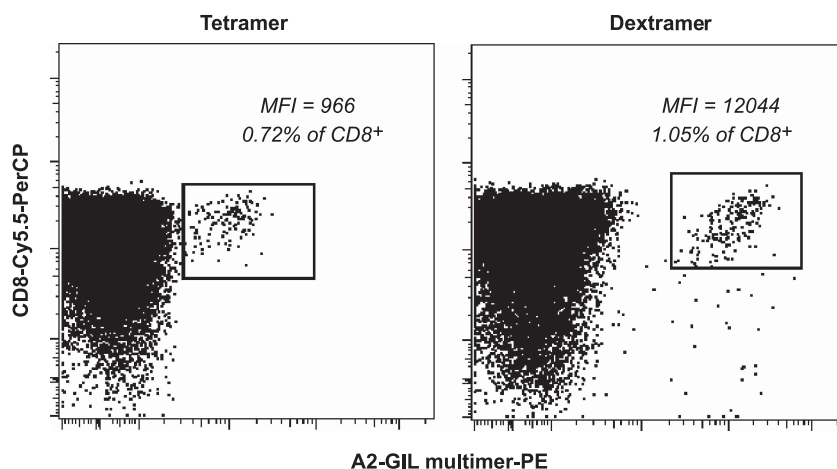
Figure 1 Comparison of multimers for use in magnetic enrichment. An HLA-A2[+] UCB sample was divided and labeled with PE-conjugated A2-GIL tetramer or dextramer, then magnetically enriched using αPE beads. Enriched cells were surface stained for CD8 expression and acquired using a FACSCanto II flow cytometer. The gating strategy is shown in Figure 2. Numbers indicate the percentage of multimer[+] cells within the enriched CD8[+] population. The median fluorescence intensity of each multimer[+] population is shown.
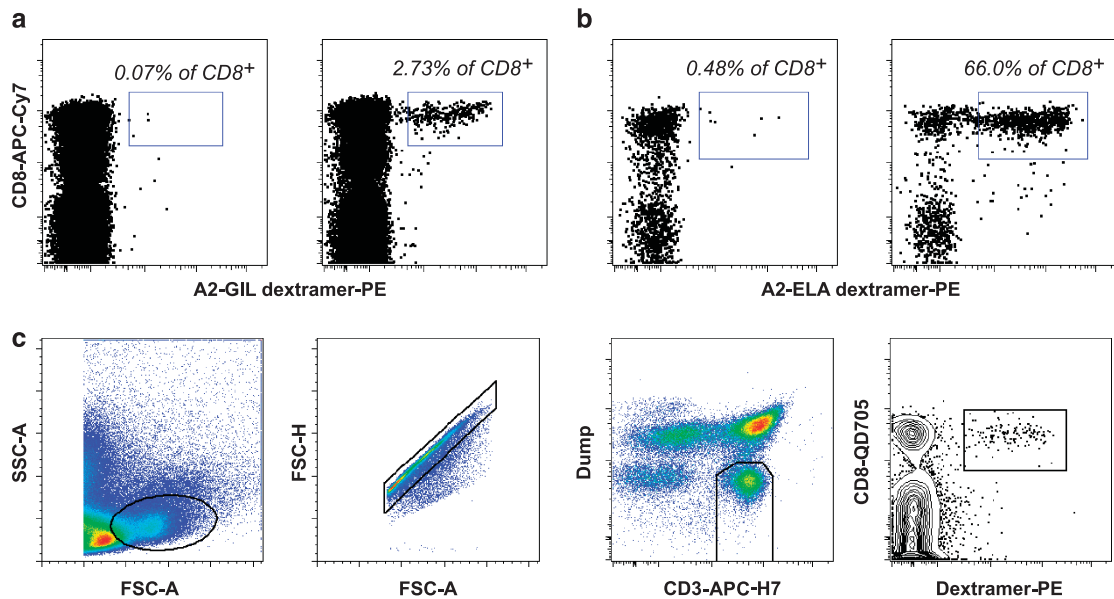
Figure 2 Magnetic enrichment and gating strategy for naive Ag-specific T-cell precursor isolation. (a and b) UCB samples were compared before and after magnetic enrichment for A2-GIL (a) and A2-ELA (b). Dextramer+ cells were either undetectable or barely detectable in unenriched UCB. In contrast, dextramer+ populations were clearly visible after magnetic enrichment. Numbers indicate the percentage of dextramer+ cells within the enriched CD8+ population. (c) Gating strategy used for high-definition flow cytometric sorting. The dump channel comprised LIVE/DEAD Fixable Violet Dead Cell Stain together with αCD14 and αCD19 mAbs conjugated to Pacific Blue. Single lymphocytes were selected based on light scatter properties. Non-dump CD3+ cells were gated and Ag-specific T cells were identified in a CD8 versus dextramer bivariate plot.

Figure 3 Ex vivo enumeration of Ag-specific precursors and memory T cells from humans. (a) The number of dextramer+ cells per 100 000 CD8+ cells was calculated from 46 UC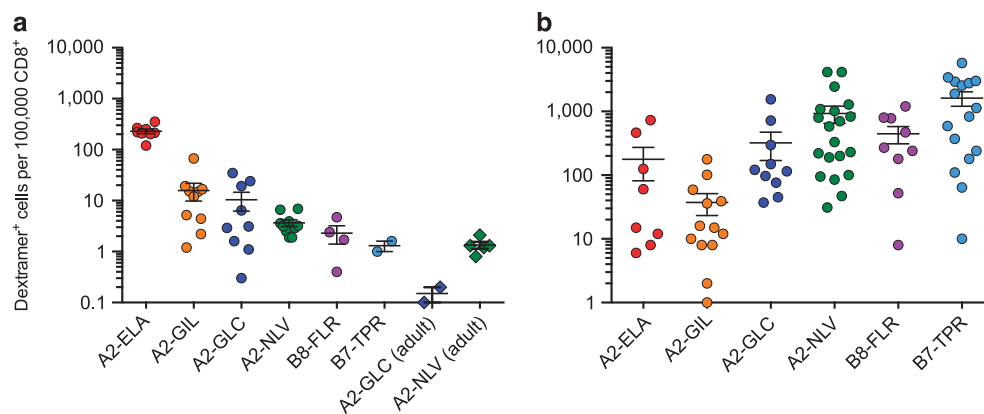B samples and seven herpesvirus-seronegative adult peripheral blood mononuclear cell (PBMC) samples. (b) The number of dextramer+ cells per 100,000 CD8+ cells was calculated from 72 herpesvirus-seropositive adult PBMC samples. Ag-specific precursor enumeration was achieved via dextramer magnetic enrichment with the exception of A2-ELA-specific cells, which were detectable in unmanipulated UCB. The flow cytometric gating strategy is shown in Figure 2c. Statistically significant differences were identified between Ag specificities (Supplementary Table 2). Among precursor populations, A2-ELA-specific T cells were significantly more frequent compared with all other Ag specificities ($P<0.001$). Among memory populations, A2-GIL-specific T cells were significantly less frequent compared with other viral specificities ($P<0.05$).

These numbers are substantially lower than the corresponding frequencies in UCB samples but comparable with those reported previously in adults.[3] Statistical analysis revealed highly significant differences between precursor frequencies across Ag specificities in UCB, most notably for A2-ELA in all comparisons with viral specificities (Supplementary Table 2A). A more nuanced picture emerged in adults, although A2-GIL-specific precursors occurred at consistently lower frequencies compared with other viral specificities (Supplementary Table 2B). Collectively, these data demonstrate that precursor frequencies cluster by Ag specificity and vary with age.

Phenotypic analysis of naive Ag-specific T-cell precursors

Multiparametric flow cytometry-based phenotypic data were acquired using an optimized panel built around ten distinct fluorochromes. The core gating strategy and representative analyses of dextramer+ T-cell precursor populations are shown in Figures 2 and 4, respectively. For all epitope specificities across all UCB samples,
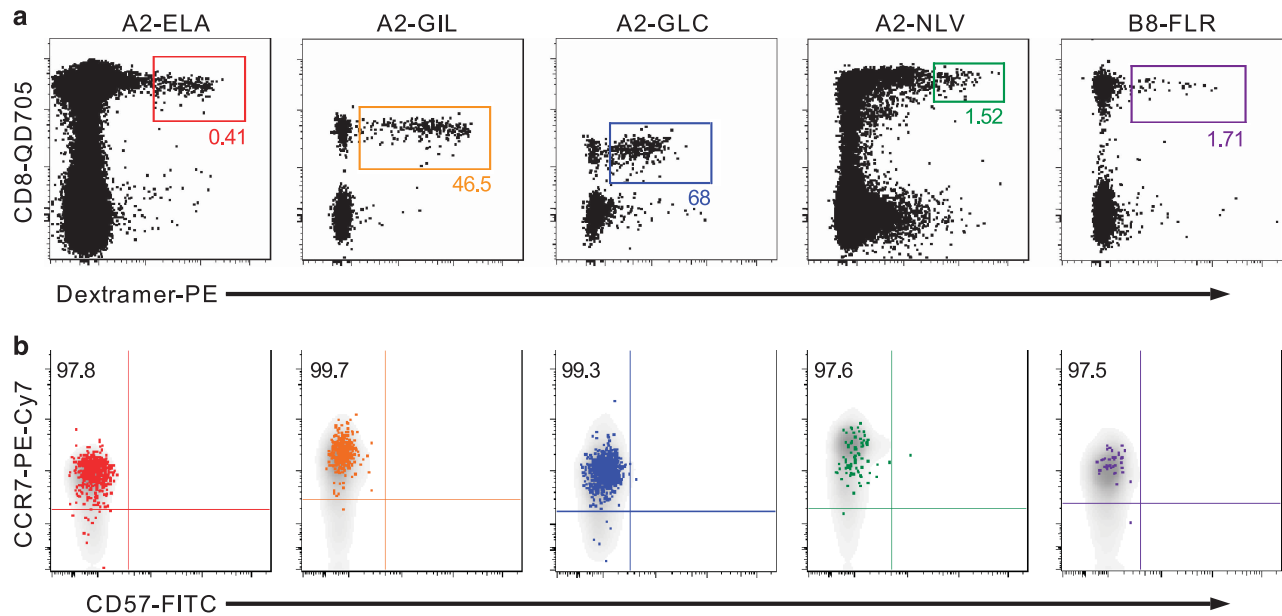
Figure 4 *Ex vivo* phenotyping of naive Ag-specific T-cell precursors. (**a**) Representative flow cytometry plots showing sort gates for naive T-cell populations across five epitope specificities as indicated. Ag-specific T cells were identified via dextramer magnetic enrichment with the exception of A2-ELA-specific cells, which were detectable in unmanipulated UCB. Numbers indicate the percentage of dextramer+ cells within the total CD8+ population. (**b**) The phenotype of dextramer+ cells (coloured) overlaid on all CD3+ T cells (gray), showing CCR7 expression and the absence of CD57. Numbers indicate the percentage of dextramer+ cells expressing this naive phenotype. Other surface marker analyses yielded similar data.

dextramer+ cells displayed the classic naive T-cell phenotype (CD27^{hi} CD45RA^{hi} CD45RO^{lo} CD57^{lo} CCR7^{hi}). These data help to verify the truly naive status of Ag-specific T-cell precursors identified in the present study.

Clonotypic analysis of naive Ag-specific T-cell precursors

Next, we examined TCR usage in naive Ag-specific T-cell precursor populations by sorting magnetically enriched dextramer+ cells at >98% purity directly into microtubes containing an RNA protectant and using a template-switch anchored PCR with reverse transcription to amplify all expressed *TR* gene transcripts without bias.[18] Final cell numbers varied between 30 and 2000 per sample depending on epitope specificity and population frequency. To contextualize the data, we compared precursor TCR transcripts (1320 sequences) with our bank of adult memory TCR transcripts covering the same specificities (6550 sequences). The similarities and differences in TCR gene usage and CDR3 length between naive precursors and memory cells are illustrated in Figure 5.

For A2-ELA-specific T cells, *TRBV* and *TRBJ* gene usage showed considerable overlap and some minor divergence (Figures 5a and c). In particular, *TRBV4* and *TRBJ2-7* were enriched in UCB, while *TRBV6* and *TRBJ1-5* were enriched in adults. The CDR3β lengths showed Gaussian distributions, with dual preference for a loop length of nine residues (Figure 5e). As A2-ELA-specific T cells in adults are known to have biased TCR α-chain architecture,[19] we examined *TRA* gene transcripts in the corresponding UCB precursor populations. Like adults, naive precursors were heavily biased toward *TRAV12* gene usage (Figure 5b); *TRAJ* gene usage was not especially ordered between populations (Figure 5d).

For A2-GIL-specific T cells, *TRBV* and *TRBJ* gene usage also showed some commonalities between adults and UCB with bias toward *TRBV19* (Figure 5a). Interestingly, in UCB precursors but not adult cells, *TRBV* gene usage was evenly split between *TRBV6-5* and *TRBV19*. Both cohorts showed enrichment for *TRBJ2* genes, with

adults displaying a strong preference for *TRBJ2-7* (Figure 5c). The CDR3β lengths in both populations were distributed in a Gaussian fashion with adult memory cells more focused on a peak loop length of eight residues (Figure 5e).

For A2-GLC-specific T cells, *TRBV* gene usage was broad across both naive and memory populations with moderate overlap (Figure 5a). Adults displayed a preference for *TRBV20-1* and *TRBV29-1*, while *TRBV12* and *TRBV27* were preferred in UCB. A similar pattern applied to *TRBJ* gene usage, with moderate overlap and population-specific bias (Figure 5c). Adults displayed a preference for *TRBJ1-2*, while UCB precursors preferentially used either *TRBJ2-3* or *TRBJ2-7*. The CDR3β lengths were superimposable between populations (Figure 5e).

For A2-NLV-specific T cells, *TRBV* gene usage was broad across both naive and memory populations with considerable overlap (Figure 5a). However, UCB precursors were enriched for *TRBV29-1* and adult memory cells were enriched for *TRBV27*. A similar pattern applied to *TRBJ* gene usage, with broad overlap and enrichment for *TRBJ2-7* and *TRBJ1-2* in UCB and adults, respectively (Figure 5c). The CDR3β lengths in both populations were distributed in a Gaussian fashion, although UCB precursors preferred a longer loop length of ten residues (Figure 5e).

For B8-FLR-specific T cells, *TRBV* gene usage was noticeably biased in both UCB precursors and adults (Figure 5a). In particular, *TRBV7-8* was used frequently in adults and, to a lesser extent, in UCB precursors. More frequently, UCB precursors used *TRBV12-4* and *TRBV27*. In both populations, *TRBJ2-7* was preferred (Figure 5c). The CDR3β lengths were different with peaks at eight and nine residues for UCB precursors and adult memory cells, respectively (Figure 5e).

Finally, we undertook a more detailed analysis of repertoire architecture by calculating basic statistics for CDR3β length, *TRBV* and *TRBJ* gene usage (Supplementary Figure 1 and Supplementary Table 3). Distribution values for CDR3β length were similar across adult repertoires (median 12–14 amino acids, range 10–20) and naive
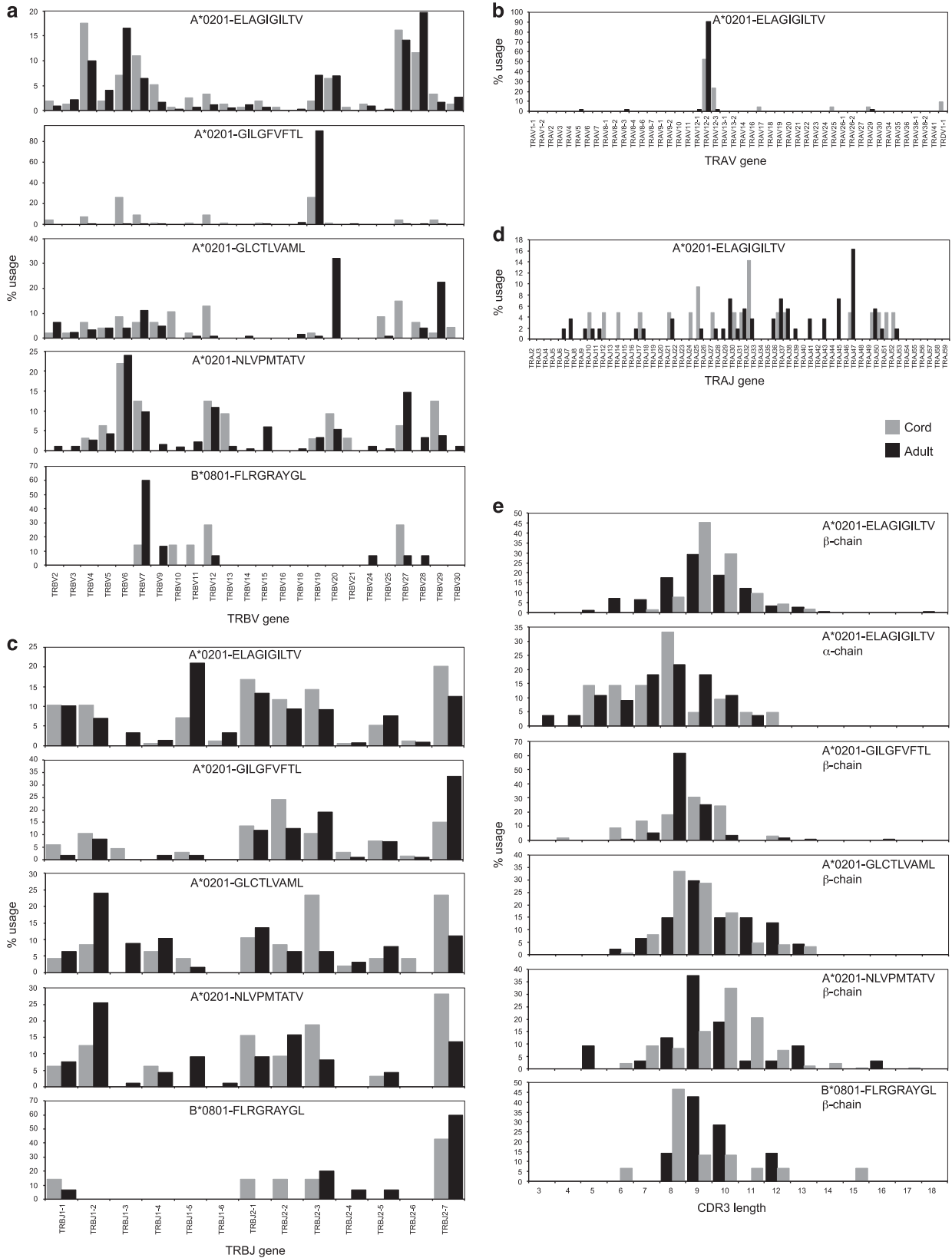
Figure 5 Comparison of *ex vivo* Ag-specific TCR repertoires between naive and memory T cells. (**a–e**) *TRBV* gene usage (**a**), *TRAV* gene usage (**b**), *TRBJ* gene usage (**c**), *TRAJ* gene usage (**d**) and CDR3 length calculated using the Chothia nomenclature (**e**). Individual clonotypes were weighted by appearance with each unique sequence counted once regardless of frequency. The database comprised 7880 sequences derived from 85 Ag-specific TCR repertoires. Naive and memory sequences, respectively, were compared for A2-ELA (612 and 2134), A2-GIL (144 and 631), A2-GLC (145 and 1127), A2-NLV (245 and 2330) and B8-FLR (174 and 338).

Table 1 Ag-specific clonotype homology between naive and memory T cells

Epitope	Donor	TRV	CDR3	TRJ
A2-ELA	Cord[3]	TRBV4-1	CASSQDGGVTQY	2-3
	Cord[1]	TRBV4-3	CASSQDGGLNGELF	2-2
	Cord[3]	TRBV4-1	CASSQDGGTSGGGQI	2-1
	Cord[6]	TRBV6-2/3	CASSWGQGDGYT	1-2
	Cord[6]	TRBV6-2/3	CASSLGQGAEQF	2-1
	Adult[1]	TRBV6	CASSLGQGQPQH	1-5
	Adult[2]	TRBV6	CASSFGQGQPQH	1-5
	Adult[2]	TRBV6	CASSPGQGFVKF	2-1
	Adult[3]	TRBV6	CASSPGQGAFSYT	1-2
	Adult[3]	TRBV6	CASSYGQGSIGELF	2-2
	Cord[7]	TRBV6-5	CASSYGTGNIQY	2-6
	Adult[4]	TRBV6-2/3	CASSYGTGQPQH	1-5
	Cord[2*]	TRBV6-5	CASSYGTGGNEQF	2-1
	Cord[1]	TRAV12-2	CAVNNARLM	31
	Adult[5]	TRAV12-2	CAVNNARLM	31
	Adult[6]	TRAV12-2	CAVNNARLM	31
	Adult[7]	TRAV12-2	CAVNNARLM	31
	Adult[6]	TRAV12-2	CAPSNARLM	31
	Cord[1]	TRAV12-2	CAVNSNTGNQFY	49
	Adult[7]	TRAV12-2	CAVPPNTGNQFY	49
	Cord[2]	TRAV12-2	CAVNGGFGNVLH	35
	Adult[8]	TRAV12-2	CAVNKGFGNVLH	35
	Adult[9]	TRAV12-2	CAVMIGFGNVLH	35
	Adult[10]	TRAV12-2	CAVSIGFGNVLH	35
	Adult[4]	TRAV12-2	CAVSIGFGNVLH	35
	Cord[1]	TRAV12-2	CAVTTDSWGKLQ	24
	Cord[6]	TRAV12-2	CAVNTDSWGKLQ	24
	Adult[11]	TRAV12-2	CAVRTDSWGKLQ	24
	Cord[1]	TRAV12-3	CAIRDDKII	30
	Adult[12]	TRAV12-2	CAVRDDKII	30
	Cord[6]	TRAV12-3	CAMSGSNFGNEKLT	48
	Adult[13]	TRAV12-2	CAMSRSNFGNEKLT	48
	Adult[14]	TRAV12-2	CAVRGSNFGNEKLT	48
	Cord[6]	TRAV12-3	CAFGGGADGLT	45
	Adult[14]	TRAV12-2	CAAGGGADGLT	45
	Adult[14]	TRAV12-2	CAGGGGADGLT	45
	Adult[14]	TRAV12-2	CALGGGADGLT	45
	Adult[14]	TRAV12-2	CARGGGADGLT	45
	Adult[14]	TRAV12-2	CASGGGADGLT	45
	Adult[14]	TRAV12-2	CAVGGGADGLT	45
	Adult[15]	TRAV12-2	CAVGGGADGLT	45

Table 1 (Continued)

Epitope	Donor	TRV	CDR3	TRJ
A2-GIL	Cord[9]	TRBV6-5	CASSYGQNTGELF	2-2
	Cord[10]	TRBV6-5	CASSYGQNTGELF	2-2
	Cord[9]	TRBV6-5	CASSYGQGTGELF	2-2
	Cord[10]	TRBV6-5	CASSYGAGTGELF	2-2
	Cord[9]	TRBV6-5	CASSYGQGTGIQY	2-4
	Cord[9]	TRBV6-5	CASSYGQGLYEQY	2-7
	Cord[9]	TRBV19	CASSPLAGGFNEQF	2-1
	Cord[10]	TRBV19	CASSPLAGLYNEQF	2-1
	Cord[10]	TRBV19	CASSPLAGFYNEQF	2-1
	Cord[9]	TRBV19	CASSIVSTGELF	2-2
	Adult[16]	TRBV19	CASSIRSTGELF	2-2
	Adult[17]	TRBV19	CASSISSTGELF	2-2
	Cord[9]	TRBV19	CASSIISTDTQY	2-3
	Adult[18]	TRBV19	CASSIRSTDTQY	2-3
	Adult[19]	TRBV19	CASSIRSADTQY	2-3
	Adult[19]	TRBV19	CASSIRSSHTQY	2-3
A2-GLC	Adult[20]	TRBV20-1	CGARDGTGNGYT	1-2
	Adult[21]	TRBV20-1	CSARDGTGNDYT	1-2
	Adult[22]	TRBV20-1	CSARDGTGNGYT	1-2
	Adult[23]	TRBV20-1	CSAREGTGNGYT	1-2
	Adult[24]	TRBV20-1	CSARDGTGNGYT	1-2
	Adult[25]	TRBV20-1	CSARDGTGNGYT	1-2
	Cord[9]	TRBV9	CASSLGTGDEQY	2-7
	Cord[9]	TRBV29-1	CSVEDGTGSYEQY	2-7
A2-NLV	Cord[12]	TRBV12-3/4	CASSLFGGGTEAF	1-1
	Cord[12]	TRBV27	CASSPTGGGTGELF	2-2
	Cord[16]	TRBV13	CASSLGGGSYEQY	2-7
	Cord[16]	TRBV13	CAGSLGGGSYEQY	2-7
	Adult[26]	TRBV7-8	CASSLGGGGFYEQY	2-7
	Adult[27]	TRBV13	CASSYGGGEDYGYT	1-2
	Cord[3*]	TRBV29-1	CSGGTGELF	2-2
	Adult[28]	TRBV29-1	CSVGGTLDTQY	2-3
	Cord[9]	TRBV20-1	CSARDQTAYYGYT	1-2
	Adult[27]	TRBV6-5	CASSYQTGAAYGYT	1-2
	Adult[28]	TRBV6-5	CASSYQTGAIYGYT	1-2
	Adult[29]	TRBV6-5	CASSYQTGTIYGNT	1-2
	Adult[30]	TRBV6-5	CASSYQTGTIYGYT	1-2
	Adult[31]	TRBV6-5	CASSYQTGAAYGYTF	1-2
	Adult[32]	TRBV6-5	CASSCQTGAACGYTF	1-2
	Adult[33]	TRBV6-5	CASSLQTGVRSYEQY	2-7

A number of A2-ELA TCR sequences were derived from concatenations of data previously reported in the literature.[19]
*Bead-amplified T-cell cultures.

precursor repertoires (median 12.5–13.5 amino acids, range 9–20). In contrast, there was a trend towards larger numbers of *TRBV* and *TRBJ* genes in the naive precursor populations. Collectively, these data suggest that naive CD8[+] T-cell precursor repertoires are architecturally more diverse than the corresponding Ag-experienced memory repertoires.

TCR bias in the naive Ag-specific T-cell precursor repertoire

TCR sequencing of Ag-specific cells revealed that UCB precursors exhibited biased chain architecture, often overlapping with the corresponding adult memory repertoires (Table 1). Identical (type III bias) and near-identical (type IV bias) TCR usage between individuals was observed within the A2-ELA specificity. Type III bias, whereby the same 'public' TCR chain is used in multiple individuals,[19,20] was observed for clonotypes with TRAV12 transcripts. Similar repertoire biases were observed for A2-GIL. Notably, a public TRBV6-5 transcript present in A2-GIL-specific UCB precursors was curiously absent in adult memory. Highly homologous TRBV19 transcripts were present in both A2-GIL-specific UCB precursors and adult memory cells. Type III and type IV bias were also observed in A2-GLC-specific populations. Type IV bias and 'TCR mosaicism', defined by alternate *TRBV* gene swapping around a common CDR3β core,[21] were observed in the A2-NLV-specific repertoires, both within and between UCB precursors and adult memory cells.

DISCUSSION

In humans, around 170 αβ *TR* genes are recombined and somatically modified to create a circulating repertoire of ~ 10[12] T cells bearing between 10[6] and 10[11] unique TCRs.[19] Precisely what fraction of this grand repertoire is capable of recognizing any given peptide-major histocompatibility complex has been a lingering question for well over two decades. The recent arrival of multimer magnetic enrichment technology[3] has provided a key to unlock the mysteries of the Ag-specific naive pool for the first time by using a direct and unmanipulated *ex vivo* approach. In the present study, we used UCB samples to obtain baseline frequency readings and found that Ag-specific T-cell precursors exist in the naive repertoire across a range between <1 and 352 cells per 100 000 CD8[+] cells. These numbers are up to two orders of magnitude higher than those recorded previously in mice and human adults.[3,10] The frequency differences between UCB and adults can be explained relatively easily by the fact that, in adults, the naive compartment is generally smaller than memory, which directly impacts precursor calculations. However, why UCB precursor frequencies are higher than those found in newborn and/or pathogen-free mice is more difficult to explain. Humans have ~ 10 000-fold more T cells than mice.[19] Assuming comparable TCR recognition degeneracy and proportionate repertoire diversity, humans should therefore encompass superior total antigenic coverage on a whole organism scale. However, when frequencies are assessed per 100 000 CD8[+] T cells, differences between species should conceivably be harder to decipher.

Despite the genetic heterogeneity associated with outbred human populations, T-cell precursor frequencies were hierarchically clustered by Ag specificity as follows: A2-ELA > A2-GIL > A2-GLC > A2-NLV > B8-FLR > B7-TPR. The neonates used in this study were not related, typically sharing only the single HLA class I allele restricting the epitope of interest. This suggests that HLA haplotype does not greatly influence final precursor frequencies in the extrathymic pool. Given previous observations of a positive relationship between HLA haplotype and the magnitude of a memory T-cell response in virus-

exposed individuals,[22] these combined data suggest that HLA bias occurs after priming.

Naive precursors specific for A2-ELA were by far the most prevalent, peaking at one cell per 285 CD8[+] thymic emigrants. What could explain these spectacularly high frequencies? Structural considerations are potentially informative in this setting. The TCR-A2-ELA ternary complex reveals TCR docking in an 'α-centric' fashion via the *TRAV12* gene segment.[23] Thus, it may be that large numbers of circulating clonotypes are capable of binding A2-ELA through minor modifications around a basic '*TRAV12* gene blueprint'. Additional evidence to support this idea comes from the observation of *TRAV12* gene-biased A2-ELA-reactive cells in HLA-A2[-] individuals.[6] An alternative but potentially synergistic explanation involves leaky tolerance as it has been shown that human medullary thymic epithelial cells misinitiate transcription of the *Melan-A/MART-1* gene.[24] High-frequency A2-GIL-specific cells can be explained similarly. These clonotypes bind in a 'β-centric' manner[25], making heavy use of germline regions encoded by the *TRBV19* gene.[19] Thus, akin to A2-ELA-specific clonotypes, there may be common A2-GIL-specific TCRs based around a '*TRBV19* gene blueprint'. Interestingly, we also identified a mysterious and equally large TRBV6-5[+] population within the A2-GIL-specific naive repertoire. Some of the corresponding transcripts were found to be residue-identical between unrelated UCB donors, yet encoded via unique nucleotide sequences. Furthermore, we identified 'near-public' clonotypes within the TRBV6-5[+] pool that differed at only one or two residues across the CDR3β loop. Although further data are required to assess the general applicability of this phenomenon, we found no evidence of these highly biased TRBV6-5[+] TCRs in the adult memory compartment. In fact, we did not identify a single *TRBV6-5* gene transcript from our entire adult database of 631 A2-GIL-specific TCRs. This intriguing observation suggests that the naive Ag-specific repertoire is much more complex than previously thought, comprising multiple layers of publicity. Some public TCRs are preferred during Ag-driven selection while others are neglected. Exactly why this occurs is unknown, but the biology of these 'lost publics' is likely to be intriguing.

TCR recognition of A2-GLC, A2-NLV and B8-FLR has been studied at the atomic level by ourselves[26,27] and others.[28] The ternary complexes reveal equally split footprints across the TCR α-chains and β-chains with the majority of bound Ag contacts focused on the CDR3 loops. These structural constraints may limit the final frequencies of cognate precursors in the naive pool. We observed considerable *TR* gene usage overlap between UCB precursors and adult memory when examining the A2-GLC, A2-NLV and B8-FLR datasets. However, in spite of these heavily biased naive repertoires, we failed to identify *bona fide* public receptors against these three Ags in UCB. One possible explanation for this unexpected finding is that public TCRs specific for these Ags may simply exist at low frequencies *in vivo*. This scenario would suggest that there is something inherently special about public TCRs, likely related to optimal Ag engagement, that allows the respective clonotypes to dominate the memory pool over decades of life.[29]

In agreement with our data, a recent study in mice found that: (i) naive T-cell precursors can cluster numerically by Ag specificity; (ii) T-cell precursor frequencies do not correlate with immunodominance hierarchies observed after pathogen encounter; (iii) public clonotypes are relatively rare in the precursor pool; and (iv) there is a distinct tapering of the TCR repertoire during transition to memory.[4] Similar investigations in HIV-seropositive adults also failed to find a correlation between precursor frequency and immunodominance.[30] These observations across species suggest that the above parameters of

632

naive cellular immunity may be intrinsic to mammalian biology. In human adult peripheral blood, very recent enumerations have been placed at one Ag-specific cell per 1000–1000 000 CD8[+] cells[7] or one cell per 100 000–1000 000 CD4[+] cells[10] for both foreign and self-derived Ags. Additionally, for CD4[+] precursors, memory-phenotype T cells were present in adults but not newborns.[10] These findings may reflect T-cell activation via crossreactive ligands. Given this collective dataset, the enumerations of two Ag-specific precursor populations in our study can be compared with the corresponding specificities in adults.[7] Such comparisons reveal that newborns have two-fold more A2-NLV-specific precursors and five-fold more A2-ELA-specific precursors.[7] The cause of this enrichment is unknown, but it is thought that UCB T cells have a highly proliferative phenotype,[31] which could conceivably increase precursor numbers. Likewise, intrinsic qualitative deficiencies in recruitment and/or division may thwart the mobilization of certain precursor clonotypes into memory.[32]

In summary, we explored the human newborn immune system to define and characterize baseline T-cell immunity against common viruses and immunogenic Ags. Despite considerable genetic hetero-geneity, we found that T-cell precursor frequencies between neonates tended to cluster by Ag specificity. These frequency clusters varied by > 300-fold, potentially as a consequence of disparate TCR docking modes. Molecular analysis of TCR usage revealed that the naive Ag-specific T-cell repertoire is typically broad and predictably structured between infants. Despite considerable overlap with the adult reper-toire, however, it was clear that the naive repertoire is not an exact copy of memory. On the contrary, we found that the naive Ag-specific T-cell repertoire was complex, comprising interlayers of structurally biased receptors. Some public clonotypes were observed in the Ag-driven memory compartment, while others were apparently not selected. Further work is required to untangle these complex repertoire dynamics. Nonetheless, the observation that 'unexploited' Ag-specific clonotypes reside in the naive T-cell precursor pool opens up new possibilities for vaccine development. Indeed, only now are we beginning to understand the pathological consequences of 'clonotype choice', which can determine host outcome in the face of pathogen challenge.[33–36] Ultimately, the ability to target and amplify 'ideal' precursors from the naive pool, possibly through the use of TCR-optimized peptides to mobilize specific clonotypes,[37–39] could be a potent advance for biomedicine.

METHODS

Blood samples and processing
Density gradient centrifugation-enriched anonymized UCB lymphocytes were obtained from The Anthony Nolan Trust, Royal Free Hospital, Hampstead, UK. Healthy adult blood was obtained from donors at the Queensland Institute of Medical Research, Brisbane, Australia. Exposure to EBV and CMV was determined serologically via ELISA. Lymphocytes were isolated by Ficoll-Paque PLUS (GE Healthcare, Little Chalfont, UK) density gradient centrifugation and cryopreserved in R10 medium (RPMI-1640 containing 10% FBS) supplemen-ted with 10% DMSO (Sigma-Aldrich, Gillingham, UK). For some UCB samples, cells were expanded by adding Human T-Activator CD3/CD28 Dynabeads (Life Technologies, Paisley, UK) as described previously.[40] All samples were obtained with written consent and all protocols were approved by ethics committees at The Anthony Nolan Trust and the Queensland Institute of Medical Research.

Antibodies, cell stains and dextramers
The following mAb conjugates were used at pre-titrated concentrations: (i) αCD3-APC-H7 (BD Biosciences, San Jose, CA, USA); (ii) αCD8-QD705, αCD14-Pacific Blue and αCD19-Pacific Blue (Life Technologies); (iii) αCD8-

PE, αCD57-FITC and αCCR7-PE-Cy7 (BD Pharmingen, San Jose, CA, USA); and (iv) αCD27-PE-Cy5, αCD45RA-ECD and αCD45RO-ECD (Beckman Coulter, High Wycombe, UK). LIVE/DEAD Fixable Aqua and Violet Dead Cell Stain Kits (Life Technologies) were used to eliminate nonviable cells from the analysis. Peptide-major histocompatibility complex dextramers conjugated separately to APC and PE (Immudex, Copenhagen, Denmark) were used for magnetic enrichment.

Dextramer-based magnetic enrichment and multiparametric flow cytometry
Cryopreserved UCB cells and Dynabead cultures were thawed into warm R10 medium. For each donor, 10[5] cells were used to determine the percentage of CD8[+] T cells in the sample. These cell aliquots were labeled with LIVE/DEAD Fixable Violet for 5 min on ice, then stained with αCD8-PE, αCD14-Pacific Blue and αCD19-Pacific Blue for a further 30 min on ice. Data were acquired using a FACSCanto II flow cytometer and analyzed with FACSDiva v6.0 software (BD Biosciences). The percentage of live CD14[-] CD19[-] CD8[+] T cells in each sample was used in conjunction with total cell counts to determine the number of CD8[+] T cells that underwent magnetic enrichment. To enrich for epitope-specific CD8[+] T cells, samples were incubated in R10 medium containing 50 nM dasatinib (Axon Medchem, Groningen, Netherlands) for 30 min at 37 °C,[17] then labeled with PE-conjugated dextramer for 30 min on ice. After washing in magnetic-activated cell sorting (MACS) buffer (degassed phosphate-buffered saline containing 0.5% bovine serum albumin and 2 mM EDTA; all Sigma-Aldrich), cells were resuspended in 80 μl MACS buffer with 20 μl αPE microbeads (Miltenyi Biotec, Bergisch Gladbach, Germany) per 10[7] cells, incubated for 15 min on ice, washed again and resuspended in 2 ml MACS buffer. Each cell suspension was then passed through a separate magnetic separation (MS) column (Miltenyi Biotec). The columns were washed three times with 500 μl MACS buffer and removed from the magnetic field. Bound cells were eluted by pushing 2 ml MACS buffer through each column with a plunger. Enriched cells were centrifuged, labeled with LIVE/DEAD Fixable Aqua for 5 min on ice, then stained for 30 min on ice with a cocktail containing the following directly conjugated mAbs: αCD3-APC-H7, αCD8-QD705, αCD14-Pacific Blue, αCD19-Pacific Blue, αCD27-PE-Cy5, αCD45RA-ECD or αCD45RO-ECD, αCD57-FITC and αCCR7-PE-Cy7. After labeling, cells were washed with MACS buffer and acquired using a custom-modified FACSAria II flow cytometer (BD Biosciences). To isolate epitope-specific T cells, single lymphocytes were selected based on light scatter properties and dump[+] (nonviable, CD14[+] or CD19[+]) cells were gated out. From this population, up to 5000 dextramer[+] CD8[+] T cells were sorted at >98% purity directly into 1.5 ml screw-cap tubes (Sarstedt, Leicester, UK) containing 100 μl RNAlater (Life Technologies). The entire sample was acquired for each donor. Sorted cells in RNAlater were then centrifuged and frozen immediately at −80 °C for subsequent molecular analysis of TCR usage. Flow cytometric data were analyzed with FlowJo v7.6.5 software (TreeStar, Ashland, OR, USA). The frequency of dextramer[+] CD8[+] T cells in each sample was calculated by dividing the number of dextramer[+] CD8[+] T cells acquired on the FACSAria II by the total number of CD8[+] T cells passed through the MS column. Frequencies were then converted to the number of dextramer[+] CD8[+] T cells per 100 000 CD8[+] T cells.

Molecular analysis of TCR usage
TCR clonotyping was performed as described previously.[18] Briefly, all expressed *TRB* or *TRA* gene products were amplified without bias using a template-switch anchored PCR with reverse transcription incorporating a 3′ TRB constant region primer (5′-TGCTTCTGATGGCTCAAACACAGCGACCT-3′) or a 3′ TRA constant region primer (5′-AATAGGCAGACAGACTTGTCACTGGA-3′). Amplicons were subcloned, sampled, Sanger sequenced and analyzed as described previously.[41]

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

1   Allen S, Turner SJ, Bourges D, Gleeson PA, van Driel IR. Shaping the T-cell repertoire in the periphery. *Immunol Cell Biol* 2011; **89**: 60–69.
2   Moon JJ, Chu HH, Pepper M, McSorley SJ, Jameson SC, Kedl RM *et al.* Naive CD4+ T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* 2007; **27**: 203–213.
3   Jenkins MK, Moon JJ. The role of naive T cell precursor frequency and recruitment in dictating immune response magnitude. *J Immunol* 2012; **188**: 4135–4140.
4   La Gruta NL, Rothwell WT, Cukalac T, Swan NG, Valkenburg SA, Kedzierska K *et al.* Primary CTL response magnitude in mice is determined by the extent of naive T cell recruitment and subsequent clonal expansion. *J Clin Invest* 2010; **120**: 1885–1894.
5   Ruckwardt TJ, Malloy AM, Gostick E, Price DA, Dash P, McClaren JL *et al.* Neonatal CD8 T-cell hierarchy is distinct from adults and is influenced by intrinsic T cell properties in respiratory syncytial virus infected mice. *PLoS Pathog* 2011; **7**: e1002377.
6   Legoux F, Debeaupuis E, Echasserieau K, De La Salle H, Saulquin X, Bonneville M. Impact of TCR reactivity and HLA phenotype on naive CD8 T cell frequency in humans. *J Immunol* 2010; **184**: 6731–6738.
7   Alanio C, Lemaitre F, Law HK, Hasan M, Albert ML. Enumeration of human antigen-specific naive CD8+ T cells reveals conserved precursor frequencies. *Blood* 2010; **115**: 3718–3725.
8   Schmidt J, Neumann-Haefelin C, Altay T, Gostick E, Price DA, Lohmann V *et al.* Immunodominance of HLA-A2-restricted hepatitis C virus-specific CD8+ T cell responses is linked to naive-precursor frequency. *J Virol* 2011; **85**: 5232–5236.
9   Kwok WW, Tan V, Gillette L, Littell CT, Soltis MA, LaFond RB *et al.* Frequency of epitope-specific naive CD4+ T cells correlates with immunodominance in the human memory repertoire. *J Immunol* 2012; **188**: 2537–2544.
10 Su LF, Kidd BA, Han A, Kotzin JJ, Davis MM. Virus-specific CD4+ memory-phenotype T cells are abundant in unexposed adults. *Immunity* 2013; **38**: 373–383.
11 Gattinoni L, Lugli E, Ji Y, Pos Z, Paulos CM, Quigley MF *et al.* A human memory T cell subset with stem cell-like properties. *Nat Med* 2011; **17**: 1290–1297.
12 den Braber I, Mugwagwa T, Vrisekoop N, Westera L, Mogling R, de Boer AB *et al.* Maintenance of peripheral naive T cells is sustained by thymus output in mice but not humans. *Immunity* 2012; **36**: 288–297.
13 Hess RD. Routine Epstein-Barr virus diagnostics from the laboratory perspective: still challenging after 35 years. *J Clin Microbiol* 2004; **42**: 3381–3387.
14 Roback JD, Drew WL, Laycock ME, Todd D, Hillyer CD, Busch MP. CMV DNA is rarely detected in healthy blood donors using validated PCR assays. *Transfusion* 2003; **43**: 314–321.
15 Batard P, Peterson DA, Devevre E, Guillaume P, Cerottini JC, Rimoldi D *et al.* Dextramers: new generation of fluorescent MHC class I/peptide multimers for visualization of antigen-specific CD8+ T cells. *J Immunol Methods* 2006; **310**: 136–148.
16 Dolton G, Lissina A, Skowera A, Ladell K, Tungatt K, Jones E *et al.* Comparison of peptide-major histocompatibility complex tetramers and dextramers for the identification of antigen-specific T cells. *Clin Exp Immunol* 2014; **177**: 47–63.
17 Lissina A, Ladell K, Skowera A, Clement M, Edwards E, Seggewiss R *et al.* Protein kinase inhibitors substantially improve the physical detection of T-cells with peptide-MHC tetramers. *J Immunol Methods* 2009; **340**: 11–24.
18 Quigley MF, Almeida JR, Price DA, Douek DC. Unbiased molecular analysis of T cell receptor expression using template-switch anchored RT-PCR. *Curr Protoc Immunol* 2011; **Chapter 10**: Unit10 33.
19 Miles JJ, Douek DC, Price DA. Bias in the alphabeta T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunol Cell Biol* 2011; **89**: 375–387.
20 Venturi V, Price DA, Douek DC, Davenport MP. The molecular basis for public T-cell responses? *Nat Rev Immunol* 2008; **8**: 231–238.
21 Miles JJ, Thammanichanond D, Moneer S, Nivarthi UK, Kjer-Nielsen L, Tracy SL *et al.* Antigen-driven patterns of TCR bias are shared across diverse outcomes of human hepatitis C virus infection. *J Immunol* 2011; **186**: 901–912.
22 Boon AC, de Mutsert G, Graus YM, Fouchier RA, Sintnicolaas K, Osterhaus AD *et al.* The magnitude and specificity of influenza A virus-specific cytotoxic T-lymphocyte responses in humans is related to HLA-A and -B phenotype. *J Virol* 2002; **76**: 582–590.
23 Cole DK, Yuan F, Rizkallah PJ, Miles JJ, Gostick E, Price DA *et al.* Germ line-governed recognition of a cancer epitope by an immunodominant human T-cell receptor. *J Biol Chem* 2009; **284**: 27281–27289.
24 Pinto S, Sommermeyer D, Michel C, Wilde S, Schendel D, Uckert W *et al.* Mis-initiation of intrathymic MART-1 transcription and biased TCR usage explain the high frequency of MART-1-specific T cells. *Eur J Immunol* 2014; **44**: 2811–2821.
25 Stewart-Jones GB, McMichael AJ, Bell JI, Stuart DI, Jones EY. A structural basis for immunodominant human T cell receptor recognition. *Nat Immunol* 2003; **4**: 657–663.
26 Miles JJ, Bulek AM, Cole DK, Gostick E, Schauenburg AJ, Dolton G *et al.* Genetic and structural basis for selection of a ubiquitous T cell receptor deployed in Epstein-Barr virus infection. *PLoS Pathog* 2010; **6**: e1001198.
27 Kjer-Nielsen L, Clements CS, Purcell AW, Brooks AG, Whisstock JC, Burrows SR *et al.* A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity. *Immunity* 2003; **18**: 53–64.
28 Gras S, Saulquin X, Reiser JB, Debeaupuis E, Echasserieau K, Kissenpfennig A *et al.* Structural bases for the affinity-driven selection of a dominant human cytomegalovirus epitope. *J Immunol* 2009; **183**: 430–437.
29 Neller MA, Burrows JM, Rist MJ, Miles JJ, Burrows SR. High frequency of herpesvirus-specific clonotypes in the human T cell repertoire can remain stable over decades with minimal turnover. *J Virol* 2013; **87**: 697–700.
30 Iglesias MC, Briceno O, Gostick E, Moris A, Meaudre C, Price DA *et al.* Immunodominance of HLA-B27-restricted HIV KK10-specific CD8+ T-cells is not related to naive precursor frequency. *Immunol Lett* 2013; **149**: 119–122.
31 Szabolcs P, Park KD, Reese M, Marti L, Broadwater G, Kurtzberg J. Coexistent naive phenotype and higher cycling rate of cord blood T cells as compared to adult peripheral blood. *Exp Hematol* 2003; **31**: 708–714.
32 Cukalac T, Chadderton J, Zeng W, Cullen JG, Kan WT, Doherty PC *et al.* The influenza virus-specific CTL immunodominance hierarchy in mice is determined by the relative frequency of high-avidity T cells. *J Immunol* 2014; **192**: 4061–4068.
33 Price DA, Asher TE, Wilson NA, Nason MC, Brenchley JM, Metzler IS *et al.* Public clonotype usage identifies protective Gag-specific CD8+ T cell responses in SIV infection. *J Exp Med* 2009; **206**: 923–936.
34 Price DA, West SM, Betts MR, Ruff LE, Brenchley JM, Ambrozak DR *et al.* T cell receptor recognition motifs govern immune escape patterns in acute SIV infection. *Immunity* 2004; **21**: 793–803.
35 Chen H, Ndhlovu ZM, Liu D, Porter LC, Fang JW, Darko S *et al.* TCR clonotypes modulate the protective effect of HLA class I molecules in HIV-1 infection. *Nat Immunol* 2012; **13**: 691–700.
36 Ladell K, Hashimoto M, Iglesias MC, Wilmann PG, McLaren JE, Gras S *et al.* A molecular basis for the control of preimmune escape variants by HIV-specific CD8+ T cells. *Immunity* 2013; **38**: 425–436.
37 Ekeruche-Makinde J, Clement M, Cole DK, Edwards ES, Ladell K, Miles JJ *et al.* T-cell receptor-optimized peptide skewing of the T-cell repertoire can enhance antigen targeting. *J Biol Chem* 2012; **287**: 37269–37281.
38 Ekeruche-Makinde J, Miles JJ, van den Berg HA, Skowera A, Cole DK, Dolton G *et al.* Peptide length determines the outcome of TCR/peptide-MHCI engagement. *Blood* 2013; **121**: 1112–1123.
39 Pentier JM, Sewell AK, Miles JJ. Advances in T-cell epitope engineering. *Front Immunol* 2013; **4**: 133.
40 Neller MA, Sewell AK, Burrows SR, Miles JJ. Tracking the repertoire of human adult and neonatal T cells during *ex vivo* amplification. *Br J Haematol* 2012; **159**: 370–374.
41 Price DA, Brenchley JM, Ruff LE, Betts MR, Hill BJ, Roederer M *et al.* Avidity for antigen shapes clonal dominance in CD8+ T cell populations specific for persistent DNA viruses. *J Exp Med* 2005; **202**: 1349–1361.

The Supplementary Information that accompanies this paper is available on the Immunology and Cell Biology website (http://www.nature.com/icb)