MHC multimer-binding CD8 T cells:
Optimized detection, automated analysis and identification of antigen specific CD8 T cells in Narcolepsy type 1

PhD Thesis

Natasja Wulff Pedersen
October 2018

DTU Nanotech
Department of Micro- and Nanotechnology

MHC multimer-binding CD8 T cells: Optimized detection, automated analysis and identification of antigen specific CD8 T cells in Narcolepsy type 1

Natasja Wulff Pedersen

PhD thesis

Kongens Lyngby 2018

DTU

Front page art: Mette Pedersen

Preface

This thesis has been submitted to the Technical University of Denmark, Nanotech Institute, as part of the requirements to obtain the PhD degree. The presented research has been conducted first in the Department of Immunology and Vaccinology, at the Veterinary Institute and later at the Nanotech Institute, at the Technical University of Denmark. The work was supervised by Professor Sine Reker Hadrup.

The presented thesis consists of an introduction followed by three research manuscripts and an epilogue. Together they introduce, describe and discuss the research that I have carried out between October 2014 and October 2018.

Kongens Lyngby, October 23, 2018

Natasja Wulff Pedersen

Abstract

Antigen specific CD8 T cells are crucial mediators of specific immunity against intracellular infections and cancer and when dysregulated can be involved in autoimmune diseases. Instrumental to the function of these cells, is the recognition of antigen presented in the context of MHC molecules on the surface of target cells, by the T cell receptor on the CD8 T cells. Disease relevant interactions between TCRs and peptide-MHC are complex and often of low affinity and frequency, complicating the detection of antigen specific CD8 T cells. In this thesis three different aspects of describing antigen specific CD8 T cells are touched upon: optimized detection using MHC multimers, automated analysis of MHC multimer-binding T cells and identification of CD8 T cells with relevance to the autoimmune disease narcolepsy type 1, using DNA barcode labelled MHC multimers.

In manuscript I, the impact of using different fluorochromes to detect virus specific CD8 T cells of both low and high frequency in MHC multimer assays, was investigated. We found that even though there was no effect on the ability to detect the T cell responses in question, the separation of MHC multimer-binding T cells from background events was affected by fluorochrome choice. This may be important for successful detection of low-affinity CD8 T cell populations such as those reactive towards self. Furthermore, we propose a bead-based strategy for optimization of fluorescence detection that is useful not only for MHC multimer assays, but for flow cytometry experiments in general.

In paper II, the feasibility of using automated gating tools for analysis of MHC multimer binding T cells is shown. Increasing complexity of flow cytometry data has now surpassed our ability to analyze these data with standard manual gating strategies and has thus prompted the development of automated gating tools. The implementation of these tools in the broad scientific community has, however, been slow, mainly due to a gap in understanding of these tools. Therefore we tested the ability of three different algorithms, FLOCK, SWIFT and ReFlow to identify virus specific CD8 T cells of varying frequency. All tools performed well when detecting MHC multimer binding T cell populations with frequencies above 0.1% whereas SWIFT was the only tool to reliably detect populations below 0.02%. Furthermore, major challenges and obstacles for the integration of automated gating tools into the broad scientific community were identified and discussed.

In paper III, the newly developed state of the art technology of DNA barcode-labelled MHC multimers was used to detect neuron specific, self-reactive T cells with

relevance to narcolepsy type 1 (NT1). NT1 is a chronic, debilitating neurological sleep disorder. It is caused by the loss of neurons in the brain that produce the neuropeptide hypocretin, which is involved in regulation of wakefulness. The loss of hypocretin neurons is thought to be the result of an autoimmune attack, although this has never been indisputably proven. Rather, the autoimmune hypothesis is based on circumstantial evidence such as the very strong correlation between NT1 and the HLA class II allele DQB1*06:02 and the increased NT1 incidence that was observed after H1N1 influenza vaccination in several countries. In our study, we investigate the ability of CD8 T cells from NT1 patients and healthy controls to recognize a library of 1183 peptides from 7 proteins expressed by hypocretin neurons that are restricted to 8 different HLA types. What we find, is a broad presence of neuron-specific CD8 T cells in blood samples from NT1 patients but also from healthy controls. We observed a difference in the number and frequency of these neuron-specific CD8 T cells between NT1 patients and healthy controls positive for HLA-DQB1*06:02, indicating that the combination of expressing the risk HLA allele and also harboring a certain level of auto-reactive CD8 T cells might be important for disease development. Even though these findings do not provide conclusive evidence for the autoimmune hypothesis of NT1, they do for the first time show the existence of hypocretin neuron-specific CD8 T cells and are one step on the way to full elucidation of the pathogenesis of narcolepsy type 1.

Dansk resumé

Antigen specifikke CD8 T-celler er afgørende komponenter i den specifikke immunitet overfor intracellulære infektioner og kræft, og kan under dysfunktionel regulering spille en rolle i autoimmunitet. Afgørende for disse cellers funktion, er genkendelsen mellem T-celle receptoren (TCR) på overfladen af CD8 T-celler og deres antigener som præsenteres på overfladen af target celler i sammenhæng med MHC molekyler. Sygdomsrelevante interaktioner mellem TCR og peptid-MHC er komplekse og ofte af lav affinitet og frekvens, hvilket komplicerer identifikationen af antigen specifikke CD8 T-celler. I denne afhandling berøres tre forskellige aspekter af beskrivelsen af antigen specifikke CD8 T celler: optimeret detektion ved brug af MHC multimerer, automatiseret analyse af MHC multimer-bindende T-celler og identifikation af CD8 T-celler med relevans for den autoimmune sygdom narkolepsi type 1, ved hjælp af MHC multimerer opmærket med DNA stregkoder.

I manuskript I, undersøges virkningen af at anvende forskellige fluorokromer til påvisning af virusspecifikke CD8 T-celler af både lav og høj frekvens i MHC multimer assays. Vi fandt, at selvom der ikke var nogen effekt på evnen til at detektere de pågældende T-celle responser, havde fluorokrom-valg indflydelse på separationen mellem MHC multimer-bindende T-celler og baggrundspopulationen. Dette kan være specielt vigtigt for vellykket detektion af CD8 T-cellepopulationer med lav affinitet, såsom dem der er reaktive mod selv. Derudover foreslår vi en bead-baseret strategi til optimering af fluorescensdetektering, der ikke kun er nyttig for MHC multimer assays, men for flowcytometri eksperimenter generelt.

I artikel II vises anvendelsen af automatiserede gating værktøjer til analyse af MHC multimer-bindende T-celler. Øget kompleksitet af flowcytometri data har nu overgået vores evne til at analysere disse data med standard manuelle gating strategier og har således afstedkommet udviklingen af automatiserede gating værktøjer. Implementeringen af disse værktøjer i det brede videnskabelige samfund har imidlertid været langsom, hovedsageligt på grund af mangelfuld forståelse af deres egenskaber. Vi testede derfor brugen af tre forskellige algoritmer, FLOCK, SWIFT og ReFlow, og deres evne til at identificere virusspecifikke CD8 T-celler med varierende frekvens. Detektion af MHC multimer-bindende T-cellepopulationer med frekvenser over 0,1% var muligt med alle gating værktøjer, mens SWIFT var den eneste algoritme der pålideligt detekterede populationer under 0,02%. Derudover blev de største udfordringer og forhindringer for integrationen af automatiserede gating værktøjer i det brede videnskabelige samfund udpeget og diskuteret.

I manuskript III, identificeres neuron specifikke, selv-reaktive CD8 T celler med relevans for narkolepsi type 1 (NT1), ved hjælp af MHC multimerer opmærket med DNA stregkoder. NT1 er en kronisk, invaliderende neurologisk søvn sygdom. Den er forårsaget af tabet af neuroner i hjernen der producerer neuropeptidet hypokretin, som er involveret i reguleringen af vågenhed. Tabet af hypocretin-neuroner antages at være resultatet af et autoimmunt angreb, selvom dette aldrig er blevet endegyldigt bevist. Den autoimmune hypotese er snarere baseret på en række indicier såsom den stærke sammenhæng mellem NT1 og HLA klasse II vævstypen DQB1*06:02 og den øgede forekomst af NT1 som blev observeret efter H1N1 influenza vaccination i flere lande. I vores studie undersøger vi CD8 T-celler fra NT1 patienter og raske kontroller og deres evne til at genkende et bibliotek af 1183 peptider fra 7 proteiner der er udtrykt i hypokretin neuroner og som binder til 8 forskellige HLA-typer. Vi finder en bred forekomst af neuron-specifikke CD8 T-celler i blodprøver fra NT1 patienter, men også fra raske kontroller. Ydermere observerede vi en forskel i antallet og frekvensen af disse neuron-specifikke CD8 T-celler mellem NT1 patienter og raske kontroller som udtrykker HLA-DQB1*06:02, hvilket kunne tyde på at en kombination af at udtrykke denne risiko-vævstype og samtidig have et vist niveau af selv-reaktive CD8 T-celler kan være vigtigt for udviklingen af NT1. Selvom vores observationer ikke giver et afgørende bevis for den autoimmune hypotese for udviklingen af NT1, viser de for første gang tilstedeværelsen af hypokretin neuron-specifikke CD8 T-celler og udgør et skridt på vejen mod fuldt ud at belyse den patologiske process der fører til narkolepsi type 1.

Acknowledgements

Publications

Papers/Manuscripts included in the thesis

I **Natasja Wulff Pedersen**, Karoline Laske, Dominik Maurer, Steffen Walter, Cecile Gouttefangeas and Sine Reker Hadrup. Optimization in detection of antigen specific T cells through differentially-labelled MHC multimers. *Manuscript in preparation.*

II **Natasja Wulff Pedersen**, P. Anoop Chandran, Yu Qian, Jonathan Rebahn, Nadia Viborg Petersen, Mathilde Dalsgaard Hoff, Scott White, Alexandra J. Lee, Rick Stanton, Charlotte Halgreen, Kivin Jakobsen, Tim Mossmann, Cecile Gouttefangeas, Cliburn Chan, Richard H. Scheuermann and Sine Reker Hadrup. "Automated Analysis of Flow Cytometry Data to Reduce Inter-Lab Variation in the Detection of MHC Multimer Binding T cells." *Front Immunol* **26** 8:858 (2017 Jul).

III **Natasja Wulff Pedersen**, Anja Holm, Nikolaj Pagh Kristensen, Anne-Mette Bjerregaard, Amalie Kai Bentzen,Tripti Tamhane, Kristoffer Sølvsten Burgdorf, Henrik Ullum, Poul Jennum, Stine Knudsen, Sine Reker Hadrup* and Birgitte R. Kornum*. "Broad presence of autoreactive CD8 T cells against narcolepsy type 1-relevant proteins." *In review in Nat Commun, October 2018.*

Papers not included in the thesis

- Rikke Andersen, Marie Christine Wulff Westergaard, Julie Westerlin Kjeldsen, Anja Müller, **Natasja Wulff Pedersen**, Sine Reker Hadrup, Özcan Met, Barbara Seliger, Bjarne Kromann-Andersen, Thomas Hasselager, Marco Donia and Inge Marie Svane. "T-cell Responses in the Microenvironment of Primary Renal Cell Carcinoma – Implications for Adoptive Cell Therapy." *Cancer Immunol Res* **6** 2:222-235 (2018 Feb)

Contents

Introduction

The immune system is an incredibly complex organ comprised of numerous cell types divided into different compartments of innate and adaptive immunity. CD8 T cells are an essential part of the adaptive immune system and their main function is to specifically recognize and react to pathogens that reside inside cells of the body. Through their T cell receptor (TCR), the CD8 T cells bind to their specific peptide antigen presented on the surface of cells, in the context of Major Histocompatibility Complex (MHC) class I molecules, in humans called Human Leukocyte Antigen (HLA). When activated upon recognition of their cognate antigen, they have the capacity to efficiently kill cells that express this antigen and are therefore also referred to as cytotoxic T cells. As CD8 T cells are the main effector cells responsible for protection against intracellular pathogens or malignant intracellular changes, they play crucial roles in fighting infectious diseases and cancer. Furthermore, they are also known to be involved in autoimmune pathogenesis where dysregulation of CD8 T cell responses can turn these powerful cells against their host, causing disease (figure 1.1).

Pivotal to the function of CD8 T cells is their specific recognition of antigen. Thus, in order to understand disease mechanisms and to develop new therapeutic strategies, it is crucial to investigate the specific interaction between the TCR and the peptide-MHC complex (pMHC). This specific interaction and the detection of it is the focus of the present PhD thesis.

1.1 The diversity of the TCR-pMHC-I interaction

There are three components of the TCR-pMHC-I interaction that contribute to its diversity and complexity, the MHC molecule, the presented peptide and the TCR.

1.1.1 HLA diversity

The genes coding for HLA-I are among the most polymorphic of the human genome. Each individual expresses 6 HLA-I alleles and with >10.000 different HLA molecules identified, the chance of being homozygous for any of them is very small (J. Robinson *et al.,* 2003) (figure 1.2). Each HLA molecule has distinct preferences for given amino acids in their peptide binding groove, so called anchor residues. The position of these anchor residues differs between HLA molecules but they are often located at
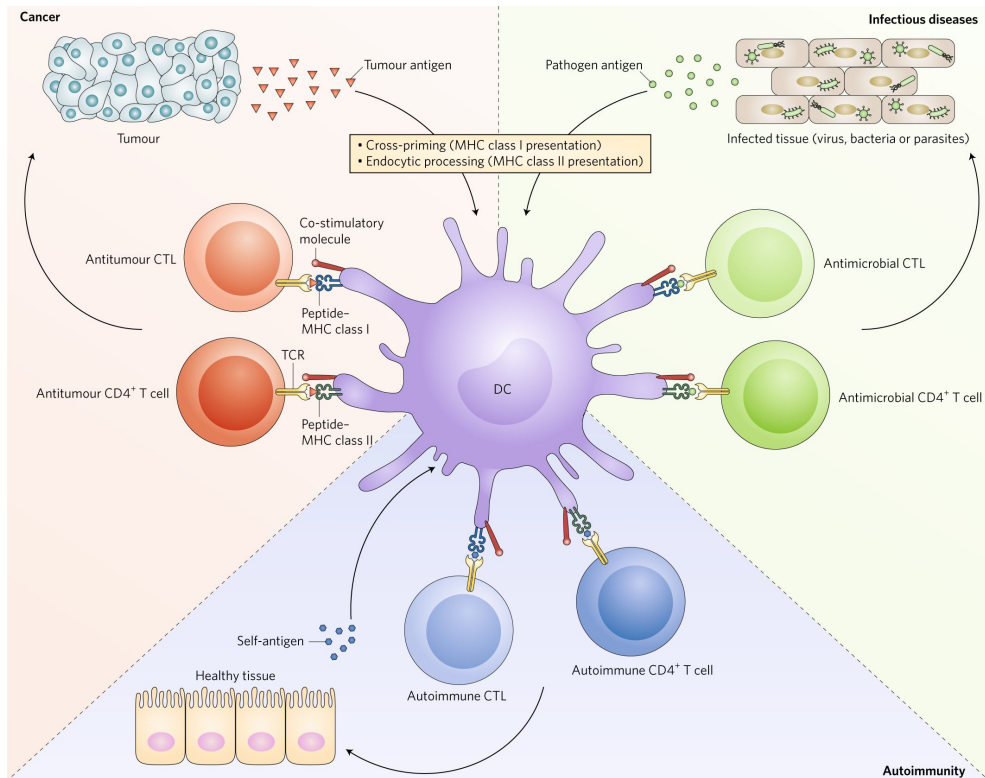
Figure 1.1: CD8 T cells, also called cytotoxic T lymphocytes (CTLs) and CD4 T cells are part of the specific response against infection and cancer and also play a role in autoimmunity. They recognize their antigen in the context of MHC class I and class II molecules respectively and upon activation exert their effector functions on the target tissue. Adapted from (Hadrup & Newell, 2017).

positions 2 and 9 in the peptide. They are important for binding to the HLA molecule whereas the remaining residues within the peptide are facing out, towards the TCR. This means that each HLA allele will present different pools of peptides with distinct characteristics, although substantial overlap does occur (Sidney *et al.*, 2008). The polymorphism of the HLA genes ensures that even if a pathogen evolves to escape presentation to the immune system by one HLA molecule, another one is there to do the job. Thus, the population as a whole is protected.
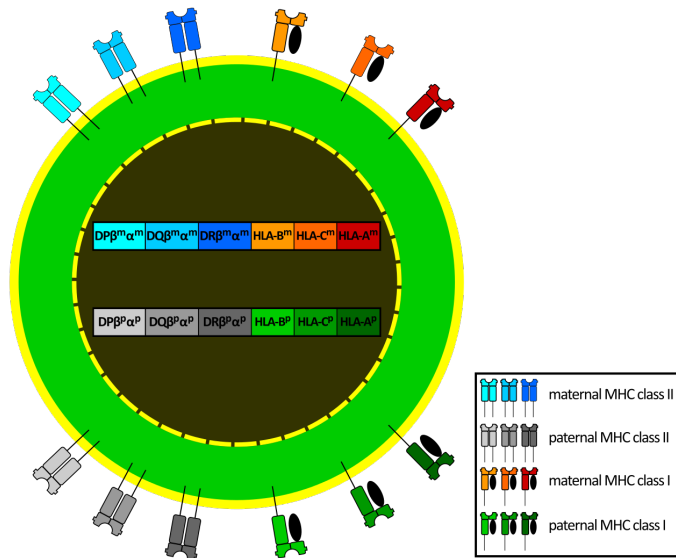
Figure 1.2: The diversity of the HLA molecule. Each individual inherits 3 different HLA class I alleles, HLA-A, -B and -C, and 3 different HLA class II alleles, HLA-DP, -DQ and -DR, from both their father and mother. 6 of each class of HLA molecules are thus expresses by each individual. From `https://en.wikipedia.org/wiki/Human_leukocyte_antigen`.

1.1.2 Peptide presentation

20 different amino acids constitute the building blocks for all proteins, foreign or self. Theoretically, $>10^{15}$ potential different peptides with a length that facilitates binding to the MHC-I, can be produced from these 20 amino acids (Sewell, 2012). This might even be underestimated as this number only relies on the possible combinations of the 20 amino acids and does not take into account post-translational modifications such as phosphorylation (Mohammed *et al.,* 2008) and citrullination (Scally *et al.,* 2013) that can alter the presented peptides. It is estimated that ~1% of all potential peptides meet the requirements needed to bind to an MHC molecule (Nielsen *et al.,* 2007; Yewdell & Bennink, 1999) and the actual number of peptides that the immune system might encounter is thus reduced, although still enormous. In the context of diseases involving human proteins, this number is substantially lower as the human proteome consists of around 20.000 proteins (Omenn *et al.,* 2016). The number of potential human derived peptides is approximately the same as the number of amino acids in the human proteome, as a peptide can start at any position along the length of the protein. If we consider an average human protein to be 1000 amino acids long, the number of potential human derived peptides is 20.000 x 1000 = 20×10^6. Only

~1% of these will be able to bind to an MHC molecule and the number of peptides that can potentially be presented in the human body is hence ~2x10^5 per MHC-I molecule.

1.1.3   TCR diversity

The complicated task of T cells in the immune system is to be able to respond to any given infection an individual might encounter. Thus, all individuals must as a default have the T cells required to fight any potential pathogen. Two important features of T cells make this possible, namely the specificity and cross-reactivity of the TCR. The specificity of the TCR arises from the random rearrangement of V(D)J gene segments in the complementarity-determining regions (CDR), of which there are 3 (figure 1.3). CDR3 is the part of the TCR that interacts with the peptide bound to the pMHC complex and is also the most variable, whereas CDR1 and 2 bind to the MHC molecule (Sewell, 2012). It has been estimated that ~10^7 different TCRs exist in an individual (Arstila *et al.,* 1999), and even though this is an enormous amount it is far from covering the >10^{15} potential foreign peptides which is estimated to exist (Sewell, 2012). In order to provide protection against all these epitopes, T cells must be cross-reactive, and it has been estimated that a single TCR has the potential to recognize up to 106 different peptide antigens (Wooldridge *et al.,* 2012).
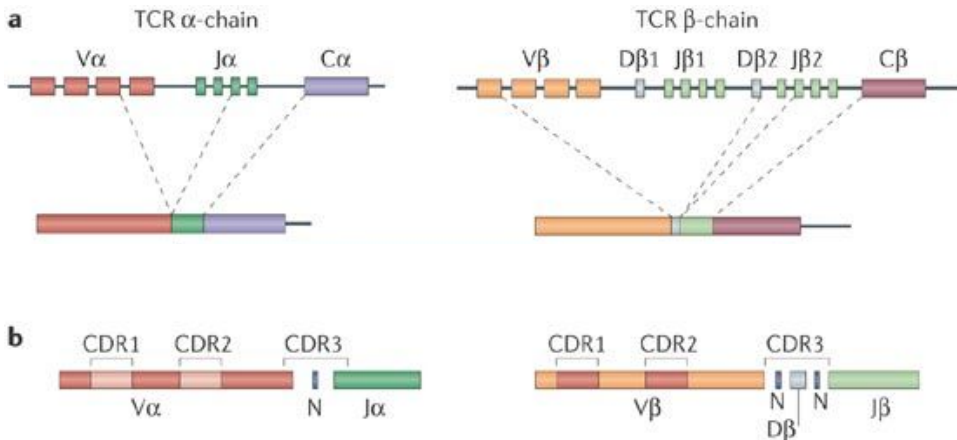
Figure 1.3: Joining of variable (V), diversity (D) and junctional (J) gene segments of the β chain and V and J gene segments of the α chain is responsible for the diversity of the TCR. In humans, there are 42 Vβ, 2 Dβ, 12 Jβ, 43 Vα and 58 Jα functional gene segments. The TCR is formed by recombination of these genes and subsequent splicing to a constant region (C). Complementarity-determining regions (CDR) 1 and 2 are encoded within the V gene segments whereas CDR3 spans parts of both the V(D)J segments. Modified from (Turner *et al.,* 2006).

The immense complexity of the interaction between CD8 T cells and their HLA-restricted antigen targets arises from the combined diversity of the TCR, HLA molecules and peptides presented. It is obvious, that if we want to get just a little glimpse into the this complexity, high-throughput screening methods are needed. Consequently, over the past decade researchers have sought to develop strategies for more comprehensive analyses of T cell recognition. Besides from the complexity described above, additionally challenging for these efforts, is the low affinity for pMHC which is often a feature of antigen specific T cells, as well as their low frequency. This is true, at least, when investigating auto- and shared cancer-antigen reactive T cells which will typically have lower affinity to self-pMHC molecules than virus specific CD8 T cells and might only represent a small fraction of the total T cell pool.

1.2   Detection of antigen specific CD8 T cells

The affinity of a single TCR for its specific pMHC target is not sufficient for stable binding, and the successful interaction between TCR and pMHC relies on the simultaneous binding of multiple molecules, increasing the overall avidity of the interaction (figure 1.4).
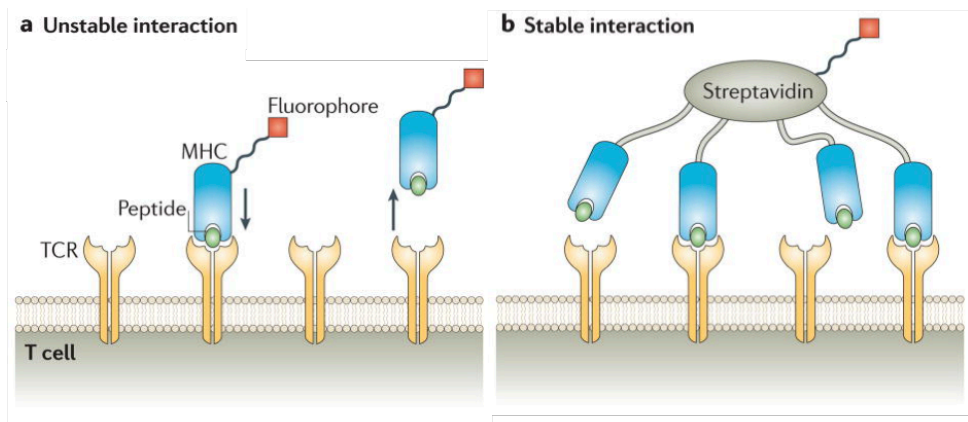
Figure 1.4: Single interactions between TCR and pMHC are unstable and not sufficiently strong to allow fluorescence based detection, whereas the simultaneous binding between several molecules increases the avidity of the interaction enough for detection to be possible. Modified from (Davis *et al.,* 2011).

Consequently, it was not until pMHC molecules were successfully multimerized around 20 years ago, that they were successfully used to detect antigen specific CD8 T cells (Altman *et al.,* 1996). MHC molecules were refolded with peptides of interest,

biotinylated and tetramers were formed by binding four identical MHC monomers to a streptavidin-conjugated fluorochrome. Using the fluorescent label, CD8 T cells that had recognized the pMHC complex and bound the tetramer could be identified by flow cytometry. This strategy offered an attractive alternative to conventional methods of detecting antigen specific T cells, as for instance the elispot assay, limiting dilution assay and intracellular cytokine staining, that all rely on the functionality of antigen specific T cells. Though these methods provide an important insight into the functional capacity of antigen specific T cells they do not provide information about the specific interaction between a T cell and its antigen.

Since its development, MHC multimer staining has become a widespread method for detection of antigen specific T cells and it has provided a great deal of new insights into the properties of these cells (Appay *et al.,* 2002; Daniels *et al.,* 2001; Lee *et al.,* 1999; Savage & Davis, 2001). The basic principle of the MHC multimer technology, developed over 20 years ago, is still used today but a number of improvements have been introduced to the method.

1.2.1   UV mediated MHC multimer generation

The requirement for generation of large pMHC libraries fostered the development of the MHC multimer assay. A major step in this direction was the introduction of the UV sensitive conditional ligand which offered an alternative to the laborious process of refolding MHC monomers with every individual peptide of interest as done initially (Rodenko *et al.,* 2006; Toebes *et al.,* 2006). UV mediated peptide exchange takes advantage of the unstable nature of the MHC molecule in the absence of a peptide in the binding groove. Upon UV exposure, the UV sensitive ligand is cleaved, leaving the MHC monomers to disintegrate. When the process occurs in the presence of another peptide, the new peptide will rescue the complex, producing the desired pMHC. This approach made high-throughput generation of large libraries of unique pMHC complexes possible. When the initial limiting step of refolding pMHC was overcome, the next challenge for high-throughput screening for antigen specific T cells, became the number of possible fluorescence labels and flow cytometer detection channels available.

1.2.2   Polychromatic flow cytometry

Over the past 30 years there has been a huge development in flow cytometry instruments and a drastic increase in the amount of new fluorochromes available. During the 1970's and 80's, 2 and 3 colors were routinely used to detect immune cells by flow cytometry, but through the 90's, 2000's and early 2010's, with the introduction of tandem dyes, Quantum Dots (Qdots) and Brilliant violet dyes respectively, alongside development in instrument complexity, the number of colors that could be used simultaneously rose to 12-18 (Chattopadhyay & Roederer, 2012). Running experiments with 12-18 colors is now standard in most labs that use polychromatic flow cytometry.

Nevertheless, efforts are continuously being put into increasing the complexity of flow cytometry experiments and this is achieved primarily via the expansion of the family of brilliant dyes and development of new instruments that have enough lasers to detect up to 50 colors simultaneously. Detecting 50 colors simultaneously will likely be possible in the very near future, but at the current stage of development state-of-the-art is detection of up to 30 colors in a single experiment (Chattopadhyay, Giles, *et al.,* 2017); significantly changing the questions that can be asked and hopefully answered about our immune system.

An alternative method for high dimensional single cell detection is mass cytometry, also called cytometry by Time-of-Flight (CyTOF), which was developed recently (Bandura *et al.,* 2009). The CyTOF method is similar to flow cytometry except antibodies for cell staining are labeled with metal tags instead of fluorochromes and the stained cells are run through a mass spectrometer for detection. This strategy enables identification of ~40 parameters and offers some advantages over flow cytometry, such as practically no overlap between metal tags as is seen between fluorochromes. However, there are also some drawbacks that might make the CyTOF less attractive. Among others, slow collection speed, low cell recovery rate, high cost of instrument and reagents and the fact that the cells are destroyed in the analysis, are worth mentioning (Spitzer & Nolan, 2016).

1.2.3   Multiplex detection of MHC multimer-binding T cells

With the introduction of new fluorochromes, it became possible to investigate an increasing number of antigen specific T cells in a single sample. Still, in the case where 10 different streptavidin conjugated fluorochromes are used for MHC multimer reagents, only 10 different antigen specificities can be investigated simultaneously. Therefore, if a library of potential antigens consists of several hundred peptides, donor material will have to be split in multiple fractions in order for T cell binding to each peptide to be investigated. This is often not possible or requires heavy culturing of the sample as patient material is typically limited.

As a means of increasing the complexity of antigen specific T cell detection, combinatorial encoding of the MHC multimers was developed (Andersen, Kvistborg, *et al.,* 2012; Hadrup, Bakker, *et al.,* 2009; Newell, L. O. Klein, *et al.,* 2009). In this strategy, dual (Hadrup, Bakker, *et al.,* 2009) or multiple (Newell, L. O. Klein, *et al.,* 2009) color codes are used to identify each antigen, enabling reuse of the same fluorochromes multiple times, each time in a new combination. This substantially increases the number of peptides that can be investigated simultaneously as the same 10 fluorochromes mentioned above, give rise to 45 different combinations that can be used in a single sample when just using dual color codes. Using varying numbers of fluorochromes, many studies of antigen specific CD8 T cells have been conducted using this method, especially in cancer (Andersen, Thrue, *et al.,* 2012; Kvistborg, Philips, *et al.,* 2014; Mc Granahan *et al.,* 2016; Rizvi *et al.,* 2016) but also in autoimmune diseases (Sabatino *et al.,* 2018; Unger *et al.,* 2011; Velthuis *et al.,* 2010).

Even though many different fluorochromes are now routinely used to detect anti-gen specific CD8 T cells, it has never been tested in-depth across T cell populations of different frequency and avidity, whether the choice of fluorochrome has an impact on the detection of MHC multimer-binding T cells. It is evident, that with the varying brightness of different fluorochromes, the separation between positive and negative cells will also vary. It is an obvious thought, that the risk of missing TCR-pMHC interactions, especially those with low avidity, would be higher the dimmer the fluorochrome used. And even though individual labs presumably test and carefully choose streptavidin conjugated fluorochromes to use in their assays, the scientific community as a whole does not have a standardized way of testing this.

Manuscript I

In this study we investigate the effect of using different fluorochromes to detect antigen specific CD8 T cells through fluorescence-labeled MHC multimers. We test the use of two bright fluorochromes – Phycoerythrin (PE) and Allophycocyanin (APC) – and two dimmer fluorochromes – Quantum Dot 605 and 705 (Qdot605 and Qdot705), and ask whether the difference in signal intensity between these fluorochromes has an impact on the ability to detect both high frequency and low frequency virus specific CD8 T cell responses. Furthermore, we test and propose a simple bead based strategy for optimization of fluorescence detection in flow cytometry instruments.

1.3   Analysis of MHC multimer-binding T cells

Analysis of flow cytometry data is done through a process called gating. Based on visual inspection of fluorochrome intensity displayed by the analyzed cells on histograms or 2D dot plots, researchers make decisions about where to distinguish between positive and negative cells for the given parameter as well as which subsets of cells to choose for further inspection in other parameters. As is the case for all human decisions, gating decisions are a matter of individual interpretation and are thus very likely to differ between different individuals. Furthermore, manual gating also often relies on what is already known, thus what a researcher might expect to find and therefore think to look for. This inherent subjectivity to the analysis of flow cytometry data is a huge obstacle for standardization and reproducibility of scientific data obtained with this method (Maecker, McCoy, *et al.,* 2012; Maecker, Rinfret, *et al.,* 2005; Mair *et al.,* 2016; White *et al.,* 2014). This is also true for the analysis of antigen specific T cells detected through MHC multimer assays. In line with this, a proficiency panel investigating the impact of manual gating on the detection of virus specific CD8 T cells, found gating to be a source of variation to the final results

(Gouttefangeas *et al.,* 2015). This was especially evident when identifying MHC multimer-binding T cells of low frequency.

1.3.1 Automated analysis

In the past 10 years, substantial efforts have been put into developing tools for automated analysis of flow cytometry data. These efforts were in part driven by the unmet need for a more objective and standardized analysis strategy. And in part by the need for new analysis tools as the complexity of flow cytometry data increased with the emergence of polychromatic flow cytometry - defined as detection of 5 or more colors (Chattopadhyay & Roederer, 2012) - as the standard in flow cytometry (Aghaeepour, Finak, *et al.,* 2013; Lugli *et al.,* 2010). Considering an antibody panel for polychromatic flow cytometry with just 10 markers that could all be expressed by the same cells, it is obvious that analyzing the 1024 unique combinations of these markers that a cell can potentially express, is both extremely time consuming and perhaps not even possible with conventional 2D manual gating. As more parameters are included the number of possible combinations grows exponentially (Lugli *et al.,* 2010). 10 or more colors are now routinely used by most research groups, thereby urging the development of new analysis tools in order for researchers to exploit the full potential that the technical advances in instrument and reagent development has provided.

Several different tools for analysis of flow cytometry data have been developed. Common to them all is that they utilize machine learning in the attempt to solve the issues addressed above. Machine learning is a field of computer science where algorithms "learn" based on examples and then transfer that "knowledge" to new data.

1.3.1.1 Supervised vs unsupervised

There are two main categories of automated gating tools: supervised and unsupervised (Saeys *et al.,* 2016; L. M. Weber & M. D. Robinson, 2016). Supervised gating algorithms are trained with data where the variables that describe the data are labeled, e.g. cells that express certain markers are labeled as belonging to a given cell type and the algorithm will learn to recognize and define this cell type again in a new data set. Another example could be that samples in an experiment might belong to two different groups, e.g. healthy and sick, and the algorithm will learn to distinguish these two groups from each other. When given new data, it will predict the classification of new samples into either of the two groups and can also identify cell populations that are correlated with this classification. Unsupervised gating algorithms are given unlabeled data and work by recognizing patterns in the data. This means that an unsupervised algorithm will sort the cells based on a pattern of similarity and give as an output all the different populations that it was able to separate (figure 1.5). As a consequence, unsupervised gating algorithms are explorative and can find patterns

in the data that are not necessarily visible to the human eye, whereas supervised algorithms will find only what the researcher will ask it to look for (Bashashati & Brinkman, 2009; Saeys *et al.,* 2016; L. M. Weber & M. D. Robinson, 2016).
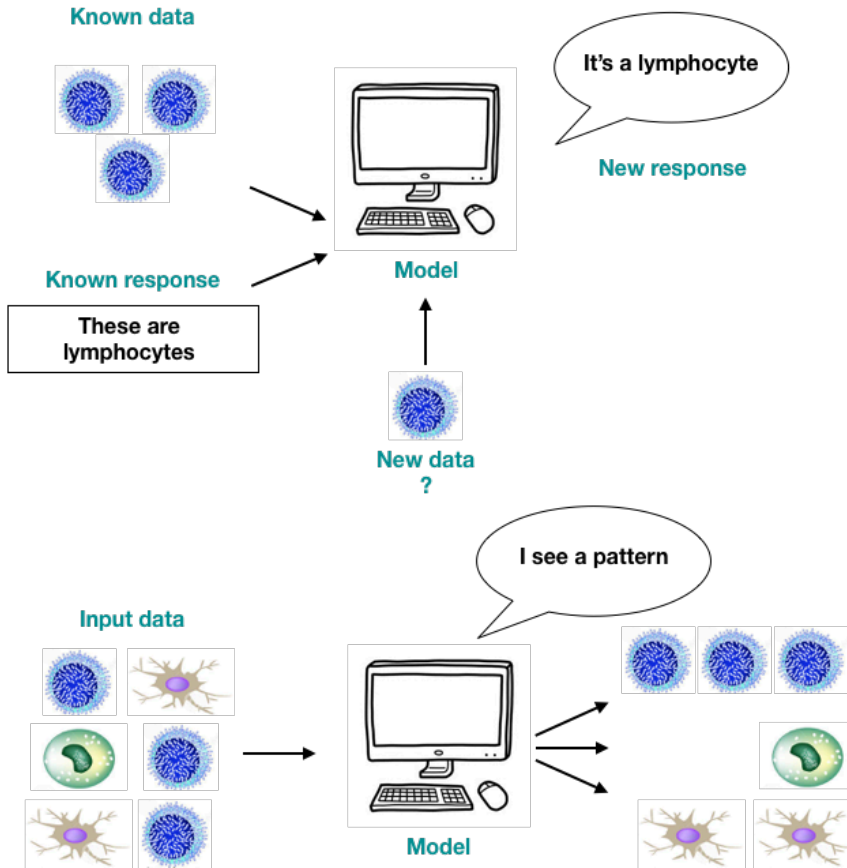
Figure 1.5: Supervised versus unsupervised learning. Supervised gating algorithms learn to recognize a certain cell type and will be able to identify this cell again in new unknown data. Unsupervised gating algorithms recognize patterns in the data and separate cells into clusters based on their similarity.

Owing to the explorative nature of unsupervised algorithms, they cannot be used for quantification of data, thus cannot tell whether observed differences between different samples are statistically significant (Lugli *et al.,* 2010).

The automated gating tools that have been developed during the past 10 years can grossly be divided into three different groups based on whether they are used for data

visualization, population identification or biomarker discovery (Saeys *et al.,* 2016). Automated algorithms for data visualization and population identification both rely on unsupervised approaches while biomarker discovery tools combine the unsupervised clustering of cell populations with supervised approaches that then correlate these populations to different parameters as for example disease status or survival.

1.3.1.2   Visualization

Some of the tools that have been most widely adopted by immunologists are visualization tools like SPADE (Qiu *et al.,* 2011) and t-SNE (Amir *et al.,* 2013; van der Maaten & Hinton, 2008). These tools offer a good alternative to manual gating for an initial inspection of the data and provide an overview of the cell populations that are present in the samples. Both tools provide a representation of the high-dimensional data in a low-dimensional space and separate the data based on similarity between cells present in the samples. SPADE does this for clusters of cells while t-SNE produces scatter plots of single cells where the distance between the cells reflects their similarity in the multidimensional space. This means that cells with similar properties are close to each other in the plot. Both SPADE and t-SNE provides the option of visualizing the expression level of the markers included in the analysis and overlaying this information on the 2D representation of the data by a color gradient. Because the SPADE output is given as entire cell populations, the average expression for each cell cluster is shown with this tool whereas t-SNE visualizes the expression at a single cell level.

   Common to both tools is the need to manually annotate the cell clusters found by the algorithms which is based on prior knowledge about the properties of known cell populations (Mair *et al.,* 2016; Saeys *et al.,* 2016) As mentioned above, the automated visualization tools are by far the most widespread of all automated analysis tools and have been used in several studies of the immune system. This goes for both SPADE (AC *et al.,* 2017; Bendall *et al.,* 2012; Qiu *et al.,* 2011) and t-SNE (Alcántara-Hernández *et al.,* 2017; Amir *et al.,* 2013; Becher *et al.,* 2014; Hartmann *et al.,* 2016; Lin, Frelinger, *et al.,* 2015). Most of these studies utilize Mass Cytometry, which enables identification of ~40 parameters at a time and so the need for automated approaches to CyTOF data analysis is even more pronounced (Bandura *et al.,* 2009; Newell, Sigal, *et al.,* 2013).

1.3.1.3   Population identification

While automated visualization of high-dimensional flow cytometry data is a necessary and helpful first step in analysis of complex data, it does not solve the issue of poor standardization and reproducibility in detection of antigen specific T cells using MHC multimers. The challenges for these experiments are perhaps not so much related to the complexity of the data, but rather to the difficulty in consistently identifying these often very rare populations, that in addition can be difficult to separate from background cells depending on the affinity of the pMHC-TCR interactions. Instead,

algorithms for automated population identification might offer solutions to these issues.

Multiple different approaches have been taken to solve the problem of automated assignment of cells into clusters or cell populations with similar properties in multidimensional space. Several tools build on various mathematical models in order to fit the data into uniform clusters, while others use different approaches to density based clustering. Common to most of them is that the algorithms go through multiple rounds of clustering until all uniform populations are identified. However, cells within a population might differ slightly on various markers, and the algorithms could in principle keep separating cells into clusters until each cluster consists of just a single cell. Therefore, most unsupervised population identification tools need information about how many different populations are expected to be in the data, so that they know how many rounds of clustering to go through. Even though some algorithms have ways of automatically identifying the number of populations, this information needs to be given manually to most tools. It is generally advised to over-cluster in order to make sure that small populations are identified as opposed to under-clustering and consequently miss rare populations, even if this means that a biologically meaningful population might be split (Saeys *et al.,* 2016).

A few examples of gating algorithms that have proven to perform well when challenged with different datasets and tasks are FLOCK (Qian *et al.,* 2010), flowMeans (Aghaeepour, Nikolic, *et al.,* 2011), flowSOM (Van Gassen *et al.,* 2015) and SWIFT (Naim *et al.,* 2014; Mosmann *et al.,* 2014), the latter two also performing well when attempting to identify rare cell populations (Aghaeepour, Finak, *et al.,* 2013; L. M. Weber & M. D. Robinson, 2016). Some tools are available with user friendly interfaces, such as FLOCK, which is a part of the Immunology Database and Analysis Portal, Immport, or the analysis framework ReFlow, which serves as both a data repository and an automated analysis platform (White *et al.,* 2014). This is, however, not the case for the majority of gating tools, and while user friendly interfaces are offered for a few, most algorithms including both flowMeans, flowSOM and SWIFT are operated through programming software, R and MatLab respectively. R especially harbours many tools, but using this interface poses a challenge for immunologists with no prior training in programming.

Paper II

Paper II is an evaluation of the ability of a few chosen automated gating tools to detect MHC binding antigen specific CD8 T cells. Automated gating tools have been tested in various competitions for their ability to perform different gating tasks, but have never been tested in MHC multimer staining experiments. Furthermore, we evaluate the feasibility of these tools to be used by an immunologist with no prior knowledge about programming and computer science. We test three tools, FLOCK, SWIFT and ReFlow which were chosen based on having a user friendly interface or requiring only a minimum of computational skills, or the reported ability to detect rare events. The ability of the tools to reduce variation to the analysis of MHC multimer detection was investigated using a proficiency panel data set with results from 16 different labs.

Common to all of the algorithms for automated population identification is the slow adoption of these tools by the broad scientific community, despite efforts to enhance awareness about the tools and their advantages, described in high impact scientific journals (Aghaeepour, Finak, *et al.,* 2013; Kvistborg, Gouttefangeas, *et al.,* 2015; Mair *et al.,* 2016; Saeys *et al.,* 2016). Besides from the obvious challenge of having to learn programming language to operate many of the tools, one possible explanation could be that even with these automated clustering tools, interpretation of the output can be difficult. This is due to the fact that they still require a manual annotation of what each identified cell cluster represents and additionally, these clusters are often not visualized in a way that offers easy and intuitive interpretation by non-computational experts.

1.3.1.4 Biomarker discovery

Another output interpretation issue is that the questions researchers often wish to answer are e.g. questions about marker expression patterns that differ between samples of different origin or identification of cell populations that correlate to disease status. These types of questions cannot be answered in an automated fashion by many of the clustering tools because of their unsupervised exploratory nature. Thus, in order to have an automated pipeline that can identify populations of interest, different approaches to automated gating need to be combined. This is attempted by biomarker discovery tools, where unsupervised clustering approaches are combined with supervised algorithms that can for instance categorize the identified cell populations from the unsupervised clustering with defined categories or correlate them to clinical outcome. Two examples are flowType-RchyOptimix (Aghaeepour, Jalali, *et al.,* 2012; Aghaeepour, Chattopadhyay, *et al.,* 2012; O'Neill *et al.,* 2014) and CITRUS (Bruggner *et al.,* 2014), the latter having been used in several studies (Gaudilliere

et al., 2014; Hansmann *et al.,* 2015), recently to identify a cytokine profile common to pediatric Systemic Lupus Erythematosus (SLE) patients (O'Gorman *et al.,* 2017).

With the introduction of analysis tools that combine exploratory analysis and identification of cell populations of interest, automated gating algorithms become increasingly attractive to non-computational experts, because they allow for a more fully automated approach. This will most likely foster the use of these tools by a broader range of research institutions.

1.4   High throughput detection of antigen specific T cells

As previously mentioned, the interaction between the pMHC and the TCR is tremendously complex with millions of different potential interactions between the huge number of different TCRs and the vast amount of epitopes that are being presented to the TCRs during disease. In a recent screening using combinatorial encoding of MHC multimers, a peptide library of 1036 peptides was investigated for CD8 T cell recognition in a melanoma patient (Cossarizza *et al.,* 2017). In this study, 8 different fluorochromes were used for MHC multimer generation, giving rise to 26 different combinations that were feasible to use. Thus, the patient sample had to be split in 55 fractions in order for the sample to be screened with all 1036 pMHC complexes. Splitting a cell sample in 55 fractions requires a large number of cells, and *ex vivo* screening for CD8 T cell recognition is hence not feasible with this strategy. At the same time, expansion of the T cells is not always desirable as the composition of antigen specific T cell populations in a given sample might change during the process of culturing (Andersen, Thrue, *et al.,* 2012).

Even if considering an MHC multimer experiment where all the 18 channels in a standard flow cytometry instrument were used, with a few of them reserved for live/dead discrimination, dump channel exclusion and CD8 detection, only just over 100 different pMHC specificities could be investigated in parallel. This is in the same range as what was recently done with combinatorial encoding of metal tags for CyTOF analysis of MHC multimer-binding T cells, where 109 specificities were investigated simultaneously (Newell, Sigal, *et al.,* 2013). With the very recent and ongoing development of new fluorochromes and flow cytometry instruments, moving towards detection of 50 parameters, the absolute state of the art experiments conducted with combinatorial encoding of MHC multimers now utilize ~25 different colors, giving rise to ~300 different dual color combinations (Kvistborg, 2018) (figure 1.6). Using this many different fluorochromes in a single experiment is not trivial as the issue of spectral overlap increases with the number of colors, not to mention the laborious process of carefully testing all of the ~300 different color combinations in order to exclude those that do not work well together.

Even though the introduction and continuous development of combinatorial encoding of MHC multimers has significantly increased the complexity of the MHC
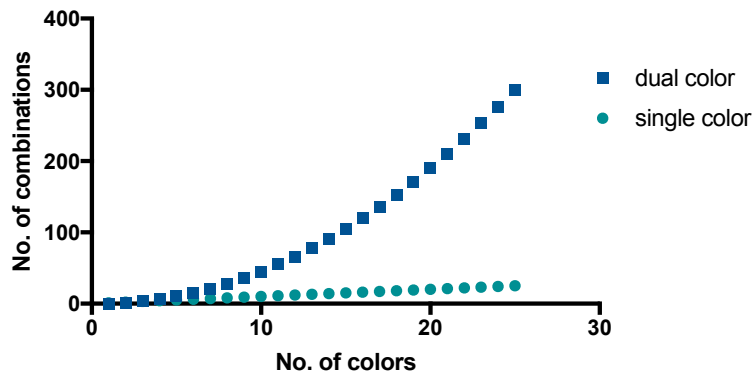
Figure 1.6: The possible number of combinations arising from a given number of fluorescence labels.

multimer assay, this strategy is still nowhere near matching the diversity of the pMHC-TCR interaction.

1.4.1 DNA barcode labeled MHC multimers

Recently, a big leap into large scale detection of MHC multimer-binding T cells was accomplished with the development of DNA barcode labeled MHC multimers (Bentzen, Marquard, *et al.,* 2016).

In this method, pMHC molecules of interest are generated through UV mediated peptide exchange, but instead of multimerizing these with fluorochromes, they are attached to dextran backbones containing a higher number of biotin-binding streptavidin sites than streptavidin conjugated fluorochromes, leaving room to also attach DNA barcodes. All dextran backbones are labeled with a common PE fluorochrome and each individual pMHC dextramer is coupled to a unique DNA barcode. All the different pMHC dextramers are mixed together and used to stain cell samples for CD8 T cell recognition. Based on the common PE signal, all cells that have bound a pMHC dextramer are sorted using fluorescence-activated cell sorting (FACS), and the identity of the specific peptides is revealed by amplification and sequencing of the attached DNA barcodes (figure 1.7). Unique molecular identifiers (UMIs) are incorporated into each individual DNA barcode in order for amplification bias to be corrected after sequencing and to control for potential contamination with amplified DNA from other sources. The sequencing data is analyzed using an online available software package, barracoda (http://www.cbs.dtu. dk/services/Barracoda). From this analysis, the number of sequencing reads for each pMHC specificity in each sample is given along with the number of reads from triplicates of a baseline sample

containing all the barcodes that the samples were stained with. Barracoda also calculates the log2 fold change of the number of reads from a given specificity to that of the baseline samples and the p value for this fold change. From the frequency of the total PE sorted population out of total CD8 T cells and the fraction of reads from a given pMHC specificity out of the total reads in the sample, the frequency of the CD8 T cell population specific for the given epitope can be estimated. This is done by calculating: (the number of reads for a given pMHC specificity/the total number of reads in the sample) x % PE sorted cells out of total CD8 T cells as described in Bentzen et al., 2016.
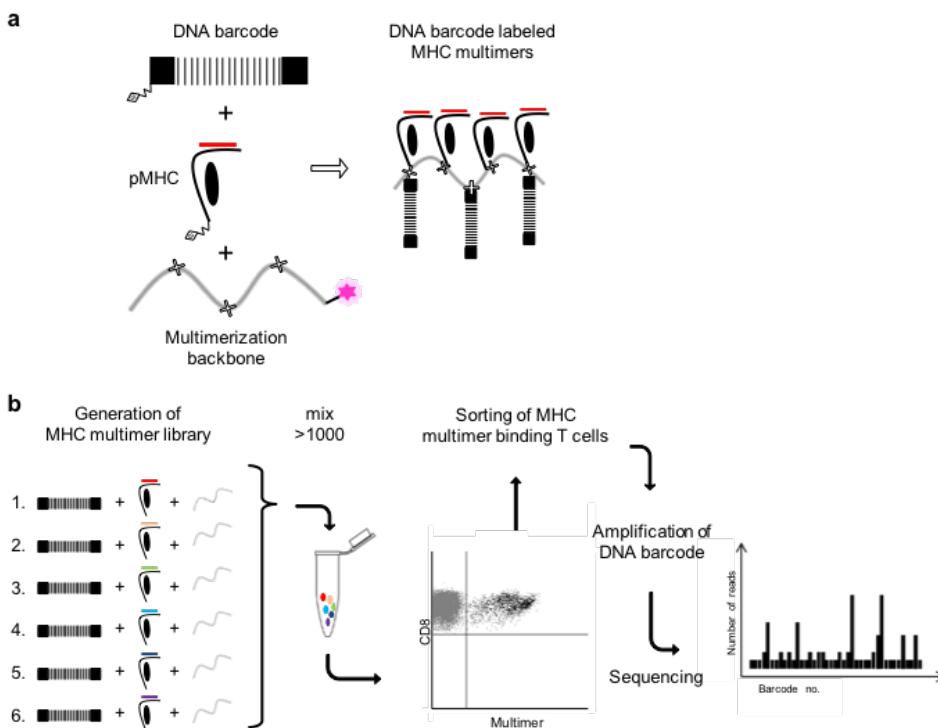
Figure 1.7: Overview of the barcoding technology. a) pMHC molecules are attached to a PE labeled dextran backbone together with unique DNA barcodes, creating a DNA barcode labeled multimer. b) Multiple MHC multimers of different specificity each have a unique barcode and more than 1000 can be mixed and used to stain a single sample. MHC multimer-binding T cells are sorted based on the common PE label, the attached DNA barcodes are amplified and the specificity of the sorted T cells is revealed through sequencing of the DNA barcodes. Modified from (Bentzen, Marquard, *et al.,* 2016).

1.4.1.1 Advantages

The DNA barcode labeled MHC multimer technology has been validated for detection of >1000 different possible epitopes simultaneously. The high complexity of the technology is achieved by removing the limiting step of availability of fluorochromes and detection channels in fluorescence based identification of MHC multimers. Instead this is replaced with detection through DNA barcodes of which more than $5x10^6$ could in theory be generated with the current design (Bentzen, Marquard, *et al.,* 2016; Xu *et al.,* 2009). Thus, the example of screening for CD8 T cell recognition using 1036 peptides described above, could with this method have been accomplished in a single sample, avoiding extensive cell culturing to reach enough cells to split in 55 fractions. The ability to do large scale screening for antigen specific CD8 T cells is the major advantage of this novel technology. Another improvement to MHC multimer detection that comes with this method relates to the detection of low affinity pMHC-TCR interactions. Identification of MHC multimer-binding T cells based on fluorescence requires a sufficiently strong binding to create a fluorescence signal that can clearly be distinguished from the background. DNA barcodes, however, can be isolated from T cells with fluorescence intensity closer to the negative population, thereby allowing detection of low avidity T cell populations. This is a clear advantage when investigating the T cell recognition of self-peptides, often being a low affinity interaction with the pMHC (Aleksic *et al.,* 2012). This feature makes the DNA barcode labeled MHC multimer technology especially desirable for investigation of mechanisms in autoimmune diseases, where recognition of self-peptides plays a major role.

1.4.1.2 Differences compared to fluorescence based detection of MHC multimers

There are a few differences between barcode- and fluorescence based detection of MHC multimer-binding T cells that are worth taking into account. Since all pMHC specificities are present in one pool of PE positive cells when using DNA barcoding, it is only possible to directly quantify the total number of antigen specific cells in a given sample. If interested in the frequency of each individual antigen specific population this number instead has to be estimated. However, it was shown that the estimated frequencies found with barcoding have a very good correlation to those found using fluorescence based detection of the same cell populations (Bentzen, Marquard, *et al.,* 2016).

More importantly, it is worth noting that the correlation between fluorescence intensity and the avidity of the pMHC-T cell interaction in a conventional fluorescence based assay, is not possible to make using the barcoding method. When fluorochromes are used for detection of MHC multimer-binding T cells, the intensity of the signal will be proportional to the amount of multimers a cell has bound and thus, will increase with increasingly strong binding between pMHC and TCR. Owing to the introduction of DNA barcodes as the identifying unit in barcoding, this information is lost. The output of a barcoding based MHC multimer assay is the number of

reads for a given barcode/peptide. This number contains no information about the amount of cells these barcodes were recovered from and thus offers no distinction between the case where a few cells have bound a high number of pMHC molecules and the one where many cells have bound only a few pMHC molecules, representing high and low avidity interactions, respectively. These two scenarios can produce the same number of sequencing reads from a given barcode, which is important to keep in mind when interpreting the results from barcoding experiments. However, the impact of this phenomenon is in practice very limited as comparisons of antigen specific T cells detected with both fluorescence and barcode based methods revealed that the correlation between observed frequencies and estimated frequencies found with the two methods respectively, was very high (Bentzen, Marquard, *et al.,* 2016).

1.4.1.3   Peptide prediction

Even though the DNA barcoding method enables detection of >1000 peptides simultaneously, the number of possible epitopes that could be of interest in a given disease vastly exceeds this number. It has been estimated that only a very small fraction, ~1%, of all potential peptides are actually able to bind to MHC molecules, making binding one of the most selective steps in antigen processing and presentation of peptides on the surface of cells (Yewdell & Bennink, 1999). Thus, ways of wisely choosing which peptides to include in a study are highly needed, in order to exclude all the potential non-binding peptides from a given protein.

For this purpose, tools that predict peptide binding to HLA molecules are very helpful. One of the first of such predictors, was the SYFPEITHI database (Rammensee *et al.,* 1999). The prediction of peptide binding affinity with this tool is based on MHC ligands eluted from the surface of cells, thus representing peptides that are naturally presented.

Another approach was taken by the developers of the NetMHC family of predictors. These algorithms use artificial neural networks to learn the rules of peptide-MHC binding based on training with *in vitro* generated binding affinity data (Buus *et al.,* 2003). The trained algorithms predict the binding affinity of all potential peptides within a given protein of interest and hence, are very valuable tools in guided selection of peptides to investigate for T cell recognition. At present they work best for prediction of peptides binding to HLA class I alleles and are widely used for this purpose. NetMHC has become state of the art for peptide prediction and can accurately predict binding between potential peptides and the HLA types for which binding data exists. Many HLA alleles are, however, not represented in the binding affinity data and as a means to expand the possibilities for peptide prediction, NetMHCpan was developed. This tool has the ability to transfer the rules that it learned from the training data containing a limited range of HLA molecules, to any known HLA allele and can, based on the sequence of the given HLA allele and peptide, successfully predict the affinity of the interaction between the two (Nielsen *et al.,* 2007). Another member of the NetMHC family is NetMHCcons, which combines the predictions from NetMHC version 3.4, NetMHCpan version 2.8 and another prediction tool called Pick-

Pocket version 1.1, achieving more accurate results than the single predictors alone (Karosiene *et al.,* 2012). Recently, however, the NetMHCpan predictor was updated to include eluted ligands in the training data in addition to the binding affinity data, combining the strengths of the two approaches. Thus this updated version now offers the most attractive tool for prediction of binding between peptide and HLA molecule (Jurtz *et al.,* 2017).

1.5   T cells in autoimmunity

The success of the immune system relies on the delicate balance of reacting to dangerous pathogens or danger signals within the body while staying unreactive to harmless foreign agents or self. The specificity of the adaptive immune system is a consequence of random rearrangement of the genes coding for the antigen receptors on B- and T cells. In T cells, the V(D)J gene rearrangements give rise to an immense amount of different TCRs and since they are generated by random combinations of V(D)J gene segments, some are bound to produce TCRs that recognize self-antigens. In order to keep these potentially auto-reactive T cells from circulating the body, negative selection during T cell development in the thymus ensures the destruction of at least the majority of such T cells. This process is, however, not complete as auto-reactive T cells do escape the selection and are found in circulation. Mechanisms of peripheral tolerance are responsible for keeping these cells in check.

1.5.1   Central tolerance

The fate of a developing T cell that has gone through successful TCR gene rearrangement is determined by its interaction with pMHC molecules expressed in the thymus. Interaction with self-pMHC above a certain threshold of affinity provides a survival signal in the T cell, whereas failure to bind to self-pMHC induces apoptosis, also termed death by neglect. This process is called positive selection and is estimated to be responsible for the death of 90-95% of developing T cells in the thymus (L. Klein *et al.,* 2014; Shortman *et al.,* 1991). However, a high affinity interaction between T cell and self-pMHC leads to clonal deletion of the given T cell. Thus, the T cells that are allowed to mature in the thymus and are released into the circulation have affinities for self-pMHC in an intermediate range that both ensures the ability of the T cells to mount an immune response towards antigen while at the same time limiting the risk of auto-reactivity. The number of T cells that are eliminated through positive selection far exceeds that of negative selection (Surh & Sprent, 1994), although this view was recently challenged (Stritesky *et al.,* 2013). A greater number of cells dying by neglect might reflect the vital importance of the ability to fight pathogens even if it comes at the price of risking autoimmunity.

In the process of negative selection, tissue restricted antigens (TRAs) are expressed in the thymus and at least two molecules have been identified that are important

regulators of this expression. The autoimmune regulator (Aire) is highly expressed in medullary thymic epithelial cells (mTECs) and promotes the expression of TRAs on these cells (Derbinski *et al.,* 2001). The forebrain-expressed zinc finger 2 (fezf2) transcription factor is likewise highly expressed in mTECs and is thought to promote the expression of TRAs that are distinct from those promoted by Aire. Defects in the Aire gene expression has been shown to cause human autoimmune disease (Nagamine *et al.,* 1997) and mouse model studies show that lack of fezf2 results in autoimmunity with distinct characteristics from Aire deficient mice (Takaba *et al.,* 2015).

Another important molecule during T cell maturation and selection is the tyrosine kinase Zap70, which is crucial for T cell signaling and is upregulated during T cell maturation. In a mouse model, a mutation in the Zap70 gene caused the kinase to have reduced but not completely ablated function. This mutation was associated with an altered threshold for both positive and negative selection, allowing auto-reactive T cells with a higher affinity for self-pMHC to escape to the periphery, thus causing autoimmunity (N. Sakaguchi *et al.,* 2003) (figure 1.8). Even though defects in crucial aspects of central tolerance mechanisms prove the importance of this process for the maintenance of a balanced immune system, they do not account for all cases of autoimmunity. Thus, cells with the ability to react to self-antigens do escape into the periphery under normal conditions.

1.5.2   Peripheral tolerance

There are several different mechanisms by which auto-reactive T cells that have escaped negative selection in the thymus can be controlled in the periphery. One such mechanism is induction of anergy. Binding between a naïve antigen specific T cell and its cognate antigen on an antigen presenting cell (APC) is not sufficient for T cell activation. Co-stimulatory signals, e.g mediated through the CD28/B7 pathway, are also required. In the absence of these co-stimulatory signals the T cell will become anergic as opposed to activated. Dendritic cells (DCs) that have been activated to mature in the absence of inflammation, in the presence of self-antigen instead of pathogen derived antigen, develop a tolerogenic phenotype. These tolerogenic DCs do not provide co-stimulatory but rather co-inhibitory signals and induce anergy in T cells that recognize the presented self-antigen (Mueller, 2010). The most studied inhibitory pathways are those of programmed death 1 (PD1) and cytotoxic T-lymphocyte antigen 4 (CTLA-4) and their ligands PDL-1 and 2 and B7 respectively. PD1 deficiency in mice causes severe systemic autoimmunity (Nishimura *et al.,* 1999)) and CTLA-4 loss is fatal due to massive lymphoproliferative disease (Tivol *et al.,* 1995; Waterhouse *et al.,* 1995), illustrating the role of these pathways in controlling immune responses and preventing auto-reactivity.

Recently Aire has been shown to be expressed in lymphoid tissues, serving as a promoter of TRA expression in these tissues through which naïve T cells circulate in order to meet their cognate antigen. Evidence suggests that just like in the thymus, clonal deletion of self-reactive T cells is induced in this process (B. Zhao *et al.,* 2018).
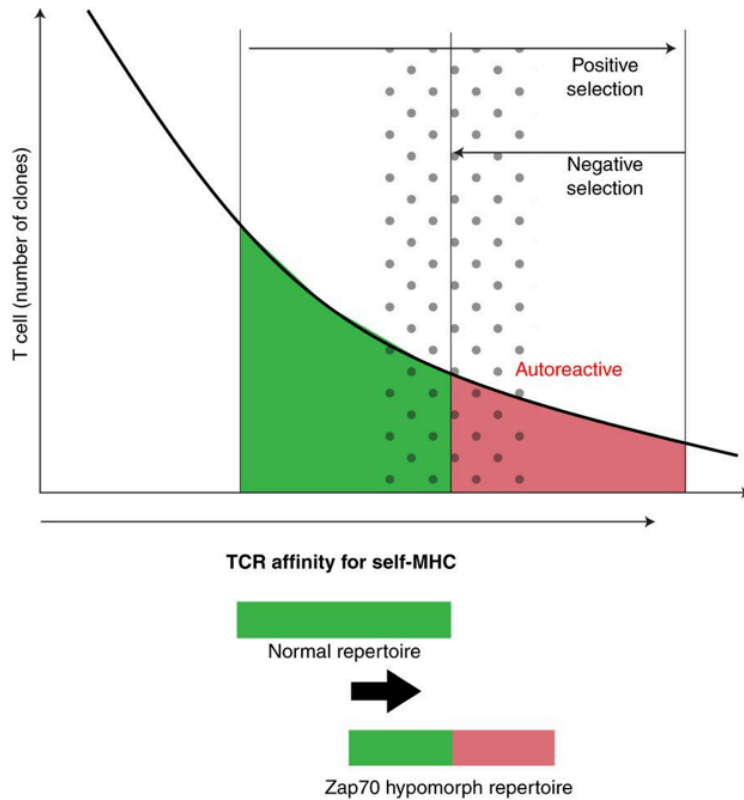
Figure 1.8: The window for positive and negative selection in the thymus. Positive selection ensures the survival of cells with the ability to react against self-pMHC while negative selection is responsible for the deletion of cells that have a high affinity for self. Dots represent the repertoire of cells that will be induced to develop into regulatory T cells. Zap70 mutations lead to a skewed window of selection, and results in the release of auto-reactive T cells with higher affinity into the periphery. Modified from (Cheng & Anderson, 2018).

Finally, regulatory T cells (Tregs) are crucial in maintaining peripheral tolerance. Developing T cells in the thymus with a certain affinity for self-pMHC are induced to express the transcription factor forkhead box P3 (FOXP3), which is a major factor for the differentiation of immature T cells into Treg cells (Ohkura *et al.,* 2013). Induction of FOXP3 expression and Treg differentiation can also occur in the periphery in conventional CD4 T cells under certain conditions as for example antigen stimulation in the presence of TGF-β (W. Chen *et al.,* 2003). Treg cells suppress T cell proliferation, effector cell differentiation and function by various mechanisms, such as release of immunosuppressive cytokines (Von Boehmer, 2005), downregulation of co-stimulatory molecules on APCs (S. Sakaguchi *et al.,* 2008), absorption of cytokines needed by effector T cells (Pandiyan *et al.,* 2007), inhibitory signaling through B7 molecules on effector T cells when bound by CTLA-4 on Tregs (Paust *et al.,* 2004) or direct killing by release of e.g. perforin/granzyme (S. Sakaguchi *et al.,* 2008). Tregs are vital for the control of immune responses and autoimmunity as the mutation of FOXP3 causes extensive autoimmunity in both mice and humans.

1.5.3   The ever present threat of auto-reactive T cells

As described above, there are multiple mechanisms by which auto-reactive T cells are deleted or kept at bay. Self-reactive T cells that escape negative selection are controlled by peripheral tolerance mechanisms. Considering these processes to be complete, implies that autoimmune diseases are by definition a result of a fault in either central or peripheral tolerance. The growing evidence of the presence of auto-reactive T cells at equal frequencies in healthy controls and individuals with autoimmune diseases is a clear indication that central tolerance mechanisms are not complete. These observations have been made for both CD4- (Cao *et al.,* 2015; Snir *et al.,* 2011) and CD8 T cells (Berthelot *et al.,* 2008; Maeda *et al.,* 2014; Culina *et al.,* 2018; Martin *et al.,* 1990) and furthermore, it has also been shown that in healthy individuals, self-specific CD8 T cells are present at the same frequency as non-self-specific T cells (Yu *et al.,* 2015). In several of these studies, self-specific T cells from healthy individuals were less functional than those from individuals suffering from autoimmune disease, suggesting that failure of peripheral tolerance mechanisms is responsible for disease development. This is, however, not always the case, as a recent study of auto-reactive CD8 T cells in type 1 diabetes (TD1) showed equal functionality of such cells from patients and healthy controls (Culina *et al.,* 2018), which were instead distinguished by their homing to the disease associated tissue. This suggests, that even in healthy individuals, self-reactive T cells with full functionality are allowed to persist in the body and are not subjected to suppression or anergy induction.

The theory of clonal ignorance addresses the question of how these cells can be present without causing disease. Naïve T cells and T cells that encounter their antigen under non-inflammatory conditions circulate between the blood and the lymph nodes, but are not infiltrating organ tissue under steady state conditions. Thus, self-specific T cells do not encounter their antigen as these are expressed inside tissue parenchyma

(Ehlers, 2018; Mueller, 2010). However, changes to the steady state situation such as inflammation or stress can trigger the infiltration of self-reactive T cells and the consequent development of autoimmune pathology in the given organ. In line with this theory, healthy mice transgenically expressing lymphocytic choriomeningitis virus (LCMV) glycoprotein (GP) in the β islet cells and a GP specific high avidity TCR in their mature, naïve CD8 T cells, do not develop autoimmune destruction of the β islet cells, thus indicating that the mere presence of both auto-antigen and auto-reactive T cells within the organism is not enough to elicit an autoimmune response. LCMV infection of the mice did however break this clonal ignorance and induced T1D onset (Ohashi *et al.,* 1991). Similar observations have been made in another more recent study utilizing transgenic OVA expression in the skin and subsequent transfer of OT-I and II cells Bianchi *et al.,* 2009, although contradictory findings also exist where OVA expression in skin led to autoimmune pathology immediately upon transfer of OT-I cells (Azukizawa *et al.,* 2003). From an evolutionary stand point, having self-reactive T cells in our bodies makes sense in the way that the deletion of all self-reactive T cells might create "holes in the T cell repertoire" that could be exploited by pathogens, as discussed by Yu et al., 2015.

Another possible mechanism by which infection can trigger autoimmunity is molecular mimicry, There are several studies showing the cross-reactive nature of the TCR, and it is well known that immunity towards one pathogen can infer immunity towards another, both in cases where the pathogens are similar and very different from each other (Su *et al.,* 2013; Welsh & Selin, 2002). Thus, the idea that similarity between pathogen and host derived epitopes causes cross-reactive T cells to attack self is obvious and also substantiated by experimental findings (Rose, 2017; Z. S. Zhao *et al.,* 1998). Whether it is by breaking clonal ignorance or molecular mimicry, infection and inflammation as a trigger for autoimmunity seems to be very plausible and has been speculated in multiple autoimmune diseases such as multiple sclerosis (Pender *et al.,* 2017), T1D (Christen *et al.,* 2016), Guillian-Barre syndrome (Wim Ang *et al.,* 2004), rheumatic fever (Chakravarty *et al.,* 2014) and narcolepsy (Partinen, Kornum, *et al.,* 2014).

1.6   Narcolepsy

Narcolepsy is a neurological disorder that affects the ability of the brain to control the sleep-wake cycle and as a consequence narcolepsy patients present with excessive daytime sleepiness as well as interrupted nocturnal sleep. Furthermore, narcolepsy is often accompanied by sleep paralysis, hypnagogic hallucinations and abnormal rapid eye movement (REM) sleep. Two different types of Narcolepsy exist, type 1 and 2, of which type 1 is by far the most frequent. Narcolepsy type 1 (NT1) is further characterized by a condition called cataplexy where muscle tone is suddenly lost, often triggered by strong emotions, and by having low or undetectable cerebrospinal fluid (CSF) levels of the neuropeptide hypocretin 1, also known under the name orexin A. Hypocretin is produced by a small population of neurons in the hypothalamus

with projections to many other parts of the brain and is implicated in diverse brain functions such as reward behavior, food intake, anxiety and other emotions, but most importantly for narcolepsy, in wakefulness (Liblau *et al.,* 2015) (figure 1.9).
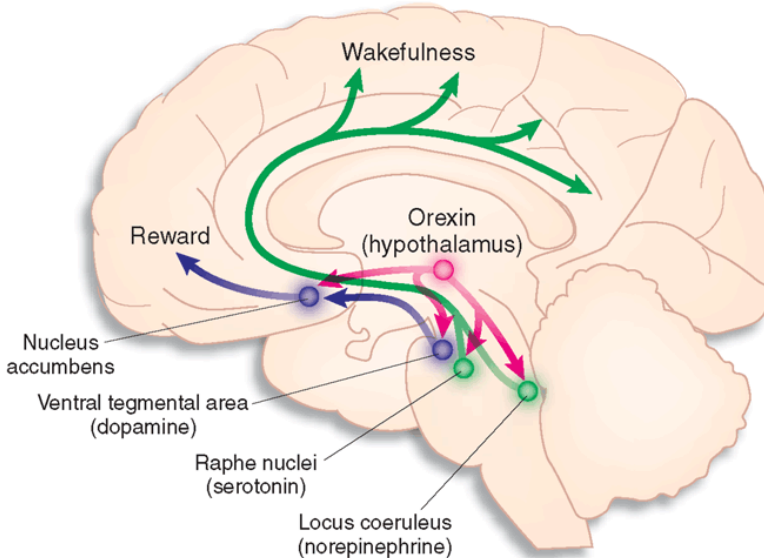
Figure 1.9: Overview of the signaling pathways of hypocretin/orexin.

Patients who have sleep dysregulation patterns consistent with narcolepsy disease but have normal hypocretin levels and no cataplexy are diagnosed with narcolepsy type 2 (NT2). For the remainder of this thesis, focus will be on narcolepsy type 1.

Narcolepsy is not exclusively found in humans, but also in animals. Canine narcolepsy has been shown to be caused by a mutation in the receptor 2 for hypocretin (HCRTR2) (Lin, Faraco, *et al.,* 1999) and studies in mice showed that knockout of the precursor protein prepro-hypocretin (HCRT), which is cleaved to give hypocretin 1 and 2, resulted in a phenotype very similar to narcolepsy (Chemelli *et al.,* 1999). Except for one reported case, no such mutations have been found in human patients (Peyron *et al.,* 2000). Instead, it was discovered that people suffering from narcolepsy have a drastic decrease in the number of hypocretin producing neurons in the hypothalamus, thus explaining the loss of signaling through this neuropeptide observed in NT1 patients. Other types of neurons that are intermingled with the hypocretin producing neurons in the hypothalamus are present at completely normal levels in narcoleptic brains, proving the neuronal loss to be very specific for hypocretin neurons (Peyron *et al.,* 2000) (Thannickal *et al.,* 2000).

The exact mechanism of the neuronal loss is not known, but it is thought to be the result of an autoimmune attack.

1.6.1   Immunology of narcolepsy

There are multiple indications to the involvement of the immune system in the pathogenesis of narcolepsy. The most prominent of these is the very strong association with the HLA class II molecule DQB1*06:02, which was reported to increase the risk for development of narcolepsy 251 times (Tafti, Hor, *et al.,* 2014). HLA-DQB1*06:02 is expressed by up to ~98% of narcolepsy patients whereas the prevalence in the healthy population is ~25% depending on ethnicity (Mignot *et al.,* 1997; Tafti, Hor, *et al.,* 2014). Furthermore, genome wide association studies (GWAS) have identified several genes related to immune regulation that are associated to narcolepsy, among them a single nucleotide polymorphisms in the TCR α locus (Kornum, Knudsen, *et al.,* 2017).

Although much less significantly than the DQB1*06:02 class II allele, a number of HLA class I alleles have been shown to be associated to narcolepsy in two different studies (Ollila *et al.,* 2015; Tafti, Lammers, *et al.,* 2016). These are HLA-A*11:01, HLA-B*51:01 and HLA-C*04:01, which were found in both studies, as well as HLA-B*18:01 and HLA-B*35:01 found by Tafti et al., and HLA-B*35:03 reported by Ollila et al.

Another indicator for a role of the immune system is the increase in narcolepsy incidence that was observed after the 2009/2010 H1N1 influenza vaccination Pandemrix in several different countries (Dauvilliers, Arnulf, *et al.,* 2013; Heier *et al.,* 2013; Partinen, Saarenpää-Heikkilä, *et al.,* 2012; Sarkanen *et al.,* 2018), as well as after H1N1 infection (Han *et al.,* 2011). For children and adolescents an overall 5-14 fold increase was observed after vaccination whereas it was 2-7 fold in adults (Sarkanen *et al.,* 2018). It could be speculated, that narcolepsy is a consequence of molecular mimicry where pathogen reactive T cells cross-react to a target expressed by hypocretin neurons, or simply that auto-reactive T cells are present in certain individuals and that these are somehow triggered by an infection. In favor of the molecular mimicry hypothesis, one study found a peptide from the H1N1 influenza virus nucleoprotein A to share residues with a fragment of HCRTR2 and antibodies that were able to cross-react to both these peptides, were found at increased levels in patients with vaccine-induced narcolepsy compared to non-narcoleptic individuals either infected with H1N1 or vaccinated with a different vaccine (Ahmed *et al.,* 2015). A number of concerns about the study were however raised, including the fact that the same antibodies were found in many healthy controls as well as in patients and furthermore the relevance of HCRTR2 antibodies for disease development was questioned (Vassalli *et al.,* 2015). In a study investigating the immune profile of narcolepsy patients, no difference was observed between patients with vaccine-induced or idiopathic narcolepsy (Hartmann *et al.,* 2016). Thus, the mechanism by which the Pandemrix vaccination or H1N1 infection itself induced narcolepsy is still not clear, as reviewed in Sarkanen

et al., 2018. It does however point to the possibility that the autoimmune process leading to narcolepsy is set off by an environmental trigger, such as infection, also in idiopathic cases. This hypothesis is supported by the observed seasonal pattern of narcolepsy onset with a yearly peak ~6 months after influenza season in China (figure 1.10) (Han *et al.,* 2011).
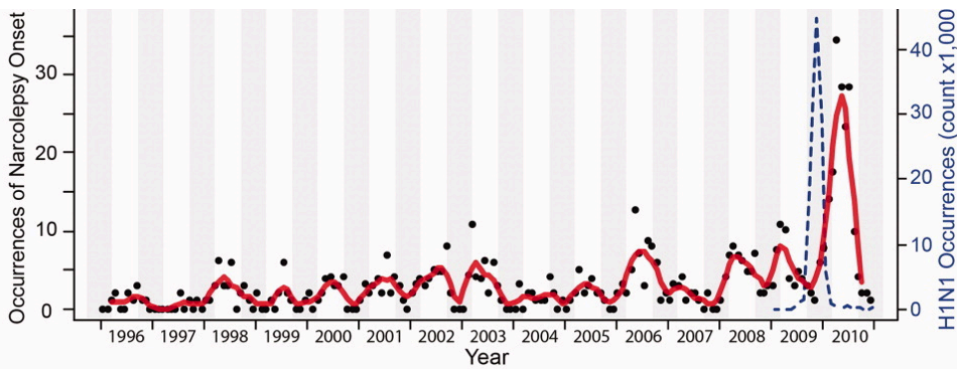
Figure 1.10: Number of monthly new occurrences of narcolepsy over a period of 15 years. Modified from (Han *et al.,* 2011).

Since the formulation of the autoimmune hypothesis for narcolepsy when the strong association with HLA-DQB1*06:02 was found and the later discovery that human narcolepsy is caused by the loss of hypocretin neurons, efforts have been put into identifying immune responses involved in this destruction. Several studies have attempted to identify auto-reactive antibodies present in narcolepsy patients, but these studies have been largely inconclusive, as autoantibodies are detected in a fraction of patients and healthy controls in some studies, whereas not in others (reviewed in Kornum et al., 2017). Immune profiling of narcolepsy patients has also been attempted, again with varying results. In one study increased activation and cytokine production by T cells was observed in patients compared to healthy controls (Hartmann *et al.,* 2016), while another study found no difference in the cytokine levels in the CSF between narcolepsy patients and healthy controls (Kornum, Pizza, *et al.,* 2015).

1.6.1.1   T cells in narcolepsy

The very strong correlation between narcolepsy and HLA-DQB1*06:02 allele could suggest that antigen specific CD4 T cells would be involved in the disease. However, one study investigating the proliferative potential of CD4 T cells from narcolepsy patients, in response to peptide pools spanning the HCRT protein, failed to show a significant difference between CD4 T cells from patients and healthy controls. HCRT

reactive CD4 T cells were only detected and verified in 3 out of 15 patients but in none of the 13 controls (Ramberger *et al.,* 2017). Very recently, another study successfully detected HCRT reactive CD4 and CD8 T cells in narcolepsy patients for the first time (Latorre *et al.,* 2018). In this extensive study, pools of peptides spanning the entire length of the HCRT protein were used to stimulate expanded memory CD4 T cells and the proliferation in response to this stimulation was measured. With this strategy, hundreds of cultures from 15 NT1 patients were screened, and responding memory CD4 T cells were found in some of the cultures from 14 out of the 15 patients but only in cultures from 3 out of 12 healthy controls. Strong responses were however only found in few patients as the majority only had very few responding cultures out of several hundred. Memory CD8 T cell reactivity was also found, although much less significantly than for the CD4 T cells as reactive cells were only detected in few cultures in 3 out of 10 patients as well as in 2 out of 9 healthy controls. CD4 and CD8 T cells reactive against tribbles homolog 2 (TRIB2), a protein kinase expressed in several tissues, including hypocretin neurons, were also found, although at the same level in both patients and healthy controls. None of the CD4 cultures specific for either HCRT or TRIB2 responded to influenza vaccine antigens, arguing against the hypothesis for molecular mimicry as a mechanism for vaccine induced narcolepsy. A few NT2 patients were included in the study and recently one of these developed cataplexy, which is a criteria for diagnosis of NT1. Interestingly, this particular patient had relatively high levels of both CD4 and CD8 T cell reactivity, thus suggesting that the T cell reactivity towards HCRT precedes the neuronal destruction leading to NT1.

Even though MHC class II expression has been detected on developing neuronal stem cells (Vagaska *et al.,* 2016) and in rare cases of disease (Hollister *et al.,* 1997), the general belief in the scientific community is that mature neurons do not express MHC II under normal physiological conditions (Barateau *et al.,* 2017; Styren *et al.,* 1990). CD4 T cells are therefore not very likely to be the main effector cells responsible for killing hypocretin neurons and instead, this part in the pathogenesis of narcolepsy has been speculated to belong to CD8 T cells. A number of findings support this hypothesis. First of all, in a post mortem case study of one patient who developed NT1 secondary to anti-Ma associated encephalitis a complete loss of hypocretin neurons was observed together with a heavy infiltration of CD8 T cells. This patient died 4 months after the onset of NT1 symptoms, thus presenting a rare opportunity to investigate a narcoleptic brain very close to disease onset (Dauvilliers, Bauer, *et al.,* 2013). In an animal model study of NT1, mice were transgenically modified to express the H1N1 influenza virus hemagglutinin (HA) exclusively in hypocretin neurons. The ability of HA specific CD4 and CD8 T cells to enter the brain, cause inflammation and neuronal destruction was then tested. The transfer of specific CD4 cells led to infiltration of these cells specifically to the hypothalamus and also to local inflammation, but no reduction in the number of hypocretin neurons was found. In contrast, the transfer and subsequent infiltration of HA specific CD8 T cells to the hypothalamus specifically, induced a substantial reduction of around 70% in the number of hypocretin neurons but not in other types of hypothalamic neurons. This

study importantly revealed the ability of cytotoxic CD8 T cells, but not CD4 T cells, to kill hypocretin neurons expressing their cognate antigen (Bernard-Valnet *et al.,* 2016).

Although the study by Latorre et al., 2018 for the first time successfully identified significant levels of CD4 T cell reactivity towards HCRT in narcolepsy patients, these findings do not offer a direct explanation for the pathogenesis of narcolepsy as CD4 T cells are thought to play an indirect role as helper cells to CD8 mediated cytotoxicity. Even though a few NT1 patients were found to harbor CD8 T cells with reactivity against HCRT, these findings were not significantly different from healthy controls. What is thus still lacking, is the identification of CD8 T cell targets expressed by hypocretin neurons that could lead to the recognition and specific destruction of these cells as well as the presence of the antigen specific CD8 T cells in NT1 patients.

1.6.2   Characterization of hypocretin producing neurons

The neuronal loss which is characteristic of NT1 is very specific to hypocretin producing neurons and this destruction is believed to be the result of an antigen specific CD8 T cell attack on these neurons. If that is in fact the underlying mechanism for narcolepsy development, the antigen target for these cytotoxic T cells must be expressed inside hypocretin neurons and presented on the surface of the cells where they can be recognized by a specific population of CD8 T cells. Several studies have characterized the hypocretin neurons in terms of their protein expression profile in order to search for proteins that regulate the function of these neurons and to search for potential targets for an autoimmune attack.

The only protein that is truly unique to the hypocretin neurons, is hypocretin itself. It is, as previously mentioned, expressed as a precursor protein which is cleaved to give hypocretin 1 and 2. There are, however, also a number of other proteins that are highly expressed by hypocretin neurons. Prodynorphin (PDYN) is the precursor of a neurotransmitter which is expressed in several regions of the brain, but within the hypothalamus, it has been found to co-localize with HCRT (Chou *et al.,* 2001; Dalal *et al.,* 2013). Mouse models with a selective loss of hypocretin neurons have been shown to also lack PDYN in the hypothalamus but not in other areas of the brain (Chou *et al.,* 2001), and in humans the concomitant lack of PDYN together with lack of hypocretin in narcoleptic brains is one of the arguments for the hypothesis that hypocretin neurons are lost and not just quiescent (Crocker *et al.,* 2005). A study in mice identified two transcription factors with importance for hypocretin neurons, Regulatory factor 4 (RFX4) and LIM Homeobox 9 (LHX9) (Dalal *et al.,* 2013). RFX4 was shown to have an almost perfectly overlapping expression with HCRT, indicating that this transcription factor is predominantly expressed in hypocretin neurons. This is consistent with the observation that RFX4 is only expressed in the hypothalamus in the Allen Brain Atlas analysis (http://portal.brain-map.org/). Interestingly, this study found LHX9 to be important for hypocretin neurons as their numbers were reduced by ~30% in LHX9 knockout mice with concomitant exces-

sive daytime sleepiness observed in these animals. The expression pattern of LHX9 overlapped with HCRT, suggesting specificity to the hypocretin neurons, although Allen Brain Atlas indicates universal brain expression. The finding of a reduction in hypocretin neurons as a consequence of LHX9 knockout was confirmed in zebrafish in another study, where it was also shown that LHX9 is sufficient for the induction of hypocretin neurons during development in both zebrafish and mice and for the continuous production of hypocretin neuropeptides (Liu *et al.,* 2015). LHX9 was also found to be necessary for the development of hypothalamic neurons expressing the pyroglutamylated RF amide peptide (QRFP). QRFP and HCRT were shown to be expressed in distinct neurons, and this was also found in another study also showing that QRFP is involved in sleep promotion in zebrafish (A. Chen *et al.,* 2016). The observed localization of QRFP was contradicted in a recent study where QRFP was found to be expressed in a fraction of the hypocretin neurons (Romanov *et al.,* 2017). Additionally, CD4 T cell reactivity against QRFP has been observed in narcolepsy (Seifinejad & Tafti, 2017).

As previously mentioned, hypocretin neurons project to several different brain regions, communicating with other types of neurons. Consequently, receptors for the hypocretin neuropeptides are expressed throughout the brain. Two types of receptors exist, hypocretin receptors 1 and 2 (HCRTR1 and HCRTR2), HCRTR1 being selective for hypocretin 1 and HCRTR2 being a receptor for both hypocretin 1 and 2 (Marcus *et al.,* 2001). It is a matter of debate whether HCRTRs are expressed on hypocretin neurons themselves, since this is found by some groups but the contrary is found by others (Vassalli *et al.,* 2015; Yamanaka *et al.,* 2010). Antibodies against HCRTR2 have been detected in narcolepsy patients, albeit at very different levels in different studies. These studies did, however, also investigate the presence of HCRTR2 antibodies in two different cohorts of narcolepsy patients with high levels of antibodies present in vaccine induced narcolepsy patients but low levels in idiopathic narcolepsy patients (Ahmed *et al.,* 2015; Giannoccaro *et al.,* 2017). Based on the finding of autoantibodies in narcolepsy patients, TRIB2 has been suggested as a target in narcolepsy. It was shown to be expressed by hypocretin neurons at higher levels than in surrounding neurons in mice and TRIB specific antibody titers were significantly higher in narcolepsy patients than in control groups (Cvetkovic-lopes *et al.,* 2010). Furthermore, CD4 T cells were shown to be reactive towards TRIB2 in the study described above by Latorre et al., although the same level of reactivity was found in narcolepsy patients and controls. TRIB2 is widely expressed throughout the brain according to the Allen Brain Atlas, but is also expressed in the periphery. The other proteins described in this section all have very limited peripheral expression (Fagerberg *et al.,* 2014), thus presenting as valid potential targets for an autoimmune attack on hypocretin producing neurons.

Manuscript III

The hypothesized autoimmune basis of narcolepsy type 1 has yet to be indisputably proven. It has been suggested that the effector cells involved in direct destruction of hypocretin neurons are the CD8 T cells. Therefore, the aim of this study was to investigate the presence of auto reactive CD8 T cells in narcolepsy type 1. We used DNA barcode-labeled MHC multimers to investigate CD8 cell reactivity towards 7 different proteins relevant to narcolepsy type 1– HCRT, HCRTR2, LHX9, PDYN, QRFP, RFX4 and TRIB2. Peptides with binding ability to 8 different HLA class I alleles – HLA-A*02:01, HLA-A*03:01, HLA-A*11:01, HLA-B*07:02, HLA-B*18:01, HLA-B*35:01, HLA-B*51:01 and HLA-C*04:01 – were predicted using the NetMHCCons prediction server. This yielded 1183 different peptides which were used to generate unique DNA barcode-labeled MHC multimers. Peripheral blood mononuclear cells (PBMCs) from 20 patients and 52 healthy controls were screened for CD8 T cells with the capacity to recognize any of the predicted peptides in our peptide library. All patients except one was positive for the narcolepsy associated HLA allele DQB1*06:02, whereas this was true for about half of the healthy controls. The patients and HLA-DQB1*06:02 positive and negative controls were compared in terms of the number, frequency and type of responses detected, in the hopes of detecting a response pattern that could shed light on the pathogenesis of narcolepsy.

CHAPTER 2

Manuscript I

Optimization in detection of antigen specific T cells through differently-labelled MHC multimers

Optimization in detection of antigen specific T cells through differently-labelled MHC multimers

Natasja Wulff Pedersen[1], Karoline Laske[2], Dominik Maurer[2], Steffen Walter[2], Cecile Gouttefangeas[3], Sine Reker Hadrup[1]

[1]Department of Micro- and Nanotechnology, Technical University of Denmark, Copenhagen, Denmark
[2]Immatics Biotechnologies, Tübingen, Germany
[3]Department of Immunology, Interfaculty Institute for Cell Biology, University of Tuebingen, Tuebingen, Germany

Abbreviations

APC – Allophycocyanin
CMV – Cytomegalovirus
EBV – Epstein barr virus
FLU – Influenza
MHC – Major Histocompatibility Complex
PBMC – Peripheral blood mononuclear cells
PE – Phycoerythrin
SI – Staining Index
QD – Quantum Dot

Abstract

Several fluorescent molecules are on a regular basis used to tag MHC multimers for detection of antigen specific T cells. We have assessed if detection of MHC multimer binding cells are dependent on the fluorescent label used to tag the MHC multimer through an exploratory proficiency panel, where detection of MHC multimer binding T cells was assessed across 16 different laboratories. The ability to detect an MHC multimer binding T cell population is correlated to the staining index of the given population, and this often depends on the fluorescent tag used, the flow cytometry settings and antibody panels used. Consequently, we describe here a strategy to identify weak detection channels and optimize the staining index for selected fluorescent molecules that can be easily employed to test and optimize fluorescent detection in relation to MHC multimer staining.

Introduction

Numerous fluorescent tags are used to detect Major Histocompatibility Complex (MHC) multimer binding T cells, either using single or combined colors to enhance the number of T cell populations that can potentially be detected in a given sample [1,2]. In the present study, we aim to assess if a collection of different fluorescent tags results in similar T cell detection when assessed across a broad range of 16 different laboratories. Detection of MHC multimer binding T cells may vary substantially across different laboratories[3], and consequently the Cancer Immune Therapy association immunoguiding work group (CIMT/CIP) has through several proficiency panels identified steps with major contribution to this variance. We have worked to harmonize and optimize protocols, and recent proficiency panels evaluating the detection of antigen-specific T cells using MHC multimers, have demonstrated a low inter-lab variability (www.cimt.eu/workgroups/CIP). There is, however, still

room for improvement. In addition to variance coming from the generation and staining protocols, differences in detection of antigen-specific T cells using MHC multimers may potentially arise as a consequence of the different fluorescent tags used, the fluorescence detection determined by the performance of the flow cytometry instrument, the instrument settings and the additional fluorescently-labelled antibodies used for detection.

In a previous proficiency panel (CIP_ID07_2010_MUL/D) we showed that differentially labelled MHC multimers could be used to detect multiple T cell populations in a single sample, with comparable efficiency as if each population was detected in separate samples (www.cimt.eu/publications). However, although no differences in T cell detection arose from mixing MHC multimer reagents for simultaneous detection of T cells, we observed a tendency towards enhanced detection frequencies with traditionally 'high intensive' fluorescent molecules, such as PE and APC, as compared to 'low intensive' fluorescent molecules, such as Quantum Dot 705 (QD705). Here, we investigate potential differences in detection of antigen responsive T cells ranging from low-frequent populations (<0.2% of CD8 T cells) to high-frequent T cell populations (1-2 % of CD8 T cells) when using four different labels: PE, APC, QD605 and QD705 for MHC multimer detections. We compare the detection rate and frequency of antigen responsive T cells detected and the SI for individual populations across 16 different laboratories participating in this exploratory proficiency panel (CIP_ID13_2012_MUL/D).

Furthermore, we have introduced a fast and easy bead-based evaluation of fluorescent detection for relevant fluorescent molecules, and show that detection of these evaluation-beads correlates with SI of the MHC multimer population. Consequently, such a bead-based system can be used to optimize fluorescent detection in selected channels.

Materials and methods

PBMC samples
Leukaphereses and buffy-coats were obtained by venipuncture from HLA-A*02 positive consenting healthy donors at the Center for Clinical Transfusion Medicine in Tübingen, Germany or at the Central Blood Bank, Rigshospitalet, Denmark. Peripheral blood mononuclear cells (PBMC) were isolated by gradient centrifugation within 8 hrs; after two washing steps in PBS, cells were counted and frozen at 10-20 x10^6 cells/ml in freezing containers, then transferred to the gas phase of a liquid nitrogen tank or to -150ºC freezers for long-term storage.

Reagents for flow cytometry
Peptide-MHC multimers: Peptide-HLA-A*02:01 monomers and multimers were produced in-house either by the classical refolding method or by UV-exchange as previously described [4]. Fluorescent multimers were generated by co-incubating monomers with streptavidin-fluorochromes (all from Life Technologies, Darmstadt, Germany), either at a 4:1 monomer/streptavidin ratio (-PE, -APC) or at a 30:1 monomer/quantum dot ratio (-quantum dot 605 (Q605) or 705 (Q705)) [2]. The following specificities were included: known epitopes derived from the viruses HCMV (pp65 495-503 NLVPMVATV, i.e. CMV), Influenza A (Flu Matrix 58-66 GILGFVFTL, i.e. FLU), and EBV (BMLF1 259-267 GLCTLVAML, i.e. EBV1 and EBV BRFL1 109-117 YVLDHLIVV, i.e. EBV2). In addition, a multimer refolded with the HLA-A*0201 UV exchangeable peptide KILGFVFJV (A2*p) was included as negative control. All multimers were frozen after addition of cryoprotectants (glycerol and serum albumin, 16% and 0.5% respectively)[4].

Fluorescent calibration beads: Quantum™ MESF and Quantum™ Simply Cellular® 6-9μm diameter microspheres (Bangs Laboratories,Inc., Fishers, IN) were used to monitor the flow cytometers´ performance in the 4 fluorescence channels also used for the multimer-detection, i.e. PE,

APC, Q605 and Q705. PE- and APC-beads were obtained from the manufacturer (Quantum™ MESF). For Q605 and Q705, microspheres coupled with anti-mouse capture antibodies (Quantum™ Simply Cellular®) were incubated with the mouse monoclonal antibodies (mAb) S3.5-Q605 or 3B5-Q705 (both Life Technologies) for 30 min at room temperature; Ab staining solutions were centrifuged 5min at 13800 g and 4°C before use in order to remove aggregates. Beads were then washed 3 times with PBS (Lonza, Cologne, Germany) at 850 g for 5 min, resuspended in PBS 0.5% BSA, 2mM EDTA and 0.02% azide, aliquoted at 4×10^4 beads in $100 – 200$ ul/ vial and frozen at -80°C until use. For all fluorochromes, unstained (blank), as well as beads labelled with increasing amount of fluorescence (dim, mid and bright) were available. Representative results for the bright-fluorescence beads are reported.

Proficiency panel design

The aim of the proficiency panel was to 1) compare the results obtained by different laboratories when staining PBMC with the same multimers coupled to four different fluorochromes, and 2) test the feasibility and utility of cryopreserved calibration beads for controlling flow cytometer performance. Participants were all experienced in multi-parametric flow cytometry and for most of them with multimer staining (15/16). All received one parcel on dry ice containing three preselected HLA-A*0201 PBMC cryovials (one vial/donor), aliquots of five different multimers (CMV, FLU, EBV1, EBV2 and A2*p) each coupled to the 4 fluorochromes tested (PE, APC, Q605 and Q705), and calibration beads for the same 4 fluorochromes. According to a pre-screening evaluation of T cell recognition, the three PBMC samples contained a total of 11 virus-specificities (CMV, FLU, EBV1 and EBV2) of high (> 1%, n=3), intermediate (0.1% - 1%, n=5) and low (<0.1%, n=3) frequencies of $CD3^+CD8^+$ (Tabel 1).

Test conditions: except for the MHC multimers and a few requirements for the procedure, all staining reagents and protocols, as well as flow acquisition and analysis conditions were free of choice. As in previous proficiency panels organized by CIP [5,6], some parameters were, however, mandatory: 1) the number of cells per stain (between 1 and 2×10^6), 2) the inclusion of at least a CD3 and a CD8 mAb, 3) the conditions of the multimer staining step (4 ug/ml of each multimer at room temperature for 30 min), 4) the acquisition of all cells contained in stain tubes, and 5) the acquisition of the calibration beads (at least 1×10^4) in the same experiment and using the same settings (PMT voltage and compensation) as for the PBMC stains. Participants were free to 1) include further mAb and/or a dead cell dye (Ab clones and dyes free of choice) and 2) use local buffers and staining protocol. A panel guideline, cell staining protocol and detailed instructions on how to handle, acquire and analyze the beads were provided. Each PBMC sample was split in 6 and each fraction was stained with the 4 different virus peptide/multimers labeled with 4 different fluorochromes. 4 out of the 6 fractions were stained with all 4 virus peptide/multimers labeled with different combinations of fluorochromes so that each multimer/fluorochrome combination was present. The last 2 fractions were used for control stains, one with the 4 different A2*p multimers and one "fluorescence minus one" (FMO), containing only antibody mix and no multimer.

This was done for all 3 donors giving a total of 18 stainings from each lab. In addition 16 calibration beads were acquired together with the cells and analyzed by each lab; blank, low, mid and high intensity for the 4 different fluorochromes. The gating strategy for the bead tests, but not for the cell stains, was standardized, with examples for both cells and beads displayed in the panel guideline.

Data reporting: the number of $CD3^+CD8^+$ and $CD3^+CD8^+multimer^+$ lymphocytes was recorded for each PBMC and each multimer specificity. In addition, the median fluorescence of the $CD3^+CD8^+multimer^+$ and of the $CD3^+CD8^+multimer^-$ as well as the standard deviation (SD) of the $CD3^+CD8^+multimer^-$ subsets were documented for all multimer fluorescence channels (PE, APC,

Q605 and Q705) and cell fractions. For the beads, median fluorescence and SD was determined. Further parameters (e.g. cell recovery after thawing, number of PBMC per stain, number of $CD3^+$ cells counted, details on staining reagents and cytometer configuration) were collected for inter-laboratory comparisons but are not presented in this article (a report of panel CIP_ID13_2012_MUL/D is available on request).

Central assessment of the panel data: All dot- or pseudocolor-plots of MHC multimer stains were analyzed and scored by 3 experienced flow users. The multimer score was given as follows: "0" = no $CD8^+$multimer$^+$ cell population, "1" = plausible $CD8^+$multimer$^+$ cell population or "2" = obvious $CD8^+$multimer$^+$ cell population, for a maximum total score of 6 (Figure 2). Only stains with a score of at least 4 were considered positive (i.e., detection of $CD8^+$multimer$^+$ cells) and hence included when the frequency of $CD8^+$multimer$^+$ cells was evaluated.

Data analysis and statistics: Cell samples: for the proficiency panel, frequencies of multimer$^+$ cells are expressed as % of $CD3^+CD8^+$ cells and were calculated using the cell numbers reported by the individual labs. To compare the brightness of the multimers coupled with the 4 different fluorochromes (Figures 1-3) or the staining with single multimers in combinations with various Ab (Figure 4), staining indexes (SI) were calculated as follows: SI_{cells} = (median fluorescence $CD3^+CD8^+$multimer$^+$ subset – median fluorescence $CD3^+CD8^+$multimer$^-$ subset) / 2 x fluorescence standard deviation of $CD3^+CD8^+$multimer$^-$ subset[4,7].

Calibration beads: we calculated for each participating lab two staining indexes that were used as sensitivity measures for flow cytometry instruments: $SI_{beads+FMO}$ was defined as = (median fluorescence of bright beads – median fluorescence of $CD3^+CD8^+$multimer$^-$ cell subset in the FMO cell staining / 2 x fluorescence standard deviation of $CD3^+CD8^+$multimer$^-$ cell subset in the FMO cell staining. $SI_{beads+Mult\ irrel.}$ is defined as = (median fluorescence of bright beads – $CD3^+CD8^+$multimer$^-$ cell subset of the A2*p multimer cell staining/ 2 x fluorescence standard deviation of $CD3^+CD8^+$multimer$^-$ cell subset of the A2*p multimer cell staining. $SI_{beads+FMO}$ can be considered as a measure of the accumulated effects of the flow cytometer settings, cell autofluorescence and Ab/fluorochromes (spreading error) included in the stain. $SI_{beads+\ Mult\ irrel.}$ can additionally provide a measure of the MHC multimer-related background.

Statistics: Statistical analyses, Spearman's rank correlation tests (Figures 2,3) and linear regressions (Figure 3) were performed with GraphPad Prism.

Bead-based optimization of MHC multimer staining

Two of the labs participating in the proficiency panel performed an extended MHC multimer optimization test. The experiment was conducted with small differences between the two labs. For lab ID04, prescreened PBMCs from 3 different donors were thawed, counted and each split in 4 fractions of ~$2*10^6$ cells. 2 fractions from each donor were stained with an antibody mix containing CD8, live dead stain (NIR – Invitrogen) and Dump channel markers CD4, CD14, CD16, CD19 and CD40 (FITC – BD Biosciences) as well as an APC- and a PE-labeled A2*p multimer at 1μg/ml (irrel multimer) and 2 fractions were stained only with the antibody mix (FMO). In order to determine the impact of using different CD8 antibodies on the multimer staining, either CD8-PerCP (Life Technologies) or CD8-Alexa Fluor 700 (BioLegend) was used in the abmix to identify CD8 cells and the 2 cell fractions in the irrel. multimer and the FMO groups were each stained with one of the two antibodies. For lab ID08 essentially the same was done, except that only one donor was used and that

it was split in 6 fractions instead of 4, in order to test the use of three different CD8 antibodies in the antibody mix. These were PE-Cy7, Qdot705 and Qdot605 (Invitrogen).

The staining procedure was performed in two successive steps, essentially following the CIP protocol (www.cimt.eu/workgroups/CIP). At first, the multimer staining was performed either at 37°C for 15 min or at room temperature for 30 min followed by a second staining with the antibody mix for 30 min on ice. Stained cells were acquired on LSR II or Canto II flow cytometers (BD Biosciences) equipped with FACSDiva software. Each cell fraction was acquired at two different PMT settings, before and after flow cytometer and PMT optimization according to Perfetto et al., 2012[8]. This optimization protocol was followed, and using the suggested cyto-cal and quantum simply cellular beads, the optimal PMT value for each specific fluorochrome was determined [8]. Before acquiring the cell samples, PMT voltages and compensations were adjusted for each fluorescence channel using unstained cells and compensation beads (BD Biosciences or Invitrogen?) labeled with antibodies or ArC Amine reactive compensation beads (Invitrogen) suitable for the NIR viability dye. The new optimal PMT values were kept as close to their optimum as possible while still being adjusted to fit with the antibody panel used, ensuring the least possible overlap between fluorochromes. In addition to the multimer stainings, the same calibration beads used in the proficiency panel were also run at the two different PMT settings.

Data analysis: A number of different SI were calculated equal to what was done for the proficiency panel data. $SI_{beads+FMO}$ was defined as = (median fluorescence of bright beads – median fluorescence of $CD3^+CD8^+$multimer$^-$ cell subset in the FMO cell staining / 2 x fluorescence standard deviation of $CD3^+CD8^+$multimer$^-$ cell subset in the FMO cell staining. $SI_{beads+Mult\ irrel.}$ is defined as = (median fluorescence of bright beads – median fluorescence of $CD3^+CD8^+$multimer$^-$ cell subset of the A2*p multimer cell staining / 2 x fluorescence standard deviation of $CD3^+CD8^+$multimer$^-$ cell subset of the A2*p multimer cell staining. These SIs were calculated from both PMT settings.

Results

Detection of MHC multimer binding T cells using differentially labelled MHC multimers
Across 16 different labs, we tested the ability to identify MHC multimer binding T cells when using four different fluorescent labels for MHC multimer detection, PE, APC, QD605 and QD705. We used 3 different donors and tested the detection of 4 different virus-epitope responsive T cell populations in each donor, CMV- NLVPMVATV, EBV- GLCTLVAML, EBV- YVLDHLIVV and FLU-GILGFVFTL (Tabel 1).

Donor	CMV	EBV1	EBV2	FLU
1	1.36%	0.18%	0.16%	0.06%
2	1.77%	0.14%	0.47%	1.95%
3	No response	0.08%	0.24%	0.06%

Table 1. Overview and size of the virus specific T cell responses present in each donor. Numbers represent the percentage of virus specific cells out of total $CD3^+CD8^+$ T cells.

Each T-cell population was detected using MHC multimers with all of the different fluorescent labels, so that no bias was introduced based on the type of T cell population detected. Based on a central evaluation of all 16 experiments by 3 gating experts, each possible multimer population with each of the four different fluorescent molecules as labels, was given a multimer score. Each expert would assign a score of 0-2 to each proposed T cell population representing whether they thought that the population was not present (0), maybe present (1) or obviously present (2). As the maximum multimer score was thus 6, a population with a score of at least 4 was considered to be a positive MHC multimer binding T cell population. Overall, we found that no difference in detection efficiency was evident when comparing the four different fluorescent labels (figure 1a). We furthermore investigated the frequency of MHC multimer binding T cells among total CD8 T cells detected for a given T cell response, when comparing the different labels. We found a minor difference in terms of decreased detection of FLU and CMV-specific T cells when analyzed using QD705-labelled MHC multimers (figure 1b, and supplementary figure 1). It is also evident that although the T cell detection across different MHC multimer-associated labels are comparable, the staining index (SI) of the particular population is largely dependent on the fluorescent label used (figure 1c). This is also evident from figure 1d showing the dot plot examples from one donor, where it is clearly seen that the separation between positive and negative events depends on the choice of label (figure 1d).

To further evaluate if difference in SI would have an impact on T cell detection, we again looked at the scores assigned to each T cell population by our gating expert panel and compared it to the SI of that particular population. Through this analysis, it was evident, that even across many different laboratories, the SI is correlated to the ability to detect a certain MHC multimer binding T cell population (figure 2a). Furthermore, it is also evident that larger populations of MHC multimer-binding T cells, are more easily detected than low frequent populations (figure 2b).
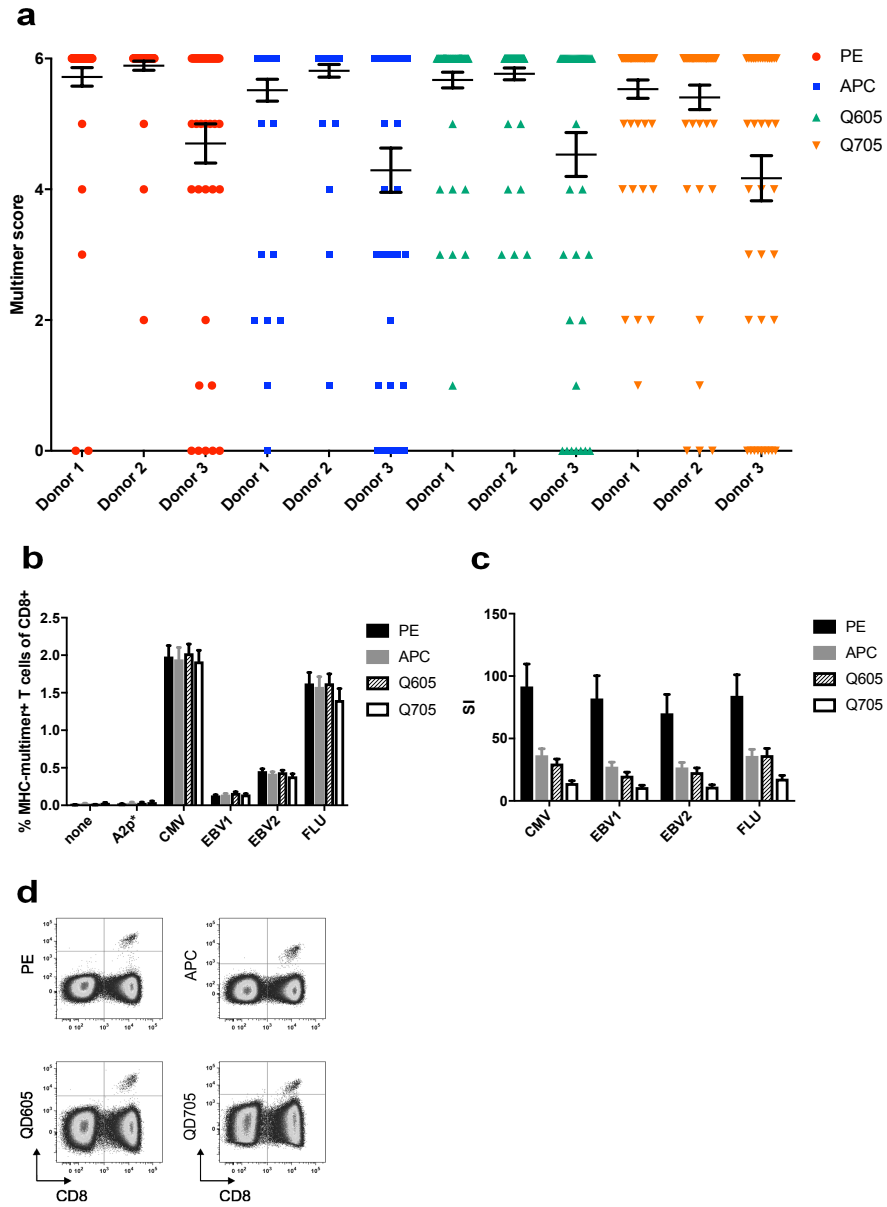
Figure 1. Influence of the fluorochrome coupled to peptide-MHC multimers on the detection of antigen-specific cells. **a)** Multimer score for all virus specific T cell populations in each donor using the multimers coupled to four different fluorochromes (n=11 cell populations per fluorochrome across the 3 donors). Each symbol represents an individual lab (n=16) and error bars indicate SEM. **b)** % $CD3^+CD8^+$multimer$^+$ and **c)** staining indices for one exemplary donor out of the 3 tested (donor 2, mean from all labs is shown (group analysis)). SEM is indicated. **d)** EBV2-multimer staining with each of the four fluorochromes is shown on donor 2 from one exemplary lab.
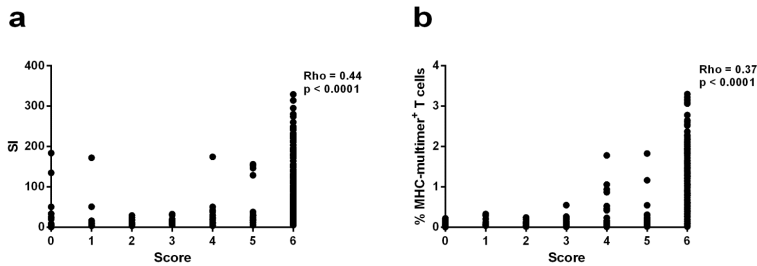
Figure 2. Correlates of multimer+ cell detection.
The positivity score attributed by the central assessment for each MHC-multimer+ T cell population (n=256 in total) is plotted in relation to **a)** the staining index and **b)** the % CD3+CD8+multimer+ of the same population. ρ and p values are indicated (spearman's rank correlation).

Evaluation and optimization of detection for selected fluorescent molecules

To further assess if differences in detection rate and SI of a given MHC multimer binding T cell population correlates to the efficacy of detection of the chosen fluorescent label, we evaluated the SI based on calibration beads labeled with each of the four fluorescent labels used, for each participating laboratory. The SI's were calculated for the beads against the negative population from both an FMO and an irrelevant multimer staining. We found that across different laboratories the SI for the different fluorescent labels varied around 3-fold (figure 3 a-d, and Supplementary figure 2 a-d). This was evident both when beads were assessed alone (data not shown), and when beads were evaluated in combination with the full detection panel (antibodies + MHC multimer negative control) as used when MHC multimer binding T cells were detected. We further analyzed the correlation between the bead-determined SI for the given fluorescent label and the SI of the MHC multimer binding T cell population. For 3 out of 4 fluorescent labels (all except APC), there was a correlation between the bead-based determined SI and the MHC multimer SI (figure 3 e-f, supplementary figure 3). Consequently, we argue that beads can be used to evaluate and optimize the SI for MHC multimer associated fluorescent labels.

We initiated an optimization procedure for two different laboratories, ID04 and ID08, with the aim to enhance the SI for APC and PE labeled multimers, respectively. We attempted to optimize two parameters to potentially enhance the SI, 1) PMT values according to the PMT performance evaluation published by Perfetto et al. and 2) the fluorescent labels used in the antibody panel to identify CD8 T cells. Following evaluation and adjustment of the PMT values, the new and old values were compared by running bright PE- and APC-labeled beads as well as a number of multimer stainings through the flow cytometer at both settings. For lab ID04, donor PBMCs were divided in four fractions. Two fractions were stained with an antibody mix (abmix) containing CD8 and dump channel markers and an irrelevant multimer (Mult irrel.), whereas the other two fractions were stained with only the abmix (FMO). One of the two fractions in each condition was stained with an Alexa Flour 700 labeled CD8 antibody, and the other fraction with a PerCP labeled CD8 antibody. This was done, in order to take the fluorescent label used to detect CD8 into account when attempting to optimize the APC multimer detection. The same was done for lab ID08, except three different CD8 antibodies were tested, PE-Cy7, Qdot705 and Qdot605, in order to optimize detection of the PE multimer. Each cell fraction was run through the flow cytometer at both the new and old PMT settings, as was the calibration beads. This was done for 3 healthy donors (fig. 4a and b) or 1 healthy
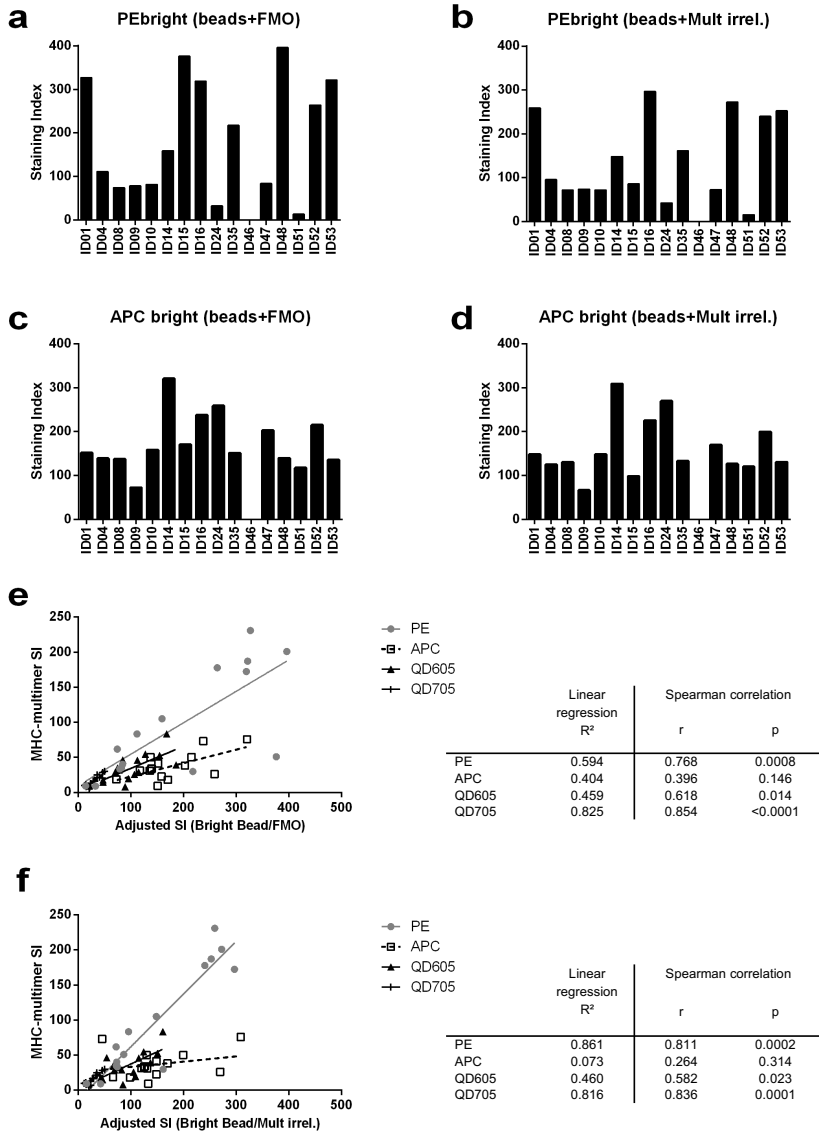
Figure 3. Interlaboratory performance based on detection of fluorescent calibration beads.
Staining indices were calculated for the bright-fluorescence beads and are shown for the PE and APC channels $SI_{beads+FMO}$ **(a,c)** and $SI_{beads+ Mult irrel.}$ **(b,d)** are shown for each individual lab. **e)** and **f)** Correlation analyses between multimer staining indices calculated from PBMC stains with each of the four multimer-flurochromes and $SI_{beads+FMO}$ **(e)** or $SI_{beads+ Mult irrel.}$ **(f)**. Statistics are shown next to the plots.
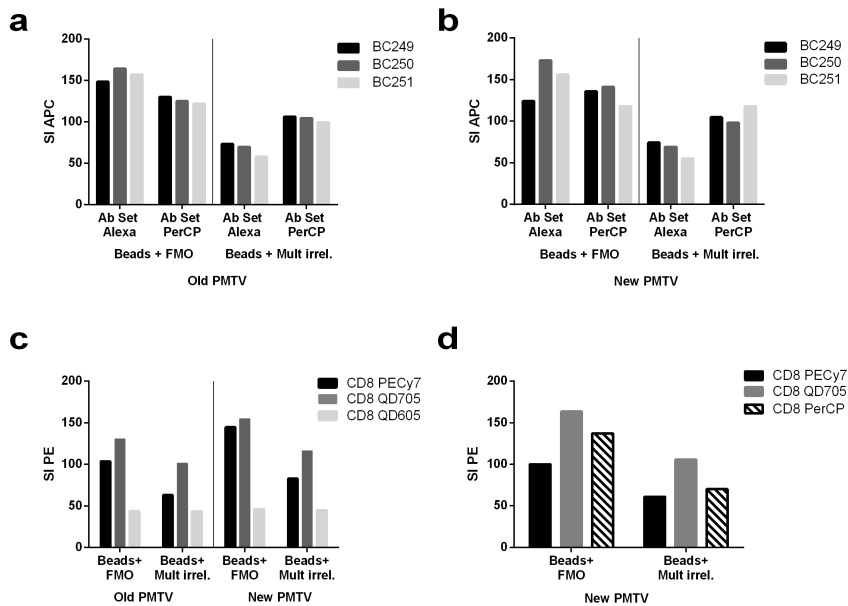
Figure 4. Effect of PMT and CD8 antibody optimization.
Staining indices for APC, calculated based on bright bead+FMO or bright bead+Mult irrel, are shown **(a,b)**. Two different Ab panel (Ab Set Alexa, Ab Set PerCP) and PMT voltages (PMTV old and new) were tested (**a** and **b**, respectively) on three different donors. **c)** and **d)** show the staining indices for PE also calculated based on bright bead+FMO or bright bead+Mult irrel. **c)** Three different Ab sets (CD8 PECy7, -QD705, -QD605) were tested and compared across two different PMTV (old, new) in one donor. In **d)** staining indices were calculated only at the new PMTV settings and a fourth Ab set (PerCP) was tested on one donor.

donor (fig. 4c and d). 2 different SIs were calculated and compared both using the bright calibration bead for the positive population. The negative population was either taken from the FMO staining – in order to see the effect of the abmix on the negative population – or from the Mult irrel staining, indicating whether the addition of the multimer resulted in unspecific staining of the negative cells and hence had an effect on the separation between the negative population and the positive bead as compared to the FMO. When looking at fig.4 it is evident that this is in fact the case, as the beads+mult irrel SIs are generally lower than FMO SIs. When comparing the APC SI between new and old PMT settings, no dramatic differences are visible (fig.4a and b). In some donors, the optimization seems to have resulted in a small rise in the SI, whereas in other donors it seems to have had the opposite effect. Changing the CD8 antibody from AlexaF700 to PerCP caused a small increase in APC SI, but only in relation to the mult irrel staining. For the PE SI, the results are shown only for one donor where we again tested the use of different CD8 labels in the abmix, this time both PE-Cy7, QD705 and QD605 were included before and after optimization (fig. 4c). Again, no dramatic effect of the PMT optimization was detectable, though there was a tendency towards a higher SI with the new PMT settings. The PE-Cy7 and QD705 CD8 antibodies were compared to the PerCP label only at the new PMT settings (fig. 4d). We found very different SIs for PE depending on the fluorescent label used to detect CD8. QD705 seems to provide a good SI increasing little after PMT optimization, whereas QD605 gives consistently and equally low SI both before and after PMT optimization and across the 2 different staining conditions. Using PE-Cy7 or PerCP yielded SIs somewhat in between these two.

Discussion

In this study, we investigate the impact of fluorescent label choice on the detection of MHC multimer-binding T cells. What we found was that the overall detection of the multimer specific cells was not affected by the choice of fluorochrome. The staining index was, however, highly dependent on the chosen label as was evident from figure 1. The SI is a measure of the separation of the positive signal from the background events as well as the spread of the negative population, and as such, a low staining index indicates a population with poor separation from the background. Such a population can be difficult to identify as being positive, which is also reflected by the correlation between the staining index and the score of each population in figure 2a. Thus, in order to gain the best possible results from an MHC multimer staining, it is crucial to optimize the experiment and take into consideration the brightness of the fluorochromes used. There are of cause other things than the SI to consider when designing an experiment. In combinatorial encoding experiments where multiple fluorochromes are used simultaneously, it is also worth taking into account the possible spillover of fluorescent labels into other channels. The strength of the dim fluorochromes used in this study, the Qdots, is that they have a very narrow emission spectrum and may thus be useful when having a complex experiment with many parameters as they are less likely to cause spillover into other channels than fluorochromes with wide emission spectra. Thus, choosing the best fluorescent labels is often a matter of tradeoff between complexity and sensitivity. With the recent increase in fluorochrome development, especially bright fluorochromes, it is becoming easier to maintain experiment complexity without inclusion of low intensity fluorochromes. In this particular study, we detect virus specific T cell populations that often has high avidity interactions between the TCR and the MHC. We found that even when the cell populations were labeled with low intensity fluorochromes they could still be detected despite the low SI. It is, however, highly likely that T cell populations with low avidity between the TCR and MHC, such as typical cancer specific cells, will be difficult to distinguish from the background if labeled with a low SI fluorochrome, as we do see a correlation between SI and detection of MHC multimer-binding T cells (fig. 2).

In this study, we propose a bead based tool to identify detection weaknesses and optimize the SI for a given multimer experiment setup. It serves as a fast and easy alternative to the somewhat laborious task of generating multimers labeled with the different fluorochromes intended to be in the experiment, followed by test stainings on donor material and calculation of the SI from each fluorochrome. As we observed a correlation between SI from multimer staining to that of the corresponding premade beads (fig. 3c and f), we suggest to use the beads to optimize the MHC multimer experiment. These premade beads can simply be taken from the freezer and run through the flow cytometer in order to give not only an idea of the intensity of the different fluorochromes, but also identify weak detection channels. In our study, the same beads run at 16 different laboratories showed great variance in SI between laboratories (fig. 3a-d). This is most likely a consequence of different detection channels and configurations on the various flow cytometers used and highlights the fact that it is crucial to optimize each experiment to the relevant instrument. Additionally, it is very likely that the different handling and staining protocols at the 16 different labs also have an impact on the SI.

Optimization of PMT is one parameter that can be adjusted to improve fitness of the flow instrument for detection of a given fluorophore. Although the PMT optimization test performed in this study had little effect, it provides a standardized set of PMT values, ensuring consistency for all users of an instrument. A major contributor to the MHC multimer channel SI, was the different fluorescent labels used to detect CD8 cells. We found that under the same conditions, different fluorescent labels gave SIs ranging from just under 50 to over 150, emphasizing again, the importance of optimizing the fluorochromes utilized in a given experiment.

Author contributions

NWP performed experiments, analyzed data, generated figures and wrote the manuscript, KL performed experiments, analyzed data, generated figures and revised the manuscript, DM conceived the concept, performed experiments, analyzed data, SW conceived the concept, performed expert gating and analyzed data, CG conceived the concept, performed expert gating, discussed data and revised the manuscript, SRH conceived the concept, performed expert gating, analyzed data, generated figures and wrote the manuscript.

Conflict of interest statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Aknowledgements

References

1.   Hadrup, S. R. *et al.* Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat. Methods* **6,** 520–526 (2009).

2.   Andersen, R. S. *et al.* Parallel detection of antigen-specific T-cell responses by combinatorial encoding of MHC multimers. *Nat.Protoc.* (2012).

3.   Britten, C. M. *et al.* The CIMT-monitoring panel: A two-step approach to harmonize the enumeration of antigen-specific CD8+T lymphocytes by structural and functional assays. *Cancer Immunol. Immunother.* **57,** 289–302 (2008).

4.   Hadrup, S. R. *et al.* Cryopreservation of MHC multimers: Recommendations for quality assurance in detection of antigen specific T cells. *Cytom. Part A* **87,** 37–48 (2015).

5.   Britten, C. M. *et al.* The CIMT-monitoring panel: a two-step approach to harmonize the enumeration of antigen-specific CD8+ T lymphocytes by structural and functional assays. *Cancer Immunol. Immunother.* **57,** 289–302 (2008).

6.   Gouttefangeas, C. *et al.* Data analysis as a source of variability of the HLA-peptide multimer assay: from manual gating to automated recognition of cell clusters. *Cancer Immunol. Immunother.* **64,** 585–98 (2015).

7.   Maecker, H. T., Frey, T., Nomura, L. E. & Trotter, J. Selecting fluorochrome conjugates for maximum sensitivity. *Cytometry. A* **62,** 169–73 (2004).

8.   Perfetto, S. P., Ambrozak, D., Nguyen, R., Chattopadhyay, P. K. & Roederer, M. Quality assurance for polychromatic flow cytometry using a suite of calibration beads. *Nat. Protoc.* **7,** 2067–2079 (2012).

Supplementary figures
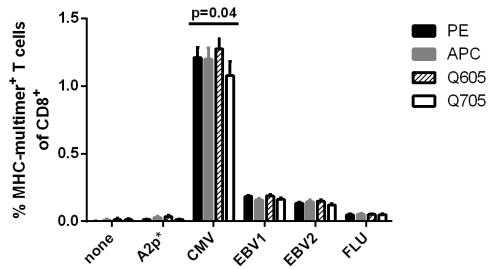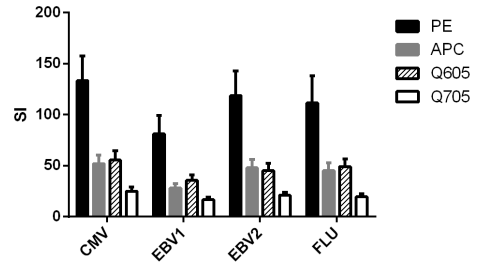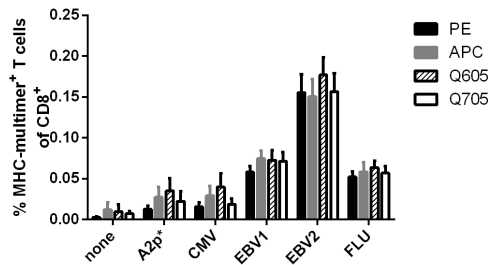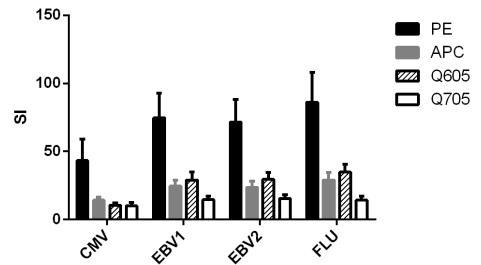
Supplementary figure 1.
a) and c) shows group analysis within two different donors (donor 1, 3) for % of MHC-multimer+ T cells. The corresponding staining index for each fluorochrome is shown in **b)** and **d)**, SEM is indicated. Next to the four different antigens (CMV, FLU, EBV1 and EBV2) two controls are included, FMO control (none) and negative control multimer (A2p*).

a

QD605 bright (beads+FMO)

b

QD605 bright (beads+Mult irrel.)

c

QD705bright (beads+FMO)

d

QD705bright (beads+Mult irrel.)
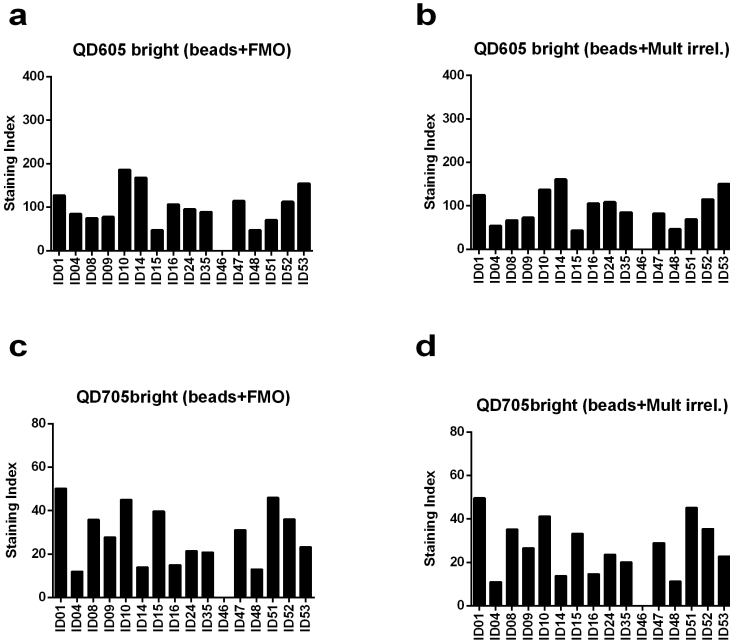
Supplementary figure 2.
Staining indices for QD605 and QD705 were calculated as bead+FMO - **(a,c)** and bead+Mult irrel.-derived staining indices **(b, d)** in individual labs. Exemplary results were shown for bright beads.

	Linear regression R^2	Spearman correlation	
		r	p
PE	0.349	0.479	0.071
APC	0.040	-0.350	0.201
QD605	0.482	0.611	0.016
QD705	0.362	0.675	0.0058
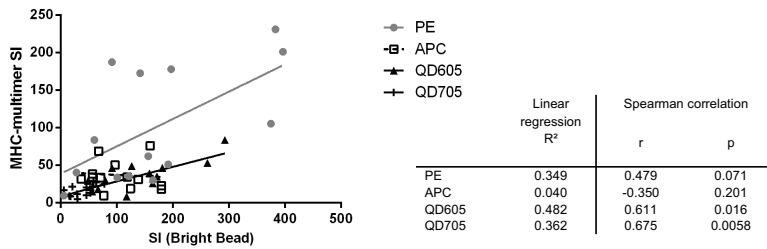
Supplementary figure 3.
Correlation between cellular MHC-multimer staining indices and bright bead-derived staining indices at each lab is shown for four fluorochromes used to label multimers. Statistics is shown next to the plot.

CHAPTER 3

Paper II

Automated Analysis of Flow Cytometry Data to Reduce Inter-Lab Variation in the Detection of MHC Multimer Binding T Cells

Automated Analysis of Flow Cytometry Data to Reduce Inter-Lab Variation in the Detection of Major Histocompatibility Complex Multimer-Binding T Cells

*Natasja Wulff Pedersen[1], P. Anoop Chandran[2], Yu Qian[3], Jonathan Rebhahn[4], Nadia Viborg Petersen[1], Mathilde Dalsgaard Hoff[1], Scott White[5], Alexandra J. Lee[3], Rick Stanton[6], Charlotte Halgreen[7], Kivin Jakobsen[7], Tim Mosmann[4], Cécile Gouttefangeas[2], Cliburn Chan[5], Richard H. Scheuermann[3,8] and Sine Reker Hadrup[1]**

[1] Division of Immunology and Vaccinology, Veterinary Institute, Technical University of Denmark, Copenhagen, Denmark, [2] Department of Immunology, Interfaculty Institute for Cell Biology, University of Tuebingen, Tuebingen, Germany, [3] Department of Informatics, J. Craig Venter Institute, La Jolla, CA, United States, [4] David H. Smith Center for Vaccine Biology and Immunology, University of Rochester Medical Center, Rochester, NY, United States, [5] Department of Biostatistics and Bioinformatics, Duke University Medical Center, Durham, NC, United States, [6] Human Longevity Inc., San Diego, CA, United States, [7] Immudex Aps, Copenhagen, Denmark, [8] Department of Pathology, University of California, San Diego, La Jolla, CA, United States

Manual analysis of flow cytometry data and subjective gate-border decisions taken by individuals continue to be a source of variation in the assessment of antigen-specific T cells when comparing data across laboratories, and also over time in individual labs. Therefore, strategies to provide automated analysis of major histocompatibility complex (MHC) multimer-binding T cells represent an attractive solution to decrease subjectivity and technical variation. The challenge of using an automated analysis approach is that MHC multimer-binding T cell populations are often rare and therefore difficult to detect. We used a highly heterogeneous dataset from a recent MHC multimer proficiency panel to assess if MHC multimer-binding CD8[+] T cells could be analyzed with computational solutions currently available, and if such analyses would reduce the technical variation across different laboratories. We used three different methods, FLOw Clustering without K (FLOCK), Scalable Weighted Iterative Flow-clustering Technique (SWIFT), and ReFlow to analyze flow cytometry data files from 28 laboratories. Each laboratory screened for antigen-responsive T cell populations with frequency ranging from 0.01 to 1.5% of lymphocytes within samples from two donors. Experience from this analysis shows that all three programs can be used for the identification of high to intermediate frequency of MHC multimer-binding T cell populations, with results very similar to that of manual gating. For the less frequent populations (<0.1% of live, single lymphocytes), SWIFT outperformed the other tools. As used in this study, none of the algorithms offered a completely automated pipeline for identification of MHC multimer populations, as varying degrees of human interventions were needed to complete the analysis. In this study, we demonstrate the feasibility of using automated analysis pipelines for assessing and identifying even rare populations of antigen-responsive T cells and discuss the main properties, differences, and advantages of the different methods tested.

Keywords: major histocompatibility complex multimers, antigen-specific T cells, automated gating, computational analysis, major histocompatibility complex dextramers, flow cytometry

INTRODUCTION

Antigen-specific T cell recognition is an essential component of the adaptive immune response fighting infectious diseases and cancer. The T cell receptor (TCR)-based recognition profile of a given T cell population can be determined through interaction with fluorescently labeled multimerized peptide major histocompatibility complexes (pMHC multimers) (1), enabling visualization of specific pMHC-responsive T cells by flow cytometry (2). This analysis has become state of the art for antigen-specific CD8[+] T cell detection and is important for pathophysiological understanding, target discovery, and diagnosis of immune-mediated diseases. Detection of pMHC-responsive T cells is challenged by the low-avidity interaction between the TCR and the pMHC, often resulting in poor separation of fluorescent signals distinguishing the MHC multimer-binding from non-binding T cells (3). Additionally, a given antigen-specific T cell population is in most cases present at low frequencies in the total lymphocyte pool (4).

Substantial effort has been applied to optimize and standardize protocols for pMHC multimer staining of antigen-specific T cells to ensure the best possible signal-to-noise ratio in such T cell assays. The Immunoguiding Program of the European Association of Cancer Immunotherapy (CIP) has been actively involved in this process, and through a series of proficiency panels, identified the parameters largely impacting the variation in such assays (5–8). Among these, individual gating strategies lead to significant variation in final results determining the frequency of pMHC-responsive T cells (9). To minimize gating-associated variation and manual handling as well as to improve standardization, several automated analysis strategies have been developed to analyze flow cytometry data based on computational assessments of the different parameters involved (10, 11). These algorithms are based on computational identification of cell clusters in multidimensional space, taking into account all the different parameters applied to a certain cell type. Hence, they consider all associated parameters simultaneously, which forms an additional advantage compared with sequential 2D determinations of "positive" or "negative" categories, and consequently leads to a potentially improved identification of a given cell population.

The performance of automated analysis tools has been investigated in a number of challenges reported by the FlowCAP consortium (11–13), but such algorithms have so far not been evaluated for identification of MHC multimer-binding T cells. The aim of the present study was to test the feasibility and to report the experience of using automated analysis tools for identification of antigen-specific T cells. Tools were selected based on (a) the requirement of a user-friendly interface, making them accessible to flow cytometry users without computational expertise and (b) the described ability to detect rare cell populations. Three software solutions were chosen based on these criteria: FLOw Clustering without K (FLOCK) (14), Scalable Weighted Iterative Flow-clustering Technique (SWIFT) (15–17), and ReFlow (18, 19), but several others may be available having similar characteristics. FLOCK is a grid-based density clustering method for automated identification of cell populations from high-dimensional flow cytometry data, which is publicly accessible through the Immunology Database and Analysis Portal (ImmPort) at http://immport.niaid.nih.gov (now moved to https://www.immportgalaxy.org/). SWIFT is a model-based clustering method that is specifically developed to identify rare cell populations. The algorithm goes through three stages of fitting the cell populations to Gaussian distributions, splitting, and merging the populations to reach unimodality. The clustered output files given by SWIFT can either be analyzed by manual cluster gating or by automatically analyzing the cluster output. It is publicly available through http://www.ece.rochester.edu/projects/siplab/Software/SWIFT.html but requires Matlab software. ReFlow is a repository and automated analysis platform for flow cytometry data that is currently available as open source with web-based access and shared GPU computation (18, 19). It employs the hierarchical Dirichlet process Gaussian mixture model that naturally generates an aligned data model to capture both commonalities and variations across multiple samples, for the identification of unique cell subsets in an automated fashion (19). We evaluated the selected algorithms for their ability to identify pMHC multimer-binding T cells compared with manual gating, using data from a recent MHC multimer proficiency panel organized by Immudex[1] in collaboration with CIP.[2] We analyzed MHC Dextramer™ staining of T cells recognizing two different virus-derived epitopes [Epstein–Barr virus (EBV) HLA-A*0201/GLCTLVAML and influenza (FLU) HLA-A*0201/GILGFVFTL] in peripheral blood mononuclear cells (PBMCs) from two healthy donors. Furthermore, data from two sets of spike-in samples were used. The overall goal was to evaluate the feasibility and limit of detection of these three different algorithms that are readily available to flow users without pre-existing computational expertise.

MATERIALS AND METHODS

Production of MHC Multimers

HLA-B*0702/TPRVTGGGAM monomers used in the spike-in 1 experiment were generated using UV-mediated peptide exchange as previously described (20). In short, HLA-B*0702 monomers carrying a UV-sensitive peptide were mixed with TPRVTGGGAM peptide in a final concentration of 100 μg/ml monomer and 200 μM peptide and kept under UV light for an hour. The resulting HLA-B*0702/TPRVTGGGAM monomers were then multimerized using phycoerythrin (PE)-streptavidin (BD Biosciences). The multimers were frozen at −80°C in freezing buffer giving a final multimer concentration of 10 μg/ml with 0.5% Bovine Serum Albumin (Sigma-Aldrich) and 5% glycerol (Fluka).

Abbreviations: APC, allophycocyanin; CIP, Immunoguiding Program of the Association for Cancer Immunotherapy; CMV, cytomegalovirus; CV, coefficient of variation; DPGMM, Dirichlet process Gaussian mixture model; EBV, Epstein–Barr virus; FLU, influenza; MHC, major histocompatibility complex; TCR, T cell receptor; PBMCs, peripheral blood mononuclear cells; PE, phycoerythrin; pMHC, peptide MHC.

---

[1]www.immudex.com/proficiency-panels.
[2]www.CIMT.eu/CIP.

For the spike-in 2 experiment, HLA-A*0201/NLVPMVATV and HLA-A*0201/GILGFVFTL monomers were generated using classical refolding (1) and multimerized using streptavidin-PE or streptavidin-allophycocyanin (APC) (Life Technologies), respectively, at a 4:1 molar ratio. After the addition of 1 mM biotin (Sigma-Aldrich), the multimers were aliquoted and frozen at −80°C in a freezing solution containing 1.7% human serum albumin (Albiomin®, Biotest, Dreieich, Germany), 0.07% sodium azide, 3.4× protease inhibitor (Complete™, Sigma-Aldrich), 42% v/v glycerol (Roth), and 7 mMTBS, such that the final mixture contained 14% (v/v) glycerol (7). The stock concentrations of PE- and APC-conjugated multimers were 310 and 485 µg/ml, respectively.

Donor Material

Peripheral blood mononuclear cells from healthy donors were obtained from buffy coats (blood products) collected at the local blood bank. All procedures were approved by the local Scientific Ethics Committee. PBMCs were isolated from buffy coats by density centrifugation on Lymphoprep (Axis-Shield PoC), and cryopreserved at −150°C in fetal calf serum (FCS; Gibco) + 10% DMSO.

Spike-in Cell Samples

FCS files from two different spike-in experiments were used in this study, spike-in 1 and spike-in 2. For spike-in 1, one PBMC sample from donor BC260 (HLA-B*0702 positive) carrying a CD8 T cell response of 1.7% of single, live lymphocytes against the cytomegalovirus (CMV) HLA-B*0702/TPRVTGGGAM epitope, was mixed into donor BC262 (HLA-B*0702 negative). Starting at 100% of the BC260 donor, a titration series was generated with fivefold dilutions going from 1.7 to 0.0001% of single, live lymphocytes. Cells were stained with PE- and APC-labeled pMHC multimers and an antibody mix containing a live/dead stain (NIR—Invitrogen), CD8 (PerCP—Life Technologies), and FITC-conjugated dump channel antibodies (CD4, CD14, CD16, CD19, and CD40—BD Biosciences) in order to identify CD8+MHC multimer+ T cells (2). For spike-in 2, one PBMC sample from donor B1054 (HLA-A*0201 positive) was mixed into donor B1060 (HLA-A*02 negative) in nine steps using twofold dilutions. Sample 1 contained only cells from B1054 with high and intermediate frequencies of T cells responsive toward the CMV HLA-A*0201/NLVPMVATV and FLU HLA-A*0201/GILGFVFTL epitopes, respectively. Sample 9 contained only cells from B1060. Cells were stained with PE-labeled CMV multimer and APC-labeled FLU MHC multimer.

MHC Multimer Proficiency Panel

FCS files used in this study were from 28 different laboratories who participated in an MHC multimer proficiency panel organized by Immudex. Originally, 51 labs participated in the proficiency panel but only 28 labs made their FCS files available for our analysis. The individual labs were anonymized and given an ID number. Each lab received two PBMC samples from each of two donors—518 and 519—and MHC Dextramers specific for EBV HLA-A*0201/GLCTLVAML, FLU HLA-A*0201/GILGFVFTL or an irrelevant peptide HLA-A*0201/ALIAPVHAV (NEG). Each lab used their own antibodies, staining protocols, and gating strategies, which

varied significantly from lab to lab. As a result, the number and type of parameters included by each lab varies to a great extent, but as a minimum all labs included CD3, CD8, and multimer staining or dump, CD8 and multimer staining, using various antibodies. The two donors used held T cell responses against the EBV and FLU-derived T cell epitopes, including both low-frequency responses (0.04 and 0.09% multimer+ CD8+ T cells), a medium (1.13% multimer+ CD8+ T cells), and a high-frequency response (5.33% multimer+ CD8+ T cells) as defined by a pretest on eight donor samples performed at two different locations with insignificant variation. All samples were run in duplicates giving a total of 12 FCS files from each lab. All labs gated their files manually and reported the percentage of identified multimer+ CD8+ T cells of the total number of CD8+ cells. The percentage of MHC multimer+ T cells was reported as the mean of the duplicate analysis. Exceptions to this were lab 104 which only provided files from one analysis run, as well as lab 235 and lab 240 where the 518-EBV and 519 FLU samples, respectively, were only included in one run. For these labs, the value from the single run was used instead of the mean value.

Central Manual Gating

A central manual gating was performed on all FCS files by one operator. SSC-A/FSC-A was used to identify lymphocytes and FSC-H/FSC-A to identify singlets. Of the 28 labs in this study, 17 labs included a live/dead stain in their analysis and 11 did not. From single, live lymphocytes or single lymphocytes the number of CD3+, CD8+, and MHC multimer+ cells were identified and reported. The percentage of multimer+ T cells was calculated both from CD8+ cells and from total single (live) lymphocytes. For lab 215, the live/dead stain was included in a dump channel stain (CD14, CD16, and CD20); thus, the percentage of multimer+ T cells was calculated from single, live, non-dump lymphocytes. The percentage of multimer+ T cells reported was the mean percentage calculated from the duplicate analysis. FACS DIVA 8.0 software (BD Biosciences) was used for manual gating and the gated FCS files were exported in FCS 2.0 format.

Manual Pregating

Prior to automated analysis in FLOCK and SWIFT, the FCS files were gated manually in order to select single lymphocytes or single live lymphocytes (when a live/dead stain was included). Throughout the study, the term pregating is used when referring to manual pregating.

Manual Postgating

SWIFT analysis was performed on raw FCS files and cluster gating was performed on the SWIFT output files to obtain single lymphocytes or single live lymphocytes (when a live/dead stain was included) before identifying the multimer population as described in the SWIFT pipeline section. Throughout the study, postgating is used when referring to manual postgating.

Automated Prefiltering

Automated prefiltering was included as an automated alternative to manual pre- or postgating. The same selection was applied

as described for manual pregating. The automated prefiltering method we developed for FLOCK and SWIFT, named Directed Automated Gating (DAG), is a 2D by 2D density-based data prefiltering method. The sequence of the 2D dot plots used in the DAG prefiltering is specified in a user-configurable file, which also includes coordinates of a rectangle gate on the 2D dot plot. DAG automatically calculates a set of density contour lines based on the data distribution on the 2D dot plot. The events that are inside the largest density contour line within the rectangle gate will be kept and passed to the next filtering step, until the sequence of the 2D dot plots is fully traversed. DAG is implemented in Matlab and is publicly accessible at Github under GPL3.0 open source license.[3] Throughout the study, the term prefiltering is used when referring to automated prefiltering.

FLOCK Pipeline

FCS files were uploaded to FLOCK at www.immport.niaid.nih.gov and joined in datasets for each individual lab. The files were then initially analyzed as a dataset using FLOCK version 1.0 with the parameters set at auto. Unused markers/channels were excluded from the FLOCK analysis as were scatter parameters and parameters that were part of the manual or automated prefiltering. All other parameters included in the stainings performed by individual labs, which were as a minimum CD3, CD8, and MHC multimer or dump, CD8, and MHC multimer, were used for clustering. FLOCK then automatically assigned the values 1–4 (1: negative, 2: low, 3: positive, 4: high) for categorizing expression levels of each marker based on the relative expression level of the given marker on each identified cell population. A file with a large and easily definable MHC multimer[+] population (in most cases the 519 EBV sample) was then chosen to be a reference sample and the centroid information for this sample was saved. Using the cross-comparison feature, the other samples were then analyzed again with the centroid from sample 519 EBV as a reference. From the output of cross comparison, the summary table was downloaded and imported into excel where the intensity level of each marker in each population was used to define the MHC multimer[+] population. In order to identify which FLOCK clusters are the CD8[+], MHC multimer[+] cells, the expression level cutoff was set at >1 for CD3 (not included in all labs), >1 for CD8, and >2 for MHC multimer. The percentage of MHC multimer[+] cells of the total single, live lymphocyte population was then calculated and noted, and the mean percentage calculated from the duplicate analysis. The same cutoff value could not be used to identify the CD8 population in samples coming from different labs most likely due to the large variation in fluorochromes used to stain for CD8 cells between individual labs. The cutoff value for the CD8 marker was consequently set very low (>1), including also cells with low CD8 expression into the CD8 population. In many samples, this lead to the inclusion of too many cells into the CD8 population, thereby skewing the frequency of MHC multimer[+] cells when calculated as a percentage of the CD8 population. As a consequence, the CD8 marker was used only for identifying the true MHC multimer-binding

population and not as the base for calculating the frequency of the population, which was instead done using the number of live, single lymphocytes. All FCS files from the 28 labs were analyzed using FLOCK. For three labs (105, 215, and 253), FLOCK analysis resulted in the identification of MHC multimer populations in the negative control samples comprising 20–50% of live, single lymphocytes, and the three labs were therefore considered to be extreme outliers and consequently removed from the analysis of the negative samples.

SWIFT Pipeline

SWIFT version 3 was downloaded through www.ece.rochester.edu/projects/siplab/Software/SWIFT.html and the SWIFT folder was placed in the Matlab folder. In Matlab, the code swift_fcs_combine was used to generate a consensus file of all samples within each lab. In the FCS combine window, 250.000 cells from each of the 12 samples were chosen to be in the concatenated sample, giving a total of 3×10^6 cells. According to SWIFT online tutorials, the optimal range of cell numbers in a sample is $2-5 \times 10^6$. For labs where the nomenclature was not consistent between samples within the given lab, the code swift_modify_channels was used to uniformly name the channels in all files, prior to creation of the consensus FCS file. The concatenated consensus file was clustered using the code swift_main, generating a template file that was then used as a reference to cluster all 12 samples from a given lab with the code swift_assign_main. All parameters contained within a given sample were used for clustering, including the parameters that were part of manual or automated prefiltering. The input cluster number was kept at default settings—100 for all labs—and all unused channels/markers or channels included in the prefiltering were unchecked in both the Dims to Cluster and Output Medians columns. The ArcSinh Factors and Percent Noise were kept at default settings for all fluorescence channels. In the end, the output clustered FCS files were analyzed manually using FlowJo version 10 (Tree star) to obtain the number of CD3, CD8, and MHC multimer[+] cells or the number of non-dump, CD8, and MHC multimer[+] cells. Twenty-seven labs were analyzed with SWIFT, lab 208 was left out due to incompatibility of the FCS format with the software. In the analysis of sample 519 FLU for **Figure 4C**, lab 133 was left out, as it was an extreme outlier.

ReFlow Pipeline

All FCS files were uploaded on ReFlow and each lab was analyzed individually. The clustering variables assigned values as follows for both Stage 1 and Stage 2; burn in: 10,000, cluster count: 32, iteration count: 1,000, and sub-sampling count: 20,000. Stage 1 clustering was performed using FSC-A, SSC-A, and live/dead marker (when available). Live lymphocyte clusters were selected manually and Stage 2 clustering was performed using the CD8 and multimer-PE parameters. Singlets were not discriminated in the ReFlow stage 1 clustering as it is not advisable to use more scatter parameters than already used to identify lymphocytes. The multimer[+] populations were chosen manually based on visual inspection of a 2D (CD8 versus multimer) representation of the clustered data. Frequency of multimer[+] clusters (sum of frequencies when more than one cluster) were exported as a .csv

---

[3]https://github.com/maxqian/DAG.

file and were used for analysis. Out of the 28 labs included in the study, ReFlow was unable to analyze labs 133, 208, 239, and 254 due to compensation issues, thus 24 labs were analyzed with ReFlow. After ReFlow clustering Lab 224 was found to be an extreme outlier and was consequently removed from the statistical analysis, giving a total of 23 labs in the final analysis.

Analysis and Statistics

The gating analysis that was performed in this study was carried out by two different immunologists. Central manual gating, FLOCK, and SWIFT analyses were performed by NWP whereas ReFlow analysis was performed by AC.

Statistical analyses were performed using GraphPad Prism 7 and R 3.3.2. A paired t-test was used to test for differences among the different algorithms, and correlations were calculated using Pearson correlations. In R, the package cvequality_0.1.1 was used to perform an asymptotic coefficient of variation (CV) equality test. For all tests, it was assumed that the data were sampled from Gaussian populations. The normal distribution was explored in R using a boxcox transformation, suggesting a log transformation of the data. All statistical tests were therefore also performed on log transformed data but gave the same results,

except for the asymptotic CV test in **Figure 4B**. When using the log transformed data, FLOCK and ReFlow software also resulted in significantly higher variation compared with manual gating for the 519 FLU population.

RESULTS

Individual Gating as a Source of Variation in the Assessment of MHC Multimer-Binding T Cells

To assess the impact of individual manual gating compared with central manual gating on specific T cell identification and quantification, FCS data files obtained from the MHC multimer proficiency panel were re-analyzed manually by the same operator. The frequency of MHC multimer[+] cells within CD8[+] cells, reported by each lab (individual manual analysis) was compared with the respective frequencies determined after central manual analysis. For all four cell populations: 518/EBV, 519/EBV, 518/FLU, and 519/FLU, no significant difference in the determined frequency was observed between manual individual and central gating (**Figure 1A**). The highest CV was observed for

FIGURE 1 | Individual versus central manual gating. **(A)** Percentage of multimer positive cells (EBV or FLU) in total CD8[+] T cells in two healthy donors (518 and 519) identified through individual or central manual gating. Each dot represents the mean value for duplicate experiments for an individual lab, $n = 28$. Line indicates mean and error bars indicate SD. No significant difference between individual gating and central gating was detected (paired t-test). **(B)** The coefficient of variation (CV = SD/mean*100) related to the identification of major histocompatibility complex multimer positive T cell populations either through individual gating (green) or central manual gating (blue) for the two virus responses and two donors. No differences are statistically significant (asymptotic CV equality test). **(C)** Correlation of the percentage of multimer positive cells found with individual and manual gating. $p < 0.0001$ (Pearson correlation), $n = 112$. Mean values from duplicate experiments are shown. Different colors represent different populations. Individual: gating is done by each individual lab. Central: gating on all files is performed by the same person. 519: healthy donor 519; 518: healthy donor 518; EBV: Epstein–Barr virus; FLU: influenza virus.

the lowest frequency (519/FLU) population, but no statistically significant difference between individual and central manual gating was found (CV = 122% and CV = 86%, respectively) (**Figure 1B**). Previous data have shown that centralizing the gating may reduce the %CV compared with individual gating (9). Furthermore, a recent publication reported a similar observation that the infrequent and poorly resolved cell populations can be highly variable across samples when individual manual gating analysis is used (21). Additionally, our results show a linear correlation between central and individual gating throughout the range of T cell frequencies analyzed (**Figure 1C**). Throughout the remaining study, the values from central manual analysis were used when comparing automated and manual flow cytometry analyses.

Performance of Automated Software

We next evaluated the ability of the three automated gating algorithms FLOCK, SWIFT, and ReFlow to identify MHC multimer-binding T cells. Each algorithm varied with respect to the processing time, additional software requirement, manual handling before or after the automated processes, and annotation requirements. Relevant features of the selected algorithms have been listed in **Table 1**. Specifically, substantial manual handling may impact both the objectivity and handling time—two parameters that we aim to improve through computational analysis. The workflow for each automated analysis tool is depicted in Figure S1 in Supplementary Material.

First, we addressed the limit of detection for the three selected algorithms, through analysis of two independent titration experiments. We used PBMCs from one donor (BC260) carrying 1.7% HLA-B0702 CMV_{TPR}-specific T cells in total live lymphocytes and mixed this in fivefold dilution steps with an HLA-B702 negative donor (BC262). A total of seven serial dilutions were used, giving a theoretical frequency of MHC multimer[+] cells ranging from 1.7 to 0.0001 out of total live, single lymphocytes, and each sample was analyzed by flow cytometry for the presence of HLA-B*0702 CMV_{TPR} multimer-binding CD8[+] T cells (**Figure 2A**). Secondly, a titration curve was generated by mixing a PBMC sample from donor B1054 holding an HLA-A*0201 CMV_{NLV} and an HLA-A*0201 FLU_{GIL} response of 0.87 and 0.13% of total lymphocytes in twofold dilution steps with donor B1060 (HLA-A*0201 negative). A "negative sample" of PBMCs from B1060 alone was also included (Figure S2 in Supplementary Material). The FCS files were analyzed, using manual analysis, FLOCK, SWIFT, and ReFlow software tools. Frequencies of MHC multimer[+] cells were not compared based on CD8[+] cells because there was no consistent CD8 expression cutoff value to use in annotating the data clusters identified by FLOCK. The same cutoff value could not be used across samples coming from different labs most likely due to the large variation in antibodies/fluorochromes used to stain for CD8 cells between individual labs. Hence, to enable comparison of results between all analysis methods, the frequency of MHC multimer-binding T cells was calculated based on live, single lymphocytes.

Our data show that all three algorithms perform equally well in comparison with central manual gating in identifying populations >0.01% of total lymphocytes (**Figure 2B**; Figure

S2 in Supplementary Material). At frequencies <0.01%, FLOCK either assigned too many cells to the MHC multimer population or did not associate any cell population with MHC multimer binding (**Figure 2B**; Figure S2 in Supplementary Material). ReFlow also assigned too many cells to the MHC multimer[+] cluster for the low-frequency populations, resulting in the assignment of approximately 0.002% MHC multimer[+] cells regardless of their true presence, as these were also assigned in the negative or very low-frequency samples (**Figure 2B**; Figure S2 in Supplementary Material). Only the SWIFT algorithm was able to identify cell populations of similar sizes as theoretically present and detected through manual analysis, down to the range of 0.0005–0.0001% of total lymphocytes, where only one to five events were present on the corresponding dot plots (**Figure 2A**). For manual analysis, a threshold of 10 events is usually applied, corresponding to 0.001% of total lymphocytes in these samples (represented by the dashed line in **Figure 2B**). However, for high avidity T cells that are very well separated based on fluorescence intensity, as in this case, the presence of MHC positive T cells can be followed at even lower frequencies.

Automated Analysis of MHC Multimer-Binding T Cells from Proficiency Panel Data

In order to reduce noise from irrelevant cell populations a preselection of live, single cell lymphocytes was performed prior to the automated analysis. We compared manual pregating to an automated prefiltering process using DAG (see footnote text 3), for its impact on the following identification of MHC multimer[+] T cells using either FLOCK or SWIFT. The final assessment of MHC multimer[+] T cells was not affected by the choice of pregating strategy, and the obtained data correlated tightly throughout the range of MHC multimer[+] T cell frequencies analyzed (Figure S3 in Supplementary Material). Since ReFlow includes a separate build-in prefiltering process, the impact of the preselection methods was consequently not compared.

Next, we compared the identification of MHC multimer-binding T cells across the three automated analysis tools to central manual analysis of the proficiency panel data. The number of relevant MHC-binding T cells was assessed for both donors: donor 518, EBV (~0.3%), FLU (~0.02%), and donor 519 EBV (~1.5%), FLU (~0.01%), all values are given as %MHC multimer-binding T cells out of total live, single lymphocytes. The coefficients of determination (R^2) for the three correlations were calculated separately for the high-frequency populations (518 and 519 EBV), for the low-frequency responses (518 and 519 FLU), and for all populations together. Overall, the three algorithms were able to identify most of the MHC multimer-binding T cell populations in a similar range as identified by manual gating (FLOCK: $R^2 = 0.977$, ReFlow: $R^2 = 0.871$, SWIFT: $R^2 = 0.982$) (**Figures 3A–C**). However, a spreading was observed for low-frequent T cell populations, especially using FLOCK and ReFlow (**Figures 3A,B**). For FLOCK, the correlation was tight for the high-frequency populations ($R^2 = 0.965$) but a significant spreading was observed for low-frequency populations ($R^2 = 0.00676$) (**Figure 3A**). There were two different issues

FIGURE 2 | Limit of detection for different automated approaches. A donor carrying ~1.7% CD8[+] T cells binding to HLA-B*0702 cytomegalovirus (TRP) was spiked into an HLA-B0702 negative donor in fivefold dilutions in order to assess the limit of detection of the four analysis approaches. The experiment was run in duplicates. **(A)** Dot plots of the spiked samples showing the theoretical frequency of multimer + cells of the total lymphocyte population and the actual detected frequency (in brackets) by manual gating. Multimer + cells are double positive for PE and APC. PE: phycoerythrin; APC: allophycocyanin. **(B)** The mean percentage of multimer positive cells out of single, live lymphocytes. Numbers represent the seven different samples. Dotted bars: the software detected zero specific cells in one of the two duplicates. #: the software was unable to detect the specific populations in both duplicates. Dashed line: a typical detection threshold for positive response in a major histocompatibility complex multimer staining.

giving rise to this observation: one was that for the low-frequency populations, FLOCK assigned background events into the true MHC multimer[+] T cell population. The other issue was related to the difficulty of annotating the data clusters identified in the FLOCK analysis. As a fully automated unsupervised clustering method, FLOCK assigned the values 1–4 (1: negative, 2: low, 3: positive, 4: high) for categorizing expression levels of each marker based on the relative expression level of the given marker on each identified cell population. In this study, an MHC multimer[+] T cell population was defined as having an expression level >1 for CD3 (not included in all labs), >1 for CD8, and >2 for the MHC multimer. The same cutoff value was used for all samples in order to have a standardized analysis pipeline, requiring a minimum of

manual intervention. The chosen cutoff value was however not suitable for all samples, as there were cases where populations that by visual inspection were defined as clearly MHC multimer[−], were identified by FLOCK as multimer[+] populations based on the cutoff values applied. These populations resulted in a false positive assignment of MHC multimer[+] T cells. This was particularly the case for samples holding low-frequency MHC multimer[+] T cell populations (Figure S3 in Supplementary Material). ReFlow showed a larger spreading throughout the range of T cell frequencies but—like FLOCK—had better performance when detecting high-frequency populations ($R^2 = 0.776$) as opposed to low-frequency populations ($R^2 = 0.138$) (**Figure 3B**). For SWIFT analysis, a tight correlation was observed for both high-frequency

TABLE 1 | Features of the three software solutions.

Feature	SWIFT	FLOCK	ReFlow
Availability	Free but requires Matlab	Free online	Free online
Program run time	~1 h	~10 min	~30 min
Template feature	Yes	No	Yes
Cross-comparison feature	Yes	Yes	Yes
Difficulties in output analysis	New gating method—centroid cluster gating	Choosing cutoff values	Easy
Automatization	+	+++	++
Sensitivity	+++	+	++
Requires common nomenclature of parameters	Yes, renaming of channels is possible	Yes	Yes, harmonized by the tool
Repository	No	No	Yes
Hardware requirement	Runs locally on the computer—analysis speed depends on local computer resources	Web access—analysis speed depends on FLOCK compute resources	Web access—analysis speed depends on ReFlow compute resources
Feasibility for non-computational experts	+	++	+++

Program run times represent the time it takes the software to analyze all files within one lab. For Scalable Weighted Iterative Flow-clustering Technique (SWIFT), it includes the clustering of a consensus sample and subsequent clustering of all samples based on the template.

and low-frequency populations ($R^2 = 0.968$ and 0.722, respectively) (**Figure 3C**).

In order to compare the automated analysis tools to each other, we determined the average frequency of the different MHC multimer-binding T cell populations identified and the CV obtained when using either central manual gating, FLOCK, SWIFT, or ReFlow (**Figures 4A,B**). Again, all evaluated tools could identify high and intermediate frequency T cell populations (518/EBV and 519/EBV) with low variance and significantly differentiate these from the negative control sample (**Figure 4A**). The low-frequency populations (518/FLU and 519/FLU) could, however, not be distinguished from the negative control samples by FLOCK. For ReFlow, a significant difference between the EBV- or FLU-specific T cell holding samples and the negative control sample was obtained; however, the assigned number of MHC multimer-binding cells in the negative samples was higher compared with both central manual analysis and SWIFT analysis (**Figure 4A**). SWIFT analysis enabled identification of the low-frequency MHC multimer-binding T cell populations at equal levels to the central manual gating (**Figure 4A**). In terms of variance, similarly, SWIFT provided comparable variance in the determination of low-frequency MHC multimer-binding T cells (FLU in 518 and 519), compared with central manual gating. In contrast FLOCK, and to a lesser extend ReFlow, resulted in increased variation for the low-frequent responses which was statistically significant only for the 518 FLU response (**Figure 4B**).

We finally assessed if the use of automated analyses could reduce the variation in identification of MHC multimer+ T cell

populations when compared with the individual manual gating conducted by the different labs involved. We chose to look at the smallest population in our study, the donor 519 FLU population as this population had the highest variance. In order to make this assessment, we needed to assign the frequency of the MHC multimer+ population based on the CD8+ T cells. Consequently, this was evaluated exclusively for ReFlow and SWIFT, as the assignment of the correct CD8+ population was challenging on this dataset using the FLOCK algorithm based on the uniform criteria's that were chosen across the full data set and the high inter-lab variations (see Materials and Methods). The variance was assessed by comparing the CV for the frequencies found with individual manual gating, central manual gating, and the two automated analysis tools (**Figure 4C**). This comparison showed that automated gating analysis using SWIFT provided significantly lower variance compared with individual gating, which is the situation applied to most data analyses. ReFlow analysis lowered the variance to the same level as central manual gating, although this was not statistically significant.

DISCUSSION

In this study, we evaluated the feasibility of using automated gating strategies for the detection of antigen-specific T cells using MHC multimers. Among the three algorithms tested, FLOCK, SWIFT, and ReFlow, all proved useful for automated identification of MHC multimer+ T cell populations from the proficiency panel at levels >0.1% which was also reflected in the high degree of correlation of all the tools with central manual analysis. Detection of responses with frequencies in the range of 0.05–0.02% within living lymphocytes was also feasible with SWIFT and ReFlow; however, only SWIFT algorithm was able to detect cell populations <0.02%. The detection limit of ReFlow was lower based on the spike-in experiments (0.002%) and one possible explanation for this discrepancy is the difference in the intensity of the pMHC positive population and the quality of the cell samples. The samples acquired during the spike-in experiment showed a very distinct MHC multimer population and almost no background, whereas the samples acquired for the proficiency panel showed a larger variation in terms of background and fluorescent separation of the MHC multimer population. This finding highlights the importance of sample quality and fluorescent separation when using automated analysis tools. The lower limit of detection of SWIFT is consistent with the results of the FlowCAP II challenge where SWIFT was one of the top performers in the identification of rare cell populations [12]. However, in a more recent study that compared automated analysis tools in a fully automated fashion (i.e., no cluster centroid gating allowed), SWIFT was outperformed by other algorithms that were not tested in this study [13]. In this particular study, all tested algorithms were compared in a fully automated fashion, which is not the way SWIFT was applied in our study. Here, SWIFT clustered output files were further gated manually on cluster centroids. This might explain the discrepancy between these and our results, and also suggests that centroid gating may improve analysis of automated clustering results. An alternative to the manual gating step could be to run the SWIFT clustered output files in another algorithm,

FIGURE 3 | Automated analyses versus central manual gating. Correlation between automated analyses and central manual gating for the identification of MHC multimer positive T cell populations, using either of the three algorithms: **(A)** FLOCK, $n = 112$, $p < 0.0001$, one data point of 0% was converted to fit the log axis (given in red); **(B)** ReFlow, $n = 92$, $p < 0.0001$; **(C)** SWIFT, $n = 108$, $p < 0.0001$. All p-values are Pearson's correlations. Different colors indicate different populations.

which could potentially also improve the automated analysis as was seen in the FlowCAP I challenge where the best results were obtained when the algorithms were combined (12). The dataset analyzed here, holds a large diversity in terms of antibodies

and fluorescent molecules used for the identification of CD8+ T cells. As such this dataset represents a "worst case scenario" for automated gating algorithms. Consequently, it was impossible to normalize staining intensities to a given standard, and cross-sample comparison could only be applied within each lab. This lack of standardization may impact the performance of the different algorithms. However, the ability to work across large differences in assay design is necessary to compare flow cytometry data between various laboratories. Obviously, when multicenter immunomonitoring projects are planned, it is advantageous to harmonize staining protocols and antibody panels across different laboratories, and such harmonization will ease the following automatic analyses and improve the outcome.

In terms of handling the three software tools, a number of relevant differences should be highlighted. FLOCK has a very user-friendly web interface with several different analysis features. The output is graphically very similar to regular dot plots and as such is well recognized by immunologists and easy to interpret by non-computational experts. An additional strength of FLOCK is the possibility to manually adjust the centroids chosen by the algorithm, in cases where they were obviously misplaced. In this study, we did not interfere with the FLOCK analysis as we aimed to obtain a standardized and fully automated approach. The ability to make manual adjustments combined with a clear graphical readout provides a sense of transparency and understanding of the analysis process, making it attractive to immunologists with limited computation expertise. Since the completion of this study, the FLOCK platform has been updated to include even more analysis features, further improving the FLOCK interface. Finally, as stated in **Table 1**, FLOCK analysis is quite fast especially compared with SWIFT. However, prior to FLOCK analysis, FCS files must be uploaded to the web interface, which can be time consuming depending on file size. The SWIFT algorithm runs locally on the computer through Matlab and consequently requires a minimal level of coding abilities. All codes are well described in the manual associated with the SWIFT installation files and simple to use. SWIFT does not require data-upload to a distant server, but may require substantial run times, depending on the local computer power. However, the slower initial clustering of a consensus file is partly compensated by the rapid assignment of individual samples to the initial cluster template. Similar to FLOCK, the SWIFT algorithm allows adjustment of parameters important for the analysis output, like input cluster number, ArcSinh Factors, and Percent Noise. These features are, however, not very intuitive for non-computational experts to understand and hence challenging to adjust in a meaningful manner. The output files generated by SWIFT, when analyzed in, e.g., FlowJo, can be displayed as either conventional dot plots, or as somewhat different dot plots in which each dot represents a full cluster rather than a single cell. This feature provides some flexibility, allowing an operator more freedom to position gates and still catch the target population across samples, even in the presence of machine noise or slight fluorescence shifts. Thus, SWIFT provides a clustering of events, but the final binning of various clusters into certain parameter-defined categories is done through manual cluster gating (in the present study) or can be accomplished by a second automated platform (17). ReFlow also

FIGURE 4 | Comparison of the different analysis methods. **(A)** Percentage of MHC multimer+ T cells out of single, live lymphocytes found using the different analysis approaches for identification of T cells recognizing two different virus-derived epitopes (EBV, FLU) in two different donors (518, 519). Error bars indicate SD. ****: $p < 0.0001$; ns: not significant (paired t-test). Central manual: $n = 28$, FLOCK: $n = 28$, ReFlow: $n = 23$, SWIFT: $n = 27$. **(B)** The coefficient of variation (CV) (SD/ mean*100) for the different analysis approaches in determining the frequency of MHC multimer+ T cells. ****: $p < 0.0001$; no line: no significant difference (asymptotic CV equality test). **(C)** The CV (SD/mean*100) specifically related to the FLU-specific response in donor 519. **: $p < 0.01$; no line: no significant difference (asymptotic CV equality test). For **(C)**, the CV is calculated based on percentage of MHC multimer+ T cells out of total CD8 T cells in order to compare with individual manual gating. 518: healthy donor 518; 519: healthy donor 519; EBV: Epstein–Barr virus; FLU: influenza virus.

has a simple and intuitive user interface that is accessible via a standard web-browser. It requires no programming knowledge to learn and operate. The FCS files have to be uploaded on to the server at speeds determined by the local internet connection. FCS files that belong together are analyzed as a group and since this is performed on shared GPUs, it is not affected by the local computational hardware. Results can be visualized graphically as 2D dot plots (showing both clusters as well as events within clusters) and in tabular format that can be further exported into a csv file. From the graphical view, clusters of interest may manually be further selected, named, and evaluated or may be selected for a further second stage analysis, as it was performed for the

current study. Live, lymphocytes were chosen for a further round of clustering to determine multimer positive clusters that are then chosen based on visual inspection of the clusters. The manual selection of clusters in ReFlow is somewhat easier than cluster gating on SWIFT output data, as it is an incorporated part of the algorithm and can be done directly from the analysis.

None of the three automated gating algorithms tested in this study provide a fully automated pipeline. Whether it is choosing cutoff values in FLOCK, cluster gating in SWIFT or choosing positive populations by visual inspection in ReFlow, the analysis of the clustering output requires some manual decision making. That being said, the manual cluster gating performed on the SWIFT

files was more laborious than what was needed for the other algorithms. In this study, the FLOCK pipeline was the most automated process as the same cutoff values were applied to all samples. In fact, it might very well have improved the FLOCK analysis if the cutoff level had been defined for each individual sample—which would have been similar to the process for SWIFT and ReFlow. With such sample-specific adjustments, at least one of the issues depicted in Figure S4 in Supplementary Material would have been eliminated. Hence, the FLOCK algorithm provides an analysis platform with higher degree of automatization, but this comes at the expense of sensitivity at least for this very diverse dataset.

A few things are worth considering if a more automated approach is desired, such as harmonization of the staining reagents and procedure, data collection, and FCS file management. In this study, we believe it would have improved the results from the FLOCK analysis had the same antibody been used for the given markers across different labs. This would have eliminated some of the discussed issues with setting an appropriate cutoff level as the fluorescence intensities could have been normalized and would also have allowed the cross-comparison feature to be applied to all samples at once instead of as current within each lab individually. Also, the procedure for SWIFT analysis could potentially have been improved by this, as all labs could have been analyzed using the same template file. Additionally, sample quality is an important issue. Just as it is difficult to manually gate samples with a lot of background due to poor cell sample quality or preparation, it makes the automated detection of specific populations equally, if not even more difficult, as the subjective distinction between background and true events based on visual inspection is removed from the analysis process. Furthermore, common parameter nomenclature between FCS files would lead to less manual intervention, eliminating the step of manual adjustment of parameter names, which is an option within most automated tools. The field of computational analysis of flow cytometry data is rapidly developing, leading to increasingly sophisticated tools that can more accurately detect the exact cell populations of interest. This development is an ongoing process dependent on feedback from actual users and exchange between the fields of software development and immunology.

In this study, we particularly aimed to evaluate automated flow cytometry analysis tools that can be used by experienced flow cytometry users with no programming skills. For all three tested algorithms, there were challenges throughout the study, and it is a problem that non-computational experts have limited possibilities to trouble-shoot data analysis in the computational space. This highlights the need for a closer interaction between the two

fields of immunology and bioinformatics/programming and also the need for immunologists to educate themselves within the field of bioinformatics in order to keep up with the development of increasingly complex data analysis in the future (10).

The data presented here shows the feasibility and potential advantage of using automated gating strategies, even across very diverse datasets. The algorithms included here, represent three user-friendly tools for such assessment, but it is by no means an exclusive list. Many computational tools for flow cytometry analyses are currently present, each having their own pros and cons and the choice of algorithm depends on the characteristics of the individual experiments and the desired outcome. Thus, it is crucial to choose carefully when deciding which algorithms to use for each purpose (10, 22).

AUTHOR CONTRIBUTIONS

NWP performed gating analysis, made figures, analyzed data, and wrote the manuscript; AC performed gating analysis, discussed data, and revised the manuscript; YQ, JR, AL, and RS performed gating analysis and revised the manuscript; KJ and CH provided data files and revised the manuscript; NVP and MH performed gating analysis; RHS, TM, CC, and SW provided technical guidance and revised the manuscript; CG analyzed and discussed data and revised the manuscript; SH conceived the concept, analyzed and discussed data, and wrote the manuscript.

ACKNOWLEDGMENTS

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://journal.frontiersin.org/article/10.3389/fimmu.2017.00858/full#supplementary-material.

REFERENCES

1. Altman JD, Moss PA, Goulder PJ, Barouch DH, McHeyzer-Williams MG, Bell JI, et al. Phenotypic analysis of antigen-specific T lymphocytes. *Science* (1996) 274:94–6. doi:10.1126/science.274.5284.94

2. Hadrup SR, Bakker AH, Shu CJ, Andersen RS, van Veluw J, Hombrink P, et al. Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat Methods* (2009) 6:520–6. doi:10.1038/nmeth.1345

3. 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. doi:10.1111/cei.12339

4. Andersen RS, Thrue CA, Junker N, Lyngaa R, Donia M, Ellebæk E, et al. Dissection of T-cell antigen specificity in human melanoma. *Cancer Res* (2012) 72:1642–50. doi:10.1158/0008-5472.CAN-11-2614

5. van der Burg SH, Kalos M, Gouttefangeas C, Janetzki S, Ottensmeier C, Welters MJ, et al. Harmonization of immune biomarker assays for

clinical studies. *Sci Transl Med* (2011) 3:108s44. doi:10.1126/scitranslmed. 3002785

6. Britten CM, Gouttefangeas C, Welters MJ, Pawelec G, Koch S, Ottensmeier C, et al. The CIMT-monitoring panel: a two-step approach to harmonize the enumeration of antigen-specific CD8+ T lymphocytes by structural and functional assays. *Cancer Immunol Immunother* (2008) 57:289–302. doi:10.1007/s00262-007-0378-0

7. Hadrup SR, Maurer D, Laske K, Frøsig TM, Andersen SR, Britten CM, et al. Cryopreservation of MHC multimers: recommendations for quality assurance in detection of antigen specific T cells. *Cytometry A* (2015) 87:37–48. doi:10.1002/cyto.a.22575

8. Welters MJ, Gouttefangeas C, Ramwadhdoebe TH, Letsch A, Ottensmeier CH, Britten CM, et al. Harmonization of the intracellular cytokine staining assay. *Cancer Immunol Immunother* (2012) 61:967–78. doi:10.1007/s00262-012-1282-9

9. Gouttefangeas C, Chan C, Attig S, Køllgaard TT, Rammensee HG, Stevanović S, et al. Data analysis as a source of variability of the HLA—peptide multimer assay: from manual gating to automated recognition of cell clusters. *Cancer Immunol Immunother* (2015) 64(5):585–98. doi:10.1007/s00262-014-1649-1

10. Kvistborg P, Gouttefangeas C, Aghaeepour N, Cazaly A, Chattopadhyay PK, Chan C, et al. Thinking outside the gate: single-cell assessments in multiple dimensions. *Immunity* (2015) 42:591–2. doi:10.1016/j.immuni.2015.04.006

11. Finak G, Langweiler M, Jaimes M, Malek M, Taghiyar J, Korin Y, et al. Standardizing flow cytometry immunophenotyping analysis from the human ImmunoPhenotyping Consortium. *Sci Rep* (2016) 6:20686. doi:10.1038/srep20686

12. Aghaeepour N, Finak G, FlowCAP Consortium, DREAM Consortium, Hoos H, Mosmann TR, et al. Critical assessment of automated flow cytometry data analysis techniques. *Nat Methods* (2013) 10:228–38. doi:10.1038/nmeth0513-445c

13. Weber LM, Robinson MD. Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. *Cytometry A* (2016) 89(12):1084–96. doi:10.1002/cyto.a.23030

14. Qian Y, Wei C, Eun-Hyung Lee F, Campbell J, Halliley J, Lee JA, et al. Elucidation of seventeen human peripheral blood B-cell subsets and quantification of the tetanus response using a density-based method for the automated identification of cell populations in multidimensional flow cytometry data. *Cytometry B Clin Cytom* (2010) 78(Suppl 1):S69–82. doi:10.1002/cyto.b.20554

15. Mosmann TR, Naim I, Rebhahn J, Datta S, Cavenaugh JS, Weaver JM, et al. SWIFT-scalable clustering for automated identification of rare cell populations in large, high-dimensional flow cytometry datasets, part 2: biological evaluation. *Cytometry A* (2014) 85:422–33. doi:10.1002/cyto.a.22445

16. Rebhahn JA, Roumanes DR, Qi Y, Khan A, Thakar J, Rosenberg A, et al. Competitive SWIFT cluster templates enhance detection of aging changes. *Cytometry A* (2016) 89:59–70. doi:10.1002/cyto.a.22740

17. Naim I, Datta S, Rebhahn J, Cavenaugh JS, Mosmann TR, Sharma G. SWIFT-scalable clustering for automated identification of rare cell populations in large, high-dimensional flow cytometry datasets, part 1: algorithm design. *Cytometry A* (2014) 85:408–21. doi:10.1002/cyto.a.22446

18. White S, Laske K, Welters MJP, Bidmon N, van der Burg SH, Britten CM, et al. Managing multi-center flow cytometry data for immune monitoring. *Cancer Inform* (2014) 13:111–22. doi:10.4137/CIN.S16346

19. Cron A, Gouttefangeas C, Frelinger J, Lin L, Singh SK, Britten CM, et al. Hierarchical modeling for rare event detection and cell subset alignment across flow cytometry samples. *PLoS Comput Biol* (2013) 9:e1003130. doi:10.1371/journal.pcbi.1003130

20. Toebes M, Coccoris M, Bins A, Rodenko B, Gomez R, Nieuwkoop NJ, et al. Design and use of conditional MHC class I ligands. *Nat Med* (2006) 12:246–51. doi:10.1038/nm1360

21. Burel JG, Qian Y, Lindestam Arlehamn C, Weiskopf D, Zapardiel-Gonzalo J, Taplitz R, et al. An integrated workflow to assess technical and biological variability of cell population frequencies in human peripheral blood by flow cytometry. *J Immunol* (2017) 198:1748–58. doi:10.4049/jimmunol.1601750

22. Aghaeepour N, Chattopadhyay P, Chikina M, Dhaene T, Van Gassen S, Kursa M, et al. A benchmark for evaluation of algorithms for identification of cellular correlates of clinical outcomes. *Cytometry A* (2016) 89:16–21. doi:10.1002/cyto.a.22732

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary figures

Fig. S1. Workflow of the three automated analysis tools SWIFT, FLOCK and ReFlow.
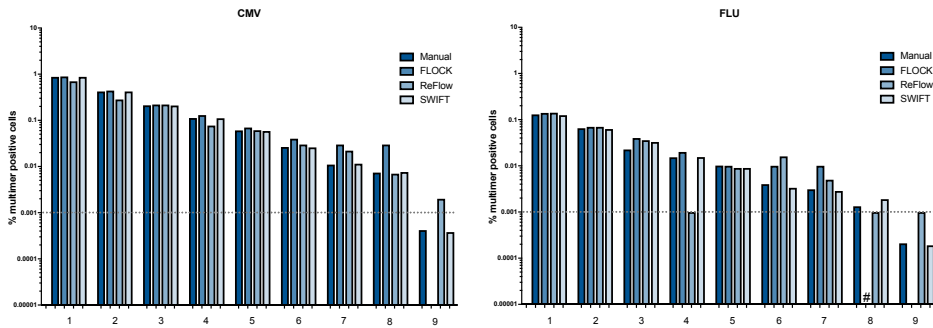
Fig. S2. Detection limit of Spike-in 2
A donor having T-cell responses against HLA-A0201 CMV(NLV) and HLA-A0201 FLU(GIL) was spiked into a negative donor in 2 fold dilutions. The last sample contains only cells from the negative donor. The percentage of MHC multimer$^+$ T cells of single, live lymphocytes was calculated with the 4 different analysis approaches. #: the software was unable to detect the positive population. CMV: Cytomegalovirus. FLU: Influenza virus.
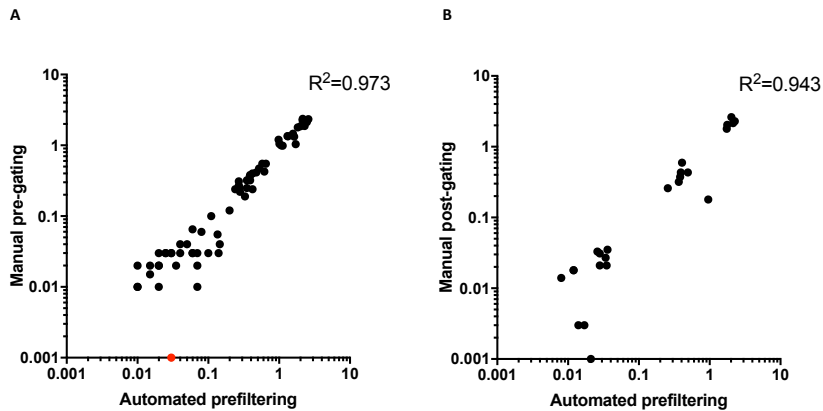
Fig. S3. Automated approach with manual gating versus automated prefiltering
A) Correlation between FLOCK analysis with a manual pre-processing step and FLOCK analysis with automated prefiltering. n=64. p<0,0001
(pearson correlation) Red dot: response of 0% that is converted to fit the log axis **B**) Correlation between SWIFT analysis with a manual post-processing step and SWIFT analysis with automated prefiltering. n=24. p<0,0001 (pearson correlation)
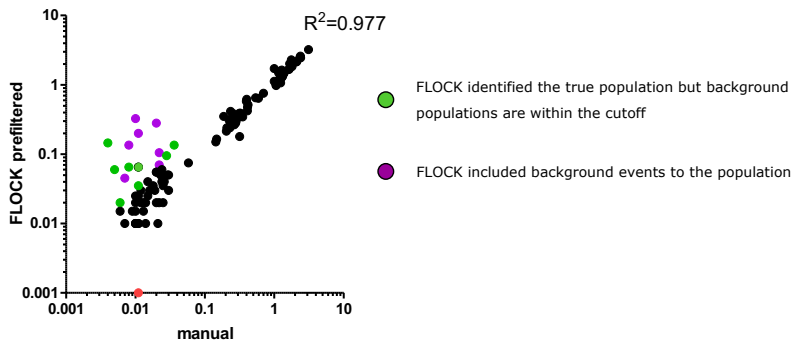
Fig. S4. FLOCK issues.
Illustration of the two different issues both contributing to the overestimation of population size with FLOCK. Correlation plot is modified from fig. 2A. Green dots: FLOCK was able to separate the true multimer population, but the cutoff value was not suitable for this sample as background populations were also included in the population. Purple dots: FLOCK was not able to separate the true multimer population from background events and thereby overestimated the size of the population.

Manuscript III

Broad presence of autoreactive CD8 T cells against narcolepsy type 1-relevant proteins

Broad presence of autoreactive CD8 T cells against narcolepsy type 1-relevant proteins

Natasja Wulf Pedersen[1], Anja Holm [2], Nikolaj Pagh Kristensen[1], Anne-Mette Bjerregaard[1], Amalie Kai Bentzen[1], Tripti Tamhane[1], Kristoffer Sølvsten Burgdorf[3], Henrik Ullum[3], Poul Jennum[4], Stine Knudsen[5], Sine Reker Hadrup[1][□], Birgitte R. Kornum[6]*[□]*

1: Department of Micro- and Nanotechnology, Technical University of Denmark, Copenhagen, Denmark.
2: Molecular Sleep Laboratory, Department of Clinical Biochemistry, Rigshospitalet, Glostrup, Denmark
3: Department of Clinical Immunology 2034, Copenhagen University Hospital, Rigshospitalet, Denmark
4: Danish Center for Sleep Medicine, Department of Neurophysiology, Rigshospitalet, Glostrup, Denmark
5: Norwegian Centre of Expertise for Neurodevelopmental Disorders and Hypersomnias (Nevsom), Oslo University Hospital, Ullevål, Norway
6: Department of Neuroscience, University of Copenhagen, Denmark.
□These authors contributed equally to this work.
*Corresponding author: Birgitte R. Kornum, kornum@sund.ku.dk and Sine Reker Hadrup, sirha@dtu.dk

Abstract

Type 1 Narcolepsy (NT1) is a severe chronic neurological sleep disorder that is characterized by the loss of hypocretin/orexin signaling in the brain. Genetic and epidemiological data support the hypothesis that NT1 is a T cell-mediated autoimmune disease that targets the hypocretin-producing neurons, but autoreactive T cells remain to be detected. Narcolepsy is tightly associated with the MHC class II allele HLA-DQB1*06:02 and, to a lesser extent, MHC class I alleles, suggesting a role for both CD4$^+$ and CD8$^+$ T cells in disease pathogenesis. We tested CD8+ T cell autoreactivity towards 1183 peptides from seven proteins associated with hypocretin neurons. Peptide-MHC-I multimers labeled with DNA barcodes were used to identify CD8$^+$ T cell reactivity in samples from 20 NT1 patients and 52 healthy controls. We observed autoreactive CD8$^+$ T cells targeting narcolepsy-relevant peptides presented primarily by HLA-B18:01, HLA-B51:01, and HLA-C04:01 in both patients and healthy controls. The frequency of autoreactive T cell clones was lower in HLA-DQB1*06:02+ controls compared to both NT1 patients and HLA-DQB1*06:02- controls, suggesting a protective role for the absence of autoreactive T cells in healthy HLA-DQB1*06:02+ individuals.

Introduction

Type 1 Narcolepsy (NT1) is a chronic, disabling neurological sleep disorder characterized by the dysregulation of the sleep-wake cycle, leading to early occurring rapid eye movement (REM) sleep, excessive daytime sleepiness and disrupted sleep during the night. Another common characteristic of NT1 is muscle tonus dysregulation during sleep and wakefulness, resulting in sudden loss of muscle tone (cataplexy). Furthermore, sleep paralysis, hallucinations and REM sleep behavior disorder/REM sleep without atonia are often seen[1–3]. NT1 is caused by disrupted signaling through the sleep-regulating neuropeptide hypocretin in the brain[4] and it has been shown that this is due to the loss of specific neurons in the hypothalamus that produce hypocretin[5, 6].

An autoimmune basis for NT1 has long been suspected based on a strong association with the common HLA-DQ haplotype, DQA1*01:02/DQB1*06:02, which encodes the MHC class II DQ0602 heterodimer[7, 8]. This HLA association is one of the highest known: up to 98% of NT1 patients with demonstrated hypocretin deficiency carry DQ0602 versus 25% of the healthy population[7, 9].

Associations between several MHC class I molecules and narcolepsy has also been suggested by two independent studies[10, 11]. HLA-A*11:01, HLA-B*51:01, and HLA-C*04:01 were found in both studies, whereas HLA-B*35:01 and HLA-B*35:03 were found in the study by Tafti *et al.* and Ollila *et al.*, respectively; the discrepancy between the two subtypes is likely due to ethnicity differences in the two cohorts. Ollila et al. further reported that HLA-B*18:01 is associated with narcolepsy, whereas HLA-B*07:02 had a weak protective effect[11].

Following the 2009/2010 H1N1 influenza vaccination campaigns with Pandemrix, as well as after the H1N1 epidemic itself, narcolepsy incidence dramatically increased in several countries[12–14], further substantiating the role of the immune system in NT1 disease development.

Remarkably, even after the discovery of hypocretin-producing neurons as the putative autoimmune target, attempts to demonstrate narcolepsy-associated autoimmune responses have been unsuccessful (reviewed in [15]). Since neurons express only MHC class I and not class II molecules, cytotoxic CD8+ T cells are the most likely effector cells in the autoimmune destruction of hypocretin neurons[16]. This is supported by the finding of post-mortem hypothalamic CD8+ T cell infiltration in a case of NT1 secondary to anti-Ma–associated diencephalitis[17]. The CD8+ T cell infiltration was associated with a complete loss of hypocretinergic neurons. Importantly, it has also been demonstrated in a mouse model that cytotoxic CD8+ T cells with reactivity towards hemagglutinin can specifically kill hypocretin neurons if these transgenically express hemagglutinin. This was not the case for CD4+ T cells targeting hemagglutinin. Even though these cells infiltrated the brain and caused local inflammation, this did not lead to loss of hemagglutinin expressing hypocretin neurons[18]. Thus, even though autoreactive CD4+ T cells might initiate the disease process, the presence of autoreactive CD8+ T cells could be necessary for the development of genuine NT1.

We use a recently developed technique for detection of antigen specific CD8+ T cells that is especially valuable in identifying CD8+ T cells populations of low frequency and affinity, as it does not rely on fluorescence signal like conventional tetramer methods[19]. Instead, it utilizes DNA barcode-labeled peptide-MHC (pMHC) complexes to identify CD8+ T cells specific for the peptide presented. Using this method, we were able to screen for recognition of 1183 peptides expressed in hypocretin neurons.

Results

Experimental strategy and peptide selection

To test the hypothesis that NT1 is mediated by CD8[+] T cells, we screened blood samples from 20 NT1 patients and 52 healthy controls for the presence of autoreactive CD8[+] T cells that targeted epitopes present in hypocretin neurons. NT1 was diagnosed according to the International Classification of Sleep Disorders, Edition 3 (ICSD-3)[20]. The core demographical, clinical, and sleep investigation parameters for the narcolepsy cohort are shown in Table 1and an overview of the experimental strategy is given in Figure 1a.

Table 1. Demographic, Clinical and Paraclinical Data of the Narcolepsy Cohort

	Narcolepsy type 1 N=20
Demography	
Gender (male), n (%)	9 (45%)
Age (years), mean (range)	25.6 (7-62)
Disease duration (years), mean (range)	5.3 (1-9)
HLA-DQB1*0602- positivity, n (%)	19 (95%)
CSF hcrt-1 ≤110 pg/ml, n (%)	20 (100%)
Daytime sleepiness, n (%)	20 (100%) N=20
Epworth Sleepiness Scale, mean±SEM	16.8 ± 0.5 N=16
Cataplexy presence, n (%)	19 (95%) N=20
Hypnagogic hallucinations, n (%)	15 (82%) N=19
Sleep paralysis, n (%)	11 (62%) N=18
Disrupted night sleep, n (%)	17 (87%) N=20
MSLT (multiple sleep latency test)	
Sleep latency (minutes) mean±SD	3.28±1.98 N=20
SOREMPs (number) mean±SD	3.95±0.91 N=20

SOREMPs (sleep onset REM periods),
CSF hcrt-1 (hypocretin-1 levels in the cerebrospinal fluid).
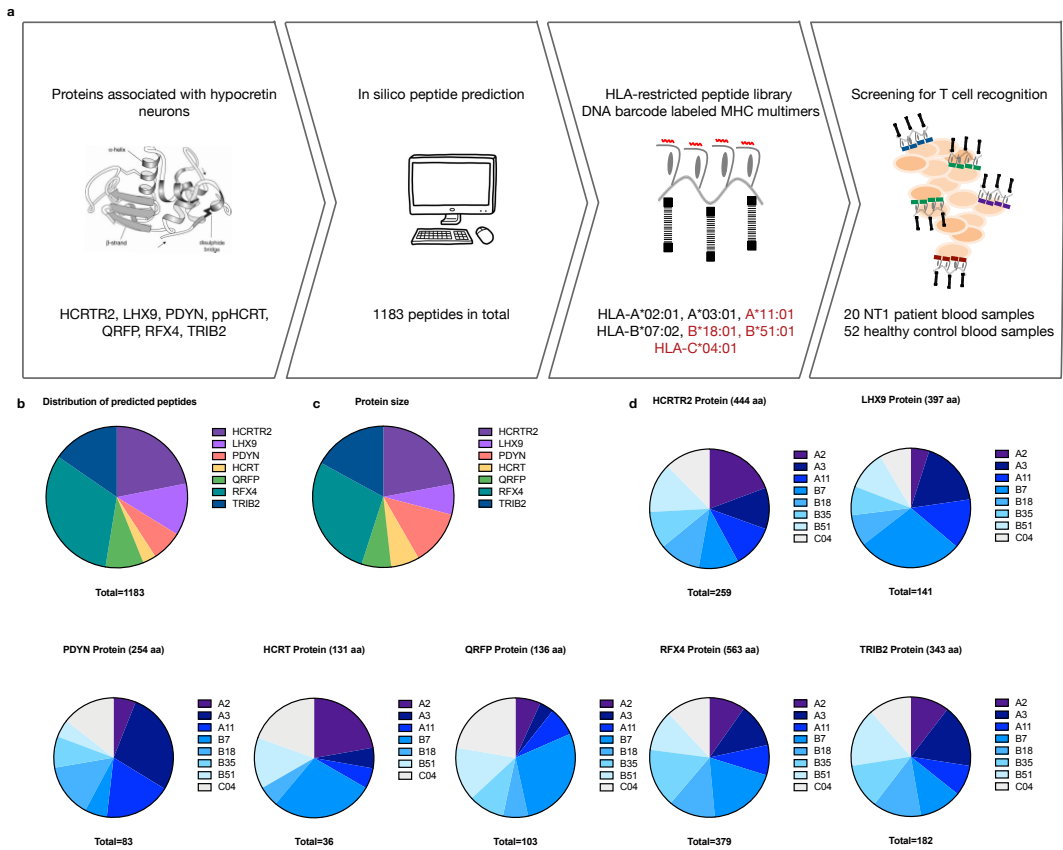
Figure 1. Experimental strategy and peptide prediction
a) From 7 different proteins, peptides with strong binding affinity to 8 different HLA types were predicted using the prediction server NetMHCcons2.8. HLA types in red are known to be associated with Narcolepsy type 1. The predicted peptides from step 2, were used to generate pMHC multimers labeled with DNA barcodes, that were in turn used to screen Narcolepsy type 1 patients and healthy controls. **b**) The distribution of total predicted peptides from each protein. **c**) Protein size. **d**) The distribution of the predicted peptides among the 8 chosen HLA types within each protein. aa: amino acid. HCRTR2: hypocretin receptor 2. LHX9: Lim homeobox 9. PDYN: prodynorphin. HCRT: hypocretin precursor protein. QRFP: pyroglutamylated RFamide peptide. RFX4: Regulatory factor x4. TRIB2: Tribbles homologue 2.

Since nearby neurons in the hypothalamus are not affected in NT1, the epitopes should be from proteins unique to hypocretin neurons, at least locally but preferably globally. Several studies have aimed to characterize the signature of hypocretin neurons[21–26]. By combining data from these studies and also taking into account the peripheral tissue expression of each target[27], we selected the following seven proteins for our screening: the precursor protein for the hypocretin peptides (HCRT), tribbles homologue 2 (TRIB2), regulatory factor x4 (RFX4), prodynorphin (PDYN), Lim homeobox 9 (LHX9), the precursor protein for the pyroglutamylated RFamide peptides (QRFP), and the hypocretin receptor 2 (HCRTR2) (details listed in Table 2).

Table 2. Overview of the protein targets included in the study and potential targets of interest

Protein	Full name	Size (amino acid)	Type	Relation to hypocretin neurons/narcolepsy	Reported in ref	Allen brain atlas	Peripheral expression*
Included in the study							
HCRT	Hypocretin neuropeptide precursor	131	Peptide precursor	Uniquely produced in the neurons	22,23	Specific	Very limited
PDYN	Prodynorphin	254	Peptide precursor	Missing post mortem in hypothalamus from NT1.	22,23	Hypothalamus + other brain areas	Very limited
QRFP	Pyroglutamylated RFamide peptide	136	Peptide precursor	Peptide related to sleep regulation	24	Only in hypothalamus	Low, gut-intestinal tract
LHX9	Lim homeobox 9	397	Transcription factor	Transcription factor regulating hcrt expression.	22,25	All brain	Very limited
RFX4	Regulatory factor x4	735	Transcription factor	Unknown but highly specific to hcrt neurons	22	Only in hypothalamus	Very limited
TRIB2	Tribbles pseudokinase 2	343	Cytosolic protein, pseudokinase	Autoantibodies detected in NT1	21	All brain	Yes
HCRTR2	Hypocretin receptor 2	444	Receptor, GPCR	Autoantibodies detected in NT1. HCRTR2 is not expressed in hcrt neurons	26	No signal	Low, kidney, heart
Other possible targets not included							
NPTX2	Neuronal pentraxin-2	431	Synaptic protein	Missing in hypothalamus from NT1. Also known as NARP	22,23	Hypothalamus + other brain areas	Medium, adrenal gland
IGFBP3	Insulin-like growth factor-binding protein 3	297	Cytosolic protein	Regulates hypocretin expression	22,23	Hypothalamus + other brain areas	High, many organs
PLAGL1	PLAG1 like zinc finger 1	463	Zinc finger protein	Unknown	21,22	Hypothalamus + other brain areas	Medium, many organs
NR6A1	Nuclear receptor subfamily 6 group A member 1	480	Nuclear receptor	Regulates hypocretin expression	21,23	Hypothalamus + other brain areas	Low, many organs

*Data from [27]

Using the NetMHCcons 1.1 in silico prediction algorithm for peptide-MHC class I binding (www.cbs.dtu.dk/services)[28], we predicted 9- to 11-mer peptides from these seven selected proteins with the ability to bind to eight different HLA-I molecules. The HLA-I molecules were chosen based on reported associations with narcolepsy in two independent studies[10, 11], i.e., HLA-A*11:01, HLA-B*18:01, B*35:01, B*51:01, and HLA-C*04:01. Furthermore, HLA-A*02:01, -A*03:01, and HLA-B*07:02 were also included due to their high prevalence in the available samples. Peptides with %Rank scores of two or below were defined as binders and included in the study, giving a total of 1183 peptides across the seven different proteins and eight different HLAs.

The distribution of peptides was not equal between the different proteins (Fig. 1b). The number of predicted peptides ranged from 36 in HCRT to 379 in RFX4. This is mostly an effect of the size differences between the proteins, ranging from 131 amino acids (aa) to 563 aa (Fig. 1c). Within each different protein, the number of predicted peptides with high affinity for each HLA also varied (Fig. 1d). For some proteins (RFX4, HCRTR2, and TRIB2), the distribution of peptides was fairly even

between the eight different HLA types included, whereas for others (LHX9, PDYN, HCRT, and QRFP) some HLA alleles were over- or underrepresented.

DNA barcode-labeled MHC multimers for screening >1000 potential auto-epitopes

The 1183 predicted peptides, were synthesized and used to generate individual peptide-MHC (pMHC) monomers using UV-mediated peptide exchange as previously described[29]. We then multimerized the pMHC monomers onto a phycoerythrin (PE)-labeled polysaccharide backbone coupled to a DNA barcode that was unique to each specific pMHC (as described in [19]). This yielded 1183 different pMHC multimers that were mixed according to donor HLA type and used to stain the relevant NT1 samples and healthy controls. CD8[+] T cells with the ability to bind to pMHC multimers were sorted based on a positive PE signal, the associated DNA barcodes were amplified and the specificity of the CD8[+] T cells could next be revealed by sequencing of the DNA barcodes. This strategy is depicted in Figure 2a along with an example of a sorting plot from a healthy donor (#138) and the subsequent analysis of the responses present in this sample (Fig. 2b).

The 1183 peptides included in this study are ligands to eight different HLA molecules. The eight HLA molecules cover the HLA haplotypes of the sample cohort for at least one and up to four HLA molecules per donor. All donors were screened only with the fraction of the library that matched their HLA type. For each protein and HLA type, Figure 2c shows whether a given sample was screened (grey) or not (white) and whether a CD8[+] T-cell response was detected to one or more of the peptides within the given protein (green/blue color gradient). We estimated the frequency of all multimer positive CD8[+] T cell populations (as previously done in [19]) as the fraction of sequencing reads specific for a given pMHC out of total reads, multiplied by the percentage of sorted multimer-binding T cells out of total CD8[+] T cells. The color gradient in figure 2c indicates the estimated frequency of the given multimer-positive population, if only one, or the total frequency if more than one T cell response was present in a sample. A full overview of all the peptides screened in each sample is given in Figure S1. As can be seen from Figure 2c, we observed CD8[+] T-cell responses to NT1-relevant peptides in both NT1 patients and healthy controls. All but one patient expressed NT1-associated HLA class II allele DQB1*06:02, and for this reason the healthy controls were divided into two groups depending on their expression of this HLA molecule. There were no visible differences in the recognition patterns of CD8[+] T cells between NT1 patients and healthy controls (Fig. 2c).

a

b

c
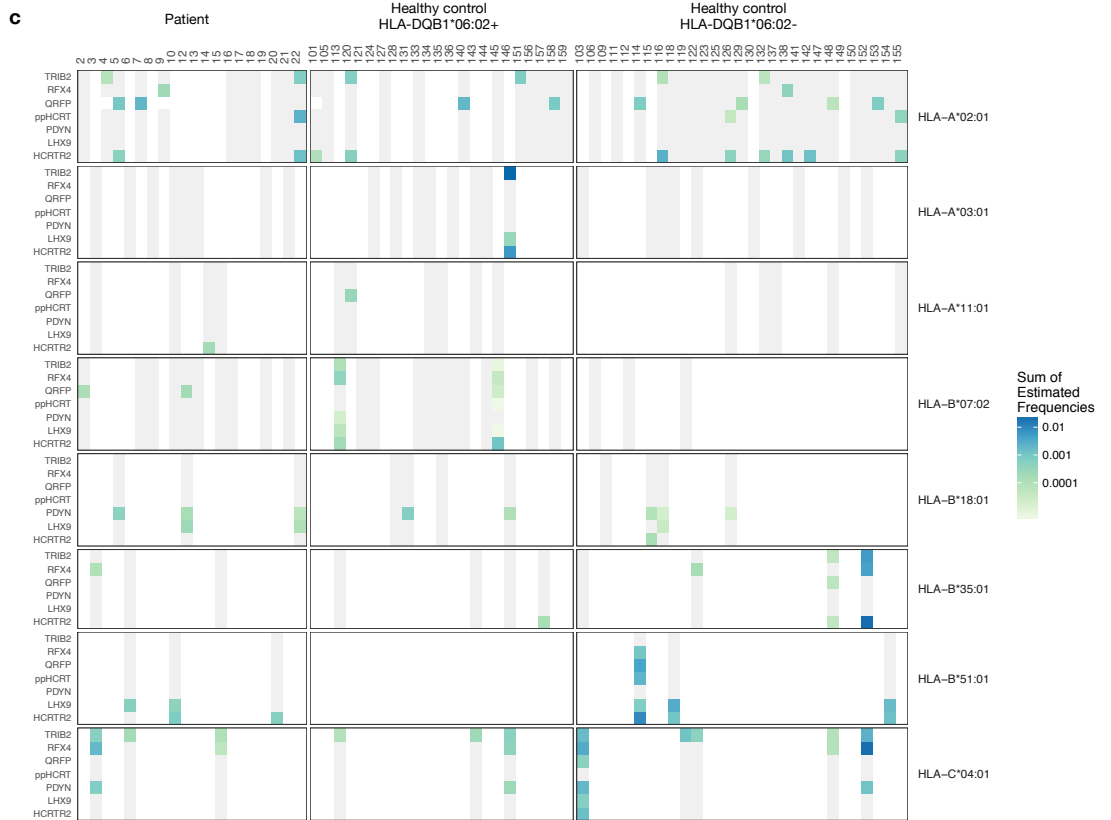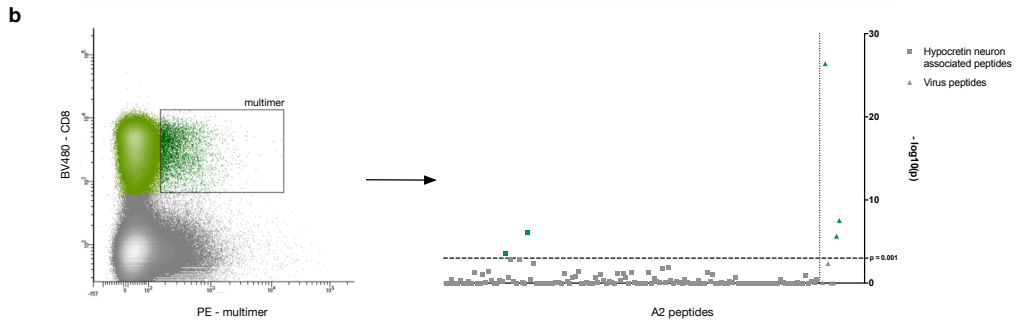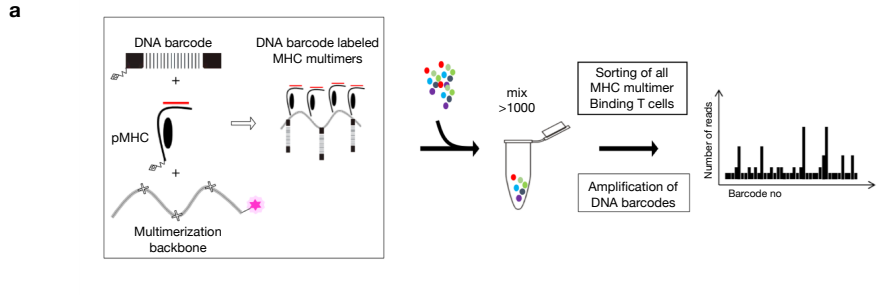
Figure 2. Overview of the experiment pipeline and T cell screening.

a) Overview of the barcode strategy (modified from Bentzen et al. 2016). pMHC monomers are coupled to a polysaccharide backbone labeled with PE. Each individual pMHC is given a unique DNA barcode in a total of 1183 different DNA barcode-labeled MHC multimers. These are mixed and used for staining of PBMCs. The common PE label is used for sorting out all specific CD8[+] T cells. DNA barcodes from the sorted cells are amplified, sequenced and the number of reads for each DNA barcode present is analyzed. **b**) Sorting plot from sample 138 showing the sorted CD8 positive PE population and the responses found in the given donor. Responses are defined based on the $\log_2$fold change of the number of reads compared to triplicate baseline samples with $p<0.001$, which equals a false discovery rate (FDR) <0.1. The axis is transformed to $-\log_{10}(p)$ for visualization. $-\log_{10}(0.001)=3$.

c) Overview of proteins and HLA types included in the screen. All donors are presented in individual columns divided in the three groups: patient and HLA-DQB1:06.02 positive and negative healthy donors. Rows show the seven proteins for each of the eight different HLA types. A green/blue box indicates that a response against one or more peptides from the protein in the specific HLA context was detected. The estimated frequency represents all responses in each donor. Grey box – the donor was screened for recognition of the given protein in the specific HLA context, but no responses were detected. White box – the donor was not screened with peptides from the given protein in the given HLA type.

pMHC: Peptide-MHC complex. PE: phycoerythrin. BV480: Brilliant Violet 480.

The frequency of individual epitope-specific CD8+ T cell populations is lower in HLA-DQB1*06:02 positive controls compared to NT1 patients.

When counting the number of T cell responses in the different cohorts, NT1 patients, HLA-DQB1*06:02-positive controls, and HLA-DQB1*06:02-negative controls, no significant difference was detected based on the average number of responses (Fig. 3a). Interestingly however, we observed a tendency for more patients (13/20, 65%) to hold a T cell response to NT1 related peptides compared to the HLA-DQB1*06:02-positive control groups (11/23, 48%).

Furthermore, when evaluating the estimated frequency for each multimer-positive CD8[+] T cell population, we found that the responses detected in the HLA-DQB1*06:02-positive controls had a significantly lower frequency than those found in both NT1 patients and DQB1*06:02-negative controls. Responses in the latter two groups had estimated frequencies of similar levels (Fig. 3b). However, when summing-up the frequencies for all different multimer-positive CD8[+] T cell populations present in a given donor, there was no difference between the groups (Fig. 3c).

In addition to the 1183 peptides potentially associated with NT1, all the samples were screened with a panel of known virus epitopes as positive controls. Similarly to the NT1 relevant peptides, these were also selected to match the donor's HLA type. For two healthy donors (#109 and #131), there was no overlap with the viral epitope-HLA complexes available and the HLA haplotype of the donors, and these samples were consequently not analyzed for T cell recognition of viral antigens. We analyzed the number and estimated frequency of the virus-specific T cell responses to test if the differences observed between the donor cohorts for NT1-relevant T cell recognition also applied to virus-specific T cell responses. For the virus-specific T cell responses no differences were observed, neither for the number of responses (Fig. 3d) nor for the estimated frequencies (Fig. 3e and f). For the summed frequencies, there was a tendency towards a higher frequency of virus responses in the NT1 patient group, but this was not statistically significant (Fig. 3f).
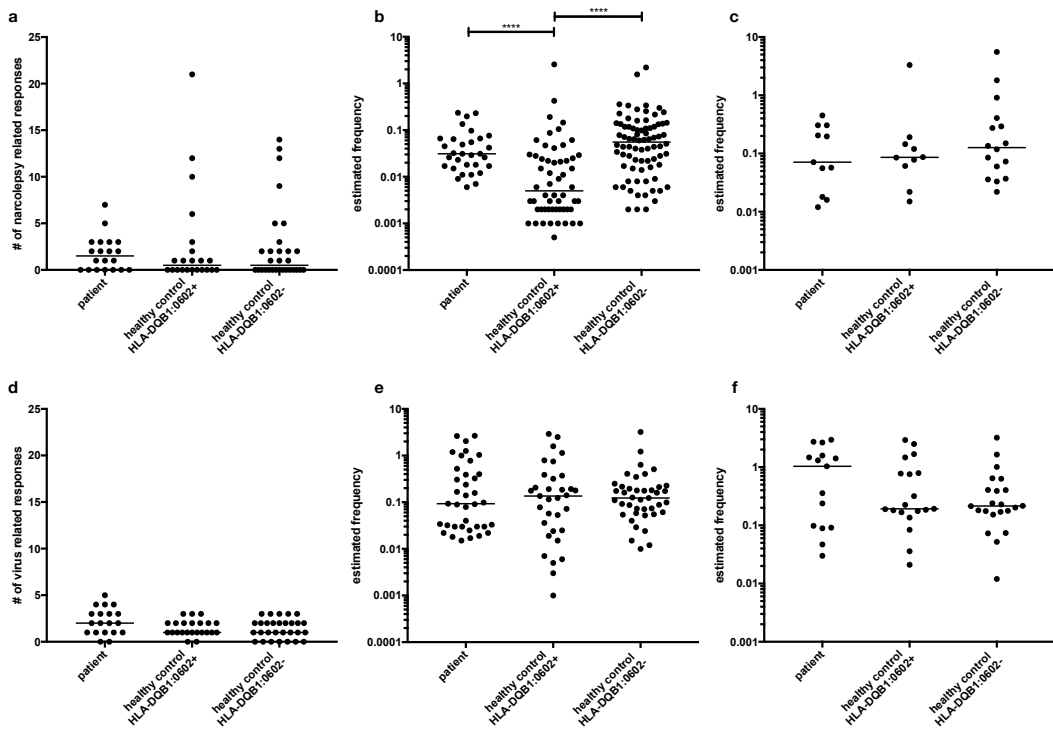
Figure 3. Number and estimated frequency of detected responses
a) The number of Narcolepsy associated responses detected in each donor. **b**) The estimated frequency of each detected response and **c**) the sum of responses in each donor. **d**), **e**) and **f**) shows the number of responses, frequency for individual responses and sum of responses in each donor respectively, for responses against known virus epitopes. Bars represent median values. **** : p< 0.0001 (Mann-Whitney test)

Both NT1 patients and healthy controls have CD8+ T cells responding to NT1 relevant autoepitiopes.

Since some individuals proved to have a CD8$^+$ T cell response to the same protein in the context of more than one HLA type, we combined all responses to determine the sum of the estimated frequency to NT1 responsive T cells for each protein in a given donor (Fig 4a, color gradient). As the library size used for T cell screening depended on the HLA expression of the donor, the response proportion was analyzed to correct for any potential biases based on differences in the pMHC library sizes (Fig 4a, circle size). The response proportion was calculated as the number of responses out of the total number of pMHCs used for screening the individual samples. Again, no clear systematic difference was observed between the three different groups. Thus, normalizing for the HLA type of the donors and the size of the pMHC library used for screening did not change the overall findings. The majority of patients (18/20; 90%) and healthy controls (42/50; 84%) displayed T cell recognition of virus-derived epitopes (Fig. 4a, bottom panel). Furthermore, these responses were generally more frequent than the NT1-related responses. The virus epitopes were selected from a pool of peptides known to frequently generate T cell responses in individuals carrying these common virus infections. Hence,

the virus epitopes served merely as a positive control and to ensure an overall similar immunological response capacity between the different cohorts. Consequently, for these highly selected peptides the response proportion was much higher than seen for the predicted NT1-derived epitopes (Fig 4a, bottom panel).

Finally, we determined the fraction of the individuals in the three different cohorts carrying a response to a given NT1-related protein for a given HLA type. Interestingly, the highest percentages of responders were found within the groups of donors expressing NT1-associated HLA types (Fig. 4b). The same pattern emerged when counting the number of responses for each protein-HLA combination and normalizing that to the total number of possible responses. The total number of possible responses was calculated as the number of samples with a given HLA times the number of peptides from the protein predicted for this HLA molecule (Fig. S2). Here again, the highest response frequency was found with the NT1-associated HLA alleles. Of the 20 patients included in this study, 9 were positive for an NT1-associated HLA allele and this was the case for 10/23 HLA-DQB1*06:02-positive controls and 15/29 HLA-DQB1*06:02-negative controls. All patients (9/9) with NT1-associated HLA types had specific T cells that recognized at least one of the NT1 related proteins in the context of the NT1-associated HLA-molecules. For the HLA-DQB1*06:02-positive and HLA-DQB1*06:02-negative healthy controls this was 60% (6/10) and 73% (11/15) respectively. When looking closer at this small cohort of donors with an NT1-associated HLA restricted CD8[+] T cell response, we found that HLA-DQB1*06:02-positive controls had significantly fewer responses than NT1 patients (Fig. 4c).
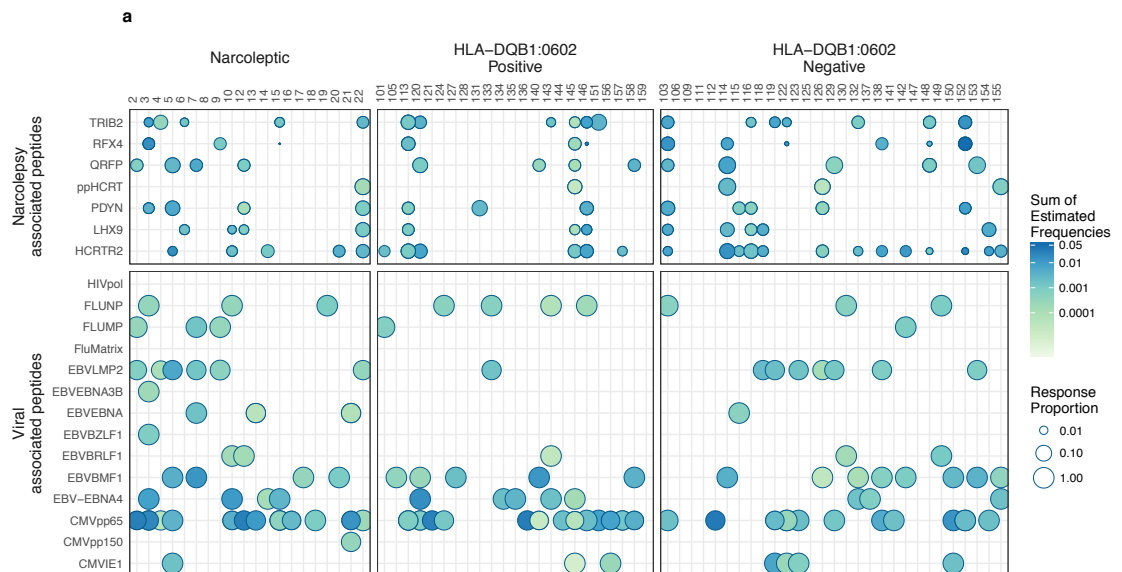
b
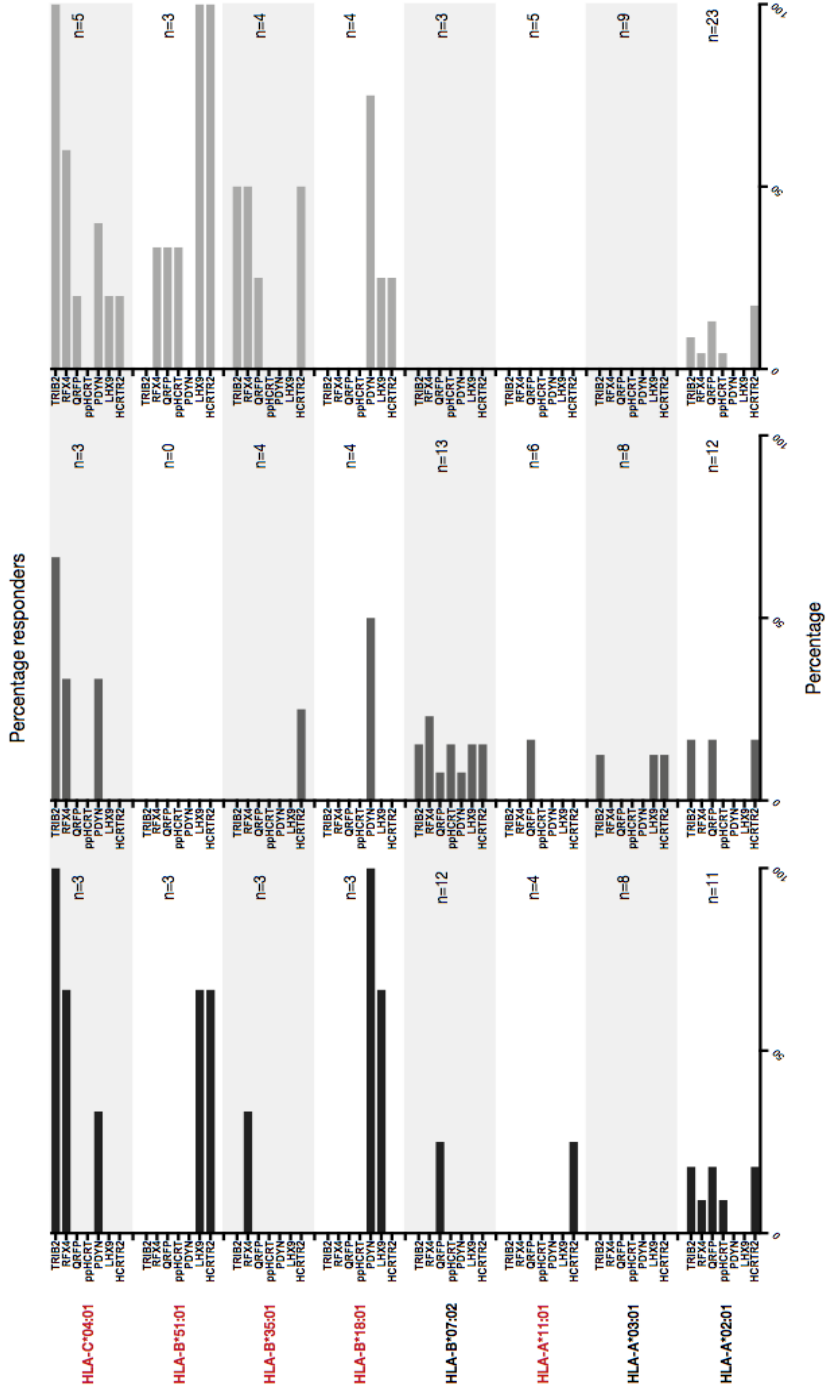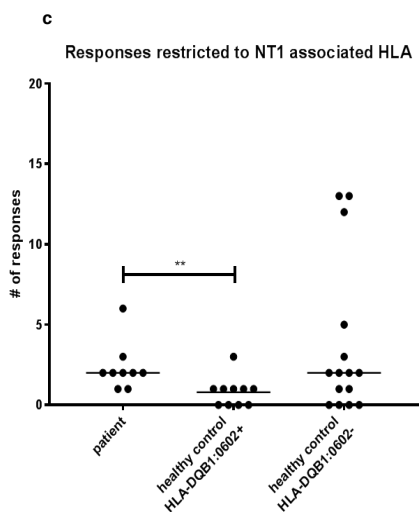
Figure 4. Response proportion and percentage.
a) Representation of all the detected narcolepsy relevant responses or responses towards known virus epitopes in the cohort. Each dot represents the full T cell recognition profile of the donor. Thus, if more than one pMHC specific T cell population was detected, these are combined. The size of the dot represents the number of peptides recognized by the given donor, relative to the total number of peptides screened for that donor. The color gradient represents the estimated frequency of the sum of responses for each donor. **b**) The percentage of donors for each protein-HLA combination with one or more positive response. n = the number of donors that were positive for the given HLA type and thus included in the group. HLA types written in red are associated to NT1. **c**) Number of responses towards a peptide restricted to an NT1-associated HLA allele. Bars represent median values. ** : p< 0.01 (Mann-Whitney test)

An NT1 relevant response towards a peptide from LHX9 was detected and confirmed with MHC multimer staining

The responses detected against NT1-related peptides were, in general, of very low frequency and therefore challenging to detect using conventional fluorescently-labeled MHC multimers, in part due to limited amounts of sample material. However, in one case an NT1 relevant response towards a peptide from LHX9 was detected and confirmed with MHC multimer staining (Fig. S3a). It was not possible to further *in-vitro* culture and expand NT1 relevant T cells, although virus-specific T cells from both patients and healthy controls could be expanded (Fig. S3b and c).

Immunogenic hotspots in NT1 relevant proteins

We mapped the location of all the possible NT1-related T cell responses observed in this study, both for patients and healthy controls. Looking at each protein included in the study, a series of immunological hotspots seems to be present in these proteins, characterized by the co-localization of peptide sequences of different HLA restrictions that were identified as immunogenic across the cohort (Fig. S4).

Discussion

Here we present a comprehensive screening for CD8$^+$ T cell recognition of NT1 relevant proteins in PBMCs from well-characterized hypocretin deficient NT1 patients and healthy donors.

We used DNA barcode-labeled MHC multimers to screen for CD8$^+$ T cell recognition of a large library of 1183 NT1 relevant peptides. These peptides were restricted to a range of HLA-molecules (HLA-A*11:01, HLA-B*18:01, B*35:01, B*51:01, and HLA-C*04:01) based on their previously described association with NT1 and three other HLA alleles (HLA-A*02:01, -A*03:01 and HLA-B*07:02) based on their high prevalence in the available samples. The samples were divided into three cohorts: NT1 patients (all but one were HLA-DQB1*06:02-positive), HLA-DQB1*06:02-positive healthy controls, and HLA-DQB1*06:02-negative healthy controls. We observed the expected levels of CD8$^+$ T cell recognition of virus-derived epitopes in all three cohorts; no differences were observed in virus response frequency or magnitude (Fig 3d-f). In all cohorts, recognition of possible NT1 associated peptide epitopes was observed, although at a significantly lower level than for virus-derived epitopes (Fig. 3). Interestingly, the estimated frequency of T cells recognizing NT1-relevant peptides was significantly higher in both the patient cohort and the HLA-DQB1*06:02-negative healthy control group, compared to the HLA-DQB1*06:02-positive healthy control group (Fig. 3b). This finding could indicate that the frequency of the individual autoreactive CD8$^+$ T cell population recognizing NT1-relevant proteins may play a role in NT1 disease development when combined with HLA-DQB1*06:02 expression. HLA-DQB1*06:02-negative individuals can carry higher levels of autoreactive T cells without the initiation of the disease whereas similar levels in HLA-DQB1*06:02-positive individuals may lead to disease development. Such a hypothesis requires further testing in a larger patient cohort.

We did not observe the same level of autoreactive responses among the different HLA-restrictions tested. Among the 20 patients included in this study, 9 carried one or more of the NT1 associated HLA class I alleles (HLA-A*11:01, HLA-B*18:01, B*35:01, B*51:01, and HLA-C*04:01). All of these 9 patients responded to one or more of the tested peptides. In contrast, among the 11 patients that did not express any of these HLA alleles, only 4 responded to one or more of the tested peptides. For the HLA-DQB1*06:02 positive controls, 10 out of 23 carried one or more of the NT1 associated HLA alleles and only 6 showed a response. Among the 15 HLA-DQB1*06:02 negative controls with an NT1 associated HLA allele (out of 29), 11 responded to NT-related peptides. Although not statistically significant in this small cohort, it is interesting that all patients with an NT1 associated HLA-molecule carry T cells that respond to NT1-related peptides in the given HLA context. Furthermore, we observed a significant difference in the number of the NT1-associated HLA allele responses, between the patients and HLA-DQB1*06:02 positive controls. Whereas 7/9 of the patients (and 9/11 HLA-DQB1*06:02-negative healthy controls) had more than a single CD8$^+$ T cell response towards a peptide restricted to an NT1-associated HLA allele, this was only the case for 1/6 of the HLA-DQB1*06:02-positive healthy controls. Thus, 5/6 HLA-DQB1*06:02-positive controls only had a single response towards an NT1 relevant peptide (Fig. 4c). This finding again suggests that HLA-DQB1*06:02-positive healthy controls could be protected from developing NT1 due to the

lower level of relevant autoreactive CD8[+] T cells and thus, that the combination of HLA-DQB1*06:02 expression and an HLA class I response is important for NT1 development.

In addition to the finding of autoreactive CD8[+] T cells in NT1 patients, the present study suggests that low level of autoreactive CD8[+] T cells of relevance for NT1 can be detected in most individuals, independent of their disease status and HLA types. Similar observations have been made in several studies of other autoimmune diseases. Disease relevant autoreactive CD8[+] T cells are found to be present in both multiple sclerosis (MS) patients and healthy controls[30–33] in several studies of in MS and the same is true for type 1 diabetes[34–36]. This point to the presence of an external trigger for the activation of autoreactive CD8[+] T cells as necessary for disease development. In NT1, external triggers, e.g., H1N1 infection or vaccination, could be hypothesized to boost preexisting autoreactive CD8[+] T cells.

Some HLA class II DQB1 alleles have been suggested to be protective in NT1. These are 05:01, 06:01, 06:03 and less strong all DQB1*02 alleles. We speculated whether the presence of these alleles differed between controls with or without CD8[+] T cell responses. Only 7 of the 23 HLA-DQB1*06:02 positive controls carried a putatively protective HLA class II type (Table S1). Three of these showed multimer positive CD8[+] T cells. Among the remaining HLA-DQB1*06:02 positive controls that were not protected by their HLA class II type, 9 individual still had multimer positive CD8[+] T cells. Furthermore, 6 of the 20 NT1 patients in this cohort carried a protective HLA class II type, indicating that the protective effect of these HLA types is most likely only minor.

A limitation in the present study is the fact that patient samples were collected from 1 to 9 years after disease onset. It is possible, that the autoimmune response in NT1 is not strong enough to establish a long lasting memory response in all patients. With this in mind, screening samples from patients very close to disease onset, would be a very interesting next step. Furthermore, the peripheral blood used for the current study may likely underrepresent a given level of CD8[+] T cell recognition associated with NT1. Detection of such T cells in the cerebrospinal fluid could be interesting and potentially more relevant as a measure of disease-initiating T cells.

Another challenge for the present study was the low level of CD8[+] T cell recognition, which compromised our ability to further study these autoreactive T cells by tracking them, e.g. using fluorescently-labeled MHC multimers. Also, autoreactive T cells are likely to have lower affinity for their target pMHC than virus-specific T cells, consequently making MHC multimer-based detection difficult. As mentioned earlier, we have previously demonstrated that the DNA barcoding-based strategy is especially valuable for detecting low affinity T cells[19], and hence the preferable choice in this case. Despite these limitations, in one case we were able to detect NT1-relevant CD8[+] T cells by a direct MHC multimer staining (Fig. S3), demonstrating the presence of such T cells.

In this study, we find that NT1 patients and HLA-DQB1*06:02 positive healthy controls are different in terms of their CD8[+] T cell recognition of NT1 relevant proteins. NT1 patients seem to have an increased frequency and breadth of their T cell response. Although the observation of a higher level of autoreactive CD8[+] T cells in patients, compared to HLA-DQB1*06:02 positive healthy controls, points to some involvement of CD8[+] T cells in disease development, the autoimmune nature of NT1 is still not indisputably proven. CD8[+] T cells are very likely only a single piece of a complex interplay

between many different cells of the immune system, and further studies are needed to fully elucidate the etiology of Narcolepsy type 1.

Materials and methods

Patient and healthy control material

20 ICSD-3[20] diagnosed Type-1 narcolepsy patients (age range: 7-62y, mean age: 25.6y; 9 male) were selected for blood donation from the known patient cohort at the Danish Centre for Sleep Medicine. Inclusion criteria were hypocretin deficiency (CSF hcrt-1 <110pg/ml) and disease duration ≤ 10 years (Table 1). Of this patient cohort, 19/20 were HLA-DQB1*06:02-positive. Patients were excluded in cases of: secondary narcolepsy; pregnancy; neurological or autoimmune co-morbidity (hay fever and asthma were accepted); severe medical or psychiatric co-morbidity; medical contraindications to blood donation; treatment with immunosuppressing drugs at any point in the past three months (asthma medication was accepted); or signs/history of infection in the two weeks prior to the blood donation.

Two patients were drug naïve, 18/20 were using stimulants (methylphenidate, modafinil) and/or anticataplectic drugs (tricyclic antidepressants (TCAs), Venlafaxin (SSRI), or Xyrem (Sodium Oxybate)). Three out of 20 patients were using asthma medication (antihistamines and/or mild inhalator/oral steroids or Montelukast (leukotriene D4 receptor antagonist)). All patients had normal white and red blood cell counts.

52 healthy controls were selected from participants in the Danish Blood Donor Study based on age and gender; 23 of them were HLA-DQB1*06:02-positive (age range: 28-60y, mean age: 42y; 9 male), and 29 of them were HLA-DQB1*06:02-negative (age range: 19-65y, mean age: 40y; 14 male). Inclusion criteria were age ≥ 18 years and fulfilling the blood bank's general criteria for health.

PBMC extraction

To collect peripheral blood mononuclear cells (PBMCs), a blood donation (450 ml) was obtained via the blood bank of Rigshospitalet, Denmark. The blood donations from the Danish blood bank were transported to the laboratory the same day and PBMCs were fractionated into plasma, buffy coat, and red blood cells. PBMCs from the buffy coat were next extracted using Ficoll gradient reagent, frozen in 10% DMSO (Sigma) in fetal bovine serum (FBS, Gibco), and cryopreserved at -140°C until use.

Peptide prediction analysis

The online available prediction algorithm NetMHCcons 1.1 (http://www.cbs.dtu.dk/services/NetMHCcons/), was used to predict 9 – 11 mer MHC class I T cell epitopes within 7 different proteins associated with hypocretin neurons. Peptides were predicted from the following proteins: HCRT (O43612-1), TRIB2 (3 variants Q92519-1, F8WA18, B5MCX4), RFX4 (10 variants Q33E94-1, Q33E94-2, Q33E94-3, Q33E94-4, R4GMS3, F8VRD4, F8VZC4, B4DZB7, F8W1T9, F8VX50), PDYN (P01213-1), LHX9 (8 variants Q9NQ69-1, Q9NQ69-2, Q9NQ69-3, Q9NQ69-4 , H0YL54, H0Y330, A0A087X083, A0A0C4DGY4), QRFP (P83859-1) and HCRTR2 (O43614-1). Eight different HLA molecules were included in the analysis: HLA-A*02:01, HLA-A*03:01, HLA-B*07:02, HLA-A*11:01, HLA-B*18:01, HLA-B*35:01, HLA-B*51:01, and

HLA-C*04:01. This yielded a total of 1183 predicted T cell epitopes spread on the eight different HLA molecules and seven different proteins. The peptides were synthesized and purchased from Pepscan (Pepscan Presto, The Netherlands) and dissolved to 10mM in DMSO.

HLA analysis across the cohort
All donors included in the study were HLA typed by next-generation sequencing (NGS) of the full exome of MHC class I HLA-A, -B, and –C. The full exome of MHC class II DQA was also sequenced, but only exon 2 and 3 of HLA-DQB1 were sequenced (GenDx, Netherlands). MiSeq from Illumina was used for all sequence analyses. The full HLA type of each donor is presented in Table S1. The highlighted HLA subtypes are associated with an increased risk of NT1[10, 11].

Generation of DNA barcodes and dextran conjugation
Attachment of 5′ biotinylated AxBy DNA barcodes to PE- and streptavidin-conjugated dextran was performed as described in [19]. Oligonucleotides containing distinct 25-mer nucleotide sequences[37] were purchased from LGC Biosearch Technologies (Denmark) and PE- and streptavidin-conjugated dextran was provided by Immudex (Denmark). All oligos carried a 6-nt unique molecular identifier (UMI)[38].

Generation of peptide-MHC class I multimers labeled with DNA barcode
A total of 1183 pMHC complexes were generated through UV-mediated conditional ligand exchange as described previously[29]. In short, HLA monomers carrying UV-sensitive ligands were mixed with matching peptides at a final concentration of 50 g/ml monomer and 100 mM peptide. The complexes were then coupled to DNA barcode- and PE-labeled dextran backbones, incubated for 20 min on ice and incubated another 20 min on ice with a freezing buffer (PBS + 0.5% BSA + 100 μg/mL herring DNA + 2 mM EDTA + 5% glycerol) supplemented with 909 nM D-biotin. The complexes were finally stored at -20 °C until use.

T cell staining with barcode-labeled MHC multimers
Cryopreserved cells were thawed and washed twice in RPMI + 10% FCS and then washed in barcode-cytometry buffer (PBS + 0.5% BSA + 100 μg/mL herring DNA + 2 mM EDTA). On the day of staining, barcode-labeled MHC multimers were thawed on ice, centrifuged for 5 min at 3,300 g, and 1.5 μl (corresponding to 0.043 μg) of each distinct pMHC was pooled according to the HLA type of the donors. The volume of the reagent pool was reduced by ultrafiltration to obtain a final volume of ~80 μL of MHC multimers. Centrifugal concentrators (Vivaspin 6, 100,000 Dalton, Sartorius) was prepared as described in Bentzen et al. An aliquot of ~5 μl of the MHC multimer reagent pool was stored at −20 °C for baseline analysis. The MHC multimer pool was added to 10×10^6 cells and incubated for 15 min at 37 °C in the presence of 50 nM dasatinib in a total volume of 80 μl. Next, a 5x antibody mix composed of CD8-BV480 (BD 566121, clone RPA-T8) (final dilution 1/50), dump channel antibodies (CD4-FITC (BD 345768) (final dilution 1/80), CD14-FITC (BD 345784) (final dilution 1/32), CD19-FITC (BD 345776) (final dilution 1/16), CD40-FITC (Serotech MCA1590F) (final dilution 1/40), and CD16-FITC (BD 335035) (final dilution 1/64)) and a dead cell marker (LIVE/DEAD Fixable Near-IR; Invitrogen L10119) (final dilution 1/1000) was added and incubated

for 30 min at 4 °C. Cells were washed three times in barcode-cytometry buffer and fixed in 1% paraformaldehyde (PFA). After 0.5 - 24 h, the cells were washed twice more and resuspended in barcode-cytometry buffer. Cells were acquired within a week after multimer staining.

Sorting of MHC multimer-positive T cells
Cells were sorted on a FACSAriaFusion (BD) into tubes containing 100 μl of barcode-cytometry buffer (tubes were saturated with PBS + 2% BSA in advance). Using FACSDiva software, we gated on single, live, CD8-positive and dump channel (CD4, 14, 16, 19, and 40)-negative lymphocytes and sorted all multimer-positive (PE) cells within this population. The sorted cells were centrifuged for 10 min at 5,000 g and the buffer was removed. The cell pellet was stored at −80 °C in a minimal amount of residual buffer (<20 μL). The gating strategy is exemplified in Figure S5.

DNA barcode amplification
DNA barcode amplification was performed as described in [19]. PCR amplification was conducted on isolated cells (in <20 μL of buffer) and on the stored aliquot of the MHC multimer reagent pool (diluted 50,000× in the final PCR), which was used as the baseline to determine the number of DNA barcode reads within a non-processed MHC multimer reagent library. PCR products were purified with a QIAquick PCR Purification kit (Qiagen, 28104). The amplified DNA barcodes were sequenced at Sequetech (USA) using an Ion Torrent PGM 316 or 318 chip (Life Technologies).

Processing of sequencing data derived from multimer-associated DNA barcodes
Sequencing data were processed by the software package 'Barracoda,' available online at (http://www.cbs.dtu.dk/services/barracoda)[19]. This tool identifies the barcodes used in a given experiment, assigns sampleID and pMHC specificity to each barcode, and counts the total number of reads and clonally reduced reads for each pMHC-associated DNA barcode. Furthermore, it accounts for barcode enrichment based on methods designed for the analysis of RNA-seq data: specifically, log_2 fold changes in read counts mapped to a given sample relative to the mean read counts mapped to triplicate baseline samples are estimated using normalization factors determined by the trimmed mean of M-values method[19, 39]. Barcodes with a log_2 fold change with a $p<0.001$ were considered to be true T cell responses.

Verification of responses with flow cytometry
In order to verify some of the detected responses, a smaller cohort of 40 selected peptides, to which specific T cells had been detected, were used to generate fluorescently-labeled MHC multimers. Again, UV-mediated peptide exchange was used and pMHC complexes were multimerized using streptavidin-conjugated PE (405204, BioLegend), APC (405207, BioLegend), PE-Cy7 (405206, BioLegend), PECF594 (562284, BD), BV421 (563259, BD), BV605 (563260, BD), BV650 (563855, BD), BUV395 (564176, BD), and BUV737 (564293, BD). The final concentrations of monomer and peptide were 100 μg/ml and 200 nM, respectively. Each pMHC was made in two different colors and mixed before the MHC multimers were stored at -20°C with 0.5% bovine serum albumin (Sigma Aldrich) and 5% glycerol (Fluka). For ex vivo T cell staining, samples from cells were thawed as described above and $2x10^6$ cells were stained with 1 μl of each pMHC, according to HLA type, for

15 min at 37 °C. Cells were then stained with a 5x antibody mix as described for staining with DNA barcode-labeled MHC multimers. Following two washing steps and fixation in PFA, cells were acquired on an LSR Fortessa (BD). The gating strategy is identical to the one described above and presented in Figure S5.

Enrichment of specific T cells

Six donors (three patients and three healthy controls) with detected associated NTI-relevant T cell responses were selected for enrichment of specific T cells. PE-labeled MHC multimers were generated in the same fashion as described for the verification. Cells were thawed and washed twice in RPMI + 10% FCS and $1x10^6$ cells were taken to a new tube and irradiated at ~20 Gy for 6 min and 15 sec. These cells were washed twice and kept for later use as feeder cells. From the remaining cells, $20x10^6$ cells were stained with 1 μl per pMHC of the PE-labeled MHC multimers for 60 min at 4 °C. Following staining, the cells were washed twice and stained with   -PE beads (Miltenyi Biotec) for 15 min at 4 °C. The cells where then washed twice again and loaded onto a magnetic column placed in the magnetic field of a MACS separator (Miltenyi Biotec), which captured the cells with a PE label through the   -PE beads. After washing the column twice, the column was removed from the separator and the specific cells were washed into a new tube containing X-vivo media (Lonza) + 5% human serum (HS; Sigma Aldrich). The specific cells were counted, washed, and cultured in X-vivo + 5% HS supplemented with 23.8 U/mL IL-15 (PeproTech, #200-15-10UG) and 100 U/mL IL-2 (PeproTech, #200-02-50UG). ~ 5000 specific cells were cultured with 50,000 feeder cells/well as well as 11,000 CD3/CD28 Dynabeads (Gibco). After 1.5 weeks of culturing, $2x10^6$ cells were stained with the above-described fluorescently-labeled MHC multimers and acquired on an LSR Fortessa (BD).

Graph generation and statistical data analysis

Graphs were generated using either GraphPad Prism 7.0b or ggplot2 in R and edited using OmniGraffle version 7.8.1. Statistical analysis of DNA barcoding data was performed using the software package 'Barracoda' as described previously. Significant differences between the number and frequency of detected responses in Figure 4 were calculated in GraphPad Prism 7.0b.

Ethical approval

All healthy donor material was collected with the approval of the local Scientific Ethics Committee and written informed consent was obtained according to the Declaration of Helsinki.

Author contribution

NWP performed experiments, data analysis, generated figures, and wrote the manuscript; AH performed experiments, data analysis and wrote the manuscript; NPK performed experiments; AMB performed data analysis and generated figures; AKB provided technical assistance, discussed data, and revised the manuscript; TT generated MHC monomers and revised the manuscript; KSB and HU provided donor material, PJ, and SK provided donor material, diagnosed and characterized the patients and revised the manuscript; and SRH and BRK conceived the concept, discussed the data, and wrote the manuscript.

References

1. Dauvilliers, Y., Arnulf, I. & Mignot, E. Narcolepsy with cataplexy. *Lancet* **369,** 499–511 (1980).

2. Overeem, S. *et al.* The clinical features of cataplexy: A questionnaire study in narcolepsy patients with and without hypocretin-1 deficiency. *Sleep Med.* **12,** 12–18 (2011).

3. Plazzi, G., Serra, L. & Ferri, R. Nocturnal aspects of narcolepsy with cataplexy. *Sleep Med. Rev.* **12,** 109–128 (2008).

4. Nishino, S., Ripley, B., Overeem, S., Lammers, G. J. & Mignot, E. Hypocretin (orexin) deficiency in human narcolepsy. *Lancet* **355,** 39–40 (2000).

5. Thannickal, T. C. *et al.* Reduced number of hypocretin neurons in human narcolepsy. *Neuron* **27,** 469–474 (2000).

6. Thannickal, T. C., Siegel, J. M., Nienhuis, R. & Moore, R. Y. Pattern of hypocretin (orexin) soma and axon loss, and gliosis, in human narcolepsy. *Brain Pathol.* **13,** 340–351 (2003).

7. Mignot, E. *et al.* Complex HLA-DR and -DQ Interactions Confer Risk of Narcolepsy-Cataplexy in Three Ethnic Groups. *Am. J. Hum. Genet.* **68,** 686–699 (2001).

8. Juji, T., Matsuki, K., Tokunaga, K., Naohara, T. & Honda, Y. Narcolepsy and HLA in the Japanese. *Ann. N. Y. Acad. Sci.* **540,** 106–114 (1988).

9. Han, F. *et al.* HLA DQB1*06:02 Negative Narcolepsy with Hypocretin/Orexin Deficiency. *Sleep* **37,** 1601–1608 (2014).

10. Tafti, M. *et al.* Narcolepsy-Associated HLA Class I Alleles Implicate Cell-Mediated Cytotoxicity. *Sleep* **39,** 581–587 (2016).

11. Ollila, H. M. *et al.* HLA-DPB1 and HLA class i confer risk of and protection from narcolepsy. *Am. J. Hum. Genet.* **96,** 136–146 (2015).

12. Partinen, M. *et al.* Increased incidence and clinical picture of childhood narcolepsy following the 2009 H1N1 pandemic vaccination campaign in Finland. *PLoS One* **7,** 1–9 (2012).

13. Han, F. *et al.* Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in china. *Ann. Neurol.* **70,** 410–417 (2011).

14. Dauvilliers, Y. *et al.* Increased risk of narcolepsy in children and adults after pandemic H1N1 vaccination in France. *Brain* **136,** 2486–2496 (2013).

15. Kornum, B. R. *et al.* Narcolepsy. *Nat. Rev. Dis. Prim.* **3,** (2017).

16. Degn, M. & Kornum, B. R. Type 1 narcolepsy: a CD8 $^+$ T cell-mediated disease? *Ann. N. Y. Acad. Sci.* **1351,** 80–88 (2015).

17. Dauvilliers, Y. *et al.* Hypothalamic immunopathology in anti-ma-associated diencephalitis

with narcolepsy-cataplexy. *JAMA Neurol.* **70,** 1305–1310 (2013).

18. Bernard-Valnet, R. *et al.* CD8 T cell-mediated killing of orexinergic neurons induces a narcolepsy-like phenotype in mice. *Proc. Natl. Acad. Sci.* **113,** 10956–10961 (2016).

19. Bentzen, A. K. *et al.* Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nat. Biotechnol.* **34,** 1037–1045 (2016).

20. *International classification of sleep disorders, 3rd edition.* (American Academy of Sleep Medicine, 2014).

21. Cvetkovic-lopes, V. *et al.* Elevated Tribbles homolog 2-specific antibody levels in narcolepsy patients. **120,** 713–719 (2010).

22. Dalal, J. *et al.* Translational profiling of hypocretin neurons identifies candidate molecules for sleep regulation. *Genes Dev.* **27,** 565–578 (2013).

23. Honda, M. *et al.* IGFBP3 colocalizes with and regulates hypocretin (orexin). *PLoS One* **4,** 1–14 (2009).

24. Romanov, R. A. *et al.* Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nat. Neurosci.* **20,** 176–188 (2017).

25. Liu, J. *et al.* Evolutionarily conserved regulation of hypocretin neuron specification by Lhx9. *Development* **142,** 1113–1124 (2015).

26. Vassalli, A., Li, S. & Tafti, M. Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2. *Sci. Transl. Med.* **7,** 2–5 (2015).

27. Fagerberg, L. *et al.* Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Mol. Cell. Proteomics* **13,** 397–406 (2014).

28. Karosiene, E., Lundegaard, C., Lund, O. & Nielsen, M. NetMHCcons: A consensus method for the major histocompatibility complex class i predictions. *Immunogenetics* **64,** 177–186 (2012).

29. Toebes, M. *et al.* Design and use of conditional MHC class I ligands. *Nat. Med.* **12,** 246–251 (2006).

30. Ota, K. *et al.* T-cell recognition of an immunodominant myelin basic protein epitope in multiple sclerosis. *Nature* **1,** 15796–15799 (1990).

31. Martin, R. *et al.* Fine specificity and HLA restriction of myelin basic protein-specific cytotoxic T cell lines from multiple sclerosis patients and healthy individuals. *J Immunol* **145,** 540–548 (1990).

32. Berthelot, L. *et al.* Blood CD8+ T cell responses against myelin determinants in multiple sclerosis and healthy individuals. *Eur. J. Immunol.* **38,** 1889–1899 (2008).

33. Crawford, M. P. *et al.* High prevalence of autoreactive , neuroantigen-specific CD8 ♂ T cells in multiple sclerosis revealed by novel flow cytometric assay. **103,** 4222–4232 (2015).

34. Sachdeva, N. *et al.* Preproinsulin specific CD8+ T cells in subjects with latent autoimmune diabetes show lower frequency and different pathophysiological characteristics than those with type 1 diabetes. *Clin. Immunol.* **157,** 78–90 (2015).

35. Culina, S. *et al.* Islet-reactive CD8+ T-cell frequencies in the pancreas but not blood distinguish type 1 diabetes from healthy donors. *Sci Immunol* **3,** 1119–1130 (2018).

36. Ehlers, M. R. Who let the dogs out? The ever-present threat of autoreactive T cells. *Sci.*

Immunol. **3,** eaar6602 (2018).

37. Xu, Q., Schlabach, M. R., Hannon, G. J. & Elledge, S. J. Design of 240,000 orthogonal 25mer DNA barcode probes. *Proc. Natl. Acad. Sci.* **106,** 2289–2294 (2009).

38. Kivioja, T. *et al.* Counting absolute numbers of molecules using unique molecular identifiers. *Nat. Methods* **9,** 72–74 (2012).

39. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11,** R25 (2010).

Supplementary figures
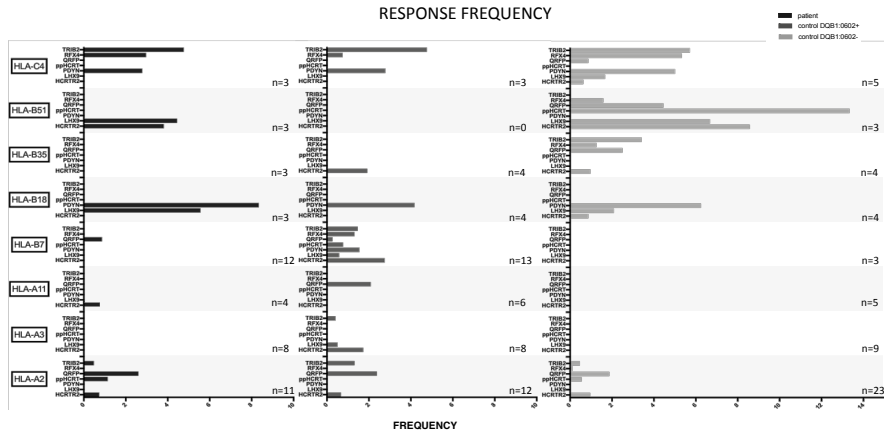
Supplementary figure 2
The response percentage within each protein and HLA type.
The number of responses across all donors within a given protein and HLA/ the total number of possible responses (number of peptides within the HLA and protein
* number of donors with the HLA type). n = the number of donors in each HLA type.

Supplementary figure 3
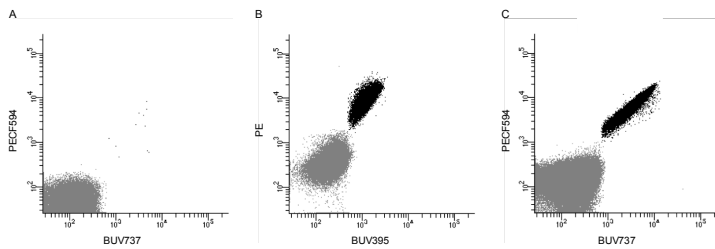
Supplementary figure 3
Verification of responses
A) Dot plot showing an ex vivo multimer staining for the LHX9 peptide 740 binding to HLA-B*51:01 in patient 6. The peptide 740 multimer was made with 2 different
fluorochromes, PECF594 and BUV737 and double positive cells are peptide specific T cells. B) and C) Examples of dot plots showing virus specific T cells after enrichment
from a Narcolepsy patient and healthy donor respectively. PE: phycoerythrin. BUV: brilliant ultra violet.
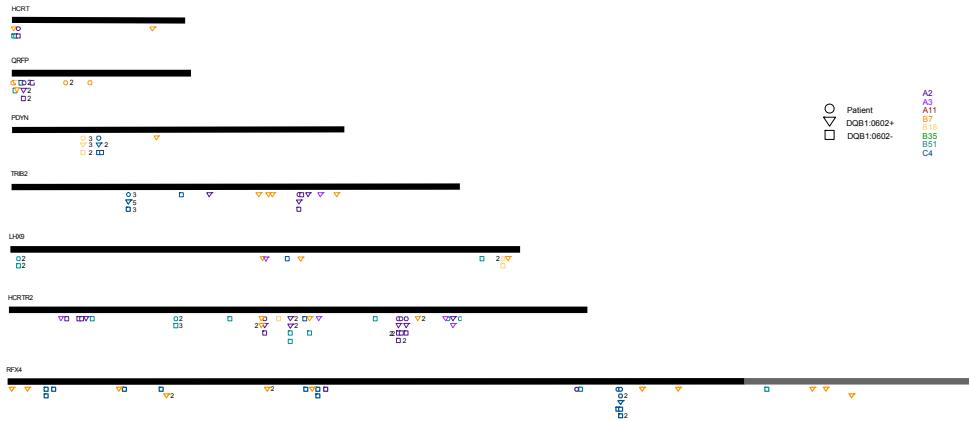
Supplementary figure 4
Immunogenic hotspots.
Overview of the position of all peptides with a detected T cell response and their position within the seven proteins. The length of the black line represents the relative size of the protein. The length of the protein represents all different isoforms of the protein. For RFX4, the grey line indicates an isoform longer than the original sequence.

Supplementary figure 5
Gating strategy used to identify single, live lymphocytes and within these, CD8 and multimer specific T cells.

Patients	HLA-A		HLA-B		HLA-C		HLA-DQA1		HLA-DQB1	
2	02:01:01	29:02:01	07:02:01	15:01:01	03:04:01	07:02:01	01:02:01	05:01:01	02:01:01	06:02:01
3	03:01:01	11:01:01	35:01:01	55:01:01	03:03:01	04:01:01	01:01:01	01:02:01	05:01:01	06:02:01
4	01:01:01	02:01:01	07:02:01	08:01:01	07:01:01	07:02:01	01:02:01		06:02:01	
5	02:01:01		18:01:01	44:02:01	05:01:01	07:01:01	01:02:01	03:03:01	03:01:01	06:02:01
6	03:01:01	68:01:02	35:01:01	51:01:01	04:01:01	15:02:01	01:01:01	01:02:01	05:01:01	06:02:01
7	02:01:01	24:02:01	07:02:01		07:02:01		01:02:01		06:02:01	
8	03:01:01		07:02:01		07:02:01		01:02:01		06:02:01	
9	02:01:01		07:02:01	15:01:01	03:04:01	07:02:01	01:02:01	03:01:01	03:02:01	06:02:01
10	03:01:01	11:01:01	07:02:01	51:01:01	07:02:01	15:02:01	01:02:01	03:01:01	03:02:01	06:02:01
12	03:01:01	25:01:01	07:02:01	18:01:01	07:02:01	12:03:01	01:02:01		06:02:01	
13	01:01:01	03:01:01	07:02:01	15:01:01	07:02:01		01:02:01	05:05:01	03:01:01	06:02:01
14	11:01:01	24:02:01	40:01:02	52:01:01	03:04:01	12:02:02	01:03:01	03:03:01	03:01:01	06:01:01
15	11:01:01	24:02:01	07:02:01	35:01:01	04:01:01	07:02:01	01:02:01	03:01:01	03:02:01	06:02:01
16	02:01:01	26:01:01:01	07:02:01	39:01:01	07:02:01	12:03:01	01:02:01	01:03:01	06:02:01	06:03:01
17	01:01:01	02:01:01	44:02:01	57:01:01	05:01:01	06:02:01	01:01:01	01:02:01	05:01:01	06:02:01
18	01:01:01	02:01:01	08:01:01	44:02:01	05:01:01	07:01:01	01:02:01	05:05:01	03:01:01	06:02:01
19	01:01:01	03:01:01	07:02:01		07:02:01		01:02:01	01:04:01	05:03:01	06:02:01
20	01:01:01	02:01:01	08:01:01	51:01:01	01:02:01	07:01:01	01:02:01		06:02:01	
21	02:01:01	03:01:01	07:02:01	15:01:01	03:04:01	07:02:01	01:02:01	03:01:01	03:02:01	06:02:01
22	02:01:01	25:01:01	18:01:01	57:01:01	06:02:01	12:03:01	01:02:01	02:01:01	03:03:02	06:02:01
Healthy controls										
101	01:01:01	02:01:01	07:02:01	08:01:01	07:01:01	07:02:01	01:02:01	03:02	03:03:02	06:02:01
103	03:01:01	23:01:01	35:01:01	44:02:01	04:01:01	07:04:01	01:01:01	01:02:01	05:01:01	06:04:01
105	01:01:01	02:01:01	39:06:02	57:01:01	06:02:01	07:02:01	01:02:01	02:01:01	03:03:02	06:02:01
106	02:01:01	32:01:01	07:02:01	14:02:01	02:02:02	08:02:01	01:02:01	02:01:01	02:02:01	05:01:01
109	25:01:01	29:02:01	18:01:01	44:03:01	12:03:01	16:01:01	01:01:01	02:01:01	02:02:01	05:01:01
111	02:01:01		15:01:01	40:01:02	03:04:01		01:02:01	03:01:01	03:02:01	06:04:01
112	24:02:01	30:04:01	07:02:01	50:01:01	06:02:01	07:02:01	02:01:01	05:05:01	02:02:01	03:01:01
113	11:01:01	23:01:01	07:02:01	35:01:01	04:01:01	07:02:01	01:01:01	01:02:01	05:01:01	06:02:01
114	02:01:01	68:01:02	44:02:01	51:01:01	05:01:01	14:02:01	01:03:01	02:01:01	02:02:01	06:03:01
115	01:01:01	03:01:01	18:01:01	37:01:01	06:02:01	12:03:01	01:01:01	05:05:01	03:01:01	05:01:01
116	02:01:01	03:01:01	15:01:01	18:01:01	03:03:01	05:01:01	05:01:01	05:05:01	02:01:01	03:01:01
118	02:01:01	03:01:01	40:02:01	51:01:01	01:02:01	02:02:02	01:03:01	05:01:01	02:01:01	06:03:01
119	02:01:01	23:01:01	07:02:01	44:03:01	04:01:01	07:02:01	01:03:01	02:01:01	02:02:01	06:03:01
120	02:01:01	11:01:01	39:01:01	44:02:01	07:02:01	07:04:01	01:02:01	02:01:01	03:03:02	06:02:01
121	24:02:01	33:03:01	07:02:01	14:02:01	07:02:01	08:02:01	01:01:01	01:02:01	05:01:01	06:02:01
122	02:01:01	03:01:01	35:01:01	44:02:01	04:01:01	05:01:01	01:02:01	03:03:01	03:01:01	06:04:01
123	01:01:01	02:01:01	08:01:01	40:01:02	07:01:01	14:02:01	01:02:01	05:01:01	02:01:01	06:04:01
124	03:01:01		07:02:01		07:02:01		01:02:01		06:02:01	
125	02:01:01	29:02:01	38:01:01	44:02:01	05:01:01	12:03:01	01:04:01	05:01:01	02:01:01	05:03:01
126	02:01:01	11:01:01	18:01:01	57:01:01	05:01:01	06:02:01	01:03:01	02:01:01	03:03:02	06:01:01
127	02:01:01	30:01:01	13:02:01	15:01:01	06:02:01	07:02:01	01:02:01	02:01:01	03:02:01	06:02:01
128	03:01:01	25:01:01	18:01:01	40:01:02	03:04:01	12:03:01	01:02:01	01:02:02	05:02:01	06:02:01
129	02:01:01	02:05:01	44:02:01	50:01:01	05:01:01	06:02:01	01:02:01	03:01:01	03:02:01	06:09:01
130	03:01:01	26:01:01:01	38:01:01	40:01:02	03:04:01	12:03:01	01:02:01	05:05:01	03:01:01	06:04:01
131	25:01:01	32:01:01	18:01:01	27:05:02	02:02:02	12:03:01	01:02:01	03:02	03:03:02	06:02:01
132	02:01:01	11:01:01	40:01:02	44:02:01	03:04:01	05:01:01	01:02:01	03:03:01	03:01:01	06:04:01
133	02:01:01	03:01:01	07:02:01	57:01:01	06:02:01	07:02:01	01:02:01	01:04:01	05:03:01	06:02:01
143	03:01:01	11:01:01	07:02:01	40:02:01	02:02:02	07:02:01	01:02:01	05:05:01	03:01:01	06:02:01
145	01:01:01	11:01:01	07:02:01	18:01:01	07:01:01	07:02:01	01:02:01	03:01:01	03:02:01	06:02:01
136	03:01:01	24:02:01	07:02:01	15:01:01	03:03:01	07:02:01	01:02:01	05:05:01	03:01:01	06:02:01
137	02:01:01	11:01:01	44:02:01	56:01:01	01:02:01	05:01:01	01:03:01	03:01:01	03:02:01	06:03:01
138	02:01:01	24:02:01	44:02:01	55:01:01	03:03:01	05:01:01	01:01:01	01:02:01	05:01:01	06:04:01
140	01:01:01	02:01:01	07:02:01	08:01:01	07:01:01	07:02:01	01:02:01	05:01:01	02:01:01	06:02:01
141	02:01:01	03:01:01	14:01:01	14:02:01	08:02:01		01:02:01	03:03:01	03:02:01	06:09:01
142	02:01:01	31:01:02	07:06:01	40:01:02	03:04:01	15:05:02	01:03:01	03:01:01	03:02:01	06:03:01
143	03:01:01	11:01:01	35:01:01	41:02:01	04:01:01	17:03	01:02:01	03:03:01	03:01:01	06:02:01
144	01:01:01	03:01:01	07:02:01	08:01:01	07:01:01	07:02:01	01:02:01	05:01:01	02:01:01	06:02:01
145	02:01:01	11:01:01	07:02:01	55:01:01	03:03:01	07:02:01	01:02:01	01:04:01	05:03:01	06:02:01
146	03:01:01	25:01:01	18:01:01	35:01:01	04:01:01	12:03:01	01:02:01		06:02:01	
147	02:01:01	31:01:02	27:05:02	40:01:02	02:02:02	03:04:01	01:02:01	03:03:01	03:02:01	06:04:01
148	02:01:01	11:01:01	35:01:01	40:02:01	02:02:02	04:01:01	01:03:01	03:01:01	03:02:01	06:03:01
149	03:01:01	24:02:01	15:01:01	38:01:01	03:04:01	12:03:01	01:03:01	03:01:01	03:02:01	06:03:01
150	02:01:01		40:01:02	44:02:01	03:04:01	05:01:01	01:01:01	01:03:01	05:01:01	06:03:01
151	02:01:01	24:02:01	40:01:02	44:02:01	03:04:01	16:04:01	01:02:01	05:05:01	03:01:01	06:02:01
152	02:01:01	03:01:01	35:01:01	39:01:01	04:01:01	12:03:01	01:02:01	05:05:01	03:01:01	06:04:01
153	02:01:01		40:01:02	44:02:01	03:04:01	05:01:01	01:01:01	01:02:01	05:01:01	06:04:01
154	02:01:01	31:01:02	40:01:02	51:01:01	03:04:01	15:02:01	01:02:01	03:01:01	03:02:01	06:04:01
155	02:01:01	11:01:01	39:06:02	40:01:02	03:04:01	07:02:01	01:02:01	03:01:01	03:02:01	06:04:01
156	02:01:01	02:05:01	07:02:01	50:01:01	06:02:01	07:02:01	01:02:01	02:01:01	02:02:01	06:02:01
157	02:01:01		35:01:01	44:02:01	03:03:01	05:01:01	01:02:01	03:02	03:03:02	06:02:01
158	01:01:01	02:01:01	07:02:01	08:01:01	07:01:01	07:02:01	01:02:01	05:01:01	02:01:01	06:02:01
159	02:01:01	02:06:01	15:01:01	40:06:01	03:03:01	08:01:01	01:02:01	01:03:01	06:01:01	06:02:01

Supplementary tabel 1
The full HLA type of each donor

CHAPTER 5

Epilogue

The overall theme of this PhD thesis is antigen specific CD8 T cells and the three research projects presented touch upon the aspects of optimized detection and analysis of these cells as well as identification of disease relevant auto reactive CD8 T cells. Common to all projects is the use of MHC multimers for identification, although different tagging strategies are used in the different papers. During the conduction of this PhD, the DNA barcode-labeled MHC multimer technology was developed and the change in detection strategy applied here reflects the introduction of this new state of the art method. Although this is now the preferred method used in our lab, routine detection of antigen specific T cells with MHC multimers relies on fluorescence signals and thus, the data presented in **manuscript I** is of interest to the general immunological scientific community.

In this study, we compare the detection of virus specific CD8 T cell responses using 4 different fluorochromes with a range of intensities. The general observation made in manuscript I was that the choice of fluorochrome did not impact the ability to detect both high and low frequency populations of virus specific T cells. There was, however, a clear impact on the staining index (SI) of the multimer populations depending on the fluorescent label used, which correlated with the ability of gating experts to confidently identify the given population. In this particular setting, where only virus specific T cells were investigated, this did not influence the detection of responses. Self-reactive T cells that have escaped negative selection in the thymus however, will likely have lower affinity towards their cognate antigen and thus be harder to separate from background events in an MHC multimer staining. SI and consequent fluorochrome choice may thus be an important aspect to consider especially when looking for self-reactive CD8 T cells, regardless of whether they are desired responses in cancer or unwanted responses leading to autoimmunity.

Since the completion of the study many new fluorochromes have been developed and with new flow cytometry instruments available, choosing only high intensity fluorochromes for MHC multimer detection while still maintaining the same assay complexity might very well be an option in the near future. Excluding dim fluorochromes from MHC multimer assays would increase the signal to noise ratio and enhance the ability to detect low-affinity T cells. However, the immense diversity of the immune system will always leave researchers lagging behind, limited by the complexity that their assays allow. Thus, the introduction of new fluorochromes will most likely not prompt researchers to include the same number of parameters as currently possible

using higher intensity fluorochromes, but rather it will foster higher dimensionality of analysis by the inclusion of more colors to flow cytometry experiments. Therefore, the careful choice and optimization of fluorochrome detection will continue to be an important issue in MHC multimer assays even in the bright fluorochrome future.

For the optimization of fluorescence detection, we propose a bead based evaluation of detector performance that is reliable and easily applicable compared to the extensive optimization protocol proposed by (Perfetto *et al.,* 2012). Importantly, the relevance for optimization of fluorescence detection extends far beyond the MHC multimer assay. For all flow cytometry experiments, optimal instrument performance and detection of each individual fluorochrome is crucial; even more so with increasing complexity of polychromatic flow cytometry.

Pertaining to the matter of increasingly high dimensionality of flow cytometry experiments is the issue of data analysis. To enable analysis of complex data and in order to increase the standardization and reproducibility, a great number of automated tools for analysis of flow cytometry data have been developed. **In paper II**, three software tools, FLOCK, ReFlow and SWIFT, are compared in their ability to automatically identify MHC multimer positive T cells specific for various virus epitopes. We show the feasibility of using these tools in the analysis of MHC multimer positive T cells and that for low frequency populations, variation in results obtained from different labs was reduced when analyzed with the SWIFT algorithm.

Furthermore, the intention of the study was to investigate the use of automated gating tools by a non-computational expert. Although this was successful in the end, there were multiple challenges on the way and it would not have been possible without collaboration with and help from the software developers. This is a huge obstacle for the integration of automated gating tools into the general flow cytometry community. It was especially true for SWIFT which was the only tool that required coding in MatLab, while FLOCK and ReFlow presented with user friendly interfaces available online. A major issue in using automated gating tools, is that troubleshooting is practically impossible to do without the knowledge required to understand how the algorithms work. As previously discussed, the widespread use of automated tools for flow cytometry data analysis has not been as rapid as hoped and probably expected by developers. One of the biggest challenges is the gap between software developers and immunologists. At first, developers were trying to convince immunologists that using these software tools in coding interfaces like R would be an easy task to learn. It did, however, quickly become apparent that the widespread use of these tools is highly dependent on easy accessibility for people without programming skills. As a result, tools are now being integrated into already known analysis platforms. Such an example is the option of using e.g. tSNE, SPADE and flowMeans directly in the analysis program FlowJo. Despite the positive effect this greater accessibility will definitely have on the reach of these tools, there are also a few precautions to be aware of. As of yet, no automated tools provide a completely automated approach. Whether it is in the form of stating the number of expected populations or interpreting output results, human input is still needed. Given the fact that different tools are developed to offer solutions to different problems in data analysis, getting a meaningful output

from the automated analysis, might very well depend on the understanding of the basic principle of the tool utilized. An example is FLOCK, which was designed to analyze high dimensional flow cytometry data by a density-based clustering method (Qian *et al.,* 2010). The specific design of this algorithm might make it especially good at unraveling diverse populations in a complex data set, but this may come at the price of poorer performance in other tasks, such as detecting rare events. This is consistent with the result we observed in paper II where FLOCK was not able to detect low frequency CD8 T cell responses below 0.1% of total CD8 T cells. Thus, if looking for low frequency populations, FLOCK may not be the best choice of analysis tool even though it might be an excellent choice for a more descriptive exploration of high-dimensional data.

The danger of the easy accessibility of automated gating tools is thus, that researchers use them without a proper introduction and understanding of what their limitations are, risking applying them to purposes they were not intended for and misinterpreting output results. Thus, there is no way around spending time and effort on getting familiar with the background of the various tools. Here, the gap between software developers and immunologists again poses a challenge. Many tools are described in a language intended for computational experts and it can be difficult as an immunologist to fully understand the description of the properties of a given tool, offered in unfamiliar terms. Luckily a few papers address this issue and provide an overview of definitions in the programming field as well as strengths and weaknesses of different categories of algorithms (Kvistborg, Gouttefangeas, *et al.,* 2015; Mair *et al.,* 2016; Saeys *et al.,* 2016). Although these are very useful for an overview of algorithms they do not provide in-depth understanding of all tools available, but rather introduce a small selection.

As described in this thesis, the complexity of flow cytometry experiments is constantly increasing and has now reached a level where conventional manual analysis is no longer feasible. Thus, as high-dimensional experiments become the standard, including automated gating tools in the experiment pipeline is going to be inevitable. There simply is no way around it, and sooner or later, immunologists will have to invest in getting familiar with the automated tools available. It would be a great help in this process if software developers could offer short descriptions of their tools with strengths and weaknesses, offered in a language easily understood by immunologists. However, the best way to close the gap between software developers and immunologists might be inter-disciplinary collaborations and education of new researchers with understanding of both worlds

In manuscript III, we investigated the presence of auto reactive CD8 T cells with relevance for narcolepsy type 1. Peptides from 7 different proteins expressed by hypocretin neurons with predicted binding affinity to 8 different HLA alleles, HLA-A*02:01, 03:01, 11:01, B*07:02, 18:01, 35:01, 51:01 and HLA-C*04:01. The peptide prediction resulted in a library of 1183 peptides, which were screened using DNA barcode-labeled MHC multimers. 19 out of 20 NT1 patients were positive for the NT1 associated HLA class II allele DQB1*06:02 whereas this was the case in 23 of the 52 controls.

We found a broad presence of epitope specific CD8 T cells in both NT1 patients and healthy controls. Although no difference in the number of responses between patients and controls was observed, we did find that responses detected in controls positive for HLA-DQB1*06:02 had a significantly lower estimated frequency than those detected in both patients and DQB1*06:02 negative controls. It could be speculated, that healthy individuals who carry the risk allele have avoided development of NT1 due to low frequency of hypocretin neuron-specific CD8 T cells. The equally higher level of responses in patients and DQB1*06:02 negative controls would then suggest that these controls are protected by the lack of the risk allele and could therefore harbor the same level of auto-reactive T cells as patients without developing disease. Within the cohort of donors which express any of the NT1 associated HLA alleles A*11:01, B*18:01, B*35:01, B*51:01 and C*04:01, there was again a difference between HLA-DQB1*06:02 positive controls and patients. In the patient group, significantly more responses were observed compared to the DQB1*06:02 positive controls, with 7/9 patients having more than one response whereas this was only true for 1/6 of the controls. Again this points to some difference in the CD8 T cell reactivity between the two donor groups, indicating that the combined effect of a certain level of CD8 T cells response and expression of HLA-DQB1*06:02 could be important for NT1 pathogenesis. It is possible that hypocretin neuron specific CD8 T cells are widely present in healthy individuals, but disease initiation requires the help of CD4 T cells, which so far have been found almost exclusively in NT1 patients (Latorre *et al.*, 2018). As reviewed in (Laidlaw *et al.*, 2016), CD4 T cell help to CD8 T cells is necessary in both the primary and secondary response to an infection and it has also been shown that CD4 T cell signaling to exhausted CD8 T cells correlates negatively with disease outcome in autoimmune diseases (McKinney *et al.*, 2015).

The recent study by Latorre et al., which is discussed in the introduction to this thesis, found substantial CD4 T cell reactivity towards the neuropeptide preprohypocretin (HCRT). Proliferation assays were used to define reactivity against pools of HCRT and TRIB2 peptides and they found that significantly more patients than controls harbored reactive CD4 T cells against HCRT but not TRIB2. TRIB2 reactive cells were detected in both patients and controls, but the responses in patients were of significantly higher magnitude than those of the controls. Latorre et al. however, did not detect significant CD8 T cell responses against HCRT or TRIB2. In our study, we detect multiple CD8 T cell responses across the different proteins included, but only a single patient had a response towards a HCRT peptide. Since the Latorre et al. study only included HCRT and TRIB2 proteins, comparisons between our studies in terms of number of responses is not meaningful. It would have been interesting to see whether CD4 T cell responses would also be present for any of the proteins, HCRTR2, LHX9, PDYN, QRFP and RFX4, that we detected CD8 T cells responses against. It cannot be expected that the same amount of CD8 T cell responses would have been found with the method used by Latorre et al. and ours, had the same protein targets been investigated. Whereas Latorre et al. would detect only CD8 T cells with a functional capacity to proliferate, we detect in our study, any CD8 T cell with the appropriate TCR regardless of whether these T cells were functional or not.

The CD8 T cells we detected could thus be self-reactive cells suppressed or induced to be anergic by peripheral tolerance mechanisms. Due to limited donor material, only few proliferation analyses were carried out in our study. Two attempts were made to specifically enrich cells from a few patients and controls with the specific peptide to which an MHC multimer response was found in the given donors. These attempts were unsuccessful for expansion of NT1-specific T cells, but demonstrated very clear expansions of control virus responses from these donors. This finding indicates that the neuron specific CD8 T cells are anergic or suppressed, although this would have to be tested on a larger cohort of donors.

The unsuccessful expansion of auto reactive CD8 T cells in our study may not be so surprising, given the fact that the number of cultures found to expand in the Latorre et al. study were in some cases just a few out of several hundred coming from the same donor. Thus, multiple attempts might be necessary in order to detect these low frequency specific CD8 T cells. No further functionality tests could be done on this patient cohort, thus, the question of functionality in the cells that we detect cannot be answered from our study. Furthermore, it would have been interesting to investigate the CD8 T cell phenotype and TCR clonality, which we did not do in this study. Whether the auto reactive CD8 T cells we detected were naïve or memory cells, and whether they were clonally expanded, might have provided a further distinction between the patient and control cohorts in our study. As indicated by others, self-reactive CD8 T cells detected in both patients and controls, might very well be distinguished by their functional capacity (Berthelot *et al.,* 2008; Maeda *et al.,* 2014) or their ability to home to the relevant tissue (Culina *et al.,* 2018). Both cases could offer plausible explanations to the existence of these cells in the circulation with no consequent pathology in healthy individuals.

In relation to this it is also worth noting, that in this study we only investigate the presence of auto reactive CD8 T cells in the blood of patients and controls. Our results suggest that NT1 relevant auto reactive CD8 T cells circulate the body in a substantial fraction of all healthy individuals, but this picture might have been different had we looked at CSF instead of blood samples. We do not know whether the cells we have detected have the ability to cross the blood-brain-barrier (BBB) and enter the CSF, which would be a prerequisite for their involvement in NT1 development.

The brain is a special organ in terms of access for activated immune cells. As reviewed by Engelhardt et al., naïve T cells do not cross the BBB, whereas activated T cells do. From the blood, activated T cells can cross into various CSF containing spaces, depending on the anatomical location, such as the subarachnoid or perivascular space. In order for a T cell to gain access to the brain parenchyma, where it can potentially cause neuronal damage, re-stimulated with its cognate antigen is required. The subsequent release of pro-inflammatory cytokines alter the characteristics of the final brain barrier, the glia limitants, and T cells as well as other immune cells can enter brain parenchyma (Engelhardt *et al.,* 2017). This is consistent with the observation from the NT1 mouse model, described in the introduction to narcolepsy, that CD4 T cells enter the brain parenchyma only when their antigen is expressed there (Bernard-Valnet *et al.,* 2016). Thus, it cannot be deducted from the presence of auto

reactive T cells in the CSF, that these are also present inside the brain parenchyma. Even so, the composition of auto-reactive T cells in the CSF is likely to be more representative for the situation within the brain as this site could maybe be considered as an intermediate between circulation and brain parenchyma. In the study by Latorre et al., CSF samples were investigated for the presence of the auto reactive TCR clonotypes identified in the blood samples. Within the CD4 T cells no such clones were present. In one case of a patient diagnosed with NT2 however, two auto-reactive TCR clones were detected at a high frequency. This patient was included in the study very close to disease onset and as such may represent another stage of the disease development where neuron destruction is only just beginning, explaining the milder symptoms of NT2 compared to NT1 (Latorre *et al.,* 2018).

Antigen specific T cells are known to be cross reactive, and in a recent study it was demonstrated how amino acid substitutions within a known epitope can affect the affinity of a TCR for its antigen (Bentzen, Such, *et al.,* 2018). While some amino acid substitutions led to the loss of TCR binding, others merely decreased the affinity and some even resulted in an increased affinity between TCR and pMHC compared to the wildtype peptide (figure 5.1).

This goes to show, that even though the CD8 T cells we have detected in our study were binding to peptides from proteins expressed in hypocretin neurons, there is no way of knowing whether these cells were in fact raised against those proteins. It is also possible that the T cell populations we have detected were raised against a foreign pathogen and that the binding of these cells to neuron specific peptides is a matter of cross-reactivity. Even so, these cells may still be important for NT1 pathogenesis. Taken together, the abundance of CD8 T cells specific towards proteins expressed in the brain across all donors in our study, and the knowledge of the inflammation induced changes to the barriers protecting the brain, these observations could suggest that the triggering event for narcolepsy development is infection and subsequent neuroinflammation. Consequently, self-reactive T cells that are either cross reactive between pathogen and hypocretin neurons, or which just happen to be present at the site of inflammation infiltrate the brain and cause disease. As previously mentioned, the hypothesis of an infectious trigger is supported by the observed seasonal pattern of narcolepsy onset (Han *et al.,* 2011) and it would also offer a plausible explanation to the increased incidence of NT1 after pandemrix vaccination.

To further elucidate the pathogenesis of NT1, a number of studies could be interesting to conduct. First of all, investigating the presence of hypocretin neuron specific CD8 T cells in patients with vaccine induced NT1 could maybe point to the mechanism behind this phenomenon, possibly also shedding light on the development of idiopathic NT1. It could also provide insight into the early mechanisms of NT1 development as many of these patients were diagnosed shortly after disease onset which is often not the case with idiopathic narcolepsy (Thorpy & Krieger, 2014). This was also not the case in our study, where the mean disease duration of the included patients was 5.3 years and it is possible that a disease causing autoimmune attack is rapid and not detectable after years of disease.

An interesting addition to the detection of MHC multimer binding CD8 T cells as
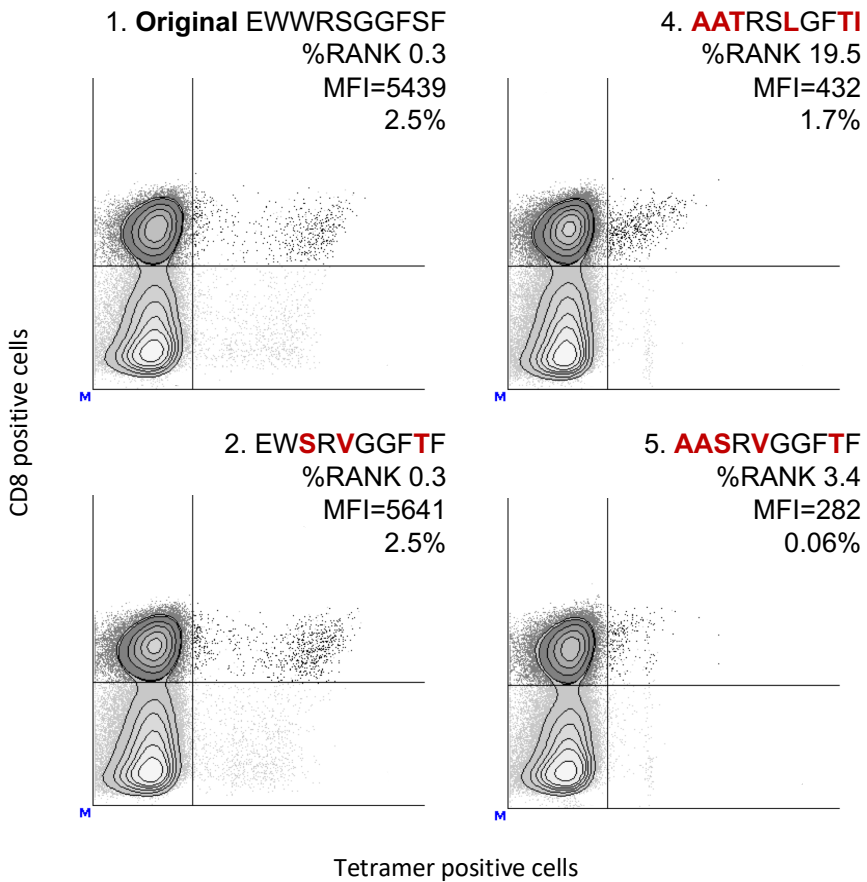
Figure 5.1: The effect of amino acid substitutions within an HLA-A*24:01 restricted peptide on the affinity of the interaction with the TCR recognizing it. MFI: mean fluorescence intensity. % rank refers to how the peptide was ranked in the prediction of binding to the given HLA. Modified from (Bentzen, Such, *et al.*, 2018).

done in our study, could be a transcriptomics analysis. This approach would overcome the challenge of characterizing these cells due to their inability to expand in culture and potentially provide insight into possible differences between auto reactive T cells present in NT1 patients and healthy controls.

As previously mentioned CD8 T cells are not the only possible mediators of narcolepsy development and CD4 T cells as well as autoantibodies are also speculated to play a role. Furthermore, the cells in the brain microenvironment, such as microglia, could be involved as well, if not directly, then by presentation of antigens or promotion of inflammation as they have been found to do in other CNS immune mediated disorders such as multiple sclerosis and Alzheimers disease (Goldmann & Prinz, 2013; Salter & Stevens, 2017). Determining the interaction of all these different cells types is likely necessary in order to dissect their roles in NT1.

The study presented in manuscript III shows for the first time the presence of auto reactive CD8 T cells specific for a range of proteins expressed by hypocretin neurons in the brain. Although this observation is significant and may add to the circumstantial evidence of an autoimmune cause for NT1, it does not provide a clear insight into the pathogenesis of NT1. If and how the specific cells that we have detected are pathogenic can only be speculated and more studies are needed to fully elucidate the involvement of antigen specific CD8 T cells in development of narcolepsy. Identification of the mechanisms underlying narcolepsy disease would enable the development of new therapeutic strategies. Immune therapy for treatment of cancer is under rapid development and the dramatic effect of blocking inhibitory pathways like those of CTLA-4 and PD-1 or adoptively transferring tumor infiltrating lymphocytes has been shown in recent years (Ribas & Wolchok, 2018; Rosenberg & Restifo, 2015). It is an obvious thought that these immune therapy strategies could be applied to autoimmune diseases, among these narcolepsy, only turned upside down. Instead of enhancing self-reactive T cell responses as done in cancer therapy, these strategies could be exploited to dampen a harmful immune response in autoimmunity. Several studies have shown a positive effect of induction of inhibitory signaling in animal models and in clinical trials of patients with autoimmune diseases (Ford *et al.,* 2014). One example is the CTLA-4 fusion protein abadacept, which is approved for treatment of rheumatoid arthritis (Kremer *et al.,* 2003; Genovese *et al.,* 2005). Adoptive transfer of Treg cells is another strategy for controlling auto-reactive T cells which has shown promising results in animal models (S. E. Weber *et al.,* 2006; Tarbell *et al.,* 2007) as well as in clinical trials, although there are still challenges to this treatment (Gliwiński *et al.,* 2017). The use of immune therapy for treatment of narcolepsy would require a rapid diagnosis and administration of therapy in order to terminate the autoimmune killing of hypocretin neurons that is speculated but still not proven to cause the debilitating disease narcolepsy type 1.

Abbreviations

APC	Allophycocyanin
APC	Antigen presenting cell
Aire	The autoimmune regulator
BBB	Blood-brain-barrier
CDR	Complementarity-Determining Regions
CTLA-4	Cytotoxic T-lymphocyte antigen 4
CSF	Cerebrospinal fluid
CyTOF	Cytometry by Time-of-Flight
FACS	Fluorescence-Activated Cell Sorting
fezf2	Forebrain-expressed zinc finger 2
FOXP3	Forkhead box P3
GWAS	Genome wide association studies
HCRT	prepro-hypocretin
HCRTR1	Hypocretin receptor 1
HCRTR2	Hypocretin receptor 2
HLA	Human Leukocyte Antigen
LCMV	Lymphocytic choriomeningitis virus
LHX9	LIM Homeobox 9
MHC	Major Histocompatibility Complex
mTEC	Medullary Thymic Epithelial Cell
NT1	Narcolepsy type 1
NT2	Narcolepsy type 2
PD1	Programmed death 1
PDYN	Prodynorphin
PE	Phycoerythrin
pMHC	peptide-MHC complex
Qdots	Quantum Dots
QRFP	Pyroglutamylated RF amide peptide
REM	Rapid eye movement
RFX4	Regulatory factor 4
SI	Staining index
TCR	T cell receptor
TD1	Type 1 diabetes
TRA	Tissue Restricted Antigens
Treg	Regulatory T cells
TRIB2	Tribbles homolog 2
UMI	Unique Molecular Identifier

Bibliography

1. AC, F. D. *et al.* Mass cytometry analysis reveals hyperactive NF Kappa B signaling in myelofibrosis and secondary acute myeloid leukemia. *Leukemia* **31,** 1962–1974 (2017).

2. Aghaeepour, N., Chattopadhyay, P. K., *et al.* Early immunologic correlates of HIV protection can be identified from computational analysis of complex multivariate T-cell flow cytometry assays. *Bioinformatics* **28,** 1009–1016 (2012).

3. Aghaeepour, N., Finak, G., *et al.* Critical assessment of automated flow cytometry data analysis techniques. *Nature Methods* **10,** 228–238 (2013).

4. Aghaeepour, N., Jalali, A., *et al.* RchyOptimyx: Cellular Hierarchy Optimization for Flow Cytometry. *Cytometry Part A* **81,** 1022–1030 (2012).

5. Aghaeepour, N., Nikolic, R., Hoos, H. & Brinkman, R. R. Rapid Cell Populations Identification in Flow Cytometry Data. *Cytometry Part A* **79,** 6–13 (2011).

6. Ahmed, S. S. *et al.* Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2. *Science translational medicine* **7** (2015).

7. Alcántara-Hernández, M. *et al.* High-Dimensional Phenotypic Mapping of Human Dendritic Cells Reveals Interindividual Variation and Tissue Specialization. *Immunity* **47,** 1037–1050 (2017).

8. Aleksic, M. *et al.* Different affinity windows for virus and cancer-specific T-cell receptors: Implications for therapeutic strategies. *European Journal of Immunology* **42,** 3174–3179 (2012).

9. Altman, J. D. *et al.* Phenotypic Analysis of Antigen-Specific T Lymphocytes. *Science* **274,** 94–96 (1996).

10. Amir, E.-a. D. *et al.* viSNE enables visualization of high dimensional single-cell data and reveals phevnotypic heterogeneity of leukemiia. *Nature Biotechnology* **31,** 545–552 (2013).

11. Andersen, R. S., Kvistborg, P., *et al.* Parallel detection of antigen-specific t cell responses by combinatorial encoding of MHC multimers. *Nature Protocols* **7,** 891–902 (2012).

12. Andersen, R. S., Thrue, C. A., *et al.* Dissection of T-cell antigen specificity in human melanoma. *Cancer Research* **72,** 1642–1650 (2012).

13.   Appay, V. *et al.* Memory CD8+T cells vary in differentiation phenotype in different persistent virus infections. *Nature Medicine* **8,** 379–385 (2002).

14.   Arstila, T. P. *et al.* A direct estimate of the human alpha beta T cell receptor diversity. *Science (New York, N.Y.)* **286,** 958–961 (1999).

15.   Azukizawa, H. *et al.* Induction of T-cell-mediated skin disease specific for antigen transgenically expressed in keratinocytes. *European Journal of Immunology* **33,** 1879–1888 (2003).

16.   Bandura, D. R. *et al.* Mass Cytometry : Technique for Real Time Single Cell Multitarget Immunoassay Based on Inductively Coupled Plasma Time-of-Flight Mass Spectrometry. *Anal. Chem* **81,** 6813–6822 (2009).

17.   Barateau, L., Liblau, R., Peyron, C. & Dauvilliers, Y. Narcolepsy Type 1 as an Autoimmune Disorder: Evidence, and Implications for Pharmacological Treatment. *CNS Drugs* **31,** 821–834 (2017).

18.   Bashashati, A. & Brinkman, R. R. A Survey of Flow Cytometry Data Analysis Methods. *Advances in Bioinformatics,* 1–19 (2009).

19.   Becher, B. *et al.* High-dimensional analysis of the murine myeloid cell system. *Nature immunology* **15,** 1181–1189 (2014).

20.   Bendall, S. C. *et al.* Single-cell Mass Cytometry of Differential Immune and Drug responses Across a Human Hematopoietic Continuum. *Science* **332,** 687–696 (2012).

21.   Bentzen, A. K., Such, L., *et al.* T cell receptor fingerprinting enables in-depth characterization of the patterns that govern peptide-MHC recognition. *Nature Biotechnology. Accepted for publlication* (2018).

22.   Bentzen, A. K., Marquard, A. M., *et al.* Large-scale detection of antigen-specific T cells using peptide-MHC-I multimers labeled with DNA barcodes. *Nature Biotechnology* **34,** 1037–1045 (2016).

23.   Bernard-Valnet, R. *et al.* CD8 T cell-mediated killing of orexinergic neurons induces a narcolepsy-like phenotype in mice. *Proceedings of the National Academy of Sciences* **113,** 10956–10961 (2016).

24.   Berthelot, L. *et al.* Blood CD8+ T cell responses against myelin determinants in multiple sclerosis and healthy individuals. *European Journal of Immunology* **38,** 1889–1899 (2008).

25.   Bianchi, T. *et al.* Maintenance of Peripheral Tolerance through Controlled Tissue Homing of Antigen-Specific T Cells in K14-mOVA Mice. *The Journal of Immunology* **182,** 4665–4674 (2009).

26.   Bruggner, R. V., Bodenmiller, B., Dill, D. L., Tibshirani, R. J. & Nolan, G. P. Automated identification of stratifying signatures in cellular subpopulations. *Proceedings of the National Academy of Sciences* **111,** E2770–E2777 (2014).

27. Buus, S. *et al.* Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. *Tissue Antigens* **62,** 378–384 (2003).

28. Cao, Y. *et al.* Functional inflammatory profiles distinguish myelin-reactive T cells from patients with multiple sclerosis. *Sci Transl Med* **7,** 287ra74 (2015).

29. Chakravarty, S. D., Zabriskie, J. B. & Gibofsky, A. Acute rheumatic fever and streptococci: The quintessential pathogenic trigger of autoimmunity. *Clinical Rheumatology* **33,** 893–901 (2014).

30. Chattopadhyay, P. K., Giles, A. & Nettey, L. Precision Immune Monitoring: New 30-parameter Flow Cytometry Uniquely and Comprehensively Defines Immune Checkpoint Expression and T Cell Phenotypes in Tumor Tissue. *Society for Immunotherapy of Cancer, meeting abstract* (2017).

31. Chattopadhyay, P. K. & Roederer, M. Cytometry: Today's technology and tomorrow's horizons. *Methods* **57,** 251–258 (2012).

32. Chemelli, R. M. *et al.* Narcolepsy in orexin knockout mice: Molecular genetics of sleep regulation. *Cell* **98,** 437–451 (1999).

33. Chen, A. *et al.* QRFP and Its Receptors Regulate Locomotor Activity and Sleep in Zebrafish. *Journal of Neuroscience* **36,** 1823–1840 (2016).

34. Chen, W. *et al.* Conversion of Peripheral CD4+ CD25− Naive T Cells to CD4+CD25+ Regulatory T Cells by TGF-β Induction of Transcription Factor Foxp3. *The Journal of Experimental Medicine* **198,** 1875–1886 (2003).

35. Cheng, M. & Anderson, M. S. Thymic tolerance as a key brake on autoimmunity. *Nature Immunology* **19,** 659–664 (2018).

36. Chou, T. C. *et al.* Orexin (hypocretin) neurons contain dynorphin. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **21,** RC168 (2001).

37. Christen, U., Bender, C. & Von Herrath, M. Infection as a cause of type 1 diabetes? *Current Opinion in Rheumatology* **24,** 417–423 (2016).

38. Cossarizza, A. *et al.* Guidelines for the use of flow cytometry and cell sorting in immunological studies. *European Journal of Immunology* **47,** 1584–1797 (2017).

39. Crocker, A. *et al.* Concomitant loss of dynorphin, NARP, and orexin in narcolepsy. *Neurology* **65,** 1184–1188 (2005).

40. Culina, S. *et al.* Islet-reactive CD8 + T cell frequencies in the pancreas, but not in blood, distinguish type 1 diabetic patients from healthy donors. *Science Immunology* **3,** 1–15 (2018).

41. Cvetkovic-lopes, V. *et al.* Elevated Tribbles homolog 2-specific antibody levels in narcolepsy patients. *The Journal of Clinical Investigation* **120,** 713–719 (2010).

42. Dalal, J. *et al.* Translational profiling of hypocretin neurons identifies candidate molecules for sleep regulation. *Genes and Development* **27,** 565–578 (2013).

43. Daniels, M. A. *et al.* CD8 binding to MHC class I molecules is influenced by T cell maturation and glycosylation. *Immunity* **15,** 1051–1061 (2001).

44. Dauvilliers, Y., Arnulf, I., *et al.* Increased risk of narcolepsy in children and adults after pandemic H1N1 vaccination in France. *Brain* **136,** 2486–2496 (2013).

45. Dauvilliers, Y., Bauer, J., *et al.* Hypothalamic immunopathology in anti-ma-associated diencephalitis with narcolepsy-cataplexy. *JAMA Neurology* **70,** 1305–1310 (2013).

46. Davis, M. M., Altman, J. D. & Newell, E. W. Interrogating the repertoire: broadening the scope of peptide-MHC multimer analysis. *Nature Reviews Immunology* **11,** 551–558 (2011).

47. Derbinski, J., Schulte, A., Kyewski, B. & Klein, L. Promiscuous gene expression in medullary thymic epithelial cells mirrors the peripheral self. *Nature Immunology* **2,** 1032–1039 (2001).

48. Ehlers, M. R. Who let the dogs out? The ever-present threat of autoreactive T cells. *Science immunology* **3,** eaar6602 (2018).

49. Engelhardt, B., Vajkoczy, P. & Weller, R. O. The movers and shapers in immune privilege of the CNS. *Nature Immunology* **18,** 123–131 (2017).

50. Fagerberg, L. *et al.* Analysis of the Human Tissue-specific Expression by Genome-wide Integration of Transcriptomics and Antibody-based Proteomics. *Molecular & Cellular Proteomics* **13,** 397–406 (2014).

51. Ford, M. L., Adams, A. B. & Pearson, T. C. Targeting co-stimulatory pathways: Transplantation and autoimmunity. *Nature Reviews Nephrology* **10,** 14–24 (2014).

52. Gaudilliere, B. *et al.* Coordinated Surgical Immune Signatures Contain Correlates of Clinical Recovery. *Science translational medicine* **6** (2014).

53. Genovese, M. C. *et al.* Abatacept for Rheumatoid Arthritis Refractory to Tumor Necrosis Factor α Inhibition. *New England Journal of Medicine* **353,** 1114–1123 (2005).

54. Giannoccaro, M. P. *et al.* Antibodies against hypocretin receptor 2 are rare in narcolepsy. *Sleep* **40,** 1–6 (2017).

55. Gliwiński, M., Iwaszkiewicz-Grześ, D. & Trzonkowski, P. Cell-Based Therapies with T Regulatory Cells. *BioDrugs* **31,** 335–347 (2017).

56. Goldmann, T. & Prinz, M. Role of Microglia in CNS Autoimmunity. *Clinical and Developmental Immunology* **2013,** 1–8 (2013).

57. Gouttefangeas, C. *et al.* Data analysis as a source of variability of the HLA-peptide mutlimer assay: from manual gating to automated recognition of cell clusters. *Cancer Immunology, Immunotherapy* **64,** 585–598 (2015).

58. Hadrup, S. R., Bakker, A. H., *et al.* Parallel detection of antigen-specific T-cell responses by multidimensional encoding of MHC multimers. *Nat.Methods* **6,** 520–526 (2009).

59. Hadrup, S. R. & Newell, E. W. Determining T-cell specificity to understand and treat disease. *Nature Biomedical Engineering* **1,** 784–795 (2017).

60. Han, F. *et al.* Narcolepsy onset is seasonal and increased following the 2009 H1N1 pandemic in china. *Annals of Neurology* **70,** 410–417 (2011).

61. Hansmann, L. *et al.* Mass Cytometry Analysis Shows That a Novel Memory Phenotype B Cell Is Expanded in Multiple Myeloma. *Cancer Immunology Research* **3,** 650–660 (2015).

62. Hartmann, F. J. *et al.* High-dimensional single-cell analysis reveals the immune signature of narcolepsy. *The Journal of Experimental Medicine* **213,** 2621–2633 (2016).

63. Heier, M. S. *et al.* Incidence of narcolepsy in Norwegian children and adolescents after vaccination against H1N1 influenza A. *Sleep Medicine* **14,** 867–871 (2013).

64. Hollister, R. D., Xia, M., McNamara, M. & Hyman, B. T. Neuronal Expression of Class II Major Histocompatibility Complex (HLA-DR) in 2 Cases of Pick Disease. *Arch Neurol* **54,** 243–248 (1997).

65. Jurtz, V. *et al.* NetMHCpan-4.0: Improved Peptide–MHC Class I Interaction Predictions Integrating Eluted Ligand and Peptide Binding Affinity Data. *The Journal of Immunology,* ji1700893 (2017).

66. Karosiene, E., Lundegaard, C., Lund, O. & Nielsen, M. NetMHCcons: A consensus method for the major histocompatibility complex class i predictions. *Immunogenetics* **64,** 177–186 (2012).

67. Klein, L., Kyewski, B., Allen, P. M. & Hogquist, K. A. Positive and negative selection of the T cell repertoire: What thymocytes see (and don't see). *Nature Reviews Immunology* **14,** 377–391 (2014).

68. Kornum, B. R., Knudsen, S., *et al.* Narcolepsy. *Nature Reviews Disease Primers* **3** (2017).

69. Kornum, B. R., Pizza, F., *et al.* Cerebrospinal fluid cytokine levels in type 1 narcolepsy patients very close to onset. *Brain, Behavior, and Immunity* **49,** 54–58 (2015).

70. Kremer, J. M. *et al.* Treatment of Rheumatoid Arthritis by Selective Inhibition of T-Cell Activation with Fusion Protein CTLA4Ig. *New England Journal of Medicine* **349,** 1907–1915 (2003).

71. Kvistborg, P. *The many colors of tumor-specific T cells* 2018. <https://www.flowcytometry-news.com/category/the-cytohub/2018/07/09/podcast-the-many-colors-of-tumor-specific-t-cells-pia-kvisborg-netherlands-cancer-institute/>.

72. Kvistborg, P., Gouttefangeas, C., *et al.* Thinking Outside the Gate: Single-Cell Assessments in Multiple Dimensions. *Immunity* **42,** 591–592 (2015).

73. Kvistborg, P., Philips, D., *et al.* Anti-CTLA-4 therapy broadens the melanoma-reactive CD8+ T cell response. *Science translational medicine* **6,** 1–9 (2014).

74. Laidlaw, B. J., Craft, J. E. & Kaech, S. M. The multifaceted role of CD4+ T cells in CD8+ T cell memory. *Nature Reviews Immunology* **16,** 102–111 (2016).

75. Latorre, D. *et al.* T cells in patients with narcolepsy target self-antigens of hypocretin neurons. *Nature* (2018).

76. Lee, P. P. *et al.* Characterization of circulating T cells specific for tumor-associated antigens in melanoma patients. *Nature medicine* **5,** 677–685 (1999).

77. Liblau, R. S., Vassalli, A., Seifinejad, A. & Tafti, M. Hypocretin (orexin) biology and the pathophysiology of narcolepsy with cataplexy. *The Lancet Neurology* **14,** 318–328 (2015).

78. Lin, L., Frelinger, J., *et al.* Identification and visualization of multidimensional antigen-specific T-cell populations in polychromatic cytometry data. *Cytometry Part A* **87,** 675–682 (2015).

79. Lin, L., Faraco, J., *et al.* The sleep disorder canine narcolepsy is caused by a mutation in the hypocretin (orexin) receptor 2 gene. *Cell* **98,** 365–376 (1999).

80. Liu, J. *et al.* Evolutionarily conserved regulation of hypocretin neuron specification by Lhx9. *Development* **142,** 1113–1124 (2015).

81. Lugli, E., Roederer, M. & Cossarizza, A. Data Analysis in Flow Cytometry: The Future Just Started. *Cytometry A* **77,** 705–713 (2010).

82. Maecker, H. T., McCoy, J. P. & Nussenblatt, R. Standardizing immunophenotypes for the Human Immunology Project. *Nature Reviews Immunology* **12,** 191–200 (2012).

83. Maecker, H. T., Rinfret, A., *et al.* Standardization of cytokine flow cytometry assays. *BMC Immunology* **6,** 1–18 (2005).

84. Maeda, Y. *et al.* Detection of self-reactive CD8+T cells with an anergic phenotype in healthy individuals. *Science* **346,** 1536–1540 (2014).

85. Mair, F. *et al.* The end of gating? An introduction to automated analysis of high dimensional cytometry data. *European Journal of Immunology* **46,** 34–43 (2016).

86. Marcus, J. N. *et al.* Differential expression of orexin receptors 1 and 2 in the rat brain. *The Journal of comparative neurology* **435,** 6–25 (2001).

87. Martin, R. *et al.* Fine specificity and HLA restriction of myelin basic protein-specific cytotoxic T cell lines from multiple sclerosis patients and healthy individuals. *J Immunol* **145,** 540–548 (1990).

88. Mc Granahan, N. *et al.* Clonal neoantigens elict T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science* **351,** 1463–1469 (2016).

89. McKinney, E. F., Lee, J. C., Jayne, D. R., Lyons, P. A. & Smith, K. G. T-cell exhaustion, co-stimulation and clinical outcome in autoimmunity and infection. *Nature* **523,** 612–616 (2015).

90. Mignot, E., Grumet, F. C. & Guilleminault, C. HLA DQB1*0602 is Associated With Cataplexy in 509 Narcoleptic Patients. *Sleep* **20,** 1012–1020 (1997).

91. Mohammed, F. *et al.* Phosphorylation-dependent interaction between antigenic peptides and MHC class I: A molecular basis for the presentation of transformed self. *Nature Immunology* **9,** 1236–1243 (2008).

92. Mosmann, T. R. *et al.* SWIFT-Scalable clustering for automated identification of rare cell populations in large, high-dimensional flow cytometry datasets, Part 2: Biological evaluation. *Cytometry Part A* **85,** 422–433 (2014).

93. Mueller, D. L. Mechanisms maintaining peripheral tolerance. *Nature Immunology* **11,** 21–27 (2010).

94. Nagamine, K. *et al.* Positional cloning of the APECED gene. *Nature Genetics* **17,** 393–398 (1997).

95. Naim, I. *et al.* SWIFT-scalable clustering for automated identification of rare cell populations in large, high-dimensional flow cytometry datasets, Part 1: Algorithm design. *Cytometry Part A* **85,** 408–421 (2014).

96. Newell, E. W., Klein, L. O., Yu, W. & Davis, M. M. Simultaneous detection of many T-cell specificities using combinatorial tetramer staining. *Nat Methods* **6,** 497–499 (2009).

97. Newell, E. W., Sigal, N., *et al.* Combinatorial tetramer staining and mass cytometry analysis facilitate T-cell epitope mapping and characterization. *Nature Biotechnology* **31,** 623–629 (2013).

98. Nielsen, M. *et al.* NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. *PLoS ONE* **2** (2007).

99. Nishimura, H., Nose, M., Hiai, H., Minato, N. & Honjo, T. Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. *Immunity* **11,** 141–151 (1999).

100. O'Gorman, W. E. *et al.* Mass cytometry identifies a distinct monocyte cytokine signature shared by clinically heterogeneous pediatric SLE patients. *Journal of Autoimmunity* **81,** 74–89 (2017).

101. Ohashi, P. S. *et al.* Ablation of "tolerance" and induction of diabetes by virus infection in viral antigen transgenic mice. *Cell* **65,** 305–317 (1991).

102. Ohkura, N., Kitagawa, Y. & Sakaguchi, S. Development and Maintenance of Regulatory T cells. *Immunity* **38,** 414–423 (2013).

103. Ollila, H. M. *et al.* HLA-DPB1 and HLA class i confer risk of and protection from narcolepsy. *American Journal of Human Genetics* **96,** 136–146 (2015).

104. Omenn, G. S. *et al.* Metrics for the Human Proteome Project 2016: Progress on Identifying and Characterizing the Human Proteome, Including Post-Translational Modifications. *Journal of Proteome Research* **91,** 165–171 (2016).

105. O'Neill, K., Jalali, A., Aghaeepour, N., Hoos, H. & Brinkman, R. R. Enhanced flowType/RcyOptimix: a Bioconductor pipeline for discovery in high-dimensional cytometry data. *Bioinformatics* **30,** 1329–1330 (2014).

106. Pandiyan, P., Zheng, L., Ishihara, S., Reed, J. & Lenardo, M. J. CD4+CD25+Foxp3+ regulatory T cells induce cytokine deprivation–mediated apoptosis of effector CD4+ T cells. *Nature Immunology* **8,** 1353–1362 (2007).

107. Partinen, M., Kornum, B. R., *et al.* Narcolepsy as an autoimmune disease: The role of H1N1 infection and vaccination. *The Lancet Neurology* **13,** 600–613 (2014).

108. Partinen, M., Saarenpää-Heikkilä, O., *et al.* Increased incidence and clinical picture of childhood narcolepsy following the 2009 H1N1 pandemic vaccination campaign in Finland. *PLoS ONE* **7,** 1–9 (2012).

109. Paust, S., Lu, L., McCarty, N. & Cantor, H. Engagement of B7 on effector T cells by regulatory T cells prevents autoimmune disease. *PNAS* **101,** 10398–403 (2004).

110. Pender, M. P., Csurhes, P. A., Burrows, J. M. & Burrows, S. R. Defective T-cell control of Epstein–Barr virus infection in multiple sclerosis. *Clinical & Translational Immunology* **6,** e147 (2017).

111. Perfetto, S. P., Ambrozak, D., Nguyen, R., Chattopadhyay, P. K. & Roederer, M. Quality assurance for polychromatic flow cytometry using a suite of calibration beads. *Nature Protocols* **7,** 2067–2079 (2012).

112. Peyron, C. *et al.* A mutation in a case of early onset narcolepsy and a generalized absence of hypocretin peptides in human narcoleptic brains. *Nature Medicine* **6,** 991–997 (2000).

113. Qian, Y. *et al.* Eluclidation of Seventeen Human Peripheral Blood B cell Subsets and Quantification of the Tetanus Response Using a Density-Based Method for the Automated Identification of Cell Populations in Multidimensional Flow Cytometry Data. *Cytometry Part B: Clinical Cytometry* **78,** 1–22 (2010).

114. Qiu, P. *et al.* Extracting a Cellular Hierarchy from High-dimensional Cytometry Data with SPADE. *Nature Biotechnology* **29,** 886–891 (2011).

115. Ramberger, M. *et al.* CD4 + T-Cell Reactivity to Orexin / Hypocretin in Patients With Narcolepsy Type 1. *Sleep* **40,** 1–9 (2017).

116. Rammensee, H. G., Bachmann, J., Emmerich, N. P. N., Bachor, O. A. & Stevanović, S. SYFPEITHI: Database for MHC ligands and peptide motifs. *Immunogenetics* **50,** 213–219 (1999).

117. Ribas, A. & Wolchok, J. D. Cancer immunotherapy using checkpoint blockade. *Science* **359,** 1350–1355 (2018).

118.   Rizvi, N. A. *et al.* Mutational landscape determines sensitivity to PD-1 block-ade in non–small cell lung cancer. *Science* **348,** 124–129 (2016).

119.   Robinson, J. *et al.* IMGT/HLA and IMGT/MHC: Sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Research* **31,** 311–314 (2003).

120.   Rodenko, B. *et al.* Generation of peptide-MHC class I complexes through UV-mediated ligand exchange. *Nature Protocols* **1,** 1120–1132 (2006).

121.   Romanov, R. A. *et al.* Molecular interrogation of hypothalamic organization reveals distinct dopamine neuronal subtypes. *Nature Neuroscience* **20,** 176–188 (2017).

122.   Rose, N. R. Negative selection, epitope mimicry and autoimmunity. *Current Opinion in Immunology* **49,** 51–55 (2017).

123.   Rosenberg, S. A. & Restifo, N. P. Adoptive cell transfer as personalized im-munotherapy for human cancer. *Science* **348,** 62–68 (2015).

124.   Sabatino, J. *et al.* Detection and Characterization of Novel Myelin-reactive CD8+ T cell Populations in Multiple Sclerosis. *14th International Congress on Neuroimmunology, meeting abstract* (2018).

125.   Saeys, Y., Van Gassen, S. & Lambrecht, B. N. Computational flow cytometry: Helping to make sense of high-dimensional immunology data. *Nature Reviews Immunology* **16,** 449–462 (2016).

126.   Sakaguchi, N. *et al.* Altered thymic T-cell selection due to a mutation of the ZAP70 gene causes autoimmune arthritis in mice. *Nature* **426,** 454–460 (2003).

127.   Sakaguchi, S., Yamaguchi, T., Nomura, T. & Ono, M. Regulatory T Cells and Immune Tolerance. *Cell* **133,** 775–787 (2008).

128.   Salter, M. W. & Stevens, B. Microglia emerge as central players in brain disease. *Nature Medicine* **23,** 1018–1027 (2017).

129.   Sarkanen, T. O., Alakuijala, A. P., Dauvilliers, Y. A. & Partinen, M. M. Inci-dence of narcolepsy after H1N1 influenza and vaccinations: Systematic review and meta-analysis. *Sleep Medicine Reviews* **38,** 177–186 (2018).

130.   Savage, P. A. & Davis, M. M. A Kinetic Window Constricts the T Cell Receptor Repertoire in the Thymus. *Immunity* **14,** 243–252 (2001).

131.   Scally, S. W. *et al.* A molecular basis for the association of the HLA-DRB1 locus, citrullination, and rheumatoid arthritis. *The Journal of Experimental Medicine* **210,** 2569–2582 (2013).

132.   Seifinejad, A. & Tafti, M. Molecular profiling of the lateral hypothalamic neu-rons identified molecules for disease and development. *World Sleep abstract* (2017).

133.   Sewell, A. K. Why must T cells be cross-reactive? *Nature Reviews Immunology* **12,** 669–677 (2012).

134. Shortman, K., Vremec, D. & Egerton, M. The kinetics of T cell antigen receptor expression by subgroups of CD4+8+ thymocytes: delineation of CD4+8+3(2+) thymocytes as post-selection intermediates leading to mature T cells. *The Journal of experimental medicine* **173,** 323–332 (1991).

135. Sidney, J., Peters, B., Frahm, N., Brander, C. & Sette, A. HLA class I supertypes: A revised and updated classification. *BMC Immunology* **9,** 1–15 (2008).

136. Snir, O. *et al.* Identification and functional characterization of T cells reactive to citrullinated vimentin in HLA-DRB1*0401-positive humanized mice and rheumatoid arthritis patients. *Arthritis and Rheumatism* **63,** 2873–2883 (2011).

137. Spitzer, M. H. & Nolan, G. P. Mass Cytometry: Single Cells, Many Features Matthew. *Cell* **165,** 780–791 (2016).

138. Stritesky, G. L. *et al.* Murine thymic selection quantified using a unique method to capture deleted T cells. *Proceedings of the National Academy of Sciences* **110,** 4679–4684 (2013).

139. Styren, S. D., Civin, W. H. & Rogers, J. Molecular, cellular, and pathologic characterization of HLA-DR immunoreactivity in normal elderly and Alzheimer's disease brain. *Experimental Neurology* **110,** 93–104 (1990).

140. Su, L. F., Kidd, B. A., Han, A., Kotzin, J. J. & Davis, M. M. Virus-Specific CD4+Memory-Phenotype T Cells Are Abundant in Unexposed Adults. *Immunity* **38,** 373–383 (2013).

141. Surh, C. D. & Sprent, J. T-cell apoptosis detected in situ during positive and negative selection in the thymus. *Nature* **372,** 100–103 (1994).

142. Tafti, M., Hor, H., *et al.* DQB1 Locus Alone Explains Most of the Risk and Protection in Narcolepsy with Cataplexy in Europe. *Sleep* **37,** 19–25 (2014).

143. Tafti, M., Lammers, G. J., *et al.* Narcolepsy-Associated HLA Class I Alleles Implicate Cell-Mediated Cytotoxicity. *Sleep* **39,** 581–587 (2016).

144. Takaba, H. *et al.* Fezf2 Orchestrates a Thymic Program of Self-Antigen Expression for Immune Tolerance. *Cell* **163,** 975–987 (2015).

145. Tarbell, K. V. *et al.* Dendritic cell–expanded, islet-specific CD4+ CD25+ CD62L+ regulatory T cells restore normoglycemia in diabetic NOD mice. *The Journal of Experimental Medicine* **204,** 191–201 (2007).

146. Thannickal, T. C. *et al.* Reduced number of hypocretin neurons in human narcolepsy. *Neuron* **27,** 469–474 (2000).

147. Thorpy, M. J. & Krieger, A. C. Delayed diagnosis of narcolepsy: Characterization and impact. *Sleep Medicine* **15,** 502–507 (2014).

148. Tivol, E. A. *et al.* Loss of CTLA-4 leads to massive lymphoproliferation and fatal multiorgan tissue destruction, revealing a critical negative regulatory role of CTLA-4. *Immunity* **3,** 541–547 (1995).

149. Toebes, M. *et al.* Design and use of conditional MHC class I ligands. *Nature Medicine* **12,** 246–251 (2006).

150. Turner, S. J., Doherty, P. C., McCluskey, J. & Rossjohn, J. Structural determinants of T-cell receptor bias in immunity. *Nature Reviews Immunology* **6,** 883–894 (2006).

151. Unger, W. W. *et al.* Discovery of low-affinity preproinsulin epitopes and detection of autoreactive CD8 T-cells using combinatorial MHC multimers. *Journal of Autoimmunity* **37,** 151–159 (2011).

152. Vagaska, B. *et al.* MHC-class-II are expressed in a subpopulation of human neural stem cells in vitro in an IFNγ-independent fashion and during development. *Scientific Reports* **6,** 1–14 (2016).

153. Van der Maaten, L. & Hinton, G. Visualizing Data using t-SNE Visualizing Data using t-SNE. *Journal of Machine Learning Research* **1,** 1–25 (2008).

154. Van Gassen, S. *et al.* FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry Part A* **87,** 636–645 (2015).

155. Vassalli, A., Li, S. & Tafti, M. Comment on "Antibodies to influenza nucleoprotein cross-react with human hypocretin receptor 2". *Science translational medicine* **7,** 2–5 (2015).

156. Velthuis, J. H. *et al.* Simultaneous Detection of Circulating Autoreactive CD8+ T-Cells Specific for Different Islet Cell-Associated Epitopes Using Combinatorial MHC Multimers. *Diabetes* **59,** 1721–1730 (2010).

157. Von Boehmer, H. Mechanisms of suppression by suppressor T cells. *Nature Immunology* **6,** 338–344 (2005).

158. Waterhouse, P. *et al.* Lymphoproliferative Disorders with Early Lethality in Mice Deficient in Ctla-4. *Science* **270,** 985–988 (1995).

159. Weber, L. M. & Robinson, M. D. Comparison of clustering methods for high-dimensional single-cell flow and mass cytometry data. *Cytometry Part A* **89,** 1084–1096 (2016).

160. Weber, S. E. *et al.* Adaptive Islet-Specific Regulatory CD4 T Cells Control Autoimmune Diabetes and Mediate the Disappearance of Pathogenic Th1 Cells In Vivo. *The Journal of Immunology* **176,** 4730–4739 (2006).

161. Welsh, R. M. & Selin, L. K. No one is naive: the significance of heterologous T-cell immunity. *Nature Reviews Immunology* **2,** 417–426 (2002).

162. White, S. *et al.* Managing multi-center flow cytometry data for immune monitoring. *Cancer Informatics* **13,** 111–122 (2014).

163. Wim Ang, C., Jacobs, B. C. & Laman, J. D. The Guillain-Barré syndrome: A true case of molecular mimicry. *Trends in Immunology* **25,** 61–66 (2004).

164. Wooldridge, L. *et al.* A single autoimmune T cell receptor recognizes more than a million different peptides. *Journal of Biological Chemistry* **287,** 1168–1177 (2012).

165. Xu, Q., Schlabach, M. R., Hannon, G. J. & Elledge, S. J. Design of 240,000 orthogonal 25mer DNA barcode probes. *Proceedings of the National Academy of Sciences* **106,** 2289–2294 (2009).

166. Yamanaka, A., Tabuchi, S., Tsunematsu, T., Fukazawa, Y. & Tominaga, M. Orexin Directly Excites Orexin Neurons through Orexin 2 Receptor. *Journal of Neuroscience* **30,** 12642–12652 (2010).

167. Yewdell, J. W. & Bennink, J. R. Immunodominance in Major Histocompatibility Complex Class I–Restricted T Lymphocyte Responses 1. *Annual Review of Immunology* **17,** 51–88 (1999).

168. Yu, W. *et al.* Clonal Deletion Prunes but Does Not Eliminate Self-Specific $\alpha\beta$ CD8+ T Lymphocytes. *Immunity* **42,** 929–941 (2015).

169. Zhao, B., Chang, L., Fu, H., Sun, G. & Yang, W. The Role of Autoimmune Regulator (AIRE) in Peripheral Tolerance. *Journal of immunology research* **2018,** 1–6 (2018).

170. Zhao, Z. S., Granucci, F., Yeh, L., Schaffer, P. A. & Cantor, H. Molecular mimicry by herpes simplex virus type 1: autoimmune disease after viral infection. *Science* **279,** 1344–1347 (1998).