Mechanisms of action of peptide immunotherapy for Type 1 diabetes Characterising T-cells induced or modified by proinsulin C19-A3

Liu, Yuk-Fun

Awarding institution:
King's College London

Mechanisms of action of peptide immunotherapy for Type 1 diabetes: characterising T-cells induced or modified by proinsulin C19-A3

Yuk-Fun Liu

A dissertation submitted to the University of London in candidature of Doctor of Philosophy

Division of Immunology, Infection and Inflammatory Disease
King's College London School of Medicine at Guy's, King's and St. Thomas' Hospitals

Abstract

Peptide immunotherapy (PIT) is a specific immunomodulatory treatment aiming to tip the balance from pro-inflammatory attack to restoration of immune tolerance. It has been successfully translated into clinical practice in the field of allergy, stimulating ongoing research in autoimmune diseases including type 1 diabetes. The safety and mechanistic effects of proinsulin C19-A3 (PI C19-A3), a HLA-DR4 restricted, naturally processed and presented peptide, was assessed through a multicentre placebo controlled double-blind clinical trial of 2 or 4 weekly intradermal doses of 10μg peptide for 6 months, with a 6 month follow-up period. Overall, treatment was well tolerated with no evidence of local or systemic hypersensitivity or disease acceleration. In some PIT-treated subjects there was retention of C-peptide. To address the question of whether and how PIT impacts upon immune function, over the 6 month treatment period, T cell receptor (TCR) clonotyping and gene expression analysis were performed on peptide-specific activated CD4 T-cells. TCR β-chain clonotyping revealed shared β-chain clonotypes between patients; however, these were not apparently linked to peptide immunotherapy. Gene expression changes were explored but no antigen-specific changes related to treatment identified.

This study concludes that peptide immunotherapy using PI C19-A3 is a safe and well-tolerated treatment in newly-diagnosed adults with type 1 diabetes; deployment of novel mechanistic studies to examine alterations in antigen-specific effector CD4 T cells do not reveal any linked changes in TCR or gene expression.

Acknowledgements

Firstly I would like to thank my supervisor Prof. Mark Peakman for his support throughout this PhD. I greatly appreciate his mentorship during my studies and the opportunities he has provided to expand my research experience.

Also I would like to thank my secondary supervisor Dr. Jake Powrie for his clinical support and guidance.

I would like to thank the following people for their help during my studies:

The MonoPepT1De and ADDRESS-2 staff, particularly Prof. Colin Dayan, Rachel Stenson and Laura Adams. I am very grateful for all your hard work in recruiting to the trial and supporting the participants.

All the MonoPepT1De participants for giving up their time to support this research.

To all my DIIID colleagues, who have patiently guided me in new fields and have taught me so much, it has been a pleasure working with you all. In particular Iria Gomez-Tourino, Hina Shariff, Megan Estorninho, Vivienne Gibson, Sefina Arif, Verity Motskin and Norkhairin Yusuf.

Tom Hayday and Richard Ellis for flow cytometry expertise.

Ania Lorenc for the bioinformatic expertise in TCR clonotyping and gene microrray analyses.

Deborah Dunn-Walters and Valerie Corrigall for their guidance through the thesis committee.

Diabetes UK for awarding the Clinical Training Fellowship to fund my role this project and the JDRF for funding of the clinical trial.

To my husband, Neil, and sons, James and Alex. Thank you for your continued love, support and understanding.

Thesis contributions

During the MonoPepT1De trial, I was the principal recruiter of subjects at Guy's (which was the most successful recruitment site numerically), and I was a key part of the trial management group (principally Colin Dayan (CD), Mark Peakman (MP), Rachel Stenson and the PIs from the other recruiting sites) which met monthly by teleconference to guide and oversee conduct of the trial. With MP, CD and Dr Mohammad Ahadj Ali I was part of the writing group for the submitted manuscript (on which I am joint first author) that reports the trial outcomes, and the key findings from this are reported here in Chapter 4 to provide context for my own laboratory studies.

I performed all of the assays described in this thesis myself, except where specified. For clarity I would like to indicate the following:

In Chapter 4: ELISPOT and antigen stimulation assays for the MonoPepT1De trial were performed within the trial laboratory team that included myself, Dr Hina Shariff, Norkhairin Yusuf, Lucas Baumard and Dr Sefina Arif.

In Chapter 5: Dr Hina Shariff and Dr Vivienne Gibson performed flow cytometry experiments to assess proliferation and IFN-γ production of antigen-specific clones. Dr Iria Gomez Tourino performed and provided data on the effects of PBMC cryopreservation on ELISPOT assay.

In Chapter 6: TCR clonotype data was compared to deep sequencing data from experiments performed by Dr Iria Gomez Tourino. Dr Anna Lorenc performed bioinformatic analysis.

In Chapter 7: Dr Anna Lorenc performed bioinformatic analysis.

Contents

List of Figures

List of Tables

Abbreviations

AAb	Autoantibodies
ADDQoL	Audit of Diabetes-Dependent Quality of Life
ADDRESS-2	After Diabetes Diagnosis REsearch Support System
APC	Antigen presenting cell
APL	Altered peptide ligands
ASI	Antigen specific immunotherapy
CFA	Complete Freund's adjuvant
CTLA-4	Cytotoxic T-lymphocyte- associated protein 4
DCCT	Diabetes Control and Complications Trial
DTSQs	Diabetes Treatment Satisfaction Questionnaire scores
DTSQc	Diabetes Treatment Satisfaction Questionnaire, change version
DUK	Diabetes UK
EAE	Experimental autoimmune encephalomyelitis
FACS	Florescence-Activated Cell Sorting
FOX-P3	Fork-head box P3
GAD	Glutamic acid decarboxylase
GAD-alum	Glutamic acid decarboxylase formulated with aluminium hydroxide
HA	Haemagglutinin
HbA1c	Glycated haemoglobin
HLA	Human leucocyte antigen
Hsp60	Heat-shock protein 60

HFS	Hypoglycaemia Fear Survey
IAA	Insulin autoantibodies
IA-2	Insulinoma antigen 2
ICA	Islet cell autoantibodies
IDAA1C	Insulin dose adjusted A1c
IFA	Incomplete Freund's adjuvant
IL2RA	Interleukin-2 receptor alpha chain gene
INS	Insulin gene
IMGT	International Immunogenetics Information System
IMP	Investigational Medicinal Product
IFN-γ	Interferon-gamma
JDRF	Juvenile Diabetes Research Foundation
MBP	Myelin basic protein
MHC	Major Histocompatibility Complex
mIAA	Micro-insulin autoantibodies
MID	Multiplex identifier
MMTT	Mixed meal tolerance test
NOD	Non-obese diabetic
NPPE	Naturally processed and presented epitopes
PBMC	Peripheral blood mononuclear cell
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PD-1	Programmed cell death protein-1

PI C19-A3	Proinsulin C19-A3
PIT	Peptide immunotherapy
PTPN22	Protein tyrosine phosphatase non-receptor type 22 gene
SEB	Staphylococcal Enterotoxin B
SCID	Severe combined immunodeficiency
SI	Stimulation index
SNP	Single nucleotide polymorphism
UCPCR	Urine C-peptide creatinine ratio
T1D	Type 1 diabetes
TCR	T-cell receptor
TGF-β	Transforming growth factor β
Treg	Regulatory T-cells
ZnT8	Zinc Transporter 8

1 Introduction

1.1 Type 1 diabetes

Type 1 diabetes (T1D) is an organ-specific autoimmune disease, which results in pancreatic β-cell destruction rendering most, if not all, individuals dependent on life-long exogenous insulin[1]. Classically T1D has been described as a T-cell mediated disease however the histological findings of pathognomonic insulitic lesion describe the presence of a combination of CD8+, CD4+ T lymphocytes, B lymphocytes and macrophages[2].

Current treatment relies on insulin replacement to restore normoglycaemia but despite great advances in insulin therapy and delivery, patients are still exposed to glucose excursions leading to complications of hyperglycaemia, such as cardiovascular disease, retinopathy, neuropathy and nephropathy, as well as the risks of hypoglycaemia. In addition, as a chronic disease of the young, T1D also carries a substantial psychological burden[3]. Peak age of diagnosis is between 10 and 14 years of age[4], however epidemiological studies have shown that age of onset is getting younger and incidence is rising worldwide[5–9] with annual increases of 3.9% in Europe[10]. There are however a significant number diagnosed as adults, estimated at over 40% of diagnosed as adults aged 31 to 60 years[11]. Although over recent decades life expectancy in Type 1 diabetes (T1D) has significantly improved, lifespan is still reduced compared to diabetes-free individuals[12,13]. Total direct and indirect costs of T1D in the UK in 2010/2011 were estimated to be £1.9bn, rising to £4.2bn by 2035/2036[14].

1.2 Genetics versus environmental determinants of Type 1 diabetes risk

Type 1 diabetes is a complex multifactorial genetic disorder and in recent years newer technologies such as genome-wide association (GWAS) studies have allowed greater elucidation of the network of genes which harbour disease potential, with over a total of forty disease-linked loci now identified[15]. Unsurprisingly genes directly linked to the immune system have been identified as key players, with the human leucocyte antigen (HLA) region on chromosome 6p21 the greatest risk factor. Predisposing *HLA* class II haplotypes *DRB1*04-DQB1*0302* and *DRB1*03* are found to be present in at least 90% of paediatric cases[16]. The *HLA* region encodes for the major histocompatibility complex (MHC) suggesting antigen presentation has an important role in the aetiology of T1D.

Single nucleotide polymorphisms (SNPs) of major non-HLA encoding genes also have an associated risk to T1D although this risk is far below MHC class II alleles. These SNPs include variants in the insulin (*INS*) gene, protein tyrosine phosphatase non-receptor type 22 (*PTPN22*), cytotoxic T-lymphocyte-associated protein 4 (*CTLA-4*) and the IL-2 receptor alpha chain (*IL2RA*)[15]. The identification of these SNPs gives clues to the pathogenesis of T1D being closely linked to T-cell signalling with insulin being a major autoantigen, *PTPN22* involved in negative regulation of TCR signalling, *CTLA-4* a co-stimulatory molecule which down-regulates T-cell activation and *IL-2RA* being a T-cell signalling receptor.

There is a lack of concordance for T1D in monozygotic twins, with a 39% probability at 40 years for a discordant monozygotic twin[17]. This highlights the influence of environmental factors in risk of T1D. Within the increasing incidence of new T1D cases there is a trend towards disease presenting in less high risk *DR3* and *DR4* genotypes[18–21]and even within protective genotypes[18], suggesting that the rise in incidence is no longer reliant on a genetic predisposition but increasingly swayed by environmental pressures.

Preclinical signs of autoimmunity appear early in life with the presence of autoantibodies (AAb) detectable in the high risk from less than 6 months of age[22]. This would suggest the importance of prenatal or early infancy environmental factors. Maternal age and heavier baby weight have been associated with T1D risk[23], with increasing weight of children a defence for the "accelerator hypothesis" which suggests the increasing demands on β-cells accelerates their decline[24].

A meta analysis of risk of T1D in children delivered by Caesarean section showed an approximate 20% increase[25]. The gut microbiome is important in shaping a developing immune system and Caesarean section appears to reduce the diversity of gut microbiota compared to vaginal birth in the early years of life[26]. Some have linked this to the "hygiene hypothesis"[27], suggesting reduced exposure to infections at a young age leads to immaturity of the immune system and higher risk of immune-modulated disorders. Comparing gut microbiome in children with or without antibody (Ab) positivity no differences were seen in bacterial diversity, microbial composition, or single genus abundances but substantial alterations in microbial interaction networks occurred up to 2 years of age in those who developed islet autoantibodies[28].

Seasonal variations in diagnosis of T1D are well known[29,30] with a peak of new cases between November and February. This implicates infective causes such as viruses, or vitamin D levels via sunlight exposure, as potential influences. Viral infections, in particular enteroviruses, echo this seasonal variation[31]. The identification of a SNP within an innate immune system receptor, interferon-induced with helicase C domain 1 (IFIH1), which encodes for a receptor for viral DNA, has focused studies onto viral triggers. Enteroviruses are being studied extensively as a pathogenic trigger of T1D and though two prospective studies failed to link enteroviral infection with T1D[32,33], other studies continue to have intriguing findings such as detection of low-grade viral infection in live islets from newly diagnosed patients[34].

As a modulator of the immune system vitamin D has been linked to autoimmune disease, especially as geographical variations in disease incidence correlate with sunlight exposure. Interestingly a plateauing of the rise in incidence of T1D in Finland after 2003 has coincided with the introduction of vitamin D fortification of dairy products[35] and a meta analysis of observational studies showed a positive effect of vitamin D supplementation[36]; nonetheless major intervention studies have yet to be reported and small scale studies have shown conflicting results[37,38].

Dietary factors in childhood have been implicated as triggers for T1D. Whilst prospective studies of breast-feeding[39] and removal of bovine insulin[40] from cow's milk formula in genetically at-risk infants showed a protective effect on progression of islet autoimmunity, feeding a hydrolysed infant formula in those at genetically high risk did not reduce risk of progression to antibody positivity over 7 years[41]. Further analysis of long-term effects of highly hydrolysed formula are awaited from the Trial to Reduce IDDM in the Genetically at Risk (TRIGR) study which will end its 10 year follow-up period in 2017[42].

Proposed environmental factors remain speculative and require very large scale studies such as the Environmental Determinants of Diabetes in the Young (TEDDY) study[43], which screened 414,714 subjects and is still actively unravelling causative factors.

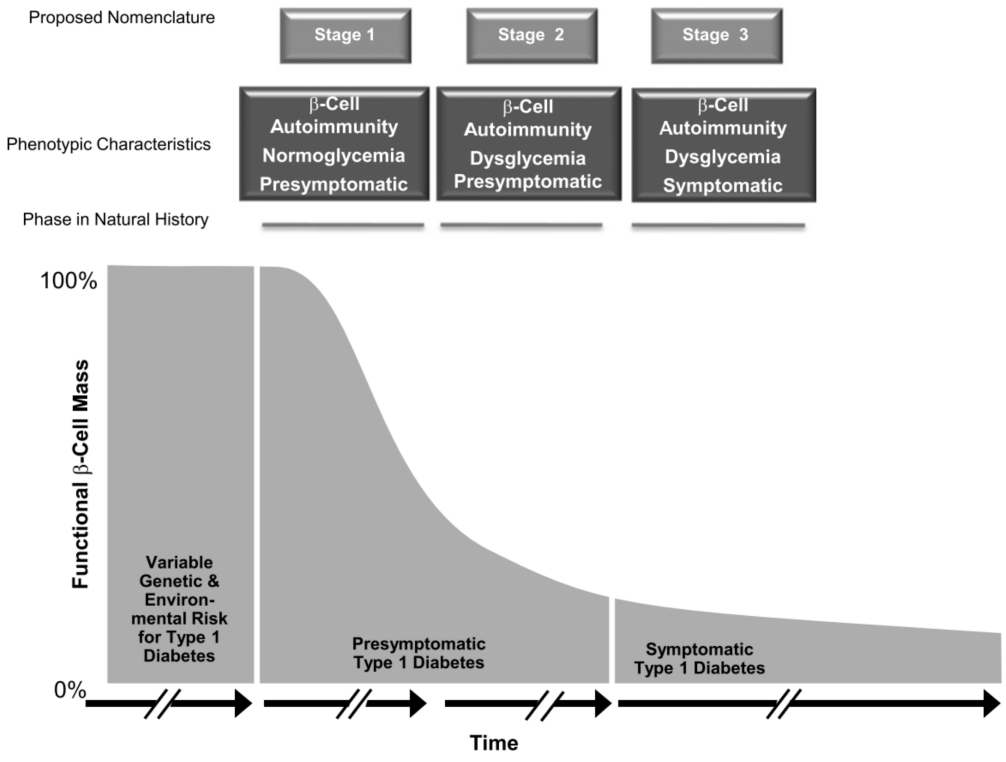
1.3 Natural history of T1D

The putative natural history of T1D was described by Eisenbarth et al.[1] in 1986, suggesting that genetic predisposition with the presence of unknown triggers leads to autoimmunity, detectable by presence of islet autoantibodies and a progressive decline in β-cell function into overt disease. What has become clearer with further natural history studies is that the

period of subclinical autoimmunity and gradient of the β-cell decline varies between individuals, which is illustrated by the effect of age. Recently the JDRF, the Endocrine Society and the American Diabetes Association issued a scientific statement clarifying the presymptomatic stages of T1D[44]. This statement highlights the need to recognise the disease continuum from the presence of two or more islet autoantibodies with normoglycaemia (presymptomatic stage 1), through to the presence of β-cell autoimmunity with dysglycaemia (presymptomatic stage 2), and onset of symptomatic disease (stage 3). Acknowledging that the presence of 2 or more autoantibodies leads to an almost inevitable decline to symptomatic disease, with a greater than 80% risk of developing type 1 diabetes within 15 years[45], allows greater focus on earlier intervention on these high risk groups in which β-cell function has yet to be compromised.

Figure 1: Early stages of type 1 diabetes

(Reproduced from Insel et al. 2015). This diagram describes the presymptomatic continuum of disease with staging beginning with the presence of autoimmunity.



Eisenbarth's original depiction of β-cell mass in T1D, showed a constant decline with the belief that no C-peptide was detectable after a number of years with diabetes. Recent evidence has shown that β-cell function can persist in individuals with T1D for far longer than

previously thought and β-cell decline is a non-linear process. Although at the time of diagnosis it is estimated that 90% of functional β-cell mass has been destroyed[46], the Diabetes Control and Complications Trial (DCCT) established that 20% of patients still had detectable insulin production at 5 years post diagnosis[47]. With newer more sensitive C-peptide assays β-cell function has been detected in 10-68%[48,49] of patients greater than 3 decades after diagnosis with 80% of those with detectable C-peptide showing a physiological rise to a meal stimulus[49]. Those older at diagnosis are more likely to have higher C-peptide levels at diagnosis[50,51] and more likely to sustain substantial β-cell function at 5 years[52], suggesting a more aggressive disease in childhood particularly at ages less than 7[51]. Pozzilli et al.[51] showed that those diagnosed post-puberty had higher C-peptide levels and required less insulin than those diagnosed at a younger age. Residual β-cell function not only appears to be protective against complications [53], but also strengthens the promise of therapies aimed at β-cell salvage. Decline in β-cell function is not necessarily a smooth progression, oral glucose tolerance test (OGTT) data from the Diabetes Prevention Trial-1 (DPT-1) showed a "ratcheting" remitting relapsing of glucose tolerance during the period from 3.8 years prior to diagnosis with some individuals developing transient hyperglycaemia post OGTT, before overt diabetes[54]. This was considered to be due to fluctuations in insulin sensitivity, as glucose fluctuations did not correlate to early C-peptide responses. The decline in stimulated C-peptide is noted to have an accelerated fall in the 3 months prior to symptomatic diagnosis[55], whilst fasting C-peptide are maintained[56]. Although immune interventions would ideally target subjects with maximal residual β-cell function before signs of autoimmunity (primary prevention) or after detection of autoimmune markers such as islet autoantibodies (secondary prevention), this can be difficult in terms of targeting a healthy, often paediatric population for large-scale immune therapy trials. Salvaging β-cell reserve at diagnosis (tertiary prevention) or beyond offers a viable option as residual endogenous secretion can still improve glycaemic control, protecting against complications and hypoglycaemia[57,58].

Study of the subclinical stage can been difficult with large, long-term screening studies required in paediatric cohorts, with the added problem of a lack of reliable biomarkers to track disease progression. A number of longitudinal studies in high-risk groups, such as the TrialNet Natural History Study[59] and the DPT-1[60] have tracked emerging autoimmunity through islet autoantibody development. Islet autoantibodies are the first markers of disease development, developing before the decline in β-cell mass triggers hyperglycaemia. The main islet autoantibodies used in disease prediction are insulin autoantibody (IAA), glutamic acid

decarboxylase antibody (GADA), tyrosine phosphatase-like protein antibody (ICA512A or IA-2A), and zinc transporter 8 antibody (ZnT8A). The most common islet autoantibody to arise in those at risk is GADA and risk of developing autoantibodies (IAA, GADA or IA-2A) declines with age[59]. Tracking the presence of autoantibodies has again highlighted heterogeneity in disease development, for example, IAA positivity is more likely to be present in children under the age of 8 and more likely to progress to multiple antibody positivity, where as GADA positive individuals who progressed to multiple antibodies were more evenly spread across age groups[61]. Such information may be helpful in targeting immunotherapy to those who may be at highest risk of progression, such as young IAA positive individuals.

1.4 Biomarkers of type 1 diabetes

Despite increasing knowledge about the genetics and natural history of T1D, biological markers of the disease are currently limited. Difficult access to the target organ hinders direct study of the immune infiltration within the pancreas at different stages. Islet autoantibodies, genotyping for high-risk alleles and HbA1c remain the mainstay of determination of risk and disease course. Emerging technologies and greater characterisation of the genetic and molecular basis of the disease aim to pinpoint biomarkers to aid diagnosis, prevention, treatment and management.

The function of future T1D biomarkers would include the following:

- **Determination of those at risk and predicting disease onset**

Currently genotyping for high-risk alleles and autoantibody testing are used in prevention trials to identify those at risk, sometimes with the addition of glucose tolerance tests. Participants in the Diabetes Autoimmunity Study of the Young (DAISY) trial were genotyped for 20 non-HLA SNPs in addition to HLA genotypes[62]. From these data they were able to identify 5 high-risk predictors: 1) a *PTPN22* gene polymorphism 2) a *UBASH3* SNP 3) an *INS* gene polymorphism 4) a family history of type 1 diabetes and 5) *HLA-DR3/4*. UBASH belongs to a family of cell regulators linked to down-regulation of receptor-induced activation in T-cells[63]. Combining these risk factors gave hazard ratios up to 16 for the development of islet autoimmunity and >40 for diabetes. Further modelling using microarray gene expression data from peripheral blood RNA in the antibody positive group from DAISY, identified and confirmed four multigene models that consistently stratify high- and low-risk subsets of antibody positive subjects with hazard ratios >6[64]. Each gene model consisted of a

combination of 5 genes, (*BACH2, IGLL3, EIF3A, CDC20, and TXNDC5*) which associated with differential progression and were implicated in lymphocyte activation and function.

In addition to current islet autoantibodies further immune markers of risk would be desirable. A small study investigated serum concentrations of a panel of cytokine and chemokine levels in first-degree relatives of T1D patients by ELISA. Significant correlation was revealed between two closely related chemokines CCL3 and CCL4 in individuals at risk with multiple positive islet autoantibodies (r=0.84, P=0.00005), not present in the autoantibody-negative group[65]. These T helper 1 (T_H1) derived chemokines were also found to be elevated and correlated to β-cell function, in serum samples collected from 256 children and adolescents, studied longitudinally for a year post diagnosis[66]. Other chemokines such as CXCL10 have been implicated in disease with this example being found at elevated levels in pancreatic biopsies of those newly diagnosed[67], as well as systemically in newly diagnosed children[68]. Such chemokine profiles are not antigen-specific and therefore subject to other pathologic influences and have yet to be validated as biomarkers.

Once risk is assessed incremental autoantibody seroconversion is used to track increasing risk of disease development[60], with the presence of two or more autoantibodies indicating over 80% risk of developing T1D within 15 years[45]. To assess pre-clinical decline in β-cell function metabolic markers such as peak C-peptide or MMTT-stimulated area under curve (AUC)[69] and first phase insulin response (FPIR)[70,71] can be assessed via OGTT and intravenous glucose tolerance tests (IVGTT). These tests are cumbersome to use routinely and data from the DPT-1 revealed that IVGTT assessment of FPIR did not appear cost effective[71].

- **Tracking β-cell function during disease**

Guidelines suggest that the current gold standard for tracking β-cell function in intervention trials is C-peptide AUC assessed by a 2 hour mixed meal tolerance test (MMTT)[72]. HbA1c and insulin dose per kg are suggested secondary outcome measures, as disease management can distort these variables. Although measurement of autoantibodies is also suggested in the same guidelines as an outcome, it is acknowledged that these do not necessarily correlate with benefit. Urinary C-peptide is becoming increasingly validated as a surrogate marker of beta-cell function[49,73], which is less invasive and time consuming than venous sampling and therefore attractive potentially for a paediatric population or out-of-hospital sampling.

Again, immune markers would be useful in this setting and in a study using multivariate logistic regression models to predict residual β-cell function up to one year after diagnosis, a meta-immunological profile measuring age, body mass index (BMI), fasting C-peptide, number of circulating CD3+CD16+CD56+ cells and the percentage of CD1c+CD19-CD14-CD303- type 1 myeloid dendritic cells (mDC1s) at disease onset had a significant predictive value[74]. This profile may be of importance when considering interventions and evaluating their impact on β-cell function.

- **Tracking disease management**

Glycaemic control is currently generally assessed with HbA1c or capillary home blood glucose monitoring (HBGM). HbA1c is able to give an approximation of glycaemic control over the previous 3 months; however glucose excursions are not taken into account, therefore lability of control can be masked. Home blood glucose monitoring is a simple way for subjects to track glucose control and gives real-time data to inform management. Nonetheless its usefulness is operator dependent and repeated measurements can become a burden. Continuous glucose monitoring systems (CGMS) are available but cost restrictions largely limit them to research tools at the moment. Flash glucose monitoring is a newer technology able to monitor continuous interstitial glucose readings at a lower set-up cost than CGMS whilst still providing large amounts of data for glucose profile analysis.

- **Determination of tolerating immune response**

Ideally an immune biomarker able to track the islet targeted immune attack would provide a gauge of response to therapy. This would allow titration of treatment duration and dose. The use of enzyme linked immunospot (ELISPOT)[75] or modified versions such as the islet-specific CD8+ T-cell interferon-γ enzyme-linked immunospot (ISL8Spot)[76] assay offer the benefit of revealing an antigen-specific functional response by cytokine output in vitro, and are therefore attractive options as biomarkers of response to antigen specific immunotherapy.

Caution must be taken as few immune biomarkers are validated and in fact immune changes may not reflect outcomes. For example in a trial of rapamycin and IL-2 in newly diagnosed patients, Treg markers rose but despite this β-cell function declined.[77] Specificity of biomarker to disease, lag-time between treatment and outcomes, functionality of immune cell changes and clinical outcome will have to be considered in development of potential biomarkers.

- **Stratification of potential response to immune therapy**

In a number of trials a subgroup of patients have shown clinical efficacy compared to the trial cohort as a whole, such as those with baseline titres of insulin autoantibodies >80 nU/mL within the oral insulin arm of the DPT-1 trial[78]. Similarly in anti-CD3 monoclonal antibody trials post hoc analysis identified sub-groups within the newly diagnosed population that benefited from anti-CD3 intervention, based on baseline metabolic control and CD8+ T cell markers[79] or insulin autoantibodies[80]. Further studies will aim to ascertain which biomarkers can be used to define those who are susceptible to immune modulation and such biomarkers are likely to be unique to each immunotherapy.

1.5 Generalised Immunotherapy

The implication of T-cells as part of the immune cell mediated pathogenesis of T1D, derived from the discovery of "insulitis", an infiltration of islets with T-cells, B-cells and macrophages[81,82]. This was reinforced by adoptive transfer studies of splenocytes or effector T-cells (Teff) from affected animals into healthy and diabetes-resistant recipients, resulting in pancreatic β-cell destruction and diabetes induction in mouse and rat preclinical models[83,84]. Both CD4+ and CD8+ T-cells are thought to be responsible for the pathogenesis of T1D[84], with autoantibodies providing an indicator of disease without direct causation. Diabetes-specific autoreactive T-cells can be found in healthy controls as well as individuals with T1D[85], signifying that regulatory factors normally present are lost in diabetic individuals leading to disease through breakdown of peripheral tolerance. This delicately held balance between disease and tolerance appears to hinge on the actions of regulatory T-cells (Tregs). Co-transfer of CD4+/CD25+ Treg cells with autoreactive T-cells from non-obese diabetic (NOD) mice into T-cell deficient NOD.SCID mice prevents induction of diabetes [86]. In preclinical models of T1D and human studies, functional defects in Tregs have been identified as possible causes of loss of tolerance [87–89].

Immunotherapy for diabetes targets the disease at its core, aiming to halt autoimmune destruction of β-cells. Early studies in the 1980s gave proof of concept that immunosuppression with agents such as cyclosporine[90], azathioprine and prednisolone[91] could lead to remission of β-cell destruction and metabolic improvements at least in the short term. However generalised immunosuppression raises concerns over non-specific targeting of the immune system. Ideal timing for immunotherapy is at or before clinical onset, the point at which there is greatest scope for β-cell salvage. Such treatment will target a young

age group with greater concerns over long-term risks of toxicity, viral reactivation and malignancy, especially if it is to replace current treatment with insulin that is relatively simple and safe.

Monoclonal antibody treatments have shown promise in autoimmune diseases, such as anti-CD20 treatment in multiple sclerosis[92]. Whilst T1D is classically thought to be a T-cell mediated disease, there is evidence that this may be over-simplistic. For example, it has been shown that B-cells can be a factor in disease pathogenesis. B-cells are part the infiltrative process seen in insulitis being the second most prominent cell type[82], and can act as antigen presenting cells using MHC class II. The role of B-cells in disease progression was strengthened with the success of selective B-cell depletion using the anti-CD20 antibody rituximab. B-cell depletion though anti-CD20 antibody (rituximab) use in T1D[93] lead to improvements in stimulated C-peptide levels compared to placebo from 3 months to 1 year[93], with follow-up data at 30 months showing treatment groups still had favourable C-peptide, HbA1c and insulin dose data compared to placebo patients, but there was still decline in C-peptide levels albeit delayed[94]. There have been concerns with a high number of infusion reactions linked to cytokine release and the potential risk of progressive multifocal leukoencephalopathy which has been seen after rituximab use in other immune-mediated diseases such as rheumatoid arthritis and systemic lupus erythematosus[95]. Anti-CD3 antibody use in T1D has been clinically studied with two humanised Fc engineered antibodies, teplizumab[96] and otelixizumab[97] showing initial promise with improvements in C-peptide levels. In order to try and reduce adverse events in phase III studies significant dose reductions were made, in the case of otelixizumab the dose was reduced from a total dose of 48-64mg from the previous study[96] to a total dose of 3.1mg in the DEFEND-1[98] and DEFEND-2[99] studies. Reducing the dose to less than one-tenth of the previous effective dose would likely have negated not only the adverse events but also the beneficial effects of treatment, with both the phase III proving to be non-efficacious. Despite the fact that the Protégé study using teplizumab failed to reach an unvalidated primary composite endpoint of insulin <0.5 U/kg/day and HbA1C <6.5% (48 mmol/mol) at 1 year, further analysis showed significant differences between change in C-peptide MMTT AUC from placebo, particularly in children aged 8-11 years[100]. This highlighted the importance of defining and validating endpoints for such studies.

Interleukin 2 (IL-2) is a critical cytokine for the survival and function of FOXP3+ Treg cells. IL-2 administration in NOD mice can suppress or reverse established T1D and increase FOXP3+ Treg levels [101,102]. Recombinant IL-2 is currently in clinical use for treatment of metastatic renal cell carcinoma, but potential benefits of IL-2 in T1D have to be balanced by toxicity of systemic IL-2. Low dose IL-2 studies in children and adults with T1D are in progress (ClinicalTrials.gov: NCT01862120 and NCT02265809). The combination of IL-2 and rapamycin (an inhibitor of mammalian target of rapamycin complex (mTORC) activation) has been successful in prevention of spontaneous diabetes and recurrent diabetes after islet transplant in the NOD mouse[103]. Rapamycin inhibits Th17 and Th1 proliferation with lesser effects on Tregs[104]. Its combination with IL-2 therapy in 9 patients up to 42 months post diagnosis of T1D resulted in a marked decrease in C-peptide levels despite an increase in FOXP3+ Tregs[77]. The finding of this study shows the pitfalls of translating treatments from animal models such as the NOD mouse, whose responses often differ to humans, as well as extrapolating success from other conditions (in this case the success of IL-2 and rapamycin combination in graft versus host disease) which can have unique responses.

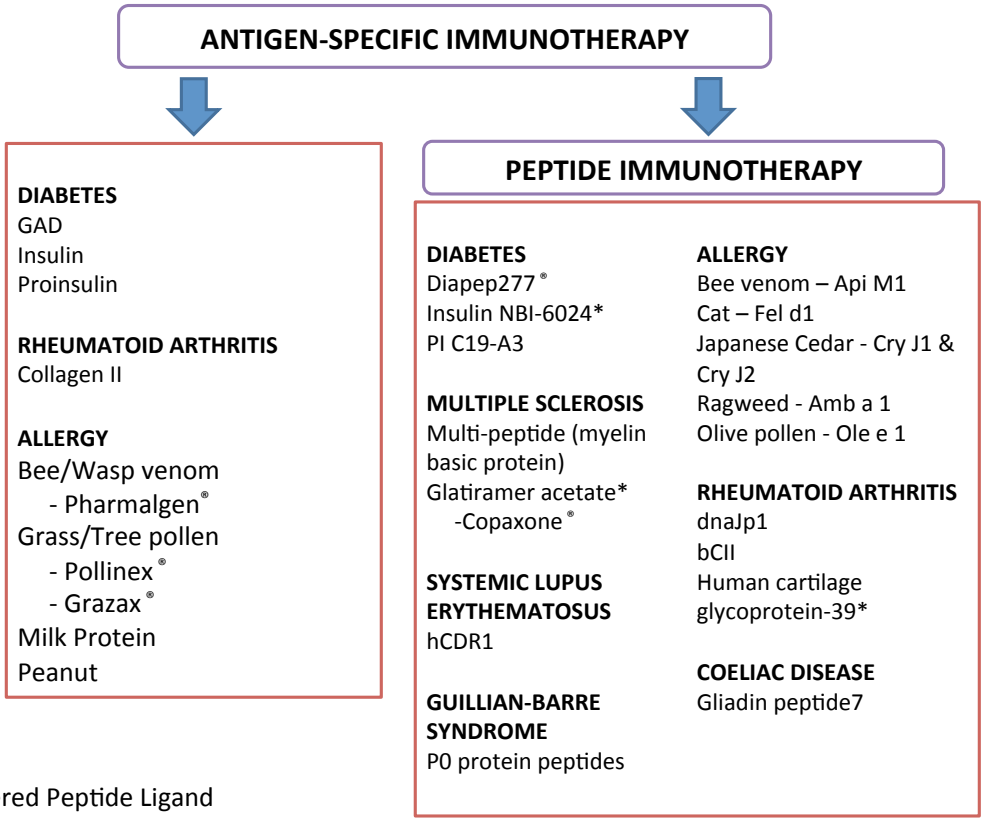
1.6 Antigen Specific and Peptide Immunotherapy

Antigen-specific immunotherapy (ASI) offers a targeted approach to modulating the autoimmune process towards tolerogenic responses compared to immune suppression. The aim is to re-establish tolerance by administration of an autoantigen, which can be in the form of whole antigen or as peptide immunotherapy (PIT), using peptides derived from antigens (Figure 2). Use of peptide epitopes has potential benefits over whole antigen administration such as ease of use, enabling a pure highly concentrated epitope to be delivered at lower cost and higher solubility than whole antigen and without the risk of complications of biological activity[105]. These properties also simplify methods to combine multiple peptides for more efficacious therapy[106]. As natural processing of the antigen is bypassed, choice of peptide is important as it is imperative to isolate a peptide which is biologically valid[107] and disease-specific. Identifying peptides that are naturally processed and presented epitopes (NPPEs) should ensure that such epitopes are able to elicit a T-cell response by directly binding to HLA, and also increase the chance that such epitopes are present in vivo in pancreatic islets or lymph nodes through MHC class II presentation. Cryptic peptides are those not naturally processed and presented and these can incur a higher risk for adverse effects. Altered peptide ligands (APLs) are a form of cryptic peptides, that have amino-acid substitutions in TCR contact positions and compete with the natural ligand or interfere with T-cell

activation[108]. ASI has been well established clinically in allergy and asthma, fuelling multiple trials in autoimmune disease (Figure 2) although this has yet to translate into viable clinical options with the exception of glatiramer acetate, a 4 amino-acid peptide based on myelin basic protein (MBP) currently used in multiple sclerosis[109].

Figure 2: Areas of research into antigen specific immunotherapy and peptide immunotherapy.

The diagram identifies antigens and peptides investigated as immunotherapy in allergy and autoimmune disease.



*Altered Peptide Ligand

1.6.1 Mechanisms of action

The use of ASI was successfully pioneered in allergy[110] and asthma[111,112] where it was noted to induce IL-10[111,112] linked to regulatory CD4 T-cell (Treg) action[111]. For example, epitopes from a major cat allergen Fel d 1 for allergic asthma have been used to induce tolerance in a mouse model[112]. A transgenic HLA-DR1 mouse was sensitized to Fel d 1 (cat allergen) by intraperitoneal administration of antigen in the presence of alum as an adjuvant, this was followed by 3 doses of allergen followed by desensitization with 1 µg Feld$^{(29-45)}$ or control

$HA^{(306-318)}$ peptides intradermally. Mice given the $HA^{(306-318)}$ peptide showed extensive peribronchial and perivascular inflammatory infiltrates not present in the $Feld^{(29-45)}$ group. The $Feld^{(29-45)}$ treatment was able to significantly increase the proportion of IL-10+ CD4+ T-cells compared to control peptide. This IL-10 would appear a crucial mediator of tolerance, as an IL-10 receptor blocker was able to negate the effects of PIT. There were also signs of linked suppression or intramolecular tolerance, with reduced proliferation of $DR1Feld1^{(29–45)}$ $tetramer^{neg}$ cells in response to other DR1-restricted epitopes of Fel d 1 after desensitisation. This treatment has now reached human trials with sustained effects on rhinoconjunctivitis symptom scores seen at 2 years after initiation of a short 16-week treatment period with Fel d 1 epitopes[113].

Mechanisms in ASI and PIT have been difficult to fully elucidate but some common pathways have come to light.

Potential mechanisms postulated in peptide immunotherapy include[114]:

- immune regulation (associated with Treg, IL-10 and TGF-β induction)
- immune deviation (from a Th1 to Th2 phenotype)
- immune deletion or functional inactivation of pathogenic cells.

Mechanisms of action may be influenced by choice of antigen, route of administration and antigen dose and frequency.

ASI is thought to induce tolerance through induction Treg populations which promote self-tolerance through secretion of cytokines such as IL-10[115], TGF-β and IL-4[116] or reduction of IFN-γ[117–119]. The importance of IL-10 production in PIT can be highlighted with neutralisation of this cytokine. The use of an anti-IL-10 monoclonal antibody in transgenic (Tg4) mice expressing a TCR specific for the N-terminal peptide Ac1–9 of MBP tolerized to EAE by PIT, abrogated the effects of PIT and rendered the animals susceptible to disease induction[120]. The concept of bystander suppression is thought to be important in PIT, by which induced Treg cells are capable of homing to the autoimmune target site and secreting cytokines such as IL-10, IL-4 and TGF-β in response to release of local autoantigens, leading to silencing of aggressive Teff also reacting to these autoantigens[116]. In a mouse model of virus induced diabetes the IL-4 and the stat-6 pathway were found to be essential for induction of these regulatory cells[116]. This Treg induction can also lead to linked suppression of T-cell response

to third party antigens presented by the same APC[121]. This phenomenon could potentially counterbalance the action of epitope spreading which propagates autoimmune targets with increasing damage to the β-cell, whilst also keeping suppression specific to the disease process and avoiding suppression of normal responses.

High doses in PIT can lead to deletion of pathogenic cells likely through activation-induced cell death[122], as well as suppression of proliferative responses to antigen[122]. This was seen with 100μg doses of myelin basic protein peptide Ac(1-9) used intranasally in EAE mice[120], and oral ovalbumin (5 doses of 5-500mg) in mice transgenic for ovalbumin-specific T-cell receptor genes[122]. It has been suggested that deletion of high-affinity TCRs in response to high doses may induce tolerance by leaving intact low-affinity T-cells that respond poorly to antigen in the absence of IL-2. Lower doses may select induction of antigen-specific cells which produce TGF-β, IL-4 and IL-10[122].

McPherson et al.[123] recently reported on the role of programmed cell death protein-1 (PD-1) after PIT in a mouse model of experimental autoimmune encephalomyelitis (EAE). PD-1, which is encoded by the *Pdcd-1* locus, acts as a co-inhibitory molecule that is transiently upregulated during CD4+ and CD8+ T-cell activation in order to restrain the primary immune response[124]. The McPherson et al. model used a peptide of myelin basic protein, Ac1-9(4Tyr), and a transgenic mouse bred to exhibit T-cells specific for this peptide, the Tg4 mouse. T-cells from Tg4 mice were introduced into B10.PLxC57BL/6 mice, then the Ac1-9(4Tyr) peptide or PBS was administered before an attempt to induce EAE with peptide in CFA adjuvant. The Ac1-9(4Tyr) peptide was able to prevent EAE after just one dose. Transferred naïve Tg4 T-cells had higher levels of PD-1 in PIT treated mice compared to PBS treated (Figure 3A) and this higher level was sustained at day 4 and day 7 (Figure 3B). T effector cells (Teff) exposed to PIT showed high PD-1 levels for over 2 weeks and when transferred into a secondary host these Teff cells failed to induce disease, which the authors considered as “silencing” of the Teff function. Interestingly although the authors saw reduced levels of IFNγ and IL-17 from Tg4 cells isolated from the spleen post PIT, they did not witness an elevation in the frequency of FoxP3+ donor Tg4 cells or in the IL-10 response. PD-1 blockade on splenocytes post PIT allowed these cells to exhibit IFNγ and IL17 responses to peptide stimulation (Figure 3C). Using a PD-1 knockout mouse Tg4.PD-1−/− as a donor prevented subsequent PIT from tolerising Teff cells**.** Interestingly the same group found that PD-1 was not critical for

induction of tolerance by PIT in a PD-1-/- knock-out mouse induced by transfer of naïve OVA(323-339)(pOVA) peptide reactive T-cells[125].

Compared to other EAE mouse model work, the lack of IL-10 response was unusual as PIT with Ac1-9 has resulted in an increase in antigen specific CD4+ T-cell release of IL-10[120].

The authors were able to illustrate a mechanism of suppression through reduction in 5-hydroxymethylcytosine (5hmC) and hypomethylation at the site of the PD-1(*Pdcd1*) promoter allowing increased sustained expression of PD-1. The cytosine demethylation of 5hmC occurred in the presence of permissive histone modifications at the promoter site and with facilitation by Ten-Eleven-Translocation (TET) proteins.

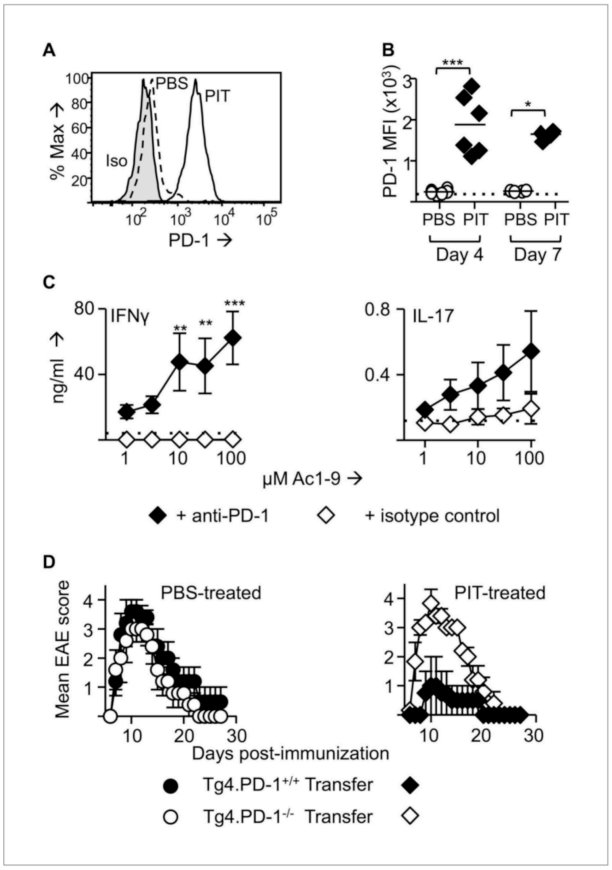


Figure 3: Role of PD-1 in tolerance. From McPherson et al. 2014[123]

PD-1 is required for the establishment and maintenance of tolerance in naïve CD4+ Tg4 cells. (A–C) B10.PL mice received PBS/PIT 1 day after transfer of naïve CD4+ Tg4 cells. (A) Representative histograms of PD-1 expression gated on CD4+ Tg4 donor cells in spleen 4 days after PBS/PIT. (B) MFI of PD-1 staining gated on CD4+ Tg4 donor cells in spleen 4 and 7 days after PBS/PIT (4–6 mice per group, from one of three experiments giving consistent results, dotted line represents MFI of isotype control staining). (C) IFN-γ and IL-17 production in response to Ac1-9 by splenocytes isolated 7 days after PIT and cultured in the presence of anti-PD-1 or isotype (four mice per group from one of three experiments giving consistent results, dotted lines represent cytokine levels for unstimulated cultures). (D) EAE in B10.PLxC57BL/6 mice that received PBS/PIT 1 day after transfer of naïve CD4+ cells from Tg4.PD-1+/+ or Tg4.PD-1−/− donors. EAE was induced 7 days after PBS/PIT by immunization with Ac1-9 (five mice per group, from one of three experiments giving PIT-drives PD-1 expression in consistent results).

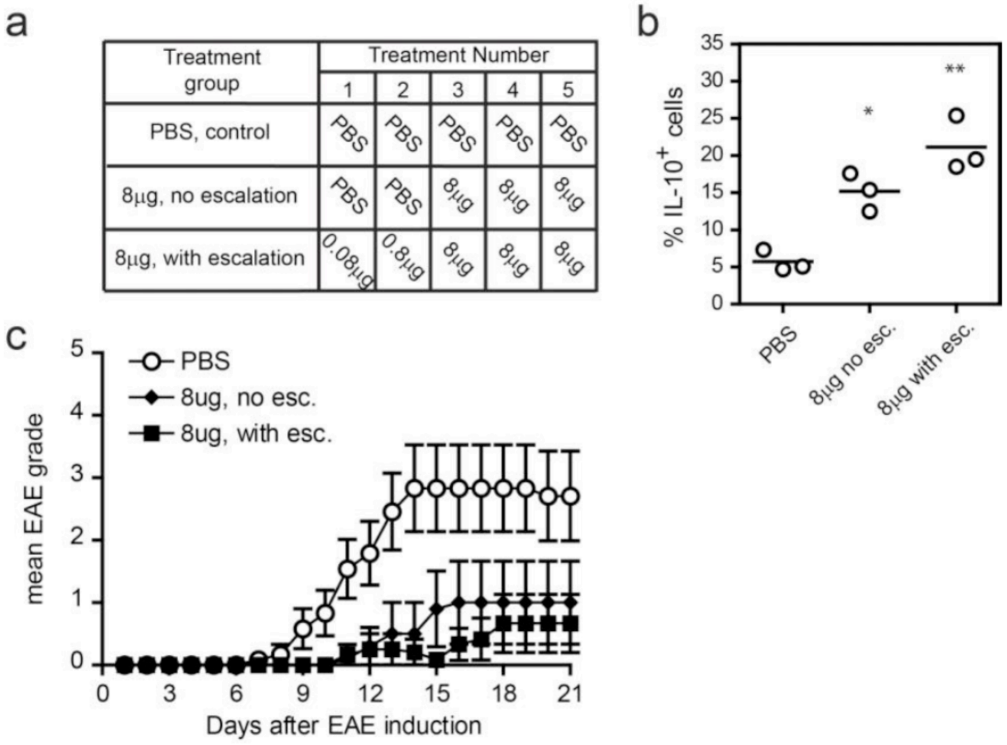
Another observation in Teff cells post PIT was that although the donor Teff cells persisted, they were not seen infiltrating the CNS and were more confined to the spleen. On examining the role of P-selectin glycoprotein ligand (PSGL-1), an adhesion molecule previously implicated as having a crucial role in T-cell trafficking to neuronal sites of EAE[126], it was found

that encephalitogenic function of the Teff cells was correlated to elevated PSGL-1 expression, with lower levels seen in PIT treated Teff cells. Antibody blockade of PSGL-1 led to attenuation of Tg4 Teff cell driven EAE, suggesting that T-cell trafficking via PSGL-1 may be disrupted by PIT.

Dose of antigen can influence mechanism of action, this was explored further recently in a study investigating sequential transcriptional changes in escalating dose immunotherapy (EDI)[127]. A transgenic Tg4 mouse EAE model was used to show that EDI was safer and induced a greater IL-10 response with greater disease suppression than a high dose regimen (Figure 4).

Figure 4: Dose escalation enhances the induction of IL-10+ CD4+ T-cells in addition to minimising the risk of adverse effects during immunotherapy. From Burton et al. 2014[127]

*Tg4 mice were treated s.c. with 3x8µg doses of MBP Ac1-9[4Y], with or without prior dose escalation, as illustrated (a). Splenocytes were cultured for five days with 10µg ml-1 MBP Ac1-9[4K] and 20U ml-1 IL-2. (b) Scatter plots show the percentage of viable CD4+ cells which are IL-10+, horizontal lines show means for each column (n=3). One of two experiments. * P ≤ 0.05, ** P ≤ 0.01 one-way ANOVA with Bonferroni post-test, comparing peptide-treated groups with PBS-treated control group. (c) Onset and severity of EAE in mice treated as illustrated in (a), immunised with MBP Ac1-9[4K]/CFA and pertussis toxin. Results of two independent experiments are pooled, showing mean disease score ± SEM (PBS group and 8µg, with escalation, both n=12. 8µg, no escalation n=10).*



Changes in T-cell transcriptome were tracked using whole genome expression arrays that found that expression of negative co-stimulatory molecules including *Lag3, Tigit* and *Havcr2* (TIM-3) were highly upregulated during treatments 1-6 of EDI and were maintained with extended high dose treatment. *Pdcd1*(PD-1) and *Ctla4* remained largely unchanged, whilst *Btla* and *Cd274*(PD-L1) decreased with extended treatment. A number of factors though to influence IL-10 production by CD4+ T-cells were upregulated by EDI, such as *Icos* and *Il10* expression as well as *Maf* and *Nfil3* transcript expression. The authors were able to immunolabel cells from $Tg4^{Il10/GFP}$ mice and demonstrate a positive correlation between IL-10 production and expression of LAG-3, TIGIT, PD-1 and TIM-3. By looking at the proportion of CD4+ cells expressing these markers they were able to discriminate induced cell populations with LAG-3 and PD-1 as good markers of an EDI induced anergic CD4+ cell population and TIGIT and TIM-3 as markers of EDI induced IL-10 secreting cells. Such findings offer potential for earlier tracking of efficacious tolerance induction through peptide immunotherapy intervention trials.

1.6.2 Safety

Many years of experience of ASI in allergy and asthma have generally supported its safety profile though potential risks of ASI are disease acceleration and hypersensitivity reactions. In asthma, the use of Fel d 1 peptides in cat allergic asthmatics can lead to isolated late asthmatic reactions which is felt to be T-cell dependent and IgE/mast cell independent[128]. Such late asthmatic reactions are felt to be more common in certain HLA groups, and repeated peptide doses can induce tolerance without further asthmatic reactions[129].

Whilst care should be taken when interpreting animal models, use of such models in ASI have shown that it disease acceleration is rare. Where this has been observed a number of factors may be involved such as route of administration, choice of peptide, dose and use of adjuvants. For example mild acceleration of disease occurred in the mouse models after intrathymic administration[130], the use of incomplete Freund's adjuvant (IFA)[131] and very large doses of oral antigen[132]. Use of GAD autoantigens in NOD mice at birth and 2 days of age was found to drive the precocious development of T-cell autoreactivity, but again this was in the presence of IFA using high doses of 200µg on 2 occasions. A clinical phase II trial in multiple sclerosis revealed local hypersensitivity and disease progression using an altered peptide ligand (APL) of myelin basic protein $MBP_{(83\text{-}99)}$ named CGP77116[133] which lead to termination of the trial. Subjects were not preselected by *HLA* genotype as the peptide ligand was

considered to be *HLA-DR* "promiscuous" and each subject was expected to have at least one *DR* allele the APL could bind to. It was also used at high doses subcutaneously of 50mg weekly for up to 9 months. Disease progression was associated with a skewing towards a Th2 phenotype with increased IFN-γ and reduced IL-4. Local skin reactions of pain, reddening and swelling up to several centimetres in diameter were seen in all 8 patients with systemic hypersensitivity reaction in one patient.

Apart from this study hypersensitivity in clinical ASI trials are rare. In preclinical models, ASI has been linked to fatal anaphylaxis in the NOD mouse[134] using repeated subcutaneous dosing of an altered peptide ligand of an insulin peptide B(9-23) (total dose >1mg). The authors were later able to avoid hypersensitivity by manipulating the isoelectric point on the same peptide[135].

Human studies of ASI have so far provided reassuring safety data with no excess adverse events with the use of insulin[78,136,137], proinsulin[138] GAD[139], heat shock protein 60[140] or an APL of insulin B(9-23)[141]. The first in man trial of intradermal proinsulin C19-A3[117] resulted in local reactions of up to 5 cm in diameter, however there was no accompanying wheal, itching or systemic symptoms to suggest hypersensitivity. These reactions appeared to be dose-dependent, with 11% of the 10μg dose group versus 56% of the 100μg dose group being affected. Although the reactions could recur on subsequent doses, they did not have a pattern of amplification. The investigators also found no evidence of induction of IgG or IgE anti-peptide antibody or T helper type 2 (Th2) IL-4, IL-5 and IL-13 cytokine responses. This led to the conclusion that these reactions were likely the result of vasodilatatory response to either the peptide or a minor contaminant from the synthesis process.

1.6.3 Clinical Trials of ASI and PIT in Type 1 diabetes

Insulin

Insulin or its precursors, (pre)proinsulin have been identified as sources of immunodominant epitopes in humans[142] and NOD mice[143]. There is significant evidence that derivatives from insulin or its precursors act as targets for the immune attack in T1D, by CD4+ T-cells in DR4 transgenic mice[144], humans[145,146] and by CD8+ T-cells in humans[147]. This in addition to the prevalence of insulin antibodies in newly diagnosed T1D suggests that insulin, its precursors or a peptide derivative of these would be a strong candidate for ASI. The Type 1 Diabetes Prediction and Prevention (DIPP) study[137] examined the use of nasal insulin versus placebo in

children (n=264) identified as a high-risk genetic group with positive antibodies, for a median follow-up period of 1.7 years (insulin) and 2.0 years (placebo). No significant difference was seen in prevention or delay in diagnosis of T1D.

The Diabetes Prevention Trial - Type 1 (DPT-1) was a large-scale trial in high-risk relatives of those with T1D aged 3-45 years. There were 2 study groups, split into treatment with oral insulin or twice daily subcutaneous insulin with annual four-day continuous intravenous infusions, with respective control groups (Table 1). Ultimately neither treatment was found to delay or prevent onset of diabetes, however a post-hoc analysis showed a significant (p=0.015) reduction in progression to T1D in a subgroup treated with oral insulin who had a baseline high titre of insulin autoantibodies >80 nU/ml compared to placebo. A further phase III trial using oral insulin (7.5mg/day versus placebo) in relatives of people with T1D who had at least 2 autoantibodies, including insulin antibodies and normal glucose tolerance was completed in 2016 with results recently reported[148]. These showed no delay overall in development in T1D over a median follow-up period of 2.7 years. A sub-analysis of strata within the trial showed that a secondary stratum (consisting of those with first phase insulin release below threshold, IAA and ICA positive, or positive for both GADA or IA2A) had a lower rate of progression (18.1%) than placebo (34.1%), HR, 0.45; 95% CI, 0-0.82; P = 0.006. There were a number of limitations in the study, which included a change in insulin antibody assay from the preceding DPT-1 study and also varying rates of adherence within participants. An exploratory analysis showed that those with better adherence rates in the primary stratum had lower progression to diabetes compared to placebo (annualized rate 2.4% with oral insulin and 6.9% with placebo (HR, 0.348; 95% CI, 0-0.855; P = 0.02)).

Table 1: Summary of Diabetes Prevention Trial-Type 1 Groups

	Treatment Group	Control Group	N	5 year risk of T1D	Criteria for risk assessment	Median follow-up period (years)
Diabetes Care, 2005[78]	Oral insulin	Blinded Placebo	372	26-50%	ICA and IAA positive FPIR above threshold* Normal glucose tolerance	4.3
NEJM 2002[149]	Subcutaneous insulin (twice daily) and annual four-day intravenous infusion	Unblinded closely monitored	339	>50%	ICA positive FPIR below threshold* and/or Abnormal glucose tolerance	3.7

FPIR=First Phase insulin response ICA= Islet cell antibody IAA=Insulin Autoantibody.

**The first phase insulin response in siblings, offspring, and second-degree relatives was considered to be below threshold if it was below the 10th percentile for this group (<100 µU per ml for subjects eight years of age or older; <60 µU per ml for subjects less than eight years of age); the response in parents was considered to be below threshold if it was below the first percentile for the group of parents (<60 µU per ml).*

GAD-Alum

The 65-kD isoform of glutamic acid decarboxylase (GAD) is a major autoantigen in patients with T1D[150] and is thought to be one of the earliest targets of T-cell reactivity before spreading of the response to other epitopes[151]. GAD antigen administered intra-peritoneally[152], subcutaneously or intravenously[151,153] to diabetes-prone NOD mice was able to inhibit disease progression.

An initial study using GAD incorporating an aluminium hydroxide adjuvant (GAD-alum) in T1D patients aged 10-18 years showed promise, with the rate of decline in fasting and stimulated C-peptide levels significantly reduced over a 30 month period, after only two 20 µg injections one month apart in a subgroup of those treated within 6 months of diagnosis[150]. *In-vitro* stimulation of PBMCs with GAD showed increased responses of IL-5, IL-10, IL-13, IL-17, IFNγ and TNFα, but not IL-6 and IL-12 between baseline and 15 months[150]. GAD-induced expression of transcription factor FOX-P3 and TGF-β was increased in GAD-alum group compared to placebo. GAD-specific induced CD4+CD25^{hi}FOX-P3+ cells had a strong positive association to secretion of Th2 and regulatory cytokines IL-5, IL-10, IL-13[118]. A follow-up study

4 years post treatment, showed evidence of long-standing preservation of β-cell function in the 6 months from diagnosis subgroup, with significant differences in fasting but not stimulated C-peptide levels in the GAD-alum treated patients compared to placebo[154]. A study of GAD in LADA patients, using placebo (n=13) or single GAD doses of 4μg (n=9), 20μg (n=8), 100μg (n=9) or 500μg (n=8) also showed fasting and stimulated C-peptide significantly increased at 24week in the 20μg group[155]. GAD has had a very good safety profile and despite theoretical concerns about stiff-man syndrome, a neurological condition linked to raised GAD antibody levels, there were have been no signs of this in GAD trials[139,154–156]. Despite early promising results, a larger scale phase 3 study using GAD-Alum failed to reproduce this β-cell preservation with no significant improvement in C-peptide levels after 12[156] or 15 months[139]. Potential reasons for the disappointing results of later trials have again called into question the limitations of extrapolating from animal models and also the optimal timing of intervention. For ASI earlier intervention may be necessary as by clinical presentation the disease is already well established and therefore the presumed mechanisms of ASI may be insufficient to control disease. Specifically authors also highlighted differences in age distribution between treatment groups and seasonal variations in the immune systems (with a greater effect seen in those treated in March-April as seen in the prior phase II study) and the effects of influenzae vaccination.

Diapep277

Diapep277 is a peptide derived from heat-shock protein 60 (hsp60). Despite being a ubiquitous protein the peptide was found to be an autoantigen linked to β-cell secretory granules in NOD mice[157]. A single peptide dose administered to NOD mice already showing signs of clinical disease led to regression of insulitis and maintenance of insulin secretion[158]. A study of male adults within 6 months of diagnosis used 3 injections of peptide in mannitol and vegetable oil based vehicle versus a placebo (mannitol in vehicle only) at baseline, 1 and 6 months[119]. This showed β-cell preservation with mean C-peptide concentrations from glucagon stimulation reduced in the placebo group but maintained in the Diapep277 group (0.26 vs. 0.93nmol/l, p=0.039).

Enzyme-linked immunosorbent spot (ELISPOT) assays of T-cell cytokine release *in vitro* 10 months after treatment, revealed that when using hsp60 compared to peptide p277 as the *in vitro* stimulus there was greater IFNγ (p=0.041) and less IL-13 (p=0.048) release, thus hsp60 induced a more Th1 response than p277 in controls. When looking at the Diapep277 treated

group there was less IFNγ (p=0.04) and more IL-10 (p=0.03) and IL-13 (p=0.04) compared to placebo.

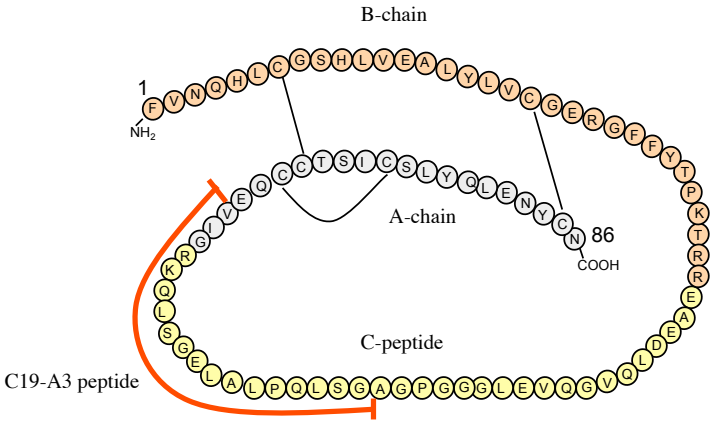
A further phase II study in a paediatric population failed to duplicate the outcome of preserved β-cell function, or show evidence of improved glycaemic control[159]. A large-scale (n=457) phase III study using Diapep277 in newly diagnosed adults was reported to have reached its primary endpoint of change from baseline in C-peptide levels at the end of the 24 month study period, with significantly more DiaPep277-treated than placebo-treated patients maintaining target HbA1c and entering partial remission[160]. Recently these results were called into question and the article was retracted due to concerns about serious misconduct and data manipulation[161]. Results are now awaiting independent third-party analysis.

1.6.4 Proinsulin C19-A3

Proinsulin C19-A3 (PI C19-A3) is an *HLA-DR4 (DRB1*0401)* restricted 18 amino acid peptide with the amino acid sequence of GSLQPLALEGSLQKRGIV (Fig 4), which matches the residues from position 19 on the C-peptide chain of proinsulin to position 3 of the A-chain. Its relevance as an autoantigen in Type 1 diabetes was supported by its presence in a panel of peptides eluted from the MHC class II molecule after presentation of proinsulin to antigen presenting cells positive for the high risk allele *HLA-DRB1*0401*[75].

As a naturally processed and presented epitope PI C19-A3 is targeted by T-cells from *HLA-DR4* positive T1D individuals. Evidence for this was shown by strong binding to soluble *HLA-DR4* in vitro and dominant immune responses detected through ELISPOT assay in T1D individuals[75].

Figure 5: Proinsulin molecule showing the 18 amino acid position of the peptide C19-A3 (marked in red)



The same authors from our laboratory, used *in vitro* ELISPOT studies to show that PI C19-A3 elicits an inflammatory IFNγ response, from subjects newly diagnosed with T1D, compared to regulatory IL-10 responses in control individuals[75]. As with previous studies in allergy, as PI C19-A3 was now identified as a target epitope, its use in ASI would be aiming to induce tolerance and reduce this polarisation of cytokine response. Theoretically a tolerisation would be detectable by an increase IL-10 response or reduction in IFNγ response.

A phase Ia study using PI C19-A3 in established T1D patients examined the effect of using 3 intradermal injections of the peptide given monthly at a dose of either 10µg (n=18) or 100µg (n=18) [117]. The subjects had been diagnosed with T1D for over 5 years (mean duration 23.8 years, range 7-45 years) and had low residual β-cell mass confirmed with glucagon stimulated C-peptide levels of <200pmol/L. A group of 12 randomly allocated subjects acted as controls. Two subjects in the lower dose 10µg group (11%) and 10 subjects (56%) in the 100µg group developed immediate local erythematous reactions up to 5cm in diameter lasting less than 2 hours. These reactions were not itchy, associated with a wheal, had no systemic features and did not progress with subsequent injections. There was no development of anti-peptide antibodies or consistent change in anti-GAD and anti-IA-2 antibodies.

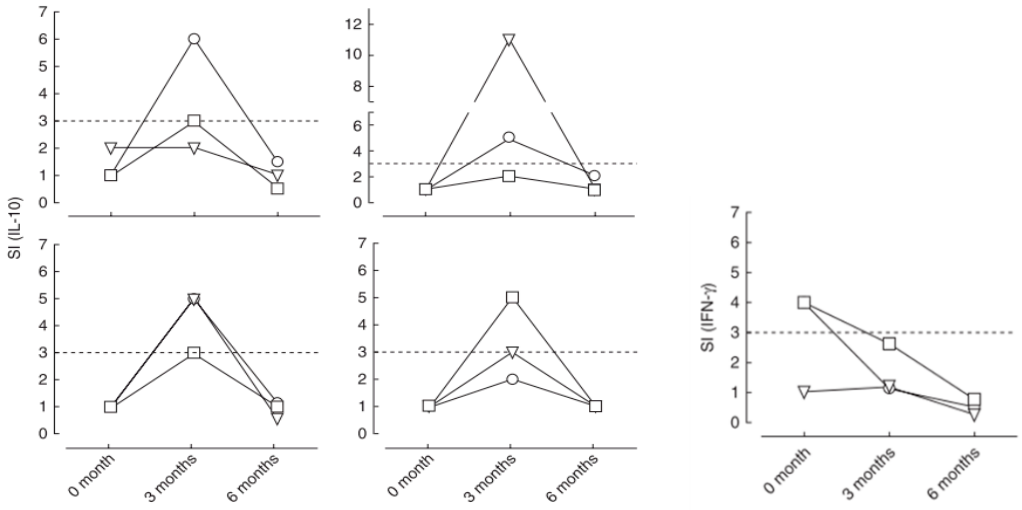
T-cell cytokine responses were evaluated by ELISPOT assays performed by the laboratory blinded to subject treatment allocation at 0, 3 and 6 months. No induction of IL-4, IL-5 or IL-13 was seen. Favourable cytokine responses were pre-determined as a loss of IFNγ or induction of IL-10 response at 3 months. This was based on the hypothesis that the C19-A3

PIT would work via mechanisms inducing C19-A3-specific CD4 IL-10 secreting cells and/or deletion or suppression of pro-inflammatory IFNγ producing C19-A3-specific cells. Such favourable responses were assessed prior to unblinding of data. Each peptide treatment group had a favourable responder who showed a significant IFNγ response at baseline (stimulation index (SI) >3.0 at two peptide concentrations), which was absent at subsequent analyses. Four subjects in the 10μg group and one control subject had a favourable response through induction of a significant IL-10 response to C19-A3 at 3 months.

With a limited number of doses tested, no clear dose response was elicited in the study, but overall favourable *in vitro* responses could be seen more commonly in the lower dose group (28%) compared to the higher dose (6%) (Figure 6). However these responses were not sustained after withdrawal of the peptide. Interestingly there was a small but significant improvement in glycaemic control, shown by a fall in HbA1c by 0.23% (95% CI -0.42 to -0.03, p=0.02), however as a stand-alone measurement without correlating other metabolic indicators such as insulin dose and within a small study this HbA1c fall should be interpreted with care.

Figure 6: Favourable responses seen in seen in ELISPOT assay during phase I trial of proinsulin C19-A3 peptide. From Thrower et al[117].

Graphs show ELISPOT IL-10 and IFN-ɣ responses to peptide at 0, 3 and 6 months measured by stimulation index (SI) in 10ug treatment group. (SI derived as number of spots in test wells/number in saline wells).



In summary the phase Ia trial provided evidence that PI C19-A3 PIT could elicit favourable *in vitro* responses more frequently in a low dose group with less risk of skin reactions. Overall

there were no systemic hypersensitivity reactions nor were there any signs of IL-4, Il-5 or IL-13 induction. No adverse antibody or T-cell changes were seen compared to the controls.

Whilst the current available data is does not give a clear indication of mechanism of action of PI C19-A3, it is proposed that its actions are through immune regulation and via IL-10 secretion and immune deviation from Th1 to Th2 phenotypes. The responder findings in the first-in-man studies indicate that there are changes in IL-10 and IFNγ to support this. One of the aims of the MonoPepT1De trial is to elucidate immune mechanisms of this therapy.

1.7 The MonoPepT1De trial

1.7.1 Rationale

Thrower et al.[117] established safety of PI C19-A3 PIT in long-standing T1D subjects with evidence pointing towards improved efficacy with lower doses. Ultimately use of immune interventions would target individuals at or even before diagnosis in order to maximise potential for salvage of β-cell mass. Should such therapies risk disease progression, newly diagnosed patients would also suffer the potential loss. The MonoPepT1De trial was designed to examine PI C19-A3 safety in a newly diagnosed patient cohort who exhibited peak stimulated C-peptide levels >200pmol/l. This level was established by the DCCT[47] as clinically relevant in prevention of long-term complications. It also introduced a placebo arm, which was lacking in the phase Ia trial.

1.7.2 Trial design

The MonoPepT1De trial was a multicentre randomised, double-blind, placebo controlled phase 1b trial to primarily assess the safety of intradermal administration of proinsulin peptide C19-A3 in newly diagnosed T1D subjects within 100 days of diagnosis. The trial protocol is summarised below.

The treatment was broken down into 3 arms, Group 1 received saline placebo every 2 weeks for 6 months, Group 2 received 10µg of peptide alternating with saline every 2 weeks for 6 months and Group 3 received 10ug of peptide every 2 weeks for 6 months. This was followed by an observation period of 6 months (Figure 7).

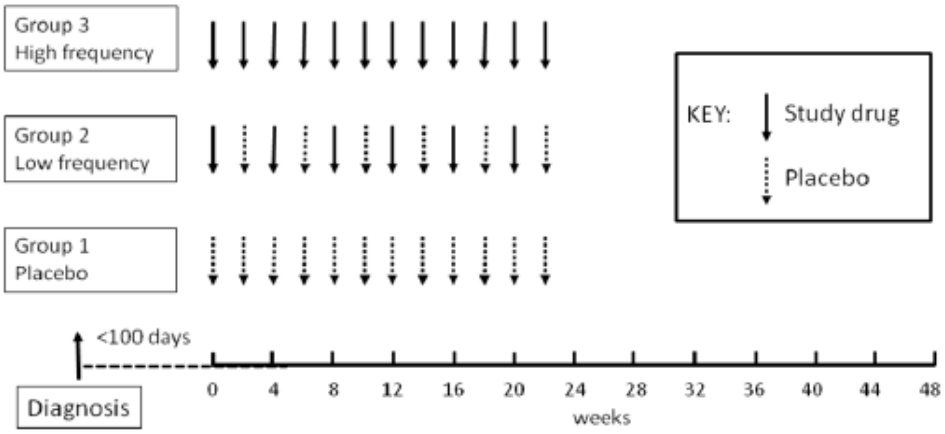


Figure 7: Chart indicating dosing schedule for trial patients and duration of study

Inclusion Criteria

For patients to have been included in the study, they must fit the following criteria:

- Age 18-45 years.
- Diagnosis of Type 1 diabetes within the last 100 days (dated from the first insulin injection).
- Possession of **0401* allele at the *HLA-DRB1* gene locus
- At least one positive islet cell autoantibody (i.e. anti-GAD65, antibodies to insulinoma-associated antigen-2 (IA-2) or zinc transporter 8 (ZnT8)).
- Peak insulin C-peptide >200 pmol/L (at any time point after stimulation with Mixed Meal Tolerance Test).
- Written and witnessed informed consent to participate.

Exclusion criteria

Criteria that prevented patients from being enrolled onto the trial are listed below.

- Females who are pregnant, breast-feeding or not using adequate forms of contraception.
- Use of immunosuppressive or immunomodulatory therapies, including systemic steroids within 1 month prior to randomisation and any monoclonal antibody therapy given for any indication.
- Any other medical condition which, in the opinion of investigators, could affect the safety of the subject's participation.

- Recent subject's involvement in other research studies which, in the opinion of investigators, may adversely affect the safety of the subjects or the results of the study.
- Subjects should not have had immunisations with live or killed vaccines or allergic desensitisation procedures less than 1 month prior to their first treatment.

1.7.3 Primary endpoint

- Assessment of the safety of PI C19-A3 peptide administration in subjects with new-onset Type 1 diabetes.

1.7.4 Secondary endpoints

- Change in stimulated C-peptide production at 12, 24, 36 and 48 weeks versus baseline and between groups.
- Change in level or quality of T lymphocyte biomarkers of β-cell specific immune response at 12, 24, 36 and 48 weeks versus baseline and between groups.
- Change in level or quality of islet cell autoantibody biomarkers of β-cell specific immune response at 12, 24, 36 and 48 weeks versus baseline and between groups.
- Change in glycated haemoglobin (as measured by % HbA1c levels), daily insulin usage, and mean amplitude of glucose excursions at 12, 24, 36 and 48 weeks versus baseline and between groups.
- Changes in the Hypoglycaemia Fear Survey (HFS), and Diabetes Treatment Satisfaction Questionnaire (DTSQs) scores at 3,6 and 12 months and ADDQoL (Audit of Diabetes-Dependent Quality of Life) at 6, and 12 months versus baseline and between groups. The DTSQc (Diabetes Treatment Satisfaction Questionnaire, change version) will also be compared between groups at 12 months.

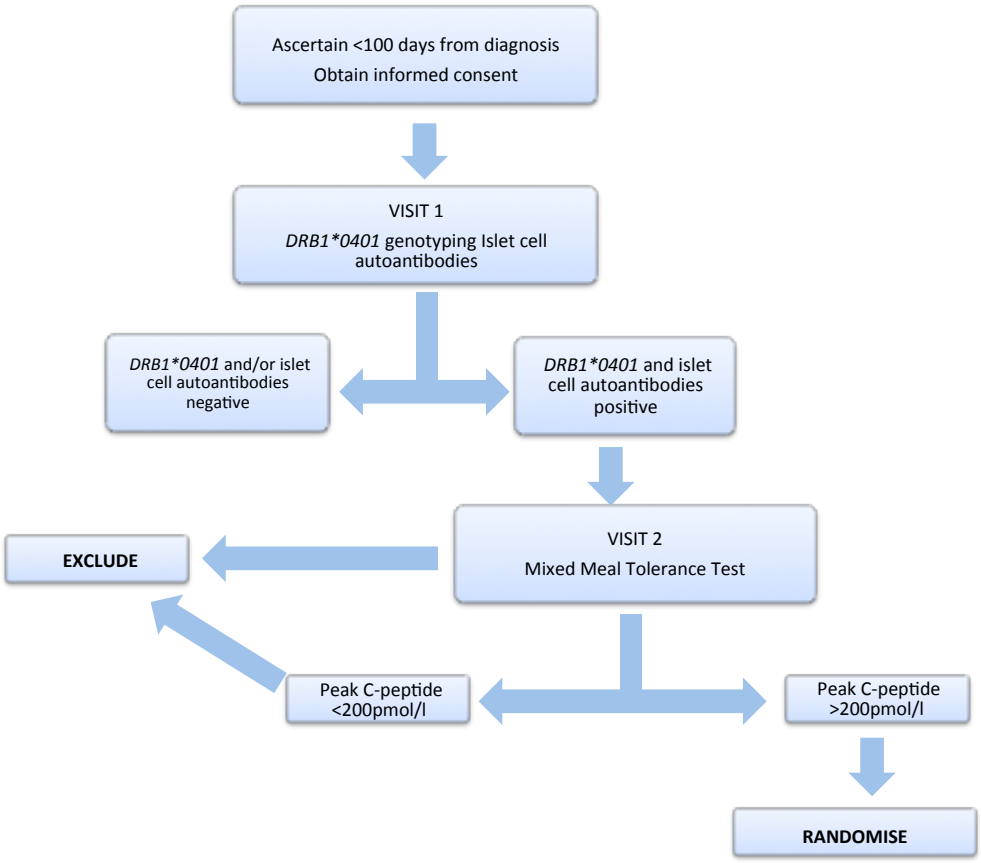
1.7.5 Screening and Recruitment

Subjects enrolled to the trial underwent a two-stage screening approach (Figure 8). Stage one involves blood sampling for *HLA DRB1*0401* genotyping and autoantibody testing for GAD, IA-2 and ZnT8 antibodies. Only if the patient was found to be *HLA DRB1*0401* positive and have at least one autoantibody positive would the subject pass onto the second stage of C-peptide testing.

The C-peptide testing consists of a fasting 2 hour MMTT looking for peak C-peptide levels of greater than 200pmol/L at any time point.

If both screening stages were passed patients would proceed to randomisation and dosing by 100 days post diagnosis.

Figure 8: MonoPepT1De screening pathway



1.8 Profiling of the TCR repertoire

The majority of T-cell receptors (TCR) are αβ receptors, composed of a heterodimer of α and β polypeptide chains linked by disulphide bonds. These αβ-TCRs are responsible for HLA-restricted antigen recognition through interaction with peptides presented by MHC molecules on antigen presenting cells. In order to allow responses to a wide range of pathogenic molecular signals, humans are estimated to have a potential αβ TCR diversity of approximately 10^{13}-10^{16} TCR post thymic selection[162,163]. This huge repertoire is formed by the stochastic recombination of 3 gene segments, the variable (V), diversity (D) and joining (J)

segments, in combination with random nucleotide additions and deletions to segment junctions. Within the α and β chains are 3 hypervariable complementarity determining regions (CDR). CDR3, formed through the recombination of the D and J regions, is the main recognition site for processed antigen. Self-reactive T-cells undergo negative selection centrally within the thymus, however those that escape this regulatory step can lead to autoimmune disease. Although CDR3 regions are formed by chance combinations, bias can occur due to convergent recombination with multiple recombination events leading to the same nucleotide sequence and multiple nucleotide sequences converging to the same amino acid sequence (Figure 9).

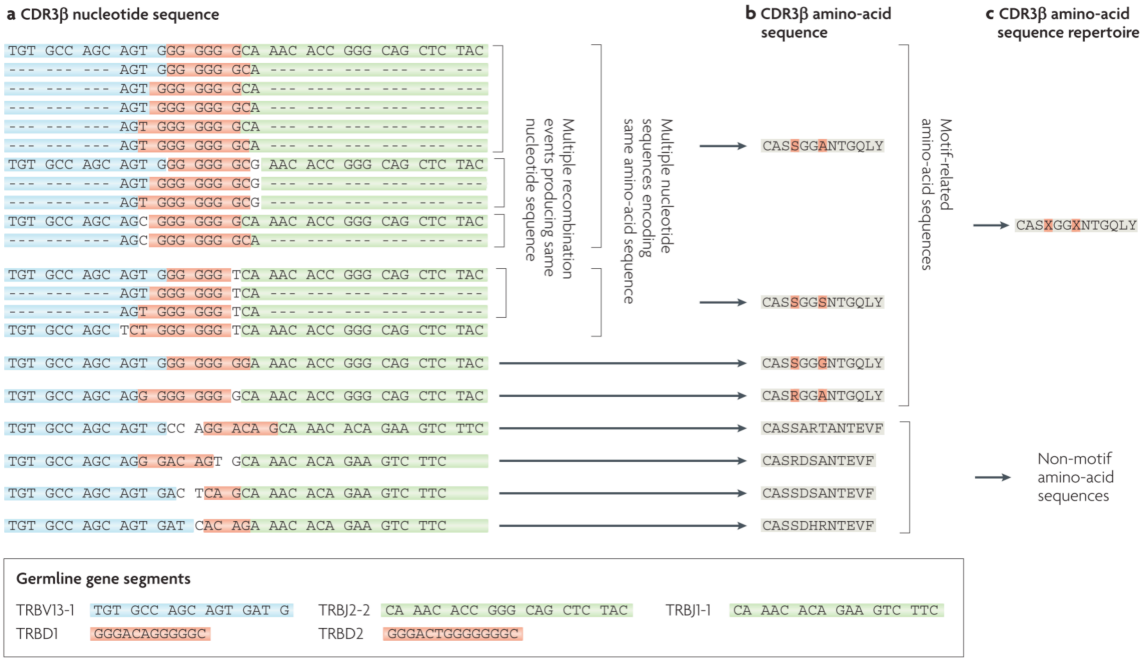
1.8.1 TCR Repertoire in preclinical models of autoimmune disease

In NOD mice thymically derived nTreg cells were found to show a lower diversity than controls with shortening of the CDR3 region[164]. Study of the NOD mouse response to a GAD65 derived peptide revealed a public response using the Vβ4, Dβ2.1 and Jβ2.7 gene segments, with a shared CDR3 motif[165]. An EAE mouse model examined with spectratyping of the CDR3 region of the β TCR chain revealed the presence of two public clonally restricted subsets in response to a MBP epitope Ac1-9, one linked to disease onset, which could transfer disease to naïve recipients, and one to recovery[166]. Similarly another group was able to identify a CD4+ Treg clonotype in EAE mice linked to recovery[167]. Using another EAE mouse model, it was shown that expansion of TCR diversity with epitope spreading is linked to relapses of EAE, not present in terminal deoxynucleotidyl transferase (TdT) knock-out mice who have a reduced diversity of the CDR3 region[168].

Using high-throughput screening (HTS) directed at islet-infiltrating memory CD4+ T-cells in pre-diabetic and diabetic NOD mice, distinct unique monoclonal expansions were found to dominate the islet infiltrating CD4+CD44^{high} TCRβ repertoire in individual diabetic mice[169]. Of these expansions a number were shared with the non-diabetic mice, which the authors suggest may be due to pathogenic clonotypes being resident in pancreas before disease onset. Of note there was a lack of a non diabetes-prone control population in this study to compare background populations. Of the islet-infiltrating CD4+CD44^{high} cells in both groups of mice, a predominant pattern of gene segment use was seen involving TRBV1 (Vβ2), TRBV13-3 (Vβ8.1), and TRBV19 (Vβ6) gene segments.

Figure 9: Schematic illustration of convergent recombination.

Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Immunology Venturi et al[170], 2008. A. The first level of convergent recombination is multiple recombination events, involving different splicings of the germline V (variable; highlighted in blue), J (joining; highlighted in green) and D (diversity; highlighted in red) β-chain gene segments and random nucleotide additions (no highlight), produce the same nucleotide sequences. B. The second level of convergent recombination involves multiple nucleotide sequences encoding the same amino-acid sequence. Protein sequences with amino acids in the V(D)J junction that are encoded by many codons can be encoded by many different nucleotide sequences. C. The third level of convergent recombination is at the level of the TCR repertoire, where some of the amino-acid sequences conform to an amino-acid motif. In this particular case, the 'XGGX' amino-acid motif (where X denotes any amino acid) is facilitated by strings of guanines encoding the two glycines, which can be spliced, often in multiple ways, from the Dβ gene segments. In most T-cell responses, some TCR sequences will also be observed responding to the epitope that do not conform to an amino-acid motif. These sequences are represented by the four non-motif amino-acid sequences.



1.8.2 Human studies of TCR repertoire in type 1 diabetes

In human studies, an early study examined peripheral TCR repertoire in people with type 1 diabetes used quantitative PCR, coamplification with radiolabelled primers and liquid scintillation in 3 pedigrees[171]. This showed there were no significant differences in the peripheral repertoire between T1D patients and their unaffected siblings and parents including a pair of discordant monozygotic twins. This was echoed in another study of monozygotic twins concordant and discordant for T1D[172]. Although it was suggested such findings indicate that TCR diversity is influenced by genetic factors more than the presence of

disease, consideration would have to be given to the limitations in the resolution of such techniques in whole CD4/CD8 subsets from peripheral blood.

Identification of TCR clonotypes in human pancreatic islets from an individual recently diagnosed with T1D showed a diverse repertoire of islet infiltrating T-cells consisting of Vβ1, Vβ7, Vβ11, Vβ17, and Vβ22 families, within which there was a biased amino acid usage[173]. One of these Vβ22 clonotypes identified could be traced back into the spleen and peripheral blood. Until recently many studies have used older technologies to analyse TCR repertoire. These technologies are reviewed by Miles et al[174]. Limitations on resolution of repertoire are a concern with older methods such as spectratyping where quantitative measures of TRV gene usage are limited to the range of CDR3 length variations and affected by variance in PCR efficiencies. In flow cytometry limitations occur as not all Vα and Vβ monoclonal antibody fluorochrome stains are available.

Research has been sparse in the field of effects of PIT on TCR repertoire, although one study demonstrated an increase in TCR diversity using spectratyping of CDR3 regions of a Tr1 population in those treated with a sublingual multiple epitopes of Cry j 1 and Cry j 2, major Japanese cedar pollen allergens, to treat pollinosis[111].

With technologies for TCR sequencing continually expanding, HTS is allowing ever deeper analysis of TCR diversity[175,176] and may allow identification of pathogenic T cell receptor motifs or clonotypes or shifts in T-cell populations. Comparing traditional Sanger sequencing with HTS such as the Roche® 454 pyrosequencing technology, HTS allows a higher sensitivity sequencing at a greater depth with a less labour intensive methodology and with the incorporation of parallel sequencing through multiplexing, the costs are significantly lower per sample. Another HTS method using the multiplex solid phase PCR and Illumina genome analyser technology, was able to sequence tens of millions of TCRs from a single sample with the sensitivity to pick out extremely rare clones doped into samples at a frequency of 1 in 100,000[177].

Clonal diversity was examined in T1D using HTS exploring CD4+ T-cells presenting markers of early activation (CD69+/CD154+) post *in vitro* stimulation with an islet antigen IA-$2_{752–775}$ or influenza HA peptide[178]. Interestingly HA and β-cell epitope specific T-cell total numbers, clonotype frequency and diversity were similar. This finding contrasted with previous work using ELISPOT data to identify autoreactive cells through functional cytokine release that

suggested that autoreactive T-cells are rare compared to pathogen responsive T-cells[75]. Without reliance on functionality the HTS approach offers a less biased estimation of antigen specific cells, which can potentially be polyfunctional. Newer techniques allow for sequencing of paired α and β TCR chains, providing a far greater resolution of TCR specificities. Eugster et al.[179] used this method to examine GAD specific TCR clonotype responses in T1D subjects and antibody positive children. Their findings were that α or β components of the TCR were often shared between individuals however the αβ pairing was always unique. These findings are discussed further in 6.4.3. Developing technologies link single cell α and β sequencing with transcriptome analysis, through qPCR[180] and more recently single cell RNAseq[181–183], allowing more in-depth information of α and β-chain clonotype pairing as well as informing on functional phenotypes[184]

1.9 Gene expression studies

With array technologies expanding and becoming more cost effective, gene expression studies have become powerful and widespread. Such studies offer an insight into complex genetic relationships linked to disease-state and can be used to identify "key players" in disease, which in turn can focus biomarker discovery and therapy targets. Whilst there are studies of gene expression changes within islets[185–188], these studies are impeded by difficulties in obtaining T1D islet samples, which can vary in age of onset and disease duration and methods of islet isolation can effect cytokine expression[186]. In T1D, HLA class I upregulation has been confirmed in islets through a number of studies[185–187], but there is little overlap in other differentially expressed genes which may reflect the differences in the T1D and controls studied.

Studying peripheral blood as whole blood or derivatives, allows for ease of sampling with the goal of identifying accessible biomarkers. Whilst larger numbers can be studied, there is still variability in T1D populations and the control populations chosen. Relevant gene expression studies are summarised in Table 2.

Focusing on T-cell specific gene expression changes can pinpoint immune-related dysregulation. In the NOD mouse differential gene expression was seen in genetic pathways (*Clec16a, Dlk1, Rnls, Sh2b3, and Zac1)* under IL-2 and induction of *Ctla4, Foxp3, Il2, Ptpn22, Sh2b3, and Zac1* induced by TGF-β signalling in T-cells[189]. Orban et al.[190] isolated unstimulated CD4+ T-cells from newly diagnosed T1D subjects (less than 6 month duration) for gene expression studies versus healthy controls and Type 2 diabetic subjects, revealing over 30

genes down-regulated in T1D with the most characteristic gene ontology linked to immune response, regulation of cell cycle, cell surface receptor linked signal transduction and electron transport.

The link of hyperglycaemia to differential gene expression has been highlighted in a number of studies[191–193] including Irvine et al.[193] who identified a gene expression signature in PBMCs for pathways controlling cellular metabolism and survival, including endoplasmic reticulum and oxidative stress (e.g. induction of *HIF1A, DDIT3, DDIT4*, and *GRP78*) which was correlated to glycaemic control in the first year after diagnosis[193]. This is an important consideration when examining treatment-linked gene expression after the onset of dysglycaemia.

Table 2: Summary of clinical and preclinical gene expression studies in T1D and antigen specific immunotherapy in autoimmune disease.

Study	Summary
Whole blood	
Reynier et al. 2010 Genes and Immunity Whole blood RNA from recently diagnosed T1D (n=19), AAb+ siblings (n=20) and controls (n=20)	IFN response genes upregulated in AAb siblings compared to T1D and controls Genes mainly associated with metabolic changes consistent with insulin deficiency and hyperglycaemia upregulated in T1D vs. AAb+ siblings and controls Suggestions of acute changes in gene expression linked to hyperglycaemia.
Han et al. 2012 Whole blood RNA from at risk (n=19), new onset (n=33), longstanding T1D (n=59) and healthy controls (n=70)	Reduction in mRNA levels for innate and adaptive immunity genes in established T1D but not new onset or at risk. Higher CTLA4:CD3G mRNA expression in longstanding T1D compared to new onset and at risk populations.
Pruul et al. 2015 Molecular and Cellular Endocrinology[194] Whole blood RNA from 2 study groups Group I: Younger age group newly diagnosed T1D (n=49) vs. control (n=31) age range 1.4−21.4 years Group II: Older age group newly diagnosed T1D (n=18) and control (n=30) age range 22.0−78.4 years).	Young onset T1D showed lower expression of BLTA and down regulation of CD226 with upregulation of CD80, which was correlated with CTLA4 expression. Adult onset group showed a higher expression of CD86 and TGF-β compared to those with younger onset.
Peripheral mononuclear cells	
Kaizer et al. 2007 JCEM[192] PBMCs from newly diagnosed children with:	Upregulated in T1D: MYC Upregulated in T1D at diagnosis but resolved by 4 months of insulin treatment: IL1B, early growth response gene 3, and prostaglandin-endoperoxide

T1D (n=43, 10.1 years ±3.8), T2D (n=12, 14.0 years ±2.3) Healthy controls (n=24,11.3 years ±4.6)	synthase 2
Padmos et al. 2008 Diabetes [195] CD14+ monocytes from established T1D (juvenile (n=30) and adult (n=30) onset), LADA (n=30) and T2D (n=30) and healthy controls (n=49)	Two clusters of monocyte differential gene expression: Cluster 1: PDE4B-associated cluster. Linked to proinflammatory cytokine or compound genes. Identified in 60% of LADA, 28% of adult-onset T1D and 10% of juvenile- onset T1D. Cluster 2: FABP5-associated cluster. Linked to chemotaxis, adhesion, motility, and metabolism genes. Detected in 43% of juvenile-onset T1D, 33% of LADA and 9% of adult-onset T1D patients.
Beyan et al. 2010 Diabetes[196] CD14+ monocytes from monozygotic twin pairs discordant for childhood onset T1D (n=10 pairs), control twin pairs (n=pairs) and healthy controls (n=51)	Quantified gene expression of gene clusters identified by Padmos et al. [195] using qPCR. Altered monocyte gene expression profile was shared between discordant twins indicating shared gene-environment interaction. Both clusters were identified in discordant twins with dominance of an upregulated FABP5-associated cluster and downregulation of a PDE4B-associated cluster.
Irvine et al. 2012 Diabetes[193] CD14+ monocytes from newly diagnosed children (n=12) and healthy controls (n=6)	Distinct monocyte gene expression signature linked to stress-induced metabolic disturbance identified through microarray with qPCR validation showing: Upregulation of ER and oxidative stress responsive genes Downregulation of genes involved in cellular energy and metabolite production
Jin et al. 2013 Diabetes Care[197] PBMCs in T1D (n=928) vs. controls (n=922) Variable duration (0-67 years 12.8 ± 12.9) and presence of complications in T1D.	Upregulated in T1D *TGFB1, SELL, PSMB3, CD74, S100A8, S100A9* and *IL12A* Downregulated *GNLY, PSMA4 and SMAD7* The presence of adult onset disease versus juvenile onset did not lead to detectable differential gene expression although a number of inflammation and myeloid cell-linked genes were found to be differentially expressed in the presence of disease complications.
Ferreira et al. 2014 Diabetes[198] PBMCs from T1D (n=64), healthy donors	Identification of a type 1 interferon signature transiently present in genetically predisposed children before the

(n=87) and BABYDIET cohort (n=109, AAb+ (n=22) and AAB- (n=87))	development of AAb, but is absent in established T1D.
Takahashi et al. 2014 Gene[199] PBMCs from established T1D (n=11) and healthy controls (n=9)	44 differentially expressed miRNAs identified between T1D and controls. Predictive target genes were compared to know literature to reveal 22 and 12 annotated KEGG pathways for the induced and repressed miRNAs, respectively. The majority of those pathways were noted to have previously been associated with T1DM: pathways in cancer, chemokine and insulin signalling pathways, axon guidance, and endocytosis.
T-cells	
Orban et al. 2007 Journal of Autoimmunity[190] CD4+ T-cells from newly diagnosed T1D (n=5), T2D (n=5) and healthy controls (n=5).	Downregulation in T1D compared to T2D and healthy controls in genes with functions in immune signalling, regulation of cell cycle, cell surface receptor linked signal transduction and electron transport. Reduced HDAC1 gene and protein expression confirmed by qPCR and Western blot. Reduced cell surface expression of P-selectin and LFA-1 confirmed by flow cytometry.
Hisanaga-Oishi et al. 2014 Endocrine Journal[189] T-cells in presence anti-CD3/CD28 and IL-2 +/- TGF-β in mouse models.	In NOD vs. B6 (TGF-β-) (n=12 in each group). Upregulated: *Clec16a, Ctla4, Dlk1, Il2ra, Il2rb, Pparg, Ptpn2, Rnls, Sh2b3, Ubash3a,* and *Vav3 expression.* Down-regulated: *Il2, Ptpn22,* and *Zac1.* In NOD vs. B6 (TGF-β+) (n=12 in each group). Upregulated: *Clec16a, Dlk1, Erbb3, Foxp3, Ikzf1, Il2rb,* and *Rnls.* Downregulated: *Glis3* and *Ptpn22.*
Antigen-specific T-cells	
Luce et al. 2011 Diabetes[200] Preproinsulin specific CD8+ T-cells isolated via tetramers and single cell PCR. HLA-A*0201 positive individuals in the following categories: adult newly diagnosed T1D (n=25), long-standing T1D (n=24).	No significant difference in gene expression levels by PPI-specific CD8+ T-cells between recent-onset and long-standing type 1 diabetic patients. Compared to CMV specific T-cells, PPI-specific T-cells showed: Downregulation of perforin, GZMA, GMZB, RANTES, MIP1b, PD1. Upregulation of CCR7. The authors speculated that this differential expression represents a central memory (TCM) phenotype of PPI specific CD8+ T-cells over naïve (TN) phenotypes, which was supported by cell surface markers on flow

	cytometry.
Serum	
Nielsen et al. 2012 Experimental Diabetes Research[201] Pooled serum samples from recently diagnosed T1D children (n=426) and controls (n=151), with subsequent qPCR (n=54) from the subsets.	Twelve upregulated human miRNAs in were identified in T1D patients; several of these were linked to apoptosis and β-cell networks. A specific miRNA (miR-25) was negatively associated with residual beta-cell function, and positively associated with glycaemic control.
Post antigen-specific immunotherapy	
Sthoeger et al. 2009 Journal of Autoimmunity[202] PBMCs from lupus patients treated with weekly tolerogenic peptide, hCDR1 (n=5) vs. placebo (n=4)	In peptide treated group downregulation of IL-1β, TNF-α, IFN-γ, and IL-10, BLyS and of the pro-apoptotic molecules caspase-3 and caspase-8. Upregulation of TGF-β and FOX-P3.
McPherson et al. 2014 eLife[123] CD4+ T-cells from mouse models of EAE treated with an altered peptide of myelin basic protein, Ac1-9 (4Tyr).	PIT exposed Teff had prolonged PD-1 expression with complete demethylation at CR-C and significantly reduced DNA methylation at the CR-B region in the Pdcd1 promoter.
Burton et al. 2014 Nature Communications[203] CD4+ T-cells from transgenic mouse model of EAE treated with escalating does of myelin basic protein, Ac1-9 (4Tyr).	Escalating doses of PIT led to: - Upregulation of negative co-stimulatory molecules including *Lag3, Tigit* and *Havcr2* (TIM-3) over the first 6 doses and then maintained with further doses. Also upregulation of *Il-10* and genes linked to IL-10 production i.e. *Icos*, *Il-21, Maf, Nfil3* and *Ahr.* - Upregulation of cell cycle and genome integrity transcripts during escalation of dose which repressed with extended high doses. - No changes in expression of *Pdcd1* (PD-1) and *Ctla4*. - Downregulation of *Btla* and *Cd274* (PD-L1) with extended treatment.
Islets	
Planas et al. 2009 Clin Exp Immunol[185] Pancreas samples from T1D donors at 5 days, 9 months, 8 and 10 years after diagnosis (n=4) with islets isolated from 2 of these. Antibody negative organ donors were used as controls (n=7).	Top differentially expressed genes were HLA and INS. No other genes found to be consistently differentially expressed in all 4 samples. The category with highest differential expression was immune response, in particular, antigenic presentation, chemotaxis, innate immunity and inflammation, complement, immunoregulation, adhesion molecules, interferon response and leucocytes. Also seen was a reduction in neural and exocrine gene transcripts.

Hopfgarten et al 2013. Acta Diabetol[186] Islets from an adult T1D organ donor (age 40 years and AAb-) who died at diagnosis and from an adult healthy volunteer	Increased expression of MCP-1 (a monocyte recruiting chemokine) and MHC class I in T1D donor and increased transcription factor ATF4, indicating increased ER stress. No signs of apoptosis on gene expression level.
Richardson et al 2016 Diabetologia[187] Pancreas samples from T1D patients (n=26) and controls (n=17)	Hyperexpression of HLA class I genes in T1D islets with elevation of STAT1 expression. HLA class I overexpression detectable in disease duration of up to 11 years and declining thereafter.
Rui et al. 2016 Diabetologia[188] Islet and β cells from NOD mice and humans	Changes gene expression of proinflammatory cytokines in whole NOD islets compared to a baseline at 3 weeks, showed a sequential rise in *Cxc110* mRNA, *Tnf* mRNA *Ifng* mRNA and *I11b* mRNA. NOD β-cells cultured with IL-1β and IL-6 or IFN-γ showed decreased expression of INS1 and INS2 as well as increased methylation.
Abbreviations: ATF4: Activating transcription factor 4, BlyS: B-lymphocyte stimulator, BTLA: B and T Lymphocyte attenuator, CMV: cytomegalovirus, FABP5: Fatty acid binding protein 5, GZMA: Granzyme A, GZMAB: Granzyme B, HDAC1: histone deacetylase 1, IFN: interferon, KEGG: Kyoto Encyclopedia of Genes and Genomes, LFA-1: lymphocyte function-associated antigen 1, PD-1: Programmed cell death protein 1, PTGS2: prostaglandin-endoperoxide synthase 2, RANTES: Regulated on Activation, Normal T Expressed and Secreted, STAT1: Signal transducer and activator of transcription 1	

Examining gene expression changes in antigen specific therapy has been performed in mouse models[123,203] and human studies[202]. In a small study of lupus patients Sthoeger et al.[202] compared a tolerogenic peptide hCDR1 (n=5) to a placebo (n=4) administered weekly for 26 weeks. Peptide treatment was associated with downregulation of IL-1β, TNF-α, IFN-γ, and IL-10, BLyS and of the pro-apoptotic molecules caspase-3 and caspase-8 with upregulation of TGF-β and FOX-P3. The study was limited by its small size, heterogeneous use of medication including steroids and lack of age matching through the subjects and there are some disease specific factors limiting its interpretation to other autoimmune diseases, such as the presence of paradoxical elevation of IL-10 in lupus patients[204]. Studies of the response of an EAE mouse model to an altered peptide of myelin basic protein, Ac1-9 (4Tyr) showed epigenetic changes in methylation in two areas of the *Pdcd1* promoter[123]. This linked to sustained PD-1 expression within T effector cells. When using an escalating dose regimen in the same mouse model, sequential changes in gene expression can be seen with up-regulation of negative co-stimulatory molecules including *Lag3, Tigit* and *Havcr2* (TIM-3) over the first 6 doses, which was maintained with further doses[203]. There was also up-regulation of *Il-10* and genes linked to IL-10 production i.e. *Icos*, *Il-21, Maf, Nfil3* and *Ahr*, with down-regulation of *Btla* and *Cd274*

(PD-L1) with extended treatment. Interestingly there were no changes in expression of *Pdcd1* (PD-1) seen however cells expressing PD1 surface markers increased in number following PIT[203].

Gene expression assays such as microarray provide a snapshot of a biological system at a particular time point under certain conditions. The examples above illustrate the complexities of gene expression studies in heterogeneity and expanse of results available, the importance of study design in interrogation of relevant genetic changes and the question of interpretation and validation of a multitude of gene "hits". Whole blood analysis allows investigation of all cell subsets and requires less manipulation with samples being stored with RNA stabilisers to minimise RNA degradation. As such cryopreservation of multiple samples and longitudinal analysis is possible. The drawback of whole blood analysis is identifying the source of differentially expressed genes from the constituent cell subsets. Pinpointing this may involve isolating or stimulating of cell subsets, the process of which could alter gene expression artificially. Whilst there have been some limited data on gene expression changes in PIT, this is mainly in mouse models or using limited clinical cases. One of the aims of this thesis is to examine the genetic profile of activated CD4+ T-cells at recent diagnosis of T1D and after 6 months with or without the presence of C19-A3 peptide immunotherapy.

1.10 Aims of Thesis

The aims of this thesis are to:

1. Draw on the practical experiences of managing a multicentre immune intervention trial into Type 1 diabetes to identify areas of strengths and weaknesses in trial design and implementation.
2. Present the clinical findings from the MonoPepT1De trial with respect to:
    a. Safety data
    b. Clinical outcome data on C-peptide levels
3. Present and examine changes in TCR clonotypes and gene expression microarray in the presence of C19-A3 peptide immunotherapy and discuss potential of these changes in elucidating mechanisms of such therapy and to identify future biomarkers.

2 Materials and Methods

2.1 Cell culture medium

Serum: Human AB serum was purchased from PAA laboratories (Yeovil, UK). Fetal calf serum (FCS) was purchased from Invitrogen Life technologies (Paisley UK). Serum was heat inactivated by agitated incubation at 56^{o}C in a water bath and aliquots were stored at -20^{o}C. Prior to use, aliquots of serum were defrosted and added to culture medium before storage at 4^{o}C. Single batches for human and FCS was used for all experiments involving patient samples.

RPMI complete medium: RPMI-1640 GlutaMAX™ containing 1.25mM HEPES, supplemented with 1% penicillin/streptomycin and 1% Fungizone® (all Invitrogen Life technologies, Paisley, UK) stored at 4^{o}C.

2.2 Islet cell autoantibody assays

Serum islet cell autoantibody titres for IA-2, ZnT8 and GAD were analysed using ELISA kits (RSR, Cardiff, UK), performed by the Diabetes Research Network, Wales Laboratory.

2.3 C-peptide assays

Serum C-peptide levels were analysed using a two-site chemiluminescent assay (Invitron, Monmouth, UK). Urine c-peptide levels were quantified using an ELISA assay (Mercodia, Uppsala, Sweden) for interpretation as urine C-peptide creatinine ratio (UCPCR). Both assays were performed by the Diabetes Research Network, Wales Laboratory.

2.4 Genotyping

Determination of *HLA DRB1*0401* genotype was performed by the Diabetes Research Network, Wales Laboratory using a PCR based technique in a Allset+ Gold SSP *HLA-DRB1*04* detection kit (Invitrogen, Paisley UK).

2.5 Peripheral blood mononuclear cell isolation

During venepuncture whole blood was stored in heparinised Vacuette™ tubes (Greiner Bio-one, Gloucestershire UK) for isolation of peripheral blood mononuclear cells (PBMC). Samples were diluted 1:1 with RPMI complete medium. Diluted blood was layered onto Lymphoprep™ (Axis Shield, Oslo, Norway) and high-speed density gradient centrifugation

was performed at 1000xg for 30min. PBMC were carefully isolated from the subsequent buffy coat using a Pasteur pipette, diluted with RPMI complete medium and centrifuged at 300xg for 10min. Resulting PBMCs were then re-suspended in fresh RPMI complete medium and centrifuged at 200xg for 10min. Cell number and viability were assessed by trypan blue (Sigma Aldrich, Poole, UK) dye exclusion using a haemocytometer. PBMCs were then placed in RPMI complete medium supplemented with 10% human AB serum.

2.6 Cryopreservation of PBMCs

For long-term storage isolated PBMCs were re-suspended in residual RPMI before drop-wise addition of ice-cold freezing medium (10% dimethyl sulfoxide (DSMO, Sigma-Aldrich, Dorset, UK) in AB serum) at a final concentration of 10-15x10^6cells/mL. Aliquots of 1mL were separated into Cryovials™ (Corning Inc. MA, USA) on ice before transfer to an alcohol free cell freezing container (Coolcell® BioCision, CA, USA) and freezing at -80^{o}C for 24 hours. Samples were then transferred to liquid nitrogen storage.

2.7 Thawing cryopreserved cells

Cryopreserved control PBMCs were used at each flow cytometry sort. Control cells were isolated from a haemochromatosis donor and cryopreserved as set out in section 2.6. These were removed from liquid nitrogen storage and transferred on ice to a 37^{o}C water bath. When samples appeared partially thawed 0.5mL of RPMI at 37^{o}C was added drop-wise to cells before transferring to a 15mL falcon tube and diluted further with RPMI up to a final volume of 10mL. Cells were centrifuged at 200xg for 10 minutes at room temperature then re-suspended in 1mL RPMI. Cells were counted with the addition of trypan blue to assess cell viability before washing once more at 200xg for 10mins at room temperature.

2.8 Enzyme linked immunosorbent spot (ELISPOT) assay.

The U-CyTech T-cell ELISPOT assay (U-CyTech Biosciences, Utrecht, The Netherlands) was used for detection of antigen-specific IL-10 and IFN-γ production. Fresh PBMCs were dispensed into 48-well plates at a density of 2 × 10^6 in 0.5 ml in complete RPMI medium supplemented with 10% (v/v) human AB, with the addition of peptide to a final concentration of 10 µM and incubated at 37°C, 5% CO2, tilted by 5°. Control wells contained TC medium with an equivalent concentration of peptide diluent alone (Dimethyl sulfoxide (DMSO) Sigma Aldrich, Poole, UK), Pediacel or HA. The peptides or antigens used and their final concentrations are summarised in Table 3. On day +1, 0.5 ml pre-warmed TC medium/10%

AB serum was added, and on day +2, non-adherent cells were re-suspended using pre-warmed TC medium/2% AB serum, washed, brought to a concentration of 10^6/300 µl, and 100 µl dispensed in triplicate into wells of 96-well ELISA plates (Nunc Maxisorp; Merck Ltd., Poole, United Kingdom) pre-blocked with 1:10 blocking buffer in PBS and pre-coated with monoclonal anti–IFN-γ or anti–IL-10 antibody. After capture at 37°C, 5% CO2 for 16-18 hours, cells were lysed in ice-cold water, plates washed in PBS/Tween 20, and spots developed according to the manufacturer's instructions. Plates were dried and spots of 80–120 µm counted in a BioReader 3000 (BioSys, Karben, Germany).

Table 3: Details of ELISPOT assay antigen or peptides concentrations and suppliers

Peptide/Antigen	Final stock concentration	Volume of reconstituted peptide/ well	Supplier
PI	10 mg/ mL	1 µL (high,10 µg) 0.5 µL (med, 5 µg) 0.1 µL (low, 1 µg)	Nova Laboratories, Leicester UK
C19-A3	10 mg/ mL	1 µL (high,10 µg) 0.5 µL (med, 5 µg) 0.1 µL (low, 1 µg)	Thermo Scientific Loughborough, UK
C13-32	10 mg/ mL	1 µL (10 µg)	Thermo Scientific Loughborough, UK
C22-A5	10 mg/ mL	1 µL(10 µg)	Thermo Scientific Loughborough, UK
IA-2 752-75	10 mg/ mL	1 µL (10 µg)	Thermo Scientific Loughborough, UK
IA-2 709-36	10 mg/ mL	1 µL(10 µg)	Thermo Scientific Loughborough, UK
IA-2 853-72	10 mg/ mL	1 µL (10 µg)	Thermo Scientific Loughborough, UK
HA 306-318	10 mg/ mL	1 µL(10 µg)	Thermo Scientific Loughborough, UK

2.9 Antigen stimulation assay

Antigens used for the stimulation assay were batched and a single batch number was used throughout the trial.

Proinsulin (PI) C19-A3 peptide (GSLQPLALEGSLQKRGIV) prepared to Good Manufacturing Practice standards was supplied by Novo Laboratories (Leicester, UK) as lyophilised peptide and stored at -20^{o}C. The peptide was > 95% pure, endotoxin-free and showed no toxicity in formal testing in two species (guinea-pigs and mice). Stocks were diluted to 10mg/mL in phosphate buffered saline (PBS, Invitrogen Life Technologies, Paisley UK). Once diluted aliquots were stored at 4^{o}C.

Haemagglutinin (HA 306-318) (PKYVKQNTLKLAT) was purchased from Thermo Scientific (Loughborough, UK) as lyophilised peptide of >95% purity, and stored at -20^{o}C. Stocks were diluted to 10mg/mL in phosphate buffered saline (PBS, Invitrogen Life Technologies, Paisley UK). Once diluted aliquots were stored at 4^{o}C.

Staphylococcal Enterotoxin B (SEB) was purchased from Sigma Aldrich (Dorset, UK). Stocks were diluted to 1mg/mL concentration in phosphate buffered saline (PBS, Invitrogen Life Technologies, Paisley UK) and stored at -20^{o}C. Once diluted aliquots were stored at 4^{o}C.

Pediacel®, a penta-vaccine consisting of purified diphtheria toxoid, purified tetanus toxoid, acellular pertussis vaccine, inactivated poliovirus and *Haemophilus influenzae* type b polysaccharide, was obtained from Sanofi Pasteur MSD (Berkshire, U.K.) and used at 1 µL/mL.

Anti-CD28 and anti-CD40 monoclonal antibodies: details of these antibodies are provided in Table 4 below. A single batch of each of these antibodies was used for patient samples throughout the trial. Aliquots of antibodies were stored at -80^{o}C.

Table 4: Details of antibodies used in cell stimulating assay

Target	Isotype	Clone	Supplier
CD28	Mouse IgG1, κ	CD28.2	BD Bioscience, Oxford UK
CD40	Mouse IgG1, κ	G28.5	Biolegend, London UK

The table states the antibody's target antigen, isotype, clone and supplier.

2.9.1 Cell staining

Antibodies used in the staining of PBMCs for flow cytometry are listed in Table 5 below. Batch numbers for each antibody were kept consistent for the length of the trial for individual patients. Lab members have determined optimal antibody staining concentrations previously.

Table 5: Antibodies used for flow cytometry staining

Target	Isotype	Clone	Fluorophore	Supplier
CD3	Mouse IgG1, κ	SK7	APC-H7	BD Bioscience, Oxford, UK
CD4	Mouse IgG1, κ	RPA-T4	APC	BD Bioscience, Oxford, UK
CD154	Mouse IgG1, κ	TRAP1	PE	BD Bioscience, Oxford, UK
CD69	Mouse IgG1, κ	FN50	FITC	BD Bioscience, Oxford, UK
CD14	Mouse IgG2a	TüK4	Pacific Blue	Invitrogen, Paisley UK
CD19	Mouse IgG1	SJ25-C1	Pacific Blue	Invitrogen, Paisley UK

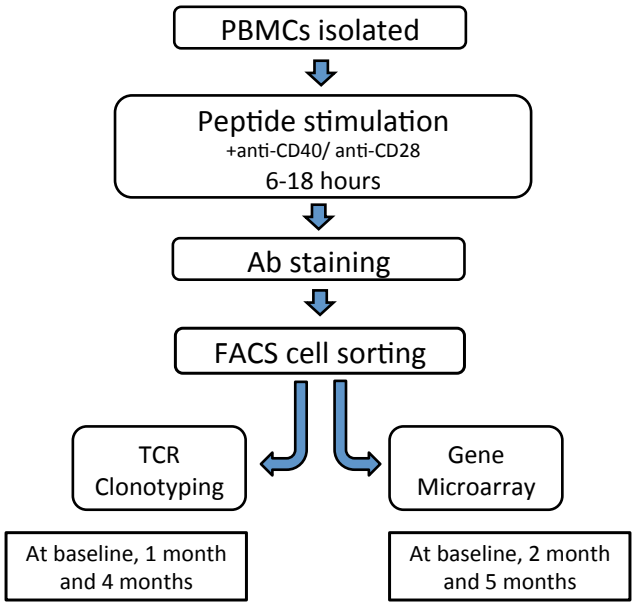
The table states the antibody's target antigen, isotype, clone and supplier. APC: Allophycocyanin, APC-H7: Allophycocyanin coupled to cyanine dye H7, PE: R-Phycoerythrin, FITC: fluorescein isothiocyanate.

LIVE/DEAD® Fixable Violet Dead Cell Stain - ViViD (Thermo Fisher Scientific, Paisley, UK) was used as a dead cell marker. A single batch was used throughout the study, stored in aliquots at -20^{o}C.

To sort live antigen-specific cells our method was based on a live-cell assay examining CD154 expression [205,206] with the addition of a co-stimulatory signal from anti-CD28 antibody . Samples were collected from trial patients at baseline, month 1, 2, 4 and 5 with captured cells allocated to either TCR clonotyping and/or gene microarray studies; the workflow is illustrated in Figure 10.

Figure 10: Diagram illustrating workflow for live cell sorts during MonoPepT1De trial.

TCR clonotyping and gene expression microarray were each performed at three timepoints.



PBMCs were isolated from whole blood using density-gradient centrifugation as outlined in section 2.5. Cells were re-suspended at a concentration of $4x10^6$ per mL in complete RPMI. Cells were cultured with anti-CD28 antibody (1µg/mL) and anti-CD40 antibody (1µg/mL) with the addition of one of the following as a stimulus: C19-A3 peptide (10µg/mL), HA peptide (10µg/mL), SEB (1µg/mL) or control PBS. For the cloning experiment (section 2.9.3) Pediacel® was used at a concentration of 1 µL/mL. In addition a control sample of cryopreserved PBMCs from a non-diabetic donor was stimulated with anti-CD28 antibody, anti-CD40 antibody and SEB (at the same final concentrations) to act as a longitudinal control. Cells were incubated for 6-18 hours at 37°C in 5% CO_2.

Cells were harvested using warmed complete RPMI supplemented with 2% AB serum and washed twice in cold PBS at 4°C centrifuging for 10min at 1400rpm. ViVid was used following manufacturer's instructions to stain dead cells before a further wash in cold PBS at 4°C, centrifuging for 10min at 1400rpm. Following this cells were stained using fluorochrome labelled antibodies (listed in Table 5). To improve uniformity of staining a mastermix of all antibodies was prepared before dispensing to each set of stimulated cells (Table 6).

Table 6: Antibody-Fluorochrome mastermix used for cell staining.

Antibody-fluorochrome	For C19-A3, HA and PBS stimulated cells ($16\text{-}20x10^6$ cells per condition)	For SEB stimulated cells ($2x10^6$ cells each for subject and control)
CD3-APC H7	64µL	22µL
CD4-APC	64µL	22µL
CD154-PE	80µL	33µL
CD69-FITC	80µL	33µL
CD14-Pac Blu	64µL	11µL
CD19-Pac Blu	64µL	11µL
	Total: 416µL made up to 460µL with 44µl PBS 140µL added to each condition	Total: 132µL made up to 160µL with 28µL PBS 70µL added to each sample

Once mastermix was made up it was briefly vortexed and centrifuged to avoid adding aggregates to cells. Following this the volumes stated were added to each sample. Cells were then incubated at 4°C for 20mins whilst on a MACSmix™ tube rotator (Miltenyi Biotec, Surrey, UK), before being washed twice in cold PBS at 1400rpm for 5min at 4°C and re-suspended for Fluorescence-Activated Cell Sorting (FACS) in cold FACS buffer (5% FCS (Invitrogen) in PBS).

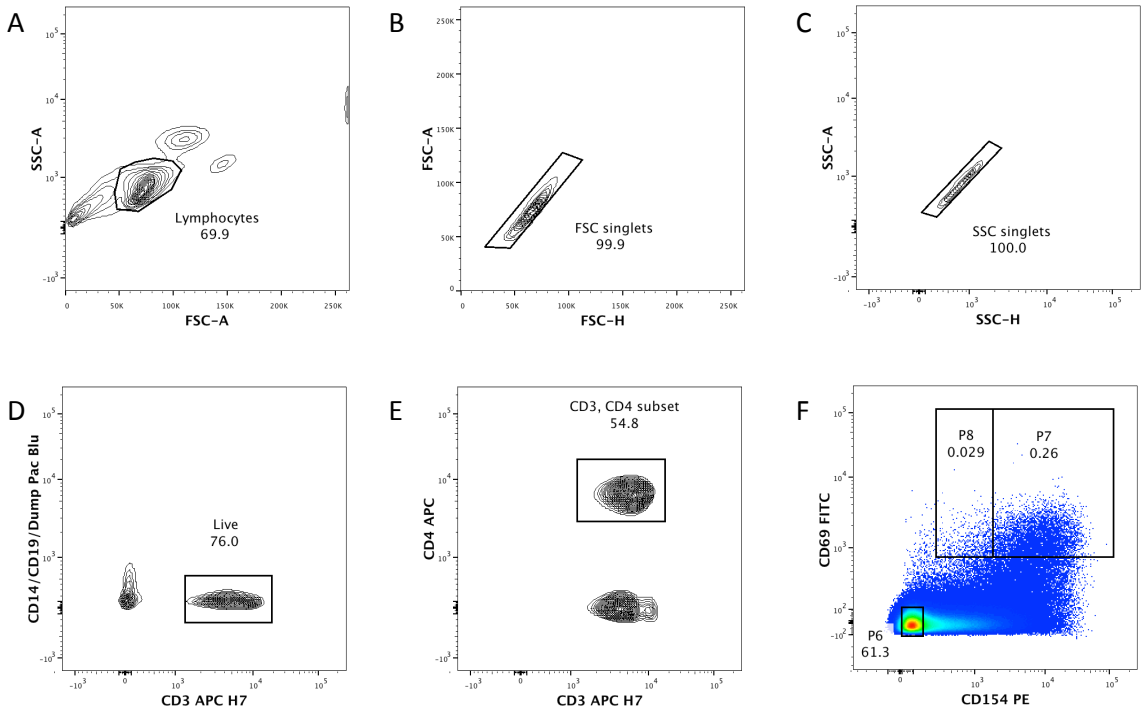
2.9.2 Cell sorting

Cells were sorted on a BD FACSAria II flow cytometer (Becton Dickinson, San Jose CA, USA) using 85µM nozzles. Two such flow cytometers were available for use and once assigned to a machine, each patient's trial samples were sorted on the same machine throughout the trial. $CD14^+$, $CD19^+$ and dead cells were excluded from the sort (dump channel). Cells were gated for CD3/CD4/CD69/CD154 positivity and sorted for gene expression studies (section 2.10) and/or TCR sequencing (section 2.11). In the case of TCR sequencing, samples were sorted into 1.7mL RNase-free microcentrifuge tubes containing 350µL RNeasy (RLT) buffer and stored at -80°C. For microarray samples were sorted into 0.2mL RNase-free PCR tubes and processed as outlined in section 2.10.1. Alternatively cells were single cell sorted for cloning (section 2.9.3).

Gating strategies are summarised in Figure 11 below.

Figure 11: Flow cytometry gating strategy for live cell sorting.

This figure shows example plots from flow cytometry sorting of cells stimulated with C19-A3. The plots illustrate (A) identification of a lymphocyte group and exclusion of doublets on (B) forward and (C) side scatter plots, (D) exclusion of CD3 negative and Pacific Blue positive (CD14, CD19 and dead cell marker exclusion), (E) isolation of the CD3+/CD4+ subgroup and (F) gating of the final CD69, CD154 plot to identify resting cells (P6), CD154+ CD69+ cells (P7) and CD154+CD69low cells (P8).



2.9.3 Cloning of T-cells by single cell sorting populations from Pediacel® stimulated PBMCs

$CD154^{+}CD69^{high}$ and $CD154^{+}CD69^{low}$ cells from a healthy donor stimulated with Pediacel (section 2.9) were single cell sorted into a NUNC™ 96 well U bottom plate containing $1.5x10^{5}$ irradiated mixed donor PBMCs in X-VIVO® media (Lonza, Castleford, UK) supplemented with 5% AB serum and 5% Cellkines (ZeptoMetrix, Belgium) and 4 µg/mL PHA-L (Biostat Diagnostic Systems, Stockport, United Kingdom). Cells were incubated at 37°C and maintained in X-VIVO® supplemented with 5% AB serum and 5% Cellkines, reducing the concentration of Cellkines by half every 5 days. Clones underwent alternate rounds of expansion every 2 weeks with irradiated mixed donor or matched donor PMBCs ($1.5x10^{5}$ cells), containing 4 µg/mL PHA-L or 1µL/mL Pediacel®, respectively. Once cells had rounded and ceased

proliferation, antigen specificity of clones was tested under 4 conditions: i. Clone only (negative control), ii. Clone and HLA matched PBMCs, iii. Clone, HLA matched PBMCs and Pediacel and iv. Clone and anti-CD3 anti-CD28 dynabeads (Life Technologies, Paisley, UK) (polyclonal stimulus/positive control). At 48 hours of culture, supernatants were removed for cytokine analysis and 0.5 mCi/well tritiated thymidine [^{3}H] (PerkinElmer, Buckinghamshire) was added. The cells were harvested 18 hours later using a MicroBeta Trilux machine (PerkinElmer) to assess ^{3}H-Thymidine incorporation as a measure of proliferation. IFN-γ levels were measured in culture supernatants taken at 48 hours using a ready-SET-Go ELISA kit (eBioscience, Hatfield, UK) in accordance with manufacturer's guidelines.

2.10 Gene expression analysis

2.10.1 Sample preparation

For the gene expression analysis, SuperAmp^{TM} lysis buffer (Miltenyi Biotec, Bergisch Gladbach, Germany) was prepared following manufacturer's instructions by incubating the buffer at 42^{o}C for 10 minutes and mixing to dissolve any precipitate formed, following storage at -20^{o}C in 6.4µL aliquots.

CD154^{+} CD69^{+} cells were collected into 0.2mL RNase-free PCR tubes and processed according to sorting volume following manufacturer's instructions: for cells collected in volumes of less than 1µL, 6.4µL of prepared SuperAmp^{TM} buffer was added directly to the sample. For samples with a volume of between 1µL and 200µL (up to 10,000 cells), cells were pelleted by centrifuging at 300xg for 5 minutes at 4^{o}C. Supernatant was carefully aspirated leaving the cells in approximately 1µL, and 6.4µL of SuperAmp^{TM} was added. All samples were mixed by pipetting up and down twice before incubation at 45^{o}C for 10 minutes. Samples were stored at -20^{o}C until further processing. The cell sample collection had an upper limit of 10,000 cells. Samples were shipped on dry ice Miltenyi Biotec (Bergisch Gladbach, Germany) where amplification using their SuperAmp technology and subsequent Aligent Whole Genome One Colour Arrays were carried out. The protocols used are briefly described below.

2.10.2 SuperAmp RNA amplification and cDNA conversion

SuperAmp RNA amplification was performed according to an undisclosed procedure by Miltenyi Biotec. Briefly, the amplification is based on a global PCR protocol using mRNA-derived cDNA. mRNA was isolated via magnetic bead technology. No adjustments were made

for cell number. Amplified cDNA samples were quantified using the ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA). cDNA integrity was checked using a Agilent 2100 Bioanalyzer platform (Agilent Technologies, Santa Clara, CA, USA). The results of the bioanalyzer run were visualised via gel images and the integrity of the cDNA was checked via the Agilent 2100 Bioanalyzer platform (Agilent Technologies). The results of the Bioanalyzer run are visualized in a gel image and an electropherogram using the Agilent 2100 Bioanalyzer expert software.

2.10.3 Aligent Whole Human Genome Oligo Microarray

A total of 96 samples were sent for microarray analysis on Agilent Whole Human Genome 8x60K v2 microarray format using a 1-color approach. 250 ng of each of the cDNAs were used as template for Cy3 labeling which was performed according to Miltenyi Biotec's undisclosed protocol. The Cy3-labeled cDNAs were hybridized overnight (17 hours, 65°C) to an Agilent Whole Human Genome Oligo Microarrays 8x60K v2 using Agilent's recommended hybridization chamber and oven. No adjustment was made for cell number at this stage. In order to avoid batch effects, samples were randomised across 13 slides each with 8 samples in 4 batches, to assign samples randomly across slides, positions and batches. Microarrays were washed once with Agilent Gene Expression Wash Buffer 1 for 1 minute at room temperature followed by a second wash with preheated Agilent Gene Expression Wash Buffer 2 at 37^{o}C for 1 minute. Cy3 fluorescence signals of the hybridised microarrays were detected using Agilent's Microarray Scanner system (Agilent Technologies) with Agilent Feature Extraction Software used to read out and process image files.

2.11 T-cell receptor β-chain sequencing

The TCR sequencing method is based on that set out by Quigley et al. 2011[207] which uses a template-switch anchored reverse transcription-polymerase chain reaction (RT-PCR) for unbiased analysis of TCR populations. The polyA tails of RNA are bound by oligo (dT) primers (5' CDS) and a complementary cDNA strand is generated by SMARTscribe reverse transcriptase (RT). SMARTscribe RT exhibits terminal transferase activity that adds 3-5 residues to the 3' end of the first strand cDNA. This elongated region allows binding of the SMARTerIIA oligonucleotide. This acts as a switching mechanism for the RT onto the 5' end of the RNA transcript, followed by rapid amplification of cDNA ends (5'RACE). Two rounds of nested PCR follow cDNA synthesis, initially using a universal primer mix (UPM) to bind the SmarterOligo and subsequently gene-specific primers (GSP) for the TCR β-chain, which are

nested within the original cDNA product. These processes are summarised in Figure 12 below. A negative control is introduced at each step. This allows visualisation of contamination at gel electrophoresis and tracking to point of introduction. Once PCR rounds are complete, amplicons can be ligated into a cloning vector that is transformed into competent *E. coli*. Upon culture blue/white colony discrimination can identify recombinant colonies with successful ligation. Alternatively PCR products can be gel extracted for high-throughput sequencing through Roche 454 GS FLX technology (Figure 13). A detailed protocol is described for each of these steps. Primer sequences used for TCR sequencing are described in Table 7.

Figure 12: TCR sequencing using nested polymerase chain reactions (based on Quigley et al. 2011[207]).

The figure illustrates the stages of cDNA synthesis via 5'RACE followed by 2 steps of PCR amplification using nested primers. GSP=Gene specific primer. UPM=Universal Primer Mix

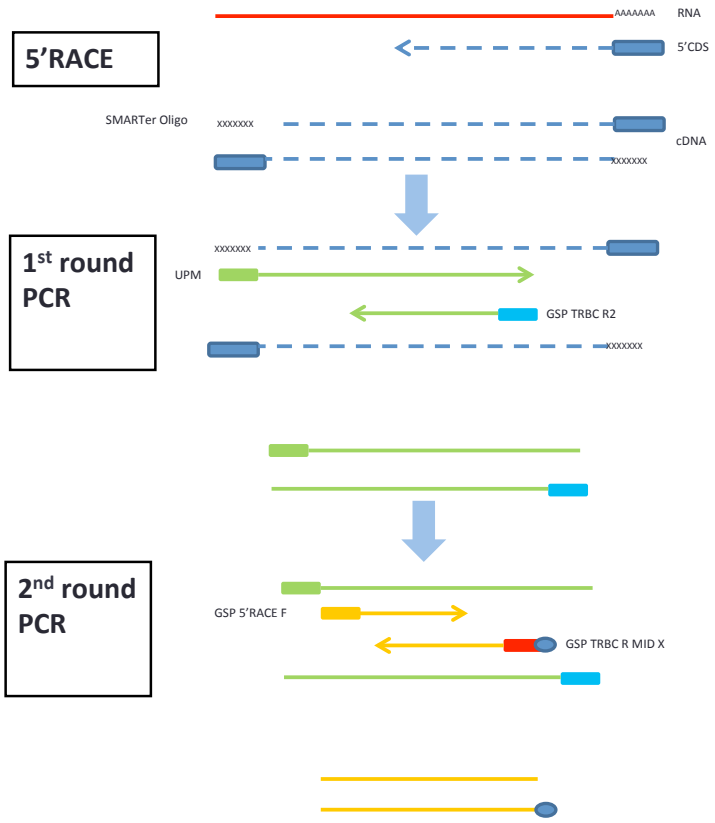


Figure 13: T-cell receptor sequencing workflow showing common and divided methodology between Sanger sequencing and high-throughput sequencing.

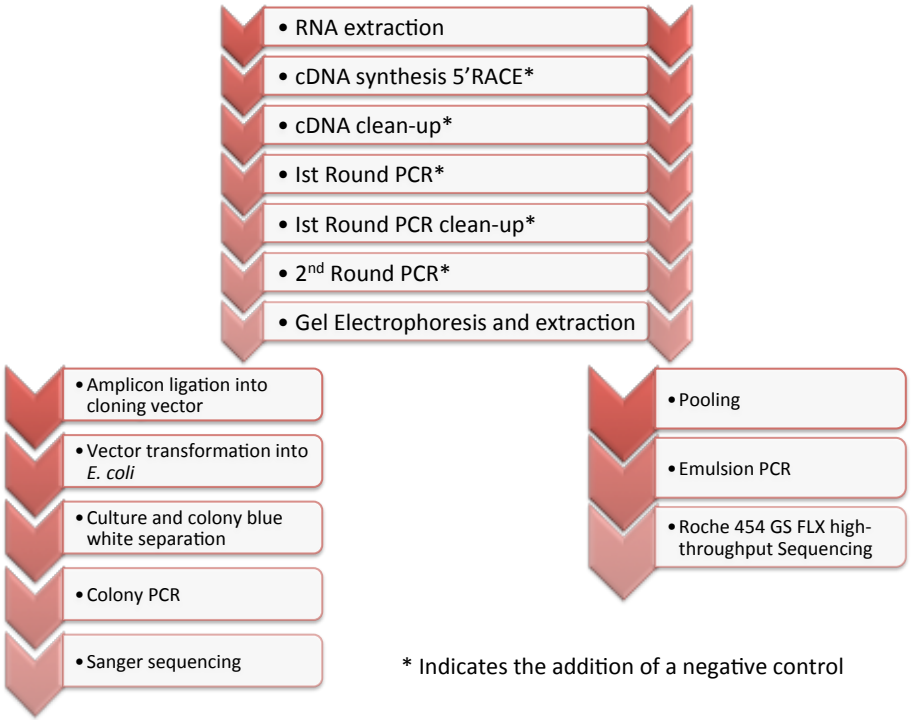


Table 7: Primer list and sequences used in amplification steps during TCR sequencing.

Sequencing step	Primer name	Primer sequence
5'RACE	5'CDS	5'-(T)25XY-3' (X=A, C, G; Y= A, C, G, T)
	SmarterOligo	5'–AAGCAGTGGTATCAACGCAGAGTACXXXXX–3' (X = undisclosed base in the proprietary SMARTer oligo sequence)
First round PCR	Universal Primer Mix (UPM)	5'CTAATACGACTCACTATAGGGC3' 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT-3'
	TRBC-R2	5'-GCTGACCCCACTGTGCACCTCCTTCCC-3'
Second round PCR	5'RACE F	5' AAGCAGTGGTATCAACGCAGAGT-3'
	MID X-TRBC-R	5'-XXXXXXXXXXCACACAGCGACCTCGGGTGGGAACAC-3' Full list of MID sequences are detailed in Table 11

2.11.1 RNA extraction

RNA was extracted from sorted cells using the RNeasy Micro Kit (Qiagen, Limburg, Netherlands) using manufacturer's protocol. Samples were defrosted on ice and all steps were carried out at room temperature. One volume (350µL) of 70% molecular biology grade ethanol was added into the cell lysate and mixed well before centrifugation through an RNeasy Minelute spin column. Flow-through was discarded. 350µL of Buffer RW1 was added and centrifugation repeated. 10µL of DNase I was added to 70µL RDD buffer and added to the column membrane for 15min. The buffer RW1 wash step was repeated before the addition of 500µL RPE buffer to the spin column and centrifuged again. 500µL of 80% ethanol was centrifuged through the column before a dry spin to remove remaining liquid from the membrane. Finally RNA was eluted with 14µL of RNase-free water and stored at -80^{o}C.

2.11.2 cDNA synthesis via reverse transcription using 5'RACE

RNA samples were taken directly from RNA extraction step or defrosted on ice from frozen. cDNA synthesis was performed using the SMARTer^{TM} RACE cDNA amplification kit (Clontech/Takara Bio, Saint-Germain-en-Laye, France). Samples were processed in triplicate and a negative control was added. Reagents and conditions are summarised in Table 8.

Table 8: cDNA synthesis by 5'RACE

	Volume added per reaction
RNA or RNase free water	3.25µL
5'CDS	1µL
72^{o}C for 3min then 42^{o}C for 2min before addition of components below.	
5x First strand buffer	2µL
DTT	1µL
dNTP mix	1µL
RNAse inhibitor	0.25µL
SMART RTase	1µL
SMARTerOligo	1µL
42^{o}C for 120min and 70^{o}C for 10min	

After amplification the triplicates were combined and an equal volume of Tricine EDTA was added to this and the control sample.

2.11.3 cDNA clean-up

This step used the Nucleospin Gel and PCR clean-up kit (Fisher Scientific, Loughborough, UK) as per manufacturer protocol; at this point another negative control was added consisting of 60µL of Tricine EDTA. Briefly, a 2:1 ratio of NT binding buffer was added to each sample, which was then transferred onto a Nucleospin column to bind DNA to membrane and centrifuged at 11,000xg for 30sec. The cDNA was washed twice with 700µL of NT3 buffer using a 30 sec 11,000xg spin with each wash. The column membrane was dried with a 1min centrifugation at 11,000xg. Clean cDNA was eluted with 20µL of NE buffer by incubating at room temperature for 1 min and centrifuging for 1min at 11,000G. Samples were stored at -80^{o}C or used to proceed to first round PCR.

2.11.4 First round PCR

A further negative control was added at this stage and reactions were set-up as outlined in Table 9 and amplified on a thermal cycler under conditions described in Table 10. cDNA yield was variable therefore PCR cycles were adjusted to sample cell number. Universal primer mix (UPM) consists of a long and a short primer that allows extension from the adapter incorporated during 5'RACE. TRBC-R2 acts as a reverse primer binding to the constant region of the β chain CDR3. Phusion High Fidelity DNA Polymerase was supplied by Fisher Scientific (Loughborough, UK).

Table 9: First Round PCR reaction composition.

	Stock concentration	Per 50µL reaction	Final concentration
5x Phusion HF Buffer	5x	10.0µL	1x
dNTP mix	10mM	1.0µL	0.2mM
UPM (Forward Primers)	10x	5.0µL	1x
TRBC-R2 (Reverse Primer)	10µM	1.0µL	0.2µM
5'RACE cDNA product (or negative control)	-	5.0µL	
Phusion polymerase	2000 units/mL	0.5µL	1 unit/50µL
RNase free water	-	27.5µL	

Table 10: First round PCR reaction conditions

	Temperature (^{o}C)	Time (Seconds)	Cycle number
Initial denature	98	30	
Denature	98	10	5
Annealing	70	30	
Elongation	72	40	
Denature	98	10	24
Annealing	65	30	
Elongation	72	40	
Final Elongation	72	60	

2.11.5 PCR product clean-up

This step was carried out using the Minelute PCR purification kit (Qiagen, Limburg, Netherlands) as per manufacturer protocol with the further addition of a negative control. Briefly, 5 volumes of binding buffer PB were added to the PCR reaction, placed onto spin column and centrifuged for 1min at 13,000rpm. Bound DNA was washed with 750µL of buffer PE by centrifugation for 1min at 13,000rpm, and the membrane was spun dry for 1 min at 13,000rpm. The PCR product was then eluted with 10µL of EB buffer after standing for 1min and then centrifuging for 1min at 13,000rpm. The clean product was stored at -80^{o}C or taken directly to second round PCR.

2.11.6 Second round PCR

Multiplex Identifier (MID) labelled primers were incorporated at this stage. These were designed to contain a unique 10-base barcode or identifier sequence, followed by a 'CA' linker between barcode and the TRBC-specific sequence (see Table 11). MID sequences were provided by Roche (Hertfordshire, UK). Primers were synthesised with salt-free purification by Eurofins Genomics (Ebersberg, Germany). MIDs were re-suspended in 1xTE buffer (Invitrogen) to 100µM and further diluted to 10µM in nuclease free water.

Table 11: Multiplex identifier sequences incorporated into TRBC-R primer.

Underlined region identifies the unique 10 base MID label attached by a CA linker to the common primer sequence.

MID	Sequence 5'-3'
MID1-TRBC-R	ACGAGTGCGTCACACAGCGACCTCGGGTGGGAACAC
MID2-TRBC-R	ACGCTCGACACACACAGCGACCTCGGGTGGGAACAC
MID3-TRBC-R	AGACGCACTCCACACAGCGACCTCGGGTGGGAACAC
MID4-TRBC-R	AGCACTGTAGCACACAGCGACCTCGGGTGGGAACAC
MID5-TRBC-R	ATCAGACACGCACACAGCGACCTCGGGTGGGAACAC
MID6-TRBC-R	ATATCGCGAGCACACAGCGACCTCGGGTGGGAACAC
MID7-TRBC-R	CGTGTCTCTACACACAGCGACCTCGGGTGGGAACAC
MID8-TRBC-R	CTCGCGTGTCCACACAGCGACCTCGGGTGGGAACAC
MID9-TRBC-R	TAGTATCAGCCACACAGCGACCTCGGGTGGGAACAC
MID10-TRBC-R	TCTCTATGCGCACACAGCGACCTCGGGTGGGAACAC
MID11-TRBC-R	TGATACGTCTCACACAGCGACCTCGGGTGGGAACAC
MID12-TRBC-R	TACTGAGCTACACACAGCGACCTCGGGTGGGAACAC
MID13-TRBC-R	CATAGTAGTGCACACAGCGACCTCGGGTGGGAACAC
MID14-TRBC-R	CGAGAGATACCACACAGCGACCTCGGGTGGGAACAC
MID15-TRBC-R	ATACGACGTACACACAGCGACCTCGGGTGGGAACAC
MID16-TRBC-R	TCACGTACTACACACAGCGACCTCGGGTGGGAACAC
MID17-TRBC-R	CGTCTAGTACCACACAGCGACCTCGGGTGGGAACAC
MID18-TRBC-R	TCTACGTAGCCACACAGCGACCTCGGGTGGGAACAC
MID19-TRBC-R	TGTACTACTCCACACAGCGACCTCGGGTGGGAACAC
MID20-TRBC-R	ACGACTACAGCACACAGCGACCTCGGGTGGGAACAC
MID21-TRBC-R	CGTAGACTAGCACACAGCGACCTCGGGTGGGAACAC
MID22-TRBC-R	TACGAGTATGCACACAGCGACCTCGGGTGGGAACAC
MID23-TRBC-R	TACTCTCGTGCACACAGCGACCTCGGGTGGGAACAC
MID24-TRBC-R	TAGAGACGAGCACACAGCGACCTCGGGTGGGAACAC

Before using MID labelled TRBC-R primers within the second round PCR, relative efficiency of these primers were assessed through real-time PCR. The cDNA yield from first round PCR and subsequent clean-up was inadequate to provide adequate substrate for MID validation by real-time PCR experiments. Therefore a previously sequenced TCR β-chain was chosen as a substrate and underwent A'tailing, ligation, transformation into *E. Coli* and culture (following methods described in sections 2.11.8 to 2.11.11). Following culture the plasmid minipreparation was performed (section 2.11.12). Serial dilutions of the plasmid miniprep (1:10, 1:100, 1:1000, 1:10,000, 1:100,000) provided the cDNA for real-time PCR. Reactions were assembled in triplicate as summarised in Table 12, using SYBR Green mastermix (Life Technologies, Paisley UK). Reactions were carried out on Applied Biosystems 7900HT Sequence Detection System under conditions in shown in Table 13 and analysed using the associated software (version 2.3). Efficiency of MID primers was calculated using Ct values from using the equation: Efficiency (%) = $10^{(-1/slope)}$-1.

Table 12: Comparison of MID labelled primer efficiency, real-time PCR reaction composition.

	Stock concentration	Per 20µL reaction	Final concentration
SYBR Green Mastermix	2x	10µL	1x
5'RACE (Forward primer)	10µM	0.4µL	0.2µM
MIDX-TRBC-R	10µM	0.4µL	0.2µM
cDNA		0.5µL	
Nuclease-free water		8.7µL	

Table 13: Comparison of MID efficiency, real-time PCR conditions.

	Temperature (^{o}C)	Time (Seconds)	Cycle number
	50	120	
Activation of polymerase within mastermix	95	600	
Denature	95	15	40
Anneal and extend	60	60	40

Once relative efficiency of MID-labelled primers had been assessed, these primers were allocated to samples for second round PCR. Second round PCR composition and conditions are summarised in Table 14 and Table 15 respectively.

Table 14: Second Round PCR composition

	Stock concentration	Per 50µL reaction	Final concentration
5x Phusion HF Buffer	5x	10.0µL	1x
dNTP mix	10mM	1.0µL	0.2mM
5'RACE F (Forward primer)	10µM	1.0µL	0.2µM
MID X-TRBC-R (Reverse Primer)	10µM	1.0µL	0.2µM
PCR product (or negative control)	-	4.0µL	
Phusion polymerase	2000 units/mL	0.5µL	1 unit/50µL
RNase free water	-	32.5µL	

Table 15: Second Round PCR conditions

	Temperature (^{o}C)	Time (Seconds)	Cycle number
Initial denature	98	30	
Denature	98	10	5
Annealing	65	30	
Elongation	72	45	
Denature	98	10	20
Annealing	68	30	
Elongation	72	45	
Final Elongation	72	60	

2.11.7 Gel electrophoresis and extraction

cDNA bands were visualised on agarose gels made at a concentration of 2% (w/v; Invitrogen) dissolved in 1xTris-acetate EDTA buffer (TAE) (Invitrogen). Once melted and cooled SYBR safe DNA gel stain (Invitrogen) was added at a concentration of 1:10,000 before pouring into moulds.

Second round PCR products and negative controls were run alongside a 100bp ladder. Products of ~500bp were cut out for extraction and negative control lanes were screened for visible contamination. DNA was extracted from gel slices using the QIAquick Gel Extraction kit (Qiagen, Limburg, Netherlands). Briefly, 3 volumes of QG buffer were added to the gel and left to dissolve at room temperature. One gel volume of isopropanol was added and the solution centrifuged through a QIAquick spin column for 1min at 13,000rpm. A further 500µL of buffer QG was added to the column and centrifuged under the same conditions. The DNA was washed by adding 750µL of buffer PE, standing for 3min and centrifugation for 1 min at 13,000rpm. A further centrifugation step to dry the membrane followed for 1min at 13,000rpm. DNA was eluted using 10µL of buffer EB. At this stage samples continued to Sanger sequencing via methods set out in sections 2.11.8 to 2.11.11 or they proceeded to

high-throughput amplicon sequencing on the Roche 454 GS FLX via Eurofins MWG Operon (Ebersberg, Germany).

2.11.8 A'tailing PCR Product

A'tailing of extracted product was performed using protocol below (Table 16) and then taken forward into ligation. Advantage 2 polymerase was supplied by Takara Bio/Clontech (Saint-Germain-en-Laye, France).

Table 16: A'tailing of DNA reagents and conditions

	Stock concentration	Per 11.4 µL reaction	Final concentration
PCR cDNA product		8 µL	
Taq advantage 2 polymerase	50x	1.14µL	5x
Advantage 2 PCR Buffer	10x	1.14µL	1x
dATP	10mM	0.456µL	0.4
Nuclease-free water		0.664µL	
Mixture placed at 70^{o}C on heat block for 30 min			

2.11.9 Ligation into vector

A'tailed product was then ligated into pGEM® Easy Vector (Promega, Madison, USA) as follows (Table 17):

Table 17: Ligation reagents and conditions

	Stock concentration	Per 15 µL reaction	Final concentration
2x Ligation buffer	2x	7.5µL	1x
A'tailed cDNA product	-	4.0µL	-
pGEM Easy Vector	50ng/µL	1.0µL	3.3ng/µL
T4 ligase	3 Weiss units/µL	2.0µL	0.4 units/µL
Nuclease-free water		0.5µL	
Kept at room temperature for an hour and then 4°C overnight			

2.11.10 Transformation of vector into *E.coli*

A 2µL aliquot of ligation reaction was added aseptically into 50µL of One Shot® Top 10 chemically competent *E. coli* cells (Invitrogen) and incubated on ice for 15 minutes, followed by heat shocking for 45 seconds at 42°C in a water bath before placing back on ice for 2 min. Then 450µL of room temperature SOC medium was added. SOC medium was made up using 2% (w/v) peptone (Sigma Aldrich), 0.5% (w/v) yeast extract (Sigma Aldrich), 10mM NaCl, 2.5nM KCl, 10nM $MgCl_2$, 10mM $MgSO_4$ 20mM glucose (Sigma Aldrich) in distilled water with aliquot stored at -20°C.

The transformation mixture was then incubated at 37°C for 2 hours.

2.11.11 Bacterial culture and blue/white colony screening

Agar plates for bacterial culture were prepared using 0.035% (w/v) Lysogeny Broth (LB) agar (Sigma-Aldrich) dissolved in distilled water before autoclaving. Ampicillin (Sigma Aldrich) was added to the molten agar at 0.1mg/mL before being poured into sterile Petri-dishes (Corning Inc.) and allowed to set before storing at 4^{o}C. For blue/white colony distinction, plates were covered with 5-Bromo-4-chloro-3-indolyl β-D-galactoside, BCIG (X-Gal) (Sigma-Aldrich), 40µL of stock solution per plate.

Transformation reactions were diluted 1:10 and 1:20 concentrations in LB supplemented with ampicillin (0.1mg/mL), spread onto X-Gal agar plates and incubated overnight at 37^{o}C.

After incubation white colonies were individually picked and sent for Sanger sequencing by GATC® Biotech. TCR sequences were then processed using International Immunogenetics Information System® (IMGT) website.

2.11.12 Plasmid Minipreparation

Following incubation of transformation reactions (2.11.10) plasmid DNA was isolated using the QIAprep Spin Miniprep Kit (Qiagen, Limburg, Netherlands). In brief, 1.5ml of culture broth was transferred to an eppendorf tube centrifuged at 9000rpm at room temperature for 1 minute, the supernatant was discarded and the process repeated. 250µL of buffer P1 was added to the pellet and the mixture was vortexed until homogenised. 250µL of buffer P2 was added and mixed thoroughly by inversion. 350µL of buffer N3 was added and again mixed by inversion. The mixture was centrifuged at 13,000rpm for 10mins and the subsequent clear supernatant was placed onto spin columns provided in the kit. The column was spun at 13,000rpm for 1 min and the eluate discarded. 0.5ml of buffer PB was placed in the spin column and centrifuged at 13,000rpm for 1 min before discarding the eluate. 0.75ml of buffer PE was placed in the spin column and centrifuged at 13,000rpm for 1 min before discarding the eluate. The column was dried by centrifuging for 2 mins at 13,000rpm. 30µL of buffer EB was added to the column and rested at room temperature for 1 minute. Eluate from the column was collected by spinning at 10,000rpm for 1 minute.

2.11.13 High throughput TCR β-chain data analysis

TCR data was received in FASTA format following sequencing using Roche 454 GS FLX technology via Eurofins MWG Operon (Ebersberg, Germany). FASTA files then underwent high-throughput analysis via the Immunogenetics Information system (IMGT/High V-Quest)[208,209]. Subsequent data was sorted in Microsoft® Excel® for Mac 2011 (Version 14.4.8) to allow for processing data through KNIME Analytics Platform (Version 2.11.3, Konstanz, Germany). KNIME data was further analysed by a biostatisician using R language[210] with diversity, sharing and clonality analysis performed with the aid of the tcR package[211] and custom R scripts.

2.12 Statistical analysis

Statistical analysis was performed using Prism 6 for Mac OS X (Version 6.0g). Clinical data was analysed using D'Agostino and Pearson omnibus normality testing. Non-parametric clinical data underwent log transformation was performed to reduce skewness followed by parametric analysis. Correlation analysis was performed using Pearson correlation. Unpaired tests were used to compare means between groups (using Welch's correction if unequal standard deviation (SD) is present) with one-way ANOVA was used if multiple groups were present. Unpaired categorical data was analysed with the Chi-square test. For TCR diversity data unmatched non-parametric data for multiple groups was analysed with Kruskal- Wallis tests.

3 Recruitment and screening data

3.1 Introduction

Optimisation of potential β-cell salvage in T1D intervention studies relies upon narrow timeframes from diagnosis as set out by Immunology of Diabetes Society (IDS) guidelines[72]. From a patient perspective, the impact of a recent diagnosis of a chronic disease can be a significant barrier to committing to clinical research. For others clinical research can provide a positive focus post-diagnosis. Recruiting a young, mainly working, adult cohort into an intensive trial such as this can raise additional logistical barriers. Those of reproductive age may also be reluctant to delay or risk pregnancy.

My role within the trial was as study doctor for the London site, based at Guy's Hospital. My responsibilities included:

1. Identification of patients close to diagnosis, consenting, screening and recruitment into the study. To ensure local clinical teams were aware of the trial I coordinated presentations to my peers at numerous local meetings, targeting diabetes teams and educational meetings for diabetes registrars.

2. Regular weekly teleconferencing with the MonoPepT1De clinical trial management team at Cardiff University, representatives of each site and our own local sponsor representatives.

3. Public engagement regarding the research trial. This encompassed identifying and liaising with interested stakeholders, presenting at meetings organised by JDRF and DUK, web events and interviews for patient publications.

4. Liaising with an expert patient panel regarding recruitment within the trial.

5. Clinical management of enrolled patients throughout the course of the trial.

6. Coordinating blood samples from trial centres for processing, including managing priority access to flow cytometry facilities for my specific research projects.

7. Performing experiments involving antigen stimulation assays, for downstream TCR sequencing and gene expression analysis.

Pathways utilised in subject identification included liaison with NHS trusts local to each centre, local research networks and the national After Diabetes Diagnosis REsearch Support System-2 (ADDRESS-2) initiative (UK Clinical Research Network Identifier: UKCRN9689). ADDRESS-2 is a register aimed to aid recruitment into T1D intervention trials by coordinating and expediting referrals of new-onset T1D cases (and, if necessary unaffected siblings) from its 148 adult and paediatric sites to participating studies. On entry to the register, ADDRESS-2 subjects agree to be contacted regarding future research projects. The register is managed by the Diabetes Research Network (DRN) jointly funded by Diabetes UK and the Juvenile Diabetes Research Foundation, with support from the Department of Health via the NIHR Clinical Research Network: Diabetes and NIHR Clinical Research Network: Children. The MonoPepT1De study was also promoted via supporting charities, and national media. A dedicated trial website and Facebook page was used for online engagement and a commercial site, www.diabetes.co.uk was also used to promote the study.

3.2 Aims

1. Evaluate initiatives used to identify and recruit people within 100 days of diagnosis of Type 1 diabetes to a multicentre interventional trial.

2. Describe screening characteristics of enrolled patients and compare to published data on newly diagnosed cohorts.

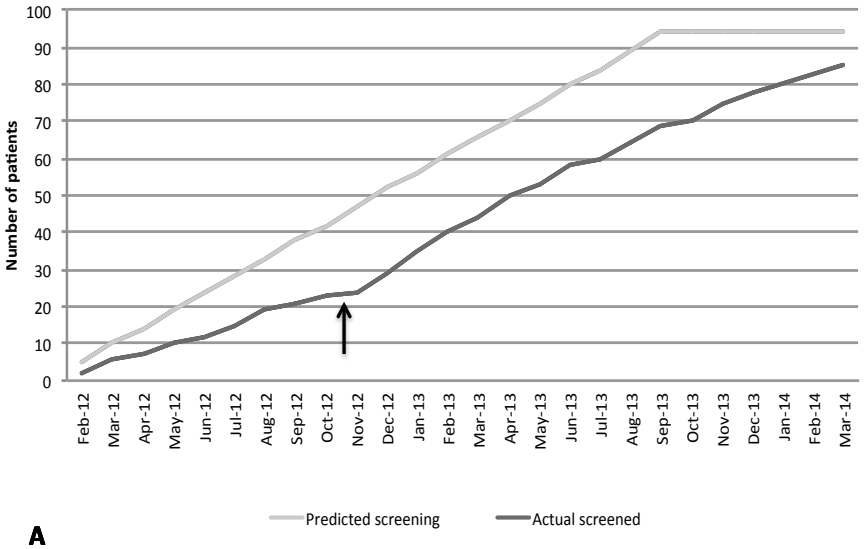
3.3 Patient identification and recruitment

Screening and recruitment for the MonoPepT1De trial began in February 2012 in 4 centres, Cardiff, London, Bristol and Newcastle. In January 2014, a fifth centre was added in Chester. A final total of 27 patients were randomised by end of recruitment in March 2014 (Figure 14). Three additional subjects were recruited over the original target of 24 due to missed doses in prior subjects, as specified according to the study protocol. Eligibility of referred patients for screening and subsequent recruitment is summarized in Table 18 for each of the five trial centres. Recruitment figures differ significantly between centres due to a number of factors including transport accessibility of sites, clustering of supporting referral centres (such as the 119 adult ADDRESS-2 sites) and Cardiff's role as coordinating referral centre. Direct calls from the public were triaged through the Cardiff centre, therefore ineligibility figures appear disproportionately high for this centre. A commercial website, www.diabetes.co.uk was used as a recruitment aid, on the basis of an established process for recruitment with NHS trusts

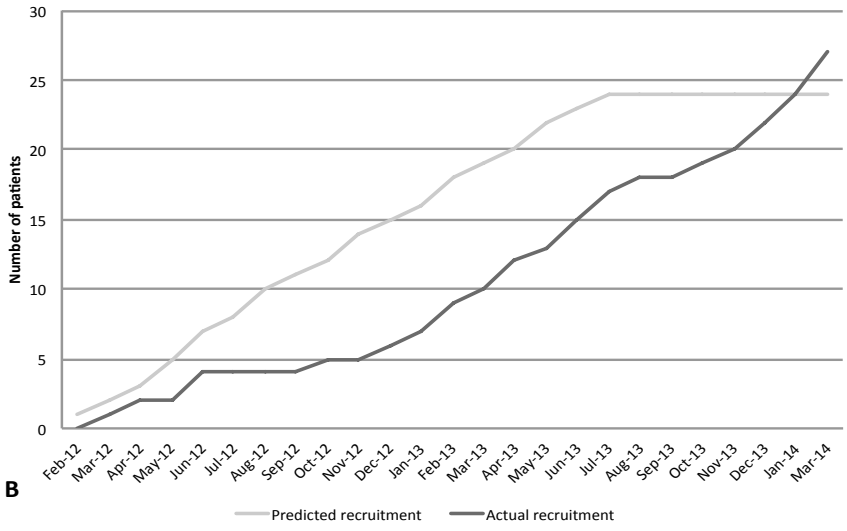
and high online interaction with Type 1 and Type 2 diabetes patients (800,000 individual user visits per month with 50,000 Facebook members). To target appropriate patients on the website details were published on Type 1 or newly diagnosed online fora and electronic newsletters. Subjects then expressed interest through an online form, self-reporting details for eligibility. Twenty-one enquiries were received, out of which 11 users wrongly reported eligibility and a high number were not contactable. None of these 21 subjects went onto screening.

Figure 14: Graphs showing expected and actual MonoPepT1De screening and recruitment rate.

Graphs show expected versus actual numbers of (A) screened and (B) recruited subjects over course of trial. Arrow in Graph A indicates timing of initiation of recruitment from ADDRESS-2 patient cohort.



A



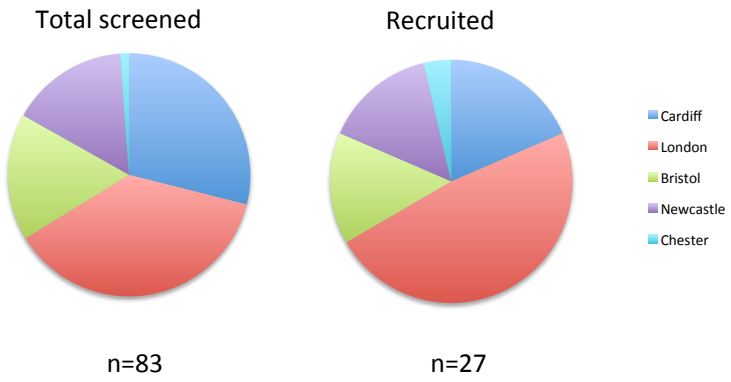
B

Table 18: Summary of recruitment eligibility and screening across 5 MonoPepT1De centres identifying factors for ineligibility and rates of screen failure.

Time expired refers to newly diagnosed subjects that could not be screened before the end of the 100 day window.

	Reason for ineligibility of referred patients												
	Time expired	Diagnosis long standing	Unable to contact	Age	Medically unsuitable / wrong diagnosis	Declined	Screen fail	Recruited	Total referrals	Total Eligible	Screened/ Eligible	Recruited/ Eligible	Recruited/ screened%
Cardiff	20	60	6	42	28	57	19	5	237	81	29.6	6.2	20.8
London	**4**	**11**	**4**	**7**	**12**	**45**	**18**	**13**	**114**	**76**	**40.8**	**17.1**	**41.9**
Bristol	1	3	1	7	2	27	10	4	55	41	34.1	9.8	28.6
Newcastle	1	1	0	4	0	15	9	4	34	28	46.4	14.3	30.8
Chester	0	1	2	0	0	6	0	1	10	7	14.3	14.3	100.0

Figure 15: Pie charts showing relative screening and recruitment across all 5 trial sites



3.4 Results

3.4.1 Expert patient panel feedback

In January 2013, a year after initiation of the trial as part of a patient and puplic involvement exercise, I presented an outline of the trial to a panel of expert patients in a “Dragon’s Den” type format. Whilst there was no structured interview, critical feedback was gained from the expert patient panel as a whole regarding trial structure, recruitment and screening. The response is summarised as follows:

- Acknowledgement of recruitment difficulties early after diagnosis. The psychological impact of the diagnosis of T1D was highlighted as an obstacle to enrollment. The panel felt the tight timeframe was a drawback, as subjects would have a greater inclination to consider research with interventional therapies when their diagnosis was more established. Only 36% (4) of the panel (n=11) would have agreed to trial screening at diagnosis compared to 64% (7) who would agree if offered the opportunity once the diagnosis was well established. In a young age group, work or study commitments can limit research involvement, the panel suggested considerations which may make the research easier such as:
    - Financial incentive
    - Flexibility with timing or location of research, such as weekend working and working from GP surgeries.
- Potential participant concerns over the trial itself being viewed as "meddling with the immune system" or for those needle-phobic having to add extra injections and blood tests to insulin injection.
- Importance of support for the newly diagnosed. The panel described experiences of isolation post-diagnosis and suggested a support system would provide reassurance to potential research candidates considering the trial. Such support could be provided formally from healthcare professionals or via volunteers with T1D already engaged in research through a "buddying" system. Alternatives suggested were the creation of a participant community either within the trial or within the wider ADDRESS-2 population. Networking through social media was felt to be an effective way to connect those requiring support.
- Improving the research experience for recruited subjects was discussed with strategies such as:
    - Educational support
    - A study or ADDRESS-2 quarterly newsletter to update on research progress
- Specific feedback regarding recruitment via ADDRESS-2 included:

- The need to stress to subjects that consent to joining the ADDRESS-2 enabled contact regarding studies, but importantly without obligation to take part. Once enrolled in ADDRESS-2 the panel felt there was no need for an interval following failed screening for a trial before another trial was offered, in fact such patients might be keen to enter trials promptly. Nonetheless repeated screening failures could lead to disappointment for subjects.

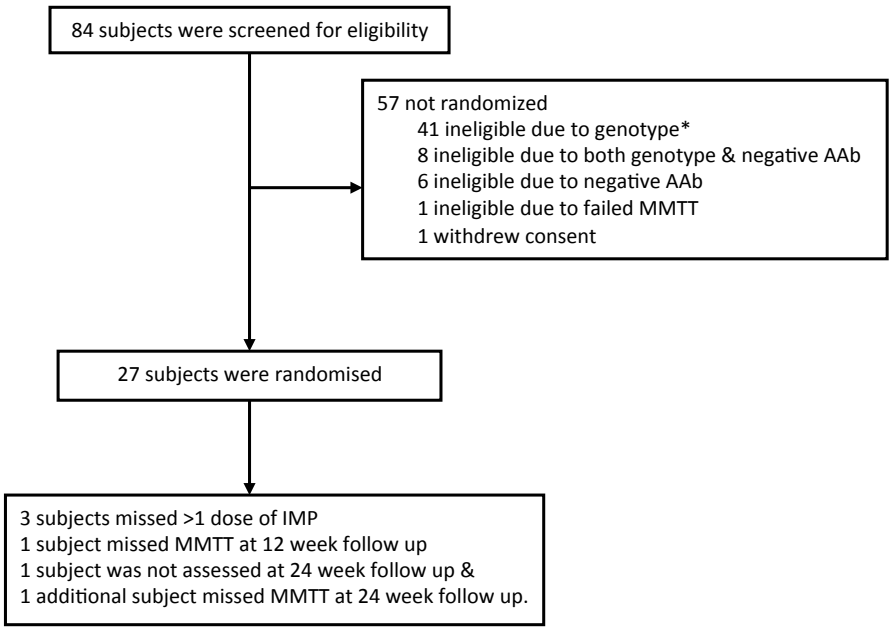
- Potential streamlining of trial recruitment by combining screening for a number of trials at one time point would be beneficial to patients and research teams.

3.4.2 Screening results

Trial screening involved 450 referred subjects across the 5 centres, with 52% of these eligible in terms of age, duration of disease and lack of exclusion criteria. Of those ineligible 11% (26) were newly diagnosed but outside of the 100-day window by the time screening would have been completed. Patient eligibility through the screening process and trial retention are summarised in Figure 16.

Figure 16: CONSORT flow diagram showing eligibility of patients in screening process and missed patient visits.

*IMP = investigational medicinal product, AAb = autoantibodies. *Two subjects excluded due to genotype were later found to be reported as falsely negative for HLA DRB1*0401.*

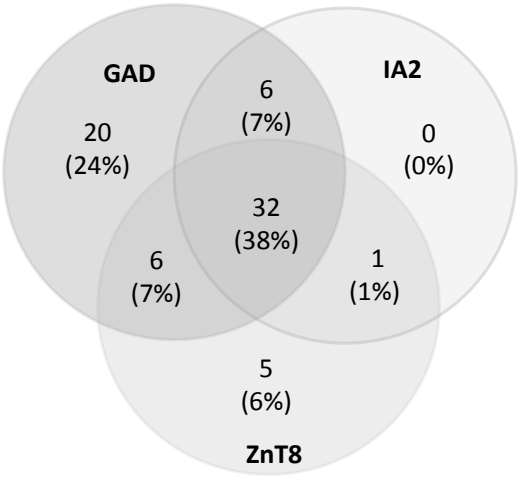


Mean (±SD) time from diagnosis to first screening visit was 46.8 (±23.75) days. Autoantibody screening was performed on 84 subjects and the results are summarised in Figure 17. An overview of genotype and autoantibody screening results are provided in Figure 18.

Figure 17: Breakdown of autoantibody status

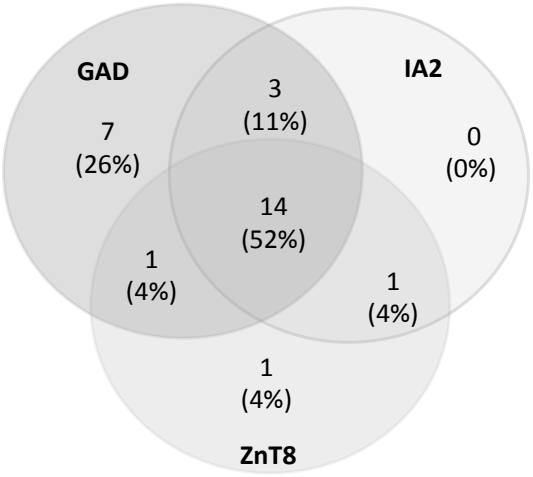
A. Screened population and B. Randomised population C. Data from Vaziri-Sani et al. 2010[212] *using the same autoantibodies to screen a 15-34 year old cohort of newly diagnosed patients D. Table showing frequency of each autoantibody in screened and randomised populations compared to Vaziri-Sani et al. data.*

A: Total screened population

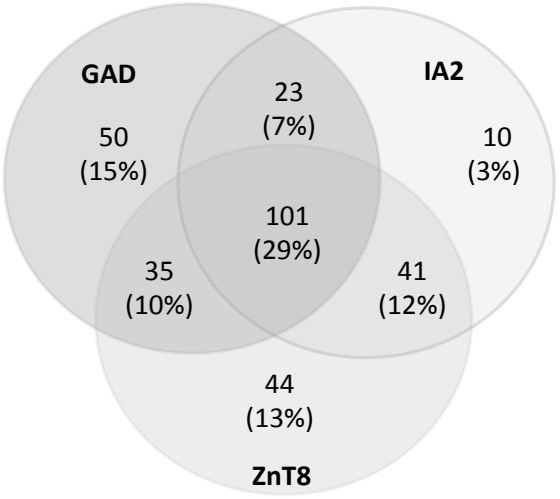


Auto-antibody negative: n=14 (17%)

B: Randomised population



C: Vaziri-Sani et al 2010



Auto-antibody negative: n=33 (10%)

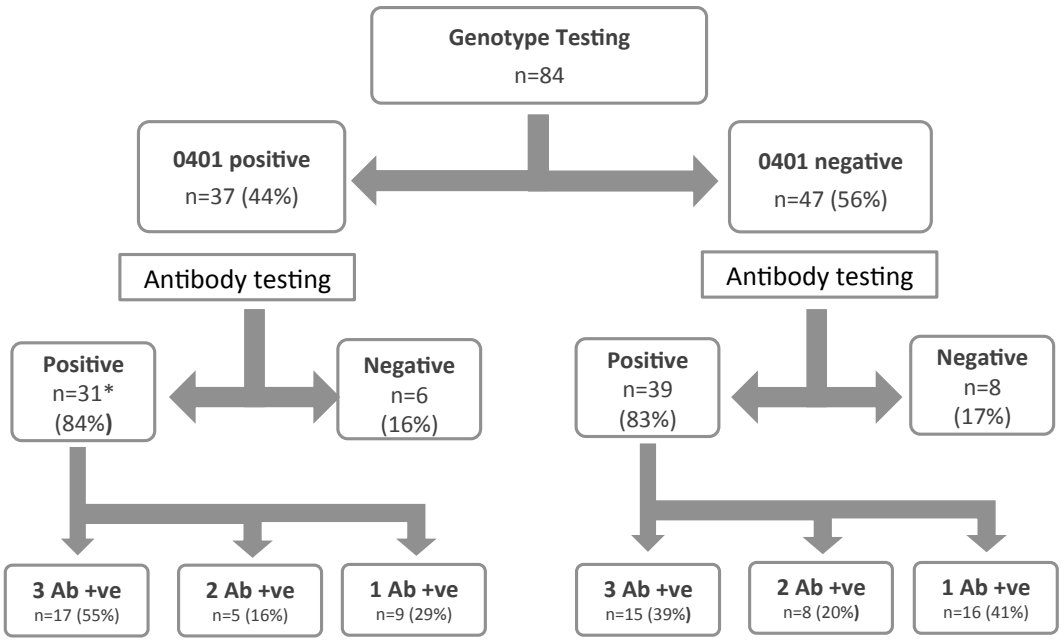
Figure 17: Breakdown of autoantibody status continued.

D: *Table showing frequency of each autoantibody in screened and randomised populations compared to Vaziri-Sani et al. data.*

	Screened n=84	Randomised n=27	Vaziri-Sani et al. 2010	Wenslau et al. 2010
Median and Range (Years)	29 (18-42)	28 (18-40)	25 (15-34)	20.3 (12.2-34.6)
GAD positive	76%	93%	61%	64%
IA2 positive	46%	67%	51%	73.8%
ZnT8 positive	52%	64%	64%	70.5%
Antibody negative	17%	-	10%	6.6%

Figure 18: Overview of screening outcome from first screening visit.

*The figure illustrates results of tests of autoantibody and genotype testing. *Out of the 31 eligible patients two subjects were mistakenly reported as false negatives, one failed the subsequent MMTT and a further subject withdrew.*



During the early window of 100-days post diagnosis, 96.4% (n=27) who had a MMTT surpassed the threshold of 0.2pmol/mL. Median and range of mean AUC of C-peptide at screening was 0.58 (0.238-2.805) pmol/mL/min. Mean AUC C-peptide between for males was (median (range)) 0.619(0.303-2.805) pmol/mL/min and females 0.525(0.238-0.992)

pmol/mL/min Figure 19. Correlates of mean C-peptide AUC were investigated and summarised in Table 19.

Figure 19: Graph showing distribution of mean AUC of C-peptide during 2 hour MMTT in all randomised patients and separated into male and female subjects.

Lines indicate median. No differences were seen between the means of log-transformed data between male and female subjects using an unpaired t-test.

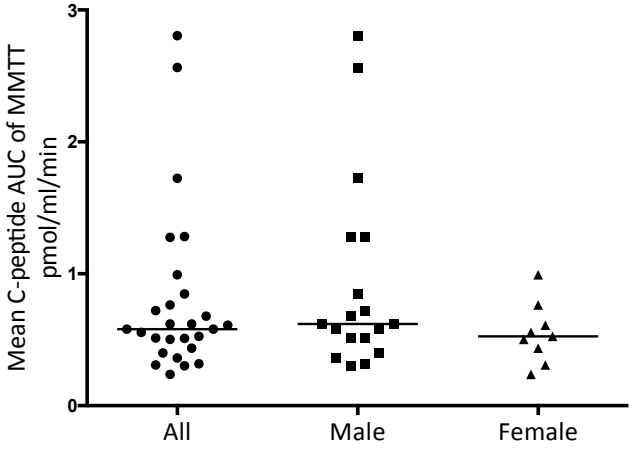


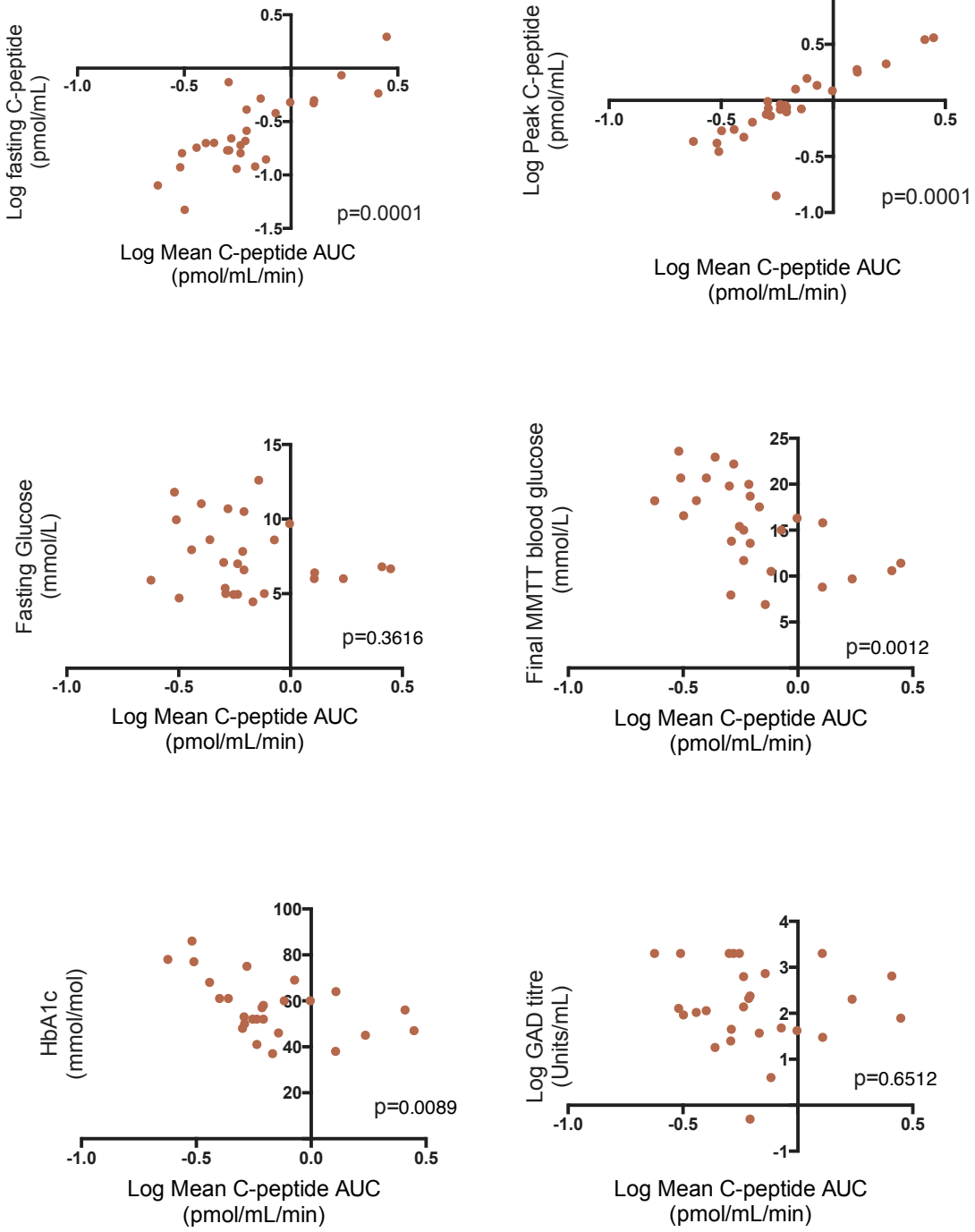
Table 19: Correlation of C-peptide mean AUC from MMTT with other baseline variables.

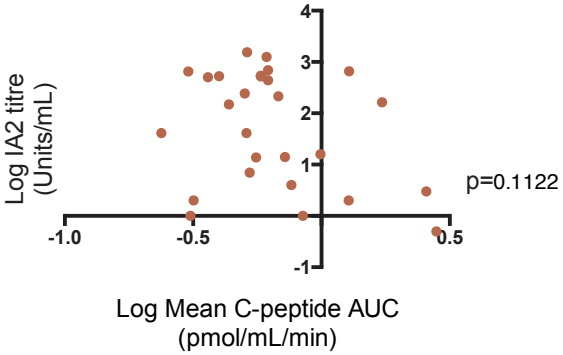
Variables were analysed for normality using D'Agostino and Pearson omnibus normality testing. Variable data which were not normally distributed were log transformed. Correlation was calculated using Pearson analysis

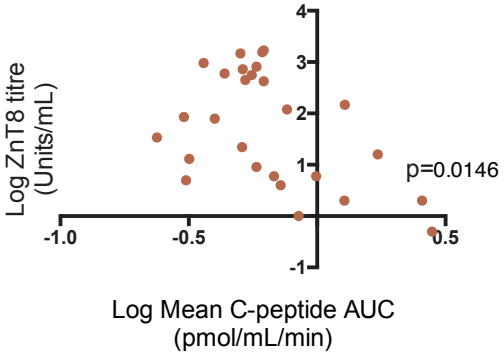
	C-peptide Mean AUC MMTT (pmol/mL/min)	
Fasting C-peptide (pmol/mL)	r (95% CI)	0.8002 (0.6037 to 0.9050)
	p value	**< 0.0001**
Peak C-peptide (pmol/mL)	r (95% CI)	0.8500 (0.6941 to 0.9297)
	p value	**< 0.0001**
Fasting glucose (mmol/L)	r (95% CI)	-0.1827 (-0.5263 to 0.2121)
	p value	0.3616
Final blood glucose (mmol/L)	r (95% CI)	-0.5897 (-0.7922 to -0.2701)
	p value	**0.0012**
HbA1c at V-1 (mmol/mol)	r (95% CI)	-0.5025 (-0.7449 to -0.1429)
	p value	**0.0089**
GAD titre (U/mL)	r (95% CI)	-0.008307 (-0.4555 to 0.2993)
	p value	0.6512
IA-2 (U/mL)	r (95% CI)	-0.3127 (-0.6192 to 0.07647)
	p value	0.1122
ZnT8 (U/mL)	r (95% CI)	-0.4646 (-0.7179 to -0.1026
	p value	**0.0146**
Insulin dose at baseline (U/kg)	r (95% CI)	-0.06812 (-0.4369 to 0.3203)
	p value	0.7356
Age at diagnosis (Years)	r (95% CI)	0.06747 (-0.3209 to 0.4364)
	p value	0.7381
BMI (kg/m2)	r (95% CI)	0.5000 (0.1480 to 0.7395)
	p value	**0.0079**
UCPCR (nmol/mmol)	r (95% CI)	-0.2213 (-0.5735 to 0.2001)
	p value	0.2988

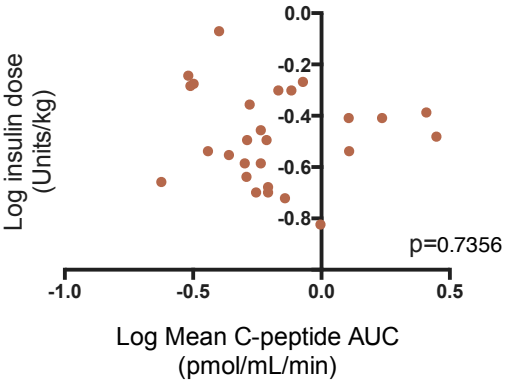
Figure 20: Graphs showing correlation of variables against mean C-peptide AUC.

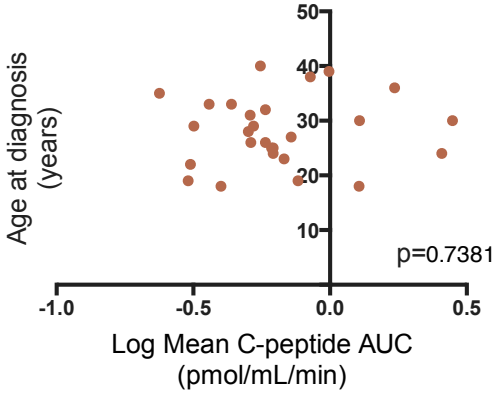
Statistical analysis as stated in Table 19.

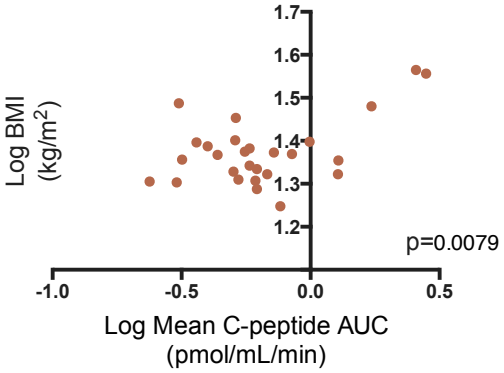


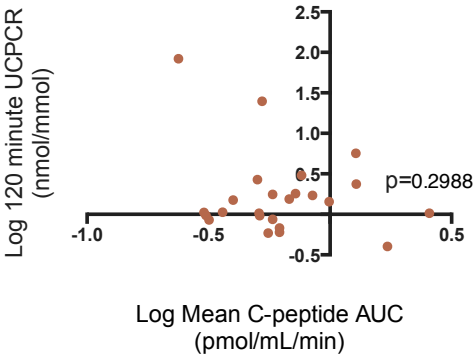

Log IA2 titre
(Units/mL)
4
3
2
1
-1
-1.0
-0.5
0.5
p=0.1122
Log Mean C-peptide AUC
(pmol/mL/min)


Log ZnT8 titre
(Units/mL)
4
3
2
1
-1
-1.0
-0.5
0.5
p=0.0146
Log Mean C-peptide AUC
(pmol/mL/min)


Log insulin dose
(Units/kg)
0.0
-0.2
-0.4
-0.6
-0.8
-1.0
-0.5
0.0
0.5
p=0.7356
Log Mean C-peptide AUC
(pmol/mL/min)

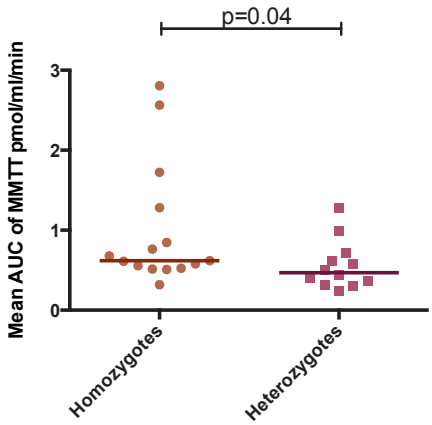

Age at diagnosis
(years)
50
40
30
20
10
-1.0
-0.5
0.0
0.5
p=0.7381
Log Mean C-peptide AUC
(pmol/mL/min)


Log BMI
(kg/m2)
1.7
1.6
1.5
1.4
1.3
1.2
-1.0
-0.5
0.0
0.5
p=0.0079
Log Mean C-peptide AUC
(pmol/mL/min)


Log 120 minute UCPCR
(nmol/mmol)
2.5
2.0
1.5
1.0
0.5
-0.5
-1.0
-0.5
0.5
p=0.2988
Log Mean C-peptide AUC
(pmol/mL/min)

Homozygotes for *HLA DRB1*0401* had higher median C-peptide AUC than heterozygotes (0.619 range: 0.318-2.805 pmol/mL/min versus 0.470 range: 0.238-1.277 pmol/mL/min) p=0.04 (Figure 21).

Figure 21: Mean AUC C-peptide in *HLA DRB1*0401* homozygous and heterozygous groups.
Line indicates median. P value calculated using unpaired t-test on log-transformed data.



There was no variation in the number of autoantibodies positive in *HLA DRB1*0401* positive and negative groups (Figure 23). When examining mean C-peptide AUC for groups with 1, 2 or 3 autoantibodies positive there was a trend towards lower AUC with higher autoantibodies number. In recruited subjects there were 4 patients who were classified as obese according to World Health Organisation criteria (BMI ≥30kg/m^2), with 2 of these in the obese class II category (BMI ≥35kg/m^2). There was a stepwise increase in C-peptide levels in each BMI band (Table 19, Figure 20. Correlations of C-peptide AUC and insulin resistance using HOMA2 (Homeostasis model assessment 2) analysis were considered but could not be performed with accuracy as many of the C-peptide readings fell out of the acceptable range for calculation of HOMA2.

Figure 22: Mean C-peptide AUC levels ranked according to BMI bands.

Bar graph showing mean C-peptide AUC of randomized patients with WHO BMI bands.

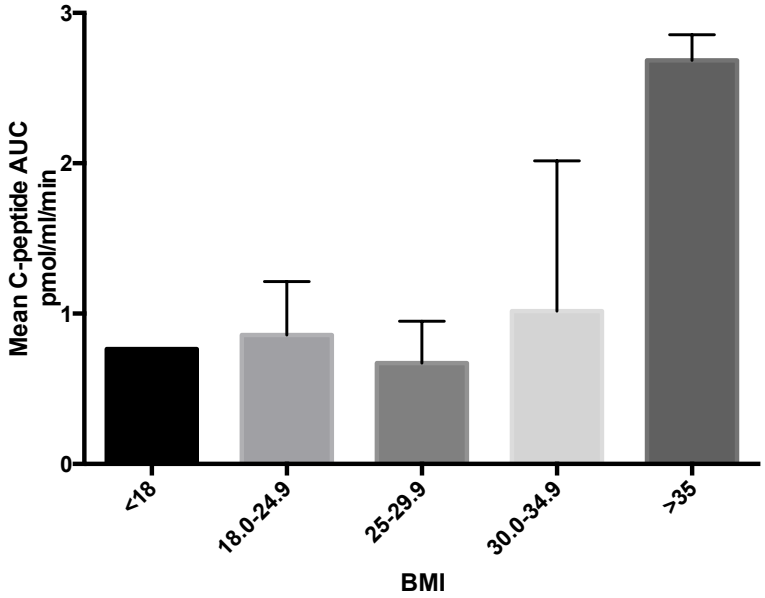


Figure 23: Number of positive autoantibodies within *HLA DRB1*0401* positive and negative groups.

(Chi squared analysis p=0.6032)

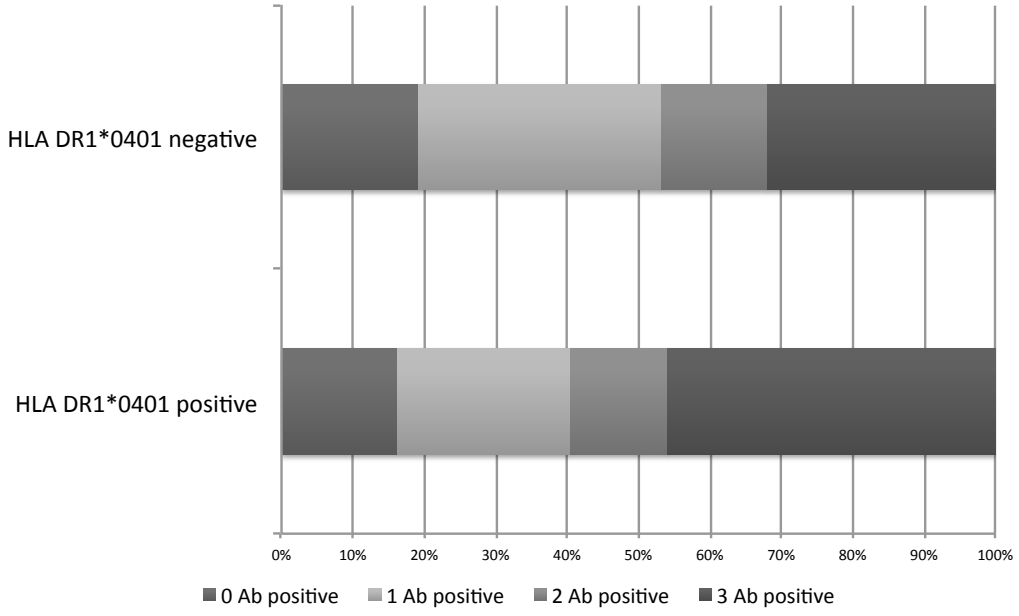


Table 20: Characteristics of randomised subjects with BMI >30kg/m^2.

Subject ID	BMI	Age at diagnosis	Autoantibodies			Family history of T1D?	Ketones present on diagnosis	Mean C-peptide AUC pmol/mL/min
			GAD	IA-2	ZnT8			
1019	36.7	24	648	-	-	-	Yes	2.56
2031	36	30	78	-	-	Aunt and Grandmother	Yes	2.81
2016	30.7	22	>2000	-	-	Sibling	Yes	0.31
2011	30.2	36	202	164	18	Sibling	Yes	1.73

Twenty-seven subjects were randomised into the trial and their baseline demographics are summarised in Table 21.

Table 21: Baseline demographics of randomised patients.

Figures denote median (range) or number (%) unless otherwise stated. V-1 =Visit -1

		Baseline
N=27		Male: 18 (67%) Female: 9 (33%)
Age (years)		28 (18-40)
Ethnicity		White British: 25 (93%) White Other: 2 (7%)
Weight (kg) (V-1)		71.5 (49.9-141.7)
BMI (kg/m^2)(V-1)		23.3 (17.7-36.7)
Peak C-peptide (pmol/ml)		0.87 (0.35-3.62)
Mean C-peptide AUC (pmol/mL/min)		0.58 (0.24-2.81)
HbA1c	(mmol/mol)	62.0 (46.0-85.0)
	%	7.8 (6.4-9.9)
Time from diagnosis to first dose (days) Mean (±SD)		82.2 (±17.4)

3.5 Discussion

Recruitment of subjects newly diagnosed with T1D into an intervention trial poses a number of unique challenges. Multiple approaches were necessary to identify relevant subjects and to highlight the active research opportunity open to these individuals. Traditional media were used through press releases in combination with the adoption of an effective online presence. Such approaches have also been successful in another T1D intervention trial of IL-2 therapy[213]. Initially lower than predicted screening numbers and levels of *HLA-DRB1*0401* positivity lead to delayed completion of recruitment. By the end of the recruitment the proportion of *HLA-DRB1*0401* positive patients equalled our expectations, suggesting that the initial run of negative results was a statistical abberation which normalised with greater numbers. The incorporation of the ADDRESS-2 register into our screening strategy in October 2012 coincided with an approximate 2-fold increase in screening (Figure 14). Further strategies to optimize screening rates included:

- Trial promotion through local and national clinical networks targeting senior and junior medical staff and diabetes teams.
- Increasing public awareness through news reports, a commercial website and a social media presence on Facebook®.
- Linking with Diabetes UK (DUK) and Juvenile Diabetes Research Foundation (JDRF). DUK and JDRF provided additional support by raising awareness of our work within their supporters through patient publications, websites, online resources and a number of patient outreach events.
- Integrating expert patient feedback from Diabetes Research Network "Dragon's Den" group.

The patient expert panel was a valuable interaction enabling incorporation of patient views into the study. There were methodological and logistical restraints on suggestions, such as widening the age or time frame criteria and opening up weekend appointments. Nonetheless, as a direct result of the meeting, a "buddying" system for individuals was integrated into the trial to provide support from research-experienced peers. Ethical approval was obtained to allow subjects to contact enrolled subjects who had volunteered to "buddy" patients. This was open to subjects from the initial screening stages with the hope that shared experiences would reassure potential subjects considering the trial. Although this did

happen on an informal basis, multiple trial sites and limited patient numbers did not facilitate communication between subjects, necessitating a more structured contact. There were important ethical considerations regarding the use of social media for communication with and between patients, in regards to consent and anonymity. In addition setting up and managing these systems would require extra staff time. Crucially the feedback from this panel and the practical experience acquired during the study was used to improve recruitment throughout the trial and continues to positively influence the design of local intervention trials planned in the future.

With 11% of eligible referrals being excluded as they would have exceeded the 100 day window by the completion of screening, this study emphasizes the time constraint on recruiting soon after diagnosis. A significant proportion, 35% (76), were not newly diagnosed, many of these being self-referrals. The initiation of recruitment via ADDRESS-2 was associated with a significant rise in screening numbers but on closer analysis other strategies may have contributed to this rise. Despite 188 subjects being identified via the ADDRESS-2 register only a total of 10 subjects (5.3%) from ADDRESS-2 were screened, with 3 proceeding to randomization. The screening figures appear low in view of high potential numbers and particularly as subjects are pre-selected as being interested in research. This experience echoes that of the DILT1D trial[213]. One of the major constraints was the ADDRESS-2 protocol specification that prohibited initial direct contact from the MonoPepT1De team. The chain of referral led from ADDRESS-2 staff to the local research team, who approached the patient for consent allowing ADDRESS-2 staff to pass contact details to the MonoPepT1De team. This prolonged process added pressure to screen within an already tight window. Local identification of subjects frequently allowed face-to-face discussion with a member of the research team, which was not possible via ADDRESS-2 and it is likely that a personal introduction to the study led to a higher uptake. Geographical factors were likely an additional consideration for ADDRESS-2 patients. The use of the www.diabetes.co.uk commercial website gave disappointing results, despite potentially far-reaching exposure. Limited details of the trial were permitted to avoid subjects directly contacting the team and bypassing the registration system. This, with lack of clinical exposure and screening of responses at source, meant that this was neither a time nor cost-worthy venture.

The recruitment centre at Guy's and St. Thomas's NHS Trust (GSTT), London recruited the highest number of subjects with good retention of patients through the trial. I attribute this to the following likely explanations:

- Research team: As clinical research fellow to this study I was able to identify and liaise directly with patients with enough dedicated time to promote the trial at different levels, and ensure wide exposure of the trial locally. My prior experience with the inpatient and outpatient teams facilitated referrals. Patients were able on the whole to have one direct contact point for queries, which allowed a more personal approach.

- Location: GSTT has a large catchment area and is located within a city with a younger population than the country as a whole. There are ADDRESS-2 sites and potential referral centres are clustered around London. Good transport links allowed patients to travel further for visits.

- Research infrastructure: Flexibility with facilities and staffing allowed us to offer earlier appointment times so patients would attend before work. With an established research history in immune studies some clinicians were already aware of the centre's ongoing interests and patients were able to identify the centre through its public profile.

The total number of subjects passing *HLA DRB1*0401* allele screening was slightly lower than estimated, with 43.0% (n=37) being positive, all of whom were Caucasian. Data from the South West newly diagnosed diabetes Type 1 collection (SWENDIC) showed, of 345 newly diagnosed adults recruited since 1999, the frequency of *HLA DRB1*0401* was 52.4% (unpublished data, C. Dayan). In control populations the proportion of the allele *HLA DRB1*0401* is estimated at between 10.3-25.1%[214,215]. In the adult onset T1D population this is higher at 48%, with similar levels in childhood groups[216]. This correlated well with US data which showed 48.8% of T1D subjects displaying the allele[217]. Unfortunately false negative results for *HLA-DRB1*0401* in our screening mistakenly excluded two subjects from further screening.

Using a panel of four autoantibodies, GAD, IA-2, IAA, and ICA or ZnT8, only 2–4% of children are autoantibody negative at diagnosis[218], but there is little data fully comparable to our screening strategy as screening panels and age criteria vary between studies. One group,

Vaziri-Sani et al.[212] used identical screening autoantibodies, GAD, IA-2 and ZnT8, in a similar age range (15-34 years) within a Swedish population at diagnosis. Their data (Figure 17) showed less antibody negativity (10% versus 17% in the MonoPepT1De group) and less GAD-centric responses compared to our group. Differences in assay techniques may account for some of the discrepancy. The assays used by Vaziri-Sani et al. were radioligand-binding assays (RBA) compared to ELISA in MonoPepT1De patients. GAD autoantibody (RBA) had lower sensitivity at 86% and specificity at 93%, compared to 92% and 98% respectively in MonoPepT1De. Similarly IA-2 autoantibody RBA sensitivity was 66%, specificity 98% compared to rates in ELISA of 72% and 99% respectively. Higher sensitivity of the ELISA assays does not explain the higher antibody negative cohort, however our age range was slightly higher and both autoantibody negativity, GAD positivity and titre rise with age[219]. Although the MonoPepT1De group was homogeneously Caucasian, there may be geographical variations of genotypes between the UK based and Swedish populations to account for the differences in autoantibody profiles. The inclusion of IA-2 autoantibody in the MonoPepT1De screening panel did not increase positive screening outcome over GAD and ZnT8, as no patients were positive for IA-2 antibody alone compared to 3% in the Vaziri-Sani et al. data. Autoantibody profile at diagnosis has been linked to low baseline C-peptide, with the presence of GAD antibody linked to more rapid C-peptide decline[220], however these data used fasting or random C-peptide and no comparable distinction was seen at baseline in the MonoPepT1De dataset. Interestingly there was a significant negative correlation between ZnT8 autoantibody titre and C-peptide AUC (Table 19, p=0.01) within this adult *HLA-DRB1*0401* population. This correlation was not found in another study of ZnT8 titres at diagnosis against MMTT stimulated C-peptide[221]. This small cohort (n=21) of newly diagnosed patients was not HLA screened and again age of subjects (mean 20.3 (12.2-34.6) years) and use of an alternative immunoprecipitation assay (incorporating a dimeric construct to combine epitopes), may account for differences. Data published on 159 children under 15 years of age which showed the inverse to our data in that positivity for the arginine subtype of ZnT8 autoantibody was associated with higher stimulated C-peptide levels[222]. Our assay did not distinguish between the three subtypes of ZnT8 autoantibodies.

The majority (96.4%) of our *HLA-DRB1*0401* and autoantibody positive subjects demonstrated a peak stimulated C-peptide of >0.2pmol/mL within 100 days of diagnosis. Residual β-cell function was not of a Gaussian distribution as shown by C-peptide AUC (Figure 19) with a small number of subjects causing a positive skewing of data. Analysing correlates

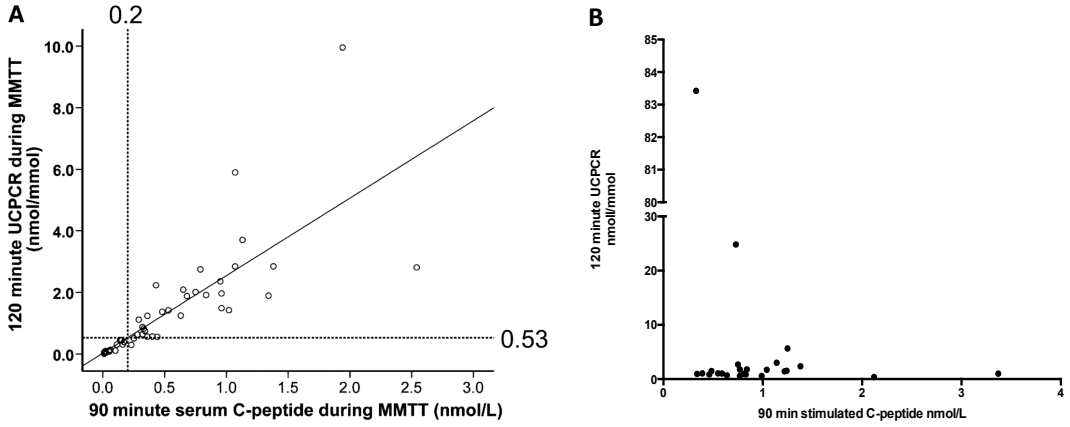
to baseline residual β-cell function as measured by mean C-peptide AUC on the MMTT showed that fasting and peak C-peptide levels had the strongest relationship. This reiterates published data findings showing fasting C-peptide correlating with C-peptide AUC when studied in a paediatric group, together with C-peptide levels at 90 minutes[223], which were considered equivalent to peak MMTT C-peptide levels[224]. Peak C-peptide (over a 120 minute MMTT) also correlated in a cohort aged between 7-46 years[225]. In the MonoPepT1De cohort C-peptide levels peaked at 90 minutes only in a third of subjects with the rest peaking at 120 minutes; again age may be a factor in the differences in time to peak C-peptide in our adult population. Correlates of β-cell reserve estimated through mean C-peptide AUC are important when considering future surrogate markers. From this baseline data fasting C-peptide appears to be a very good surrogate for mean C-peptide AUC and would have the advantage of negating the need for a MMTT, which would be particularly beneficial in children. Nevertheless fasting C-peptide requires further validation to identify a comparator to the peak 0.2pmol/mL threshold as a marker for significant residual β-cell reserve. The dynamics of C-peptide release change as patients approach clinical diagnosis, with fasting C-peptide climbing in the 6 months prior to diagnosis whilst other C-peptide parameters such as AUC fall[226]. Therefore although it would be clinically beneficial to have a simpler test for β-cell reserve, it may be judicious to gather wider data on fasting and stimulated responses to gauge longitudinal changes. Peak C-peptide also needs further investigation if the definition is to be reliably narrowed to a set time point in the MMTT, with differences in timing of peaks between adult and paediatric populations to be fully elucidated.

Urine C-peptide creatinine ratio (UCPCR) has been assessed as an alternative to stimulated C-peptide as a measure of endogenous insulin secretion with high correlation of a 120 minute post MMTT UCPCR with a 90 minute MMTT C-peptide ($r=0.91$; $p<0.0001$) in patients with T1D of a median duration of 6.5 (2.3-32.7 years)[73]. The benefits of this test are its non-invasiveness, simplicity and stability, potentially allowing patients to test at home following meals. In our sample there failed to be significant correlation of baseline 120 minute post MMTT UCPCR with 90 minute MMTT C-peptide or C-peptide MMTT AUC, which remained non-significant despite matching the statistical non-parametric analysis used in the Besser et al. paper. Of the 72 T1D subjects examined in this paper the upper limit of UCPCR was approximately 10nmol/mmol compared to the wide range between 0.40 and 83.4 nmol/mmol (Figure 24). Significant correlations between 120 minute UCPCR and 90 stimulated MMTT C-peptide has been seen in children[73], adults[73,227], Type 1 diabetes[73,227],

Type 2 diabetes[185] and islet transplant recipients[228], however there is currently little data on the use of UCPCR at time of diagnosis with the lowest stated median duration of diabetes in these studies being 6.5 years[73]. Wide variability of C-peptide levels at diagnosis could limit the use of UCPCR as a marker of β-cell reserve[229].

Figure 24: Scatter diagrams showing correlation between 120 minute UCPCR and 90 minute serum C-peptide during MMTT

A. From Besser et al. 2011[73] correlation of stimulated C-peptide and UCPCR in subjects with established Type 1 diabetes. Dotted lines show equivalence of >0.2nmol/L C-peptide and >0.53nmol/mmol UCPCR levels calculated through linear regression (with 94% sensitivity and 100% specificity). B. Correlation in newly diagnosed MonoPepT1De cohort.



I compared MonoPepT1De data with large scale published data on T1D characteristics at diagnosis from the T1D TrialNet group, an international multicentre clinical trial consortium which has collated large cohorts of cases. Table 22 shows composite data from two TrialNet publications looking at C-peptide levels at baseline from a number of their intervention trials against the MonoPepT1De trial. The table outlines variations in trial inclusion criteria, such as age and screening autoantibodies. The MonoPepT1De data shows slightly higher mean C-peptide AUC and peak C-peptide than data from Greenbaum et al.[225] and Bollyky et al[230], both of which datasets include paediatric patients. When children are excluded in the breakdown of the Bollyky et al. data, the figures for mean C-peptide AUC (MonoPepT1De 0.80 (±0.64), Bollyky et al. 0.82 (±0.37) pmol/mL/min) and peak C-peptide (MonoPepT1De 1.154 (±0.83), Bollyky et al. 1.06 (±0.57)) are comparable.

Table 22: Comparison of C-peptide measurements in randomised MonoPepT1De subjects compared to TrialNet cohorts.

	MonoPepT1De	TrialNet data	
		Greenbaum at el, 2012[225]	Bollyky et al. 2015[230]
N	27	191	883
Time from diagnosis	Within 100 days		
Peak C-peptide criteria	≥ 0.2pmol/mL*		
Use of *HLA DRB1*0401* inclusion criteria	Yes	No	No
Autoantibody criteria	At least one autoantibody positive out of GAD, IA-2 or ZnT8 Ab	At least one autoantibody positive out of IAA, GAD, IA-2 or ICA Ab	One autoantibody positive out of GAD, IAA, IA-2, ZnT8 Ab
Age group	18-45	7-45	3-45
2-hour AUC mean C-peptide pmol/mL/min Mean (SD)	0.80 (0.64)	0.71 (0.33)	All ages: 0.74 (0.38) Age>17: 0.82 (0.37) Age 12-17: 0.79 (0.36) Age 8-12: 0.63 (0.32) Age<8: 0.40 (0.18)
2-hour peak C-peptide pmol/mL	1.154 (0.83)	0.93(0.44)	All ages: 0.92 (0.49) Age>17: 1.06 (0.57) Age 12-17: 0.98 (0.44) Age 8-12: 0.76 (0.40) Age<8: 0.50 (0.22)
Trial data	MonoPepT1De trial only	Data from intervention trials	Data from intervention trials (All subjects)

		Mycofenolate mofetil and daclizumab (All patients) Abatacept (Placebo patients only) Rituximab (Placebo patients only)	Mycofenolate mofetil and daclizumab. Abatacept Rituximab GAD-alum Canakinumab

** to ensure metabolic stability MMTT not done before 21 days.*

The inclusion criteria of peak stimulated C-peptide ≥0.2pmol/L was failed by 2.9% (1/35 patients) of the MonoPepT1De subjects compared to 2.0% (18/883 subjects) in the Bollyky et al. group, which included children. Recent data from Lachin et al[58] revealed an almost linear relationship between decline in stimulated C-peptide to undetectable levels and outcomes such as retinopathy and hypoglycaemia risk. This is in keeping with earlier data showing peak stimulated C-peptide levels of >0.2 pmol/mL were still protective against complications of retinopathy and nephropathy[57]. Whilst only a minority of our screened population was excluded, such data suggest that potential for immunotherapy intervention could be widened. This would be of greatest benefit in paediatric groups, where TrialNet demonstrated that 9% of patients less than 8 years of age fail the >0.2 pmol/mL cut off[230]. In these cases lowering the C-peptide cut-off would not only open immunotherapy options to those with greatest need, i.e. with those more aggressive disease and greater prospective disease burden, but also target the population with the fastest rising incidence[10]. Furthermore the use of immunotherapy could also be expanded later into the disease course.

Subject data from Greenbaum et al.[225] were broadly categorized as *HLA DR3, DR4* positive and *DQB1*0602* negative, with no correlation with C-peptide changes over time. The effect of *HLA* genotypes on baseline levels was not reported. Data on the link between genotype and β-cell function in T1D is contrasting. In one study lower fasting C-peptide levels were linked to higher risk genotypes (*DRB1*03-DQB1*0201/DRB1*04-DQB1*0302 (DRB1*04* different from *0403)*), compared to moderate or low risk, and this difference is sustained for 12 months after diagnosis[231]. Another study found that DR3 but not DR4 alleles were significant parameters in determining C-peptide levels. In our data subjects homozygous for *HLA*

*DRB1*0401* alleles had significantly lower C-peptide AUC than heterozygotes, though in view of the small subject numbers, it is difficult to draw firm conclusions from this.

Just over a quarter of TrialNet subjects in interventional trials are overweight or obese[230], which compares well with MonoPepT1De data (25.9%). Our class II obese patients retained higher levels of C-peptide, which may be due to earlier disease presentation as a result of insulin resistance increasing insulin requirements. These patients showed clear autoimmunity with positive antibody titres, a high proportion of family history of T1D and presence of ketosis at diagnosis, which would suggest insulin deficiency. Obesity is not fully concordant with higher C-peptide levels. Three out of 4 obese subjects had C-peptide AUC figures in the top quartile however one was in the lowest quartile.

The MonoPepT1De trial baseline data results show heterogeneity in terms of stimulated C-peptide, BMI and autoantibody profiles. This is despite narrowing of the targeted newly diagnosed population to a stimulated C-peptide threshold, antibody positivity and presence of the *HLA DRB1*0401* allele. Comparisons to other trials that do not limit genotype and include paediatric groups widen this heterogeneity further. With many clinical immunotherapy trials yielding disappointing results there is a need to focus on this heterogeneity, particularly when considering future intervention trials. Previous examples of distinct subgroups in disease progression or response to treatment have been reported. In disease prevention, *post-hoc* analysis of the Diabetes Prevention Trial of Type 1 diabetes (DPT-1) showed that the presence of IAA >80 nU/mL in high-risk subjects reduced the risk of developing T1D following oral insulin treatment (hazard ratio 0.566, 95% CI 0.361–0.888; P=0.015) and delayed disease onset by 4.5 years[78]. This benefit substantially increased when IAA levels of >300nU/mL were considered, delaying disease for nearly 10 years[232]. IAA positivity was also linked to outcome in treatment with Otelixizumab, with multivariate analysis showing the presence of IAA (P=0.022) or the interaction of IAA and C-peptide (P=0.013) independent predictors of outcome to treatment[80]. Mamchak et al[233] combined *in-silico* disease modelling and preclinical clinical models to look at predictors of outcome following combination treatment with an anti-CD3 monoclonal antibody and oral insulin. These modelling techniques also identified IAA levels as a positive predictor. In the anti-CD3 teplizumab trial[79], neither autoantibodies, BMI, age, duration of disease nor sex distinguished responders from non-responders to therapy, but the authors were able to define responders as having lower insulin requirements and HbA1c at baseline despite having similar C-peptide

AUC as non-responders. Immunologically they were able to delineate shifts in T-cell populations between the two groups, with lower frequency of CD4+CCR4+ memory and naïve T-cells, CD4+CCR6+ naïve CD4+ T-cells, naïve CCR4+ CD8+ T-cells, and IFN-γ-producing CD8+ T-cells at baseline in responders. Higher number of activated CD8+ terminally differentiated effector and CD8+ effector memory T-cells were also seen in responders versus non-responders.

The HLA composition of patient groups has not been found to influence intervention outcomes, with the exception of Diapep277 therapy[234] where adults with low and moderate risk HLA groups appeared to benefit most. This finding will require validation.

This chapter has outlined the strategies and potential pitfalls in recruiting to an immune intervention trial in T1D and the need to exploit avenues of information dissemination whilst retaining the integrity of responses within a narrow timeframe. Baseline data have highlighted disease heterogeneity despite restriction to a newly diagnosed adult cohort with the presence of an *HLA DRB1*0401* allele. Such heterogeneity appears to be a common factor in T1D studies, with suggestions that endotypes of disease are defined by discrete pathophysiological mechanisms[235]. Such endotypes have implications for potential treatment efficacy. Careful stratification of subjects needs to be considered in the design of intervention studies, to strike a fine balance between isolating subpopulations most suited to individual treatments without limiting wider applicability.

4 MonoPepT1De trial outcomes

4.1 Introduction

The MonoPepT1De trial began screening in February 2012 with completion of the last patient visit in February 2015. Unblinding of treatment allocation began in April 2015. This chapter presents the clinical and immunological outcomes from the trial.

4.2 Results

Of the 24 patients randomised, 3 did not complete a minimum of 11 out of 12 treatments (one from the low frequency and 2 from the high frequency group) and additional replacement subjects were recruited. All treated subjects were included in the analysis. In total there were 10 subjects allocated to the low frequency treatment group, 9 subjects in the high frequency treatment group and 8 subjects in the placebo group. Two subjects were lost to follow-up and two declined follow-up visits. Baseline characteristics of the treatment groups are described in Table 23. Baseline mean glycated haemoglobin was significantly different between the high frequency and placebo groups (p=0.02).

Table 23: Baseline characteristics of subjects within low frequency, high frequency and placebo groups.

		C19-A3 Peptide Immunotherapy		
		Low frequency	High frequency	Placebo
Number of subjects		10	9	8
Mean age (years±SD)		26.6±5.5	30.0±5.7	28.9±8.2
Gender		4 Female: 6 Male	3 Female: 6 Male	2 Female: 6 Male
BMI (kg/m^2±SD)		24.2±5.5	25.6±5.4	23.1±2.6
Number of antibodies (GAD, IA2, ZnT8)	1	50.0%	11.1%	12.5%
	2	30.0%	11.1%	25.0%
	3	20.0%	77.8%	62.5%
Mean time from diagnosis to first dose (days±SD)		82.5±16.0	91.0±15.5	95±22.8
Mean glycated haemoglobin (mmol/mol±SD)		58.4±14.9	51.7±6.83	62.5±13.7
Average total daily insulin dose (IU/kg/day±SD)		0.38±0.18	0.30±0.07	0.42±0.20
Stimulated C-peptide AUC (nmol/L/min±SD)		0.81±0.76	0.99±0.73	0.58±0.25

4.2.1 Safety of C19-A3 peptide immunotherapy

The primary endpoint was the safety assessment of PI C19-A3 treatment, which was well tolerated with no hypersensitivity reactions. No serious adverse events were considered linked to treatment. Local erythematous reactions previously described with C19-A3 peptide[117] were present in both treatment and placebo groups. These skin reactions were not associated with local wheal or swelling, and did not change in size or quality with repeated peptide administration. Skin reactions were present in 8/9 subjects in the high frequency, 10/10 subjects in the low frequency and 4/8 subjects in the placebo group. Five serious adverse events (SAEs) were reported which were classified as mild or moderate in severity and either unrelated or unlikely related to the randomisation drug (Table 24).

Table 24: Summary of serious adverse events reported during the MonoPepT1De trial

Placebo Group		
Event description	Severity	Relationship of event to study treatment?
Mild DKA resulting in hospitalisation.	Mild	Unlikely
Low Frequency Group		
Event description	Severity	Relationship of event to study treatment?
Patient admitted to hospital following collapse. Investigated by cardiologists with a diagnosis of vasovagal event.	Mild	Unlikely
DKA - hospitalised	Moderate	None
High Frequency Group		
Event description	Severity	Relationship of event to study treatment?
Chronic pilonidal sinus treated prior to recruitment to trial. Reviewed by surgeons and hospitalised for urgent treatment.	Moderate	None
Subject's wife fell pregnant	Mild	None

4.2.2 C-peptide responses

Stimulated C-peptide production was assessed using the normalised area under the curve of a 2 hour MMTT at 12, 24, 36 and 48 weeks versus baseline and between groups (Table 25). The pre-determined analysis plan specified a multilevel model repeated measures (MMRM) analysis adjusted for baseline value of AUC, which showed no significant treatment-related effects. Change in stimulated C-peptide over time was a pre-specified seconday endpoint and the decline in stimulated C-peptide was different between study groups, such that at the 3 month timepoint mean loss of C-peptide in the placebo group significantly more than in the high frequency ($p=0.03$) and low frequency groups. This difference in C-peptide decline was evident in individual data plots: compared with baseline, C-peptide levels in subjects receiving placebo showed a decline at every timepoint in every subject (apart from one subject, at 6, 9 and 12 months) and mean values declined significantly in paired analyses to baseline. This contrasted with findings in the treatment groups, in which the mean percent change was more modest, fewer individual subjects showed actual loss of C-peptide, and

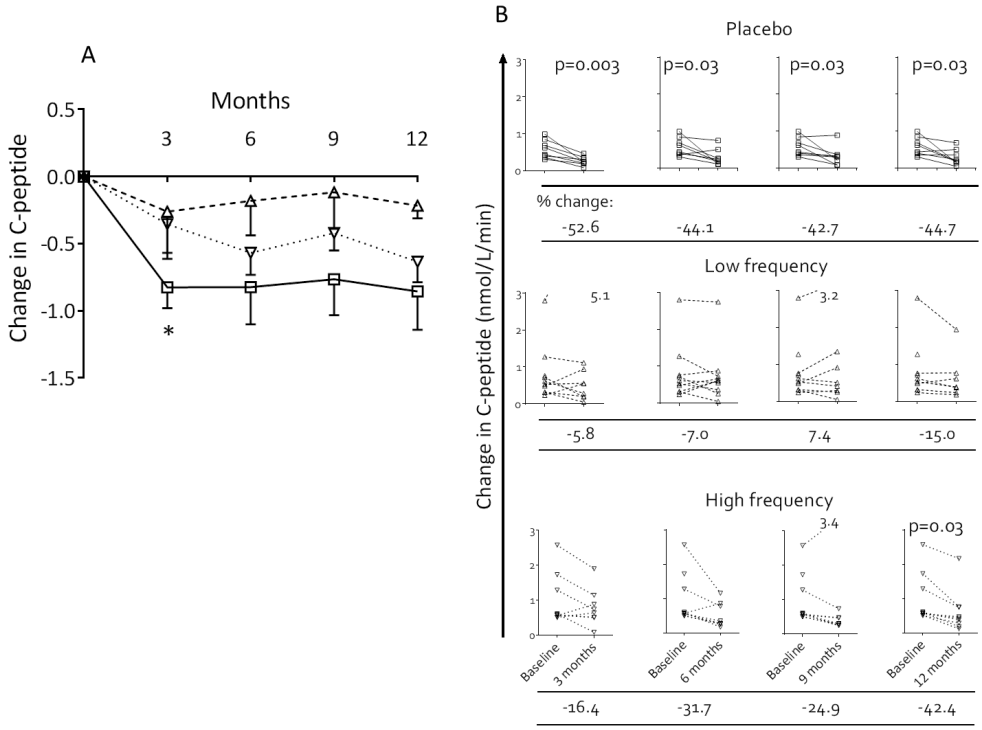
significant changes in mean values were only seen when comparing baseline with 12 month levels in the high frequency group. Thus in this study, patients on placebo manifest an early decline of approximately half of measurable C-peptide production, whilst this is not seen in the active treatment groups during administration of proinsulin C19-A3 peptide injections.

Table 25: Normalised C-peptide area under curve (pmol/mL/min) at baseline and 12, 24, 36 and 48 weeks after initiation of treatment.

Percentage change in AUC from baseline is shown in shaded columns. Missing data is indicated by a crossed through cell.

	Patient No.	Baseline C-peptide AUC	Week 12 C-peptide AUC	% change from baseline at week 12	Week 24 C-peptide AUC	% change from baseline at week 24	Week 36 C-peptide AUC	% change from baseline at week 36	Week 48 C-peptide AUC	% change from baseline at week 48
HIGH-FREQUENCY	1017	0.51	0.64	24.70	0.26	-49.82	0.26	-49.82	0.12	-77.10
	1019	2.57	1.89	-26.31	1.16	-54.74	3.38	31.32	2.16	-15.87
	2010	1.28	0.74	-42.19	0.78	-38.87	0.74	-42.19	0.74	-41.97
	2011	1.72	1.14	-34.02		NA		NA	0.76	-56.11
	2017	0.50		NA	0.29	-42.64	0.30	-41.15	0.28	-43.58
	2021	0.58	0.50	-13.25	0.87	49.30	0.31	-46.71	0.43	-26.62
	2025	0.58	0.51	-12.20	0.28	-52.47	0.48	-18.17	0.38	-34.19
	3015	0.56	0.88	57.87	0.36	-36.30	0.48	-14.89	0.47	-16.01
	4005	0.61	0.08	-86.11	0.18	-70.72	0.25	-58.67	0.18	-70.52
LOW-FREQUENCY	1024	0.32	0.20	-36.11	0.61	90.50	0.25	-19.99	0.24	-25.29
	2001	0.62	0.28	-53.90	0.56	-9.32	0.51	-17.27	0.37	-40.02
	2006	1.27	1.12	-12.46	0.73	-42.74		NA		NA
	2016	0.31	0.05	-83.89	0.05	-83.89	0.05	-83.89		NA
	2018	0.53	0.56	5.38	0.37	-29.95	0.41	-22.03	0.39	-26.11
	2020	0.72	0.19	-73.52	0.26	-63.28		NA		NA
	2027	0.49	0.95	94.03	0.66	36.37	0.91	87.68	0.62	26.30
	2031	2.80	5.12	83.06	2.74	-1.97	3.21	14.75	1.94	-30.47
	4007	0.76		NA	0.88	15.26	1.35	76.23	0.77	0.57
	5001	0.24	0.55	129.83	0.62	158.74	0.29	23.39	0.18	-25.36
PLACEBO	1008	0.30	0.19	-37.20	0.11	-63.10	0.09	-70.21	0.05	-83.47
	1025	0.36	0.08	-77.30	0.50	39.60	0.30	-15.77	0.36	-0.52
	2019	0.44	0.20	-53.44	0.19	-56.24	0.30	-30.34	0.50	15.78
	3002	0.99	0.23	-76.29	0.10	-89.93	0.21	-78.59	0.15	-85.25
	3005	0.68	0.36	-46.57	0.26	-61.20	0.37	-46.30	0.19	-71.95
	3012	0.85	0.47	-44.94	0.75	-11.68	0.89	4.95	0.69	-18.92
	4001	0.62	0.24	-61.39	0.18	-70.87	0.08	-87.30	0.18	-71.52
	4012	0.40	0.30	-23.80	0.24	-39.70	0.33	-18.17	0.23	-41.74

Figure 25: Change in C-peptide at 3, 6, 9 and 12 months from baseline.
A. Natural Log of mean change in normalised mixed meal tolerance test (MMTT) stimulated AUC C-peptide from baseline is shown in groups receiving placebo (open squares), low (triangle) and high (inverted triangle) frequency C19-A3 peptide over 12 months. There was decline in stimulated C-peptide in the placebo arm; the level significantly exceeded that in the high frequency arm at 3 months (; p=0.03). Error bars represent standard error of the mean. B. Change in normalised MMTT stimulated AUC C-peptide values from baseline versus level at 3, 6, 9 and 12 months in groups receiving placebo (open squares), low (triangle) and high (inverted triangle) frequency C19-A3 peptide over 12 months. There is a significant reduction in the stimulated C-peptide level in the placebo group at each time point. No significant decline was seen in the low and high frequency groups during the treatment phase; during follow-up a significant change in the high frequency group was seen at 12 months compared to baseline. Comparisons were made using paired t-tests. The mean % change was calculated at each time point for each study group.*



4.2.3 Metabolic responses

Insulin dose (units/kg/day) and HbA1c were measured longitudinally as indicators of metabolic control. Mean change in average insulin dose progressively rose in subjects in the placebo arm. In contrast, in the high and low frequency arms of the study, the mean change in average insulin dose was minimal or showed a decline. As a result, mean changes in insulin use were significantly lower in the high frequency arm at 6, 9 and 12 months (p=0.03, p=0.04 and p=0.01, respectively) and significantly lower in the low frequency arm at 12 months

(p=0.009) compared with placebo, with an overall difference between the treatment and placebo groups across all time points in multilevel model repeated measures analysis (p= 0.01) (Figure 26A).

Glycaemic control was intensively managed in the study, with a target HbA1c of less than 48mmol/mol (6.5%). Differences in HbA1c between study groups would not be expected, therefore, and significant changes were not seen; however, there was a trend for increased HbA1c levels in the placebo group, whereas in the treatment groups change was minimal or values declined (Figure 26B).

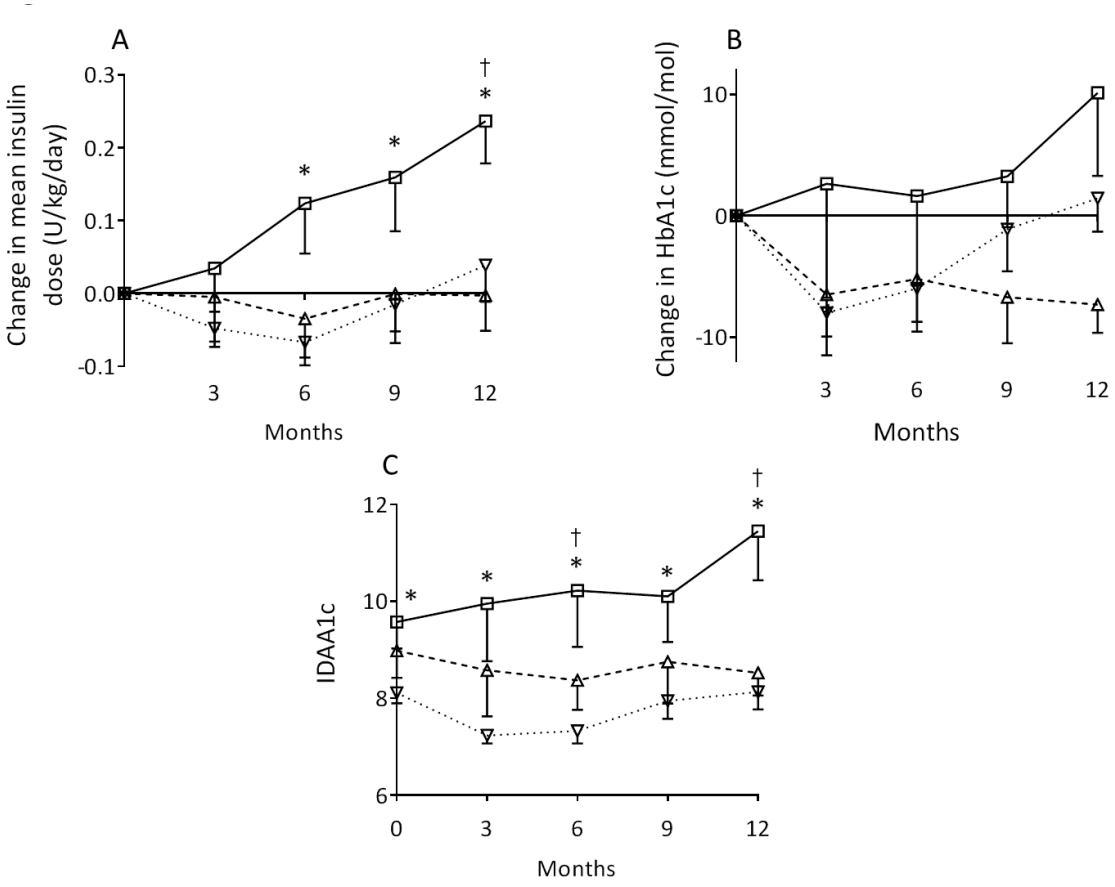
Using either mean insulin dose or HbA1c alone as a reflection of residual β-cell function has a number of limitations. As the goal of clinical management is glycaemic control, both these measures are influenced by exogenous insulin and the many factors that guide insulin dose. Mortensen et al.[236] introduced the insulin adjusted A1c (IDAA1C) as a marker of residual β-cell function. IDAA1C is calculated as A1C (percent) + [4 x insulin dose (units per kilogram per 24 h)]. An IDAA1C ≤9% was found to correspond to a predicted stimulated C-peptide >300 pmol/l and was used to define partial remission[236]. The placebo group had consistent increases in IDAA1c over 12 months consistent with a decline in endogenous insulin production, but IDAA1c was maintained at baseline levels in the intervention groups (Figure 26C).

4.2.4 Defining a clinical responder group.

Clinical response to treatment was defined using criteria described by Herold et al.[97] in which responders were characterised as those with an increase (a positive change from baseline) or maintenance (change of <50% of the interassay coefficient of variation of the C-peptide assay) in C-peptide response to a MMTT. There were 10 such “C-peptide responder” subjects identifiable during the treatment period (6 months), 1 in the placebo group, 6 in the low frequency and 3 in the high frequency groups. C-peptide responders were significantly more frequent in the low frequency group than in the placebo group at 3 months (p=0.03) months.

Figure 26: Changes in mean insulin, HbA1c and insulin dose adjusted HbA1c (IDAA1c)

A. Mean change in average insulin dose from baseline is shown in groups receiving placebo (open squares), low (triangle) and high (inverted triangle) frequency C19-A3 peptide over 12 months. There was a progressive increase in insulin requirement in the placebo arm. In contrast there was a reduction or stabilization in the low and high frequency groups, with significant differences in change compared with placebo in the high frequency group at month 6 (; p=0.03), month 9 (*; p=0.04) and month 12 (*p=0.01) and low frequency group at month 12 (†; p=0.009). Error bars indicate standard deviations. B. Mean change in HbA1c from baseline is shown in groups receiving placebo (open squares), low (triangle) and high (inverted triangle) frequency C19-A3 peptide over 12 months. There was a reduction observed over the 6-month treatment period in the high and low frequency groups. Error bars indicate standard deviations.* ***C.*** *Insulin dose adjusted HbA1c (IDAA1c) values following treatment were significantly lower in the high frequency arm at baseline, 3, 6, 9 and 12 months (*p=0.02, p=0.001, p=0.003, p=0.01 and p=0.002, respectively) and significantly lower in the low frequency arm at 6 and 12 months (†p=0.05 and p=0.01, respectively) compared with placebo.*



4.2.5 Immune assays

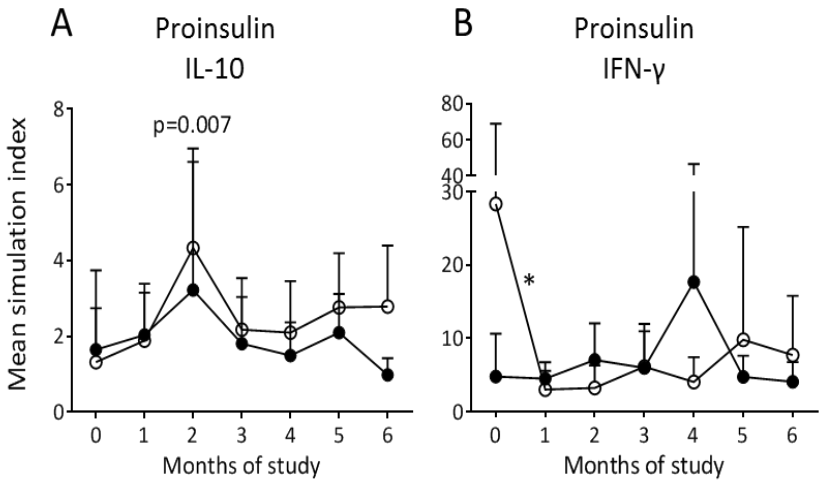
Over the duration of the treatment period, cumulative CD4 T-cell IL-10 responses to proinsulin stimulation were significantly higher in the blood of the high frequency compared with placebo (p=0.015) and low frequency groups (p=0.003). There were no significant differences in CD4 T-cell IFN-γ responses, circulating subsets of regulatory T-cells (Tregs) or

activated CD8 T-cells specific for β-cell target peptides between these groups (data not shown).

Using the responder definition as those with an increase (a positive change from baseline) or maintenance (change of <50% of the interassay coefficient of variation of the C-peptide assay) in C-peptide response to a MMTT to compare responders/non-responders, at 2 months, IL-10 responses against PI were significantly higher in peptide-treated C-peptide responders (p=0.007) and higher levels of IL-10 responses were maintained in the responder group. Peptide-treated C-peptide responders (but not non-responders) showed a trend for IFN-γ response levels against PI to decline between starting therapy and the first assay performed at 1 month (p=0.08).

Figure 27: Analysis of ELISPOT T-cell responses in peptide-treated C-peptide responders and non-responders.

*Amplitude of CD4 T-cell responses through (A) IL-10 and (B) IFNγ production as measured by stimulation index after in vitro culture with whole recombinant human proinsulin. Open circles represent peptide treated C-peptide responders and filled circles represent non-responders. Across the treatment period IL-10 responses to proinsulin stimulation rose and were significantly higher in the C-peptide responder group compared with non-responders (p=0.007). Peptide-treated C-peptide responders (but not non-responders) showed a trend for IFN-γ response levels against PI to decline between starting therapy and the first assay performed at 1 month (*p=0.08). Bars and symbols represent mean stimulation index at each timepoint and error bars are the 95% confidence intervals. This was a post-hoc analysis using longitudinal measurements of the SI which were transformed using the natural logarithm (Ln) and were analyzed with linear models having visit and treatment as main factors and a repeated-measures error structure. Estimates of the mean SI across visits were computed using model-based estimates ("least-squares means").*



Additional immune assays, performed by colleagues, analysed Treg expression of FoxP3 (Dr Jennie Yang) and CD8 antigen experienced (CD57+) T-cells (Dr Lorraine Yeo).

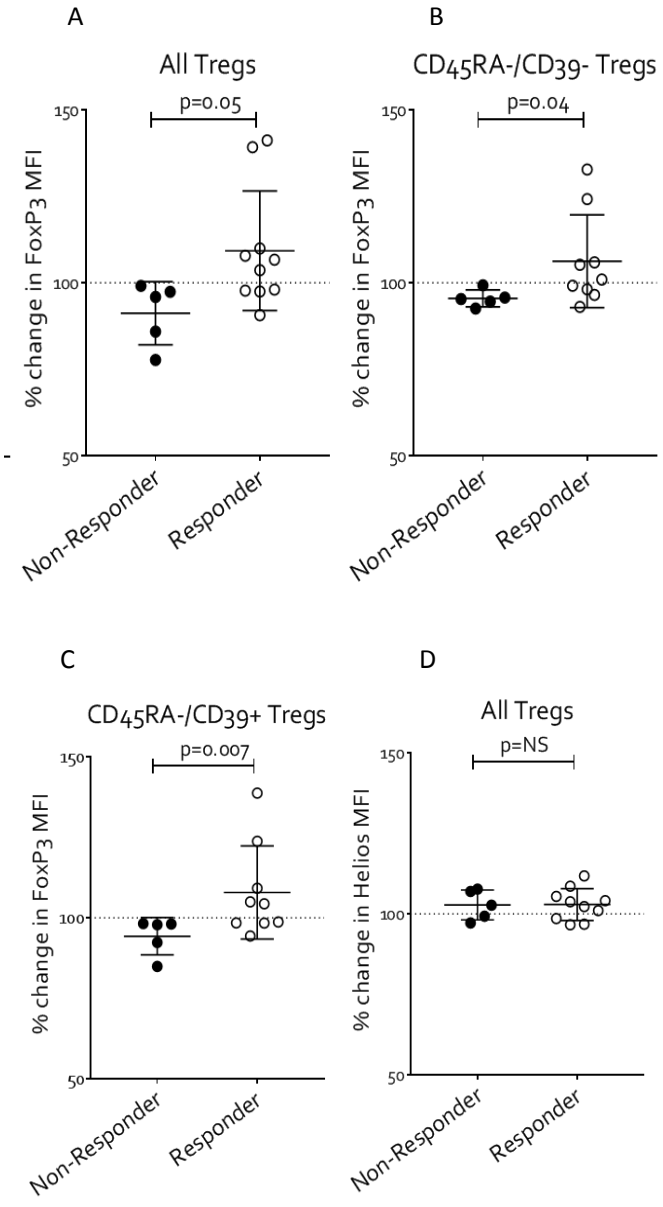
Immunophenotyping of Tregs was performed on cryopreserved PBMCs at baseline, 6 and 12 months in batches each consisting of all 3 visit samples from 4 subjects selected at random. Thawed PBMCs were live/dead blue (Thermo Fisher Scientific, London, UK) and then surface stained using anti-CD4-APC-Cy7 (RPA-T4), anti-CD25-PE (2A3 and M-A251), anti-CD27-BV605 (L128), anti-CD39-PE-Cy7 (A1; BioLegend), anti-CD45RA-PE-CF594 (HI100), anti-CD278-BV711 (DX29) and anti-HLA-DR-BV786 (G46-6; all BD Biosciences, Oxford, UK unless specified) and intracellularly stained using anti-Ki67-FITC (B56; BD), anti-FOXP3-Alexa Fluor 647 (259D/C7; BD Biosciences) and anti-Helios-Pacific blue (22F6; BioLegend, San Diego). Data acquisition was performed using a BD Biosciences LSRFortessa cell analyzer.

CD57+ (antigen-experienced) effector memory β-cell peptide-specific CD8 T-cells were detected using peptide-HLA-A*0201 tetramers loaded with preproinsulin 15-24, insulin B chain 10-18, and IA-2 797-805 as previously described by Skowera et al.[237] and expressed as a percentage of the parent tetramer population. Identical populations of CD8 T-cells specific for common viral peptides CMVpp65 495-503, EBV BMLF-1 280-288 and influenza matrix 58-66 were measured as controls.

Levels of Treg expression of FoxP3 increased between baseline and 6 months, with a greater increase in the peptide-treated C-peptide responders compared to non-responders (Figure 28A). Levels returned to baseline at 12 months and were unchanged throughout the study in C-peptide non-responders. The greatest fold-change in FoxP3 expression was seen in CD45RA- (memory) Treg sub-populations that lacked Helios expression, especially those co-expressing CD39 (Figure 28B-C). In contrast, levels of Helios expression by Tregs did not change in either study group (Figure 28D).

Figure 28: Change in expression of FoxP3 between peptide-treated C-peptide responders and non-responders at 6 months compared to baseline.

(A) Change within all Treg subsets. (B) Change within memory (CD45RA-) adaptive Tregs (Helios-negative). (C) Change within memory (CD45RA-) CD39+ Treg subset. (D) Change in Helios expression within all Treg subsets. Error bars show mean and SEM. This was a post-hoc analysis using one-way ANOVA (analysis of variance) with Tukey post hoc correction.

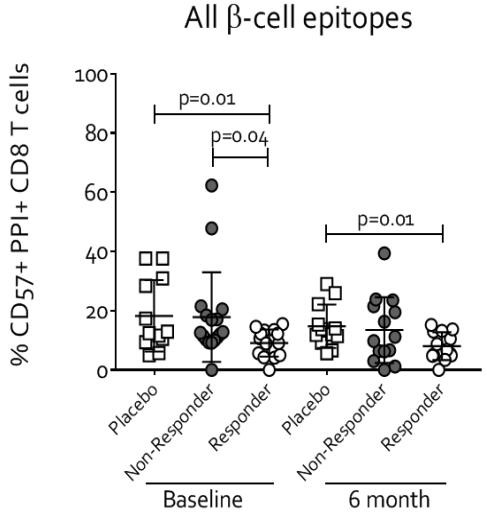


The proportion of CD8 T-cells specific for β-cells that expressed the disease-associated activated (CD57+) effector memory phenotype was significantly lower in treated C-peptide responders compared to placebo and non-responders at baseline, and remained lower than

placebo at 6 months. As observed in our previous studies in new onset type 1 diabetes patients, T-cell responses to C19-A3 at baseline were present in a small minority of patients; significant treatment and responder-related changes were not observed. There were no treatment- or response-related changes in autoantibodies.

Figure 29: Percentage of antigen experienced (CD57+) CD8 T-cells stained with peptide-HLA tetramers loaded with beta cell peptides at baseline compared to 6 months in placebo, and peptide treated responder and non-responder groups.

Error bars show mean and SEM.



4.3 Discussion

Tolerance induction through peptide immunotherapy in allergy and autoimmune disease has shown evidence of benefit in preclinical models[110,113,238,239], with PIT regimens for allergy already being well established in clinics. The MonoPepT1De trial aimed to establish the safety and immunological effects of a naturally processed and HLA-DR4 presented proinsulin peptide C19-A3 in new onset T1D. Repeated dosing of 10µg of peptide at 2 or 4 weekly intervals for 6 months was well tolerated and not associated with hypersensitivity or acceleration of β-cell loss.

C-peptide decline occurred early and was maintained in the placebo group, whereas those receiving peptide treatment had comparably less early C-peptide loss, with loss being significantly lower at 3 months in the high frequency group. Whilst being cognisant of

imbalances of baseline metabolic data present between treatment groups, particularly in the baseline HbA1c and insulin dose in the high frequency treatment group and small subject numbers, HbA1c, mean insulin dose and IDAA1c data appear to support the improved metabolic control and residual β-cell function in treated groups over placebo. In a small study such as this there are difficulties in balancing out baseline inequalities, which may exist despite the randomization process. Those receiving peptide treatment had stable mean insulin doses without compromise in glycaemic control as HbA1c fell or stabilised, compared to rising HbA1c in placebo despite increasing insulin administered. Whilst high frequency dosing was safe it did not confer additional benefits over low frequency administration, despite showing higher cumulative IL-10 responses. The increment in IL-10 response was not sufficient to infer any additional effect on metabolic parameters, and to explore the mechanistic responses further dose response studies would be beneficial with a wider span of doses than used in this study.

The definition adopted in this study was one previously used in an immunotherapy intervention trial by Herold et al.[97] equating to a threshold of 92.5% or more of baseline. As such this responder status may include those with slowed β-cell decline as well as those with preservation, which may be an analysis more suited to a small phase I study. Difficulties defining clinical response to treatment can be due to lack of an accepted consensus definition of responder, intra-subject variability in C-peptide AUC and lack of studies linking C-peptide with definitive clinical endpoints. Beam et al.[240] reviewed these issues when considering a definition of responders within insulin preservation studies, where they conclude that a prespecified definition of a clinical responder should be maintenance or increase of C-peptide from baseline, i.e. a threshold of 100% or more of baseline C-peptide AUC values. The responder status should be considered a supportive analysis, which in our case drove a number of interesting findings of immunological differences between responder and non-responders, which were not present between treatment groups. Importantly tightening up the threshold to 100% would not have changed the divide between responders and non-responders in this study, which reinforces confidence in allocation of responder status. The limitations of responder analyses have been previously reported by Greenbaum et al.[225] highlighting natural variations in the decline of C-peptide leading to some individual fluctuating in and out of responder status over time.

Immunological responses were examined through a number of methods. IL-10 responses are considered a key mechanism of tolerance induction in ASI and PIT. ELISPOT data showed higher level of IL-10 responses to proinsulin in association with high frequency treatment with trends for higher IL-10 responses to proinsulin in peptide-treated responders. A higher fold-change in Treg expression of FoxP3 was seen in peptide-treated responders; in particular this was most notably in memory Tregs co-expressing CD39. Such cells have been associated with controlling inflammation through IL-10 secretion[241]. The memory subsets showing FoxP3 upregulation were also Helios negative, suggesting that these were peripherally generated, adaptive Tregs. It is proposed that autoantigen-specific CD4 T-cells with immunoregulatory properties are induced and suppress bystander inflammatory responses to the same epitope, autoantigen or related autoantigens being presented in cis by the same antigen-presenting cell (APCs)[121]. There is also evidence that under these conditions APCs are licensed to induce new cohorts of regulatory T-cells in a process termed "infectious tolerance"[242]. Such infectious tolerance may be a mechanism of action by which proinsulin C19-A3 could induce IL-10+ proinsulin-specific CD4 T-cells and/or adaptive Tregs, leading to amelioration of disease. Lower levels of CD57 expressing CD8 T-cells were found in peptide-treated responders, with CD57 expression an indicator of recent antigen exposure[237]. This finding may indicate that subjects who show restricted activation of autoreactive cytotoxic T lymphocytes may be more amenable to immune effects induced by PIT.

In summary this study has established the safety of proinsulin C19-A3 in adults newly diagnosed with T1D. This approach would be appealing for prevention strategies in T1D, which would on the whole be targeting children, therefore necessitating very positive safety profiles. In this arena, PIT offers significant advantages over biologic agents that carry short and long-term risks. Future studies are needed to explore efficacy and responses from both adult and paediatric cohorts.

5 Optimisation of an antigen stimulation assay for use in a longitudinal clinical trial.

5.1 Introduction

The antigen stimulation assay described in section 2.9 was based on work carried out by two groups refining capture of an antigen-specific response in live CD4+ T-cells after *ex-vivo* stimulation through detection of induced surface CD154 expression[205,206,243,244]. The methods have been shown to be robust and sensitive in detecting potentially rare activated CD4+ cells[244] and, with CD154 upregulation via bystander activation being low and having low constitutive levels, this marker is more specific than other activation markers such as CD25, CD134, CD69, CD137 or HLA-DR[243]. The assay was validated further within our own laboratory by comparing antigen-specific populations enriched via this method versus antigen specific responses detected by ELISPOT[178]. Using sensitive ELISPOT techniques in an index donor ~0.012% PBMCs were detected to secrete IFNγ in response to short-term (16 hour) stimulation with $HA_{306-318}$ peptide. In comparison, using $HA_{306-318}$ in the antigen stimulation assay and sorting CD154+/CD69+ CD4+ cells, 24.5% of 49 clones obtained and tested were specific for $HA_{306-318}$ when using IFN-γ production as the read-out, which indicates ≥20,000-fold enrichment. True enrichment is likely to be even higher as CD154+/CD69+ cells are known to be polyfunctional (i.e. IFN-γ secretion alone may not capture all specific clones).

This assay was used within the MonoPepT1De trial to isolate the antigen-specific CD4 T-cells for downstream use in both the TCR sequencing and gene expression assays, therefore was important to optimise the assay for longitudinal use in a clinical trial. To ensure that the assay was standardised throughout the course of the clinical trial a number of measures were taken, including:

- Working with clear up-to-date standardised operating procedures (SOP) to good laboratory practice (GLP) principles.

- Samples were couriered and processed on the same day as the patient visit from all 5 clinical sites.

- Prior batching and allocation of compensation beads, co-culture antibodies (anti-CD28 and anti-CD40) and staining antibodies to individual subjects for all time points. Co-culture antibodies were separated into aliquots and frozen at -80^{o}C until use.

- Use of a single batch of AB serum and FCS within the assay at all time points.
- Use of a single flow cytometer and the use of the same machine for all samples from a patient, where possible. Two flow cytometers were available for the study, one within the department and another external.
- Consistent cytometer set-up and tracking (CST) settings throughout.
- Addition of a control sample at each assay, using cryopreserved PBMCs from a single donor stimulated with SEB.

5.2 Aim

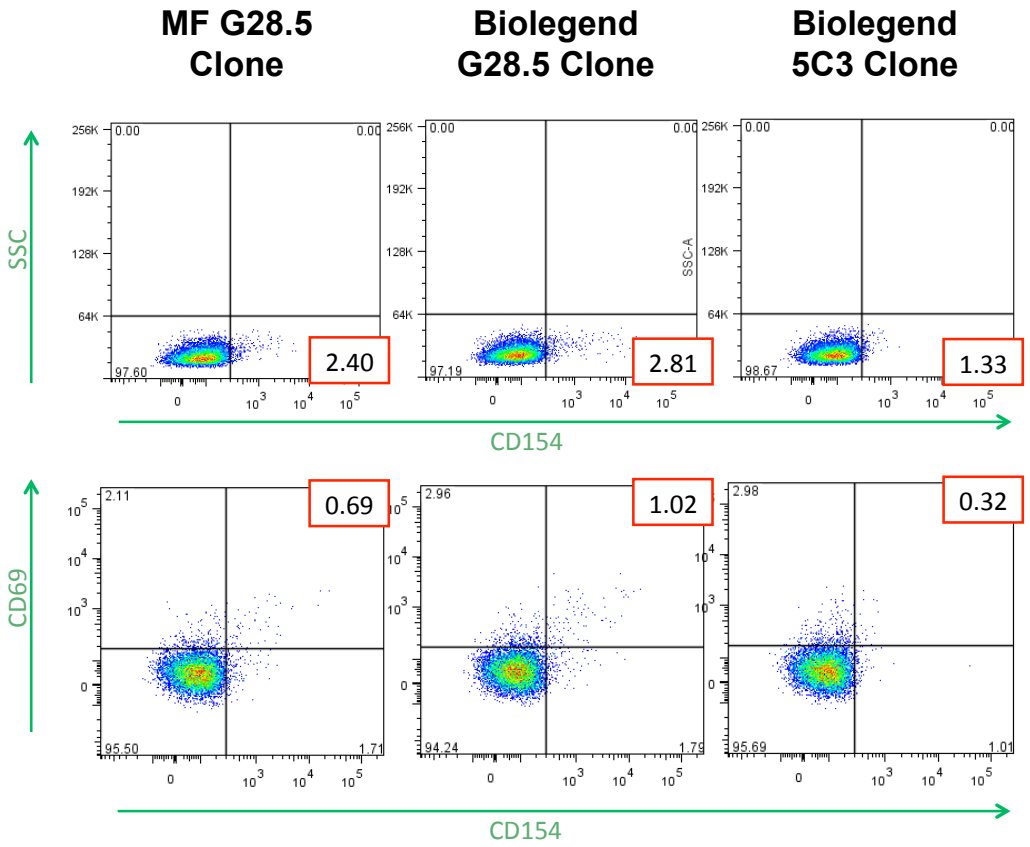
Describe measures taken to optimise the antigen stimulation assay and to standardise its use across a longitudinal clinical study.

5.3 Comparison of anti-CD40 antibody clones in antigen stimulation assay.

Addition of anti-CD40 blocking antibody prevents internalisation of CD154 allowing optimised collection of live antigen-responsive CD4+ T-cells[206]. The assay was previously performed successfully within our laboratory using an anti-CD40 (G28.5 clone) antibody donated from M. Frentsch (Berlin). In order to find a suitable replacement for the duration of the trial, this antibody was compared to two commercially available antibodies; another G28.5 clone and a 5C3 clone (both Biolegend, London, UK). The stimulation assay (as outlined in section 2.9) was carried out using Pediacel® as an antigen stimulus with cells from a healthy donor with the three versions of anti-CD40 antibody stained with anti-CD4, anti-CD154 and anti-CD69 fluorochromes. Comparative flow-cytometry plots are shown in Figure 30.

Figure 30: Selection of anti-CD40 clone.

PBMCs from a healthy donor were cultured in the presence of anti-CD28, one of 3 anti-CD40 antibodies and Pediacel®. Figure shows flow-cytometric plots showing comparative frequency (shown in red box) of CD154+ cells (top panel) and CD154+/CD69+ cells (bottom panel) using each anti-CD40 antibody.



5.4 Assessing functionality of antigen specific CD69+/CD154+ cells

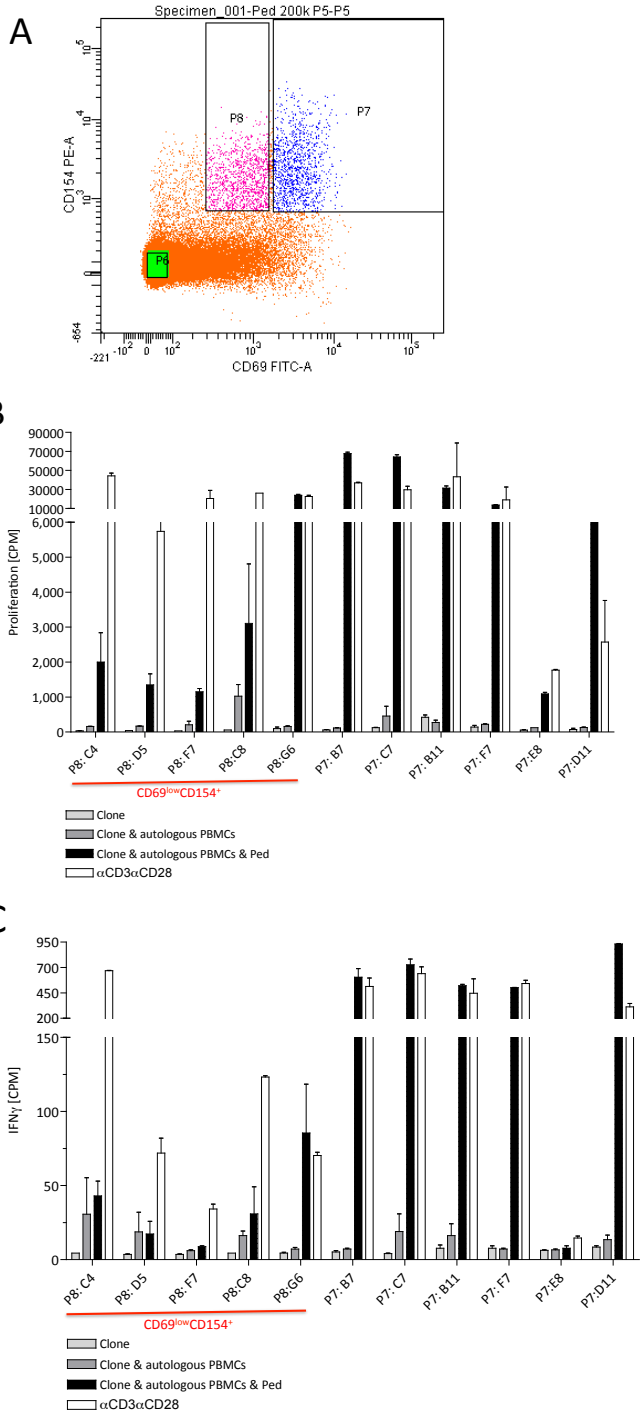
TCR and gene expression experiments rely on the isolation of CD4+ T-cells exhibiting the activation markers of CD69 and CD154 in response to antigen stimulation. To identify the optimal flow cytometry gating strategy, experiments were performed in collaboration with lab members, Dr H. Shariff and Dr V. Gibson. The aim was to assess antigen specific clones in terms of antigen-induced proliferation and IFN-γ production, across two gates i. $CD69^{high}/CD154^{+}$ gate (named P7) and ii. $CD69^{low}/CD154^{+}$ (named P8) (Figure 31A). This would allow the collection gate to be clearly defined by both phenotype and functionality.

PBMCs from a healthy donor were single cell sorted after antigen stimulation with Pediacel® as outlined in section 2.9, and antigen-induced proliferation and IFN-γ production were quantified (section 2.9.3). 5 clones which were responsive to Pediacel® stimulation were

identified from the $CD69^{low}$ $CD154^{+}$ (P8) gate and 6 positive Pediacel® clones from $CD69^{+}CD154^{+}$ (P7) gate. A significantly higher rate of proliferation was seen in response to antigen in clones from $CD69^{high}/CD154^{+}$ (P7) gate than the $CD69^{low}/CD154^{+}$ (P8) clones (Figure 31B). IFN-γ production was also significantly higher in the $CD69^{high}/CD154^{+}$ (P7) clones (Figure 31C).

Figure 31: Assessing functionality of antigen-specific clones exhibiting early activation markers.

A. Plot showing gating of flow cytometry CD154/CD69 staining populations (see Figure 11 *for full gating strategy). B. Graph showing proliferation across clones sorted from gates P7 and P8. Proliferation is divided into isolated clones, or in the presence of autologous PBMCs with or without Pediacel® stimulus and a positive control with the addition of beads coupled with anti-CD3 and anti-CD28 antibodies. C. Graph showing IFNγ production analysed via ELISA under conditions set out in Graph B.*



5.5 Longitudinal SEB stimulated donor controls

At each flow cytometry experiment, a positive control was added in the form of a thawed cryopreserved PBMC sample from a control donor stimulated with SEB. The aim of this control was to provide an assessment of the reproducibility of the antigen simulation assay over the course of the trial. The PBMCs were cryopreserved from 3 venesection collections from a single haemochromatosis donor over the course of 3 months. The SEB stimulated control samples were incorporated into the assay during a period of over 2 years, between July 2012 and August 2014. The percentage of $CD69^{+}/CD154^{+}$ cells (from the parent CD4+ population) from all SEB stimulated cryopreserved samples are plotted in Figure 32. The data are also divided into two further groups depending on which of two flow cytometers were used to analyse the sample, internal departmental equipment versus external. Coefficients of variation (CV) values of these data were high, with CV being 58.8% in total data and 22.1% in the external group versus 61.8% in the internal group. In view of high CV values, potential outlying results were examined. Technical errors (marked in Figure 32) were suspected for two very high data points, identified possibly due to:

1. Excess SEB stimulation - supported by disproportionately high $CD69^{+}/CD154^{+}$ cell frequency in both SEB control group and the SEB patient samples analysed on the same day.

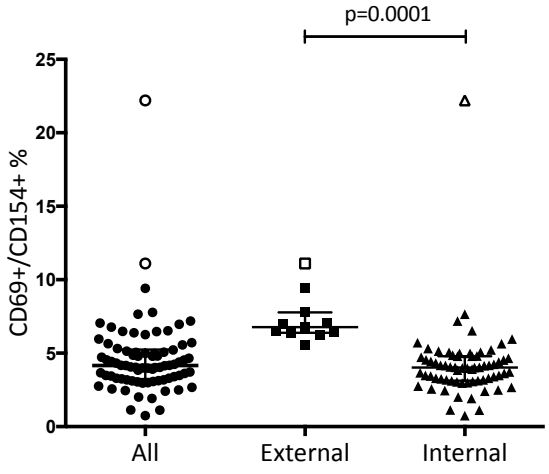
2. Excess co-stimulation with anti-CD28 antibody - supported by over 10-fold increase from baseline in mean cell numbers across all stimuli.

Cryopreserved cells were used to enable repeated sampling from a single control over the course of the trial, as fresh control samples would not be feasible in this setting. To analyse the long-term effect of cryopreservation on these samples the frequency of $CD69^{+}/CD154^{+}$ post SEB stimulated was plotted chronologically (Figure 33). Figure 33 shows the variation in CV values during 6-monthly intervals across the trial. Cell viability was assessed using cell yield from thawed cryopreserved PBMCs with a mean (SD) of 74.0% (±12.2%), CV 16.5%. This was also analysed across 6 monthly intervals, but there was no significant difference between mean yields across the duration of the trial (not shown). The method of cryopreservation has been standardised throughout the trial as such methods can influence antigen-specific responses, through assays such as ELISPOT[245]. Supplementary experiments performed by a lab colleague, Dr. Iria Gomez-Tourino, illustrate the fact that cryopreservation can have significant effects on activation of antigen-specific cells. These examined the effects of a 3.3-

month cryopreservation period, on the antigen-specific response of PBMCs to GAD and PI peptides, control protein, Pediacel® and PBS in terms of IFN-γ and IL-10 production (Figure 34, Figure 35). As changes in ELISPOT SI response were bidirectional, results from statistical analysis with paired t-test or Wilcoxon rank test were non-significant and may not illustrate the variability of SI following cryopreservation. An alternative way to evaluate two assays is the Bland Altman analysis (Figure 36), which plots the difference between the SI values from the two assays versus the average. Also plotted are the mean biases and the 95% limits of agreement, defined as mean bias ±2SD. The plots shown in Figure 36, illustrate that there are significant discrepancies between the fresh and thawed assays with no clear relationship between results. Wide bands of 95% limits of agreement would be of clinical significance; therefore interpretation of results from thawed cells would be unreliable.

Figure 32: Percentage of $CD69^{+}/CD154^{+}$ cells from parent CD4 population within SEB stimulated control samples.

Total data are shown in addition to data divided by internal or external flow cytometry facility. Lines indicate median and interquartile range. P value calculated using unpaired t-test with Welch's correction for unequal SD, on log-transformed data. Unfilled symbols indicate samples with identified technical error.



	CV	CV	CV
Including open symbols	57.8%	22.1%	61.3%
Excluding open symbols	37.5%	15.2%	33.1%

Figure 33: Variation in frequency of CD69+/CD154+ cells post SEB stimulation of cryopreserved cells over time.

Each interval indicated equals a 6-month period, with corresponding CV for that period labelled above. Unfilled circles indicate samples where technical error has been identified.

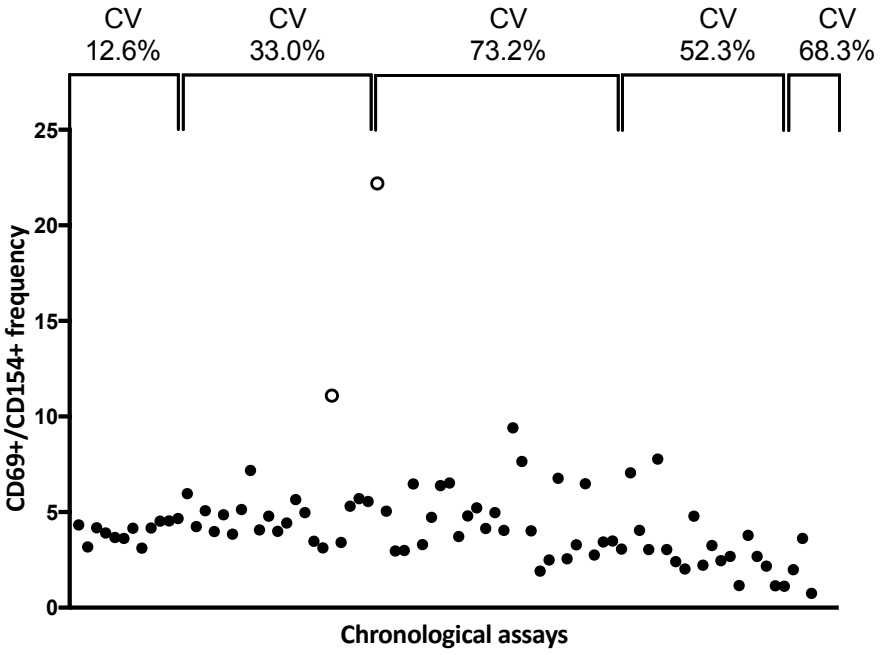
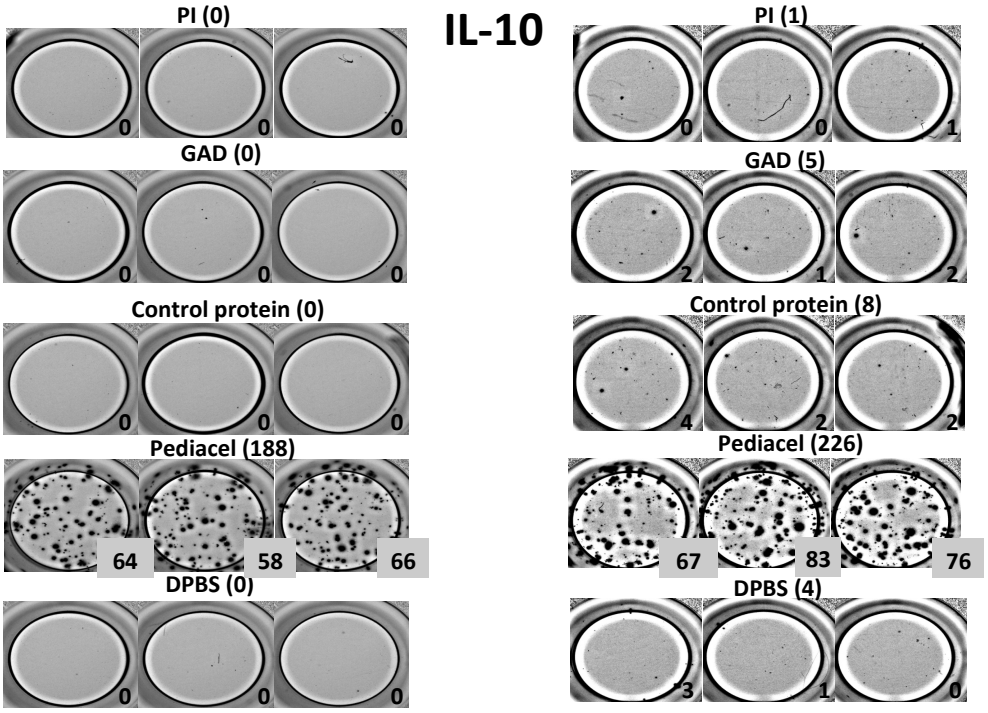


Figure 34: ELISPOT images of IL-10 and IFN-γ antigen specific responses using fresh and cryopreserved PBMCs from a single individual. Courtesy of Dr. Iria Gomez-Tourino.

Images show triplicate wells with number of responsive cells in each well labelled in the right-hand bottom corner. The total number of responsive cells is shown in brackets next to the stimulating antigen.



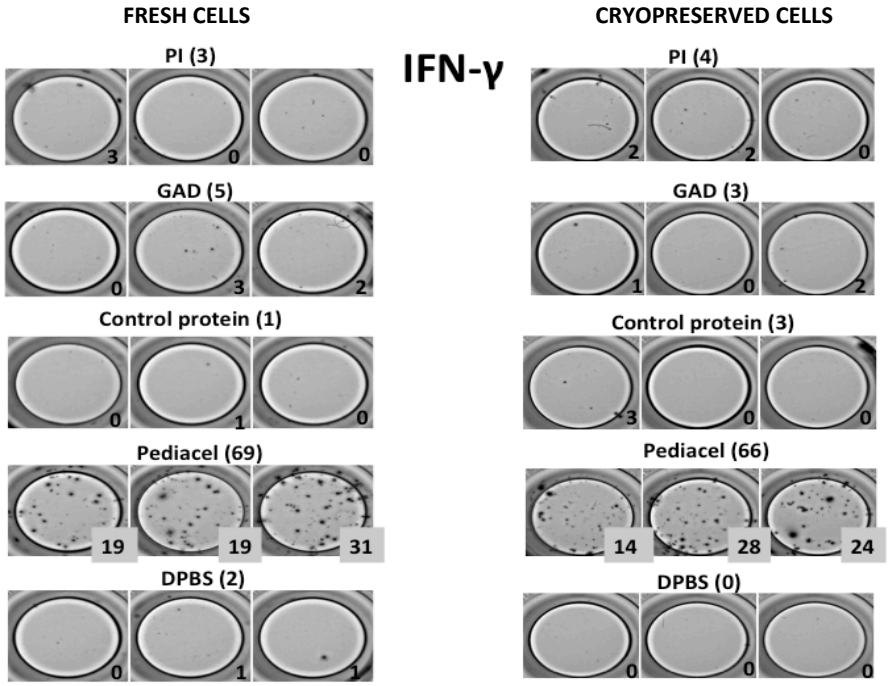


Figure 35: Differences in stimulation index (SI) in ELISPOT assays performed on fresh and thawed cryopreserved cells. Data courtesy of Dr Iria Gomez-Tourino.

All conditions showed non-significant differences using either paired t-test or Wilcoxon rank test.

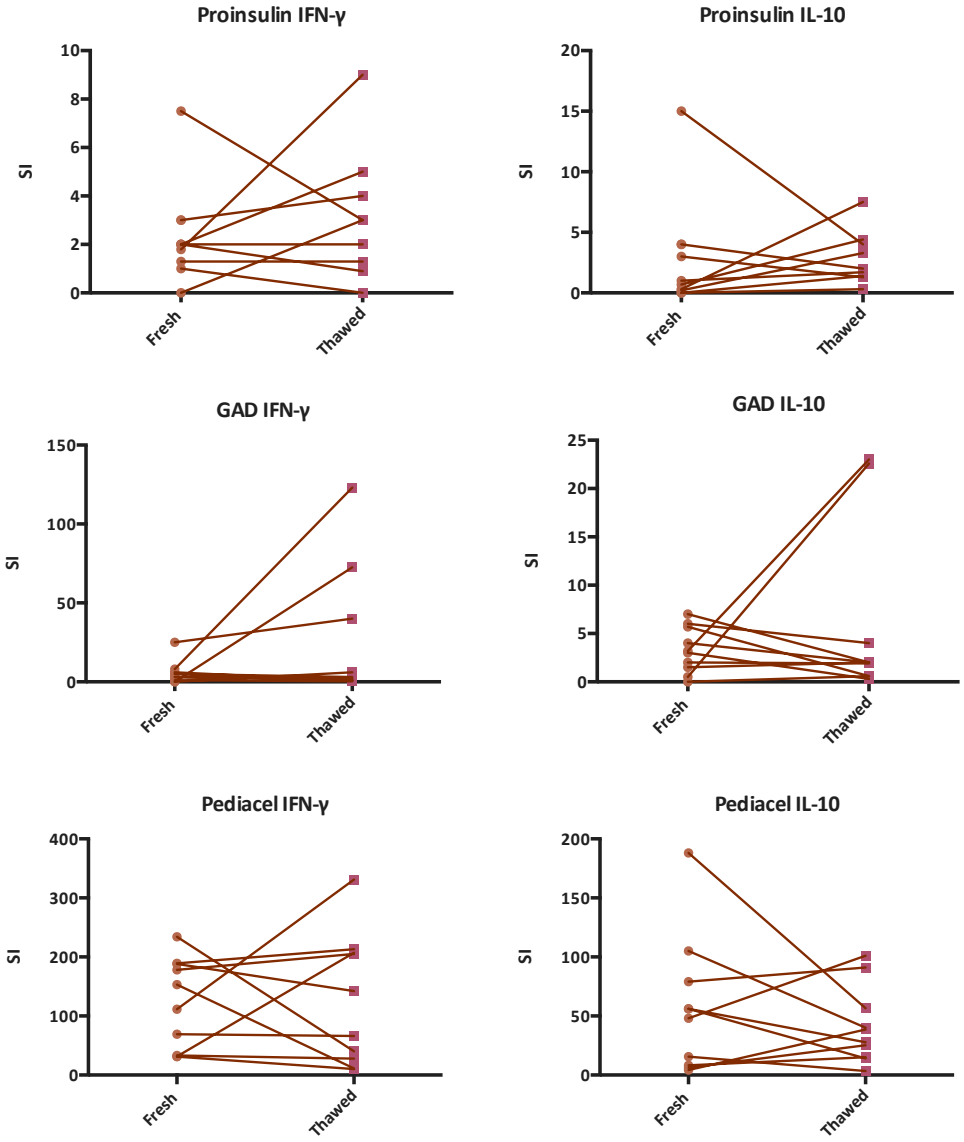
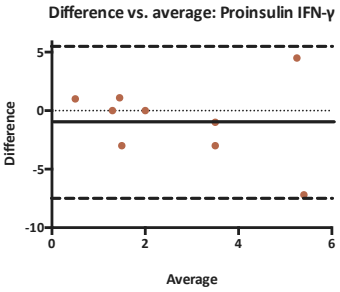
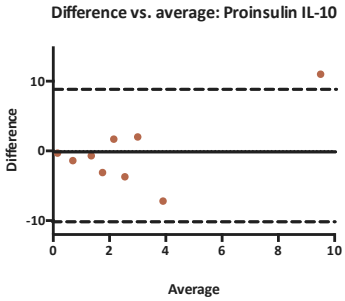
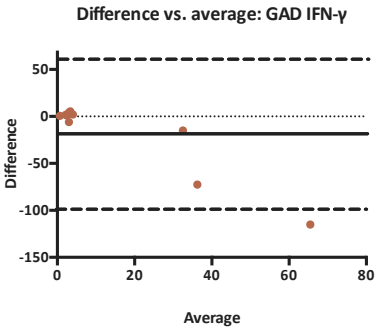


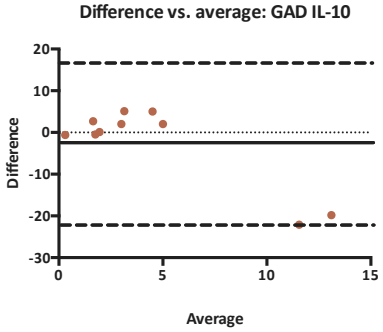
Figure 36: Bland-Altman analysis of ELISPOT assays on fresh and thawed cryopreserved cells. Data courtesy of Dr. Iria Gomez-Tourino.

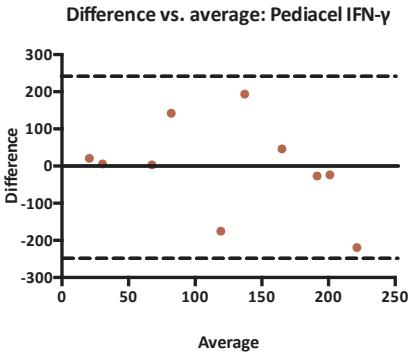
The solid lines mark mean bias and dashed lines mark the 95% limits of agreement for stimulation index.

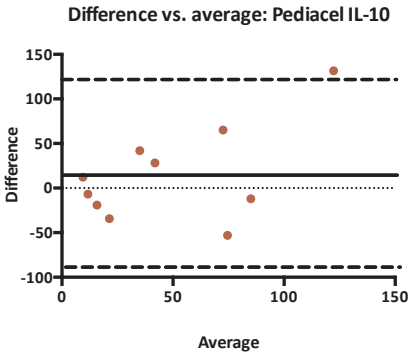












5.6 Discussion

Standardisation of protocols is critical in the implementation of immunological assays within longitudinal clinical trials. To ensure consistency of the stimulation assay within the MonoPepT1De trial, measures to control variation in reagents and cytometer configurations were considered from the outset.

Optimal identification of antigen-specific activation through detection of *de novo*-synthesised and expressed CD154 is reliant on extracellular stabilisation using an unconjugated anti-CD40 blocking antibody. The presence of CD154 is suggested to be a functional correlate of cognate immune response as it identifies cells which may provide B-cell help or activate APC function[205]. The added advantage of this assay over others is that live cell collection is possible, enabling further downstream analysis of such cells, in this case with TCR clonotyping and gene expression analysis. The results in section 5.3 demonstrate efforts to compare the action of anti-CD40 antibody clones. The trial adopted the use of the Biolegend G28.5 clone, which detected a higher frequency of $CD154^+$ cells than the previously used clone.

When examining cells for CD69 and CD154 positivity, there is a clearer distinction between CD154 positive and negative cells than CD69 positive and negative cells. This necessitated further functional identification of CD69+ cells in order to gate these cells effectively. The clones identified within the P8 gate showed less antigen-specific proliferation and IFN-γ production than those in the P7 gate. In view of these findings the cells were collected from the P7 gate without expanding collection to the P8 gate, to isolate cells with the greatest antigen-specific response. The functionality in terms of IFN-γ production was assessed via ELISA. Chattopadhyay et al[205] were able to demonstrate that CD154 cells showed multiple diverse T-cell function through intracellular staining and polychromatic flow cytometry.

Repeated stimulation of cryopreserved control samples using the super antigen SEB was performed to examine reproducibility of the assay longitudinally. The design and the duration of the trial required multiple operators to process assays, and during each experiment multiple operators shared variable points in the assay, therefore defining individual operator effects is not possible. Intra-assay variability was not assessed, as only one SEB control was present per experiment. Inter-assay variability was measured using the percentage of $CD69^+/CD154^+$ cells from the parent CD4 population. Although there was a significant difference in CV values between different flow cytometer facilities (Figure 32), only a limited number of experiments were performed externally (14%) which may have contributed to the

lower CV value. With overall CV values from the SEB control being high, figures were examined longitudinally to reveal deterioration in CV values from an initial acceptable value of 12.6% up to a peak of 73.2%. Exclusion of two incidences of clear technical error can more clearly show the effects of this deterioration, presumably due to length of cryopreservation. In doing so stepwise increments in CV can be seen with each 6-month period, from 12.6% to 20.7%, 39.9%, 52.3% and 68.3%. Cryopreservation being used in control cells alone would suggest that the high CV levels do not reflect on the reproducibility of the stimulation assay, but would have implications on use of patient cryopreserved samples.

Importantly these changes in response were not predictable with rise or falls varying between individuals. PBMCs which have been cryopreserved for a period of one month, have been shown to have comparable results to fresh PBMCs when used in peptide stimulation cytokine flow cytometry assays[246]. Supporting work by Dr Gomez-Tourino shows unpredictable discrepancies between antigen-specific responses pre and post a cryopreservation period of 3.3-months. The samples used within the trial were cryopreserved for approximately 2 years with deterioration in reproducibility being present from 6 months. There is little published data on the effects of long-term cryopreservation on PBMC responses, however Owen et al.[247] examined the effects of cryopreservation on a control group (n=8) against HIV positive samples, which were banked for over 300 days. They showed that PBMC samples were able to retain CD4+ and CD8+ T-cell IFN-γ responses to CMV pp65 after long-term cryopreservation regardless of HIV status. Although there was a decline in CD4+ responses in controls, the authors felt that the rank order of the responses was retained and consistent enough to have comparable results on a population basis. Nonetheless HIV- samples were only tested for IFN-γ response against CMV pp65 leaving the possibility of individual responses to other antigens being affected as illustrated by the ELISPOT example (Figure 34) where antigen-specific responses can vary depending on antigen stimulus and cytokine response.

The use of a stimulation assay within the MonoPepT1De trial has required generalised standardisation of reagents and lab procedures as well as assay specific considerations. I have described a number of stages of development, including antibody comparison and identifying functionality of cells phenotyped through flow cytometry. The choice of a stable control comparator (Section 5.5) in this case was flawed due to the effects of cryopreservation on

CD4+ SEB responses detected by flow cytometry, and may have implications on the use of cryopreserved samples for assessing assay drift in future studies.

6 T-cell receptor clonotyping

6.1 Introduction

It is postulated that changes in antigen-specific TCR populations occur in response to PIT with shifts towards populations with a tolerogenic phenotype and/or deletion of pathogenic clonotypes (Section 1.6.1). High-throughput sequencing of β-chain clonotypes specific for PI C19-A3 was performed with the aim of detecting changes in dominant populations. Isolating CD4+/CD69+/CD154+ cells post stimulation with PI C19-A3 allowed β-chain sequencing on an enriched rare antigen-specific population. The TCR sequencing methods used were previously validated within our laboratory, to show a strong correlation between observed and expected clonotype frequency (R2=0.98, p=0.001) when identifying a known clonotype spiked into a background polyclonal population[248]. At the lowest spike-in frequency (1/500; 0.2%), the assay has a mean 80.0% sensitivity, and precision is also highest at this level of detection.

6.2 Aims

1. To discuss validation of a multiplex primer system for optimisation of HTS of TCR β-chain repertoire.

2. To investigate the size and structural diversity of an antigen-specific β-chain TCR response to PI C19-A3 in patients newly-diagnosed with T1D in response to PIT or placebo.

3. To identify β-chain clonotypes (defined as distinct TCR sequences) which are public (or shared) versus private (only found in the individual) that may be linked to pathological immunodominant epitopes of disease or those that may be induced by PIT.

6.3 Validation of multiplex identifier labelled primer panel

6.3.1 Introduction

Barcoding of PCR products optimises efficiency of high-throughput TCR sequencing. Within these experiments a multiplex identifier (MID) tags was assigned to each sample. These MID

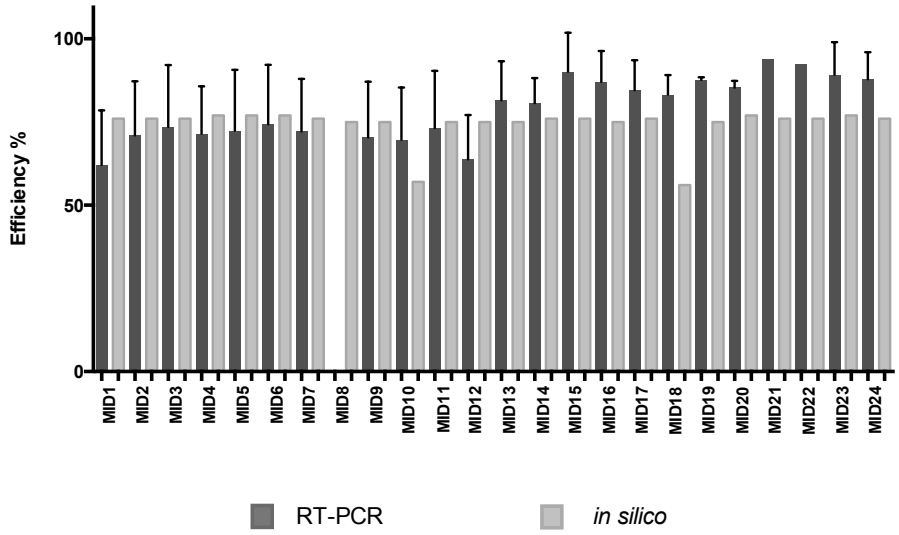
tags were engineered by Roche (Sussex, UK) to complement their Roche® 454 pyrosequencing technology, with the aim of avoiding mis-assignment of reads whilst being tolerant to several types of errors, including those often introduced during primer synthesis. The use of 24 MID label sequences linked with our TRBC-R reverse primer was validated in a number of ways. Each MID-labelled primer sequence underwent a NCBI nucleotide sequence BLAST[249] search to identify target sequences. The secondary structures of MID-labelled primers were interrogated *in silico* for hairpin formation, dimerization to self or to the forward primer, 5'RACE F using the PrimersList webtool[250]. Subsequently *in vitro* testing of relative efficiencies of MID labelled primers was performed through RT-PCR (2.11.6). Each PCR run contained triplicates of 12 MID samples and was repeated in triplicate. Due to lack of first round PCR product to use for MID validation, RT-PCR was performed directly using cDNA cloned using a plasmid vector with transformation into *E. Coli*. The clone chosen was a HA specific T-cell clone which was previously successfully sequenced. Samples with an R^2 value of <0.99 were excluded.

6.3.2 Results

The NCBI BLAST search confirmed high specificity for the human T-cell receptor beta chain remained despite MID labelling, with expect values of <0.001.

MID-labelled primer efficiencies ranged from 56-77% when estimated *in silico*, whilst estimates derived from RT-PCR were 62-94% (Figure 37). The unlabelled TRBC-R primer had an efficiency of 74% *in silico*.

Figure 37: MID labelled primer efficiencies estimated through RT-PCR or *in silico*



In silico analysis was able to estimate melting temperatures (Tm) of the primers, with 5'RACE F primer having a Tm of 58.7^{o}C and TRBC-R primer a Tm of 65.3^{o}C. The addition of MID labels changed the Tm to a range between 68.6-72.7^{o}C. The target Tm for primer design is 52-58^{o}C with Tm of primer pairs being within 5^{o}C of each other. Tm of >65^{o}C can also lead to greater secondary annealing.

MID8 had the highest number of potential self-dimers in the in-silico analysis, one of which being a potential G/C-quadruplex. Reactions using MID 8 failed to show exponential amplification on RT-PCR on multiple attempts. Further experiments to optimise primer concentrations for MID 8 using a primer matrix were also unsuccessful.

6.4 TCR β-chain clonotyping of patient samples

6.4.1 Introduction

Collection of CD69+/CD154+ cells after PI C19-A3, HA, SEB and PBS stimulation took place at baseline, 1 and 5 months of treatment with subsequent TCR β-chain sequencing as described in 2.11. The number of samples that could be processed were limited by time, cost and number of valid MID tags, therefore the longitudinal samples of PI C19-A3 stimulated cells were prioritised for sequencing. Unblinding of treatment groups was not performed until sequencing was completed. Analysis of high-throughput sequencing data is complex and

potentially can incorporate bias from sample size. Data analysis was performed with bioinformatics support (Dr. Anna Lorenc, KCL).

Diversity measures in TCR populations can determine whether there are detectable shifts in clonotypes. For the purposes of this thesis, diversity refers to structural diversity resulting from the recombination of the VDJ regions, in contrast to functional diversity which refers to the range of different effector functions from activated T-cells which share specificity for a single epitope. When assessing diversity of a TCR population, diversity measures can consider the number of different clonotypes present as well as distribution of clone size (number of copies of a distinct sequence) or relative abundance of clonotypes as the number of clonotypes can be biased by sample size. Measures of diversity have been adapted from the field of ecology where diversity of species within an environment equates to clonotypes and number of individuals within a species equates to clone size. Such measures can help identify clonal dominance over identification of absolute clonal numbers. Three measures of diversity were used in this TCR data, the true diversity, the Gini index and the Chao1 index. Using a combination of diversity measures allows for assessment of differing aspects of diversity. True diversity is defined as the number of equally abundant types needed for the average proportional abundance of the types to equal that observed in the dataset of interest where all types may not be equally abundant. The Gini index is a measure of inequality derived from economic data analysis of distribution of wealth, a score of 0 equating to a perfectly even distribution of species and higher scores suggesting greater inequality. The Chao1 index indicates the richness of a population. True diversity is the reciprocal of the weighted generalized mean of the proportional abundances of the types in the dataset.

Table 26: Measures of TCR diversity used in analysis with corresponding mathematical formulas.

True Diversity	$\frac{1}{(\sum \sqrt{frequency})^{-\frac{1}{2}}}$
Gini index	$\frac{1}{n} * (n + 1 - 2 * (\sum (n + 1 - 1 : n) * frequency / \sum frequency))$
Chao1 index	$S_{chao1} = S_{obs} + \frac{n_1 (n_1 - 1)}{2 (n_2 + 1)}$

During sequencing the presence of low clone size can indicate sequencing errors or the presence of rare clones, which may skew results especially in larger samples. To reduce this effect single reads were excluded from diversity analysis for true diversity and Gini index calculations, however as an estimator of total richness, Chao1 calculation included all data.

Diversity was measured at both the nucleotide levels and amino acid level and there were no significant differences between these in any of the diversity measures. Examining changes in diversity across treatment groups, there were no significant changes detected in any of the diversity measures over time.

6.4.2 Results

A number of samples required repeated amplification steps to enable adequate final DNA concentrations. Final DNA concentrations were variable as initial RNA concentration could not be predicted from cell number and small volumes of eluates at each stage made quantification of nucleic acid difficult without significant loss of product. Only at gel electrophoresis stage would a visible band at approximately 500bp provide an indication of adequate product and even beyond this step a number of samples failed quality control by fragment analysis prior to pooled multiplex sequencing. With the added issue of missed patient samples, final results were limited in number with gaps in longitudinal data (Table 27).

Table 27: TCR sequencing sample summary.

Samples were divided into two datasets that were sequenced on separate runs. Left hand columns indicate patient numbers with treatment groups of placebo (P), high frequency dosing (H) and low frequency dosing (L). Visits 1 (baseline), 3 (1 month) and 9 (5 months) are indicated by V1, V3 and V9 respectively. MID allocation is indicated for each patient and visit number. Lower half of table indicates number of patients and samples in each treatment allocation.

DATASET 1				DATASET 2			
	V1	V3	V9		V1	V3	V9
1008 P	MID 4		MID 11	2025 H	MID 15		MID 22
1017 H	MID 10	MID 22	MID 2	2027 L		MID 5	MID 2
1025 P	MID 4	MID 13	MID 7	2031 L	MID 4	MID 11	MID 7
2001 L	MID 23		MID 24	3005 P	MID 23		MID 24
2010 H	MID 3	MID 14	MID 6	3012 P	MID 3	MID 17	MID 6
2017 H	MID 18		MID 17	3015 H	MID 14	MID 18	MID 13
2018 L			MID 21	4005 H	MID 20	MID 19	MID 21
2019 P			MID 16	4007 L	MID 10		
2020 L	MID 12			5001 L	MID 12		MID 1
2021 H	MID 15	MID 20					

	Patients	Samples		Patients	Samples
Placebo	3	6	Placebo	2	5
High Freq	4	10	High Freq	3	8
Low Freq	3	4	Low Freq	4	8

In total both datasets yielded data for 78,107 sequences. Normalisation of data was considered in view of variations in cell numbers between samples that could have influenced the diversity of sequences identified and bias towards greater diversity in larger samples. Median cell number per sample was 711 (range 80-12,968). No correlation was found between cell number and total number of sequences (all or only productive), nor unique CDR3 sequences (amino acid or nucleotide) (Table 28). Correlations were also not seen when data was split into visits.

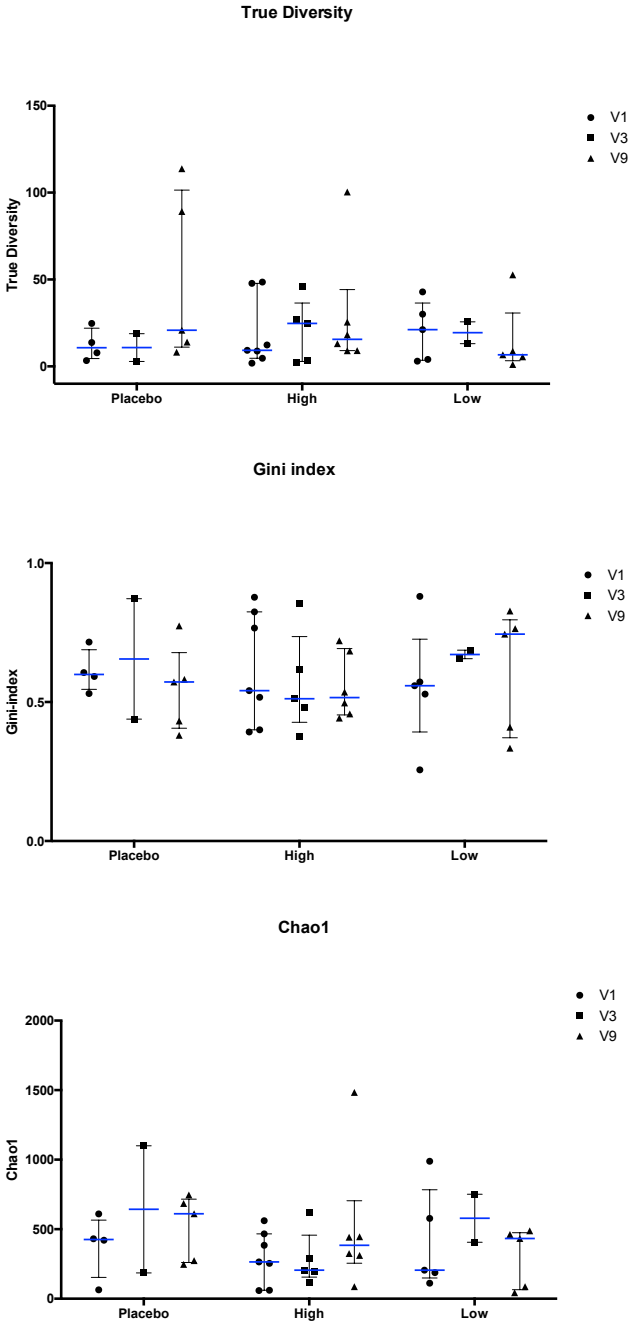
Table 28: Pearson correlation of sample cell number to sequence numbers

Pearson Correlation with sample cell number	R	95% confidence interval	R-squared	P value
Total number of sequences	-0.064	-0.365 to 0.248	0.004	0.690
Total number of productive sequences	-0.092	-0.389 to 0.2219	0.008	0.567
Number of unique productive CDR3 amino acid sequences	0.202	-0.112 to 0.480	0.041	0.2046
Number of unique productive CDR3 nucleotide sequences	0.199	-0.115 to 0.478	0.040	0.211

Figure 38 summarises the 3 diversity measures calculated for amino acid CDR3 sequences grouped by treatment and visit. There were no significant differences in calculated medians over time in each treatment group, nor at each time point between treatment groups.

Figure 38: Changes in diversity measures of amino acid level clonotypes in treatment groups over time

Graphs show true diversity, Gini-Simpson index and Chao 1 measures for individual samples in treatment groups over visits 1, 3 and 9. True diversity and Gini indexes were calculated excluding unique single sequences. Overlaying bars indicated median and interquartile range of for corresponding visit and treatment. No significant differences were found within treatment groups over time using these diversity measures using Kruskal-Wallis testing.



Sharing of β-chain CDR3 sequence between patients

Identification of identical β-chain CDR3 regions could be seen between multiple samples, both within and between patients. 125 (102 productive) sequences were shared at an amino acid level and 109 (90 productive) shared at a nucleotide level. In identifying the degree of convergent recombination involved in sharing of sequences, 18 of the shared amino acid level CDR3 sequences showed convergence from differing nucleotide sequences (Table 29: Shared amino acid beta-chain CDR3 regions showing convergent recombination. Most (8, 44.4%) of these differed by only 1 nucleotide but 9 (50.0%) differed by 2, 3 or 4 nucleotides and 1 (5.5%) differed by 5 nucleotides. There was a fall in number of convergent clonotypes with increasing number of alternative convergent nucleotides (Figure 39A). Similarly there was a fall in the number of samples sharing a convergent clonotype as the number of alternative nucleotides increased (Figure 39B); neither of these correlations was significant. Out of the 18 sequences that showed convergent recombination 13 were found in C-peptide responders, although this appears to be driven largely by a single patient (1025). This patient was in the placebo group and was the only clinical responder within this group. None of the shared sequences were germline, defined by the lack of insertions or deletions in the VDJ region. 3 germline sequences were identified in the total data.

Figure 39: Nucleotide variation in convergent clonotypes

(A) Graph showing distribution of number of convergent clonotypes with increasing number of alternative nucleotides within convergence. (B) Graph showing number of samples with convergent clonotypes with differing number of alternative nucleotides. Lines indicate mean and standard deviation. r and p values are derived from Pearson correlation co-efficient calculations.

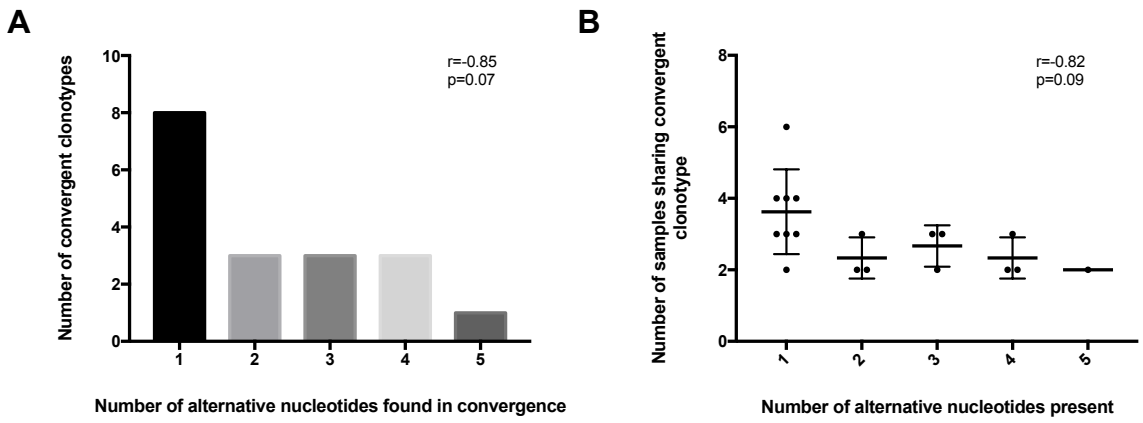


Table 29: Shared amino acid beta-chain CDR3 regions showing convergent recombination.

Underlined CDR3 nucleotide sequences show alternative nucleotide placement resulting in identical amino acid sequence. Functionality: P indicates productive whilst U indicates unproductive sequence.

Sample	CDR3 nucleotide sequence	Functionality	N	% of all seqs	No. of differing nuc	No. of samples sharing seq	No. of patients sharing seq
1017_V1_PI_MID10	TGTGCCAGCAGCCAAGGGGGCACCGGGGCCAACGTCCTGACTTTC	P	4	0.46	4	3	3
1025_V3_PI_MID13	TG**C**GCCAGCAGCCAAGG**A**GG**G**AC**T**GGGGCCAACGTCCTGACTTTC	P	2	0.08			
2019_V9_PI_MID16	TGTGCCAGCAGCCAAGGGGGCACCGGGGCCAACGTCCTGACTTTC	P	1	0.06			
1025_V1_PI_MID04	TGCGCCAGCAGCCCCGAGGGGCAGCCCCGTGAAAAACTGTTTTTT	P	96	10.04	1	3	2
1025_V1_PI_MID04	TGCGCCAGCAGCCCCGAGGGGCAGCCCCGTGAAAAACTGTTTTTT	U	1	0.10			
1025_V3_PI_MID13	TGCGCCAGCAGCCCCGAGGGGCAGCCCCGTGAAAAACTGTTTTTT	P	208	8.28			
1025_V3_PI_MID13	TGCGCCAGCAGCCCCGAGGGGCAGCCCCGTGAAAAACT**A**TTTTTT	P	5	0.20			
1025_V3_PI_MID13	TGCGCCAGCAGCCCCGAGGGGCAGCCCCGTGAAAAACTGTTTTTT	U	4	0.16			
5001_V1_PI_MID12	TGCGCCAGCAGCCCCGAGGGGCAGCCCCGTGAAAAACTGTTTTTT	P	1	0.04			
1008_V1_PI_MID05	TGTGCCAGCAGCCGGACAGGGGGCACTGAAGCTTTCTTT	P	31	6.64	1	2	2
1008_V1_PI_MID05	TGTGCCAGCAGCCGGACAGGGGGCACTGAAGCTTTCTTT	U	2	0.43			
2018_V9_PI_MID21	TGTGCCAGCAGCCG**C**ACAGGGGGCACTGAAGCTTTCTTT	P	13	0.92			
1008_V9_PI_MID11	TGTGCCAGCAGTTACAGCGGCCCGGGACTAGCGGGACGCACAGATACGCAGTATTTT	P	120	5.80	1	2	2
1008_V9_PI_MID11	TGTGCCAGCAGTTACAGCGGCCCGGG**G**CTAGCGGGACGCACAGATACGCAGTATTTT	P	2	0.10			
1008_V9_PI_MID11	TGTGCCAGCAGTTACAGCGGCCCGGGACTAGCGGGACGCACAGATACGCAGTATTTT	U	1	0.05			
2021_V1_PI_MID15	TGTGCCAGCAGTTACAGCGGCCCGGGACTAGCGGGACGCACAGATACGCAGTATTTT	P	1	0.04			

Table 29 continued: Shared amino acid beta-chain CDR3 regions showing convergent recombination.

Underlined CDR3 nucleotide sequences show alternative nucleotide placement resulting in identical amino acid sequence. Functionality: P indicates productive whilst U indicates unproductive sequence.

1017_V1_PI_MID10	TGTGCCAGCAGTTTCTTACAGGGAGGCTATGGCTACACCTTC	P	15	1.71	1	2	2
1017_V1_PI_MID10	TGTGCCAGCAGTTTCTTACAGGGAGGCTATGGCTACACCTTC	U	2	0.23			
1017_V1_PI_MID10	TGTGCCAGCAGTTTCTTACAGGGAGGCTATGGCTACAC**G**TTC	P	1	0.11			
1017_V3_PI_MID22	TGTGCCAGCAGTTTCTTACAGGGAGGCTATGGCTACACCTTC	P	2	0.45			
1017_V1_PI_MID10	TGTGCCAGCAGCTTGGCGGAGACCCAGTACTTC	U	2	0.23	3	2	2
1017_V1_PI_MID10	TGTGCCAGCAGCTTGGCGGAGACCCAGTACTTC	P	2	0.23			
5001_V1_PI_MID12	TG**C**GCCAGCAGCTT**A**GC**C**GAGACCCAGTACTTC	P	14	0.59			
1017_V9_PI_MID02	TGTGCCAGCAG**C**TTAGG**A**GGGGG**G**GAGACCCAGTACTTC	P	1	0.31	3	2	2
2018_V9_PI_MID21	TGTGCCAGCAGTTTAGGGGGGGGTGAGACCCAGTACTTC	P	9	0.64			
1025_V1_PI_MID04	TGTGCCAGCAGCGGAACAGGGGGCACCGGGGAGCTGTTTTTT	P	2	0.21	1	2	1
1025_V3_PI_MID13	TGTGCCAGCAGCGG**T**ACAGGGGGCACCGGGGAGCTGTTTTTT	P	7	0.28			
1025_V3_PI_MID13	TGTGCCAGCAAATTGCAAGGTCCGGGAAACACCATATATTTT	P	1	0.04	4	2	2
2017_V9_PI_MID17	TGTGCCAGCAA**G**TTGCA**G**GG**C**CC**T**GGAAACACCATATATTTT	P	10	0.20			
1025_V3_PI_MID13	TGTGCCAGCAGCTT**A**CAGGGGGC**C**ACAGATACGCAGTATTTT	P	5	0.20	2	2	2
2021_V3_PI_MID20	TGTGCCAGCAGCTTGCAGGGGGCAACAGATACGCAGTATTTT	P	33	1.19			
2021_V3_PI_MID20	TGTGCCAGCAGCTTGCAGGGGGCAACAGATACGCAGTATTTT	U	13	0.47			
1025_V3_PI_MID13	TGTGCCAGCAGTTT**A**AGGGA**T**AC**G**CAGTA**T**TT**T**	P	5	0.20	5	2	2
2010_V9_PI_MID06	TGTGCCAGCAGTTTGAGGGACACCCAGTACTTC	P	1	0.07			
1025_V3_PI_MID13	TGTGCCACCAGCAGGCGACAGGCGTCAGATACGCAGTATTTT	P	9	0.36	1	2	2
1025_V3_PI_MID13	TGTGCCACCAGCAGGCGACAGGCGTCAGATACGCAGTA**C**TTT	P	1	0.04			
2021_V3_PI_MID20	TGTGCCACCAGCAGGCGACAGGCGTCAGATACGCAGTATTTT	P	28	1.01			

Table 29 continued: Shared amino acid beta-chain CDR3 regions showing convergent recombination.

Underlined CDR3 nucleotide sequences show alternative nucleotide placement resulting in identical amino acid sequence. Functionality: P indicates productive whilst U indicates unproductive sequence.

1025_V3_PI_MID13	TGTGCCACCAGCAGTTACCCGGGACTTGCCAAAAACATTCAGTACTTC	P	1	0.04	3	2	2
1025_V3_PI_MID13	TGTGCCACCAGCAGTTACCCGGG**CT**T**A**GCCAAAAACATTCAGTACTTC	P	1	0.04			
2021_V3_PI_MID20	TGTGCCACCAGCAGTTACCCGGGACTTGCCAAAAACATTCAGTACTTC	P	13	0.47			
1025_V9_PI_MID07	TGTGCCAGCAGCTTAGTAGCGGGGGGTGGTGATACGCAGTATTTT	P	1	0.05	2	2	2
2021_V3_PI_MID20	TGTGCCAGCAGCTTAGTAGCGGGGGG**G**GG**A**GATACGCAGTATTTT	P	2	0.07			
2017_V9_PI_MID17	TGTGCCAGCGACAGGGGGCTGAACACTGAAGCTTTCTTT	P	1	0.02	2	2	2
4005_V9_PI_MID21	TGTGCCAGCGACAGGGG**C**CT**T**AACACTGAAGCTTTCTTT	P	98	3.57			
2019_V9_PI_MID16	TGTGCCAGCAGCTTAGGGCAGGGAGGTCAGCCCCAGCATTTT	P	1	0.06	4	2	2
3015_V1_PI_MID14	TGTGCCAGCAGC**C**T**T**GG**A**CAGGG**C**GGTCAGCCCCAGCATTTT	P	103	2.81			
2021_V3_PI_MID20	TGTGCCAGCAGCCAAGATCAAGCTCAGCCCCAGCATTTT	P	4	0.14	1	3	2
2021_V3_PI_MID20	TGTGCCAGCAGCCAAGATCAAGCTCAGCCCCAGCATTTT	U	1	0.04			
5001_V9_PI_MID1	TGTGCCAGCAGCCAAGATCAAGC**C**CAGCCCCAGCATTTT	P	5	0.19			
2031_V1_PI_MID4	TGCGCCAGCAGCAGTATGACAGGGAACACTGAAGCTTTCTTT	P	3	0.09	1	2	2
4005_V1_PI_MID20	TGCGCCAGCAGCAGTATGACAGGGAACACTGAAGCTTTCTTT	P	1715	59.65			
4005_V1_PI_MID20	TGCGCCAGCAGCAGTATGACAGGGAACACTGAAGCTTTCTTT	U	12	0.42			
4005_V1_PI_MID20	TGCGCCAGCAGCAGTATGACAGGGAACAC**C**GAAGCTTTCTTT	P	1	0.03			

When considering only the TCR sequences shared between C-peptide responders who received peptide treatment, 7 sequences were shared (Table 30). Five sequences were shared within samples for a single patient whilst 2 were shared between patients. Two sequences were found to have the presence of one or more frameshifts and were unproductive, whilst the others were all largely productive.

Table 30: Relative number of shared clonotypes between peptide treated C-peptide responders.

Table shows all TCR samples from C19-A3 treated patients who classed as C-peptide responders during the 6 months of treatment. Treatment groups are indicated by H (high frequency) and L (low frequency). Presence of shared amino acid TCR beta chain sequences are indicated as a percentage of the number of shared sequences from the total number of raw sequences for the sample. # indicates frameshift.

Relative number of shared clonotypes.	2021 V1	2021 V3	2027 V3	2027 V9	2031 V1	2031 V3	2031 V9	4007 V1	5001 V1	5001 V9
CASRPQGPGNTIYF	0.65	0.02								
CASSFTVGLMKN##F									2.05	5.87
CASSFTVGLMKNC#F									0.13	0.14
CASSFYGWANEKLFF									0.08	0.04
CASSLKGTGGSGANVLTF	0.40	0.16								
CASSQDQAQPQHF		0.10								0.18
CSGTGGRRTYNEQFF							0.04		0.84	
Treatment groups	H	H	L	L	L	L	L	L	L	L
Total number of shared sequences	2	3	0	0	0	0	1	0	4	4

CDR3 amino acid lengths of clonotype which were shared within all patients ranged from 11 to 22 amino acids with a mode of 15 amino acids. This distribution of CDR3 amino acid lengths in all patients and peptide treated responders is shown in Figure 40. CDR3 lengths shared between all patients were normally distributed, but those found only in peptide treated responders were too small in number to analyse. Frequencies of shared and private CDR3 clonotypes, as well as CDR3 clonotypes shared only within peptide treated responders are shown in Figure 41. There was no significant difference in frequencies between these groups.

Figure 40: Graph shows CDR3 amino acid length in all clonotypes shared between patients and also clonotypes which were shared between peptide treated responders.

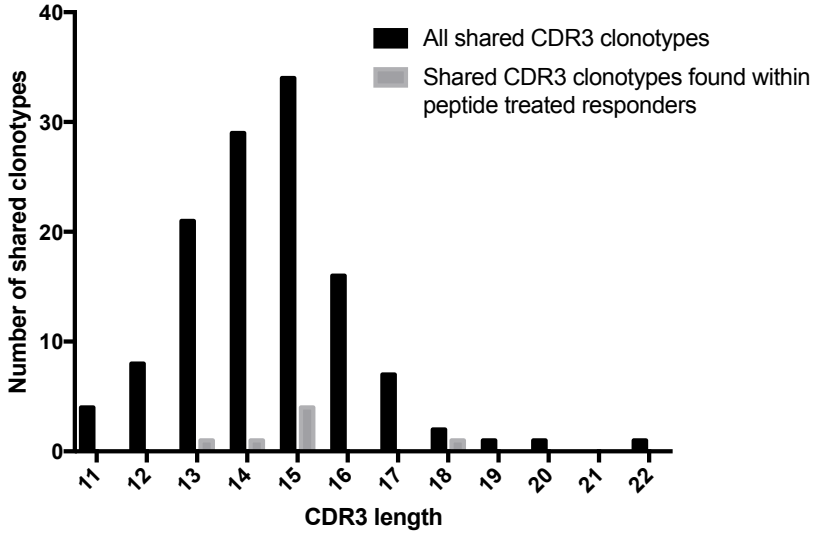
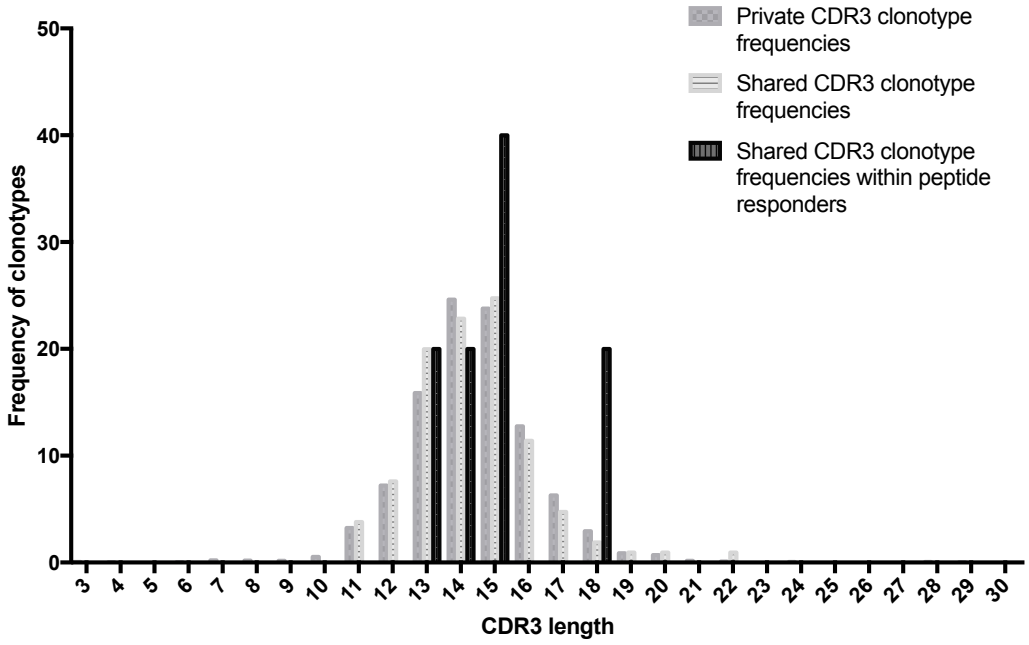


Figure 41: Graph shows frequencies of CDR3 clonotypes grouped as private, shared between patients and shared within peptide treated responders.



To identify differences in amino acid patterns between shared and private CDR3 sequences, iceLogo an open source application was used[251]. CDR3 sequences were divided into lengths with the exclusion of stop codons and frameshifts. Shared CDR3 sequences were then compared with private sequences of the same amino acid length. Using private CDR3

sequences as a reference set, iceLogo can generate sequence logos which show amino acids which are under or over represented compared to the reference set in each sequence position (Figure 42).

Figure 42: IceLogo derived sequence logos comparing amino acid sequences of shared versus private CDR3 sequences of varying lengths.

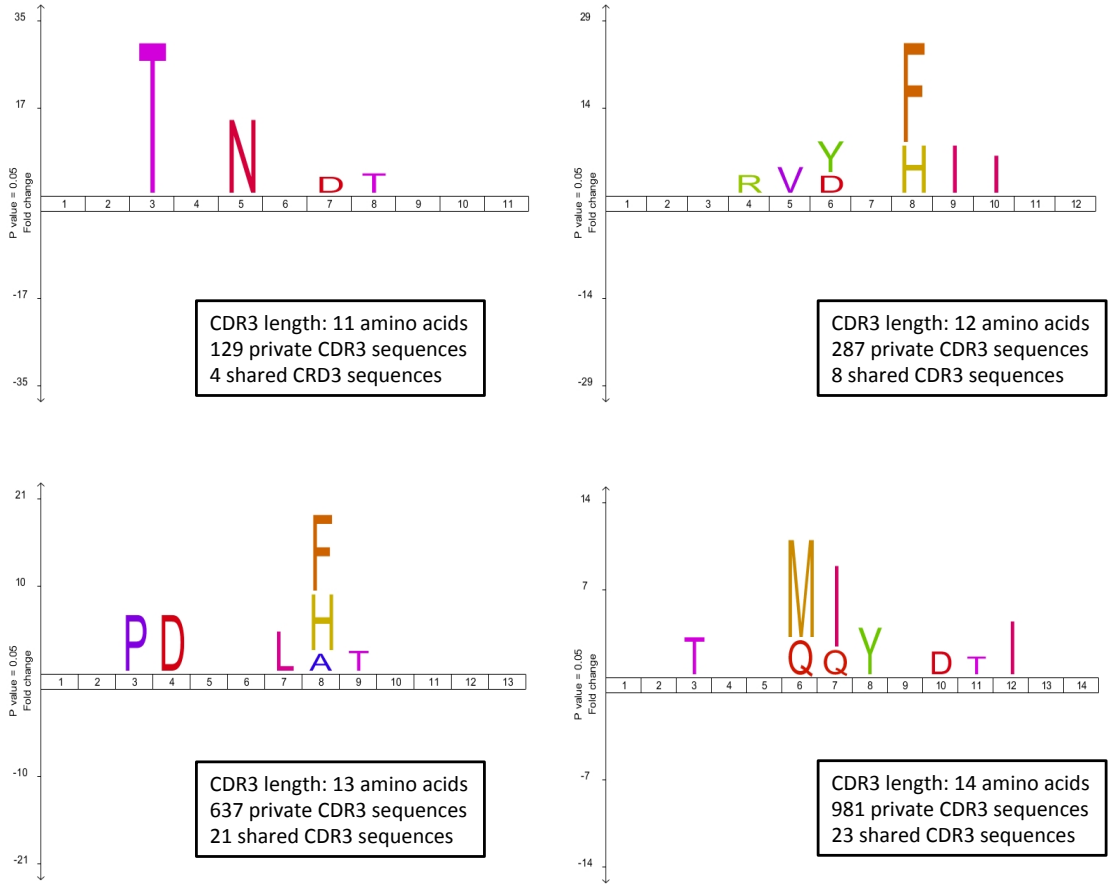
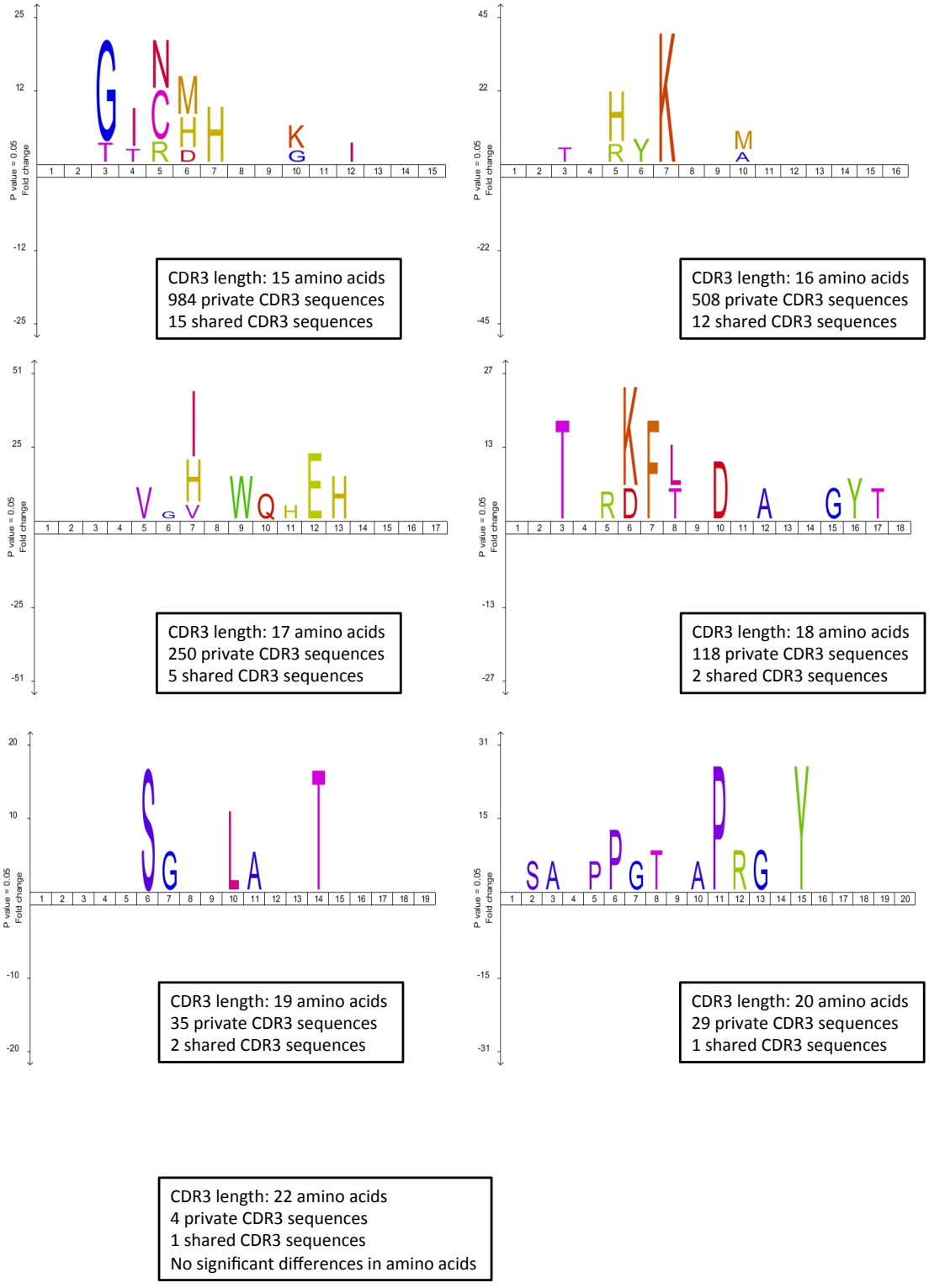


Figure 42 **continued: IceLogo derived sequence logos comparing amino acid sequences of shared versus private CDR3 sequences of varying lengths.**



6.4.3 Discussion

The incorporation of MID labelling within these experiments allows for high-throughput sequencing of samples. The effect of a multiplex label on primer efficiency was assessed in a number of ways. Although NCBI BLAST interrogation revealed that targeting of the primer was not influenced, further validation stages incorporated *in silico* and RT-PCR testing. Both of these methods showed suboptimal efficiencies, which may be due to a number of factors. Amplicons should ideally be 50-150bp in length for optimal RT-PCR efficiency, although for standard PCR this can be up to 500bp. At approximately 500bp, the TCR amplicon could reduce amplification efficiency when assessed through the means of RT-PCR. This is supported by the fact that the unlabelled TRBC-R primer also showed suboptimal efficiencies. Primer Tm was raised above optimal levels by the addition of the MID label, which can increase secondary annealing. Primers were chosen to sequence full-length TCR VDJ regions that could not be achieved with shorter amplicon length and therefore redesign of these primers was not an option. The choice of target cDNA would also have an influence on PCR efficiency. As volumes of cDNA for use in the second round PCR were very limited, the choice was taken to use an E.coli colony transformed with a TCR containing vector for direct PCR use. Inhibition of PCR can occur through bacterial cell debris due to high colony concentration or presence of LB agar. Such inhibition may reduce PCR efficiency but this would not be a factor in the final experiments.

The evaluation of MID efficiencies allowed the allocation of MIDs with approximately equal efficiency to longitudinal samples. In addition the RT-PCR experiments identified MID 8 as a primer to exclude from our high-throughput sequencing, this was not anticipated through *in silico* analysis. Although *in silico* analysis can be a useful initial screen for PCR design it can fail to account for all the components within an *in vitro* PCR reaction. The high number of self-dimers and the presence of a potential G/C quadruplex in the MID 8 labelled primer may have been a limiting factor and reflects the high GC content (66.7%, ideal range: 40-60%).

During data analysis the effects of varying cell numbers on the final sequencing was carefully considered. The samples were limited by differing cell number collection and the need for repeated amplification steps to provide adequate DNA, statistical analysis was used to undertand any bias introduced through these limitations and if adjustments were necessary. When encountering unequal sample sizes it is suggested that normalisation to smallest sample size can overcome some introduced bias[252], however within these samples there was

no correlation of cell number to sequence numbers and normalisation to cell number was not considered appropriate. There was also a large variation in cell number and as such normalisation to smallest cell number samples would further substantially limit data. Absence of correlation was also seen in other work using the same stimulation assay and sequencing methods[178]. The effect of larger cell numbers would nonetheless risk bias towards identification of rare clonotypes therefore exclusion of single clonotypes was used in diversity analysis. Diversity measures were limited by small numbers and missing data. There were no detectable significant differences between groups or over time, although there was a non-significant trend for a rise in Gini index mirrored by a fall in true diversity in the low frequency treatment group, suggesting the emergence of more dominant clonotypes over time.

Within the clonotype data, sharing of β-chain clonotypes between samples was commonly seen. To ensure that this was not due to contamination between samples, strict laboratory protocol was followed including individual sample analysis with aseptic practices, individual locations for pre and post-PCR steps with dedicated equipment, and introduction of negative controls for each assay step.

In a healthy adult population, frequency of shared CD4+ TCR β chain CDR3 clonotypes for an individual (n=8) was found to be between 5.2 and 11.4%[253]. When investigating the GAD-65 antigen specific response in CD4+ T-cells, sharing of isolated TCR α-chains in subjects with T1D was commonly found by Eugster et al.[179]. This study showed one individual with T1D shared 5% of their GAD-65 specific TCR α-chain clonotypes with all four other patients tested and shared 35% of GAD-65 α-chain clonotypes with at least one other patient. However pairing analysis revealed sharing of GAD-65 α-chain clonotypes did not correspond to a shared β-chain and all samples which shared either an α-chain or β-chain were found to have a unique TCR identity[179]. Identical α and β-chain pairings were found when testing two sequential samples from the same subject, but there was significant variation in the percentage found from 0.9% to 7.7% in the two patients sampled. Similarly in a study by the same group looking at tetanus toxoid (TT) specific antigen CD4+ T-cell responses, shared α or β-chain clonotypes, at a nucleotide level, were seen between individuals but these TCRs did not share the accompanying paired α or β-chains[180]. The presence of identical β-chains matched with varying α-chains within samples from the same individual may be predicted through the temporal relationship of α and β-chain recombination. As successful β-chain VDJ recombination results in the productively rearranged gene being expressed and β-chain

protein associated with pTα (a surrogate 33kDa α-chain), this complexes with CD3 and expression of this pre-T-cell receptor triggers a proliferative burst of division cycles before α-chain recombination occurs[254]. This suggests that cells with identical β-chain clonotypes but differing α-chain may have a common precursor but differentiation of α-chain identity occurred following this proliferative burst.

Published literature was reviewed for previous identification of the shared β-chain sequences found within these experiments. The Network of Pancreatic Organ Donors with Diabetes (nPOD) has collated TCR sequencing data from donor tissue from people with T1D, type 2 diabetes (T2D) and healthy controls into its TCR/BCR Clonesearch database[255], which was screened for the presence of shared sequences. Of the shared clonotypes, 39 individual amino acid sequences were found in the nPOD database in 33 donors of which 18 had a history of T1D, but such sequences were also found in donors classified with no diabetes (n=8), T2D (n=3), gastric bypass (n=1), AAb positive (n=1) and "other" (n=2). Sequences with the database have been isolated from intraislet, pancreatic (pLN) lymph node, non-pancreatic (iLN) lymph node, PBMC and spleen. Shared sequences were found across all these compartments with the following total sequence numbers intraislet (n=2), pLN (n=667), iLN (n=184), PBMC (n=26) and spleen (n=286). The two intraislet sequences were from a single nPOD donor with T1D and consisted of a single CD8+ amino acid sequence CASSLGQGNQPQHF from two separate nucleotide sequences. However this sequence was found in a total of 30 donors including those who did not have T1D. This sequence was also found in unpublished deep sequencing of data (courtesy of Dr Iria Gomez-Tourino) again identifying the sequence in T1D donors and healthy controls and was present through various T-cells subsets, namely true naïve, central memory and stem cell memory T-cells. The large number of non-shared sequences limited individual nPOD database searches, however in a small sample there was a high degree of matches (9 out of 15 sequences) to both T1D and non-diabetic donors. The presence of shared clonotypes across treatment groups and in healthy controls or T2D and the presence of a normally distributed CDR3 length would suggest that these clonotypes do not indicate immunodominance through either presence of PIT or disease state.

Amino acid sequences logos from iceLogo show that there is significant preferential amino acid usage in the shared CDR3 clonotypes. There is evidence to show that hydrophobic amino acid usage in positions P6 and P7 of the CDR3 region are associated with self reactivity and

are found in TCRs from NOD mice over control mice[256]. This was not demonstrated when comparing shared and private CDR3 clonotypes in this experiment.

The final TCR data was restricted by a number of factors with limited number of samples and lack of matched samples longitudinally significantly impacting on the power of the statistical analysis. In addition analysis was also limited by the identification of solely β-chain sequences. Without information on α chain usage and αβ pairing we cannot fully identify whether shared clonotypes are truly identical. Also the functional diversity of the sequenced antigen-specific CD4+ cells is unknown and therefore expansion of specific clonotypes cannot be assigned to specific sub-populations, nor can cytokine and chemokine responses be predicted. With sequencing data only the PI C19-A3 specific cells, there is no comparable data available on control CD69+/CD154+ cells and therefore we cannot test if these findings are truly antigen-specific. In short, there were no treatment related changes in TCR repertoire identifiable through these methods.

Previously TCR data has had limited depth in sequencing relatively small peripheral samples of a vast T-cell repertoire but with the growing influence of HTS and efforts by networks such as nPOD there is growing data to elucidate compartmental T-cell populations and importantly comparable healthy control data. Using qPCR, Eugster et al.[180] were able to follow the gene expression phenotype of TT specific cells expressing identical TCR receptor clonotypes, and were able to quantify distinct cytokine profiles with differential expression of IFNγ, IL-4 and FOXP3 genes. Newer technologies such as these, are now expanding on the capabilities of HTS with techniques such as Fluidigm capable of elucidating combined structural and functional identities of T-cells with sequencing of paired αβ TCR chains linked to transcriptional profiles[182]. This enables very detailed evaluation of small T-cell subsets but as throughput capabilities improve and cost fall, such powerful tools will significantly further knowledge in both autoimmune disease and response to therapies such as peptide immunotherapy.

7 Microarray analysis

7.1 Introduction

As outlined in section 1.9, gene expression studies of immunotherapy in autoimmune disease are limited, with no clinical studies in T1D. The MonoPepT1De study included longitudinal CD4+/CD69+/CD154+ antigen-specific cells for gene expression studies through microarray.

7.2 Aims

1. Identify differential gene expression (DGE) patterns present either (i) in response to PI C19-A3 treatment or (ii) longitudinally in newly diagnosed T1D patients

7.3 Results

A total of 96 samples were analysed including from healthy controls (Table 31). Healthy control samples were additional samples to the samples provided by the MonoPepT1De patients. Patient samples were chosen prior to unblinding of the trial and visit 1 and visit 11 samples were chosen to look at longitudinal changes. The healthy controls consisted of 4 samples from a single donor comprising of CD69+/CD154+ cells post HA, SEB and C19-A3 stimulation and CD69-/CD154- resting cells post C19-A3 stimulation. Also, to assess the input of cell number, activated cells following SEB stimulation were collected in populations of 200, 400, 800 and 10,000 cells from 3 further controls. Following SuperAmp RNA amplification and cDNA conversion, amplified cDNA samples were quantified using the an Aligent Bioanalyzer instrument. Six samples were found to have profiles showing low cDNA detection, with 4 out of 6 of these samples having small cell number (≤200 cells) but all samples proceeded to hybridization. Hybridisation was repeated in a number of these samples to rule out technical issues but signal intensities remained low. Bioinformatic analysis of microarray data was performed by Dr Anna Lorenc (King's College London). This confirmed this low signal intensity in these samples despite background correction. These samples were not included in the main analysis to avoid skewing of normalisation. This omission of data had impact on the data sat as a whole as pairing of samples between visit 1 and 11 was not possible as a result.

Table 31: Details of microarray gene expression samples.

*Samples are labelled according to patient number and where applicable visit number. Numbers indicate cell number collected. * indicates samples that with low signal intensity.*

	1025 V1	1025 V11	2010 V1	2010 V11
Treatment	Placebo		High Freq	
C19-A3 (CD69+/CD154+)	2824	*2723	305	743
HA (CD69+/CD154+)	2840	3481	381	857
C19-A3 (CD69-/CD154-)	1000	1000	1000	1000
SEB (CD69+/CD154+)	9372	15207	5116	5221

	2017 V1	2017 V11	2021 V1	2021 V11
Treatment	High Freq		High Freq	
C19-A3 (CD69+/CD154+)	161	2175	658	1397
HA (CD69+/CD154+)	*128	2729	777	1122
C19-A3 (CD69-/CD154-)	1000	1000	1000	1000
SEB (CD69+/CD154+)	922	10000	10000	10000

	3012 V1	3012 V11	3015 V1	3015 V11
	Placebo		High Freq	
C19-A3 (CD69+/CD154+)	422	2734	1696	664
HA (CD69+/CD154+)	704	3047	2380	*1024
C19-A3 (CD69-/CD154-)	1000	1000	1000	1000
SEB (CD69+/CD154+)	10000	9682	10000	10000

	4001 V1	4001 V11	4005 V1	4005 V11
Treatment	Placebo		High Freq	
C19-A3 (CD69+/CD154+)	522	952	*89	756
HA (CD69+/CD154+)			*124	863
C19-A3 (CD69-/CD154-)	1000	1000	1000	1000
SEB (CD69+/CD154+)	10000	10000	10000	10000

	4012 V1	4012 V11	5001 V1	5001 V11
Treatment	Placebo		Low Freq	
C19-A3 (CD69+/CD154+)	803	2849	867	1014
HA (CD69+/CD154+)	980	3188	932	968
C19-A3 (CD69-/CD154-)	1000	1000	1000	1000
SEB (CD69+/CD154+)	10000	10000	9872	10000

	1008 V1	1008 V11	CON G432C
Treatment	Placebo		Control
C19-A3 (CD69+/CD154+)			1782
HA (CD69+/CD154+)			1576
C19-A3 (CD69-/CD154-)			1000
SEB (CD69+/CD154+)	10000	10000	10000

Table 31 continued

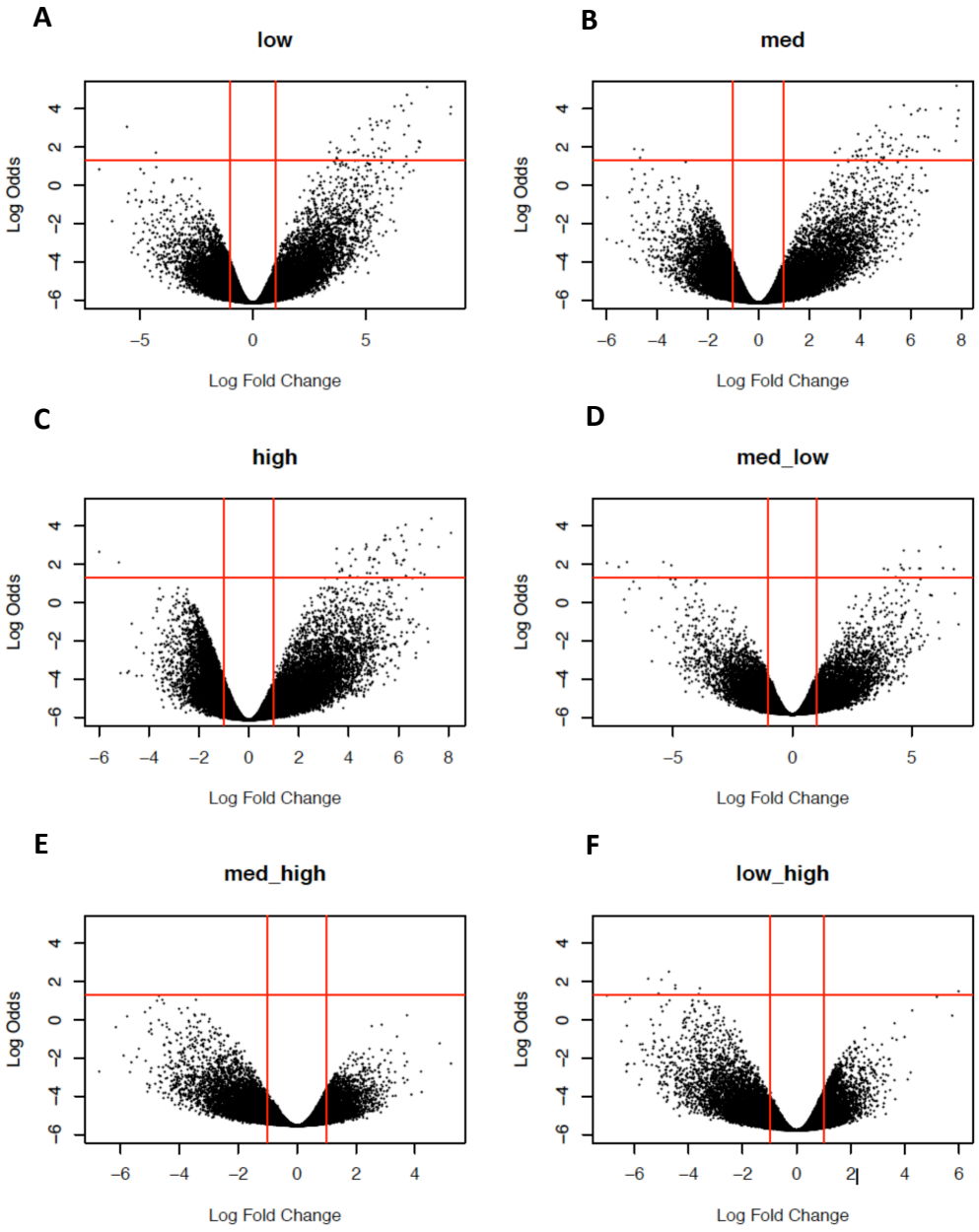
	MP031	MP033	MP104
Treatment	Control	Control	Control
SEB P7 (CD69+/CD154+)	200	200	*200
SEB P7 (CD69+/CD154+)	400	400	400
SEB P7 (CD69+/CD154+)	800	800	800
SEB P7 (CD69+/CD154+)	10000	10000	10000

7.3.1 Comparing the effect of cell number on DGE.

Table 31 illustrates the variability in cell numbers collected in samples. The control SEB samples, which were collected in samples consisting of 200, 400, 800 and 10,000 cells, were incorporated into the microarray to identify the influence of cell number on gene expression. To interpret the effect samples were split into 4 bins labelled as, extremely low (200 cells), low (400 cells), medium (800 cells) and high (10,000 cells). Figure 43 shows the comparison of the gene expression through volcano plots, with genes appearing to be more highly expressed in samples with more cells.

Figure 43: Volcano plots representing gene expression between CD69+/CD154+ SEB stimulated control samples of varying sizes.

A-C represent genes differentially expressed between the extremely low cell number samples and the low (A), medium (B) and high(C) cell number groups. D-F represent genes differentially expressed between medium and low (D), medium and high (E) and low and high (F).



7.3.2 Gene expression changes linked to peptide immunotherapy

Subsequent analysis was performed using 4 methods to model the effect of cell number on DGE pre and post peptide immunotherapy:

A. Ignoring the effect of cell number

B. Modelling the effect of cell number. Using the control SEB stimulated cell populations, a model was fitted where gene expression was dependent on individual and cell number. It was noted that splitting samples into 3 levels of cell number or bins (very low (200 cells), low (400 cells) high (800 or 10,000 cells) in the analyses gave the best fit per gene over other groupings as selected using Bayesian Information Criterion.

C. Adjustment of DGE only in genes that are susceptible to cell number variability. With the assumption that not all gene expression is influenced by cell number, calculations were performed ignoring cell number and expression of genes affected by cell number in the SEB samples were corrected.

Using all treatment samples that had paired visit 1 and V11, but excluding low signal samples, left only 3 pairs of samples. An standard adjusted p-value of <0.05 was considered significant. Method A resulted in no significant adjusted p-values. Method B showed significant differences only in C19-A3 stimulated populations (Table 32). Comparable data using model C are shown in Table 33. Changes in gene expression were also seen in the placebo group, so to identify whether the DGE described above is limited to the treatment group only gene expression was modelled dependent on patients before treatment (peptide or placebo) and after peptide in C19-A3 stimulated cells. This results in non-significant adjusted p-values in all genes, with or without correction for cell number. The 6 genes found have DGE in both methods B and C and were submitted to KEGG[257] Pathway Mapper and the Database for Annotation, Visualization and Integrated Discovery (DAVID)[258] for annotation of biological pathways and processes which these genes are linked to (Table 34). DGE in SEB stimulated cells pre and post treatment versus placebo showed significant adjusted p-values in two genes (Table 35). DGE in C19-A3 CD69-/CD154- resting cells before and after peptide treatment versus placebo showed only one gene with a significant adjusted p-value (Table 36).

Table 32: Model B - Top 10 differentially expressed genes between pre and post treatment C19-A3 stimulated CD69+/CD154+ cells.

Gene name	Log fold change	Adjusted p-value
ATP6V1B1	-1.846	0.000
IL20	-2.185	0.000
ATP5G2	5.780	0.000
JPH2	-2.215	0.000
LOC100130856	-4.629	0.001
DDX3Y	-2.449	0.002
FAM221A	-1.810	0.002
RAB3A	-2.594	0.003
CSTF3	-5.908	0.003
SMG1	5.476	0.003

Table 33: Model C - Top 10 differentially expressed genes between pre and post treatment C19-A3 stimulated CD69+/CD154+ cells

Gene name	Log fold change	Adjusted p-value
JPH2	-2.582	0.000
ATP6V1B1	-1.945	0.000
IL20	-2.337	0.000
ATP5G2	5.926	0.000
BOLA3-AS1	-1.759	0.000
RAB3A	-2.825	0.001
LOC400684	-2.468	0.002
LOC100507670	-1.982	0.002
FAM221A	-1.734	0.002
LOC100130856	-4.309	0.002

Table 34: Biological processes and pathways linked to genes with significant DGE between pre and post treatment C19-A3 stimulated CD69+/CD154+ cells, identified using both methods B and C.

Gene (Gene ID)	Up or down regulation	Gene name	KEGG Pathway	GO Biological Process
ATP6V1B1 *(525)*	Down regulated	ATPase H+ transporting V1 subunit B1	hsa01100 Metabolic pathways hsa00190 Oxidative phosphorylation hsa04150 mTOR signalling pathway hsa04145 Phagosome hsa04966 Collecting duct acid secretion hsa04721 Synaptic vesicle cycle hsa05323 Rheumatoid arthritis hsa05110 Vibrio cholerae infection hsa05120 Epithelial cell signalling in Helicobacter pylori infection	Ossification, regulation of pH, excretion, sensory perception of sound, insulin receptor signalling pathway, ATP hydrolysis coupled proton transport, proton transport, regulation of macroautophagy, transferrin transport, ion transmembrane transport, inner ear morphogenesis, pH reduction, ATP metabolic process, calcium ion homeostasis, phagosome acidification.

Table 34 continued: Biological processes and pathways linked to genes with significant DGE between pre and post treatment C19-A3 stimulated CD69+/CD154+ cells, identified using both methods B and C.

IL20 *(50604)*	Down regulated	interleukin 20	hsa04630 Jak-STAT signalling pathway hsa04060 Cytokine-cytokine receptor interaction	Immune response, positive regulation of tyrosine phosphorylation of Stat3 protein, positive regulation of epidermal cell differentiation, positive regulation of keratinocyte differentiation, positive regulation of osteoclast differentiation, regulation of inflammatory response,
ATP5G2 *(517)*	Up regulated	ATP synthase, H+ transporting, mitochondrial Fo complex subunit C2 (subunit 9)	hsa01100 Metabolic pathways hsa00190 Oxidative phosphorylation hsa05010 Alzheimer's disease hsa05012 Parkinson's disease hsa05016 Huntington's disease	ATP biosynthetic process, ATP synthesis coupled proton transport, ATP hydrolysis coupled proton transport, mitochondrial ATP synthesis coupled proton transport.

Table 34 continued: Biological processes and pathways linked to genes with significant DGE between pre and post treatment C19-A3 stimulated CD69+/CD154+ cells, identified using both methods B and C.

JPH2 *(57158)*	Down regulated	junctophilin 2	Not found	Release of sequestered calcium ion into cytosol, regulation of cardiac muscle tissue development, calcium ion homeostasis, regulation of ryanodine-sensitive calcium-release channel activity, positive regulation of ryanodine-sensitive calcium-release channel activity, calcium ion transport into cytosol,
LOC100130856 *(100130856)*	Down regulated	uncharacterized	Not Found	Not found
FAM221A *(340277)*	Down regulated	family with sequence similarity 221 member A	Not Found	Not found

Table 35: Differentially expressed genes between SEB stimulated CD69+/CD154+ cells in treatment versus placebo groups showing adjusted p-value of <0.05.

Gene name	Log fold change	Adjusted p-value
LOC441268	2.450	0.019
ZNF581	-2.123	0.029

Table 36: Differentially expressed genes between C19A3 resting CD69-/CD154- cells pre and post treatment versus placebo showing adjusted p-value of <0.05.

Gene name	Log fold change	Adjusted p-value
IFI35	4.797	0.010

7.3.3 Gene expression changes linked to clinical responders

Of the peptide treated patients we compared the gene expression changes seen in clinical responders versus non-responders. Within the microarray data there were 3 patients in each of these groups all of which had a high cell number. Figure 44 summarises the presence and relationship of DGE in the subsets of cell populations between clinical responders and non-responders who received peptide across all visits. In the C19-A3 resting cells there were 11 genes with greater expression in non-responders and 5 with greater expression in the responders (Table 37), and in the HA stimulated cells there were 4 genes with greater expression in the responders (Table 38).

Figure 44: Diagram summarising differentially expressed genes found in metabolic responders versus non-responders within different cell populations: C19A3 CD69-/CD69- resting cells, and CD69+/CD154+ cells following C19-A3, HA and SEB stimulation.

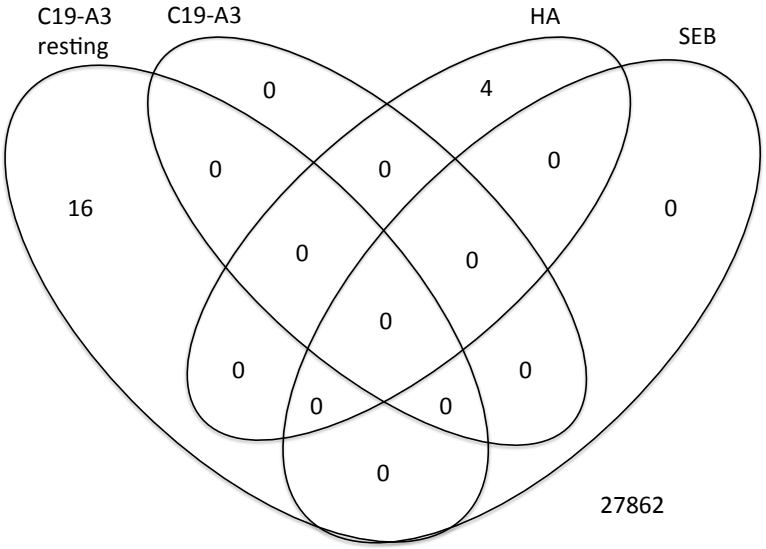


Table 37: Differential gene expression in C19-A3 CD69-/CD154- resting cells between peptide treated clinical responders and non-responders.

Negative log fold change indicates genes with greater expression in non-responders, positive log fold changes indicates greater expression in responders.

Gene name	Log fold change	Adjusted p-value
CYSLTR1	3.238	0.007
COX7B	-5.229	0.007
ITPA	-4.706	0.007
NEK11	2.443	0.007
CHMP2A	5.743	0.032
ZFYVE21	-5.614	0.032
DYNLL1	-5.149	0.032
FANCA	-5.806	0.032
CES4A	-3.631	0.032
EHMT2	-5.050	0.032
ENTPD4	-3.053	0.032
GCC2	2.946	0.032
ATG13	-6.045	0.035
SLC25A37	4.879	0.035
PHIP	-4.684	0.038
C1orf122	-4.134	0.041

Table 38: Differential gene expression in HA stimulated CD69+/CD154+ cells between peptide treated clinical responders and non-responders.

Negative log fold change indicates genes with greater expression in non-responders, positive log fold changes indicates greater expression in responders.

Gene name	Log fold change	Adjusted p-value
GJD3	2.700	0.000
BOD1L1	5.077	0.001
LINC00944	2.017	0.001
LSM11	3.192	0.042

7.3.4 Comparing genes with DGE to established datasets.

Gene of interest identified within these experiments were compared to those genes identified in Burton et al.[203] as genes linked to peptide immunotherapy in a mouse model of EAE. Comparisons of 81 genes identified by these authors were examined within this microarray, however no significant DGE was present when examining C19-A3, HA stimulated cells or C19-A3 resting cells.

ImmuNet (www.immunet.princeton.edu) is a public resource providing an immunological-pathway focused Bayesian integration of over 35,000 genome-scale experiments to help elucidate mechanisms of human immunological disease[259]. With the aim of identifying immune relationships between genes of interest, ImmuNet was used to analyse:

a. DGE pre and post treatment in C19-A3 CD69+/CD154+ cells (Table 30),

b. Genes with greater expression in metabolic responders

c. Genes with greater expression in metabolic non-responders

In the case of genes with DGE in metabolic responders and non-responders the 5 genes with smallest adjusted p-value were chosen to fit with ImmuNet analysis. A number of immune specific functional networks can be analysed within ImmuNet and 3 such networks were chosen for analysis: immune global, antigen processing and presentation and T-cell receptor signalling.

Figure 45: Functional relationships of genes found to have DGE post peptide treatment, ATP6V1B1, IL20, ATP5G2, JPH2 within ImmuNet immune global network

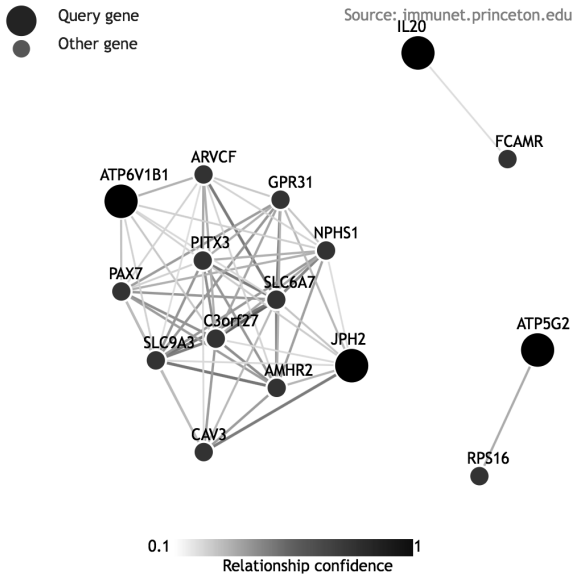


Figure 46: Functional relationships of genes found to have DGE post peptide treatment, ATP6V1B1, IL20, ATP5G2, JPH2 within ImmuNet T-cell signalling pathway network

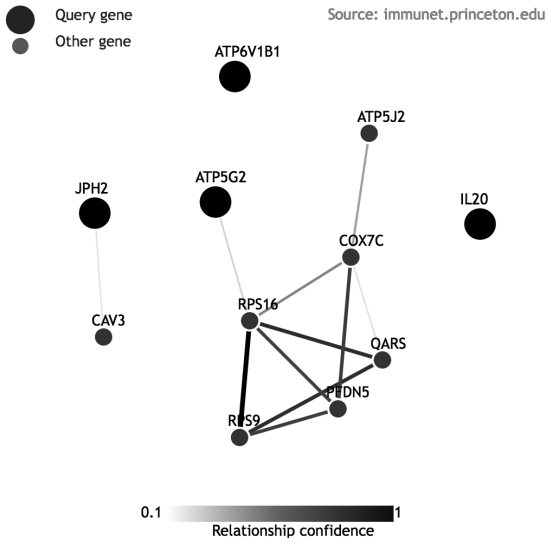


Figure 47: Functional relationships of genes found to have DGE post peptide treatment, ATP6V1B1, IL20, ATP5G2, JPH2 within ImmuNet antigen processing and presentation network

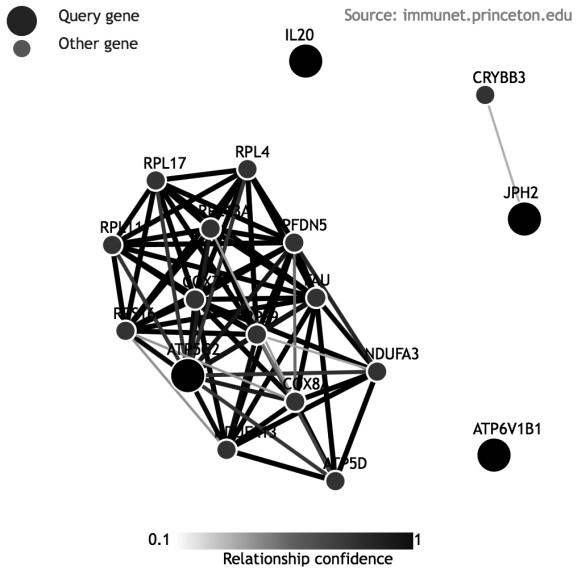


Figure 48: Functional relationships of genes upregulated in responders, CYSLTR1, NEK11, CHMP2A, GCC2 and SLC25A37 within ImmuNet global immune network

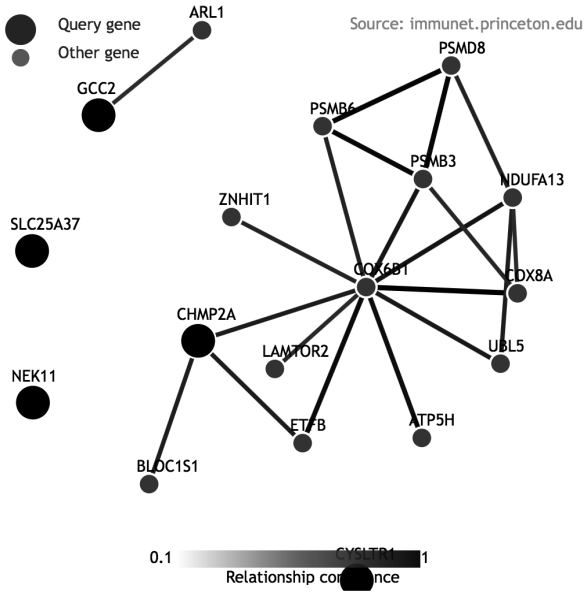


Figure 49: Functional relationships of genes upregulated in responders, CYSLTR1, NEK11, CHMP2A, GCC2 and SLC25A37 within ImmuNet T-cell signalling pathway network

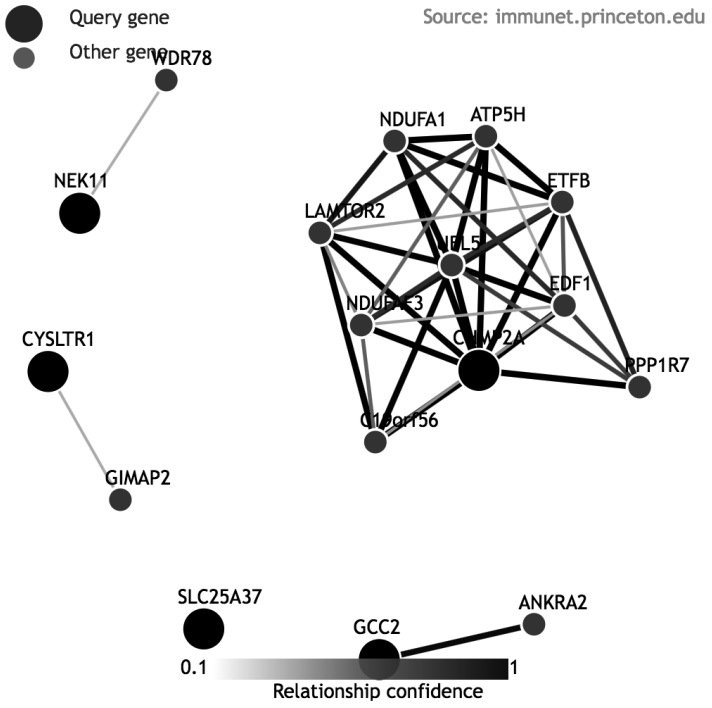


Figure 50: Functional relationships of genes upregulated in responders, CYSLTR1, NEK11, CHMP2A, GCC2 and SLC25A37 within ImmuNet antigen presentation and processing network

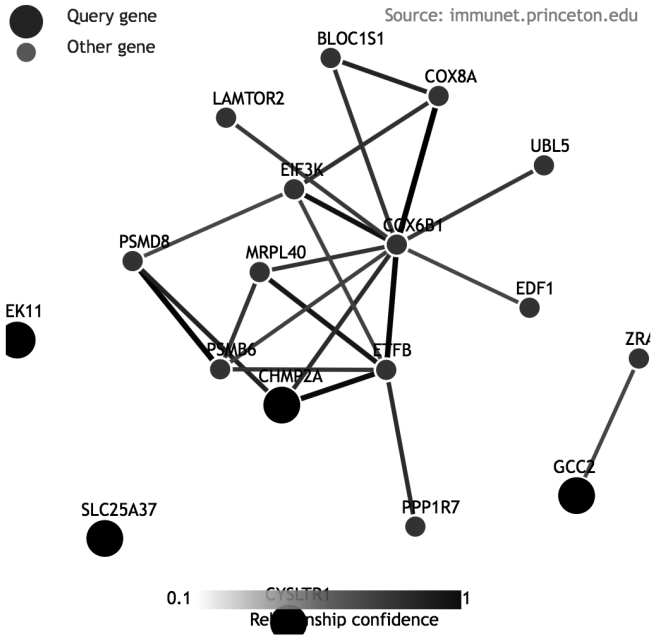


Figure 51: Functional relationships of genes upregulated in non-responders, COX7B, ITPA, ZFYVE21, DYNLL1 and FANCA within ImmuNet global immune network

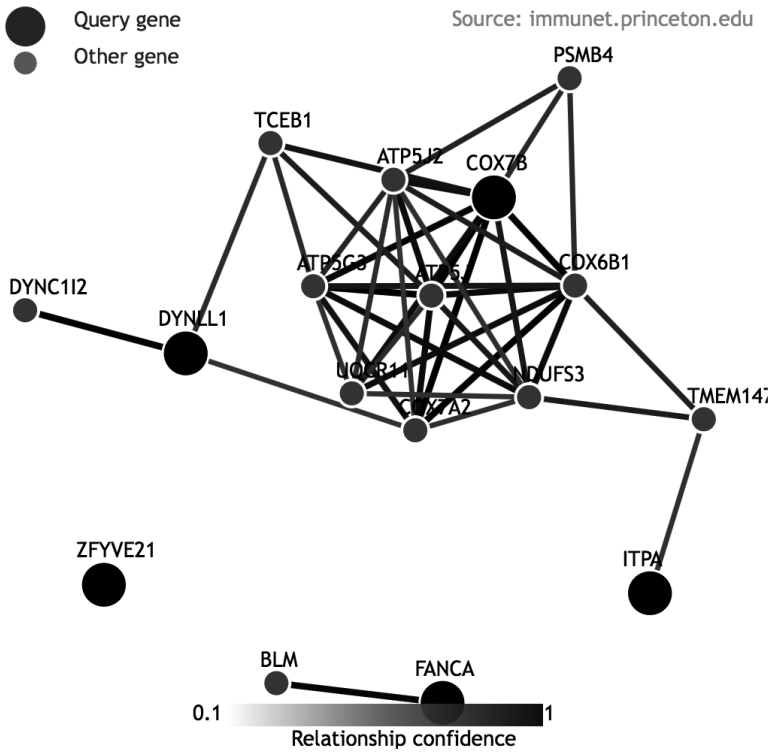


Figure 52: Functional relationships of genes upregulated in non-responders, COX7B, ITPA, ZFYVE21, DYNLL1 and FANCA within ImmuNet T-cell signalling network

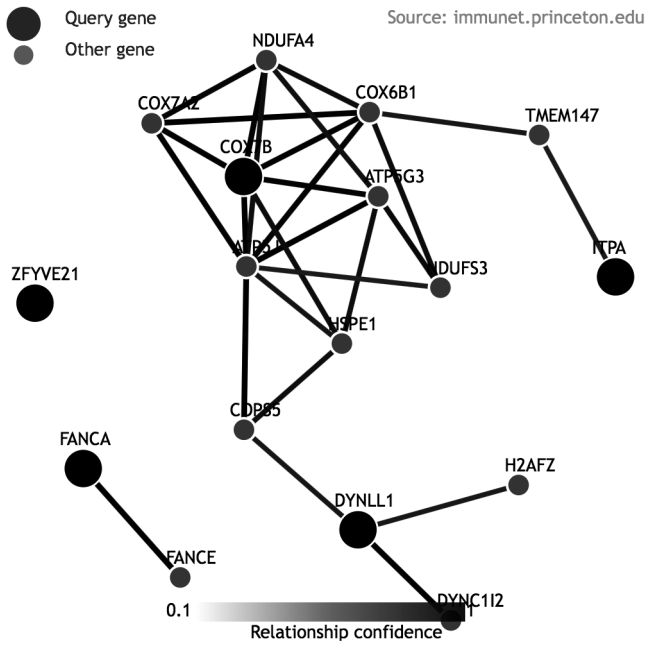
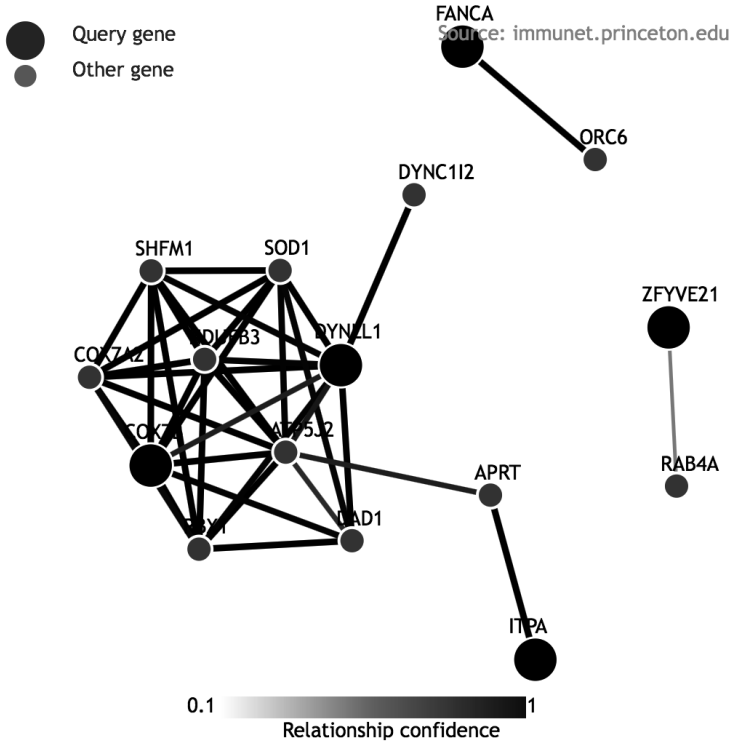


Figure 53: Functional relationships of genes upregulated in non-responders, COX7B, ITPA, ZFYVE21, DYNLL1 and FANCA within ImmuNet antigen presentation and processing network



7.4 Discussion

This microarray experiment was limited by some samples having extremely small number of antigen specific CD69+/CD154+ cells, which not only lead to low signal intensity on the microarray but could also bias interpretation of gene expression. Attempts to quantify this using SEB stimulated cells of varying size cell populations showed that gene expression was reduced in the smaller cell samples, suggested that RNA from rare cells was less likely to be captured in samples of low cell number. This sample size effect was only studied in the SEB stimulated group with the assumption that there would be a similar effect with other stimuli but this was not tested.

DGE in C19-A3 stimulated cells pre and post treatment did not differ significantly from placebo treated individuals. Whilst this suggests that gene expression changes are not linked to peptide treatment it may identify changes occurring through the early period after clinical diagnosis. Six genes were found to have DGE in both methods B and C to adjust for cell number, ATP6V1B1, IL20, ATP5G2, JPH2, LOC100130856 and FAM221A. Of these genes, LOC100130856 and FAM221A have unknown functions.

The biological pathways in which these genes are involved would be in keeping with expected processes early after diagnosis, i.e. changes in metabolic and signalling pathways associated with the resolution of hyperglycaemia. IL20 is expressed on keratinocytes, monocytes, granulocytes, dendritic cells and fibroblasts, and is linked to psoriasis and rheumatoid arthritis.[260] It has previously been identified in GWAS studies as one of 40 loci associated with risk of T1D[15]. It has also been found to be elevated in plasma of autoantibody positive children[261]. Junctophilins are junctional membrane complex proteins known to regulate calcium signalling. Whilst the junctophilin 4 has been recently discovered to be expressed in T-cells[262], junctophilin 2 is felt to be largely expressed in cardiac and skeletal muscle[263]. For the resting C19-A3 CD69-/CD154- cells there was only DGE of the interferon produced protein 35 gene, IFP35, which is involved in type 1 interferon signalling. Although there was DGE in the SEB stimulated CD69+/CD154+ cells this gene differed from the C19-A3 subset in showing a downregulation of ZNF581, a gene involved in transcription.

Differences in HbA1c at baseline highlight the fact that subjects may be identified at different stages of hyperglycaemia ranging from asymptomatic levels picked up on routine screening to admissions for diabetic ketoacidosis. Such variation will have impact on gene expression levels presumably through effects of hyperglycaemia and potentially through differing levels of immune activation. Between visit 1 and visit 11 there was a significant rise in cell numbers in grouped CD69+/CD154+ HA and C19-A3 populations ($p=0.005$) though this rise was non-significant when the populations were considered separately. The proliferative capabilities of PBMCs may be reduced by hyperglycaemia[264], nonetheless within this group there was a non-significant fall in HbA1c between visit 1 and V13.

Responder gene expression analysis showed no DGE in C19-A3 CD69+/C154+ cells but DGE was found in resting C19-A3 CD69-/CD154- cells and HA CD69+/CD154+ cells, indicating that peptide treatment specific changes could not be identified within these limited samples. Of the genes with significant DGE, 3 KEGG pathways were shared between genes, purine metabolism (greater expression of ENTPA and ITPA in non-responders), metabolic pathways (greater expression of COX7B and ITPA in non-responders) and longevity regulating pathways (greater expression of ATG13 and EHMT2 in non-responders).

ImmuNet was chosen to analyse the gene expression data as it offers a validated, simple open-access web interface to explore immune-related functional relationship networks, which offers advantages over non-immune focused networks[259]. Other similar options such as

the Immunological Genome Project (ImmGen)[265] and Immuno-Navigator[266], but these are established using mouse data, whereas ImmuNet has the advantage of using human datasets. The ImmuNet analysis revealed no direct functional relationships between genes identified as having differential gene expression except in between DYNLL1 and COX7B, which were upregulated in non-responders within the antigen presentation and processing pathways. When looking at networks identified in pre and post treatment DGE there appears to be greater relationship confidence within the antigen processing and presentation network linked to the ATP5G2 gene.

8 Summation

This thesis reports the results of a multicentre placebo-controlled peptide immunotherapy trial examining the use of proinsulin C19-A3 in adults newly diagnosed with type 1 diabetes. In addition, the validation and optimisation of immunological assays for the isolation of antigen specific CD4 T-cells and subsequent TCR clonotyping and gene microarray are described.

Modern treatments for T1D are still sub-optimal and there is a clear rationale for peptide immunotherapy as an intervention in this autoimmune disease. With rising numbers of mainly paediatric cases being diagnosed, there is an unmet need for safe, targeted treatments aimed at amelioration of disease over the current focus of management, which is optimisation of insulin replacement. The success of antigen-specific immunotherapy in the field of allergy has led to translation into clinical therapies; and prevention trials in children such as the LEAP study for prevention of peanut allergy in high-risk children, provide encouraging evidence of safety and efficacy in a young age group[267]. ASI also offers potential safety benefits over theoretical worries of the use of immune modulation using monoclonal antibodies in paediatric cohorts. Establishing the safety of PIT in T1D would offer potential for expansion of trials into stage 1 presymptomatic disease (normoglycaemia with 2 or more islet autoantibodies positive)[44] or even earlier in those who have high genetic risk.

With a target population of adults of working and reproductive age, who are generally well and stable on insulin, unique recruitment issues have been highlighted, particularly early in the course of T1D. Importantly, patient perspectives on the study design were sought which have proved helpful for design of further trials. A number of methods to enhance engagement and drive screening have been described, notably the using established networks such as ADDRESS-2 and social media platforms. This study has highlighted that screening and recruiting to a phase 1 study within a tight 100-day window from diagnosis is feasible using a small number of centres within the UK, but also it has quantified a significant proportion missing the screening window due to this tight timeframe. Since the development of this study, the establishment of the T1DUK Consortium has promoted and strengthened the presence of T1D immunotherapy in the UK and there has been a wide expansion of T1D social media communities. One of the aims of this is to aid self-referral and clinical team identification for immunotherapy studies. To further streamline recruitment, protocol design

for immunotherapy trials must balance the need for regular assessment with the practicalities of work, study or travel for people with diabetes.

Comparisons of the MonoPepT1De baseline data with published data were interpreted in the context of the patient populations studied, as there were no data on cohorts that were fully matched with our inclusion criteria. Our data showed a higher percentage of GAD, IA2 and ZnT8 antibody negative subjects with more frequent GAD responses, which could be explained by differences in cohort age, assay methods and possibly geographical genotype distribution. In the screening population, ZnT8 assays were able to distinguish an extra 7% of antibody positive individuals over GAD alone but IA2 did not confer any additional discrimination of antibody positive subjects. ZnT8 titre was also significantly negatively correlated to C-peptide AUC (p=0.01) in the recruited *HLA-DRB1*0401* positive which has not been published in the literature previously.

The 100-day time window from diagnosis aimed to recruit subjects with significant residual C-peptide, defined as a peak stimulated C-peptide of >0.2pmol/mL, and 96.4% of our antibody positive *HLA-DRB1*0401* individuals exceeded this threshold. This suggests that within our inclusion criteria, 3.6% are rapid progressors in which C-peptide reserve declines precipitously within this narrow window. Evidence suggests children progress even more rapidly in terms of C-peptide loss than this adult cohort with 9% of those under 8 falling below the peak stimulated C-peptide of >0.2pmol/mL at diagnosis[230]. This should strengthen the argument for early identification and prevention strategies, as there is a significant proportion of rapid progressors who may be at greater risk of complications due to early decline[58].

Whilst stimulated C-peptide AUC is the most commonly used parameter to assess β-cell function, other simpler, less invasive biomarkers are being sought. In our data we saw fasting C-peptide correlating well with C-peptide AUC figures but further validation is necessary as fasting C-peptide levels may not fully reflect dynamic changes in C-peptide response. C-peptide AUC was not normally distributed and was positively skewed possibly due to the BMI distribution. Over 25% of our patient group was overweight or obese, with BMI being correlated with C-peptide levels. This prevalence of high BMI was similar to other studies[230] and whilst those with higher BMI may reach C-peptide thresholds more readily, the influence of insulin resistance on glycaemic control may have longer term prognostic importance. Urine C-peptide creatinine ratios, which may offer a convenient method to monitor C-peptide levels

in established T1D did not correlate well with stimulated C-peptide levels in the early post-diagnosis period due to high variabiity[229].

The peptide proinsulin C19-A3 was chosen due to its immunological relevance to T1D and previous data in established T1D subjects. Being part of a panel of peptides from proinsulin naturally processed and presented by the high risk HLA DR4 molecule, proinsulin C19-A3 elicited a strong polarisation of ELISPOT responses for inflammatory IFNγ in T1D subjects and IL-10 in healthy controls[75]. A subsequent clinical trial established safety in people with long-standing T1D with the suggestion of favourable immunological responses in certain individuals who showed increased in IL-10 or decreases in IFNγ stimulation indexes[117]. The MonoPepT1De trial strengthens the safety profile of this approach, as it has established that there were no significant risks from high or low frequency treatment in terms of hypersensitivity or acceleration of disease. This is in keeping with the safety profile of PIT in allergy. Decline in C-peptide was significantly lower in the high frequency PIT group at 3 months than placebo and other metabolic measures were reflective of this difference. These suggested that those receiving peptide treatment showed no differences in HbA1c over time or compared to baseline, but increases in insulin doses from baseline were significantly higher in the placebo group at 6, 9 and 12 months despite a non-significant trend for increasing HbA1c. IDAA1c, a composite measure of residual β-cell function was higher at baseline in the placebo and continued to be significantly higher throughout the trial. Immunological assays were able to identify higher IL-10 levels in peptide-treated responders across the treatment period, with a greater increase in Treg expression of FoxP3 from baseline to 6 months in peptide treated-responders than non-responders. There were also higher cumulative CD4 T-cell IL-10 responses in the high frequency group over the 6 month treatment period. These changes are in keeping with postulated mechanisms of immune regulation through peptide immunotherapy, however as with metabolic data, baseline immune differences in response may be present and we were unable to show clear changes in IL-10 or IFNγ from baseline in treatment groups. Our experiments were able to identify changes in Treg subsets with significant changes in memory Tregs expressing CD39+ and negative for Helios within peptide treated responders. These findings indicate such Tregs could be linked to clinical response through peripheral generation of adaptive Tregs and IL-10 mediated immunoregulation. The presence of lower CD57+ CD8 T-cells in the peptide treated responder group is a novel finding suggesting that presence of activation in CD8 T-cells could be a predictor of outcome in PIT. Therefore the MonoPepT1De study has opened up potential

for further exploration of mechanisms of action through immune regulation as well as possible predictors of responder status.

Whilst the MonoPepT1De trial offers satisfactory primary assessment of safety of PI C19-A3 PIT in new-onset T1D, the limited number of subjects in this trial and despite randomisation differences in baseline data limit interpretation of secondary outcomes data and further larger studies would be needed to fully explore efficacy and immunological outcomes. In addition, as the drug as PI C19-A3 been examined only using a limited range of doses, further study of optimal dosing for immunological and metabolic changes would be desirable.

When coordinating a multicentre clinical trial, mechanistic assays pose a number of logistical problems. I have described how planning of the immunological assays was constructed to limit variance in the results and how this was tracked. This procedural knowledge has helped in the development of a further trial using multiple peptides.

Preclinical mouse model experiments have shown evidence that multiple proinsulin peptide immunotherapy may offer advantages over single peptide[239]. Using a transgenic mouse expressing human DR4 as a translational model of T1D, Gibson et al.[239] were able to break tolerance to proinsulin by sensitising with complete Freund's adjuvant, in addition to two doses of pertussis toxin at day 0 and 1. Subsequently the HLA-DR4 transgenic mice show autoinflammatory responses, including proinsulin-specific T-cell proliferation, IFNγ and autoantibody production. This model could then be used to explore PIT regimens with treatment pre- and post-induction reducing autoinflammatory responses. Levels of proliferating FoxP3 Tregs isolated from lymph nodes were significantly higher following multiple peptide imunotherapy than control peptide which was not seen with single peptide PIT. Similarly multiple peptide immunotherapy was effective at down modulating the response to proinsulin protein *in vivo*, by reducing serum levels of proinsulin-specific IgG to a greater degree than single peptides. Whilst preclinical animal studies can be helpful, experience in immunotherapy in T1D shows us that translating results to human studies can often yield differing outcomes, suggesting fundamental differences between these models and human pathophysiology. A phase I clinical trial of PIT using multiple islet autoantigens, the MultiPepT1De trial (ClinicalTrials.gov identifier: NCT02620332) is currently in progress in people who have a duration of T1D of up to 4 years.

Within the study, steps were taken to optimise assays, which for TCR clonotyping involved validation of multiplexing techniques for high-throughput sequencing. Whilst TCR clonotyping was limited to solely β-chain sequencing the findings support other literature where there is common sharing of β-chain clonotypes but this is not limited to T1D or autoimmunity. Evidence suggests that sharing of β-chain clonotypes does not reflect identical αβ pairing[179]; therefore disease or treatment-specific interpretation of the significance of such public β-chain clonotypes is limited without wider data. Improving technologies have expanded the capabilities of TCR sequencing with subsequent recent studies in our laboratory having the capacity for single-cell paired αβ chain sequencing incorporating cell phenotyping through flow cytometry and transcriptional analysis. The microarray data generated in this thesis showed no difference in gene expression in C19-A3 stimulated cells post PIT within treated patients as a whole or in responders. Whilst peptide specific changes were not present, changes in gene expression were present in peptide and placebo treated groups and may indicate changes present early after clinical diagnosis. Also within responders there was DGE in resting C19-A3 CD69-/CD154- and HA CD69+/CD154+ cell subsets, but interpretation of such findings can be limited in the presence of small cell numbers and further validation and evaluation would be necessary.

This thesis has built on the current knowledge of peptide immunotherapy in type 1 diabetes and established the safety of this approach at clinical diagnosis over a 6 month treatment period, with very favourable safety outcomes in both hypersensitivity and importantly absence of disease exacerbation. The approach offers clear advantages to other therapies and its safety indicates it is a valid option for children, as an intervention at diagnosis or potentially in disease prevention. With autoreactive CD8 T-cell CD57 expression present at lower levels in responders, this would suggest that PIT would be ideally suited as an agent in primary prevention. Optimisation of PIT may require multiple peptides as supported by a preclinical T1D model[239] and a further clinical study incorporating use of C19-A3 peptide into a multiple PIT is currently ongoing. Combination treatments with non-antigen specific treatments may offer synergistic effects with ASI. The HLA restriction of C19-A3 PIT currently limits the target population for treatment. With evidence that lower risk HLA groups are increasingly found in emerging T1D cases[21,268], expansion of similar approaches to other HLA groups would require further proof of efficacy and HLA-specific antigen discovery. Significant heterogeneity exists in T1D linked to variations in genetic and environmental factors between individuals, and this has hindered biomarker discovery. In view of this heterogeneity, the

future for immunotherapy suggests a need for a tailored personalised medicine approach and a drive towards preventative strategies. Whilst the difficulties of recruiting adults at diagnosis of T1D have been discussed earlier, recruitment into prevention trials in high-risk but otherwise healthy children poses its own unique challenges. Recent nomenclature clarifying presymptomatic phases of T1D[44] highlights the continuum of disease, identifiable many years prior to symptoms. Such classification aims to highlight to the clinical community a broader window of opportunity for interventions with immunomodulatory agents. The safety profile of PIT would promote its early initiation and further studies would be needed to assess use of the approach in children.

References

1. Eisenbarth, G. S. Type 1 diabetes mellitus. A chronic autoimmune disease. *N. Engl. J. Med.* **314,** 1360–1368 (1986).
2. Willcox, A., Richardson, S. J., Bone, A. J., Foulis, A. K. & Morgan, N. G. Analysis of islet inflammation in human type 1 diabetes. *Clin. Exp. Immunol.* **155,** 173–181 (2009).
3. Ortham, E. L. A. N. Psychosocial Well-Being and Functional Outcomes in Youth With Type 1 Diabetes 12 years After Disease Onset. **33,** (2010).
4. Department of Health Diabetes Policy Team. Making Every Young Person with Diabetes Matter: Report of the Children and Young People with Diabetes Working Group. (2007).
5. Dabelea, D. *et al.* Prevalence of type 1 and type 2 diabetes among children and adolescents from 2001 to 2009. *JAMA* **311,** 1778–86 (2014).
6. Lipman, T. H. *et al.* Increasing Incidence of Type 1 Diabetes in Youth: Twenty years of the Philadelphia Pediatric Diabetes Registry. *Diabetes Care* 1–7 (2013).
7. Ziegler, A.-G. *et al.* Accelerated progression from islet autoimmunity to diabetes is causing the escalating incidence of type 1 diabetes in young children. *J. Autoimmun.* **2,** 6–10 (2011).
8. Derraik, J. G. B. *et al.* Increasing Incidence and Age at Diagnosis among Children with Type 1 Diabetes Mellitus over a 20-Year Period in Auckland (New Zealand). *PLoS One* **7,** e32640 (2012).
9. Onkamo, P., Väänänen, S., Karvonen, M. & Tuomilehto, J. Worldwide increase in incidence of Type I diabetes--the analysis of the data on published incidence trends. *Diabetologia* **42,** 1395–403 (1999).
10. Patterson, C. C., Dahlquist, G. G., Gyürüs, E., Green, A. & Soltész, G. Incidence trends for childhood type 1 diabetes in Europe during 1989-2003 and predicted new cases 2005-20: a multicentre prospective registration study. *Lancet* **373,** 2027–33 (2009).
11. Thomas, N. J. *et al.* Frequency and phenotype of type 1 diabetes in the first six decades of life: a cross-sectional, genetically stratified survival analysis from UK Biobank. *Lancet Diabetes Endocrinol.* **8587,** 1–8 (2017).
12. Miller, R. G., Secrest, A. M., Sharma, R. K., Songer, T. J. & Orchard, T. J. Improvements in the life expectancy of type 1 diabetes: The Pittsburgh Epidemiology of Diabetes Complications Study cohort. *Diabetes* **61,** 1–5 (2012).
13. Livingstone, S. J. *et al.* Estimated Life Expectancy in a Scottish Cohort With Type 1 Diabetes, 2008-2010. *JAMA Intern. Med.* **313,** 37–44 (2015).
14. Hex, N., Bartlett, C., Wright, D., Taylor, M. & Varley, D. Estimating the current and future costs of Type 1 and Type 2 diabetes in the UK, including direct health costs and indirect societal and productivity costs. *Diabet. Med.* **29,** 855–862 (2012).
15. Barrett, J. C. *et al.* Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. *Nat. Genet.* **41,** 703–7 (2009).
16. Todd, J. A. Etiology of type 1 diabetes. *Immunity* **32,** 457–67 (2010).
17. Redondo, M. J. *et al.* Heterogeneity of Type I diabetes: Analysis of monozygotic twins in Great Britain and the United States. *Diabetologia* **44,** 354–362 (2001).
18. Hermann, R. *et al.* Temporal changes in the frequencies of HLA genotypes in patients with Type 1 diabetes--indication of an increased environmental pressure? *Diabetologia* **46,** 420–5 (2003).
19. Vehik, K. *et al.* Trends in high-risk HLA susceptibility genes among colorado youth with type 1 diabetes. *Diabetes Care* **31,** 1392–1396 (2008).
20. Gillespie, K. M. *et al.* The rising incidence of childhood type 1 diabetes and reduced contribution of high-risk HLA haplotypes. *Lancet* **364,** 1699–1700 (2004).

21. Fourlanos, S. *et al.* The rising incidence of type 1 diabetes is accounted for by cases with lower-risk human leukocyte antigen genotypes. *Diabetes Care* **31,** 1546–9 (2008).
22. Kimpimäki, T. *et al.* The first signs of β-cell autoimmunity appear in infancy in genetically susceptible children from the general population: The finnish type 1 diabetes prediction and prevention study. *J. Clin. Endocrinol. Metab.* **86,** 4782–4788 (2001).
23. Cardwell, C. R. *et al.* Birthweight and the risk of childhood-onset type 1 diabetes: A meta-analysis of observational studies using individual patient data. *Diabetologia* **53,** 641–651 (2010).
24. Wilkin, T. J. The accelerator hypothesis: a review of the evidence for insulin resistance as the basis for type I as well as type II diabetes. *Int. J. Obes. (Lond).* **33,** 716–26 (2009).
25. Cardwell, C. R. *et al.* Caesarean section is associated with an increased risk of childhood-onset type 1 diabetes mellitus: A meta-analysis of observational studies. *Diabetologia* **51,** 726–735 (2008).
26. Jakobsson, H. E. *et al.* Decreased gut microbiota diversity, delayed Bacteroidetes colonisation and reduced Th1 responses in infants delivered by caesarean section. *Gut* **63,** 559–66 (2014).
27. Strachan, D. P. Hay fever, hygiene, and household size. *BMJ* **299,** 1259–1260 (1989).
28. Endesfelder, D. *et al.* Compromised gut microbiota networks in children with anti-islet cell autoimmunity. *Diabetes* (2014).
29. Patterson, C. *et al.* Seasonal variation in month of diagnosis in children with type 1 diabetes registered in 23 European centers during 1989-2008: little short-term influence of sunshine hours or average temperature. *Pediatr. Diabetes* (2014).
30. Haynes, A., Bulsara, M. K., Bower, C., Jones, T. W. & Davis, E. A. Cyclical Variation in the Incidence of Childhood Type 1 Diabetes in Western Australia (1985-2010). *Diabetes Care* 1–3 (2012).
31. Knip, M. *et al.* Environmental triggers and determinants of type 1 diabetes. *Diabetes* **54,** S125–S136 (2005).
32. Graves, P. M. *et al.* Prospective study of enteroviral infections and development of beta-cell autoimmunity: Diabetes autoimmunity study in the young (DAISY). *Diabetes Res. Clin. Pract.* **59,** 51–61 (2003).
33. Füchtenbusch, M., Irnstetter, A., Jäger, G. & Ziegler, A. G. No evidence for an association of coxsackie virus infections during pregnancy and early childhood with development of islet autoantibodies in offspring of mothers or fathers with type 1 diabetes. *J. Autoimmun.* **17,** 333–340 (2001).
34. Krogvold, L. *et al.* Detection of a low-grade enteroviral infection in the islets of Langerhans of living patients newly diagnosed with type 1 diabetes. *Diabetes* **64,** 1682–1687 (2015).
35. Harjutsalo, V., Sund, R., Knip, M. & Groop, P.-H. Incidence of type 1 diabetes in Finland. *JAMA* **310,** 427–8 (2013).
36. Zipitis, C. S. & Akobeng, A. K. Vitamin D supplementation in early childhood and risk of type 1 diabetes: a systematic review and meta-analysis. *Arch. Dis. Child.* **93,** 512–517 (2008).
37. Ataie-Jafari, A. *et al.* A randomized placebo-controlled trial of alphacalcidol on the preservation of beta cell function in children with recent onset type 1 diabetes. *Clin. Nutr.* **32,** 911–7 (2013).
38. Walter, M. *et al.* No effect of the 1-alpha, 25-dihydroxyvitamin D3 on beta-cell residual function and insulin requirement in adults with new-onset type 1 diabetes. *Diabetes Care* **33,** 1443–1448 (2010).

39. Lund-Blix, N. A., Stene, L. C., Rasmussen, T., Torjesen, P. A. & Andersen, L. F. Infant feeding in relation to islet autoimmunity and type 1 diabetes in genetically susceptible children : The MIDIA Study. *Diabetes Care* 1–7 (2014). doi:10.2337/dc14-1130
40. Vaarala, O. *et al.* Removal of bovine insulin from cow's milk formula and early initiation of beta-cell autoimmunity in the FINDIA pilot study. *Arch. Pediatr. Adolesc. Med.* **166,** 608 (2012).
41. Knip, M. *et al.* Hydrolyzed infant formula and early β-cell autoimmunity: A randomized clinical trial. *JAMA* **311,** 2279–87 (2014).
42. Åkerblom, H. K. *et al.* The trial to reduce IDDM in the genetically at risk (TRIGR) study: Recruitment, intervention and follow-up. *Diabetologia* **54,** 627–633 (2011).
43. Hagopian, W. a *et al.* The Environmental Determinants of Diabetes in the Young (TEDDY): genetic criteria and international diabetes risk screening of 421 000 infants. *Pediatr. Diabetes* 1–11 (2011).
44. Insel, R. A. *et al.* Staging Presymptomatic Type 1 Diabetes: A Scientific Statement of JDRF, the Endocrine Society, and the American Diabetes Association. *Diabetes Care* **38,** 1964–1974 (2015).
45. Ziegler, A. G. *et al.* Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA* **309,** 2473–9 (2013).
46. Coppieters, K. T., Harrison, L. C. & von Herrath, M. G. Trials in type 1 diabetes: Antigen-specific therapies. *Clin. Immunol.* **149,** 345–55 (2013).
47. The DCCT Research Group. Effect of intensive therapy on residual beta-cell function in patients with type 1 diabetes in the diabetes control and complications trial. A randomized,. *Ann Intern Med* **128,** 517–523 (1998).
48. Wang, L., Lovejoy, N. F. & Faustman, D. L. Persistence of Prolonged C-peptide Production in Type 1 Diabetes as Measured With an Ultrasensitive C-peptide Assay. *Diabetes Care* **35,** 465–70 (2012).
49. Oram, R. A. *et al.* The majority of patients with long-duration type 1 diabetes are insulin microsecretors and have functioning beta cells. *Diabetologia* **57,** 187–191 (2014).
50. Thunander, M. *et al.* Levels of C-peptide, body mass index and age, and their usefulness in classification of diabetes in relation to autoimmunity, in adults with newly diagnosed diabetes in Kronoberg, Sweden. *Eur. J. Endocrinol.* **166,** 1021–9 (2012).
51. Pozzilli, P. *et al.* Metabolic and immune parameters at clinical onset of insulin-dependent diabetes: a population-based study. IMDIAB Study Group. Immunotherapy Diabetes. *Metabolism* **47,** 1205–10 (1998).
52. The DCCT Research Group. Effects of age, duration and treatment of insulin-dependent diabetes mellitus on residual beta-cell function: observations during eligibility testing for the Diabetes Control and Complications Trial (DCCT). The DCCT Research Group. *J. Clin. Endocrinol. Metab.* **65,** 30–36 (1987).
53. Panero, F., Novelli, G., Zucco, C. & Fornengo, P. Fasting plasma C-peptide and micro- and macrovascular complications in a large clinic-based cohort of type 1 diabetic patients. *Diabetes Care* **32,** (2009).
54. Sosenko, J. M. *et al.* Glucose excursions between states of glycemia with progression to type 1 diabetes in the Diabetes Prevention Trial-Type 1 (DPT-1). *Diabetes* **59,** 2386–2389 (2010).
55. Sosenko, J. M. *et al.* Glucose and C-Peptide Changes in the Perionset Period of Type 1 Diabetes in the Diabetes Prevention Trial – Type 1. *Diabetes Care* **31,** 2188–2192 (2008).
56. Sosenko, J. M. *et al.* Patterns of metabolic progression to type 1 diabetes in the

diabetes prevention trial-type 1. *Diabetes Care* **29,** 643–649 (2006).
57. Steffes, M. W., Sibley, S., Jackson, M. & Thomas, W. Beta-cell function and the development of diabetes-related complications in the Diabetes Control and Complications Trial. *Diabetes Care* **26,** 832–6 (2003).
58. Lachin, J. M., McGee, P. & Palmer, J. P. Impact of C-peptide preservation on metabolic and clinical outcomes in the Diabetes Control and Complications Trial. *Diabetes* **63,** 739–748 (2014).
59. Vehik, K. *et al.* Development of autoantibodies in the TrialNet Natural History Study. *Diabetes Care* **34,** 1897–901 (2011).
60. Sosenko, J. M. *et al.* A longitudinal study of GAD65 and ICA512 autoantibodies during the progression to type 1 diabetes in Diabetes Prevention Trial-Type 1 (DPT-1) participants. *Diabetes Care* **34,** 2435–7 (2011).
61. Bosi, E. *et al.* Impact of age and antibody type on progression from single to multiple autoantibodies in type 1 diabetes relatives. *J. Clin. Endocrinol. Metab.* **102,** 2881–2886 (2017).
62. Steck, A. K. *et al.* Effects of non-HLA gene polymorphisms on development of islet autoimmunity and type 1 diabetes in a population with high-risk HLA-DR,DQ genotypes. *Diabetes* **61,** 753–758 (2012).
63. Newman, T. N. *et al.* Members of the novel UBASH3/STS/TULA family of cellular regulators suppress T-cell-driven inflammatory responses in vivo. *Immunol. Cell Biol.* **92,** 837–850 (2014).
64. Jin, Y. *et al.* Risk of type 1 diabetes progression in islet autoantibody-positive children can be further stratified using expression patterns of multiple genes implicated in peripheral blood lymphocyte activation and function. *Diabetes* **63,** 2506–2515 (2014).
65. Hanifi-Moghaddam, P. *et al.* Altered chemokine levels in individuals at risk of Type 1 diabetes mellitus. *Diabet. Med.* **23,** 156–63 (2006).
66. Pfleger, C. *et al.* Relation of circulating concentrations of chemokine receptor CCR5 ligands to C-peptide, proinsulin and HbA1c and disease progression in type 1 diabetes. *Clin. Immunol.* **128,** 57–65 (2008).
67. Uno, S. *et al.* Expression of chemokines, CXC chemokine ligand 10 (CXCL10) and CXCR3 in the inflamed islets of patients with recent-onset autoimmune type 1 diabetes. *Endocr. J.* **57,** 991–6 (2010).
68. Antonelli, A. *et al.* Serum Th1 (CXCL10) and Th2 (CCL2) chemokine levels in children with newly diagnosed Type 1 diabetes: a longitudinal study. *Diabet. Med.* **25,** 1349–53 (2008).
69. Schatz, D. *et al.* Preservation of C-peptide secretion in subjects at high risk of developing type 1 diabetes mellitus--a new surrogate measure of non-progression? *Pediatr. Diabetes* **5,** 72–9 (2004).
70. Sosenko, J. M. *et al.* Acceleration of the Loss of the First-Phase Insulin Response During the Progression to Type 1 Diabetes in Diabetes Prevention Trial-Type 1 Participants. *Diabetes* **62,** 4179–4183 (2013).
71. Xu, P. *et al.* Prognostic Accuracy of Immunologic and Metabolic Markers for Type 1 Diabetes in a High-Risk Population: Receiver operating characteristic analysis. *Diabetes Care* **35,** 1975–1980 (2012).
72. Greenbaum, C. J. & Harrison, L. C. Guidelines for intervention trials in subjects with newly diagnosed type 1 diabetes. *Diabetes* **52,** 1059–65 (2003).
73. Besser, R. E. J. *et al.* Urine C-peptide creatinine ratio is a noninvasive alternative to the mixed-meal tolerance test in children and adults with type 1 diabetes. *Diabetes Care* **34,** 607–9 (2011).
74. Galgani, M. *et al.* Meta-immunological profiling of children with type 1 diabetes

identifies new biomarkers to monitor disease progression. *Diabetes* **62,** 2481–2491 (2013).
75. Arif, S. *et al.* Autoreactive T cell responses show proinflammatory polarization in diabetes but a regulatory phenotype in health. *J. Clin. Invest.* **113,** 451–463 (2004).
76. Mallone, R. *et al.* CD8+ T-cell responses identify beta-cell autoimmunity in human type 1 diabetes. *Diabetes* **56,** 613–21 (2007).
77. Long, S. A. *et al.* Rapamycin/IL-2 Combination Therapy in Patients with Type 1 Diabetes Augments Tregs yet Transiently Impairs b -Cell Function. *Diabetes* **2,** 1–9 (2012).
78. Skyler, J. S. *et al.* Effects of oral insulin in relatives of patients with type 1 diabetes: The Diabetes Prevention Trial-Type 1. *Diabetes Care* **28,** 1068–1076 (2005).
79. Herold, K., Gitelman, S. & Ehlers, M. Teplizumab (Anti-CD3 mAb) Treatment Preserves C-Peptide Responses in Patients With New-Onset Type 1 Diabetes in a Randomized Controlled Trial Metabolic and. *Diabetes* (2013). doi:10.2337/db13-0345.
80. Demeester, S. *et al.* Preexisting Insulin Autoantibodies Predict Efficacy of Otelixizumab in Preserving Residual β-Cell Function in Recent-Onset Type 1 Diabetes. *Diabetes Care* 1–8 (2015).
81. Gepts, W. & Lecompte, P. The pancreatic islets in diabetes. *Am J Med* **70,** 105–15 (1981).
82. In't Veld, P. Insulitis in human type 1 diabetes: The quest for an elusive lesion. *Islets* **3,** 131–138 (2011).
83. Like, A. A. *et al.* Adoptive transfer of autoimmune diabetes mellitus in biobreeding/Worcester (BB/W) inbread and hybrid rats. *J. Immunol.* **134,** 1583–1587 (1985).
84. Bendelac, A., Carnaud, C., Boitard, C. & Bach, J. F. Syngeneic transfer of autoimmune diabetes from diabetic NOD mice to healthy neonates. Requirement for both L3T4+ and Lyt-2+ T cells. *J. Exp. Med.* **166,** 823–832 (1987).
85. Monti, P. *et al.* Evidence for in vivo primed and expanded autoreactive T cells as a specific feature of patients with type 1 diabetes. *J. Immunol.* **179,** 5785 (2007).
86. Szanya, V., Ermann, J. & Taylor, C. The subpopulation of CD4+ CD25+ splenocytes that delays adoptive transfer of diabetes expresses L-selectin and high levels of CCR7. *J. Immunol.* **169,** 2461–2465 (2002).
87. Kukreja, A., Cost, G. & Marker, J. Multiple immuno-regulatory defects in type-1 diabetes. *J. Clin. Invest.* **109,** 131–140 (2002).
88. Lindley, S., Dayan, C., Bishop, A. & Roep, B. Defective suppressor function in CD4+ CD25+ T-cells from patients with type 1 diabetes. *Diabetes* **54,** (2005).
89. You, S. *et al.* Autoimmune diabetes onset results from qualitative rather than quantitative age-dependent changes in pathogenic T-cells. *Diabetes* **54,** (2005).
90. Feutren, G. *et al.* Cyclosporin increases the rate and length of remissions in insulin-dependent diabetes of recent onset. Results of a multicentre double-blind trial. *Lancet* **19,** 119–24 (1986).
91. Silverstein, J. *et al.* Immunosuppression with azathioprine and prednisone in recent-onset insulin-dependent diabetes mellitus. *N. Engl. J. Med.* **319,** 599–604 (1988).
92. Hauser, S. *et al.* B-Cell Depletion with Rituximab in Relapsing–Remitting Multiple Sclerosis. *N. Engl. J. Med.* **358,** 676–688 (2008).
93. Pescovitz, M. D. *et al.* Rituximab, B-lymphocyte depletion, and preservation of beta-cell function. *N. Engl. J. Med.* **361,** 2143–2152 (2009).
94. Pescovitz, M. D. *et al.* B-Lymphocyte Depletion with Rituximab and Beta-Cell Function: Two-Year Results. *Diabetes Care* **37,** 453–459 (2013).
95. Carson, K. R. *et al.* Progressive multifocal leukoencephalopathy after rituximab

therapy in HIV-negative patients: a report of 57 cases from the Research on Adverse Drug Events and Reports project. *Blood* **113,** 4834–40 (2009).
96. Keymeulen, B. *et al.* Insulin needs after CD3-antibody therapy in new-onset type 1 diabetes. *N. Engl. J. Med.* **352,** 2598–2608 (2005).
97. Herold, K. C. *et al.* A Single Course of Anti-CD3 Monoclonal Antibody Responses and Clinical Parameters for at Least 2 Years after Onset of Type 1 Diabetes. *Diabetes* **54,** 1763–1769 (2005).
98. Aronson, R. *et al.* Low-Dose Otelixizumab Anti-CD3 Monoclonal Antibody DEFEND-1 Study: Results of the Randomized Phase III Study in Recent-Onset Human Type 1 Diabetes. *Diabetes Care* **37,** 2746–54 (2014).
99. Ambery, P. *et al.* Efficacy and safety of low-dose otelixizumab anti-CD3 monoclonal antibody in preserving C-peptide secretion in adolescent type 1 diabetes: DEFEND-2, a randomized, placebo-controlled, double-blind, multi-centre study. *Diabet. Med.* **31,** 399–402 (2014).
100. Sherry, N. *et al.* Teplizumab for treatment of type 1 diabetes (Protégé study): 1-year results from a randomised, placebo-controlled trial. *Lancet* **378,** 487–97 (2011).
101. Grinberg-Bleyer, Y. *et al.* IL-2 reverses established type 1 diabetes in NOD mice by a local effect on pancreatic regulatory T cells. *J. Exp. Med.* **207,** 1871–1878 (2010).
102. Johnson, M. C. *et al.* β-Cell-Specific IL-2 Therapy Increases Islet Foxp3+Treg and Suppresses Type 1 Diabetes in NOD Mice. *Diabetes* **62,** 3775–84 (2013).
103. Rabinovitch, A., Suarez-Pinzon, W. L., James Shapiro, a. M., Rajotte, R. V. & Power, R. Combination therapy with sirolimus and interleukin-2 prevents spontaneous and recurrent autoimmune diabetes in NOD mice. *Diabetes* **51,** 638–645 (2002).
104. Delgoffe, G. M. *et al.* mTOR differentially regulates effector and regulatory T cell lineage commitment. *J. Immunol.* **30,** 832–844 (2010).
105. Peakman, M. & Dayan, C. M. Antigen-specific immunotherapy for autoimmune disease: fighting fire with fire? *Immunology* **104,** 361–366 (2001).
106. Zhong, M., Kerlero de Rosbo, N. & A, B.-N. Multiantigen/multiepitope – directed immune-specific suppression of 'complex autoimmune encephalomyelitis' by a novel protein product of a synthetic gene. *J. Clin. Invest.* **110,** 81–90 (2002).
107. Peakman, M. *et al.* Naturally processed and presented epitopes of the islet cell autoantigen IA-2 eluted from HLA-DR4. *J. Clin. Invest.* **104,** 1449–57 (1999).
108. Miller, S. D., Turley, D. M. & Podojil, J. R. Antigen-specific tolerance strategies for the prevention and treatment of autoimmune disease. *Nat. Rev. Immunol.* **7,** 665–77 (2007).
109. Aharoni, R. Immunomodulation neuroprotection and remyelination - The fundamental therapeutic effects of glatiramer acetate: A critical review. *J. Autoimmun.* **54,** 81–92 (2014).
110. Oldfield, W., Larche, M. & Kay, a. Effect of T-cell peptides derived from Fel d 1 on allergic reactions and cytokine production in patients sensitive to cats: a randomised controlled trial. *Lancet* **360,** 47–53 (2002).
111. Yamanaka, K. *et al.* Induction of IL-10-producing regulatory T cells with TCR diversity by epitope-specific immunotherapy in pollinosis. *J. Allergy Clin. Immunol.* **124,** 842–5.e7 (2009).
112. Campbell, J. D. *et al.* Peptide immunotherapy in allergic asthma generates IL-10-dependent immunological tolerance associated with linked epitope suppression. *J. Exp. Med.* **206,** 1535–47 (2009).
113. Couroux, P., Patel, D., Armstrong, K., Larché, M. & Hafner, R. P. Fel d 1-derived synthetic peptide immuno-regulatory epitopes show a long-term treatment effect in cat allergic subjects. *Clin Exp Allergy* **45,** 974–81 (2015).

114. Peakman, M. & von Herrath, M. Antigen-specific immunotherapy for type 1 diabetes: maximizing the potential. *Diabetes* **59,** 2087–93 (2010).
115. Fousteri, G., Chan, J., Zheng, Y. & Whiting, C. Virtual optimization of nasal insulin therapy predicts immunization frequency to be crucial for diabetes protection. *Diabetes* **59,** (2010).
116. Homann, D., Holz, A., Bot, A., Coon, B. & Wolfe, T. Autoreactive CD4+ T Cells Protect from Autoimmune Diabetes via Bystander Suppression Using the IL-4/Stat6 Pathway. *Immunity* **11,** 463–472 (1999).
117. Thrower, S. L. *et al.* Proinsulin peptide immunotherapy in type 1 diabetes: report of a first-in-man Phase I safety study. *Clin. Exp. Immunol.* **155,** 156–65 (2008).
118. Hjorth, M. *et al.* GAD-alum treatment induces GAD65-specific CD4+CD25highFOXP3+ cells in type 1 diabetic patients. *Clin. Immunol.* **138,** 117–26 (2011).
119. Raz, I. *et al.* Beta-cell function in new-onset type 1 diabetes and immunomodulation with a heat-shock protein peptide (DiaPep277): a randomised, double-blind, phase II trial. *Lancet* **358,** 1749–53 (2001).
120. Burkhart, C., Liu, G. Y., Anderton, S. M., Metzler, B. & Wraith, D. C. Peptide-induced T cell regulation of experimental autoimmune encephalomyelitis: a role for IL-10. *Int. Immunol.* **11,** 1625–1634 (1999).
121. Tree, T. I. M. *et al.* Naturally Arising Human CD4 T-Cells That Recognize Islet Autoantigens and Secrete Interleukin-10 Regulate Proinflammatory T-Cell Responses via Linked Suppression. *Diabetes* **59,** 1451–1460 (2010).
122. Chen, Y. *et al.* Peripheral deletion of antigen-reactive T cells in oral tolerance. *Nature* **376,** 177–180 (1995).
123. Mcpherson, R. C. *et al.* Epigenetic modification of the PD-1 (Pdcd1) promoter in effector CD4+ T cells tolerized by peptide immunotherapy. *Elife* **1,** 1–20 (2014).
124. Keir, M. E., Butte, M. J., Freeman, G. J. & Sharpe, A. H. PD-1 and its ligands in tolerance and immunity. *Annu. Rev. Immunol.* **26,** 677–704 (2008).
125. Konkel, J. E. *et al.* PD-1 signalling in CD4+ T cells restrains their clonal expansion to an immunogenic stimulus, but is not critically required for peptide-induced tolerance. *Immunology* **130,** 92–102 (2010).
126. Deshpande, P., King, I. L. & Segal, B. M. IL-12 driven upregulation of P-selectin ligand on myelin-specific T cells is a critical step in an animal model of autoimmune demyelination. *J. Neuroimmunol.* **173,** 35–44 (2006).
127. Burton, B. R. *et al.* Sequential transcriptional changes dictate safe and effective antigen-specific immunotherapy. *Nat. Commun.* **5,** 1–13 (2014).
128. Haselden, B. M., Kay, A. B. & Larché, M. Immunoglobulin E-independent major histocompatibility complex-restricted T cell peptide epitope-induced late asthmatic reactions. *J. Exp. Med.* **189,** 1885–94 (1999).
129. Oldfield, W. L. G., Kay, A. B. & Larche, M. Allergen-Derived T Cell Peptide-Induced Late Asthmatic Reactions Precede the Induction of Antigen-Specific Hyporesponsiveness in Atopic Allergic Asthmatic Subjects. *J. Immunol.* **167,** 1734–1739 (2001).
130. Cetkovic-Cvrlje, M. *et al.* Retardation or acceleration of diabetes in NOD/Lt mice mediated by intrathymic administration of candidate β-cell antigens. *Diabetes* **46,** 1975–1982 (1997).
131. Tian, J., Olcott, A. P. & Kaufman, D. L. Antigen-Based Immunotherapy Drives the Precocious Development of Autoimmunity. *J. Immunol.* **169,** 6564–6569 (2002).
132. Blanas, E., Carbone, F. & Allison, J. Induction of autoimmune diabetes by oral administration of autoantigen. *Science (80-.).* **274,** 1707–1709 (1996).
133. Bielekova, B. *et al.* Encephalitogenic potential of the myelin basic protein peptide (amino acids 83-99) in multiple sclerosis: results of a phase II clinical trial with an

altered peptide ligand. *Nat. Med.* **6,** 1167–1175 (2000).
134. Liu, E. *et al.* Anti-peptide autoantibodies and fatal anaphylaxis in NOD mice in response to insulin self-peptides B:9-23 and B:13-23. *J. Clin. Invest.* **110,** 1021–1027 (2002).
135. Liu, E. *et al.* Preventing peptide-induced anaphylaxis: Addition of C-terminal amino acids to produce a neutral isoelectric point. *J. Allergy Clin. Immunol.* **114,** 607–613 (2004).
136. Bonifacio, E. *et al.* Effects of High-Dose Oral Insulin on Immune Responses in Children at High Risk for Type 1 Diabetes. *Jama* **313,** 1541 (2015).
137. Näntö-Salonen, K. *et al.* Nasal insulin to prevent type 1 diabetes in children with HLA genotypes and autoantibodies conferring increased risk of disease: a double-blind, randomised controlled trial. *Lancet* **372,** 1746–55 (2008).
138. Roep, B. O. *et al.* Plasmid-Encoded Proinsulin Preserves C-Peptide While Specifically Reducing Proinsulin-Specific CD8+ T Cells in Type 1 Diabetes. *Sci. Transl. Med.* **5,** 191ra82-191ra82 (2013).
139. Ludvigsson, J. *et al.* GAD65 antigen therapy in recently diagnosed type 1 diabetes mellitus. *N. Engl. J. Med.* **366,** 433–42 (2012).
140. Raz, I., Avron, A. & Tamir, M. Treatment of new- onset type 1 diabetes with peptide DiaPep277® is safe and associated with preserved beta-cell function: extension of a randomized, double blind, phase II trial. *Diabetes/metabolism …* **23,** 292–298 (2007).
141. Walter, M., Philotheou, A., Bonnici, F., Ziegler, A. G. & Jimenez, R. No effect of the altered peptide ligand NBI-6024 on β-cell residual function and insulin needs in new-onset type 1 diabetes. *Diabetes Care* **32,** 2036–2040 (2009).
142. Kent, S. C. *et al.* Expanded T cells from pancreatic lymph nodes of type 1 diabetic subjects recognize an insulin epitope. *Nature* **435,** 224–8 (2005).
143. Nakayama, M. *et al.* Prime role for an insulin epitope in the development of type1 diabetes in NOD mice. *Nature* **435,** 220–223 (2005).
144. Congia, M., Patel, S., Cope, A. P., De Virgiliis, S. & Sønderstrup, G. T cell epitopes of insulin defined in HLA-DR4 transgenic mice are derived from preproinsulin and proinsulin. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 3833–8 (1998).
145. Mannering, S. I. *et al.* The A-chain of insulin is a hot-spot for CD4 + T cell epitopes in human type 1 diabetes. *Clin. Exp. Immunol.* **156,** 226–231 (2009).
146. Pathiraja, V. *et al.* Proinsulin-Specific, HLA-DQ8, and HLA-DQ8-Transdimer-Restricted CD4+ T Cells Infiltrate Islets in Type 1 Diabetes. *Diabetes* **64,** 172–82 (2015).
147. Kronenberg, D. *et al.* Circulating, Preproinsulin Signal Peptide-Specific CD8 T Cells Restricted by the Susceptibility Molecule HLA-A24 Are Expanded at Onset of Type 1 Diabetes and Kill β-Cells. *Diabetes* 1–8 (2012).
148. Krischer, J. P., Schatz, D. A., Bundy, B., Skyler, J. S. & Greenbaum, C. J. Effect of Oral Insulin on Prevention of Diabetes in Relatives of Patients With Type 1 Diabetes. *Jama* **318,** 1891 (2017).
149. The Diabetes Prevention Trial - Type 1 Study Group. Effects of insulin in relatives of patients with type 1 diabetes mellitus. *N. Engl. J. Med.* **346,** 1685–91 (2002).
150. Ludvigsson, J. *et al.* GAD treatment and insulin secretion in recent-onset type 1 diabetes. *N. Engl. J. Med.* **359,** 1909–20 (2008).
151. Kaufman, D., Clare-Salzler, M. & Tian, J. Spontaneous loss of T-cell tolerance to glutamic acid decarboxylase in murine insulin-dependent diabetes. *Nature* **366,** 69–72 (1993).
152. Tian, J., Clere-Salzler, M. & Herschenfeld, A. Modulating autoimmune responses to GAD inhibits disease progression and prolongs islet graft survival in diabetes-prone

mice. *Nat. Med.* (1996).
153. Tisch, R., Liblau, R. S., Yang, X. D., Liblau, P. & McDevitt, H. O. Induction of GAD65-specific regulatory T-cells inhibits ongoing autoimmune diabetes in nonobese diabetic mice. *Diabetes* **47,** 894–9 (1998).
154. Ludvigsson, J. *et al.* Extended evaluation of the safety and efficacy of GAD treatment of children and adolescents with recent-onset type 1 diabetes: a randomised controlled trial. *Diabetologia* **54,** 634–40 (2011).
155. Agardh, C. D. *et al.* Clinical evidence for the safety of GAD65 immunomodulation in adult-onset autoimmune diabetes. *J. Diabetes Complications* **19,** 238–246 (2005).
156. Wherrett, D. K. *et al.* Antigen-based therapy with glutamic acid decarboxylase (GAD) vaccine in patients with recent-onset type 1 diabetes: a randomised double-blind trial. *Lancet* **6736,** 1–9 (2011).
157. Brudzynski, K., Martinez, V. & Gupta, R. S. Secretory granule autoantigen in insulin-dependent diabetes mellitus is related to 62 kDa heat-shock protein (hsp60). *J. Autoimmun.* **5,** 453–63 (1992).
158. Elias, D. & Cohen, I. R. Peptide therapy for diabetes in NOD mice. *Lancet* **343,** 704–6 (1994).
159. Lazar, L. *et al.* Heat-shock protein peptide DiaPep277 treatment in children with newly diagnosed type 1 diabetes: a randomised, double-blind phase II study. *Diabetes. Metab. Res. Rev.* **23,** 286–91 (2007).
160. Raz, I. *et al.* Treatment of recent-onset type 1 diabetic patients with DiaPep277: results of a double-blind, placebo-controlled, randomized phase 3 trial. *Diabetes Care* **37,** 1392–400 (2014).
161. Raz, I. *et al.* Statement of retraction: Treatment of Recent-Onset Type 1 Diabetic Patients With DiaPep277: Results of a Double-Blind, Placebo-Controlled, Randomized Phase 3 Trial. Diabetes Care 2014;37:1392–1400. DOI: 10.2337/dc13-1391. *Diabetes Care* **38,** 178 (2015).
162. Nikolich-Zugich, J., Slifka, M. K. & Messaoudi, I. The many important facets of T-cell repertoire diversity. *Nat. Rev. Immunol.* **4,** 123–32 (2004).
163. Shortman, K., Egerton, M., Spangrude, G. J. & Scollay, R. The generation and fate of thymocytes. *Semin. Immunol.* **2,** 3–12 (1990).
164. Ferreira, C. *et al.* Non-obese diabetic mice select a low-diversity repertoire of natural regulatory T cells. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 8320–5 (2009).
165. Quinn, A. *et al.* T cells to a dominant epitope of GAD65 express a public CDR3 motif. *Int. Immunol.* **18,** 967–979 (2006).
166. Menezes, J. S. *et al.* A public T cell clonotype within a heterogeneous autoreactive repertoire is dominant in driving EAE. *J. Clin. Invest.* **117,** 2176–2185 (2007).
167. Madakamutil, L. T., Maricic, I., Sercarz, E. E. & Kumar, V. Immunodominance in the alpha TCR repertoire of TCR peptide-specific CD4+ Treg population that controls experimental autoimmune encephalomyelitis. *J. Immunol.* **180,** 4577–85 (2008).
168. Fazilleau, N. *et al.* T cell repertoire diversity is required for relapses in myelin oligodendrocyte glycoprotein-induced experimental autoimmune encephalomyelitis. *J. Immunol.* **178,** 4865–4875 (2007).
169. Marrero, I., Hamm, D. E. & Davies, J. D. High-throughput sequencing of islet-infiltrating memory CD4+ T cells reveals a similar pattern of TCR Vbeta usage in prediabetic and diabetic NOD mice. *PLoS One* **8,** e76546 (2013).
170. Venturi, V., Price, D. a, Douek, D. C. & Davenport, M. P. The molecular basis for public T-cell responses? *Nat. Rev. Immunol.* **8,** 231–8 (2008).
171. Malhotra, U., Spielman, R. & Concannon, P. Variability in T cell receptor V beta gene usage in human peripheral blood lymphocytes. Studies of identical twins, siblings, and

insulin-dependent diabetes mellitus patients. *J. Immunol.* **149,** 1802–8 (1992).
172. Fozza, C. *et al.* T-cell receptor repertoire analysis in monozygotic twins concordant and discordant for type 1 diabetes. *Immunobiology* 6–11 (2012). doi:10.1016/j.imbio.2012.01.002
173. Codina-Busqueta, E. *et al.* TCR Bias of In Vivo Expanded T Cells in Pancreatic Islets and Spleen at the Onset in Human Type 1 Diabetes. *J. Immunol.* (2011). doi:10.4049/jimmunol.1002423
174. Miles, J. J., Douek, D. C. & Price, D. A. Bias in the αβ T-cell repertoire: implications for disease pathogenesis and vaccination. *Immunol. Cell Biol.* **89,** 375–87 (2011).
175. Robins, H. S. *et al.* Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. *Blood* **114,** 4099–107 (2009).
176. Freeman, J. D., Warren, R. L., Webb, J. R., Nelson, B. H. & Holt, R. a. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. *Genome Res.* **19,** 1817–24 (2009).
177. Robins, H. *et al.* Ultra-sensitive detection of rare T cell clones. *J. Immunol. Methods* **375,** 14–19 (2012).
178. Estorninho, M. *et al.* A Novel Approach to Tracking Antigen-Experienced CD4 T Cells into Functional Compartments via Tandem Deep and Shallow TCR Clonotyping. *J. Immunol.* **191,** 5430–5440 (2013).
179. Eugster, A. *et al.* High Diversity in the TCR Repertoire of GAD65 Autoantigen-Specific Human CD4+ T Cells. *J. Immunol.* **194,** 2531–2538 (2015).
180. Eugster, A. *et al.* Measuring T cell receptor and T cell gene expression diversity in antigen-responsive human CD4+ T cells. *J. Immunol. Methods* **400,** 13–22 (2013).
181. Eltahla, A. A. *et al.* Linking the T cell receptor to the single cell transcriptome in antigen-specific human T cells. *Immunol. Cell Biol.* **94,** 1–8 (2016).
182. Stubbington, M. J. T. *et al.* Simultaneously inferring T cell fate and clonality from single cell transcriptomes. *Nat. Methods* **13,** 329–332 (2016).
183. Redmond, D., Poran, A. & Elemento, O. Single-cell TCRseq: paired recovery of entire T-cell alpha and beta chain transcripts in T-cell receptors from single-cell RNAseq. *Genome Med.* **8,** 80 (2016).
184. Han, A., Glanville, J., Hansmann, L. & Davis, M. M. Linking T-cell receptor sequence to functional phenotype at the single-cell level. *Nat. Biotechnol.* **32,** 684–692 (2014).
185. Planas, R. *et al.* Gene expression profiles for the human pancreas and purified islets in Type 1 diabetes: New findings at clinical onset and in long-standing diabetes. *Clin. Exp. Immunol.* **159,** 23–44 (2010).
186. Hopfgarten, J. *et al.* Gene expression analysis of human islets in a subject at onset of type 1 diabetes. *Acta Diabetol.* 199–204 (2013). doi:10.1007/s00592-013-0479-5
187. Richardson, S. J. *et al.* Islet cell hyperexpression of HLA class I antigens: a defining feature in type 1 diabetes. *Diabetologia* **59,** 2448–2458 (2016).
188. Rui, J. *et al.* Methylation of insulin DNA in response to proinflammatory cytokines during the progression of autoimmune diabetes in NOD mice. *Diabetologia* **59,** 1021–1029 (2016).
189. Hisanaga-Oishi, Y., Nishiwaki-Ueda, Y., Nojima, K. & Ueda, H. Analysis of the expression of candidate genes for type 1 diabetes susceptibility in T cells. *Endocr. J.* **61,** 577–588 (2014).
190. Orban, T. *et al.* Reduced CD4+ T-cell-specific gene expression in human type 1 diabetes mellitus. *J. Autoimmun.* **28,** 177–187 (2007).
191. Reynier, F. *et al.* Specific gene expression signature associated with development of autoimmune type-I diabetes using whole-blood microarray analysis. *Genes Immun.* **11,** 269–78 (2010).

192. Kaizer, E. C. *et al.* Gene expression in peripheral blood mononuclear cells from children with diabetes. *J. Clin. Endocrinol. Metab.* **92,** 3705–3711 (2007).
193. Irvine, K. M. *et al.* Peripheral Blood Monocyte Gene Expression Profile Clinically Stratifies Patients With Recent-Onset Type 1 Diabetes. *Diabetes* **61,** 1281–1290 (2012).
194. Pruul, K. *et al.* Differences in B7 and CD28 family gene expression in the peripheral blood between newly diagnosed young-onset and adult-onset type 1 diabetes patients. *Mol. Cell. Endocrinol.* **412,** 265–271 (2015).
195. Padmos, R. C. *et al.* Distinct monocyte Gene-Expression profiles in autoimmune diabetes. *Diabetes* **57,** 2768–2773 (2008).
196. Beyan, H. *et al.* Monocyte gene-expression profiles associated with childhood-onset type 1 diabetes and disease risk: A study of identical twins. *Diabetes* **59,** 1751–1755 (2010).
197. Jin, Y. *et al.* The expression of inflammatory genes is upregulated in peripheral blood of patients with type 1 diabetes. *Diabetes Care* **36,** 2794–2802 (2013).
198. Ferreira, R. C. *et al.* A type I Interferon transcriptional signature precedes autoimmunity in children genetically at risk for type 1 diabetes. *Diabetes* **63,** 2538–2550 (2014).
199. Takahashi, P. *et al.* MicroRNA expression profiling and functional annotation analysis of their targets in patients with type 1 diabetes mellitus. *Gene* **539,** 213–223 (2014).
200. Luce, S. *et al.* Single Insulin-Specific CD8+ T Cells Show Characteristic Gene Expression Profiles in Human Type 1 Diabetes. *Diabetes* **60,** 3289–3299 (2011).
201. Nielsen, L. B. *et al.* Circulating levels of MicroRNA from children with newly diagnosed type 1 diabetes and healthy controls: Evidence that miR-25 associates to residual beta-cell function and glycaemic control during disease progression. *Exp. Diabetes Res.* **2012,** (2012).
202. Sthoeger, Z. M. *et al.* Treatment of lupus patients with a tolerogenic peptide, hCDR1 (Edratide): Immunomodulation of gene expression. *J. Autoimmun.* **33,** 77–82 (2009).
203. Burton, B. R., Britton, G. J., Fang, H., Verhagen, J., Smithers, B., Sabatos-Peyton, C. A., Carney, L. J., Gough, J., Strobel, S. and Wraith, D. . Sequential transcriptional changes dictate safe and effective antigen-specific immunotherapy of autoimmune disease. *Nat Commun* **5,** 1–13 (2014).
204. Yuan, W., Dimartino, S. J., Redecha, P. B., Ivashkiv, L. B. & Salmon, J. E. Systemic lupus erythematosus monocytes are less responsive to interleukin-10 in the presence of immune complexes. *Arthritis Rheum.* **63,** 212–218 (2011).
205. Chattopadhyay, P. K., Yu, J. & Roederer, M. A live-cell assay to detect antigen-specific CD4+ T cells with diverse cytokine profiles. *Nat. Med.* **11,** 1113–7 (2005).
206. Chattopadhyay, P. K., Yu, J. & Roederer, M. Live-cell assay to detect antigen-specific CD4+ T-cell responses by CD154 expression. *Nat. Protoc.* **1,** 1–6 (2006).
207. Quigley, M., Almeida, J., Price, D. & Douek, D. Unbiased molecular analysis of T cell receptor expression using template-switch anchored RT-PCR. *Curr. Protoc. Immunol.* **94,** 10.33.1-10.33.16 (2011).
208. Alamyar, E., Giudicelli, V., Li, S., Duroux, P. & Lefranc, M.-P. IMGT/HIGHV-QUEST: THE IMGT® WEB PORTAL FOR IMMUNOGLOBULIN (IG) OR ANTIBODY AND T CELL RECEPTOR (TR) ANALYSIS FROM NGS HIGH THROUGHPUT AND DEEP SEQUENCING. *Immunome Res.* **8,** 2–15 (2012).
209. Li, S. *et al.* IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. *Nat. Commun.* **4,** 2333 (2013).
210. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2015).

211. Nazarov, V. I. *et al.* tcR: an R package for T cell receptor repertoire advanced data analysis. *BMC bioinformatics* **16,** 175 (2015).
212. Vaziri-Sani, F. *et al.* ZnT8 autoantibody titers in type 1 diabetes patients decline rapidly after clinical onset. *Autoimmunity* **43,** 598–606 (2010).
213. Heywood, J. *et al.* Effective recruitment of participants to a phase I study using the internet and publicity releases through charities and patient organisations: analysis of the adaptive study of IL-2 dose on regulatory T cells in type 1 diabetes (DILT1D). *Trials* **16,** 1–13 (2015).
214. Klitz, W. *et al.* New HLA haplotype frequency reference standards: high-resolution and large sample typing of HLA DR-DQ haplotypes in a sample of European Americans. *Tissue Antigens* **62,** 296–307 (2003).
215. Gonzalez-Galarza, F. F., Christmas, S., Middleton, D. & Jones, A. R. Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations. *Nucleic Acids Res.* **39,** D913-9 (2011).
216. Windsor, L. *et al.* Does a central MHC gene in linkage disequilibrium with HLA-DRB1*0401 affect susceptibility to type 1 diabetes? *Genes Immun.* **6,** 298–304 (2005).
217. Lipner, E., Tomer, Y. & Noble, J. HLA Class I and II Alleles are Associated with Microvascular Complications of Type 1 Diabetes. *Hum. Immunol.* **74,** 538–544 (2013).
218. Bingley, P. J. Clinical applications of diabetes antibody testing. *J. Clin. Endocrinol. Metab.* **95,** 25–33 (2010).
219. Bingley, P. J. *et al.* Prediction of IDDM in the general population: Strategies based on combinations of autoantibody markers. *Diabetes* **46,** 1701–1710 (1997).
220. Törn, C. *et al.* Combinations of beta cell specific autoantibodies at diagnosis of diabetes in young adults reflects different courses of beta cell damage. *Autoimmunity* **33,** 115–20 (2001).
221. Wenzlau, J. M. *et al.* Kinetics of the post-onset decline in zinc transporter 8 autoantibodies in type 1 diabetic human subjects. *J. Clin. Endocrinol. Metab.* **95,** 4712–4719 (2010).
222. Andersen, M. L. M. *et al.* Association between autoantibodies to the Arginine variant of the Zinc transporter 8 (ZnT8) and stimulated C-peptide levels in Danish children and adolescents with newly diagnosed type 1 diabetes. *Pediatr. Diabetes* **13,** 454–462 (2012).
223. Besser, R. E. J., Shields, B. M., Casas, R., Hattersley, A. T. & Ludvigsson, J. Lessons From the Mixed Meal Tolerance Test: Use of 90-minute and fasting C-peptide in pediatric diabetes. *Diabetes Care* (2012). doi:10.2337/dc12-0836
224. Greenbaum, C. *et al.* Mixed-meal tolerance test versus glucagon stimulation test for the assessment of β-cell function in therapeutic trials in type 1 diabetes. *Diabetes Care* **31,** (2008).
225. Greenbaum, C. J. *et al.* Fall in C-peptide During First 2 Years From Diagnosis: Evidence of at Least Two Distinct Phases From Composite TrialNet Data. *Diabetes* 1–8 (2012). doi:10.2337/db11-1538
226. Sosenko, J. M., Skyler, J. S., Herold, K. C. & Palmer, J. P. The Metabolic Progression to Type 1 Diabetes as Indicated by Serial Oral Glucose Tolerance Testing in the Diabetes Prevention Trial-Type 1. *Diabetes* **61,** 1331–1337 (2012).
227. Jones, a. G. *et al.* Urine C-peptide creatinine ratio is an alternative to stimulated serum C-peptide measurement in late-onset, insulin-treated diabetes. *Diabet. Med.* **28,** 1034–1038 (2011).
228. Home Urine C-Peptide Creatinine Ratio Can Be Used to Monitor Islet Transplant Function.
229. Tatovic, D. *et al.* Stimulated urine C-peptide creatinine ratio vs serum C-peptide level

for monitoring of β-cell function in the first year after diagnosis of Type 1 diabetes. *Diabet. Med.* n/a-n/a (2016). doi:10.1111/dme.13186
230. Bollyky, J. B. *et al.* Heterogeneity in recent onset type 1 diabetes - a clinical trial perspective. *Diabetes. Metab. Res. Rev.* n/a-n/a (2015). doi:10.1002/dmrr.2643
231. Spoletini, M. *et al.* Low-risk HLA genotype in Type 1 diabetes is associated with less destruction of pancreatic B-cells 12 months after diagnosis. *Diabet. Med.* **24,** 1487–1490 (2007).
232. Skyler, J. S. Update on worldwide efforts to prevent type 1 diabetes. *Ann. N. Y. Acad. Sci.* **1150,** 190–196 (2008).
233. Mamchak, A. a *et al.* Preexisting Autoantibodies Predict Efficacy of Oral Insulin To Cure Autoimmune Diabetes in Combination With Anti-CD3. *Diabetes* (2012). doi:10.2337/db11-1304
234. Buzzetti, R. *et al.* C-peptide response and HLA genotypes in subjects with recent-onset type 1 diabetes after immunotherapy with Diapep277: an exploratory study. *Diabetes* **60,** 3067–72 (2011).
235. Arif, S. *et al.* Blood and islet phenotypes indicate immunological heterogeneity in type 1 diabetes. *Diabetes* **63,** 2–35 (2014).
236. Mortensen, H. *et al.* New Definition for the Partial Remission Period in Children and Adolescents With Type 1 Diabetes. *Diabetes Care* (2009). doi:10.2337/dc08-1987.
237. Skowera, A. *et al.* β-Cell–Specific CD8 T Cell Phenotype in Type 1 Diabetes Reflects Chronic Autoantigen Exposure. *Diabetes* **64,** 916–925 (2015).
238. Larché, M. & Wraith, D. C. Peptide-based therapeutic vaccines for allergic and autoimmune diseases. *Nat. Med.* **11,** 17–20 (2005).
239. Gibson, V. B. *et al.* Proinsulin multi-peptide immunotherapy induces antigen-specific regulatory T cells and limits autoimmunity in a humanized model. 251–260 (2015). doi:10.1111/cei.12687
240. Beam, C. a., Gitelman, S. E. & Palmer, J. Recommendations for the Definition of Clinical Responder in Insulin Preservation Studies. *Diabetes* **63,** 3120–3127 (2014).
241. Kochetkova, I. *et al.* IL-35 Stimulation of CD39 + Regulatory T Cells Confers Protection against Collagen II-Induced Arthritis via the Production of IL-10. *J. Immunol.* **184,** 7144–7153 (2010).
242. Kleijwegt, F. S. *et al.* Transfer of Regulatory Properties from Tolerogenic to Proinflammatory Dendritic Cells via Induced Autoreactive Regulatory T Cells. *J. Immunol.* **187,** 6357–6364 (2011).
243. Meier, S., Stark, R., Frentsch, M. & Thiel, A. The influence of different stimulation conditions on the assessment of antigen-induced CD154 expression on CD4+ T cells. *Cytom. Part A* **73,** 1035–1042 (2008).
244. Frentsch, M. *et al.* Direct access to CD4+ T cells specific for defined antigens according to CD154 expression. *Nat. Med.* **11,** 1118–1124 (2005).
245. Tree, T. I. M., Roep, B. O. & Peakman, M. Enhancing the sensitivity of assays to detect T cell reactivity: The effect of cell separation and cryopreservation media. *Ann. N. Y. Acad. Sci.* **1037,** 26–32 (2004).
246. Maecker, H. T. *et al.* Impact of cryopreservation on tetramer, cytokine flow cytometry, and ELISPOT. *BMC Immunol.* **6,** 17 (2005).
247. Owen, R. E. *et al.* Loss of T cell responses following long-term cryopreservation. *J. Immunol. Methods* **326,** 93–115 (2007).
248. Estorninho, M. *et al.* A novel approach to tracking antigen-experienced CD4 T cells into functional compartments via tandem deep and shallow TCR clonotyping. *J. Immunol.* **191,** 5430–40 (2013).
249. Altschul, S. *et al.* Gapped BLAST and PSI- BLAST: a new generation of protein database

search programs. *Nucleic acids Res* **25,** 3389–3402 (1997).
250. Kalendar, R., Lee, D. & Schulman, A. H. FastPCR software for PCR, in silico PCR, and oligonucleotide assembly and analysis. *Methods Mol. Biol.* **1116,** 271–302 (2014).
251. Maddelein, D. *et al.* The iceLogo web server and SOAP service for determining protein consensus sequences. *Nucleic Acids Res.* **43,** W543–W546 (2015).
252. Venturi, V., Kedzierska, K., Turner, S. J., Doherty, P. C. & Davenport, M. P. Methods for comparing the diversity of samples of the T cell receptor repertoire. *J. Immunol. Methods* **321,** 182–95 (2007).
253. Li, H. M. *et al.* TCR repertoire of CD4+ and CD8+ T cells is distinct in richness, distribution, and CDR3 amino acid composition. *J. Leukoc. Biol.* **99,** jlb.6A0215-071RR-(2015).
254. Murphy, K. *Janeway's Immunobiology*. (Garland Science, 2011).
255. nPOD TCR/BCR Search. Available at: http://clonesearch.jdrfnpod.org/. (Accessed: 9th September 2016)
256. Stadinski, B. D. *et al.* Hydrophobic CDR3 residues promote the development of self-reactive T cells. *Nat. Immunol.* **17,** 946–55 (2016).
257. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44,** D457–D462 (2016).
258. Huang, D. W., Lempicki, R. a & Sherman, B. T. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4,** 44–57 (2009).
259. Gorenshteyn, D. *et al.* Interactive Big Data Resource to Elucidate Human Immune Pathways and Diseases. *Immunity* **43,** 605–614 (2015).
260. Rutz, S., Wang, X. & Ouyang, W. The IL-20 subfamily of cytokines--from host defence to tissue homeostasis. *Nat Rev Immunol* **14,** 783–795 (2014).
261. Yeung, W. C. G. *et al.* Children with islet autoimmunity and enterovirus infection demonstrate a distinct cytokine profile. *Diabetes* **61,** 1500–1508 (2012).
262. Woo, J. S. *et al.* Junctophilin-4, a component of the endoplasmic reticulum-plasma membrane junctions, regulates Ca2+ dynamics in T cells. *Proc. Natl. Acad. Sci. U. S. A.* **113,** 2762–7 (2016).
263. Nishi, M., Mizushima, a, Nakagawara, K. I. & Takeshima, H. Characterization of human junctophilin subtype genes. *Biochem. Biophys. Res. Commun.* **273,** 920–927 (2000).
264. Daoud, a K., Tayyar, M. a, Fouda, I. M. & Harfeil, N. A. Effects of diabetes mellitus vs. in vitro hyperglycemia on select immune cell functions. *J. Immunotoxicol.* **6,** 36–41 (2009).
265. Heng, T. S. P. & Painter, M. W. The Immunological Genome Project: networks of gene expression in immune cells. *Nat. Immunol.* **9,** 1091–4 (2008).
266. Vandenbon, A. *et al.* Immuno-Navigator, a batch-corrected coexpression database, reveals cell type-specific gene networks in the immune system. *Proc. Natl. Acad. Sci. U. S. A.* **113,** E2393-402 (2016).
267. Radulovic, S., Santos, A. F., Brough, H. A., Phippard, D. & Ph, D. Randomized trial of peanut consumption in infants at risk for peanut allergy. *N. Engl. J. Med.* **372,** 803–813 (2015).
268. Steck, A. K., Armstrong, T. K., Babu, S. R. & Eisenbarth, G. S. Stepwise or Linear Decrease in Penetrance of Type 1 Diabetes With Lower-Risk HLA Genotypes Over the Past 40 Years. *Diabetes* **60,** 1045–1049 (2011).