Implications of HCV genotype 3 specific immunity on cross-reactive vaccine design

Annette von Delft
Christ Church
University of Oxford
Trinity Term 2014

A thesis submitted in partial requirement for the degree of
Doctor of Philosophy in Clinical Medicine

Nuffield Department of Medicine
Supervisor: Dr Eleanor Barnes

Acknowledgements

First of all, I would like to thank my supervisor, Ellie Barnes, who has been most supportive and inspirational during the course of my DPhil. She left me the freedom to discover my own scientific ideas, and wander off and fail on several of them; an invaluable experience for any scientist. I am also very honoured that she trusted me with this very interesting project. I value her helpful and sharp scientific comments, and am especially grateful for her flexibility with meeting times.

This study would not have been possible without the HCV infected patients who agreed to donate blood, and the clinical team caring for them at the John Radcliffe hospital Oxford: Jane Collier (consultant), Ellie Barnes (consultant), Lizzie Stafford, Denise O'Donnell, Mark Ainsworth (research nurses).

The scientific environment in Oxford has been both challenging and inspiring. I would like to thank the Nuffield Department of Medicine and Scatchert European Scholarship for funding and making my stay in Oxford possible. I am indebted to the combined Barnes/Klenerman groups for scientific input at Labmeetings and Journal clubs and to Paul for his support and comments on my work. In particular, I would like to thank Tony, Isla, Ayako, Leo and Chris Willberg for help with experiments and technical questions, and Jo and Ruth for proofreading this thesis.

To Ruth, the best colleague ever, with whom I shared endless discussions in our coffee breaks and designed the lab database Pedra: Thank you for always offering your support, and being available for brain dumps and walks whenever needed.

I would like to thank everyone who kept me grounded by enduring my intonation and inaccurate setting. Having a baby during my PhD, I quickly realized that my scientific success would depend on a wide support network: I would like to thank Claire for her wonderful childcare and willingness to work flexible hours, Genevieve, Ruth and the petals for babysitting, and the social support group from the lab:
Ruth, Alba, Cbel, Jo, Ayako, Lian Ni and Ellie.
This thesis would not have been possible without the continuous support of my family: thank you for always encouraging me to reach out for challenges.

And finally, Frank, my most supporting and loving husband:
Thank you for introducing me to the exciting world of science. Thanks for giving up countless days to look after baby Clara and enable me to finish this opus.
Ach, und danke fürs Necken. You never fail to make me smile.

Abstract: Implications of HCV genotype 3 specific immunity on cross-reactive vaccine design.

Annette von Delft, Christ Church
Submitted for a DPhil in Clinical Medicine, Trinity Term 2014

Hepatitis C virus (HCV) is a major global pathogen that infects an estimated 170 million people worldwide, and for which currently no vaccine is available. HCV is a highly diverse viral pathogen and exists as 6 major genotypes sharing only 75% sequence homology; developing a vaccine that is cross-reactive between genotypes is a major challenge. Defining immune responses that target different HCV genotypes will facilitate pan-genotypic T cell vaccine development.

HCV genotype 3 (gt3) is now the most common infecting genotype in the United Kingdom and large parts of Asia; however, data regarding the T cell antigenic targets of this genotype is very limited. In this thesis, HCV gt3 specific T cell targets were defined in acute, chronic and spontaneously resolved infection: in chronic gt3 infection, T cell responses were low in magnitude and narrowly focused in specificity, similar to those previously reported for gt1; in contrast, resolved infection was associated with a higher magnitude and broader specificity of CD4+ and CD8+ T cell responses across the genome. Overall, T cell specificity in gt3 infection was markedly different to that previously described for gt1, confirming that sequence differences between genotypes result in distinct immunological profiles.

Previous work from our laboratory demonstrated that, though T cell responses induced by a potent T cell vaccine containing HCV gt1b non-structural regions do target epitopes dominant in natural infection, induced T cells show limited cross-reactivity against other genotypes. In this thesis, it was assessed whether T cells primed in natural gt3 infection are able to recognize viral sequence variants at dominant epitopes, which would make these potential targets in cross-reactive vaccine design. For seven gt3-specific T cell epitopes identified here as dominant, major sequence variability was observed within and between genotypes, and limited T cell cross-reactivity observed against identified viral variants. This suggests that regions frequently targeted in natural infection may not serve as attractive targets for cross-reactive vaccine design

These results informed the subsequent design of a cross-reactive vaccine based on fragments of HCV that are conserved between genotypes. A generic algorithm was developed to define viral regions conserved between major HCV genotypes (for 1a/1b, 1/3a, 1-6), and these were joined to form immunogens between 819 and 1543 AA long. Possible artificial, non-HCV epitopes formed by junctions were identified using online epitope prediction servers, and abrogated through the insertion of 2-6 amino acid linkers. To address the concern that conserved regions may not be immunogenic, epitopes described in natural HCV infection were mapped on HCV sequences, showing that conserved segments are well populated with epitopes; additionally, strong binding peptides were predicted for conserved segments using online epitope prediction programs, suggesting potential *in vivo* immunogenicity.

In conclusion, HCV T cell specificity is distinct between genotypes, with limited T cell cross-reactivity between viral variants. Leading from this result, vaccine immunogens were designed entirely based on conserved viral regions. This work paves the way for future studies of novel HCV immunogens based on conserved viral segments between genotypes.

1 Introduction 18
1.1 Why study subtype specific HCV immune responses? or: Is a cross-reactive T cell vaccine necessary? 18
1.2 Hepatitis C virus 20
1.2.1 HCV prevalence 20
1.2.2 HCV structure and life cycle 20
1.2.3 HCV genotypes and variability of the HCV genome 24
1.3 Natural history of HCV infection 26
1.3.1 HCV transmission, symptoms and outcome 26
1.3.1.1 Chronic HCV infection 28
1.3.1.2 Spontaneous clearance of HCV infection 28
1.3.2 Treatment of HCV infection 30
1.3.2.1 Interferon/Ribavirin 30
1.3.2.2 Directly acting antivirals (DAAs) 31
1.4 Innate immune response to HCV infection 35
1.4.1 Interferon production and interferon signalling 35
1.4.2 Immunomodulatory and antiviral effects of interferons 38
1.4.3 NK cells in HCV infection 40
1.4.4 HCV subverts the innate immune response 41
1.4.5 Polymorphisms in the innate immune genes are associated with HCV infection and treatment outcome 42
1.5 Adaptive immune response to HCV infection 44
1.5.1 Antigenic peptide presentation - MHC class-I and -II processing 44
1.5.2 T cell responses in acute infection 46
1.5.3 T cell responses in spontaneously resolved HCV infection 48
1.5.4 T cell responses in chronic infection 49
1.5.4.1 Viral escape 50
1.5.4.2 CD8+ dysfunction 51
1.5.4.3 Lack of CD4+ help 52
1.5.4.4 Suppression by regulatory T cells 53
1.5.5 Humoral responses 54
1.5.6 The influence of the immunotolerant environment of the liver on HCV specific immune responses 55
1.6 Distinctive features of HCV genotype 3 infection 57
1.6.1 Distinct pathophysiological features of HCV genotype 3 infection 58
1.6.2 HCV genotype 3a and response to antiviral therapy 59
1.6.3 Innate immune responses in HCV gt3a infection 63
1.6.4 Adaptive immune responses in HCV gt3a infection 64
1.7 Aims of this work 66
2 Materials and Methods 67
2.1 Patient cohort 67
2.1.1 Patient recruitment 67
2.1.2 Treatment 67
2.2 Molecular Methods 67
2.2.1 Isolation of Plasma 67
2.2.2 Viral RNA extraction from Plasma 68
2.2.3 Primers 68
2.2.4 One-step Transcription First Round PCR 68
2.2.5 Second Round PCR 70
2.2.6 Gel electrophoresis and extraction of PCR products 70

2.2.7 Sequencing and DNA precipitation 71
2.2.8 Sequence analysis 71
2.2.9 DNA Purification 71
2.2.10 IL28b genotyping using real time PCR 71
2.2.11 HLA-typing 72
2.3 T cell assays 72
2.3.1 Isolation of PBMCs from whole blood 72
2.3.2 Freezing and thawing of cells 72
2.3.3 Interferon-γ ELISpot assays 72
2.3.4 CD8 depletion of whole PBMCs 74
2.3.5 T cell lines (for intracellular staining) 74
2.3.6 Proliferation assays 74
2.3.7 Fluorescence-activated cell sorting (FACS) analysis 75
2.3.7.1 Intracellular staining assays 75
2.3.7.2 HLA class I Pentamer assays 75
2.3.8 Data analysis 76
2.4 Peptides sets 76
2.4.1 Subtype 3a screening peptides 76
2.4.1.1 HCV gt3 overlapping peptides spanning the whole HCV genome .. 76
2.4.1.2 Peptide screening strategy 77
2.4.2 HLA predicted peptides spanning HCV non-structural regions 78
2.4.2.1 Identification of polymorphic sites associated with HLA alleles 78
2.4.2.2 Evaluation of putative CD8 HLA associated epitopes 78
2.4.3 Genotype 1 screening peptides 80
2.5 HLA restriction experiments 80
2.5.1 B cell lines 80
2.5.2 T cell lines (for HLA restriction experiments) 80
2.5.3 HLA restriction experiments 81
2.6 Statistical and data analysis 81
2.6.1 Analysis of epitopes described in natural HCV infection (IEDB) 81
2.6.2 Statistical methods 82
2.7 Work done by others 82
3 Identification of HCV specific T cells responses in HCV genotype 3a infection 83
3.1 Abstract 83
3.2 Background 84
3.2.1 Evidence for distinct features of HCV genotype 3 infection 84
3.2.2 A novel peptide screening approach to identify HCV gt3 specific T cell targets 85
3.3 Aims 87
3.4 Patient characteristics and study outline 88
3.5 T cell responses to overlapping peptide pools 92
3.5.1 T cell responses in resolved HCV infection 92
3.5.2 T cell responses in chronic HCV genotype 3a in infection 94
3.5.3 T cell responses in acute HCV genotype 3a infection 97
3.5.4 Comparison of magnitude of T cell responses and targeted viral regions using overlapping peptide pools in spontaneous resolvers, acutely and chronically infected patients 98
3.6 T cell responses to HLA predicted peptides 101
3.6.1 Responses to HLA predicted peptides in spontaneous resolvers 101

3.6.2 T cell responses to HLA predicted peptides in chronically infected patients 102
3.6.3 Responses to HLA predicted peptides in acutely infected patients 103
3.6.4 Responses to single peptides and viral variants 103
3.7 Comparison of T cell responses to overlapping and HLA matched peptide sets 105
3.8 High resolution analysis of HCV genotype 3 specific T cell responses .. 110
3.8.1 Confirmation of T cell responses measured by IFNγ ELISpot assay 110
3.8.2 Individual peptides targeted by T cells in HCV gt3 infection 113
3.8.3 Viral sequence analysis at individual peptides targeted by T cells 117
3.9 Comparison of T cell specificity in HCV gt1 and gt3 infection 121
3.9.1 Were T cell targets detected in the Oxford HCV gt3 cohort previously described in the literature? 121
3.9.2 Are CD8+ T cell epitopes dominant in HCV gt1 infection targeted in HCV gt3 infection? 128
3.9.3 Limited HCV gt1 and gt3 cross-reactivity on single peptide level 130
3.10 Associations of magnitude of HCV gt3 specific T cell responses with clinical predictors of spontaneous resolution of infection and treatment response 132
3.10.1 Associations of treatment outcome with known predictors of treatment response132
3.10.2 Magnitude of T cell response 134
3.10.3 IL-28B status and outcome of primary infection 136
3.11 Discussion and conclusions 137
4 Characterization of dominant T cell responses in HCV genotype 3a infection 148
4.1 Abstract 148
4.2 Introduction 149
4.3 Selection of frequently detected T cell responses in HCV gt3 infection in the Oxford HCV genotype 3 cohort. 152
4.4 CD4+ restricted T cell responses to HCV core frequently detected in chronic genotype 3a infection 153
4.4.1 Core responses in HCV gt3 infection evolve over time in acutely infected patients developing chronic infection 154
4.4.2 Sequence variability at CD4+ epitopes frequently targeted in HCV gt3 infection 156
4.4.3 Comparison of CD4+ T cell specificity against HCV core in HCV gt1 and gt3 infection 160
4.4.4 Cross-reactivity of CD4+ T cells targeting HCV gt3 core peptides against HCV gt1 sequence variants 162
4.4.5 Cytokine profile of core specific CD4+ responses 163
4.4.5.1 Ex-vivo ICS responses – cytokine profile 163
4.4.5.2 Short-term incubation of core specific CD4+ T cells at high cell concentrations to assessed cytokine production 165
4.5 Frequently detected CD8+ restricted T cell responses targeting HCV genotype 3 non-structural regions 169
4.5.1 Sequence variability at dominant CD8+ restricted T cell targets 170
4.5.1.1 HCV gt3 sequence variability at dominant CD8+ T cell targets on population level 170

4.5.1.2 Sequence variability at CD8+ T cell targets frequently detected within the Oxford HCV gt3 cohort, in context of T cell responses detected by IFNγ ELISpot and patients' HLA type 173
4.5.1.3 Sequence variability between HCV gt3 and gt1 at CD8+ T cell targets frequently detected in HCV gt3 infection 179
4.5.2 Cross-reactivity of T cell responses primed in HCV genotype 3 infection against common viral variants 181
4.5.3 Function of frequently detected HCV gt3 specific CD8+ T cell responses 184
4.5.3.1 Phenotypic analysis of CD8+ responses frequently detected by HLA predicted peptides 184
4.5.3.2 Phenotypic characterization of a CD8+ restricted T cell response frequently detected using overlapping peptides...188
4.6 Discussion...193
5 Immunogen design 198
5.1 Abstract..198
5.2 Rationale 199
5.2.1 Do we need a vaccine for HCV?...199
5.2.2 Current HCV vaccine candidates...199
5.2.3 Why develop a T cell vaccine for HCV? 201
5.2.4 Cross-reactivity of HCV genotype 1b vaccine-induced responses........203
5.2.4.1 Genotype cross-reactivity of vaccine-induced responses 204
5.2.4.2 Cross-reactivity against quasi-species in a single host 204
5.2.5 Limited cross-reactivity of HCV gt3 and gt1 subtype specific T cell responses in natural infection..207
5.2.6 Rationale for a HCV vaccine based on conserved viral regions............207
5.2.7 Aims of this work...210
5.3 Programming approach for conserved immunogen design 211
5.3.1 Sequence selection for HCV genotypes 1 to 6......................................211
5.3.2 Calculating sequence diversity within a sequence dataset.....................212
5.3.3 Approaches for immunogen design 213
5.3.4 Comparison of conserved sequence segments for HCV gt1, HCV gt1/3a and HCV gt1-6 datasets 216
5.4 Sequence selection for final immunogens 219
5.4.1 Comparison of within-genotype consensus sequences to all-genotype consensus 219
5.4.2 Selection of patient sequences for final immunogen constructs............224
5.5 Evaluation of potential immunogenicity of designed immunogens 227
5.5.1 Epitopes in conserved regions described in natural infection...............228
5.5.1.1 Do CD8+ epitopes described in natural HCV gt1 infection fall into conserved regions?..228
5.5.1.2 Are HCV gt1 epitopes previously linked to protection included in the immunogens? 230
5.5.1.3 Epitopes described in HCV gt3 infection 231
5.5.2 Evaluation of immunogenicity using epitope prediction programs.......233
5.5.2.1 Evaluation of HLA type distribution 233
5.5.2.2 Choosing an epitope prediction server...234
5.5.2.3 Are the designed immunogens immunogenic? - MHC class-I epitope prediction using the NetMHC server 236
5.5.3 Putative artificial MHC class-I epitopes in junction areas....................238

5.5.3.1 Avoiding potential artificial epitope: Usage of linkers 242
5.6 Adaptations of the designed immunogens 245
5.6.1 TPA leader sequence and invariant chain 245
5.6.2 BLAST assessment for selected immunogens 245
5.6.3 Safety 246
5.7 Summary of results and discussion 247
5.7.1 Final immunogens 247
5.7.2 Discussion of results 248
5.7.3 Cross-reactive HCV vaccine – the future 254
6 Concluding remarks 260
7 References 268
8 Appendix 288
8.1 Supplementary figures 288
8.2 Comparative analysis of CD8+ restricted epitopes predicted by three different online epitope prediction servers in junction regions between conserved epitopes 291
8.2.1 NetMHC – Syfpeithi 291
8.2.2 NetMHC - BIMAS 296
8.2.3 BIMAS – Syfpeithi 301
8.3 Linker design at junction regions between conserved segments to avoid predicted strong binders 305

Table of Figures

Figure 1-1: Phylogenetic analysis of Flaviviridae based on polyprotein sequences ...21
Figure 1-2: Schematic representation of the HCV polyprotein. 21
Figure 1-5: Phylogenetic tree of the seven HCV genotypes. 25
Figure 1-6: Global distribution of HCV genotypes. 26
Figure 1-7: Natural history of HCV infection..27
Figure 1-8: Directly acting antivirals: Protease Inhibitors, NS5A Inhibitors and Polymerase Inhibitors. 32
Figure 1-10: Binding of different types of Interferons to their receptors and following intracellular pathways. 38
Figure 1-11: Peptide loading of MHC class-I and -II complexes. 44
Figure 1-12: Schematic of the clinical course and adaptive immune responses in acute HCV infection in self-limiting and chronically evolving HCV infection...........47
Figure 1-13: Distinctive features of HCV genotype 3 infection 57
Figure 1-14: Interferon free trials including Sofosbuvir for HCV genotype 3 subjects. 62
Figure 2-1: Strategy for HCV sequencing reactions..70
Figure 2-3: Representative gating strategy for analysis of FACS samples 76
Figure 2-4: Peptide mapping strategy 77
Figure 3-1: Study outline for the assessment of HCV gt3 T cell responses 91
Figure 3-2: T cell responses in spontaneous resolvers to HCV genotype 3a specific peptides 92
Figure 3-3: T cell responses in spontaneous resolvers to HCV genotype 1b specific peptides 93
Figure 3-4: Comparison of T cell responses in spontaneously resolved patients against HCV genotype 3a and 1b peptide sets 93
Figure 3-5: T cells responses against a HCV gt3a peptide set in chronically infected HCV gt3a patients...95
Figure 3-6: T cell responses against a HCV gt1b peptide set in chronically infected HCV gt1 patients...96
Figure 3-7: Comparison of T cell responses to HCV viral regions in HCV gt1 and gt3 infection. 97
Figure 3-8: Magnitude of T cell responses did not differ in acutely infected patients who spontaneously cleared infection or developed chronic infection. 98
Figure 3-9: Magnitude of T cell responses in spontaneously resolved patients and acute chronic HCV genotype 3a infected patients 99
Figure 3-10: T cell responses of spontaneously resolved, acutely and chronically infected HCV patients target different viral region 100
Figure 3-11: T cell responses in spontaneous resolvers to HLA predicted peptides .101
Figure 3-12: T cell responses in chronically infected patients to HLA predicted peptides 102
Figure 3-13: T cell responses to HLA predicted peptides in acutely infected HCV gt3a patients. 103
Figure 3-14: Details of all responses to HLA predicted peptides. 104
Figure 3-15: Comparison of responses to overlapping pools and HLA predicted peptides in spontaneous resolvers and acutely and chronically infected patients. 106
Figure 3-16: HCV specific T cell responses to HLA predicted peptides. 107
Figure 3-17: Representative T cell responses measured by IFNγ ELISpot assay. 111

Figure 3-18: CD4+ and CD8+ specific T cell responses measured by ICS. 112
Figure 3-19: CD4/CD8 responses in spontaneous resolvers and chronic patients. ...116
Figure 3-20: GT3 sequence variants of commonly targeted epitopes in HCV genotype 1 infection. 129
Figure 3-21: Association between age and sex and treatment outcome with peg-IFNα and Ribavirin...132
Figure 3-22: Pre-treatment viral load and treatment outcome. 133
Figure 3-23: IL28B and influence on treatment response and baseline viral load. ...133
Figure 3-24: Associations of magnitude of T cell response and treatment outcome, age or sex...134
Figure 3-25: Magnitude of T cell response and IL28B genotype. 135
Figure 3-26: IL28B genotype and clinical outcome of primary infection. 136
Figure 4-1: Experimental overview for Chapter 4. 152
Figure 4-2: Viral load and HCV specific T cell responses in acutely infected patients measured over time. 155
Figure 4-3: Shannon entropy of core sequences. 157
Figure 4-4: Sequence information at frequently detected T cell targets PKARRSEGRSWAQPGYPW and PVGGVARALAHGVRAL in the HCV core region in HCV gt3 patients with and without a T cell response. 159
Figure 4-5: Comparison of CD4+ T cell responses to the core region in HCV genotype 1 and 3 infection. 160
Figure 4-6: Experimental cross-reactivity against HCV gt3 and gt1 sequence variants at dominant CD4+ restricted HCV core T cell targets. 162
Figure 4-7: Verification of HCV gt3a specific core responses. 163
Figure 4-8: Cytokine profile of core specific HCV gt3a responses. 164
Figure 4-9: Representative FACS plots for T cell responses ex-vivo and after short-term incubation. 166
Figure 4-10: Cytokine release of core specific T cells after short-term incubation...167
Figure 4-11: Comparison of polyfunctionality of HCV gt3 core specific T cells after short-term incubation. 167
Figure 4-12: Sequence heterogeneity on population level at frequently detected CD8+ restricted T cell targets within HCV non-structural regions. 171
Figure 4-13: Sequence information at dominant epitope $NS2_{886}$ (L)LYPSLIFDI in HCV gt3 patients with and without a T cell response. 174
Figure 4-14: Sequence information at dominant epitope $NS3_{1443}$ ATDALMTGY/F (A*01) in HCV gt3 patients with and without a T cell response. 175
Figure 4-15: Sequence information at dominant epitope $NS4b_{1853}$ RVLLDILAGY (A*26) in HCV gt3 patients with and without a T cell response. 176
Figure 4-16: Sequence information at frequently detected T cell targets $NS3_{1520}$ (RPSGMFDSVVLCECYDAGCSWYDL) and $NS5a_{2126}$ (AEFFTEVDGVRLHRYA) in HCV gt3 patients with and without a T cell response. 178
Figure 4-17: Comparison of HCV gt1 and gt3 sequences at frequently detected HCV gt3 specific CD8+ T cell targets. 180
Figure 4-18: T cell cross-reactivity against HCV gt3 variants at CD8+ T cell targets frequently detected in HCV gt3 infection. 181
Figure 4-19: T cell cross-reactivity against HCV gt1 sequence variants at frequently detected CD8+ T cell targets..183
Figure 4-20: T cell responses to HLA predicted peptides frequently detected in IFNγ ELISpot assays assessed by intracellular staining assay..................................184

Figure 4-21: Magnitude of T cell responses in IFNγ ELISpot to HLA predicted peptides, which were further tested in ICS assays. 185
Figure 4-22: Phenotypic profile of ADTAMTGY in acute and chronic HCV gt3 infection. 187
Figure 4-23: Responses to peptide $NS3_{1520}$ were recovered by stimulation with IL-2. 188
Figure 4-24: HLA type distributions in patients with and without T cell response to peptide $NS3_{1520}$. 189
Figure 4-25: HLA restriction experiments at peptide $NS3_{1520}$.................................190
Figure 4-26: Strategies of optimal mapping for overlapping peptide $NS3_{1520}$ and representative ELISpot responses to different optimal mapping strategies.......191
Figure 5-1: HCV vaccine candidates 200
Figure 5-2: Magnitude of T cell responses to different vaccine approaches. 201
Figure 5-3: HCV viral load and liver parameters in animals vaccinated with a prophylactic vaccine and control animals...202
Figure 5-4: HCV genotype 1b insert...203
Figure 5-5: Cross-reactivity of immune responses against HCV genotype 1b, 1a, 3a and 4a peptide sets in patients vaccinated with a HCV genotype 1b insert.......204
Figure 5-6: Cross-reactivity of vaccine-induced T cell responses against circulating patient variants. 205
Figure 5-7: Sequence diversity of the entire HCV genome between HCV genotypes 1 and 3. 212
Figure 5-8: Approaches for immunogen design 213
Figure 5-9: Pilot epitope prediction analysis for immunogens for HCV gt1a, 1b and 3a designed by approaches A to D. 215
Figure 5-10: Comparison of conserved segments for HCV gt1, HCV gt1/3a and HCV gt1-6 datasets (Figure next pages) 216
Figure 5-11: Selection process for sequences to be included in the final immunogen 220
Figure 5-12: Comparison of HCV genotype 1a and 1b consensus sequences to the overall HCV genotype 1 consensus sequence 221
Figure 5-13: Comparison of HCV genotype 1a, 1b and 3a consensus sequences to the overall HCV genotype 1/3a consensus sequence...222
Figure 5-14: Comparison of HCV genotype 1, 2, 3, 4, 5 and 6 consensus sequences to the overall HCV genotype 1-6 consensus sequence 223
Figure 5-15: Comparison of subtype consensus (all segments) to overall consensus. 224
Figure 5-16: Number of patient sequences of each subtype included in the final immunogen. 227
Figure 5-17: HLA type distribution of published epitopes as derived from the IEDB. 229
Figure 5-18: Distribution of epitopes previously described in natural HCV gt1 infection in relation to HCV gt1 sequence variability 230
Figure 5-19: Distribution of epitopes previously described in natural HCV gt3 infection in relation to HCV gt1/3a sequence variability 232
Figure 5-20: Frequency of HLA types worldwide..233
Figure 5-21: Numbers of predicted strong binders to MHC class-I epitopes. 237
Figure 5-22: HLA type distributions of predicted epitopes in different analyses and immunogens. 237
Figure 5-23: Potential artificial epitopes generated at segment junctions 238

Figure 5-24: Detailed numbers of potential artificial epitopes in junction regions for all immunogen designs. 239
Figure 5-25: Strategy for comparing different epitope prediction servers 240
Figure 5-26: Examples for abrogated predicted strong binders by insertion of linkers 245
Figure 5-27: Schematics of final immunogens 247
Figure 8-1: Sequence diversity of full-length genotype-3a sequences. 288
Figure 8-2: T cell responses in HCV gt1 infection measured by IFNγ ELISPot assays using HCV subtype specific peptide sets. 289
Figure 8-3: Sequence heterogeneity (Shannon entropy) of NS3 sequences 290
Figure 8-4: HLA type distribution of predicted strong binders forming artificial epitopes in junction areas using the Syfpeithi prediction server........................291
Figure 8-5: HLA type distribution of predicted strong binders forming artificial epitopes in junction areas using the BIMAS prediction server..........................296

Table of Tables

Table 1-1: Predictors of treatment response 31
Table 1-2: Drug profiles of directly acting antivirals 33
Table 1-3: Response rates to combination therapy in HCV genotype 3a patients.......60
Table 1-4: Leading protease inhibitors are highly active against multiple HCV genotypes except HCV gt3 61
Table 2-1: HCV genotype 3a specific forward primers..........69
Table 2-2: HCV genotype 3a specific reverse primers..........69
Table 2-3: Characteristics of phenotypic markers analysed by HLA class-I tetramer staining...........75
Table 2-4: HCV genotype 3a specific peptides - overlapping peptide pools 77
Table 2-5: HLA-associated peptides – polymorphisms within predicted epitope.......79
Table 2-6: HLA-associated peptides – polymorphics flanking predicted epitopes.....80
Table 3-1: Patient characteristics of spontaneously resolved patients..........88
Table 3-2: Patient details chronic HCV gt3a cohort..........89
Table 3-3: Patient details acute HCV gt3a cohort. 90
Table 3-4: Details of T cell responses to HLA predicted peptides in spontaneously resolved, acute and chronic patients. 108
Table 3-5: Sequence difference between overlapping peptide pools and HLA predicted peptides. 109
Table 3-7: T cell responses to individual peptides in HCV non-structural regions mapped from overlapping peptide pools..........114
Table 3-8: Viral consensus sequence analysis at T cell targets detected in HCV structural regions..........118
Table 3-9: Viral sequence analysis at T cell targets detected in non-structural regions. (Table on the next page). 119
Table 3-10: CD8+ restricted T cell targets previously described in the literature in the region of CD8+ restricted T cell targets defined in this study...........122
Table 3-11: Sequence comparison for HCV gt3 and gt1 at newly described HCV gt3 CD8+ restricted T cell targets...........123
Table 3-12: CD4+ T cell targets previously described in the literature in the region of CD4+ T cell targets defined in HCV gt3 infection in this study. 125
Table 3-13: HCV gt1 T cell targets previously described in the literature in the region of HCV gt3 T cell targets with non-defined CD4/CD8 restriction...........127
Table 3-14: Were dominant HCV gt1 CD8+ restricted epitopes detected in the Oxford cohort?..........128
Table 3-15: HCV gt3a and gt1 peptide cross-reactivity. 130
Table 4-1: Frequently detected HCV gt3 specific T cell responses..........153
Table 4-2: HCV gt1 and 3 sequence comparison at commonly targeted core peptides. 161
Table 5-1: Comparison of vaccine epitope sequences at immunodominant epitopes with circulating viral sequences in patients. 206
Table 5-2: Sequence selection for immunogen design, overview. 211
Table 5-3: Potential advantages/ disadvantages for different immunogen designs...214
Table 5-4: Patient sequences selected for final HCV gt1 immunogen. 225
Table 5-5: Are epitopes linked to protection in HCV genotype 1 included in the conserved segments. 231
Table 5-6: Evaluated MHC class-I epitope prediction servers and computational methods..........234

Table 5-7: Evaluated MHC class-II epitope prediction servers and computational methods 235
Table 5-8: Inserted linker regions for HCV gt1 immunogen (approach D) 243
Table 8-1: MHC class-I types assessed for epitope prediction using the Syfpeithi prediction server....292
Table 8-2: Predicted strong binders in junction areas excluded from the comparative analysis, because of non-matching HLA types in the comparative analysis between NetMHC and Syfpeithi servers....293
Table 8-3: Comparison of MHC class-I prediction results using NetMHC and Syfpeithi servers (table on the next pages). 293
Table 8-4: MHC class-I types assessed for epitope prediction using the BIMAS prediction server....297
Table 8-5: Excluded predicted strong binders in junction areas in comparative analysis, because of non-matching HLA types in comparative analysis between NetMHC and BIMAS servers. 298
Table 8-6: Comparison of MHC class-I prediction results using NetMHC and BIMAS servers. 299
Table 8-7: MHC class-I types shared between Syfpeithi and BIMAS servers. 301
Table 8-8: Excluded predicted strong binders in junction areas in comparative analysis, because of non-matching HLA types in comparative analysis between Syfpeithi and BIMAS servers. 302
Table 8-9: Comparison of MHC class-I prediction results using Syfpeithi and BIMAS servers (table on next pages). 302
Table 8-10: Prediction of artificial epitopes in junction regions by three different servers and modification through linkers (table depicted on the next 3 pages). 305

Abbreviations

AA	Amino acid
ALT	amino
ANN	artificial neural network
AST	aspartate aminotransferase
BSA	Bovine serum albumine
CD	Cluster Differentiation
cDNA	double stranded DNA
CMV	Cytomegalovirus
ConA	concavidin A
CTL	cytotoxic T lymphocytes
DAA	directly acting antivirals
DCs	dendritic cells
DMSO	Dymethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsRNA	double stranded RNA
E1	HCV envelope protein 1
E2	HCV envelope protein 1
EBV	Epstein Bahr virus
FCS	Fetal calf serum
FACS	Fluorescence-activated cell sorting
FEC	Flu, EBV and CMV CD8+ specific epitopes
gt	genotype
HAART	highly active antiretroviral therapy
HBV	Hepatitis B virus
HCC	hepatocellular carcinoma
HCV	hepatitis C virus
HIV	human immunodeficiency virus
HLA	human leukocyte antigen
ICS	intracellular staining
IEDB	immune epitope dabatase (www.iedb.com)
IFN	Interferon
IFN/RIB	combination treatment with pegylated-IInterferon-α and Ribavirin
IFNγ	Interferon gamma
IL	Interleukin
ISG	interferon stimulated genes
IVDU	intravenous drug users
JAK/STAT	Janus kinase/signal transducers and activators of transcription
MHC	major histocompatibilty complex
MSM	men who have sex with men
NK	natural killer cells
NKT	natural killer T cells
NS	not significant
PBMC	peripheral blood mononuclear cells
PD-1	programmed death 1
Pt	patients
RIG-I	retinoic acid–inducible gene I
RBV	Ribavirin
RNA	Ribonucleic acid
RSV	respiratory syncytial virus
SEM	standard error of means
SFU	Spot forming units
STAT	Signal transducer and activator of transcription
SVR	Sustained virological response
ssRNA	single-stranded RNA
TCR	T cell receptor
Th1	T helper 1 CD4+ T cells
Tim-3	T cell immunoglobulin and mucin domain–containing molecule 3
tPA	Tissue plasminogen activator
TRAIL	tumor necrosis factor-related apoptosis-inducing ligand
UK	United Kingdom
WHO	world health organization

1 Introduction

1.1 Why study subtype specific HCV immune responses? or: Is a cross-reactive T cell vaccine necessary?

Hepatitis C virus (HCV) is a major pathogen leading to a significant burden of disease, including hepatocellular cancer, liver transplantation and death. Up to now, treatment of chronic infection was long and associated with multiple side effects, but major advances have been achieved over recent years with an arsenal of effective drugs approved or expected to be approved soon. However, most patients infected worldwide will not have access to the new expensive treatment options, or do not know that they are infected in the first place, preventing them from being treated. Therefore, an effective vaccine for HCV virus remains a pressing need.

A major challenge for HCV vaccine design is the extensive diversity of the virus, which is one of the most diverse pathogens known. HCV has seven known major genotypes that vary in 20-30% of their sequence, and have different local distributions. An ideal HCV vaccine would not only prevent infection with a single strain, but also with different genotypes of the virus. This is of particular importance in risk groups like intravenous drug users, where patients might be exposed to multiple strains over time. Vaccines being currently developed are exclusively based on HCV genotype 1 (gt1) immunogens, which is the most common infecting HCV genotype.

A strong and broad T cell response to HCV infection has been linked to spontaneous resolution of HCV infection, providing a rationale for a T cell based vaccine design. We have shown previously that potent T cell vaccines are capable of inducing HCV specific responses that might protect against infection. However, T cell responses primed against the subtype 1b insert were not cross-reactive against other HCV genotypes and subtypes *in vitro* (Swadling et al., manuscript in preparation).

Therefore, the main aim of this thesis was to design a cross-reactive vaccine covering all HCV genotypes. To achieve this, detailed knowledge about immunity in natural infection is necessary. Extensive research has been published on T cell immunity in HCV gt1 infection, but data on T cell immunogenicity in HCV genotypes other than HCV gt1 is scarce. HCV genotype 3, now the most common infecting genotype in the United Kingdom, is a particularly interesting genotype to study: major inter-genotypic differences between HCV genotype 1 and 3 have been described, regarding different rates of spontaneous resolution of infection, disease progression and development of

steatosis and fibrosis. In addition, a distinct treatment response to both combination therapy with pegylated Interferon-α and Ribavirin (IFN/RIB) and newer HCV treatment options like directly acting antivirals has been observed.
In this thesis, I initially assessed T cell immunity in acute, chronic and spontaneously resolved HCV genotype 3 infection, and compared it to known data on HCV genotype 1 infection. Subsequently, these data informed the design of vaccine constructs based on conserved viral regions, with the aim to develop a HCV vaccine that generates cross-reactive T cell responses against multiple genotypes.

In the introduction to this thesis, I cover general knowledge on HCV infection (including HCV prevalence, structure, life cycle and variability), and then discuss features specific to HCV genotype 3 infection. I include recent data on new treatment options for HCV, particularly focusing on their efficiency in HCV genotype 3, and recent genotype specific discoveries regarding differences in innate immune responses to the different genotypes. Next, I focus on HCV adaptive immunity, and the role of the immunomodulatory environment of the liver as the place of primary infection. In addition to the overall introduction, specific background is discussed at the start of each chapter.

1.2 Hepatitis C virus

1.2.1 HCV prevalence

Hepatitis C virus (HCV) is a major viral pathogen. The world health organization (WHO) estimates that about 3% of the world's population has been infected with the virus and that there are about 150 million chronic carriers, with 3-4 million new infections every year (WHO 2014). A recent study greatly exceeded the numbers published by WHO, with estimates ranging at over 185 million chronic carriers (Mohd Hanafiah et al. 2013). Estimating the prevalence of HCV remains complex, as HCV is asymptomatic for most of its incubation time, and injecting drug users, who are at greatest risk for infection, are a difficult group to study.

Overall, prevalence increased from 2.3% (1990) to 2.8% (2005) (Mohd Hanafiah et al. 2013). Prevalence ranges from 0.5% to over 5% in some countries (WHO 2014; Negro & Alberti 2011), with the highest prevalence in Egypt, Pakistan, Russia and South America. In the UK, the prevalence of HCV is approximately 0.67%, or 203.000 individuals, with similar numbers observed for Northern Europe, the US and Canada (Health Protection Agency 2012; R. J. Harris et al. 2012). The UK prevalence is much higher in immigrant communities, reflecting higher prevalence rates between 5 to 10% in developing countries, with rates peaking in Egypt between 15 to 20% (Hajarizadeh et al. 2013).

1.2.2 HCV structure and life cycle

HCV belongs to the family of *Flaviviridae*, which also includes other classic flaviviruses such as dengue and yellow fever (Figure 1-1). HCV is grouped into the class *Hepacivirus* with the animal homologs GB virus B (GBV-B), a virus recovered from laboratory housed tamarins (Simons et al. 1995), and recently identified rodent, equine and canine hepaciviruses (Kapoor et al. 2011; Burbelo et al. 2012; Kapoor et al. 2013; Stapleton et al. 2011; D. B. Smith et al. 2013; Drexler et al. 2013). Several other GB viruses (GBV-A, GBV-C and GBV-D do not cause Hepatitis and are classed as Pegi viruses (Stapleton et al. 2011).

The origin of HCV in human hosts remains unclear: similarly to the zoonotic origin of HCV-1 from chimpanzees in Central Africa, non-human primates have long been suspected as a source, however, no published evidence to date has been obtained supporting this hypothesis(Simmonds 2013).

Figure 1-1: Phylogenetic analysis of Flaviviridae based on polyprotein sequences

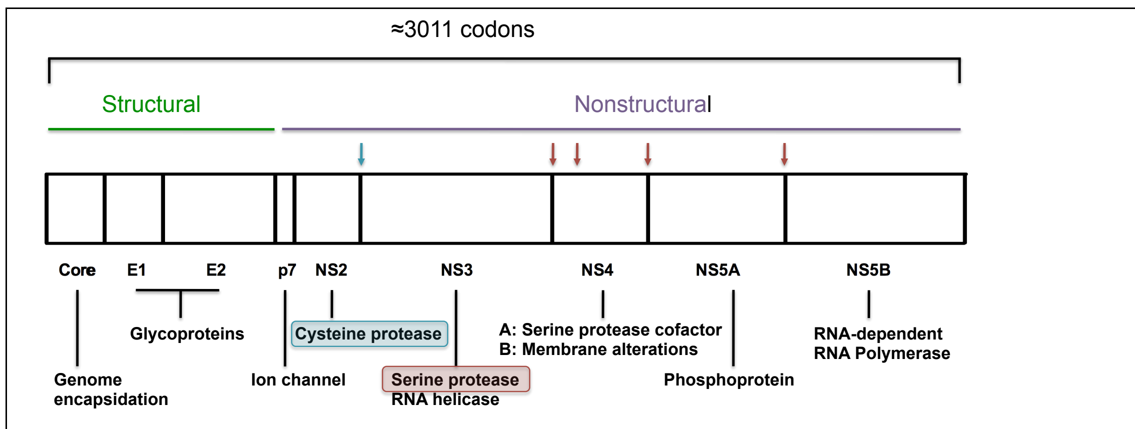
Figure deleted due to copyright:

Figure material can be found in Drexler et al., 2013. Figure included in thesis has been derived from information shown in Figure 4a in this paper.

Bayesian phylogeny of the Flaviviridae NS5B gene, modified from (Drexler et al. 2013). Representatives of all *Flaviviridae* genera and five novel rodent viruses are depicted. Scale bar corresponds to genetic distance. A tombusvirus (Lisianthus necrosis virus, GenBank accession number NC_007983) was chosen as an outgroup.

The enveloped virus contains a positive sense single stranded virus, and has a 9.6 kilobase (kb) long genome, flanked by two un-translated regions (5' and 3' UTR). It consists of an open reading frame, which translates into a poly-protein of approximately 3000 codons, containing structural (core, E1, E2) and non-structural proteins (p7, NS2, NS3, NS4A/B, NS5A/B, Figure 1-2) (Scheel & Rice 2013).

Figure 1-2: Schematic representation of the HCV polyprotein.



Schematic representation of the structure of HCV, with structural and non-structural regions marked. Gene products and functions of the structural and non-structural proteins are shown. Arrows mark cleaving points of the cystein protease (blue) and the serine protease (red).

The HCV life cycle within the human host is depicted in Figure 1-3.

Figure 1-3: HCV life cycle

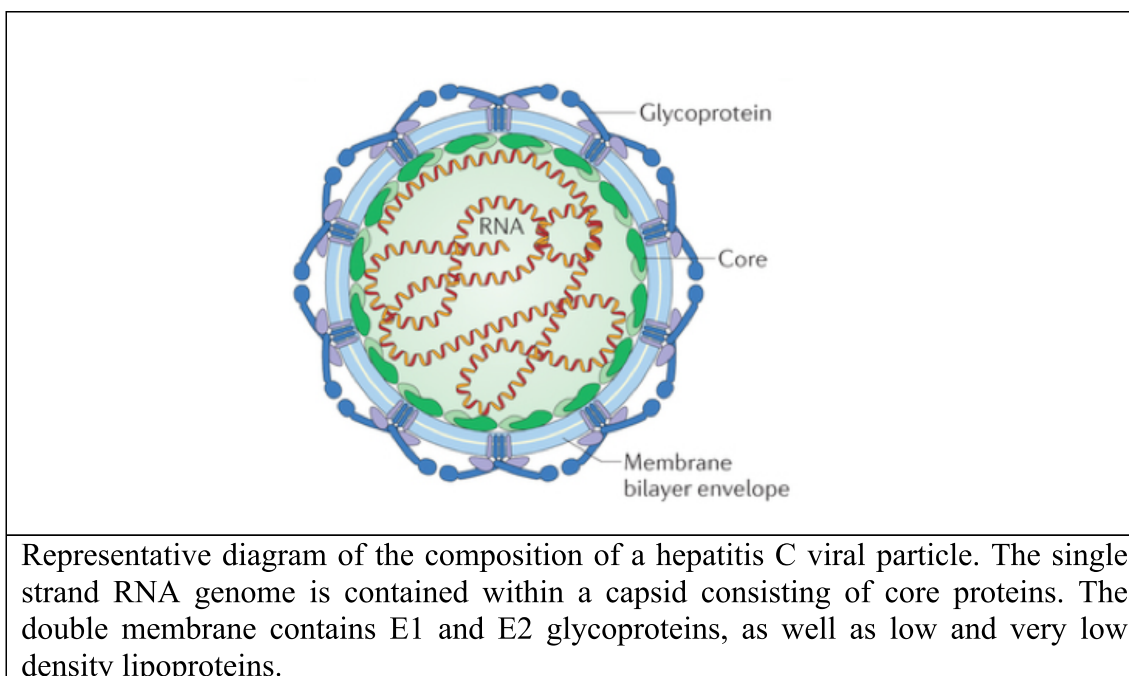
Figure deleted due to copyright:

Figure material can be found in Scheel and Rice., 2013. Figure included in thesis has been derived from information shown in Figure 3 in this paper.

Depiction of the HCV life cycle. Figure modified from Scheel and Rice, 2013.
(1) The lipoviral particle (LVP) interacts with cellular surface receptors. (2) Entry process: Binding to cellular surface receptors (CD81/SR-B1 and Claudin/Occludin). (3) After pH-dependent fusion and uncoating, the incoming HCV genome is translated and the resulting polyprotein processed (bottom inset). (4) Replication takes place in ER-derived membrane spherules (membranous web, bottom right inset). (5) In the assembly and release process (top right inset), core protein is transferred from cytoplasmatic lipid droplets (cLDs) to form nucleocapsids that, assisted by NS5A, are loaded with RNA. The p7, NS2 and NS3-NS4A proteins are also involved in coordination of assembly. HCV virion morphogenesis is coupled to the VLDL pathway, and particles are produced as LVPs.
EphA2, ephrin receptor type A2; EGF, epidermal growth factor receptor; LDLR, low density lipoprotein receptor; GAG, glycosaminoglycans; CLDN1 Claudin 1 OCLN Occludin; ER, Endoplasmatic reticulum; cLD cytoplasmatic lipid droplets luLD luminal lipid droples; PL, phospholipids; TG, triglycerides; MTP, microsomal transfer protein; LVP, lipoviral particle, VLDL, very low density lipoprotein.

HCV infects the host in form of a lipoviral particle, incorporating mainly low and very low density lipoproteins (LDL and VLDL) (S. Nielsen et al. 2006; Bassendine et al. 2013) (Figure 1-4). Incorporated into the particle membrane are viral envelope proteins E1 and E2 as heterodimers, with a nucleocapsid (core) containing the viral RNA (Figure 1-3 (1)).

Figure 1-4: Model of the HCV particle



Representative diagram of the composition of a hepatitis C viral particle. The single strand RNA genome is contained within a capsid consisting of core proteins. The double membrane contains E1 and E2 glycoproteins, as well as low and very low density lipoproteins.

The virus enters the cell via a combination of specific receptors (Figure 1-3 (2)): the LDL receptor and glycosaminoglycans are responsible for the low-affinity binding (McLauchlan et al. 2002; Lorenz et al. 2006). Then E1 and E2, highly variable glycoproteins forming the surface of the virus, form bonds with scavenger receptor class B member 1 (SRB1) and CD81 (P Pileri et al. 1998; Agnello et al. 1999; Barth et al. 2003). Other proteins required for HCV entry are Claudin, which induces clathrin-mediated endocytosis (M. J. Evans et al. 2007; Farquhar et al. 2012), and Occludin, which is also essential for HCV entry, but its precise role remains unknown (Ploss et al. 2009; Sourisseau et al. 2013). The late stages of viral entry, including fusion with host membranes and un-coating, are poorly understood (Scheel & Rice 2013), but seem to be dependent on a low pH (Sharma et al. 2011).

After the virus enters the cytoplasm, HCV translation of the virus into a polyprotein is initiated via an internal ribosome entry site (IRES) (B. Hoffman & Q. Liu 2011) (depicted in Figure 1-3 (3)). Individual proteins are released co- and post-translationally from the polyprotein through both cellular (signalase and signal peptide peptidase) and viral (NS2-NS3 cysteine protease, NS3-NS4a serine protease) proteases (McLauchlan et al. 2002; Lorenz et al. 2006). The viral helicase encoded by NS3 facilitates viral RNA processing.

HCV replication (Figure 1-3 (4)) is believed to replicate in close proximity to the endoplasmatic reticulum (ER) in single and double membrane vesicles also termed

'membranous web', derived from the rough ER, early and late endosomes, COP vesicles, mitochondria and lipid droplets (Romero-Brey et al. 2012). The main actor in viral replication is the RNA-dependent RNA polymerase (NS5B), responsible for the synthesis of new viral genomes (Appel et al. 2006). The lack of proof-reading function in this polymerase contributes to the high variability of HCV. Other non-structural proteins are known to play a pivotal role in viral replication: NS4B as an organizer of the viral replication complex and membranous web, inducing the formation of ER derived vesicles (D. Egger et al. 2002; Romero-Brey et al. 2012), and NS5A as a regulator of viral replication (Lim & Hwang 2011; Lemay et al. 2013). Viral assembly and release (Figure 1-3 (5)) is a process intimately coupled with the cell's lipid cycle (Miyanari et al. 2007). After cleavage, the core protein relocates to the lipid droplets or (after translocation of cytoplasmatic lipid droplets) to the ER (McLauchlan et al. 2002; Boulant et al. 2006; Miyanari et al. 2007), where the nucleocapsid incorporating viral RNA is formed under the assistance of NS5A (Appel et al. 2006). Then, HCV 'hitches a ride' on the VLDL pathway (H. Huang et al. 2007; Gastaminza et al. 2008), a strategy that possibly enhances hepatotropism, and exits the cell through the Golgi (Coller et al. 2012; Scheel & Rice 2013).

1.2.3 HCV genotypes and variability of the HCV genome

HCV is a genetically diverse RNA virus with seven major genotype groups and 67 subtypes (Figure 1-5) (Simmonds 2004; D. B. Smith et al. 2013). Genotypes have identical complement of genes (as shown in Figure 1-2) of similar or identical size within the open reading frame. There is significant divergence between genotypes, which differ by approximately 30% at nucleotide level, and subtypes, differing by 20-25% (Simmonds et al. 2005).

Figure 1-5: Phylogenetic tree of the seven HCV genotypes.

Figure deleted due to copyright:
Figure material can be found in Simmonds, 2013. Figure included in thesis has been derived from information shown in Figure 3 and includes the phylogenetic tree of seven HCV genotypes.

Evolutionary tree of NS5B sequences of HCV genotypes 1–7 (positions 8276–8615 as numbered as in the H77 reference sequence). The scale bar depicts an evolutionary distance of 0.05. Figure modified from (Simmonds 2013)

Genotype distribution varies across the globe (Figure 1-6) and amongst risk groups. Genotype 1 is the most prevalent genotype, accounting for approximately 60% of infections world wide, with a high frequency of genotype 1b in Europe and 1a in the United States. Genotype 2 is mainly found in Japan, Western Africa, the Mediterranean region and in some parts of America (Candotti et al. 2003), whereas genotype 3 is mostly found in Southern Asia, and is also prevalent in intravenous drug user (IVDU) communities in Europe. Genotype 4 is highly diverse, and commonly found in central Africa, the Mediterranean region and especially Egypt, were a mass treatment campaign against Schistosomiasis using inadequately sterilized needles led to a wide spread of this genotype (L.-Z. Xu et al. 1994; Oni & T. J. Harrison 1996; Angelico et al. 1997; Sánchez-Quijano et al. 1997; Habib et al. 2001; Matera et al. 2002; Njouom et al. 2003). Genotype 5 seems to be mainly concentrated in central and southern Africa, but also has been observed in ethnically diverse regions such as the Benelux countries (Smuts & Kannemeyer 1995). HCV genotype 6 is found mainly in South East Asia, is very diverse and seems to be related to HCV genotype 3 (Simmonds et al. 1996; Mellor et al. 1996; N. J. Adams et al. 1997). It is further suspected, although with very limited data, that HCV genotype 7 is concentrated in Central and Southern Africa (one single isolate: QC69, EF108306). In this thesis, HCV genotype 7 is not further discussed, since only one single sequence of this

genotype was reported, which was isolated from an emigrant from the Congo (D. B. Smith et al. 2013).

In Europe, genotypes 1a and 1b are the most commonly identified types, with a high frequency of genotype 3a within intravenous drug user communities (IVDU), and over 90% of patients are either infected with HCV genotype 1 or 3 in the UK (R. J. Harris et al. 2012; Health Protection Agency 2013).

Figure 1-6: Global distribution of HCV genotypes.

Figure deleted due to copyright:
Figure material can be found in Messina et al., 2014. Figure included in thesis has been derived from information shown in Figure 1 in this paper (permission obtained).

Global distribution of 6 major HCV genotypes is depicted. Size of pie charts is proportional to the number of seroprevalent cases as estimated by (Mohd Hanafiah et al. 2013). Image obtained from: Messina et al., 2014, in submission .

Within the HCV genome, different regions have different levels of sequence homology. The 5' un-transcribed region (UTR), has a high degree of sequence conservation and therefore often used for genotyping in clinical assays in combination with NS5B (Germer et al. 1999). Other regions vary in their degree of variability, with the envelope proteins and NS2 being highly variable (Simmonds et al. 2005; Humphreys et al. 2009) and non-structural proteins NS3 to NS5 comparatively conserved.

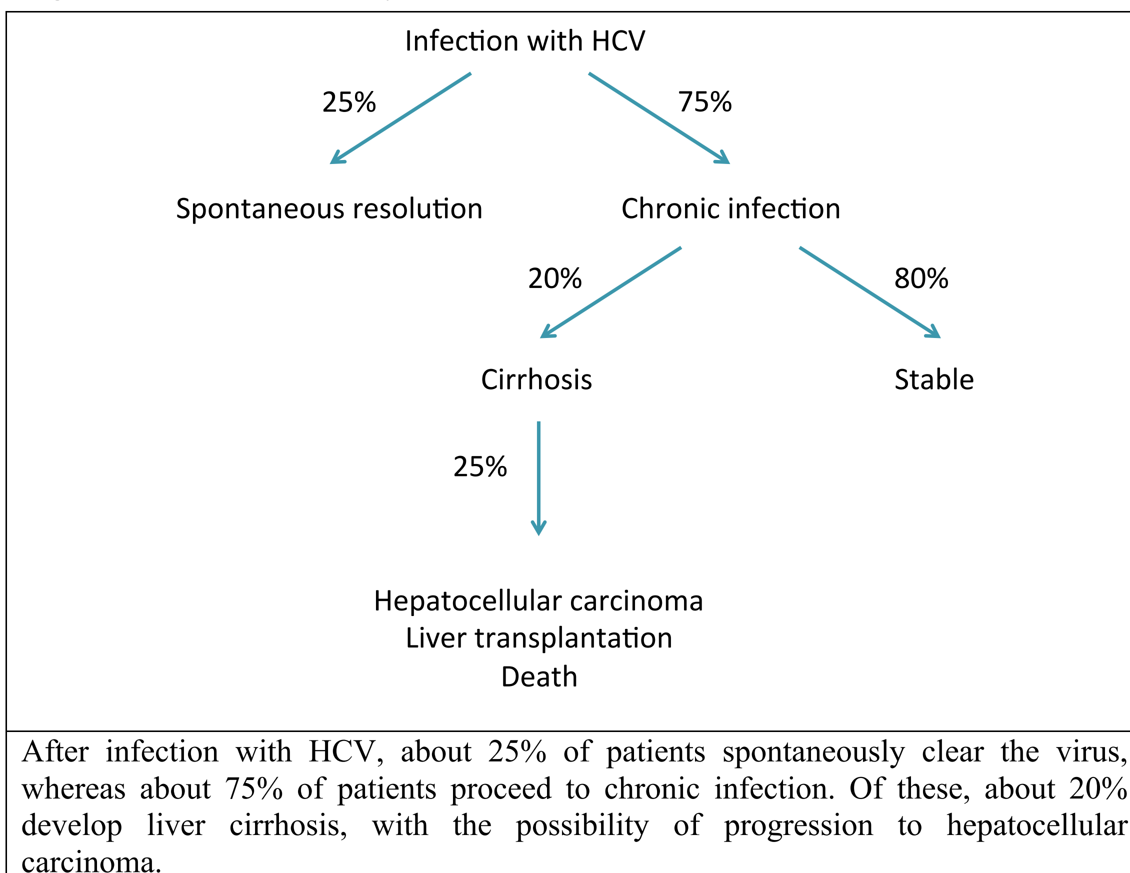
1.3 Natural history of HCV infection

1.3.1 HCV transmission, symptoms and outcome

HCV is transmitted through direct exposure to infectious blood. In developed countries, the virus is most commonly spread through injecting drug use. In the UK, about 40% of intravenous drug users (IVDU) are infected with HCV, with about half of them becoming infected within the first 3 years of injecting drug use (R. J. Harris

et al. 2012). Historically, the receipt of contaminated blood products accounted for a significant amount of infections, causing single strain outbreaks of the virus after injection of anti-D immunoglobulins in Ireland (J. P. Power et al. 1995; Casino et al. 1999) and Eastern Germany (Dittmann et al. 1991; Höhne et al. 1994). Injections with contaminated needles have caused another mass outbreak of HCV gt4 in Egypt, where as part of an extensive anti-schistosomiasis treatment campaign, potassium antimony tartarate was administered as a series of intravenous injections, using syringes and needles that were reused without being properly sterilized (Angelico et al. 1997; Frank et al. 2000; Habib et al. 2001; Pybus et al. 2003; Y. Tanaka et al. 2004). Other possible risk factors include transplantations of infected organs and needle prick injuries. Sexual transmission is not considered a major risk factor apart from a subset of HIV positive men who have sex with men (Thomson et al. 2011). The risk of mother-to-child transmission is estimated to be about 5%, with a higher risk in immunosuppressed women (mainly those infected with HIV), and a lower risk in women with low HCV RNA levels (European Paediatric Hepatitis C Virus Network 2005; Indolfi et al. 2006; Canadian Paediatric Society 2008).

Figure 1-7: Natural history of HCV infection.



After infection with HCV, about 25% of patients spontaneously clear the virus, whereas about 75% of patients proceed to chronic infection. Of these, about 20% develop liver cirrhosis, with the possibility of progression to hepatocellular carcinoma.

The clinical course of disease varies widely amongst infected individuals. While about 25% of patients with acute Hepatitis C infection clear the virus spontaneously, the majority of patients become chronically infected (Figure 1-7).
During acute infection, the majority of infected individuals are asymptomatic or have mild or unspecific symptoms (A. E. Mitchell et al. 2010), and therefore remain undiagnosed. However, some individuals develop an onset of acute hepatitis with jaundice, malaise and nausea, usually about 7 to 8 weeks after infection (Lauer & B D Walker 2001).

1.3.1.1 Chronic HCV infection

The majority of patients infected with HCV develop chronic infection. The clinical burden of HCV infection is associated with morbidity caused by chronic HCV infection, with major clinical problems like liver cirrhosis, liver failure and hepatocellular cancer. Of those who become chronically infected (about 75%), about 20% develop liver cirrhosis, with a risk of hepatocellular carcinoma of 1-4% per year (Seeff 2002). HCV is the leading cause for liver cirrhosis, and the primary indication for liver transplantation (G. L. Davis et al. 2003; Biggins et al. 2012). Chronic infection with HCV is also associated with an increased likelihood for the development of membranoproliferative glomerulonephritis (7-fold increase), cryoglobulinemia (11-fold increase), and skin disease such as porphyria cutanea tarda (12-fold increase) and lichen planus (2-fold increase) (El-Serag et al. 2002). Estimates indicate that about 350 000 deaths world wide occur each year due to all HCV related causes (Perz et al. 2006).
Morbidity and mortality of liver disease due to chronic HCV infection is expected to rise considerable over the coming decades, when individuals infected during the 1980 and 1990's develop chronic liver disease. In the UK, hospital admissions and deaths due to HCV have more than tripled between 1996 and 2011, even with antiviral therapy available during that time (Health Protection Agency 2013).

1.3.1.2 Spontaneous clearance of HCV infection

Spontaneous clearance is defined as clearance of HCV infection within the first 6 months of infection without further treatment. In a meta-analysis of 31 published reports including 675 study objects, the rate of spontaneous clearance was estimated at 26% (Micallef et al. 2006). HCV clearance has been associated with several host

and viral factors, including gender, age, ethnicity, gene polymorphisms, behavioural factors, route of infection and viral genotype and variability.

An association with gender has been reported separately in various studies, with about 40% of females clearing the infection compared to 19% of males (Yamakawa et al. 1996; Kenny-Walsh 1999; Inoue et al. 2000; Alric et al. 2000). The strong association of ethnicity with spontaneous HCV clearance (M. P. Busch et al. 2006) was recently explained through an association of IL28B polymorphisms with both spontaneous clearance and treatment outcome (discussed further under 1.4.5 Polymorphisms in the innate immune genes are associated with HCV infection and treatment outcome, page 42). Rates of spontaneous clearance in different ethnic groups reflect the distribution of favourable and un-favourable IL28B gene alleles; with high rates of clearance and occurrence of favourable IL28B alleles in Asians, low rates of clearance and frequent occurrence of the un-favourable IL28B allele in African-Americans and intermittent rates in Caucasians (Ge et al. 2009; D L Thomas et al. 2009). This association has been validated in a recent genome wide association study assessing HCV spontaneous clearance using data from multiple cohorts; in this study, IL28B status and HLA class-II were factors independently associated with spontaneous clearance (Duggal et al. 2013). In addition, spontaneous clearance of HCV infection has been associated with a symptomatic onset of disease, which is thought to be a result of a vigorous immune response (Barrett et al. 1999; Thursz et al. 1999; Thimme et al. 2001; Larghi et al. 2002; Gerlach et al. 2003; Hofer et al. 2003; Wawrzynowicz-Syczewska et al. 2004).

Behavioural factors associated with viral clearance were low alcohol intake and long term monogamy (Piasecki et al. 2004; Wawrzynowicz-Syczewska et al. 2004). However, the duration of monogamy could be a surrogate marker for other reduced risky behaviors and exposures.

Other factors associated with viral clearance where published evidence is inconclusive include age and route of infection. High spontaneous resolution rates have been observed in children (75-100%) (Meky et al. 2006), and a younger age of infection has been associated with clearance in some studies (Messick et al. 2001; Eyster et al. 2004). However, several other groups have shown a beneficial effect of older age at infection (Wawrzynowicz-Syczewska et al. 2004), or that clearance is independent of age (A. L. Cox, Mosbruger, Mao, et al. 2005; S Keating et al. 2005).

Viral factors associated with spontaneous resolution of infection are the infecting viral genotype and the variability of the virus. Higher rates of spontaneous clearance have been described for HCV gt3 in comparison to HCV gt1 infection (M. Lehmann et al. 2004; Bortolotti et al. 2005). Lehmann et al. determined the viral genotype in young male IVDU positive for HCV antibodies, finding that HCV gt3 specific antibodies were more prevalent among spontaneous resolvers than in chronically infected patients (23% overall; 86% HCV gt3 vs. 7% HCV gt1) (M. Lehmann et al. 2004). Similar results were reported in an Italian paediatric study where 22% of HCV gt3 patients had self-limited disease but less than 3% of the genotype 1 patients. In addition to the association with genotype, high viral mutation rates in hypervariable regions of viral envelope genes were predictive of subsequent development of chronicity (P. Farci et al. 2000; Laskus et al. 2004).

1.3.2 Treatment of HCV infection

The aim of HCV treatment is the clearance of infection to prevent the progression of liver disease and subsequent complications. Successful treatment is defined by measuring undetectable viral loads in patient plasma 6 months after the end of treatment. This is referred to as sustained virological responses (SVR). An additional aim of treatment is the prevention of viral transmission to others.

1.3.2.1 Interferon/Ribavirin

Until recently, the standard of care for HCV treatment was a combination of pegylated interferon alpha (IFNα) injected subcutaneously and oral ribavirin (IFN/RIB). The outcome of treatment is defined as either successfully clearing the virus (SVR), or treatment failure through virological relapse or non-response. The timeline of viral clearance in response to treatment has been associated with treatment success: a decrease in viral load greater than 2 logs by treatment week 4 is defined as rapid virological response (RVR), and has been associated with successful viral clearance. Early virological response (EVR) is defined as an undetectable viral load after 12 weeks of treatment. If the virus is not cleared by treatment week 24 (in gt1 infected individuals) or week 16 in HCV gt3 infected individuals, treatment response is unlikely and treatment is discontinued.

Similarly to HCV spontaneous resolution, treatment response success is associated with several host and viral factors (Table 1-1).

Table 1-1: Predictors of treatment response

Viral factors	Host factors	
	Fixed	**Modifiable**
HCV genotype	Age	Body mass index
Baseline HCV RNA levels	Sex	Insulin resistance
Viral variability	Race	Coffee consumption
Specific viral polymorphisms associated with treatment response (ISDR, PePHD)	Fibrosis	Vitamin D status
	IL28B genotype	
	ISG pre-activation	

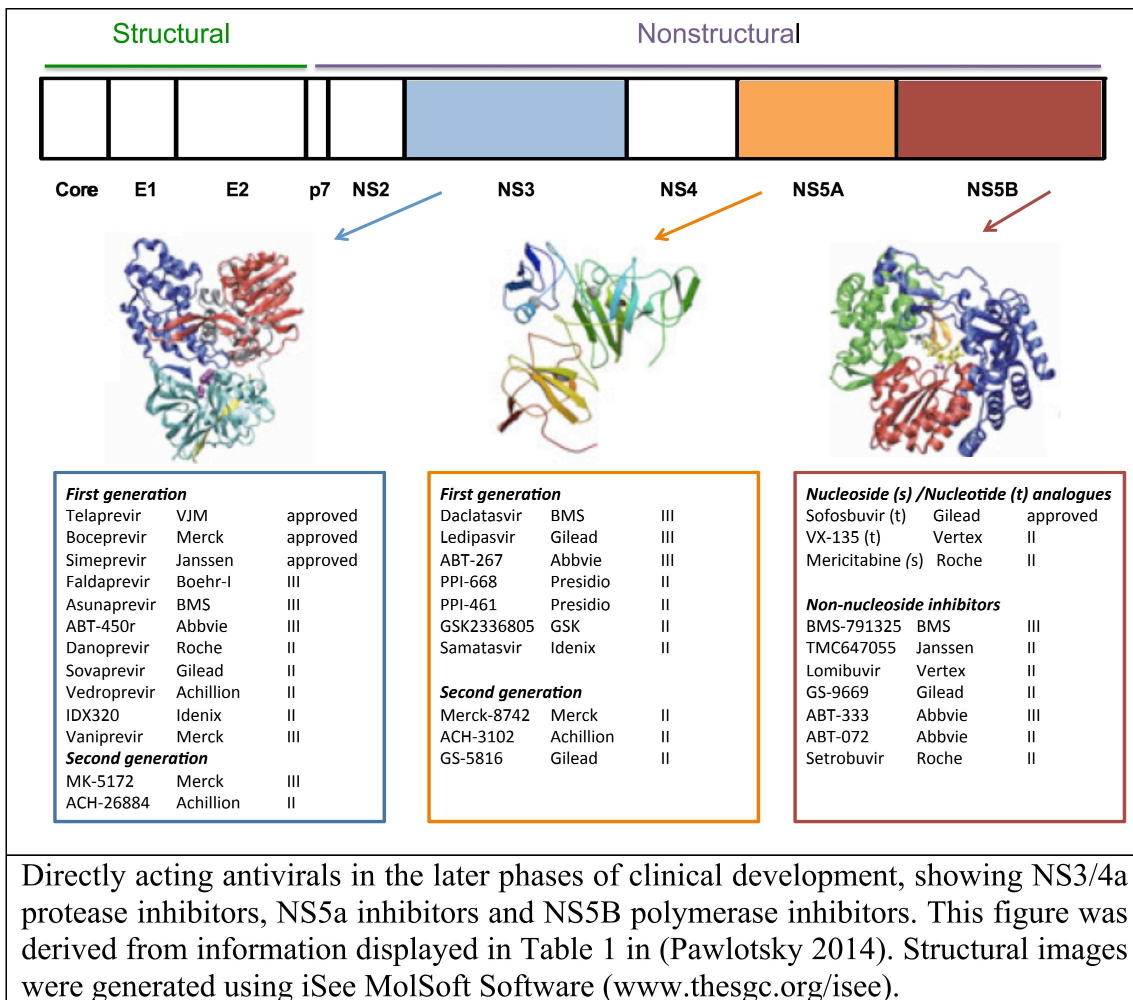
More recently, Interferon-λ (IFNλ) has been administered instead of IFNα in clinical trials (Zeuzem et al. 2012; BMS press release 2013) . The rationale of these trials has been based on basic research linking polymorphisms close to the interferon-λ gene to spontaneous resolution and treatment response (discussed in detail under 1.4.5 Polymorphisms in the innate immune genes are associated with HCV infection and treatment outcome, page 42). IFNλ has a more favourable side effect profile in comparison to IFNα, which is thought to be linked to restricted receptor distribution (Sommereyns et al. 2008; K. Witte et al. 2009).

Unfortunately, the standard of care treatment with a combination of pegylated-IFNα and Ribavirin is long (24 - 48 weeks), expensive (approximately £11,000/ therapy), has strong side effects and is not effective in 10-40% of patients, depending on the infecting viral genotype (NICE Guidelines 2014), which made new therapies for HCV a pressing need.

1.3.2.2 Directly acting antivirals (DAAs)

A major advance in HCV treatment was facilitated through structure based drug design, which enabled the discovery of a new drug class in the treatment of HCV, directly acting antivirals (DAA). Structurally derived inhibitors were designed for NS3 (Protease inhibitors), NS5A (NS5A inhibitors) and NS5B (Polymerase inhibitors) (Figure 1-8). The clinical development of DAAs is currently an extremely fast moving field, with multiple pharmaceutical companies testing several agents inhibiting non-strucural proteins NS3, NS5a and NS5b in different phases of clinical development (recently reviewed by (Schinazi et al. 2014; Pawlotsky 2014; Wendt et al. 2014), Figure 1-8).

Figure 1-8: Directly acting antivirals: Protease Inhibitors, NS5A Inhibitors and Polymerase Inhibitors.



Directly acting antivirals in the later phases of clinical development, showing NS3/4a protease inhibitors, NS5a inhibitors and NS5B polymerase inhibitors. This figure was derived from information displayed in Table 1 in (Pawlotsky 2014). Structural images were generated using iSee MolSoft Software (www.thesgc.org/isee).

The first DAAs available approved for HCV gt1 infected patients were the protease inhibitors Telaprevir and Boceprevir, now accompanied by Simeprevir, with multiple other first generation NS3 inhibitors in phase II or phase III clinical trials (reviewed by Jean-Michel Pawlotsky 2014). All first generation protease inhibitors show limited pan-genotypic efficiency and frequently induce resistance mutations (Kuntzen et al. 2008; G. R. Foster et al. 2011; B. R. Bacon et al. 2011; Poordad et al. 2011; I. M. Jacobson et al. 2011) (Table 1-2). They have to be administered every 8 to 12 hours. Second generation NS3/4a protease inhibitors (MK-5172, ACH 26884) have a better pan-genotypic coverage and induce less adverse events. They also induce fewer resistance mutations, although some mutations seen with the second generation inhibitors have been associated with viral breakthrough (Summa et al. 2012).

Inhibitors of the RNA dependent polymerase NS5B are split into nucleoside/nucleotide analogue inhibitors that act as false polymerase substrates, and

non-nucleoside inhibitors binding to the NS5B surface and thereby altering the NS5B active site (Figure 1-8). Because of their mechanism of action, nucleoside/nucleotide inhibitors have a favourable profile with pan-genotypic coverage and have the highest barrier to resistance of all DAAs (Gane et al. 2010; Membreno & Lawitz 2011; Gane et al. 2013; Lawitz et al. 2013). They depend on 3 (nucleoside) or 2 (nucleotide) phosphorylations for activation, which explains faster action of nucleotide analogues. The nucleotide inhibitor Sofosbuvir is now approved, and is part of most interferon/DAA and non-interferon/DAA combinations in the pipeline for all HCV genotypes (reviewed by (Pawlotsky 2014; Wendt et al. 2014) Table 1-2).

First generation non-nucleoside polymerase inhibitors act primarily against HCV gt1 and have a low barrier to resistance; however, broader drugs which are less prone to mutations are in pre-clinical development (Larrey et al. 2012).

NS5a inhibitors block viral RNA replication and virion assembly by interaction with the replication complex, however, the exact mechanism of action still remains unclear (Guedj et al. 2013). First generation NS5a inhibitors, with 7 drugs in phase II and III clinical development, have a low barrier to resistance mutations, and are only effective against genotype 1 and 4 (Figure 1-8 and Table 1-2) (M. Gao et al. 2010; M. Huang et al. 2011; R. Colonno et al. 2011; J. Walker et al. 2014; Nettles et al. 2011; Link et al. 2014; DeGoey et al. 2014). Second generation NS5a inhibitors seem to have a higher barrier to resistance, and show a broader coverage of genotypes, however, some are still show less pan-genotypic efficiency against genotypes 2 and 3 (G. Yang et al. 2012; Coburn et al. 2013).

Table 1-2: Drug profiles of directly acting antivirals

DAAs	NS3 inhibitors		NS5a inhibitors		NS5B inhibitors	
	1st Generation	2nd Generation	1st Generation	2nd Generation	1st Generation	2nd Generation
Efficiency						
Resistance Profile						
Pangenotypic efficiancy						
Adverse events						
Drug-drug Interactions						

Clinical profile of first and second generation directly acting antivirals (DAAs). Green: good profile, yellow: average profile, red: non-favourable profile. Table adapted from (Schinazi et al. 2014).

Apart from the structure-based antivirals, another drug class currently exploited are host-targeting antivirals that target cell components of the HCV life cycle. Examples are cyclophilin inhibitors (Alispovir [Novartis], SCY-635 [Scynexis]) and antagonists of microRNA 122 (Miravirsen [Santaris]) (Lanford et al. 2010; S. Hopkins et al.

2010; Pawlotsky et al. 2012; Quarato et al. 2012; Janssen et al. 2013). Alisporivir is an analogue of cyclosporine A and represents the prototype of a new class of non-immunosuppressive cyclophilin inhibitors (Quarato et al. 2012).

When used as a monotherapy, DAAs cause a rapid emergence of escape mutations (Hezode et al. 2009; McHutchison et al. 2010). Therefore, most drugs will be used as combination treatments. Several combinations of new DAAs, such as triple combinations of a NS3/4a inhibitor, a NS5a inhibitor and a non-nucleoside NS5b inhibitor, currently trialled in clinic with promising results (Abbvie press releases 2013; Abbvie press releases 2014). Double combinations of drugs with a low barrier to resistance, such as NS3/4a inhibitor Asunaprevir and NS5a inhibitor Daclatasvir, only yield high SVR rates in easy-to-cure patient groups (Chayama et al. 2012). Several combination therapies include the potent nucleotide polymerase inhibitor Sofosbuvir, which is now licenced (Gilead press releases 2013; Gilead press releases 2014b).

Interferon will stay a backbone of treatment combinations until new IFN-free combinations are approved, of which several are currently in clinical development (Gane et al. 2010; Zeuzem et al. 2011; Chayama et al. 2012). IFN is then expected to disappear from HCV treatment guidelines - at least in countries that can afford the new expensive drugs: added costs for Boceprevir and Telaprevir are £8,400 to £22,398 per treatment in addition to the costs for interferon (NICE guidelines for Boceprevir and Telaprevir 2013). However, administration of both Telaprevir and Boceprevir has been described as cost effective in terms of life years gained (LYG) and quality adjusted life years (QALY), in both previously treated and untreated HCV gt1 patients (Cammà et al. 2012; Cammà et al. 2013). Ribavirin will still be used in combination with the DAAs, since it prevents viral relapses by an unknown mechanism.

Overall, DAAs have caused a significant improvement in HCV therapy by significantly increasing the treatment response rate in HCV infected patients without previous therapy, but also in ‘difficult-to-treat’ groups like prior relapsers and null responders. However, even with an arsenal of effective drugs available, most HCV infected patients worldwide will not have access to the expensive new therapies, making the development of an efficient vaccine a pressing need. In the following section, I will discuss basic knowledge on innate and adaptive immunity in response to viral infections, and in particular to HCV infection.

1.4 Innate immune response to HCV infection

The innate immune response targets HCV as the first line of defence, attacks the virus in a non-specific manner and does not confer long lasting immunity to the host (Janeway et al. 2001). Innate immune responses are activated shortly after HCV infects the cell, and have been repeatedly shown to play an important role in HCV clearance. The role of the interferon system and NK cells in HCV infection are discussed below.

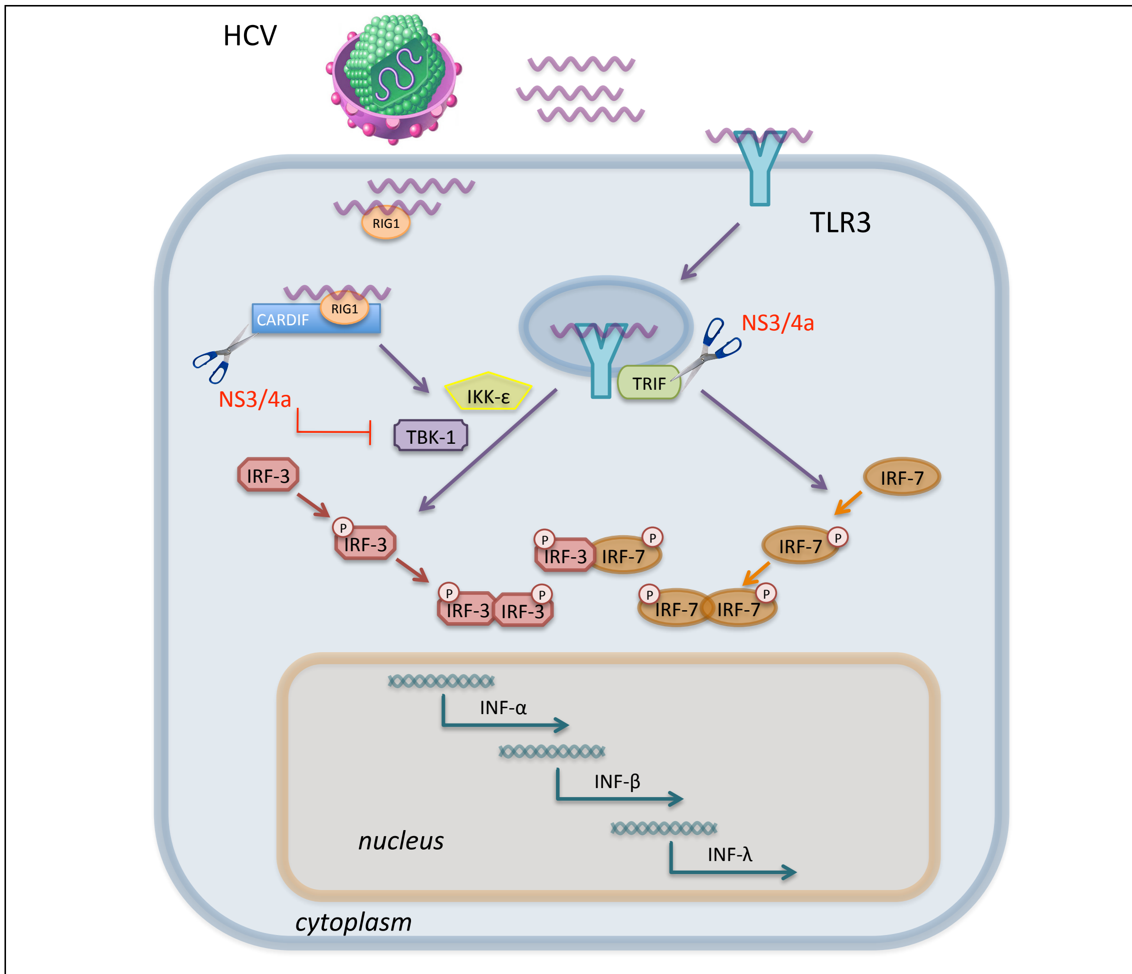
1.4.1 Interferon production and interferon signalling

The precise mechanisms of induction of interferon stimulated genes in response to viral infections continue to evolve (Stone et al. 2013; Rosen 2013); current understanding is that the virus is recognized by the host cells via two pathways, resulting in the induction of type I and type III interferons.

Cytosolic pathway: Followed by the binding to cell surface proteins and incorporation of the viral particle, HCV translation and replication takes place in the cytosol. The cell recognizes the virus by noticing pathogen associated molecular patterns (PAMPs) in the HCV 5' and 3' UTR through pattern recognition receptors (PRRs), such as retinoic acid inducible gene-I (RIG-I) and melanoma differentiation antigen 5 (MDA5). Viral RNA binds to RIG-I (retinoic acid-inducible gene-I), which directs a stable interaction and then activates the interferon signaling pathway through CARDIF (CARD adaptor inducing Interferon-β (IFNβ), also known as MAVS, IPS-I or VISA) (Saito et al. 2008). Activation of NFκB and the interferon regulatory factor 3 (IRF-3) and translocation of these factors into the nucleus result in increased transcription of interferons (Dragan et al. 2007) (Figure 1-9). In addition, dsRNA sensor protein kinase R (PKR) binds to the HCV internal ribosomal entry site (IRES) as early as 2 hours after infection and prior to the interaction with RIG-I, triggering early induction of several IRF-3 dependent genes (Arnaud et al. 2010; Arnaud et al. 2011).

TLR pathway: In addition to the cytosolic pathway, the toll-like receptor 3 (TLR-3) recognizes extracellular double-stranded RNA (dsRNA) and re-localizes to an endosome, where it activates TRIF (Toll/IL-1 receptor domain-containing adaptor inducing IFNβ). Downstream signalling pathways ultimately lead to the transcription of interferon coding genes (Figure 1-10).

Figure 1-9: Proposed model of CARDIF and TRIF interference by the NS3/4 viral protease.



Interferons are induced via two main pathways in response to HCV infection. Left: Cytosolic pathway. After HCV enters the cell via specific receptors, RNA is detected by pattern recognition receptors such as RIG-I, which binds to CARDIF (or MAVS). Right: Endosomal pathway. TLR3 detects double stranded RNA, initiating downstream signaling via TRIF and other kinases. Both pathways lead to activation of downstream pathways (IRF-3) and ultimately transcription of interferons.
HCV interferes with the induction of interferons. Cellular proteins CARDIF (MAVS) and TRIF are cleaved by the HCV NS3-4A protease (depicted as scissors). TBK1 is also inhibited by NS3/4a (Jin 2007), rendering IRF-3 capability to translocate to the nucleus. Additionally, HCV inhibits the translation of transcribed IFNs through activation of PKR (not depicted).
CARDIF: CARD adaptor inducing IFNβ (also mitochondrial antiviral signaling protein MAVS), IKK-ε: inhibitor of NFk-B kinase e, IRF: *interferon* regulatory transcription factor, RIG: Retinoic acid-inducible gene I, STAT: Signal transducer and activator of transcription protein, TLR: Toll like receptor, TRIF: Toll/IL-1 receptor domain-containing adaptor inducing IFNβ.

There are three classes of Interferons (IFNs): Type I IFNs consist of IFNα and β, and other suptypes (κ, δ, ε, ρ, π, ω, ξ), acting through a common receptor, the IFNα/β

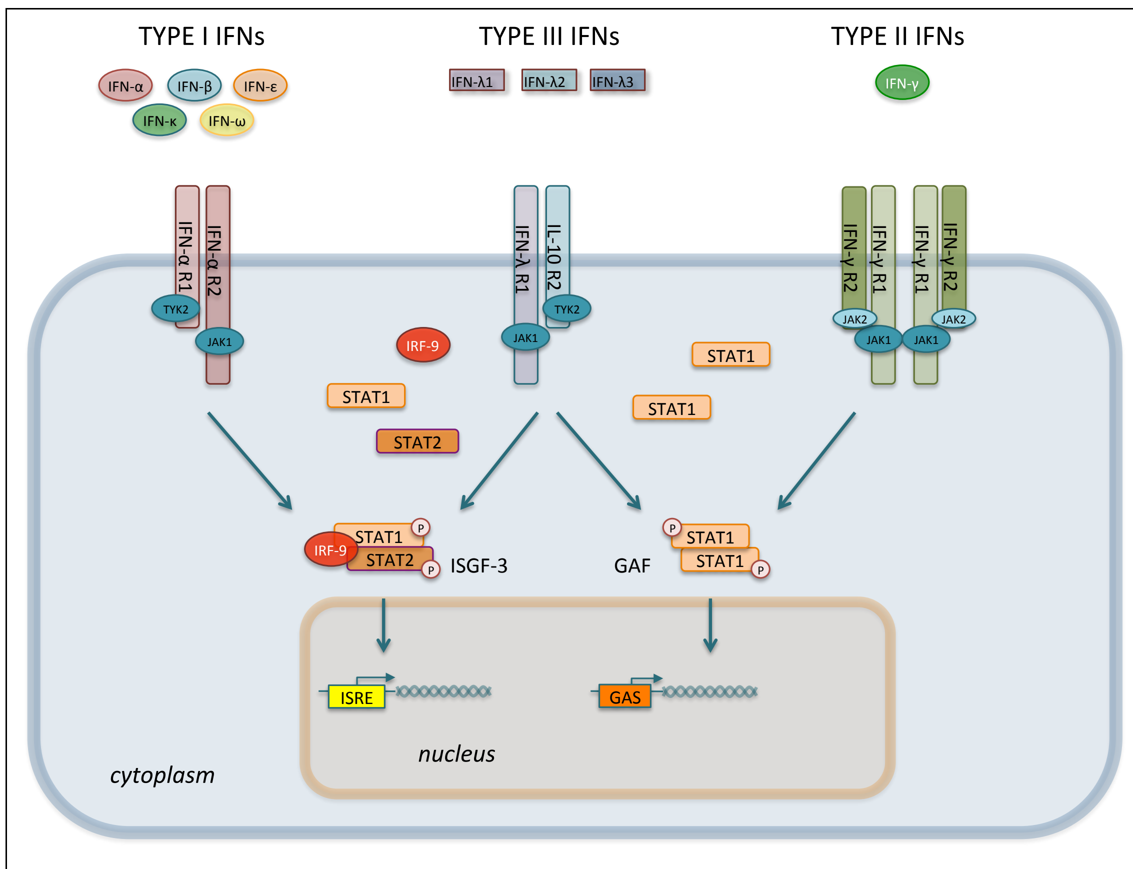
receptor. Type II IFNs consist of a single cytokine, IFNγ, essential for cell-mediated immunity by activating macrophages and promoting the development of CD4+ and CD8+ T cells (Gattoni et al. 2006). Type 3 IFNs include 4 members, IFN λ1, λ2, λ3, and IFN-λ4. IFN λ1, λ2, and λ3 are also known as IL29, IL28A, IL28B (Ank et al. 2006; Ank & Paludan 2009; Prokunina-Olsson et al. 2013).

The different types of IFNs are produced by different cell types: IFNα can be produced by virtually all types of nucleated human cells (Biron 1999). As part of the innate immune response, IFNγ is predominantly produced by natural killer (NK) and natural killer T (NKT) cells. Once antigen-specific immunity develops, CD4+ and CD8+ cytotoxic T lymphocyte (CTL) effector cells produce IFNγ (Schoenborn & C. B. Wilson 2007). IFNλ is mainly produced by plasmacytoid dentritic cells, however, other cell types like macrophages, monocyte derived dentritic cells and hepatocytes have been shown to produce this cytokine (Coccia et al. 2004; Wolk et al. 2008; Sommereyns et al. 2008; Yin et al. 2012).

Subsequently, IFNs bind to their receptors (Figure 1-10), and induce interferon stimulated gene (ISG) transcription via the JAK/STAT pathway. Type I interferons signal through a receptor composed of low- (IFNα-R1) and high-affinity (IFNα-R2) components. Type III interferons signal through a high-affinity receptor (IFNλ-R1) and an receptor (IL10-R2), which is shared by IL-10, IL-22 and IL-26 (Kotenko et al. 2003; Sheppard et al. 2003). In contrast, type II IFNs signal through a complex composed of four transmembrane receptors, two chains of a high-affinity receptor (IFNγ-R1) and two chains of a low-affinity receptor (IFNγ-R2) (Pestka et al. 2007).

After binding to their receptors, there is a great deal of redundancy in downstream signalling pathways activated (JAK/STAT), particularly seen for Type I and III IFNs. However, through distribution of different receptors on different cell types, the effects of IFN types are restricted (reviewed by (De Weerd & T. Nguyen 2012)). Receptors for the type I and II IFNs, and the IL10RB involved in type III signaling are widely distributed and found on the surface of most cell types. In contrast, the distribution of the high-affinity receptor for type III IFNs, IFNλ-R1, is more restricted. It can be found on cells of epithelial origin (Ank et al. 2008; Donnelly & Kotenko 2010), in particular keratinocytes, and mucosal cells in the gastrointestinal tract, lungs and kidneys (Sommereyns et al. 2008).

Figure 1-10: Binding of different types of Interferons to their receptors and following intracellular pathways.



Interferons bind to their specific receptor, activating downstream JAK/STAT signaling pathways. STAT proteins form phosphorylated homo- or heterodimers and translocate to the nucleus, to activate gene regulatory elements (ISRE and GAS), ultimately leading to transcription of IFN stimulated genes.
GAF/GAS: gamma-*Interferon* activation factor/site, IFN: Interferon, IRF: *Interferon* regulatory transcript, IRSE: Interferon-stimulated response element, ISGF: *Interferon stimulated* gene factor e, JAK: Janus kinase, STAT: Signal transducer and activator of transcription protein, TYK: tyrosine kinase.

1.4.2 Immunomodulatory and antiviral effects of interferons

Interferons have important immunomodulatory and antiviral effects at different levels of the immune system, including antigen presentation, T cell differentiation, and modulation of innate immunity, of which some are discussed in the following paragraphs.

Immunomodulatory effects

Type I IFNs have an important role in the differentiation of both CD4+ and CD8+ T cells. IFNs enhance proliferation and maturation of dendritic cells, ensuring efficient antigenic peptide presentation (Montoya et al. 2002). Additionally, Type I IFNs

facilitate cross-presentation in dendritic cells (Diamond et al. 2011; Schiavoni et al. 2013), and trigger the production of IL-15 by antigen presenting cells, which is responsible for stimulating the generation of memory T cells (Boyman et al. 2007). Class I IFNs up-regulate antigen processing and presentation on both MHC class I and class II molecules (Y. Yang et al. 1995), and subsequently sustain the proliferation of antigen specific CD8+ cells (Kolumam et al. 2005; K. Honda et al. 2005). Type I IFNs also drive T_H1 cell development by inducing IL-12, and the phosphorylation and activation of STAT4 (S. S. Cho et al. 1996; Szabo et al. 2002). In addition, type I IFNs synergize with other cytokines like IL-17 and IL-21 (S Matikainen et al. 2001; Strengell et al. 2004). Type I IFNs have been associated with the suppression of T_H2 and T_H17 responses (Moschen et al. 2008; J. P. Huber & Farrar 2011). Therefore, it seems that IFNα/β ensures the proper differentiation of T_H1 cells by restricting the development of alternative subsets (J. P. Huber & Farrar 2011). Other immunomodulatory effects of IFNs include the activation of NK cells through complex processes, leading to upregulation of perforin and granzymes (Bolitho et al. 2007).

Antiviral effects of IFNs.

The interferon system induces extremely powerful antiviral responses, capable of controlling most viral infections in the absence of adaptive immunity (R. E. Randall & Goodbourn 2008). Treatment of cells with IFNα/β induces an antiviral state via the induction of several hundred genes. An example is a molecule that acts as a direct antiviral, the interferon stimulated gene 15 (ISG15) (Lenschow et al. 2005), which activates a large number of proteins involved in positive feedback mechanisms for the interferon systems itself and other antiviral effector proteins. Overall, interferon induced genes encode proteins with antiviral activity that have been extensively studied, such as the dsRNA dependent protein kinase R (PKR) (Clemens 2005; Silverman 2007), 2’5’-oligoadenylate synthetase (OAS), the Mx family (Martens & Howard 2006; Haller et al. 2007) viperin and the RNAseL, which restricts HCV replication within the hepatocyte.

A major antiviral component of IFNs is the up-regulation of genes involved in antigen processing and presentation, particularly the expression of MHC class I genes and associated endocytic proteins (Le Bon et al. 2003; Le Bon et al. 2006). In addition, IFNα/β enhance CD8+ expansion and their cytolytic activity.

In HCV infection, higher hepatic ISG expression is predictive of a non-response to IFNa/Ribavirin therapy (Sarasin-Filipowicz et al. 2008; Sarasin-Filipowicz et al. 2009; Afdhal et al. 2011). A proposed mechanism to explain this somewhat paradoxical finding, is that type III IFNs, which are uniquely upregulated in hepatocytes in response to HCV infection (E. Thomas et al. 2012), may generate a refractory state to further type I IFN action in hepatocytes. This concept was supported by the finding that blocking type III IFNs enhances the antiviral activity of IFNa in cell culture (E. Thomas et al. 2012).

1.4.3 NK cells in HCV infection

NK cells belong to the innate immune system and form about 5%-15% of peripheral blood mononuclear cells (PBMC). They are enriched in the liver (30%-50%) (Doherty & O'Farrelly 2000), where they form an important part of the innate immune response against hepatotropic viruses (Golden-Mason & Rosen 2006). They are cytolytic and release cytokines, and can directly induce apoptosis of virus-infected hepatocytes (Tay & Welsh 1997; Guidotti et al. 1999; Z. X. Liu et al. 2000). Unlike T and B cells, NK cells can lyse without the need to recognize a major histocompatibility complex (MHC), and can be activated by cytokine secretion without priming (Cheent & Khakoo 2009). Effector mechanisms of NK cells include direct cytotoxicity through cytotoxic granules or induction of apoptosis through interaction of FAS and tumour necrosis factor related apoptosis inducing ligand (TRAIL) with upregulated receptors on target cells; and indirect mechanisms like dendritic cell activation through IFNγ and TNFα, which then activate T cells (Z. X. Liu et al. 2000). NK cells are regulated by polymorphic inhibitory receptors (KIR).

NK cells seem to be involved in HCV clearance, either directly by targeting infected hepatocytes, or indirectly through regulating other cell types. Evidence for their role in HCV infection arises from a study showing that the genetic combination of a specific KIR molecule on NK cells (KIR2DL3) and HLA class C is associated with spontaneous and treatment-induced resolution of HCV infection (Khakoo et al. 2004). In addition, expression of TRAIL on NK cells has been associated with the control of HCV infection (Stegmann et al. 2010).

1.4.4 HCV subverts the innate immune response

HCV has developed multiple mechanisms of subverting the cell's innate immune response (reviewed by (Rehermann 2009; Rosen 2013)).

HCV inhibits the action of IFNs at multiple levels, interfering with the induction, actions of IFNs, and subsequent functions of interferon stimulated genes.

Via the NS3/4a protease HCV has a strong influence on the induction of IFNs (Jouan et al. 2010). The protease cleaves CARDIF and TRIF (Figure 1-9) (Foy et al. 2003; M. Gale & Foy 2005; Foy et al. 2005), two components by which a cell triggers the induction IFNs (Rebsamen et al. 2008; Saito et al. 2008), thus inhibiting the expression of Interferon stimulated genes (ISG) (X.-D. Li et al. 2005). Additionally, it cleaves the IRF-3, preventing its translocation into the nucleus. This effect is reversible by use of protease inhibitors (Jouan et al. 2010).

Subsequently, HCV interferes with the actions and downstream signalling of IFNs. HCV core binds to STAT1, thereby inhibiting its activation and inducing its degradation (W. Lin et al. 2006). HCV also induces suppressor of cytokine signalling proteins (SOCS), in particular SOCS3, an inhibitor of the JAK/STAT pathway (Bode et al. 2003). The interaction of HCV with SOCS7 is particularly pronounced in HCV gt3 infection, where HCV core has been shown to interfere with STAT signalling in cell culture (Pazienza et al. 2010). In addition to the SOCS proteins, the protein phosphatase 2a (PP2A) is specifically induced by HCV infection; leading to an additional inhibition of STAT1 and resulting in reduced transcriptional activation of IFN stimulated genes (Duong et al. 2004).

Finally, HCV interferes directly with IFN induced proteins. These effects are primarily related to HCV protein NS5A, which has been shown to inhibit the 2′-5′ oligoadenylate synthetase (OAS), and induce IL-8, resulting in a decrease of overall ISG expression (S J Polyak et al. 2001). In addition, NS5A forms heterodimers with PKR, thereby decreasing its inhibitory effect on the translation of viral proteins (M. J. Gale Jr et al. 1997; Pflugheber et al. 2002). Another viral protein that interferes with PKR in a similar manner is E2. The two genomic regions are named PKR-eukaryotic transcription factor (elF2-alpha) phosphorylation homology domain (PePHD) in E2 (Puig-Basagoiti et al. 2001), and the IFN sensitivity-determining region (ISDR) in NS5A (Yokozaki et al. 2011), and both regions were associated with the treatment response to INFα (Sarrazin et al. 2000).

Another route by which HCV subverts the innate immune response is through NK cells. In HCV infection, NK cells overexpress inhibitory receptors and produce cytokines attenuating the adaptive immune response, such as IL-10 and TGF-β (Jinushi et al. 2004). The mechanisms by which this inhibition is induced are controversial: Initially, it was thought that HCV glycoprotein E2 inhibits NK cells directly by cross-linking CD81 (Crotta et al. 2002; Tseng & Klimpel 2002). However, these results could not be verified when E2 was part of infectious lipoviral particles. In addition to E2, HCV core encoded peptides are also thought to inhibit NK cell function (Nattermann et al. 2005).

1.4.5 Polymorphisms in the innate immune genes are associated with HCV infection and treatment outcome

A major breakthrough in HCV research was the finding that polymorphisms in the gene linked with Interferon lambda (IFNλ) type III (also known as Interleukin 28B [IL28B]) are determinants of clinical outcome in HCV. This association has been established in large genome wide association studies (Ge et al. 2009; Rauch et al. 2010) and candidate gene association studies (Mangia, A. J. Thompson, et al. 2010; McCarthy et al. 2010; Montes-Cano et al. 2010; Mosbruger et al. 2010; D L Thomas et al. 2009), establishing a key role in determining clinical outcome of HCV infection. Polymorphisms in the IL28B gene are associated with the outcome of primary infection (Rauch et al. 2010; D L Thomas et al. 2009) and the treatment success of IFNα/Ribavirin combination therapy (Suppiah et al. 2009; Y. Tanaka et al. 2009). There are two polymorphisms that have been studied extensively, rs12979860 and rs8099917, which are in partial linkage equilibrium with each other and located about 3kb upstream of the IL28B gene (Ge et al. 2009; Z. Jia et al. 2012). The favourable alleles are more frequently found in Caucasian and Asian populations compared to African-Americans (D L Thomas et al. 2009; Ge et al. 2009).

IFN-λ3 is a member of the IFN-λ family, composed of IFN-λ1 (IL-29), IFN-λ2 (IL28A) and IFN-λ3 (IL28B) and IFN-λ4 (Kotenko et al. 2003; Sheppard et al. 2003; Prokunina-Olsson et al. 2013). Although the IFN-λ family is closer to the IL-10-related cytokines in terms of gene structure, protein structure, and receptor usage, they display type I interferon-like anti-viral and cytostatic activities (K. Witte et al. 2009; K. Witte et al. 2010). Unlike IFN-α and β, the target cell population of the IFN-λ

family is restricted and mainly includes epithelial cells and hepatocytes, suggesting that the IFN-λ family of cytokines evolved specifically to protect the epithelium from viral invasions (Sommereyns et al. 2008). Indeed, IFN-λ3 has been shown to inhibit viral replication *in vitro* (Marcello et al. 2006; S. E. Doyle et al. 2006), as well as acting as an antiviral cytokine in humans: IFN-λ3 has been shown to exhibit comparable activity to IFNα in combination with Ribavirin and directly acting antivirals (Friborg et al. 2013).

IFN-λ was trialled in clinical studies and showed considerable activity and a favourable side effect profile in comparison to IFNα (Muir et al. 2010; Zeuzem et al. 2012), which is thought to be due to the restricted expression of the IFN-λ receptor (Sommereyns et al. 2008).

Attempts to reveal the basic mechanism leading to the association of the IFN-λ3 related polymorphisms with HCV clearance were largely unsuccessful (D. R. Booth & Jacob George 2013). One potential proposed mechanism was a differential expression of IFN-λ3 in hepatocytes, but several groups failed to show an effect *in vitro* and *in vivo* (T. J. Urban et al. 2010; M. Honda et al. 2010; Dill et al. 2011; Naggie et al. 2012). Two recent publications suggested possible alternative mechanisms of action: a di-nucleotide polymorphism causing a frameshift influencing the expression of Interferon-λ4 may mediate the effects on HCV spontaneous clearance (Prokunina-Olsson et al. 2013). The SNP lies between the IFNλ2 and IFNλ3 gene, and either creates (polymorphism ΔG) or aborts (TT) the expression of the newly described transcript, IFNλ4. In a repeated association study, IFNλ4 was more strongly related to spontaneous clearance in patients with African ancestry, but provided similar information in Asians and Caucasians. However, the exact influence of IFNλ4 on HCV spontaneous resolutions remains to be defined. A second publication (McFarland et al. 2014) describes a functional polymorphism (rs4803217) in the 3′ untranslated region (UTR) of IFN-λ3, that was associated with the clearance of HCV infection. A T-to-G substitution at this SNP resulted in increased expression of IFNλ through escape from both AU-rich element (ARE)-mediated decay and post-transcriptional regulation by HCV-induced miRNAs. The authors suggest that these pathways mediate a robust repression of the unfavorable *IFNL3* polymorphism, by which HCV attenuates the antiviral response.

1.5 Adaptive immune response to HCV infection

A coordinated adaptive immune response involving both T cell responses and antibodies is normally required for pathogen control. Since the major part of this work described T cell responses towards HCV infection, and the development of a immunogen potentially priming cross-reactive T cell responses, I first describe basic facts on antigen presentation to T cells, and then review T cell responses in HCV infection, including mechanisms of T cell failure in chronic infection, followed by a short overview on humoral responses against HCV. Finally, I discuss the particular impact of immunotolerant environment of the liver on adaptive immunity in HCV infection.

1.5.1 Antigenic peptide presentation - MHC class-I and -II processing

T cell epitopes are peptides that induce immune responses when bound to major histocompatibility complexes (MHC) and presented on the cell surface for recognition by T cells. Peptides can be presented on two major classes of MHC complexes: MHC class-I and class-II complexes (Figure 1-11).

Figure 1-11: Peptide loading of MHC class-I and -II complexes.

Figure deleted due to copyright:

Figure material can be found in Neefjes et al., 2011. Figure included in thesis has been derived from information shown in Figure 1 and 3 in this paper.

Peptide loading and presentation of MHC class-I and class-II complexes. Left: Peptide presentation to CD8+ restricted cells by MHC class-I complexes. Intracellular material is degraded by proteasomes. Resulting peptides are transported into the ER by transporter associated with antigen presentation (TAP), where loading onto MHC class-I complexes takes place. MHC class-I/peptide complexes are presented on the cell surface. Figure modified from (Neefjes et al. 2011).

MHC class-I complexes present intracellular material of cytosolic and nuclear origin to the T cell receptor (TCR) of $CD8^+$ T restricted cells (Figure 1-11, left). MHC class-I molecules are expressed on almost all nucleated cells, and primarily present intracellular material degraded by proteasomes. Proteasomes are large multi-protein complexes, with the housekeeping function of degrading damaged or improperly folded proteins, marked with an ubiquitin cap. Proteins are usually degraded at the end of their functional lives, but additionally, about 30-70% of newly produced cellular proteins are immediately degraded after synthesis (Reits et al. 2000; Schubert et al. 2000).

Proteasomes are capable of producing peptides of different length, and are able to adapt their composition and function in response to an activated innate immune system. In cells treated with IFN, more peptides with hydrophobic or basic residues are generated, which preferentially bind to MHC class-I complexes (K. Tanaka 1994; Cardozo & Kohanski 1998; Griffin et al. 1998). Interestingly, proteasomes can even produce new peptides through antigenic peptide splicing (Vigneron et al. 2004; Michaux et al. 2014).

Peptides generated by the proteasome are subsequently transported from the cytoplasm to the endoplasmatic reticulum (ER) via the transporter associated with antigen presentation (TAP). In the ER, peptides are loaded onto MHC class-I complexes, assembled from a heavy chain and a light chain called β_2m microglobulin (Bjorkman et al. 1987; Madden et al. 1992). The peptide is required for the stability of the MHC complex, the groove of which accommodates peptides of 8-9 amino acids (AA) in length (Rammensee et al. 1993). Before peptide loading, the MHC class-I complex is stabilized through ER chaperones, such as Tapasin, calreticulin and ERp57 (Sadasivan et al. 1996; Morrice & Powis 1998; Farmery et al. 2000). Tapasin also couples with the TAP, leading peptides transported into the ER to the MHC complex for loading (Peh et al. 2000). Fully assembled and stabilized MHC class-I/peptide complexes then leave the ER for presentation.

In contrast to MHC class-I complexes, MHC class-II complexes primarily present extracellular material incorporated via the endosomic route to the TCR of CD4+ restricted T cells (Figure 1-11, right). Expression of MHC class_II molecules is restricted to professional antigen-presenting cells (APCs), including dendritic cells (DCs), B cells and macrophages.

During the maturation of MHC class-II molecules in the ER, they are prevented from binding any intracellular molecules by the invariant chain (Ii) (Roche & Cresswell 1990; Teyton & P. A. Peterson 1992). In addition to pre-mature binding of peptides in the ER, structural motifs in the invariant chain also control the intracellular transport of the MHC class-II complexes, either directing the complex to remain in the ER, or via the Golgi to MHC class-II compartment vesicles in the cell periphery (either directly or via the cell membrane), or to a degradation pathway circumventing the Golgi (Lotteau et al. 1990). Here, in the early endosome, resident proteases cleave both endocytosed proteins and the invariant chain, which is degraded to a class II-associated invariant-chain peptide (CLIP). With the help of a dedicated chaperone, HLA-DM, CLIP is exchanged for the antigenic peptide. Class-II molecules have a binding groove that is open at both ends, allowing longer peptides (usually 12 to 25 amino acids) to bind (Chicz et al. 1992; J. H. Brown et al. 1993; Jardetzky et al. 1994). Loaded MHC class-II complexes are then transported to the plasma membrane for presentation to CD4+ restricted cells.

In addition to the pathways described above, certain antigen presenting cells have the ability to cross-present exogenous antigens onto MHC class-I molecules. This property is atypical, but essential for the initiation of immune responses to viruses that do not infect antigen presenting cells (Neefjes et al. 2011).

1.5.2 T cell responses in acute infection

It is generally accepted that adaptive immune responses play a central role in the outcome of HCV infection (Rehermann 2009; Thimme et al. 2012). Most likely, multiple components (B cells, T cells, regulatory T cells) of the adaptive immune system are involved in viral control and finally clearance of HCV.

T cell responses in acute HCV infection are of particular immunological interest, since during this time the hallmarks for the outcome of infection are set: the patient either clears the infection or proceeds to chronic infection. However, immune responses in the acute phase of HCV infection are particularly difficult to study, since >80% patients do not show any, or non-specific and mild clinical symptoms during the acute phase, and if, typically not before the rise of ALT levels 8-12 weeks after initial infection. Additionally, a large proportion of acute infections is due to intravenous drug use (IVDU), and therefore in a patient group which is difficult to study.

Therefore, much of the knowledge about the early phase of infection is derived from the chimpanzee model. Most data on acute HCV infection in humans comes from closely followed intravenous drug using cohorts, where T cell responses in acute infection were described in patients both clearing infection and developing chronic infection (Lauer et al. 2005; A. L. Cox, Mosbruger, Lauer, et al. 2005).

One of the key characteristics of HCV infection is the delayed adaptive immune response during acute infection. An increase in HCV titer is detectable early and an induction of ISGs happens soon; however, T cell responses are typically detectable only 5-9 weeks after infection (Thimme et al. 2001), coinciding with the onset of clinical hepatitis (Figure 1-12). Following the initial rise in proliferative capacity around 12 weeks after infection, the fate of T cell responses depends on the outcome of infection (spontaneous resolution or chronic infection), and is discussed in the following paragraphs.

Figure 1-12: Schematic of the clinical course and adaptive immune responses in acute HCV infection in self-limiting and chronically evolving HCV infection.

Figure deleted due to copyright:

Figure material can be found in Rehermann 2009. Figure included in thesis has been derived from information included in Figure 3 in this paper.

Acute hepatitis C followed by (A) spontaneous resolution of infection or (B) chronic infection. **(A)** After the incubation phase, with HCV viral RNA detectable in the blood, ALT levels rise in most patients (top panel), although only 20-30% of patients develop clinical symptoms such as jaundice. The rise in ALT coincides with an HCV specific T cell response in the liver (bottom panel). T cell responses in spontaneously resolved infection can be detected for years after the clearance of infection. (**B**) Chronically evolving acute hepatitis C. Chronic HCV infection is characterized by relatively stable HCV RNA levels, which are lower than in the acute phase (top panel). HCV-specific T cell reactivity decreases over time, and *in vitro* responses to HCV antigens are typically weak to undetectable (bottom panel). Figure modified from (Rehermann 2009).

1.5.3 T cell responses in spontaneously resolved HCV infection

Cellular immune responses have been shown to play a critical role in the spontaneous resolution of HCV infection (Grakoui et al. 2003; Shoukry et al. 2003). As shown in Figure 1-12, T cell responses in spontaneously resolved infection have been described as broad and vigorous (Takaki et al. 2000; Grüner et al. 2000; S. Cooper et al. 1999; Lechner et al. 2000), and persist after the virus is cleared from the blood. However, the exact determinants distinguishing a successful from an unsuccessful immune response remain unclear. Evidence for the role of CD4+ and CD8+ specific T cell responses for spontaneous resolution of HCV infection is discussed below.

HCV specific CD4+ cells are essential for the generation of an effective immune response. Vigorous CD4 responses are detected in the blood of patients who clear the virus, whereas a loss of initially strong CD4+ T cell responses has been associated with recurrent viraemia in humans and chimpanzees (Gerlach et al. 1999; Nascimbeni et al. 2003). Similar to the associations of HCV clearance with class-I MHC types (discussed below), class-II MHC types have been associated with viral clearance: In a cohort of 758 US women infected with HCV, DRB1*0101 was associated with viral clearance, and DRB1*0301 with viral persistence (Kuniholm et al. 2010). A meta-analysis of 11 studies proposes a significant influence of class-II types DQB1*0301 and DRB1*1101 on viral clearance (X. Hong et al. 2005), and DRB1*01, DRB1*0401 and DRB1*15 were associated with protection in the well defined Irish cohort (McKiernan et al. 2004).

The role of CD8+ T cells has been thoroughly studied. Collectively, published data shows an important role for CD8+ restricted T cells. The clearance of infection is supported by several lines of evidence (reviewed by (Thimme et al. 2012)), including:

(1) CD8+ T cells have been shown to control viral replication in a chimpanzee model *in vivo* (Shoukry et al. 2003).

(2) A kinetic correlation between CD8+ T cell responses and viral clearance, where strong and broad intrahepatic CD8+ responses are associated with the onset of clinical disease and viral clearance (S. Cooper et al. 1999; Grüner et al. 2000; Thimme et al. 2001; A. L. Cox, Mosbruger, Lauer, et al. 2005).

(3) An association between HLA type and viral clearance, shown for alleles A*03, B*27, B*57 and Cw*01, whereas B*08 and Cw*04 were associated with viral

persistence (Thio et al. 2002; McKiernan et al. 2004; Kuniholm et al. 2010; A. Y. Kim et al. 2010; Fitzmaurice et al. 2011).
(4) CD8+ T cells inhibit viral replication *in vitro* (Jo et al. 2009), using a HCV infected hepatocyte cell line expressing MHC class-I molecules for HLA type A*02, and HCV specific T cells exerting strong antiviral effects.
Even though most patients achieve spontaneous resolution of the virus within 6 months of infection, there is published evidence that spontaneous resolution of the virus might occasionally occur after chronic infection is established (Scott et al. 2006); a phenomenon particularly common amongst Alaskan natives.

1.5.4 T cell responses in chronic infection

Up to now, it was thought that adaptive immune responses in patients developing chronic infection are initially weak and narrowly focused. However, a recent paper by Lauer and co-workers showed that in individuals acutely infected with HCV and proceeding to chronic infection, T cell responses are initially strong, but decrease quickly during the phase of acute infection (Schulze Zur Wiesch et al. 2012). The assumption that responses in acutely infected patients proceeding to chronic infection are weak, might have been based on patchy data: patients evolving chronic infection rarely exhibit clinical symptoms during acute infection, and are therefore diagnosed later compared to patients with clinical symptoms who more likely clear infection (Thimme et al. 2001). Due to late diagnosis, initial strong responses therefore might have been missed in patients analysed with acute infection. With a careful selection of sample time points early after initial infection, as done by Schulze Zur Wiesch *et al.*, strong T cell responses were observed in acute patients irrespective of outcome (Schulze Zur Wiesch et al. 2012).
Once chronic infection is established, it is difficult to detect a broad range of T cell responses in peripheral blood (Lechner et al. 2000; Thimme et al. 2001; Shoukry et al. 2003; Grakoui et al. 2003). However, adaptive immune responses are often present in chronic HCV infection and are enriched in the liver (Kang et al. 2012) and are thought to contribute to the progression of liver disease, but fail to control infection. A combination of host and viral factors are thought to contribute to the lack of T cell response, such as down-regulation of T cell responses, T cell exhaustion and suppression by regulatory T cells, potentially leading to viral persistence (reviewed by

(Neumann-Haefelin et al. 2007; Thimme et al. 2012)) and are discussed in detail below.

1.5.4.1 Viral escape

An efficient adaptive immune response relies on the ability of the host to detect the pathogen and target appropriate viral infected cells. An important escape mechanism is viral variation in T cell epitopes, where polymorphisms in or flanking targeted epitopes can result in viral escape from HLA specific immune responses (reviewed by (Bowen & C M Walker 2005b)).

This escape is facilitated by the unusually fast replication rate of HCV, estimated at 10^{12} virions per day in each infected individual (Neumann et al. 1998; Zeuzem et al. 1998; Ramratnam et al. 1999). The viral RNA dependent RNA polymerase lacks a proofreading function (Steinhauer et al. 1992), resulting in many co-circulating viral quasi-species in an infected individual (Martell et al. 1992; M Kurosaki et al. 1994). This characteristic set of diverse viral sequences is not completely random, but seems to be defined by two restrictors: the fitness of the virus and the HLA type of the infected host (S. C. Ray et al. 1999; Mao et al. 2001; A. L. Cox, Mosbruger, Mao, et al. 2005). At the same time, the high number of viral quasi-species and variants facilitates the escape from HCV T cell pressure at immunogenic sites.

Viral escape has first been experimentally demonstrated in patients chronically infected with HCV (K. M. Chang et al. 1997), and has been associated with chronicity (Erickson et al. 2001). Already early during the acute phase of infection, viral escape from CD8+ T cell responses has been observed in patients who develop chronic infection, but not in individuals who clear the virus (Timm et al. 2004; A. L. Cox, Mosbruger, Lauer, et al. 2005; Tester et al. 2005). The escape pattern relates to the infected patients' HLA type, and polymorphic footprints have been used in several association studies to identify new T cell epitopes in chronic infection (Gaudieri et al. 2006; Rauch et al. 2009; Ruhl et al. 2011). In addition to mutations within T cell epitopes, compensatory mutations outside of T cell antigenic targets often correspond to the changes within an epitope (Tester et al. 2005; Ruhl et al. 2011; Ruhl et al. 2012).

The effectiveness of viral control through T cell mediated immune pressure relates to the cost of viral fitness at a certain epitope, limiting the replication efficiency of the virus (Wölfl et al. 2008). In epitopes linked to protection, such as the B*27 restricted

epitope (AA 2841–2849) (McKiernan et al. 2004; Neumann-Haefelin et al. 2006), it has been shown that mutations at anchor positions lead to a substantial loss of fitness (Dazert et al. 2009). However, in chronically infected patients, mutations spare the anchor positions, forming clusters of mutations around them that lead to an efficient escape from T cell pressure (Dazert et al. 2009). Additionally, viral escape does not occur universally in all T cell targets, but only in about 50-70% (A. L. Cox, Mosbruger, Mao, et al. 2005; Neumann-Haefelin et al. 2008), suggesting that other mechanisms contribute to T cell failure (Thimme et al. 2012).

1.5.4.2 CD8+ dysfunction

T cell exhaustion has been demonstrated for a variety of chronic viral infections (like HIV, HCV, HBV, CMV), and is associated with a failure of effector functions like the production of IL-2 and IFNγ, limited proliferation capacity, and sustained expression of inhibitory molecules (Wherry 2011). CD8+ T cells are thought to become exhausted through a prolonged phase of antigenic stimulation without sufficient CD4+ help (Wherry & R. Ahmed 2004; Virgin et al. 2009). Both cell intrinsic negative regulatory pathways (such as Programmed cell death protein 1 (PD-1)) and extrinsic pathways (such as regulatory cytokines from regulatory T cells) seem to play an important role (Wherry 2011). Dysfunctional T cells seem to lose CTL effector functions in a hierarchical order: first, the ability to produce IL-2, then TNFα and last IFNγ (Lechner et al. 2000; Semmo et al. 2005; Wherry 2011).

In HCV infection, exhausted T cells are commonly observed in chronically infected patients (Penna et al. 2007; Radziewicz et al. 2007; Radziewicz et al. 2008; N. Nakamoto et al. 2009) and express a variety of inhibitory molecules. A well described inhibitory molecule expressed on exhausted T cells is programmed-death 1 (PD-1), which is up-regulated in HCV specific T cells (Urbani, Amadei, Tola, et al. 2006; Penna et al. 2007; Radziewicz et al. 2008). Particularly high PD-1 expression can be found on HCV specific T cells in the liver and spleen (V. Kasprowicz et al. 2008; Sumida et al. 2013). Other co-expressed markers on exhausted CD8+ cells include 2B4, CD160 and KLRG1 (Bengsch et al. 2010). T cells with up-regulated exhaustion markers also display low levels of CD127, suggesting that they do not belong to a memory T cell subgroup (Radziewicz et al. 2007; Bengsch et al. 2010). Another exhaustion marker up-regulated in HCV chronic disease is the T cell Ig and mucin domain protein-3 (Tim-3) (McMahan et al. 2010; Moorman et al. 2012), with

expression correlating with HCV disease progression (Vali et al. 2010). Tim-3 has been shown to negatively regulate effector T cells, and was overexpressed on both effector and regulatory T cells in HCV infection (Mengshol et al. 2010). Additionally, Tim-3's natural ligand, Galectin-9, has been shown to be up-regulated in HCV infection, inducing apoptosis on HCV specific CTLs (Mengshol et al. 2010).
Blockade of PD-1 increases CD8+ effector function *in vitro* (Radziewicz et al. 2007; Penna et al. 2007; Golden-Mason et al. 2007; N. Nakamoto et al. 2008). A similar singular and synergistic effect was observed for a different inhibitory molecule highly expressed on HCV specific T cells, cytotoxic T lymphocyte-associated antigen (CTLA-4) (N. Nakamoto et al. 2009), which is also described as a marker of regulatory T cells.
However, PD-1 expression on HCV specific T cells does not necessarily identify exhausted T cells and was shown to be highly expressed on HCV specific T cells during acute infection irrespective of clinical outcome (Bowen et al. 2008; V. Kasprowicz et al. 2008), suggesting other pathways play an important role in CD8+ T cell exhaustion. Taken together, these results imply multi-factorial causes for T cell exhaustion, with the activation of several inhibitory pathways, and additional factors like the lack of CD4+ cell help and CTL suppression through CD4+ regulatory T cells.

1.5.4.3 Lack of CD4+ help

While CD8+ T cells have a major role in the elimination of pathogens, they depend on appropriate CD4+ T cell help. CD4+ specific mechanisms are thought to contribute to CD8+ T cell dysfunction in chronically infected individuals, and a dysfunctional phenotype in CD4+ cells in persistent viral infections comparable to CD8+ exhaustion has been described (Oxenius et al. 1998; D. G. Brooks et al. 2005; Urbani, Amadei, Fisicaro, et al. 2006).
In patients chronically infected with HCV, CD4+ cells have been shown to mount weak and functionally impaired responses (Semmo et al. 2005; Lucas et al. 2007). CD4+ T helper cell responses have been associated with the outcome of acute HCV infection, with weak and narrowly focused CD4+ responses that are functionally impaired in patients who evolve chronic infection (Urbani, Amadei, Fisicaro, et al. 2006). In addition, strong CD4 responses during acute infection were primed irrespective of clinical outcome, but rapidly disappear from the blood in patients

developing chronic infection, with early functional defect (e.g. low levels of secreted IL-2) (Schulze Zur Wiesch et al. 2012).

1.5.4.4 Suppression by regulatory T cells

Regulatory T cells (Tregs) suppress the activation, differentiation and effector functions of immune cells such as CD4+ and CD8+ cells (Shevach et al. 2006), B cells (D.-M. Zhao et al. 2006), NK cells (Ghiringhelli et al. 2006) and dendritic cells (Mahnke & Enk 2005). Currently, no single marker can be used to characterize Tregs. In most published studies, regulatory T cells are characterized by an expression of CD25. However, this expression is not exclusive, since CD25 expression has been observed on activated effector T cells (Bluestone & Abbas 2003). However, in chronic infection the CD4+CD25+ subset most likely represents a regulatory T cell subset (Manigold & Racanelli 2007). Another marker of regulatory T cells is the transcription factor forkhead box P3 (FOXP3), which has been used in multiple studies to characterize T regulatory subsets (S. M. Ward et al. 2007; Heeg et al. 2009). However, FoxP3 can be expressed transiently by nearly all activated human CD4+ and CD8+ T cells after exposure to strong antigens (M. R. Walker et al. 2003; Pillai et al. 2007; J. Wang et al. 2007). It was proposed, though, that stable expression may be related to suppressor activity (J. Wang et al. 2007). Other distinguishing markers for T regulatory subsets include a high CD39 expression and low CD127 expression (Bengsch et al. 2007).

In acute and spontaneously resolved HCV infection, FoxP3+ cells Tregs were rarely detected (Heeg et al. 2009; Langhans et al. 2010). In chronically HCV infected individuals however, high frequencies of CD4+CD25+ cells were observed compared to spontaneously resolved patients or healthy controls (Sugimoto et al. 2003; Cabrera et al. 2004; Boettler et al. 2005), with CD4+CD25+ cells exerting immunomodulatory functions on HCV specific CD4+ and CD8+ cells. Additionally, it was shown that regulatory T cells suppress *in vitro* proliferation and perforin expression of CD8+ cells (Rushbrook et al. 2005).

A strikingly high proportion of CD4+FoxP3+ regulatory T cells was observed in liver samples from HCV infected patients compared to livers of patients with other causes of liver disease and healthy livers (S. M. Ward et al. 2007), but conflicting data exists on the influence of T regulatory cells on liver damage. Some studies suggested that high numbers of CD4+FoxP3+ cells limit the extent of fibrosis, but could not find an

association with ALT levels (Claassen 2010); however, other groups saw that increased amounts of CD4+CD25+ cells were associated with abnormal ALT levels (Bolacchi et al. 2006). In a different study, CD4+CD25+FoxP3+ T regulatory cells correlated inversely with ALT levels, but directly with MELD score and disease progression (Ferri et al. 2011).

In addition to the CD4+ T regulatory subset, a subset of CD8+ restricted HCV specific regulatory T cells has been observed in the liver, which attenuated hepatocellular necrosis, liver fibrosis and inflammation (M. Abel et al. 2006). These CD8+ T cells express high levels of IL-10, and their suppression of CTLs was reversible *in vitro* through IL-10 neutralizing antibodies (Accapezzato et al. 2004).

The mechanisms by which HCV specifically activates regulatory T cells remains unclear. Recent studies suggest that HCV antigen presentation by hepatic stellate cells preferentially induces functional regulatory T cells (Dunham et al. 2013). Others have shown an influence of HCV core on the generation of regulatory T cells, with HCV core-transduced CD4+ cells showing a regulatory phenotype (expression of regulatory cytokines CTLA-4 and FoxP3, and production of IL-10) compared to non-transduced CD4+ cells (Dominguez-Villar et al. 2012; Fernandez-Ponce et al. 2014). Additionally, Langhans et al. identified functionally active HCV core-specific Treg-cells in patients with chronic hepatitis C, which require the continued presence of antigen to maintain their functional differentiation (Langhans et al. 2010).

1.5.5 Humoral responses

Whether humoral responses play an important role in HCV infection remains controversial. B cells do not seem to be essential for the clearance of HCV infection, since patients with hypoalbumineamia can spontaneously eradicate HCV (Christie et al. 1997). Even though not necessary for HCV clearance, antibody (Ab) responses in HCV infected patients against structural and non-structural proteins evolve in all patients, regardless of outcome (Kaplan et al. 2007; Lin Liu et al. 2010), but are not neutralizing or cross-reactive in most patients (Logvinoff et al. 2004).

In contrast to these data, it has been shown more recently that a rapid induction of antibodies during the early phase of infection may contribute to HCV control (Pestka et al. 2007) and that virus-specific neutralizing Abs (nAbs) drive sequence evolution and correlate with the outcome of infection (Dowd et al. 2009).

HCV envelope proteins are well-accepted targets for humoral responses, and most of the host antibody response is formed to E1 and E2 (Prentoe et al. 2011). In this viral region, humoral antibodies seem to be driving envelope gene evolution (J. C. Booth et al. 1998; Pestka et al. 2007; T. von Hahn et al. 2007; Dowd et al. 2009), and recent studies have suggested nAb to be causative in the evolution of HCV envelope proteins (Lin Liu et al. 2010). However, these viral regions are particularly variable, impairing the potential for viral neutralization (Prentoe et al. 2011).

1.5.6 The influence of the immunotolerant environment of the liver on HCV specific immune responses

Even with an optimal vaccine-mediated induction of T cell responses, protection against HCV infection might be hampered via the immunotolerant environment of the liver, that might lead to inefficient antigen presentation and therefore suboptimal recognition of presented viral antigen by vaccine-induced T cells.

The fact that the liver is an unusually immunotolerant organ was first noted in transplantation: for liver transplants, no HLA match of donor and recipient is necessary (Calne et al. 1967; F. Poli et al. 1998). In addition, patients with liver transplants or other organs transplanted in conjunction with a liver require less immunosuppression (Calne et al. 1969), suggesting that the liver can induce relative immunotolerance. However, hepatocyte transplants alone are acutely rejected (Bumgardner et al. 1998; Bumgardner & Orosz 2000), suggesting an important role for non-parenchymal liver cells in the induction of immunotolerance.

The liver receives blood from both the systemic circulation and the intestine via the portal vein, which carries digestive as well as microbial products from the gut, exposing the liver to huge amounts of antigens. Arterial and venous blood mixes in the liver sinosoids, where high numbers of immune cells are found and antigens are presented. Immune responses to the liver are primed, but they are frequently ineffective. This is potentially caused by ineffective antigen presentation through non-professional APCs (Zinkernagel 1996), which has been also described for other viruses like HPV, where immunogens presented by keratinocytes do not activate a significant immune response until they are released from lysed cells and presented by professional antigen presenting cells (Tindle & Frazer 1994).

Multiple subsets of antigen presenting cells (APC) are abundant in the liver, including several subsets of dentritic cells and mononuclear phagocytes in the form of Kupffer

cells (Crispe 2009). In addition, there are two forms of liver specific APCs; liver sinusoidal endothelial cells (LSECs) and hepatic stellate cells (HSCs) (Knolle & G Gerken 2000). These cells are thought to mediate the liver specific immunotolerance by secretion of anti-inflammatory cytokines and induction of T cell apoptosis. They also down-regulate MHC molecules and co-stimulatory molecules, and up-regulate co-inhibitory receptors such as PD-1 (Tiegs & Lohse 2010). A particular immunosuppressive role has been described for HSCs that mediate T cell apoptosis via the Fas/FasL pathway and the regulation of TGF-β and IL-10 production (Z. Jiang et al. 2013). In addition to the particular properties of antigen presenting cells in the liver, activated CD8+ T cells can be functionally inactivated in the liver or undergo apoptosis (Knolle & Thimme 2014). A further liver specific feature is the lack of CD4+ activation, due to a paucity of MHC class II presenting cells and lack of co-stimulation via antigen presenting cells.

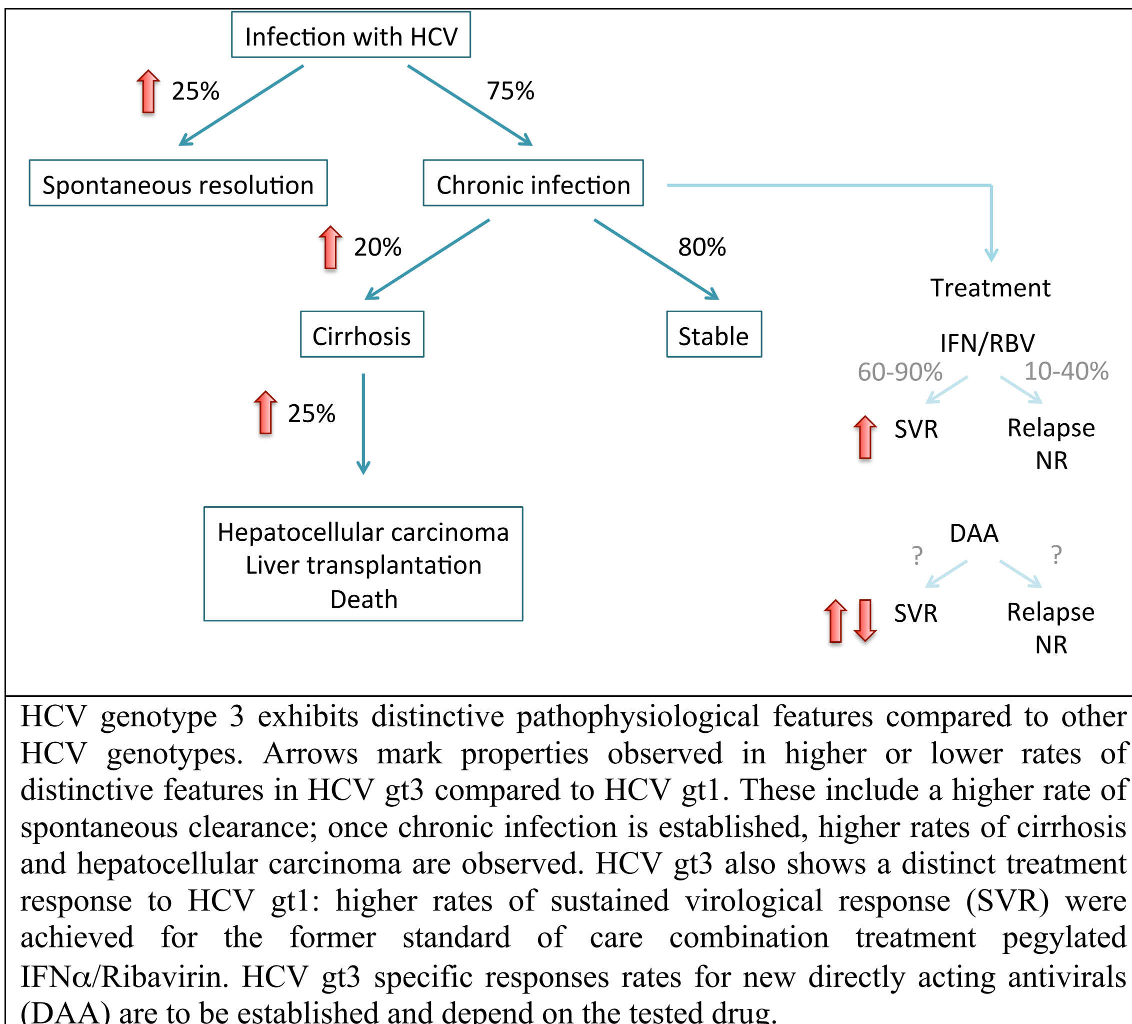
Based on these particular immunotolerant properties, the liver is the site of several infections where the immune response fails to clear the virus and the infection contributes to the disease by causing tissue damage and destruction of liver function. HCV is one of the pathogens utilizing the comparatively immunoquiescent zone of the liver to establish chronic infection, as well as other hepatotropic pathogens such as Hepatitis B virus, Plasmodium spp. and Echinococcus (Protzer et al. 2012).

In addition to the liver specific properties described above, HCV has particular characteristics that are employed in the liver. It has the ability to hide from immune cells by infecting liver cells by direct cell to cell transmission, avoiding neutralization of the virus (Timpe et al. 2008). HCV also escapes cytosolic recognition by helicases by blocking the signalling molecule IFNB-promoter stimulator 1 (IPS1), and can mask HCV specific surface glycans to escape antibody related immune response (Helle et al. 2007).

1.6 Distinctive features of HCV genotype 3 infection

In this thesis, I comprehensively analysed T cell immunity to HCV gt3 infection. The rationale for a detailed analysis of HCV gt3 specific responses was based on the fact that this genotype represents the second most common genotype globally, but also on the observation that HCV gt3 carries several distinct pathophysiological properties, which are introduced below. HCV genotype 3a accounts for roughly 30% of infections in the Western world (Wartelle-Bladou et al. 2012), and 35 to 80% of infections in the Indian Subcontinent, South East Asia and Australia (Shepard et al. 2005; Hissar et al. 2006) (see 1.2.3, page 24). Besides the feature that infections with HCV gt3 are more readily cleared spontaneously (see 1.3.1.2, page 28), genotype 3 carries several pathophysiological patterns, which are also shared by other genotypes, but which may be particularly pronounced in this genotype (Figure 1-13).

Figure 1-13: Distinctive features of HCV genotype 3 infection



HCV genotype 3 exhibits distinctive pathophysiological features compared to other HCV genotypes. Arrows mark properties observed in higher or lower rates of distinctive features in HCV gt3 compared to HCV gt1. These include a higher rate of spontaneous clearance; once chronic infection is established, higher rates of cirrhosis and hepatocellular carcinoma are observed. HCV gt3 also shows a distinct treatment response to HCV gt1: higher rates of sustained virological response (SVR) were achieved for the former standard of care combination treatment pegylated IFNα/Ribavirin. HCV gt3 specific responses rates for new directly acting antivirals (DAA) are to be established and depend on the tested drug.

These include a higher rate of steatosis, fibrosis and HCC. In addition, HCV genotype 3 shows a distinct response to HCV treatments, reported both for combination therapy with pegylated IFNα/Ribavirin and the newer treatment options including DAAs.

1.6.1 Distinct pathophysiological features of HCV genotype 3 infection

Steatosis: It is a well known clinical finding that HCV genotype 3 is associated with higher rates of steatosis (Hui et al. 2002). A proposed mechanism for these clinical findings is that HCV genotype 3 interferes with the late cholesterol synthesis pathway, evidenced by lower distal sterol metabolites and preserved lanosterol levels (Clark et al. 2012). This distal interference resolves after successful treatment, with a normalization of distal sterol metabolite levels. In addition, a clinical association between viral load and steatosis in HCV gt3, but not in other studied genotypes, was observed (Adinolfi et al. 2001).

Lipid metabolism: HCV interferes with the lipid metabolism at key stages of its life cycle such as viral entry, viral replication and virion release. It seems that HCV hitches a ride on the cell lipid cycle: HCV enters hepatocytes via the low density lipoprotein (LDL) receptor (Thompson et al. 2010), in form of a lipoviral particle, incorporating mainly low and very low density lipoproteins (LDL and VLDL) (S. Nielsen et al. 2006; Bassendine et al. 2013). In cell culture it was shown that HCV core and NS5a interact with the lipid cycle (Perlemuter et al. 2002; Shi et al. 2002), and that the lipid droplet is essential for viral assembly and virion release (Miyanari et al. 2007). A particular role of HCV genotype 3 in lipid metabolism was described in cell culture: it was shown that when expressing core protein of different genotypes in Huh7 cells, genotype 3 showed the highest accumulation of triglycerides (Abid et al. 2005). A possible mechanism for this could be that HCV gt3 core activates the promoter for the fatty acid synthethase to a much higher level than HCV gt1 (Jackel-Cram et al. 2010). Several lines of clinical evidence support these experimental findings. It was an early observation that HCV patients have a low LDL and cholesterol levels (Corey et al. 2009). Especially the betalipoprotein fraction is decreased, a finding particularly pronounced in patients with HCV genotype 3 (Serfaty et al. 2001). Finally, the degree of interference with the cell's lipid cycle is influential for HCV treatment success (S. A. Harrison et al. 2010), with higher LDL levels associated with better treatment response.

Fibrosis progression: In a large study from Switzerland enrolling 1189 patients to identify risk factors for fibrosis progression, independent risk factors identified included male sex, age at infection, and, with the strongest statistical association (odds ratio of 1.89), HCV genotype 3a infection (P.-Y. Bochud et al. 2009). In a meta-analysis from the same group, HCV genotype 3a was associated with faster fibrosis progression (Probst et al. 2011). It was proposed that an additional cytopathic effect of steatosis causes this accelerated fibrosis progression unique to genotype 3a (Westin et al. 2002).

Development of HCC: Steatosis can contribute to the development of hepatocellular carcinoma (HCC). In a study prospectively screening for the development of HCC in 353 patients with proven hepatic cirrhosis from 1994 to 2007, infection with HCV genotype 3 proved to be an independent predictor for the development of HCC in a multivariate analysis (Nkontchou et al. 2011).

1.6.2 HCV genotype 3a and response to antiviral therapy

Standard treatment for HCV infection up to now was a combination treatment with peg-IFNα and Ribavirin for 24 to 48 weeks depending on HCV genotype and rapid virological response. Until now, HCV gt1 infection has been considered to have the most unfavourable treatment response, with SVR rates of approximately 50% on a dual therapy of peg-INF-α/Ribavirin, whereas genotypes 2 and 3 were considered to be relatively easy to treat genotypes, with SVR rates of 70-80% (Manns et al. 2001). Treatment success is expected to rise considerably for HCV gt1 infection using the new DAAs, but evidence is now arising that genotype 3 might not be as easy to treat and its pathogenicity might be more troubling than previously suspected (Goossens & Negro 2013; Tapper & Afdhal 2013): In most published studies, HCV genotypes 2 and 3 have been analysed together. On closer inspection of clinical trials, however, it seems that HCV genotype 3 patients have a much poorer response to combination therapy with peg-IFNα/Ribavirin than previously thought (Tapper & Afdhal 2013). Tapper and Afdhal extracted data on HCV gt3a patients from a selection of studies pooling gt2 and gt3 patients, showing that HCV genotype 2 has response rates of 80%, whereas rates in HCV genotype 3a patients range between 65-70%, respectively (Table 1-3). This difference might be due to a higher rate of steatosis and fibrosis in HCV genotype 3 infected patients (Aghemo et al. 2009). Tapper and Afdhal conclude that HCV genotype 3 has always been an intermediate IFN responsive strain, but with

a larger subset of patients that attain RVR and have shortened duration of therapy. However, HCV gt3 should be evaluated separately from genotype 2. Another consideration when analysing HCV genotype 3 cohorts is varying composition of patient groups with major differences in host factors like ethnicity, IL28B status and presence of cirrhosis (Mangia et al. 2013): a young European patient with HCV gt3 infection acquired from IVDU might respond very differently to INFα/Ribavirin therapy than South-East Asian patients, who are often older and have more advanced disease.

Table 1-3: Response rates to combination therapy in HCV genotype 3a patients

Trial	No. of gt3 pts	Duration of therapy (weeks)	RVR (%)	SVR (%)	SVR in cirrhosis (%)
(Zeuzem et al. 2004)	182	24	75.3	79	N/A
(Dalgard et al. 2004)	99	14 24 (without RVR)	74.7	80.1	N/A
(Mangia, Bandiera, et al. 2010)	70	12 vs 24	58.6	65.7	N/A
(M. von Wagner et al. 2005)	113	16 vs 24	92	76 vs 75 (RVR) 40 (non-RVR)	N/A
(M. L. Shiffman et al. 2007)	727	16 vs 24	52.3	62.2 vs 67	43 vs 49
(Dalgard et al. 2008)	343	14 vs 24 (RVR) 24 (without RVR)	68.8	84 vs 91 (RVR) 56.3 (non-RVR)	
(Lagging et al. 2008)	276	12 vs 24	N/A	58 vs 78	30 vs 57
(Mecenate et al. 2010)	94	12 vs 24 (if RVR) 24 (non-RVR)	68.1	78.1 (RVR) 43.3 (non-RVR)	N/A
(Mangia, A. J. Thompson, et al. 2010)	414	24 vs 12 vs 36	63.1	71.5 vs 74.9	N/A

Abbreviations: gt3: HCV genotype 3, RVR: rapid virological response, SVR: sustained virological response, No. of gt3 pts.: number of HCV genotype 3 patients included in the study. N/A: not available. Table adapted from (Tapper & Afdhal 2013).

With the newly developed directly acting antivirals (DAAs), response rates in HCV genotype 1 infection have been increasing to over 70% (see 1.3.2.2 Directly acting antivirals (DAA, page 31). Multiple new targets have been tested in clinic, but most of the first generation DAAs, especially protease inhibitors, show only limited antiviral efficiency in HCV genotype 3a (G. R. Foster et al. 2011), and are only partly active against genotypes 4, 5 and 6 (Table 1-4) (Schinazi et al. 2014). Several other drug classes, like polymerase inhibitors, NS5A inhibitors and host targeting antivirals show a stronger cross-genotypic activity.

Table 1-4: Leading protease inhibitors are highly active against multiple HCV genotypes except HCV gt3

Protease inhibitor	Company	Clinical phase	Chemical structure	Genetic barrier	Drug-Drug interaction	Active against HCV genotypes
Boceprevir	Merck	approved	linear	low	moderate	1,2
Telaprevir	VJM	approved	linear	low	high	1,2
Simeprevir	Janssen	approved	macrocyclic	moderate	low	1,2,4,5,6
Faldaprevir	Boehringer-Ingelheim	III	macrocyclic	moderate	low	1,2
Asunaprevir	BMS	III	macrocyclic	moderate	low	1,4
ABT-450	Abbvie	III	macrocyclic	moderate	low	1
Danoprevir	Roche	II	macrocyclic	moderate	low	1,2,4
Vedroprevir	Achillion	II	macrocyclic	moderate	low	1
Vaniprevir	Merck	II	macrocyclic	high	low	1,2,4,5,6

Antiviral activity of leading protease inhibitors against HCV genotypes. For each protease inhibitor, the developing company, the clinical phase, chemical properties, genetic barrier and drug-drug-interactions are shown. Table adapted from (Schinazi et al. 2014).

For HCV gt3, NS5B **polymerase inhibitors** nucleotide analogues with pangenotypic activity are the leading drug class in clinical trials. Sofosbuvir recently was approved for clinical use in HCV genotypes 1 to 6 (Gilead press releases 2014a).

Initially, clinical data from a non-inferiority trial (Fission) in HCV gt2 and gt3 infected patients showed no increase of SVR rates compared to peg-Interferon/Ribavirin; SOF/RIB only showed a benefit in HCV gt2 patients (SVR 97% vs. 78%), but not in HCV gt3 patients (SVR 56% vs 63%) (Lawitz et al. 2013). Further trials (Positron and Fusion) enrolled patients who were interferon intolerant or ineligible (Positron) or had a prior non response (Fusion) to treatment with IFN/RIB; and treated patients with SOF/RIB for 12 or 16 weeks with a low response rate in HCV genotype 3 infected individuals (I. M. Jacobson et al. 2013). However, SVR rates in these trials were increased considerably by a longer duration of treatment (12 weeks 30%, 16 weeks 62%) in HCV gt3 patients. Since all patients initially achieved RVR, the low SVR rates were due to viral relapse specific to HCV gt3.

This issue was then addressed in a recent European trial (Valence) with longer treatment regimens of 24 weeks SOF/RIB in HCV gt3, which then lead to an impressive increase of SVR rates to over 90% in patients with and without cirrhosis (Figure 1-14, left). Only in the patient group most difficult to treat (cirrhotic HCV gt3 patients with previous treatment failure), longer treatment duration did not lead to increased SVR rates (Figure 1-14, right).

Figure 1-14: Interferon free trials including Sofosbuvir for HCV genotype 3 subjects.

Figure deleted due to copyright:

Figure material can be found in Pawlotsky 2014. Figure included in thesis has been derived from information included in Figure 1 (c and d) in this paper.

SVR rates of HCV genotype 3 patients in phase III trials. Patients received sofosbuvir plus ribavirin for either 12/16 weeks (FUSION) or 24 weeks (VALENCE). ***Left:*** SVR rates in treatment-naive patients according to fibrosis stage (cirrhosis vs no cirrhosis) and treatment duration (12 or 24 weeks). ***Right:*** SVR rates in treatment-experienced patients according to fibrosis stage and treatment duration. Figure modified from (Pawlotsky 2014).

NS5A inhibitors Daclatasvir, the first clinical available NS5A inhibitor, has potent activity against HCV genotype 1 in clinical studies (Nettles et al. 2011; Chayama et al. 2012; Lok et al. 2012). When tested in HCV gt2 and gt3 patients, higher SVR rates were achieved in genotype 2 than genotype 3, due to a high rate of viral relapse in genotype 3 patients, who also showed a higher rate of cirrhosis (G.J. Dore et al. 2013). Some novel second generation NS5A inhibitors have potent *in vitro* activity against HCV genotype 3a, with *in vivo* activity to be tested in clinical trials.

Other drugs showing a good pan-genotypic profile are **cyclophilin inhibitors**, which also have a high barrier to resistance. In a recent clinical trial Alispovir (Novartis) achieved high SVR rates (88%) in combination with Ribavirin, with no differences seen between HCV gt2 and 3 patients (Pawlotsky et al. 2012).

Most recently, high response rates in HCV genotype 3 have been reported for several combinations of new directly acting antivirals, such as Sofosbuvir/Daclatasvir (Sulkowski et al. 2014) and Sofosbuvir/Lepidasvir (Gane et al. 2014). These and other combinations will most likely lead to good treatment responses even in HCV gt3 patient that are currently difficult to treat, such as patients with liver cirrhosis and treatment experienced patients.

In conclusion, with the new DAAs HCV genotype 1 has become a genotype much easier to treat. Currently, research effort has turned to optimising new treatment regimens for HCV gt3 infection. It is to be expected that new interferon-free treatment options for HCV gt3 will be available in the near future.

1.6.3 Innate immune responses in HCV gt3a infection

The IL28B gene codes for interferon (IFN)-λ3 and polymorphisms linked to this gene have been shown to predict spontaneous resolution of infection and treatment response to combination therapy with pegylated Interferon-α and Ribavirin (IFN/RIB) in HCV gt1 infection. The role of the IL28B polymorphisms in HCV gt3 infection is less clear, and was evaluated in several independent cohorts with partially conflicting results: In a retrospective analysis of Mangia's *et al*. trial (Mangia, A. J. Thompson, et al. 2010), IL28B polymorphisms predicted rapid virological response (RVR), and sustained virological response (SVR) in patients who did not achieved RVR. In a study from Switzerland, SVR was not associated with IL28B genotype (Rauch et al. 2010). In a retrospective study reviewing 281 chronic HCV gt3 infected patients from a Swedish cohort IL28B status predicted RVR but not SVR (Moghaddam et al. 2011); a finding that was replicated in our cohort (Bucci et al. 2013). However, it is difficult to compare results from these studies, since in some studies gt2 and 3 patients were grouped (Lindh et al. 2011; Mangia, A. J. Thompson, et al. 2010; Rauch et al. 2010; Sarrazin et al. 2011), and treatment lengths or regiments varied (14 weeks IFN/RBV (Moghaddam et al. 2011) vs 24 weeks as standard of care).

Remarkably, IL28B genotype has been specifically associated with liver injury in HCV gt3 infection. In several publications (Moghaddam et al. 2011; Rembeck et al. 2012; Ydreborg et al. 2013), high ALT and liver injury has been associated with the infecting genotype, with higher liver injury seen in patients carrying the favourable IL-28B genotype in HCV gt3, but not HCV gt1 infected individuals.

Other genotype specific effects on the innate immune system relate to the interference of HCV viral proteins, leading to differences in induction levels of proteins of the cellular innate immune response. Inter-genotypic differences have been described for the cleavage activity of the HCV NS3 protease, cleaving MAVS and TRIF and thereby inhibiting the release of type I interferons (IFN-α and IFN-β). A higher cleavage activity of the NS3/4a protease has been described in HCV genotypes 2 and 3 compared to genotype 1 and 4 by detecting higher amounts of cleaved MAVS in liver biopsies of genotype 2 and 3 samples (Bellecave et al. 2010). Given the role of MAVS in Interferon-β production, this result would predict a lower induction of interferon stimulated genes (ISGs) in genotype 2 and 3, a phenomenon reported by

the same group (Sarasin-Filipowicz et al. 2008). However, inter-genotypic differences have not been assessed for other targets of the NS3/4a protease, such as TRIF (K. Li et al. 2005), the T cell protein tyrosine phosphatase (Brenndörfer et al. 2009) and interferon regulatory factor 3 (Foy et al. 2003).

1.6.4 Adaptive immune responses in HCV gt3a infection

Historically, T cell responses in HCV gt3 infection have been described in cohorts with mixed genotypes, mostly using HCV gt1 peptide sets. One of these early studies evaluated T cell responses in proliferation assays using recombinant protein NS3 in gt1, gt2 and gt3 patients (Hultgren et al. 2004). Hultgren et al. described that the magnitude of NS3-specific T cell responses was related to the viral genotype, with HCV gt3 patients exhibiting higher responses, and more rapid clearance of HCV RNA during antiviral therapy.

Two further studies directly compared cross-reactivity between HCV gt1 and gt3; however, both studies are limited in their approach: one only assessed T cell responses against a small region of the HCV genome (NS3) (Giugliano et al. 2009), the other only used sequence data to determine subtype specific escape mutations (Rauch et al. 2009). Results from both publications are discussed in detail below.

Evidence that immune responses may vary considerably between HCV gt1 and gt3 infection comes from a study assessing sequence polymorphisms due to T cell pressure in HCV gt1 and gt3 infection (Rauch et al. 2009). In this study, HCV non-structural regions (NS2 to NS5) were sequenced in 187 HCV gt1 and 136 HCV gt3 infected individuals. Viral sequence polymorphisms were subsequently assessed for associations with patient HLA types at each amino acid. In HCV gt1 infection, 32 viral HCV sequence polymorphisms were associated with patients' HLA type distributions. In HCV gt3 infection, 19 sequence polymorphisms were associated with patients' HLA types. However, minimal overlap of HLA-associated viral polymorphisms was observed between HCV gt1 and gt3 infection: only two viral polymorphisms were common for both genotypes. This was the case for a polymorphism in the NS3 region associated with HLA type A*0101 (Y1444F), previously described as immune mediated escape within epitope ATDALMTG**Y/F** in both HCV gt1 and gt3 infection (Neumann-Haefelin et al. 2008; Giugliano et al. 2009; Humphreys et al. 2012), and for a sequence polymorphism in NS5B (NS5B-2467) associated with HLA type B*1501, not previously published as an epitope, but

with a high predicted binding score in Syfpeithi and BIMAS. Following from these results, the authors conclude that the pattern of viral evasion is genotype specific and based on inter-genotypic sequence variation.

The first study to assess HCV gt3 specific T cell responses using a gt3 specific peptide set was a study on T cell immunity that was limited to the HCV NS3 region (Giugliano et al. 2009). Two peptide sets covering the NS3 regions of HCV gt1 and gt3 that were based on consensus sequences generated from the cohort viral strains were tested. IFNγ ELISpot assays were performed in HCV gt1 and gt3 infected patients and individuals with spontaneously resolved infection. Major differences in T cell specificity were observed for both genotypes, with only 22% of targets overlapping between HCV genotype 1 and 3. Of these shared targets, only 11% had an identical sequence in gt1 and gt3. Amongst the targets shared between the two genotypes were well-characterized epitopes like ATDALMTGF, IPFYGKAI, and HSKKKCDEI (Lauer et al. 2002; Wertheimer et al. 2003; Lauer et al. 2004; Kuntzen et al. 2007; Yerly et al. 2008).

In conclusion, studies investigating HCV T cell responses in HCV gt3 using specific gt3 cohorts and specific gt3 peptide sets are limited to one investigating T cell responses to the NS3 protein (Giugliano et al. 2009) and a recent study from our laboratory (Humphreys et al. 2012), although numerous publications included patients infected with HCV gt3. Also, not much is known on T cell responses in patients with spontaneous resolution of HCV infection; since it is difficult determine the infecting genotype once infection is cleared.

1.7 Aims of this work

This thesis aims to examine the implications of HCV subtype specific immunity on the design of a HCV vaccine with cross-genotypic action. I aim to assess T cell targets in HCV genotype 3 infection; this data on subtype specific immunity will then inform rational immunogen design. Next, I aim to design vaccine immunogens with the potential to induce T cell responses cross-reactive against multiple HCV genotypes.

1. To determine whether T cell responses differ between HCV subtypes:
    a. To comprehensively assess T cell targets in HCV genotype 3 infection in acute, chronic and spontaneously resolved infection
    b. Compare the results to previously published data in HCV genotype 1 infection and to define subtype specific T cell targets
2. To explore T cell cross-reactivity at frequently detected HCV genotype 3 specific T cell targets in the context of a highly variable HCV sequence population. I hypothesise that
    a. T cell targets will vary between HCV genotype 1 and 3 infection
    b. Limited T cell cross-reactivity exists at frequently targeted HCV gt3 specific T cell targets and sequences variants common in other genotypes, underlining the need of a cross-reactive vaccine
3. To design vaccine immunogens with the aim of inducing responses cross-reactive between HCV viral genotypes
    a. To define HCV sequence segments that are conserved between all HCV genotypes
    b. To assess whether these conserved segments are potentially capable of inducing cross-reactive immune responses, by
        i. analysing epitopes falling into conserved segments previously described in natural HCV genotype 1 infection, and HCV genotype 3 specific epitopes discovered in this thesis, and
        ii. predicting strong binders using online epitope prediction server
    c. To generate immunogens from conserved regions and modify constructs to optimize potential immunogenicity.

2 Materials and Methods

Standard techniques established in our laboratory were used for most of the procedures described. All experiments involving blood, or blood products like plasma or peripheral blood mononuclear cells (PBMC) from HCV infected individuals were performed in a biological containment area (derogated CAT3) in the Peter Medawar Building of Pathogen Research. This involved the handling of infectious agents within biological safety cabinets and wearing appropriate protective clothing.

2.1 Patient cohort

2.1.1 Patient recruitment

HCV gt3a infected individuals (108 chronically and 16 acutely infected) were recruited from different sites: John Radcliffe Hospital, Oxford, UK (acute=1, chronic 104), Massachusetts General Hospital, Boston, USA (acute=9, chronic=4), and John Hopkins Hospital, Baltimore, USA (acute=4, chronic=1). Additionally, 16 patients with spontaneously resolved HCV infection and 44 patients with HCV gt1a/b infection were recruited from the John Radcliffe Hospital, Oxford, UK. Informed consent and local ethical approval was obtained for all patients.

2.1.2 Treatment

Patients were treated with standard of care treatment, a combination of peg-Interferonα (peg-IFNα, 180μg/week) and Ribavirin (800-1200mg/day adjusted by body weight). Sustained virological response (SVR) was defined as negative HCV RNA 6 months after the end of treatment. Treatment non-response (NR) was defined as HCV RNA positive throughout treatment, and treatment relapse (REL) as HCV RNA negative during treatment, but RNA positive within the first 6 months after end of treatment.

2.2 Molecular Methods

2.2.1 Isolation of Plasma

For separation of plasma, whole blood was collected in an Ethylediaminetetraacetic acid (EDTA) tube and centrifuged at 2200 rpm for 10 min. The supernatant was aspirated and stored at -80°C.

2.2.2 Viral RNA extraction from Plasma

Patient plasma was thawed on ice. 500 µl of plasma were concentrated by high-speed centrifugation (23,600g) for 1 hour at 4°C. 360 µls were carefully removed from the top, and pellet at the bottom resolved in the remaining 140 µls. Viral RNA was extracted using the QIAmp Viral RNA Mini Kit (Qiagen) according to manufacturers instructions. RNA was eluted into 60 µls Buffer AVE (Qiagen) and stored in 20 µl aliquots at -80°C.

2.2.3 Primers

For HCV gt3a viral sequencing, primers were used as previously described (Humphreys et al. 2012). A full list of primers can be found in Table 2-1 and Table 2-2. Subtype-3a specific primers included both previously described and in house newly designed primers.

2.2.4 One-step Transcription First Round PCR

Reverse-transcription and first round PCR was carried out in a single reaction using the Superscript III One-Step Reverse Transcription - PCR system with Platinum Taq High Fidelity (Invitrogen) as previously described (Humphreys et al. 2012). 2.5µls of extracted viral RNA were added to a mixture of primers (10 µM each, MWG), dNTPs (25mM each, Bioline), Superscript III/Taq polymerase and buffer (Roche), and sterile distilled water to a final volume of 25 µls. For first round reactions, either a 4kb product encoding HCV structural proteins or a 7kb product encoding the non-structural proteins were amplified (Figure 2-1). Primer combinations 277-For/F4-Rev were used for the 4kb product (Core, E1 and E2) and 2412F/9192R for the 7kb product (p7, NS2-5 proteins). Samples were heated in a thermocycler with RT-PCR cycling conditions at 55°C for 30 min, 94°C for 2 min, 39 cycles of 94°C for 15 s, 58°C (4kb reactions)/ 60°C (7kb reactions) for 30 s, 68°C for 1min/kb, and a final extension step of 68°C for 10 min. After PCR reaction, the samples were kept at 4°C until further use.

Table 2-1: HCV genotype 3a specific forward primers

	Name	Primer Sequence (5'-3')	Tm	Position	Reference
5' UTR	156-For	GGTGAGTACACCGGAATCGCT	61.8	156	(Moghaddam et al. 2006)
	277-For	CCTTGTGGTACTGCCTGATAG	59.8	277	(Barnes et al. 2002)
	286-For	ACTGCCTGATAGGGTGCTTGC	61.8	286	(Humphreys et al. 2009)
Core	745-For	TACATCCCGCTCGTCGGC	60.5	745	
	887-For	CTTGCTTAATYCATCCAGC	53.4	887	
E1	1082-For	TCARGTAYGTCGGAGCAACCACCGC	67.9	1082	
	1416-For	AGCCTATTACTCCATGCAGGG	62.1	1416	
	1435-For	GGCAACTGGGCCAAGGTCGC	65.5	1435	
E2	2237-For	TCAAGGTGAGGATGTTTGTG	55.3	2237	
	2289-For	CAACTGGACCAGGGGGGAGCG	67.6	2289	
	F4 For	TGGGATGGGCGCTGAAATGGGA	60	2485	(Gaudieri et al. 2009)
	2412F2	CACCTCCACCARAACATYGT	57.3	2412	(Lucas et al. 2007)
NS2	3018-For	CATCACTAAGCTGCTGATAGC	57.9	3018	(Humphreys et al. 2009)
NS3	F5 For	ACAGCATACGCCCAGCAAACTAGG	60	3444	(Gaudieri et al. 2009)
	3892-For	ATCTTTAGGGCTGCTGTGTGC	59.8	3892	(Humphreys et al. 2009)
	F6 For	GATGAATGTCATGCCCAAGACGCTAC	62	4302	
	4832-For	GCCGAGGTAGACTCGGTACG	63.5	4832	
	F7 For	TGTCTCGTGCGGCTTAAGCCAA	59	5184	
NS4b	5712-For	TGCTGTGGCGTCTCTTATGGC	61.8	5712	
	6076-For	CTGTGCARTGGATGAACMG	56.7	6076	
	F8 For	GGAGGGAGCGGTACAGTGGATGA	63	6083	
NS5a	6573-For	TTGCCCATCACCCAACTACACTCGC	66.3	6573	
	F9 For	TGAGCTAGTGGACGCCAACTTGTTATG	62	7028	(Gaudieri et al. 2009)
	7507-For	GAGTCCGACTCAGAGTCATGC	61.8	7507	(Humphreys et al. 2009)
NS5b	F10 For	CGAAGTTCGGGTATAGTGCGAAGGA	61	7924	(Gaudieri et al. 2009)
	8500-For	AACACAATCACTTGYTACATCAAGGC	60.9	8500	(Humphreys et al. 2012)
	8848-For	TCCTGGTTRGGCAACATCATCATGTACGC	65.6	8848	
	8872-For	TACGCGCCTACCATCTGGGTGC	65.8	8872	

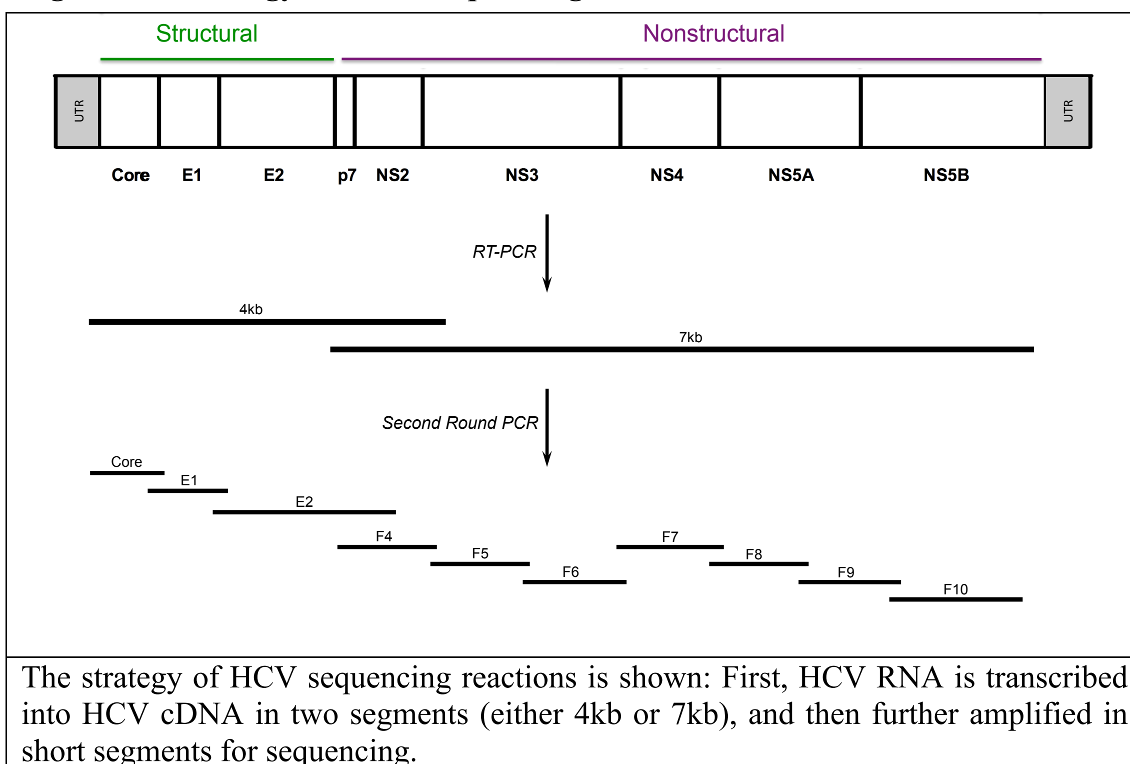
Table 2-2: HCV genotype 3a specific reverse primers

	Name	Primer Sequence (5'-3')	Tm	Binding	Reference
E1	1105-Rev	ACATGACTGCGTATCGAAGC	57.3	1105	(Humphreys et al. 2009)
	1207-Rev	GAGGTCTGAACGTGAAGG	58.8	1207	
E2	1585-Rev	ATGTGCCACGAGCCATTGGT	59.4	1585	(Barnes et al. 2002)
	1625-Rev	CGGTGTTTATGGACTCATTGC	57.9	1625	(Humphreys et al. 2009)
	2319-Rev	TCGCTGCGGTCACGATCCTCG	65.7	2319	
	2340-Rev	GAATGCAGCAGCGGATGTTGC	61.8	2340	
NS2	2837-Rev	GTGATAAGGTAAAGAAGC	49.1	2837	
	2982-Rev	ATAAAGCAGGCTTGTTAG	52.4	2982	
NS3	F4 Rev	CTGGGTAGCCGTAGAAAGCACCT	58	3558	(Gaudieri et al. 2009)
	3904-Rev	GCTACACCTCTGGTGCACACAGC	66.0	3904	(Humphreys et al. 2009)
	F5 Rev	TAGAATGTGGCACAGTGATGCTGC	58	4438	(Gaudieri et al. 2009)
	4880-Rev	ACCGAGTCAAACATYCCAGACG	61.2	4880	In house
	F6 Rev	GCCATGATGTATTTTGTGATGGGGTGTG	62	5297	
NS4b	5713-Rev	TGAACGCCATAAGAGACGCCACAGC	66.3	5713	
	F7 Rev	GTGACAGTTAGAGAACTCAGCAATG	57	6218	
NS5a	6684-Rev	TTGGCACGGACACTTGAGC	58.8	6684	
	6823-Rev	TCCGGTTCGGGCTCRCANG	63.1	6823	
	F8 Rev	CACAACCTTTGTTTCAGACTCCACCCG	61	7110	
	7578-Rev	TGTCACTAACGGTGGACCAAGAGTCG	66.4	7578	
NS5b	F9 Rev	GTTCTTCGCCATGATGGTGGTTGGAAT	60	8055	(Gaudieri et al. 2009)
	8443-Rev	CAACGGCGATAACCACACTGGGC	66.0	8443	(Humphreys et al. 2012)
	F10 Rev	TGCCCGATGTCTCCAAGCTCGTA	59	9141	(Gaudieri et al. 2009)
	9192R	GGAGTGAGTTTGAGCTTGGT	57.3	9192	(Lucas et al. 2007)
3' UTR	9550-Rev	AGGTTATTGGGAAGTTTGGGAAGG	61.0	9550	In house
	H3'X45	ACAGCTAGCCGTGACTAGGGC	63.7	9600	(Yanagi et al. 1997) adapted

2.2.5 Second Round PCR

Second round PCR was performed using 2.5 μls of cDNA template generated by Reverse transcriptase One-Step PCR reaction. A PCR reaction mix was prepared using the Expand High Fidelity PCR System (Roche) according to manufactures instructions, containing sdH_2O, dNTP (25 mM each, Bioline), forward and reverse primers (10 μM each), buffer and Taq DNA polymerase High Fidelity enzyme. Primers were designed to amplify the cDNA template in multiple nested PCR reactions of approximately 1kb each (Figure 2-1), with thermocycler programs dependent on the primer combinations used for individual reactions (Core: 277-For/977-Rev; E1: 745-For/1585-Rev; E2: 1435-For/2982-Rev; F4 – F10 primer pairs, see Table 2-1 and Table 2-2.

Figure 2-1: Strategy for HCV sequencing reactions



The strategy of HCV sequencing reactions is shown: First, HCV RNA is transcribed into HCV cDNA in two segments (either 4kb or 7kb), and then further amplified in short segments for sequencing.

2.2.6 Gel electrophoresis and extraction of PCR products

All second round PCR products were visualized on a 1% agarose gel and products of the correct size were gel extracted and purified using Qiagen Gel Extraction kit (Qiagen). As specified in the manufacturers instructions, 500 μls of Buffer QG were added to the extracted gel fragment and heated to 52°C for 10 mins to dissolve the gel. The mixture was added to an enclosed spin column and spun at 1300 rpm for 1 min

(this step was repeated if necessary). 500 µls of Buffer QG were added to the spin column to remove remainders of the gel and spun at 1300rpm for 1min. 750 µls of wash buffer PE were added to spin column, incubated at room temperature for 2-5 mins and spun at 1300 rpm for 2 mins. PCR products were eluted into 50 µls of sdH_2O and sequenced immediately or stored at -20°C until needed.

2.2.7 Sequencing and DNA precipitation

PCR products were sequenced on an ABI 3100 DNA automated sequencer bidirectionally with 2nd round PCR primers using Prism Big Dye (Applied Biosystems). Additional inner primers were added to ensure adequate coverage. After the sequencing reaction, PCR fragments were precipitated with sodium acetate (final concentration 10 mM, Sigma) in 80 % ethanol for 30 min. Samples were spun for 1 hour at 4000 rpm at 4°C and then washed twice with 70% ethanol. Plates were dried at 95°C for 1 min to evaporate any remaining ethanol.

2.2.8 Sequence analysis

Sequences were edited using the Sequencher 4.8 Software (Gene Codes Corp.), and aligned using Se-Al (http://tree.bio.ed.ac.uk). Sequence entropy was calculated using the Shannon information entropy score (http://evolve.zoo.ox.ac.uk/software, Shannon 1948).

2.2.9 DNA Purification

DNA was extracted using the DNeasy Blood And Tissue Kit (Qiagen) from PBMCs or from whole blood using the Gentra Puregene kit (Qiagen) following the instructions and subsequently used for HLA and IL28B typing.

2.2.10 IL28b genotyping using real time PCR

IL28B rs8099917 typing was performed as previously described (Bucci et al. 2013). In brief, the rs8099917 SNP was genotyped using a commercially validated TaqMan SNP genotyping assay (Applied Biosystems; assay ID_11710096_10). 1-20 ng/ml of DNA was used in an LC480 PCR thermocycler (Roche) and each sample was run in triplicate using internal positive and negative control samples. The genotypes at this SNP are GG, GT or TT. The homozygote TT allele as been associated with increased viral clearance in acute HCV infection and a favourable response to IFN/RBV treatment in comparison to the G allele.

2.2.11 HLA-typing

HLA typing was performed at the Transplant Immunology Laboratory in collaboration with Martin Barnado (Oxford Radcliffe Hospitals).

2.3 T cell assays

2.3.1 Isolation of PBMCs from whole blood

PBMC were isolated from whole blood by density gradient centrifugation (Lymphoprep, Stemcell technologies) at 2500g for 24 minutes at room temperature without brakes. The layer containing PBMC was removed carefully and placed in RPMI media containing penicillin (100U/ml), streptomycin (100µg/ml) and L-glutamine (2mM) and washed 3 times. Cells were resuspended in 10mls R10 media (RPMI media containing 10% heat-inactivated fetal calf serum (FCS), penicillin (100U/ml), streptomycin (100µg/ml) and L-glutamine (2mM)) and counted using Quava Personal Cell Analyser (Guava Technologies Inc., California) or Scepter Automated Handheld Cell Counter (Millipore).

2.3.2 Freezing and thawing of cells

Cells were frozen at 10 million PBMC per vial in a freezing mix containing 500 µl FCS, 400 µl RPMI and 100 µl DMSO. All reagents were chilled on ice, filled vials placed in a chilled isopropranolol container immediately and frozen in a -80°C freezer. Cells were transferred into liquid nitrogen the following day.

For thawing, RPMI containing penicillin (100U/ml), streptomycin (100µg/ml) and L-glutamine (2mM), R10 and CTL supplement wash (GibcoBRL) were warmed to room temperature. Thawing media was prepared by diluting CTL wash supplement 1:10 in RPMI medium, adding DNase (Roche, 50U/ml). Vials were thawed at 37°C in a water bath and immediately transferred into a 15ml falcon tube. Thawing media was added gradually to avoid osmotic shock. Subsequently, cells were washed in RPMI twice and resuspended in R10.

2.3.3 Interferon-γ ELISpot assays

T cell responses were tested by Interferon-γ (IFNγ) ELISpot assays as previously described (Humphreys et al. 2012). 96-well ELISpot plates (Multiscreen HTS, Millipore) were coated with 100 µl of mouse anti-human IFN-γ monoclonal antibody (Mabtech; final concentration 5µg/ml) in PBS. Plates were incubated overnight at 4°C.

The next day, plates were washed 4x with 200μl/well with sterile PBS; 200 μl R10 was added to each well and the plates incubated at 37°C and 5% CO_2 for 2 hours. 50 μl of antigen was added to each well; as HCV peptides (final concentration 3 μg/ml, Mimotopes, BEI peptides and Proimmune), cytomegalovirus lysate (CMV, final concentration 0.05 μg/ml, Chiron), concanavalin-A (ConA, final concentration 5 μg/ml, Sigma), Influenza, Epstein Barr virus and CMV CD8+ epitopes (FEC, 3μg/ml final concentration, BEI resources) and DMSO (3μg/ml final concentration, Sigma), which served as a negative control. PBMC were prepared at $4x10^6$/ml and 50μls plated in duplicate to a final concentration of 200,000 cells/well. Plates were incubated at 37°C and 5% CO_2 for 18-20 hours.

ELISpot plates were washed 7 times with phosphate buffered saline (PBS, GibcoBRL)/Tween20 (Sigma) (200 μl/well, Sigma). 50 μ of biotinylated mouse anti-human IFN-γ antibody (Mabtech) diluted 1:2000 in 5% Bovine serum albumin (BSA, Sigma)/PBS was added to each well and incubated at room temperature for 3 hours. Plates were washed 4x. 50 μl alkaline phosphatase-conjugated anti-biotin antibody diluted 1:750 in 5% BSA/PBS was added to each well and incubated at room temperature for 2 hours. Plates were washed 4x, and 50 μl of BCIP/NBT substrate solution (Pierce) was added to each well and incubated at room temperature for 7 mins, when plates were rinsed with deionized water to stop the reaction. Plates were left to air dry.

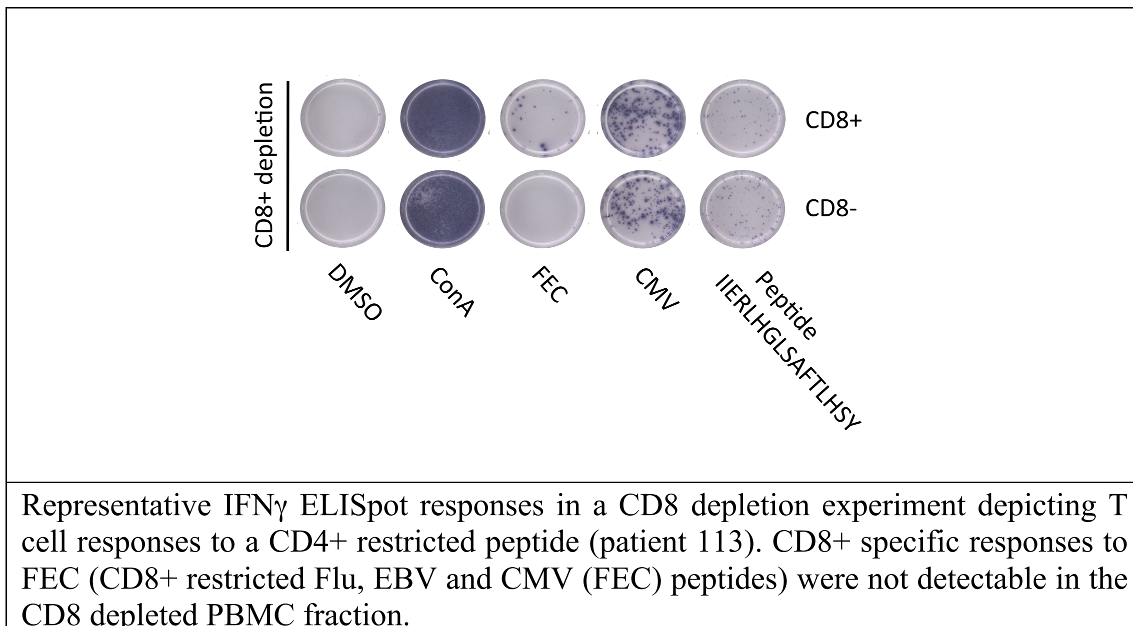
Spot-forming units (SFU) were counted using an automated ELISpot plate reader (AID). To define a positive cut off for the assay, ELISpot assays using peptides for both genotypes were run in 12 healthy subjects. The cut off was based on the following calculation; mean number of SFU/10^6 PBMC in test wells minus negative control plus 3x standard deviation: Positive responses were counted as >40 SFU/10^6PBMC for HCV genotype 3a peptides, and >43 SFU/10^6PBMC for HCV genotype 1b peptides.

The pass/fail criteria for ELISpot assays stated that the negative control wells should ideally be 0 spots/10^6PBMC, but always have less than 10 spots per well and in wells containing only R10 (no cells) should have less than 5 spots per well. For positive controls (PMA) there should be greater than 20 spots per well (100SFU/10^6PBMC). If any of these criteria were not met, then the plate was failed and repeated.

2.3.4 CD8 depletion of whole PBMCs

CD8 cells were depleted from whole PBMC using magnetic bead separation (CD8 Dynabead, Invitrogen) following manufacturers instructions. In brief, for each sample (5 million PBMC) 50 µl of beads were washed 3x in chilled PBS containing 2% FCS to remove any traces of sodium azide. Cells were incubated with 25 µl of beads for 30 min in chilled PBS/2% FCS on a MACS mixer at 4°C, after which beads with bound CD8 positive cells were removed using a magnet. This incubation step was repeated. To ensure complete removal of CD8 positive cells, Influenza, Epstein Barr virus and CMV CD8+ epitopes (FEC, 3µg/ml final concentration, BEI resources) were included as controls in subsequent ELISpot assays on full PBMC and the CD8 depleted fraction (Figure 2-2).

Figure 2-2: CD8 depletion ELISpot



Representative IFNγ ELISpot responses in a CD8 depletion experiment depicting T cell responses to a CD4+ restricted peptide (patient 113). CD8+ specific responses to FEC (CD8+ restricted Flu, EBV and CMV (FEC) peptides) were not detectable in the CD8 depleted PBMC fraction.

2.3.5 T cell lines (for intracellular staining)

PBMCs (5×10^6) were stimulated with 2.5 µg/ml antigen supplemented with 50 IU/ml recombinant IL-2 on day 2, 5, and 8. After 10–14 days of culture, T cell lines were rested for 24 hours in R10 without IL-2 and subsequently used for intracellular staining assays.

2.3.6 Proliferation assays

In vitro proliferation assays were performed on freshly isolated PBMCs plated in triplicate at 2×10^5 PBMCs per well with conventional [^{3}H] thymidine incorporation

and HCV antigens (1 µg/ml) (Mikrogen). Data are displayed as SI (stimulation index; fold change above background). A positive response is defined as SI ≥ 3.

2.3.7 Fluorescence-activated cell sorting (FACS) analysis

2.3.7.1 Intracellular staining assays

Intracellular cytokine stains were performed as previously published (Barnes et al. 2012). In brief, rested thawed PBMCs were stimulated with single peptides or peptide pools or unstimulated (controlled for DMSO) or PMA/ionomycin. After overnight stimulation (Brefeldin A (final concentration 10µg/ml) was added after 2 hour at 10 mg/ml), cells were permeabilized (BD Perm) and stained with the following antibodies: Near-Infrared Life/Dead, CD3-PO, CD4–Qdot 605, CD8–PB, IFN-γ–Alexa Fluor 700, IL-2–APC (allophycocyanin), TNFα–PE–Cy7, and CD107α–PECy5, and Mip1b–PE. Flow cytometry was performed with a BD LSRII.

2.3.7.2 HLA class I Pentamer assays

HLA class I pentamer analyses were performed as previously published (Barnes et al. 2012), PE-labeled pentamers loaded with HCV NS3 (ATDALMTGY, HLA-A*0101) were obtained from ProImmune. The cells were co-stained with combinations of the following antibodies: CD3-PO, CD8-PB, CCR7–PE-Cy7, CD45RA-FITC (fluorescein isothiocyanate), CD127-APC, CD38–PerCP-Cy5.5, HLA-DR–Alexa Fluor 700, perforin-FITC, GzB–Alexa Fluor 700, GzA–PerCP-Cy5.5, CD161-APC, and PD-1–PECy7. For details on phenotypic characteristics of stained markers refer to Table 2-3. Flow cytometry and analysis were performed as above.

Table 2-3: Characteristics of phenotypic markers analysed by HLA class-I tetramer staining.

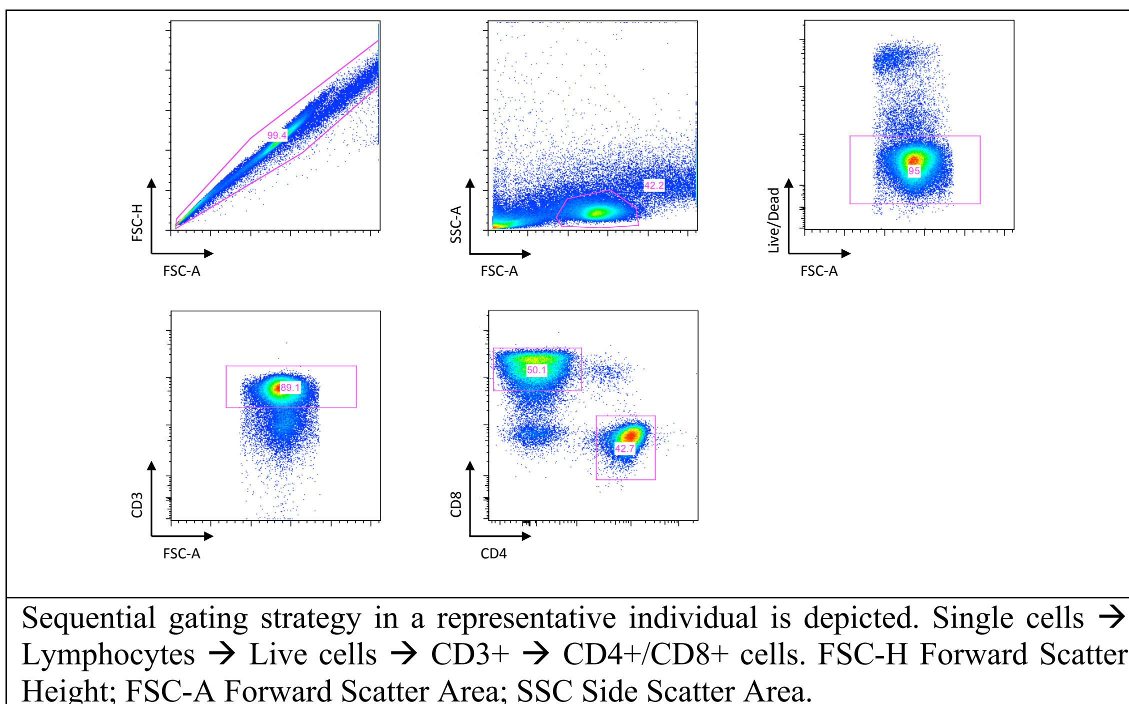
	PECy7 1:50	FITC* 1:50	APC 1:50	PERCPCy5.5§ 1:50	Alexa700 1:50
Panel 1	CCR7	CD45RA	CD127	CD38	HLA-DR
	Central memory, LN homing	Antigen exposure, Re-expressed on terminally differentiated CD8 T cells	Memory and antigen exposure	Activation	Activation
Panel 2	PD-1	Perforin	CD161	Granzyme A	Granzyme B
	Exhaustion or activation	Cytotoxicity	Liver homing	Cytotoxicity	Cytotoxicity

*FITC = fluorescein isothiocyanate; §PerCP = peridinin chlorophyll protein

2.3.8 Data analysis

Data analysis was performed using the FlowJo software (TreeStar). Samples were gated sequentially (see representative gating strategy in Figure 2-3). Analysis of polyfunctionality of T cells was performed with SPICE (Roederer et al. 2011).

Figure 2-3: Representative gating strategy for analysis of FACS samples



Sequential gating strategy in a representative individual is depicted. Single cells → Lymphocytes → Live cells → CD3+ → CD4+/CD8+ cells. FSC-H Forward Scatter Height; FSC-A Forward Scatter Area; SSC Side Scatter Area.

2.4 Peptides sets

2.4.1 Subtype 3a screening peptides

2.4.1.1 HCV gt3 overlapping peptides spanning the whole HCV genome

A previously described HCV genotype 3a specific peptide set was used (Humphreys et al. 2012). To generate this set, full-length viral sequencing was performed (amino acids 1-2929) on 18 genotype-3a patients with chronic infection (Accession numbers: GQ356200-GQ356215, GQ356217, JF509175-JF509177). Peptides were designed to be 15-19 amino acids in length in order to detect both CD4+ and CD8+ HCV specific T cells, overlapping by 11 amino acids (n=460, Mimotopes, Australia). Subtype 3a peptides with amino acids shown to affect peptide presentation at the C-terminus (glycine, proline, glutamic acid, aspartic acid, glutamine, asparagine, threonine, serine, cysteine) were either shortened or lengthened until a tolerated amino acid occurred. If after shortening or lengthening the C-terminal position was still occupied by an non-

tolerated amino acid, the original 18-mer peptide was retained. Peptides were grouped into 10 pools corresponding to the individual viral proteins as shown in Table 2-4, with 29 to 72 peptides in each pool. Peptides were tested at final concentration of 3ug/ml for each single peptide.

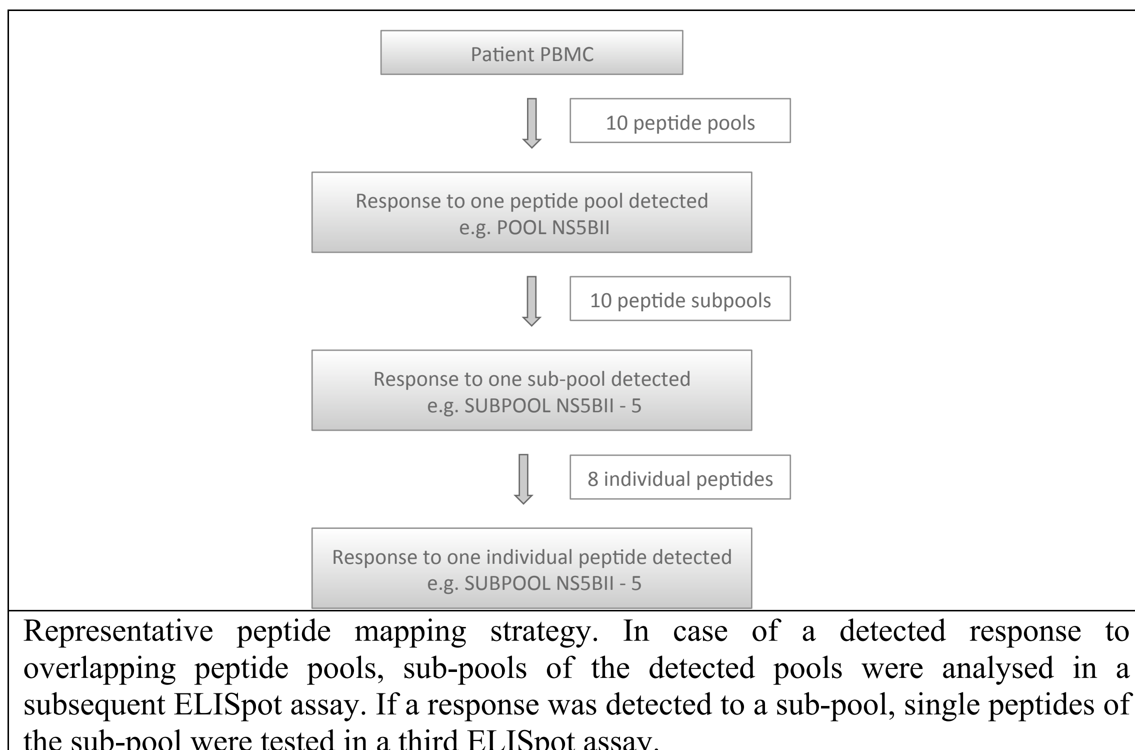
Table 2-4: HCV genotype 3a specific peptides - overlapping peptide pools

Peptide Pool	Amino acids (aa)	Number of peptides
Core	1-191	29
E1	192-383	28
E2	384-752	57
p7 & NS2	753-1032	41
NS3 protease	1033-1359	49
NS3 helicase	1349-1663	48
NS4	1664-1978	46
NS5A	1979-2430	72
NS5B I	2431-2726	45
NS5B II	2716-3021	45

2.4.1.2 Peptide screening strategy

If a response to a peptide pool was detected, responses were mapped to single peptides following the peptide mapping strategy as depicted below (Figure 2-4). First, subpools of the pool with the detected response were assessed using ELISpot assays, and then single peptides of the responding subpool were further analysed.

Figure 2-4: Peptide mapping strategy



Representative peptide mapping strategy. In case of a detected response to overlapping peptide pools, sub-pools of the detected pools were analysed in a subsequent ELISpot assay. If a response was detected to a sub-pool, single peptides of the sub-pool were tested in a third ELISpot assay.

2.4.2 HLA predicted peptides spanning HCV non-structural regions

A HLA predicted peptide set was based work done by the Gaudieri group in Perth (Royal Perth Hospital, Murdoch University, Perth, Australia) and Isla Humphreys in our group.

2.4.2.1 Identification of polymorphic sites associated with HLA alleles

This work was done by Silvana Gaudieri's group (Rauch et al. 2009): HCV genotype 3a specific HLA class I restricted peptides were based on a novel, sequence-led approach (Rauch et al. 2009). Polymorphic sites in HCV non-structural regions (NS2-NS5B) in a sequence data set comprising 136 HCV genotype 3a sequences were identified and assigned as consensus or non-consensus. Polymorphisms were associated with patients' HLA alleles using Fisher's exact test. For a detailed discussion of statistical methods used, refer to the Discussion of Chapter 3 (page 137).

2.4.2.2 Evaluation of putative CD8 HLA associated epitopes

This work was done by Isla Humphreys: Viral regions containing polymorphic sites were assessed using online epitope prediction servers to identify putative epitopes. These epitope prediction servers predict possible binders for certain given HLA types, with polymorphic regions within or flanking the predicted epitope.

For this analysis, two independent epitope prediction servers were used: BIMAS (http://www-bimas.cit.nih.gov/molbio/hla_bind/), and Syfpeithi ((Rammensee et al. 1999), http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm).

A cut off value of ≥20 (Syfpeithi) and ≥50 (BIMAS) was determined by assessing known HCV epitopes using Syfpeithi. Using these cut-offs, 65 putative T cell epitopes were predicted, with 55 containing the polymorphic site within the epitope (Table 2-5), and 10 additional putative T cell epitope were predicted with the polymorphic site flanking the epitope (Table 2-6). These peptides were subsequently made by Proimmune and used in IFNγ ELISpot assays matched to the patients HLA types.

Table 2-5: HLA-associated peptides – polymorphisms within predicted epitope

Protein	HLA	Wildtype	Variant	Amino Acids	Peptide	Predicted Epitope
NS2	A0101	V	A	879-887	001-A0101	**V**ILLTSLLY
	A02	L	P	862-871	049-A02	ALQVWVPPL**L**
		L	P	870-879	048-A02	L**L**ARGSRDGV
		Y	H	881-890	052-A02	LLTSLL**Y**PSL
		Y	H	885-894	047-A02	LL**Y**PSLIFDI
	A0301	V	A	926-935	006-A0301	RLCMLVRS**V**M
		V	A	929-938	004-A0301	MLVRS**V**MGGK
		V	A	930-939	005-A0301	LVRS**V**MGGKY
	A2402	Y	H	886-894	007-A2402	L**Y**PSLIFDI
	B1501	T	A	878-887	008-B1501	GVILL**T**SLLY
	B15	I	V	946-954	055-B15	S**I**GRWFNTY
	B2705	R	K	940-949	010-B2705	FQMIILSIG**R**
		R	K	948-956	009-B2705	G**R**WFNTYLY
		H	Y	962-971	011-B2705	MQ**H**WAAAGLK
	B4402	I	V	822-831	012-B4402	ATLGAG**I**LVL
	B44	I	V	826-835	057-B44	AG**I**LVLFGFF
	B5101	S	G	871-880	013-B5101	LARG**S**RDGVI
	C03	L	M	829-838	050-C03	I**L**VLFGFFTL
	C04	V	I	981-990	051-C04	IFSPMEIK**V**I
NS3	A0101	Y	F	1442-1450	014-A0101	ATDALMTG**Y**
	A0201	A	D	1389-1398	015-A0201	**A**LLKGGRHLI
	A0301	V	I	1138-1146	016-A0301	L**V**TRDADVI
	B1501	K	R	1296-1305	017-B1501	**K**LTYSTYGKF
	B2705	L	I	1379-1388	018-B2705	IPFYGKAIP**L**
		V	I	1632-1641	019-B2705	YRLGP**V**QNEI
	B4402	G	S	1407-1416	020-B4402	DEIASKLR**G**M
	B4403	L	S	1639-1647	021-B4403	NEIC**L**THPI
	B5101	A	D	1388-1397	023-B5101	I**A**LLKGGRHL
NS4B	A02	A	T	1873-1882	062-A02	KIMGGELPT**A**
		A	T	1880-1889	065-A02	PT**A**EDMVNLL
	A0301	I	V	1901-1910	026-A0301	GV**I**CAAILRR
	A2601	R	K	1852-1861	027-A2601	**R**VLLDILAGY
	A68	A	I	1736-1744	067-A68	EK**A**LGLLQR
	B27	R	K	1948-1957	061-B27	A**R**VTALLSSL
	B4001	I	V	1847-1855	028-B4001	G**I**GLGRVLL
	B51	A	T	1879-1888	066-B51	LPT**A**EDMVNL
		A	T	1881-1889	064-B51	T**A**EDMVNLL
	C0401	V	I	1733-1742	025-C0401	QFKEK**V**LGLL
NS5A	A0201	G	S	2321-2330	029-A0201	ALPPR**G**APPV
	A0301,	V, P	A, S	2382-2391	030-A0301	K**V**PPS**P**GGES
	A2601	D	G	2268-2276	033-A2601	ET**D**AELSVA
	A68	T	S	1989-1998	070-A68	WVC**T**VLSDFK
	B0702	V	E	2309-2317	036-B0702	APDY**V**PPTV
		V	E	2313-2322	035-B0702	**V**PPTVHGCAL
		T	A	2332-2341	034-B0702	PPRRKR**T**IQL
	B0801	I	V	2251-2259	037-B0801	ESETKVV**I**L
	B44	R	Q	2097-2105	071-B44	VEVR**R**VGDF
NS5B	A0201	T	I	2489-2498	038-A0201	VLDDHYK**T**AL
		N	D	2540-2549	040-A0201	SLSSKAI**N**QI
		N	D	2544-2552	039-A0201	KAI**N**QIRSV
	A1101	R	K	2500-2509	041-A1101	EVKERAS**R**VK
	A2601	K	R	2537-2545	042-A2601	DVRSLSS**K**A
	B0801	I	M	2507-2515	043-A2601	RVKARMLT**I**
	B1501	Q	L	2476-2484	044-B1501	S**Q**RQKKVTF
	B5101	K	R	2474-2482	045-B5101	SASQRQ**K**KV

Table 2-6: HLA-associated peptides – polymorphics flanking predicted epitopes

Protein	HLA	Polymorphic Site	Peptide	Amino Acids	Predicted Epitope and Wild-type Residue ()	Variant Residue
NS2	A0201	943	002-A0201	934-942	VMGGKYFQM(**I**)	V
		943	003-A0201	944-953	(**I**)ILSIGRWFNT	V
	A02	871	054-A02	861-870	AALQVWVPPL(**L**)	P
		887	053-A02	878-886	GVILLTSLL(**Y**)	H
NS3	A68	1046	058-A68	1047-1056	(**L**)GTIVTSLTGR	F
	B5101	1389	022-B5101	1379-1388	IPFYGKAIPI(**A**)	D
		1389	024-B5101	1390-1398	(**A**)QLKGGRHLI	D
NS5A	A1101	2383	031-A1101	2374-2382	DTQSSTTSK(**V**)	A
	A2402	2148	032-A2402	2138-2147	RYAPPCKPLL(**D**)	E
NS5B	B5101	2855	046-B5101	2846-2854	APTIWVRMV(**M**)	L

Polymorphic site flanking epitope is indicated in bold within brackets ().

2.4.3 Genotype 1 screening peptides

HCV genotype 1 screening peptides were based on the HCV gt1b sequence J4 and HCV gt1a sequence H77 and obtained from BEI resources (www.beiresources.org). Similar to the HCV gt3 overlapping peptide set, peptides were 15 to 19 amino acids in length, and overlapped by 11 amino acids. HCV gt1 peptides were adapted to remove any un-tolerated amino acids at the C terminus similarly to the HCV gt3 peptide set, and used in 10 peptide pools at a peptide concentration of 3ug/ml for each peptide.

2.5 HLA restriction experiments

2.5.1 B cell lines

B cell lines (BCL) were transformed from fresh PBMC. 5 million PBMC/well (24 well plate) were incubated in 3ml containing 1ml Epstein-Bahr-Virus stock (EBV, grown from EBV infected B95.8 cell line, Sigma Aldrich), 1ml FCS and 1ml R10 containing Cyclosporin A (CSA, final concentration 0.1 μg/ml, Sigma Aldrich) for 7 days without disturbing the cultures. Subsequently, half of the media in each well was replaced with fresh R10 every 3-4 days and cultures were grown for 3-6 weeks until dense yellow colonies on the plate were seen by eye, and brown masses of cells observed under the inverted microscope.

2.5.2 T cell lines (for HLA restriction experiments)

5 million PBMCs per well (24 well plate) were stimulated with 2.5 μg/ml antigen in R10, cultures were supplemented with 50 IU/ml recombinant IL-2 (every 3-4 days,

starting on day 2). After 21 days of culture, T cell lines were rested for 24 hours in R10 without IL-2 and subsequently used for restriction experiments.

2.5.3 HLA restriction experiments

HLA restriction experiments were performed as previously described (Payne et al. 2010). Briefly, Epstein-Barr-Virus (EBV) transformed B cell lines (BCL) were matched to one HLA class-I type of the CD8+ T cell lines being tested. B cell lines were incubated with the peptide of interest for 1h at 37°C and then washed 5x before being incubated with the according T cell line for 6 hours. Brefeldin A (final concentration 10 μg/ml) was added after 1 hour of incubation. Unstimulated BCLs, as well as unstimulated (controlled for DMSO), peptide and phorbol 12-myristate 13-acetate (PMA)/ionomycin (50 and 500 ng/ml, respectively) stimulated TCLs were used as controls. Cells were permeabilized and stained with Near-Infrared Life/Dead, CD3-PO, CD8–PB, IFN-γ–PeCy5.5, TNFα–PE(phycoerythrin)–Cy7, CD19-APC-Cy7 and CD69-FITC. Flow cytometry was performed with a MACSQuant® Analyser (Miltenyi) and analysis by FlowJo (TreeStar).

2.6 Statistical and data analysis

2.6.1 Analysis of epitopes described in natural HCV infection (IEDB)

We obtained epitopes described for HCV gt1 and gt3 from the immune epitope database resource (IEDB, www.iedb.com, search date 19/02/2014). 8587 T cell epitopes were described for HCV gt1, and 218 epitopes for HCV gt3. To ensure data quality, epitopes were crosschecked with the primary publications. Epitopes described in non-human organisms were excluded, as well as duplications of epitopes and variants described in the same publication (included in the IEDB as duplicates due to multiple performed assays with the same epitope). For several publications, inaccurate information was included in the IEDB, e.g. full screening peptide sets were described as detected positive epitopes; these peptides were excluded from the dataset. Refined epitope lists were included in two separate analyses in this thesis; (i) a comparison of cross-reactivity between T cell epitopes described in HCV gt3 infection in the Oxford cohort and epitopes from the IEDB in Chapter 3 (3.9.1, page 121), or (ii) in an analysis of potential immunogenicity of HCV conserved sequences in Chapter 5 (5.5.1, page 228).

(i) Comparative analysis of HCV gt3a epitopes from the Oxford cohort with epitopes previously described in the literature.

Comparative Overlap between an epitope described in the Oxford cohort and a published epitope from the IEDB was defined as sequence overlap of more than 3 amino acids. If the CD4/CD8 restriction was defined for the epitopes defined in the Oxford cohort, published epitopes for the according restriction (CD4 or CD8) and epitopes with non-defined restriction were included in the comparison. If the CD4/CD8 restriction for an epitopes defined in the Oxford cohort was not defined, references for both CD4 and CD8 epitopes were included in the analysis. In cases where the CD4/CD8 restriction was not defined in the published study referenced in the IEDB, the epitope was included for all subgroups (CD4/CD8 restricted epitopes, and epitopes without defined restriction).

2.6.2 Statistical methods

Nonparametric tests were used throughout, paired for within-individual comparisons (Wilcoxon) and unpaired for group comparisons (Mann-Whitney). T cells targeting structural and non-structural viral genomic regions in genotypes-1 and -3 HCV infected patients (Chapter 3) were assessed using Fisher's exact test. Comparisons between T cell responses over time (Chapter 3 and 4) were assessed using paired t-test. A p value <0.05 was considered significant. Prism (v4.0 for Mac) was used throughout.

2.7 Work done by others

Isla Humphreys designed the HCV genotype 3a peptide sets, performed full-length HCV sequencing of 15 HCV gt3 infected patients to generate a consensus sequence for the overlapping peptide set, and performed IFNγ ELISpot assays screening in 18 chronic patients using HLA predicted peptides and in 12 patients using overlapping peptides, as well as selected cross-reactivity assays (Table 3-15, Figure 8-1). Anthony Brown performed the IFNγ ELISpot assays on HCV genotype 1 patients (Figure 3-8, Figure 8-2). Rachel Townsend performed IL28B typing assays. Clinical care of patients and sample collection was provided by the clinical Hepatology team at the John Radcliffe Hospital, Oxford. Work described in section 5.3 (page 214) was done in collaboration with Jose Lourenco (Zoology, Oxford).

3 Identification of HCV specific T cells responses in HCV genotype 3a infection

3.1 Abstract

Background: Defining protective immunity across all Hepatitis C virus (HCV) genotypes will facilitate pan-genotypic T cell vaccine development. HCV subtype 3a (gt3a) is now the commonest infecting genotype in the United Kingdom and large parts of Asia; however, data regarding the T cell antigenic targets of this subtype is currently very limited. The aims of this study were to define T cell targets in acute, chronic and spontaneously resolved HCV gt3a infection using two complementary peptide sets.

Methods: We used two different approaches to identify T cell targets in acute (n=16), established chronic (n=108) and spontaneously resolved (n=17) HCV gt3a infection: (i) overlapping peptides (spanning the whole viral genome, 15-18 amino-acids (AA), overlapping by 11 AA), based on a consensus sequence derived from 18 full-length gt3a sequences and (ii) a novel, sequence-led approach using 46 wild-type and variant peptides (non-structural proteins, 9-10 AA) corresponding to putative HLA class-I restricted epitopes under T cell selection.

Results: Using two parallel methodologies we have identified multiple new T cell epitopes in gt3a infection. Overall, the two peptide sets used in this study are complementary and identify distinct gt3a T cell targets. In acute infection multiple CD8+ T cell epitopes were identified that predominantly targeted non-structural proteins. In chronic infection, T cell responses mainly targeted two dominant epitopes in the core (CD4+) and NS3 (CD8+) region. In contrast, resolved infection is associated with broad distribution of strong CD4+ and CD8+ T cell responses across the genome that may confer protection. T cell targets in HCV gt3 infection were markedly different to those previously described for HCV gt1 infection, with immunodominant epitopes detected in HCV gt1 infection not targeted in HCV gt3 infection. The magnitude of adaptive immune responses was not associated with age, sex, IL28B genotype and treatment outcome.

Conclusions: We have identified multiple new T cell targets specific for HCV gt3 infection. These observations will inform rational vaccine design for HCV gt3a infection and cross-reactive vaccine development.

3.2 Background

Phylogenetic analysis has shown that HCV has existed in human hosts for thousands of years, leading to distinct genotypes endemic in different geographic locations that only share 70-80% of sequence homology (Pybus et al. 2009). HCV exists in 6 major infecting subtypes and also displays significant variation within each host. One of the major challenges in HCV vaccine development for HCV is the extensive diversity of the virus: a better understanding of genotype specific immune responses will aid the development of vaccines active against multiple genotypes.

HCV genotype 3a is now the major infecting subtype in the United Kingdom, and particularly prevalent within drug using communities (Health Protection Agency 2013). This subtype is also endemic in parts of Asia and Western Europe. However, data on HCV gt3a antiviral immune responses is currently very limited, and will have to be further assessed for informed vaccine design for this subtype. A comprehensive assessment of HCV gt3a specific anti-viral immunity may also lend insights into the distinct clinical phenotype associated with gt3.

3.2.1 Evidence for distinct features of HCV genotype 3 infection

The classification of HCV by viral genotype has been proven to be highly informative, not only in terms of the assessment of viral evolution, but also in distinguishing treatment outcome and duration of therapy. For a detailed discussion of clinical and immunological features specific to HCV genotype 3 refer to the introduction (1.6, page 57).

Large randomized studies have shown that HCV gt3 has a high rate of spontaneous clearance (M. Lehmann et al. 2004), a more favourable outcome of Interferon based treatments compared with gt1 (Fried et al. 2002), a high rate of viral relapse after interferon/ribavirin therapy (Manns et al. 2001) and higher rates of hepatic steatosis and fibrosis (Adinolfi et al. 2001; P.-Y. Bochud et al. 2009). In addition, a striking association of the IL28B genotype with viral clearance and treatment response was shown in HCV gt1 (D L Thomas et al. 2009; Ge et al. 2009), but not gt3 infection. However, higher rates of liver injury have been found in chronically infected patients carrying the favourable IL28B genotype, a phenomenon only observed in HCV gt3 infected patients(Moghaddam et al. 2011; Rembeck et al. 2012; Ydreborg et al. 2013). Regarding the adaptive immune response to HCV genotype 3 infection, data remain scarce. Several strands of evidence point to a substantial difference between HCV

genotype 1 and 3 specific T cell responses, including an analysis of sequence polymorphisms in HCV gt1 and gt3 infection in association with HLA types (Rauch et al. 2009), discussed in detail below. A single study analysed HCV gt3 specific T cell responses to the HCV NS3 region in HCV gt1 and gt3 infection, observing different T cell targets for both genotypes (Giugliano et al. 2009). However, T cell specificity to the whole HCV genome remains unknown.

3.2.2 A novel peptide screening approach to identify HCV gt3 specific T cell targets

Identification of viral T cell targets is often based on IFNγ ELISpot assays using overlapping peptide pools. However, there are several issues with testing large numbers of peptides in overlapping pools for a comprehensive screening approach. These include:

(1) Competition for binding on MHC complexes: there may be potential competition for binding space at MHC complexes with other peptides specific for the same HLA type.

(2) Peptide length: It was previously shown that the peptide length matters for both presentation on MHC complexes and recognition by T cells (Falk et al. 1991; Rock et al. 1992). In our experiments, overlapping peptide pools contained peptides 15 to 18 amino acids in length, ensuring peptide presentation on MHC class-II complexes (peptides up to 15mers are presented). However, the optimal presentation length on MHC class-I complexes ranges between 8 and 11 amino acids. Therefore, overlapping peptides may not be of the right binding length, and an amino acid crucial for presentation might be at the wrong position preventing peptide loading on MHC complexes. Consequently, T cell responses may be missed with a screening approach using overlapping peptides. Theoretically, one would expect a certain amount of peptide degradation in each sample associated with smaller peptides, which might aid peptide presentation.

(3) Peptide identification: To cover the whole genome of a virus, several hundred peptides are necessary, which are grouped into pools for the initially screening. Once a response to a pool is detected, responses are mapped to individual overlapping peptides, a laborious, time consuming and resource intense step (Precopio et al. 2008).

To avoid these constraints, alternative approaches have been developed. One previously applied approach has been based on the identification of T cell targets using sequence-led approaches that associate viral sequence polymorphisms with patient HLA types, based on the rationale that immune pressure exerted by the T cell responses leads to polymorphic sites within HCV sequences. This approach has been used to identify T cell epitopes associated with HLA in well-defined cohorts of women infected with HCV from a single source (Neumann-Haefelin et al. 2006; Fitzmaurice et al. 2011).

In this thesis, I sequence-led approach was used to identify new HCV gt3 specific T cell targets (termed "HLA predicted peptides"). The approach introduced above entirely focused on HLA types linked to protection, whereas here an approach screening all patient HLA types for HLA-associated sequence polymorphisms was deployed.

Identification of polymorphic sites associated with HLA types: HLA predicted peptides were developed as a collaborative effort between Silvana Gaudieri's laboratory and the Barnes group (Rauch et al. 2009). Viral sequence polymorphisms were associated with patients HLA types in a cohort of 136 HCV gt3 infected patients recruited from the UK, Switzerland and Australia. Viral polymorphic sites at each amino acid were assigned as consensus and non-consensus based on amino acid frequency, and associations with patient HLA class-I types were assessed using Fisher's exact test at each residue of the non-structural HCV proteins. Significant associations with an odds ratio >1 were considered as mutational escape from HLA restricted cellular immune pressure; meaning that viral polymorphisms abrogating HLA restricted CD8+ T cell responses are overrepresented in the presence of the restricting HLA allele.

The identification of putative T cell epitopes at viral polymorphic sites associated with HLA type: This work was performed by Isla Humphreys in our laboratory. Viral polymorphic sites associated with patient HLA types were evaluated regarding putative T cell epitopes using online epitope prediction programs. Freely available online prediction programs Syfpeithy and BIMAS were used, and a prediction score over 20 (Syfpeithi) and over 50 (BIMAS) was used as a cut-off. The epitope prediction software allowed the prediction of several possible epitopes that may bind the specific HLA allele, therefore multiple putative epitopes of varying length were

predicted for some polymorphic residues. In total, 65 CD8 class-I associated peptides were predicted. Epitopes were predicted that contained the polymorphic sites *within* the peptide (55 peptides; Table 2-5) and also epitopes where the polymorphic site was *flanking* the predicted epitope (10 peptides; Table 2-6) as a mutation at the flanking sites may affect the processing of the epitope and subsequent T-cell response.

Assessment of putative T cell epitopes in IFNγ ELISpot assays: This work was performed by myself. I aimed to assess whether the associations between sequence polymorphisms and patient HLA types and epitope prediction at polymorphic sites identified T cell targets. 65 peptides at putative T cell epitopes were synthetically synthesized (Proimmune) and subsequently assessed in IFNγ ELISpot assays. Results are described in this Chapter.

3.3 Aims

The aims of this Chapter were to

(1) Comprehensively assess T cell targets in three different patient groups: patients with spontaneously resolved HCV infection, and those acutely and chronically infected with HCV gt3a.

(2) Compare two different methodologies for the assessment of T cell responses: overlapping peptides spanning the whole genome, and HLA predicted peptides.

(3) Define T cell responses to single peptides, and assess sequence polymorphisms within T cell targets detected in HCV gt3 infection

(4) Assess and compare T cell specificity in HCV gt1 and gt3 infection. To do this, we aimed to analyse T cell responses published in the literature for HCV genotype 1 infection, and compare them to responses detected in the Oxford HCV gt3 cohort.

(5) Establish whether the magnitude of T cell responses is associated with (A) spontaneous resolution of HCV infection, (B) treatment response and (C) known predictors of treatment response (HCV pre-treatment viral load, IL28B genotype, age, sex and cirrhosis) in chronic HCV genotype 3 infection.

Data published in Humphreys, von Delft et al (joint first authorship) is included in Chapter 3 of this thesis.

3.4 Patient characteristics and study outline

HCV gt3a infected individuals (108 chronically and 16 acutely infected) were recruited from different sites: (John Radcliffe Hospital Oxford [acute=1, chronic 104], MIH Boston [acute=10, chronic=4], Baltimore [acute=5]). Additionally, 16 patients with spontaneously resolved HCV infection and 44 patients with HCV gt1a/b infection were recruited from the John Radcliffe Hospital, Oxford. Informed consent and local ethical approval was obtained for all patients.

Spontaneously resolved patients (Table 3-1) were HCV antibody positive and HCV RNA negative, and sampled at different time points after resolved infection. Spontaneously resolved patients were all Caucasian and predominately male. The main risk factor was intravenous drug use and the majority of patients had the favourable IL28B genotype (TT) (14/16 patients, 87.5%).

Table 3-1: Patient characteristics of spontaneously resolved patients

number of patients		16	%
sex	female	5	**31.3**
	male	11	**68.8**
IL28B genotype	TT	14	**87.5**
	GT	1	**6.3**
	GG	0	**0.0**
	not typed	1	**6.3**
Risk factor	intravenous drug use	11	**68.8**
	tattoo	2	**12.5**
	needlestick	1	**6.3**
	sexual	1	**6.3**
	unkown	1	**6.3**
Treatment	not treated	16	**100.0**
	treated	0	**0.0**
Ethnicity	Caucasian	16	**100.0**
Age	>40	7	**43.8**
	<40	9	**56.3**

Patient characteristics of chronically infected individuals are detailed in Table 3-2. Patients were mainly male (71.3%), Caucasian (80.6%), and the majority was infected through intravenous drug use (64.8%). Of 98 patients typed for the IL28B polymorphism (90.7%), the majority had the favourable IL28B genotype TT (57.1%),

whereas 39% had the IL28B type GT, and only 4% the unfavourable IL28B genotype GG. 78 patients (72.2%) were treated with pegylated Interferon-α and Ribavirin (IFN/RBV), with 53.8% achieving sustained virological response (SVR). After the first treatment, 25 patients relapsed (32.1%), six patients did not respond to treatment (7.7%) and three patients (3.8%) did not complete the treatment course due to interferon related side effects. Eight patients were re-treated with IFN/RBV, of which two achieved SVR, four patients relapsed, one patient did not respond to treatment and one patient discontinued treatment.

Table 3-2: Patient details chronic HCV gt3a cohort

number of patients		108	%				
	female	31	**28.7**				
	male	77	**71.3**				
IL28B genotype	typed	98	**90.7**				
	TT	56	57.1				
	GT	38	38.8				
	GG	4	4.1				
	not typed	10	**9.3**				
Risk factor	IVDU	70	**64.8**				
	unkown	14	**13.0**				
	BP	12	**11.1**				
	cocaine	6	**5.6**				
	tattoo	3	**2.8**				
	needlestick	2	**1.9**				
	vertical	1	**0.9**				
Treatment	not treated	30	**27.8**				
	treated	78	**72.2**				
		outcome		Tx2	outcome	Tx3	outcome
	SVR	42	**53.8**		2		
	REL	25	**32.1**	5	4	2	2
	NR	6	**7.7**	1	1		
	incomplete	3	**3.8**	2	1		
	not known	2	**2.6**				
Ethnicity	Caucasian	87	**80.6**				
	Asian	16	**14.8**				
	South American	1	**0.9**				
	not known	4	**3.7**				
Age	>40	35	**32.4**				
	<40	73	**67.6**				

IVDU intravenous drug use; Tx treatment; SVR sustained virological response; REL relapse; NR not responding to treatment.

16 patients with acute infection were recruited from different sites (Table 3-3). The majority of patients were infected using intravenous drugs (9/16, 56%). Clinical outcome within the acute cohort was variable: of 16 patients analysed, 2 were lost to

follow up, 4 cleared the infection spontaneously, 4 were treated during the acute phase with 3 clearing under therapy and 1 patient not responding, and 4 patients developed chronic infection. All patients had a confirmed infection with HCV genotype 3a. The majority of patients were <40 (69%), Caucasian (100%), male (56%), and were infected through intravenous drug use (56%).

Table 3-3: Patient details acute HCV gt3a cohort.

sample ID	Age	Sex	Race	Risk factor	GT	HCV VL	ALT	Treatment	outcome
						Iu/ml			
7-41	23	F	W	Sexual	3	>700000	68	naïve	cleared spontaneously
50163	31	M	W	IVDU	3a	6290		naïve	cleared spontaneously
10256	23	M	W	IVDU	3a	8930000	46	naïve	cleared spontaneously
50017	21	M	W	IVDU	3	15900	72	naïve	cleared spontaneously
7-31	19	M	W	IVDU	3	168000	229	naïve	chronic
7-06	20	F	W	IVDU	3	16100	38	naïve	chronic
3-43	46	M	W	blood exposure	3	204000	592	naïve	chronic
6-23	28	M	W	Sexual	3	136000	698	naïve	chronic
8-10	45	F	W	Tattoo	3a	>700000	201	naïve	chronic
70230	21	F	W	IVDU	3a	600000		naïve	chronic
7-75	18	F	W	IVDU	3	102000	88	naïve	lost in <6 months
70351	25	F	W	IVDU	3a		21	naïve	lost in <6 months
6-40	40	F	W	Sexual	3a	>700000	464	treated	cleared on Tx
822	49	M	W	Sexual	3a	46603	1441	treated	cleared on Tx
7-16	41	M	H	Sexual	3a	67535	250	treated	chronic
6-56	20	M	W	IVDU/Tattoo	3	>700000	383	treated	chronic (non-responder)

F female; M male; W White; H Hispanic; IVDU intravenous drug use; Tx treatment.

A schematic for the study outline is depicted in Figure 3-1. Two different approaches were used to identify HCV gt3 specific T cell targets in IFNγ ELISpot assays:

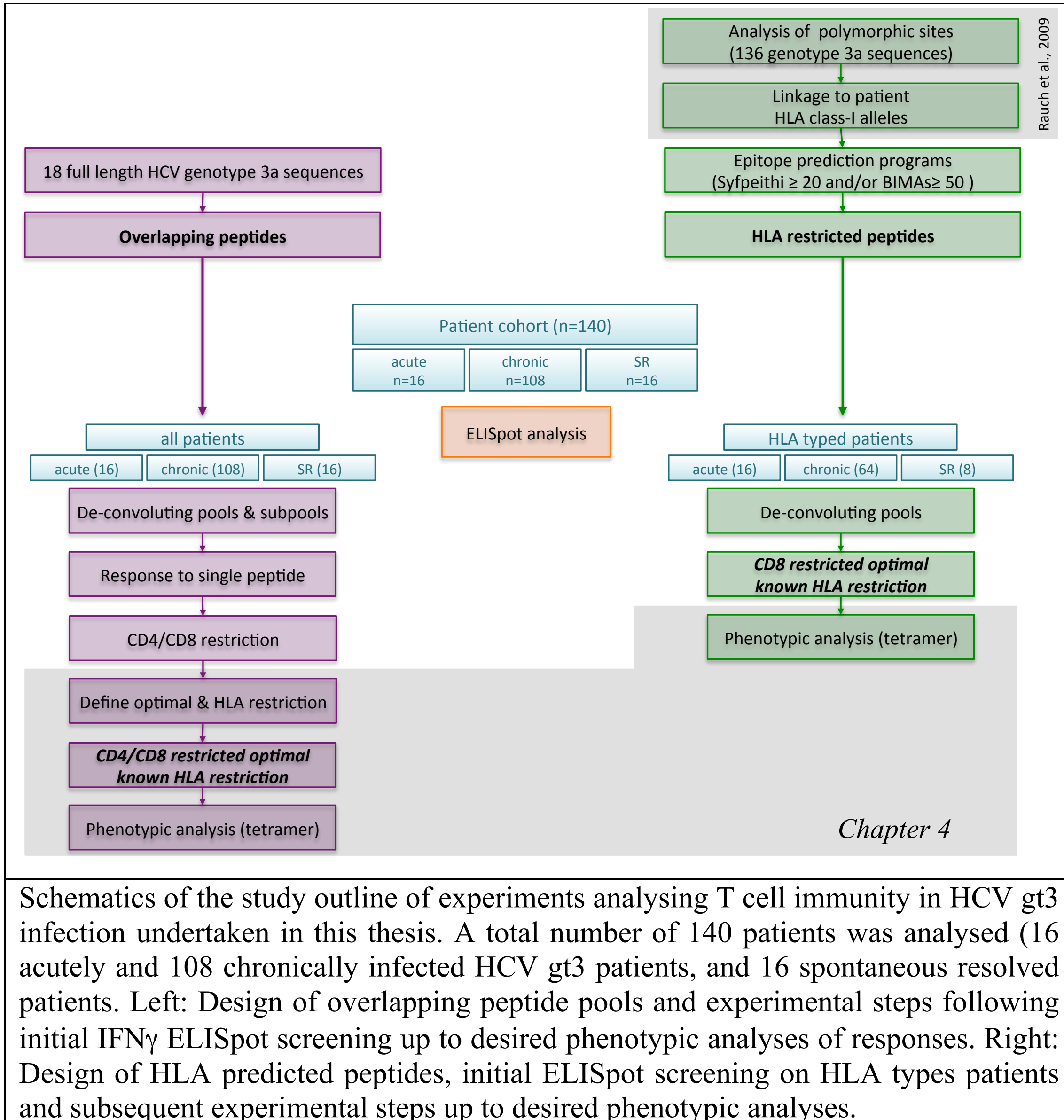
(i) Overlapping peptides (spanning the whole viral genome, 15-18 amino-acids (AA), overlapping by 11 AA), based on a consensus sequence derived from 18 full-length gt3a sequences (left) and

(ii) A novel, sequence-led approach using 46 wild-type and variant peptides (non-structural proteins, 9-10 AA) corresponding to putative HLA class-I restricted epitopes under T cell selection (right).

A detailed description of peptide design can be found in the Methods (page 76).

All patients (acute (n=16), established chronic (n=108) HCV gt3a infection and spontaneously resolved infection (n=16)) were screened with overlapping peptides; whereas only HLA typed patients were tested using the HLA predicted peptide set (acute (n=16), established chronic (n=64) HCV gt3a infection and spontaneously resolved infection (n=8)).

Figure 3-1: Study outline for the assessment of HCV gt3 T cell responses



Schematics of the study outline of experiments analysing T cell immunity in HCV gt3 infection undertaken in this thesis. A total number of 140 patients was analysed (16 acutely and 108 chronically infected HCV gt3 patients, and 16 spontaneous resolved patients. Left: Design of overlapping peptide pools and experimental steps following initial IFNγ ELISpot screening up to desired phenotypic analyses of responses. Right: Design of HLA predicted peptides, initial ELISpot screening on HLA types patients and subsequent experimental steps up to desired phenotypic analyses.

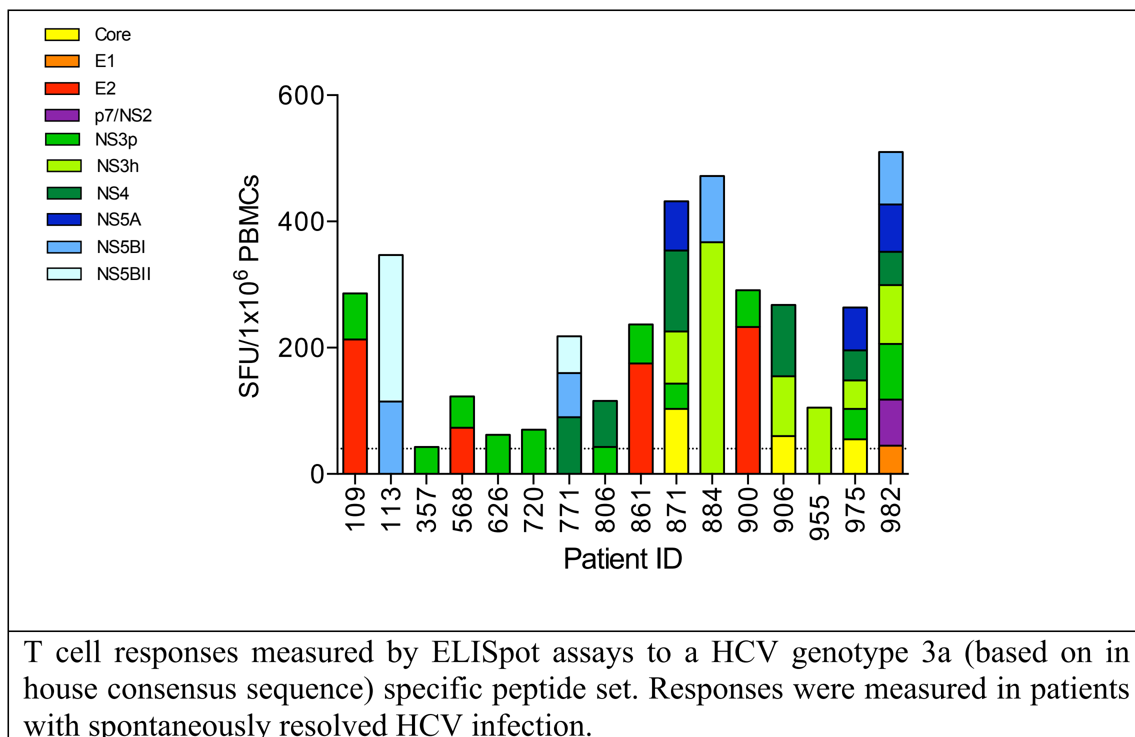
3.5 T cell responses to overlapping peptide pools

3.5.1 T cell responses in resolved HCV infection

To analyse responses to HCV genotype 3a specific peptides in spontaneously resolved infection, we identified 16 patients with previously resolved infection. Patients were recruited from the John Radcliffe Hospital in Oxford, and plasma and PBMC were derived as described in the methods section (see 2.2.1 page 67, and 2.3.1 page 72).

To assess the HCV gt3a specific immune response, T cell responses were tested in ELISpot assays using a genotype 3a specific peptide set (see 2.4.1, page 76). T cell responses to overlapping peptide pools were found in all tested patients. The majority of patients (14 out of 17) elicited a T cell response against NS3 peptides, with 11 patients responding to NS3 protease and 6 to NS3 helicase. Strong responses were also seen against E2 and NS5 peptides in several patients (Figure 3-2).

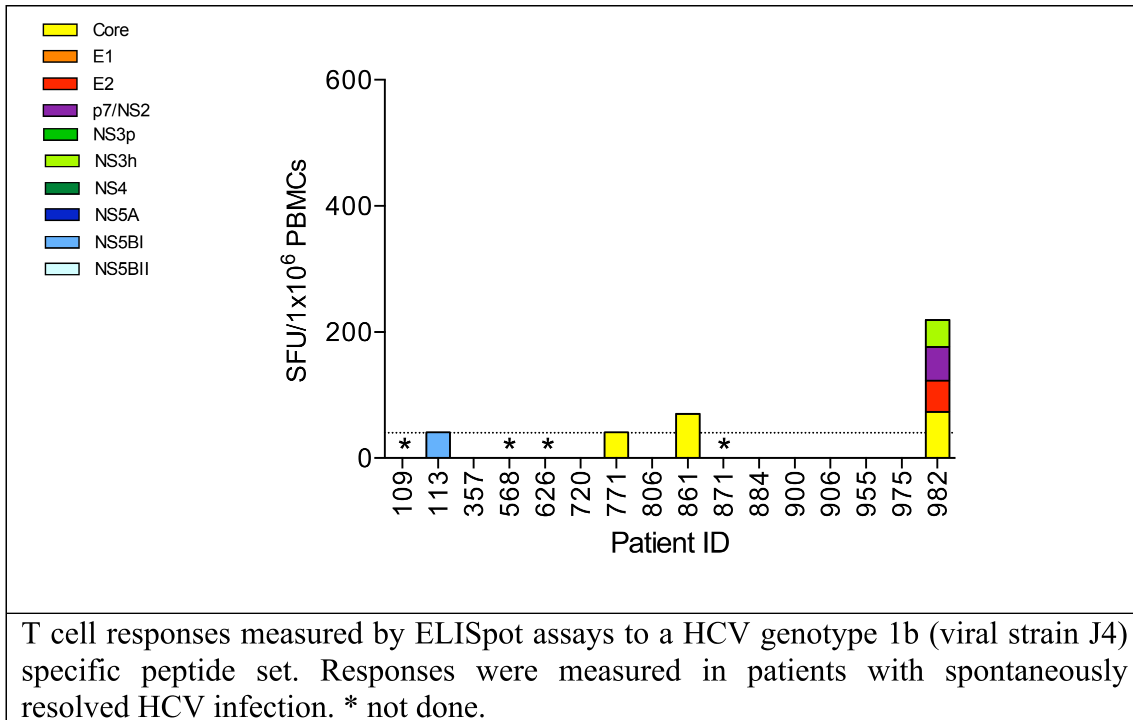
Figure 3-2: T cell responses in spontaneous resolvers to HCV genotype 3a specific peptides



T cell responses measured by ELISpot assays to a HCV genotype 3a (based on in house consensus sequence) specific peptide set. Responses were measured in patients with spontaneously resolved HCV infection.

To assess cross-reactivity between HCV gt3 and gt1 responses in spontaneously resolved patients, T cell responses to HCV genotype 1b peptides were assessed in 13 out of 17 patients (Figure 3-3). Responses to HCV gt1 peptides were detected in 5 patients only and generally low, and the majority targeted HCV core (4/5).

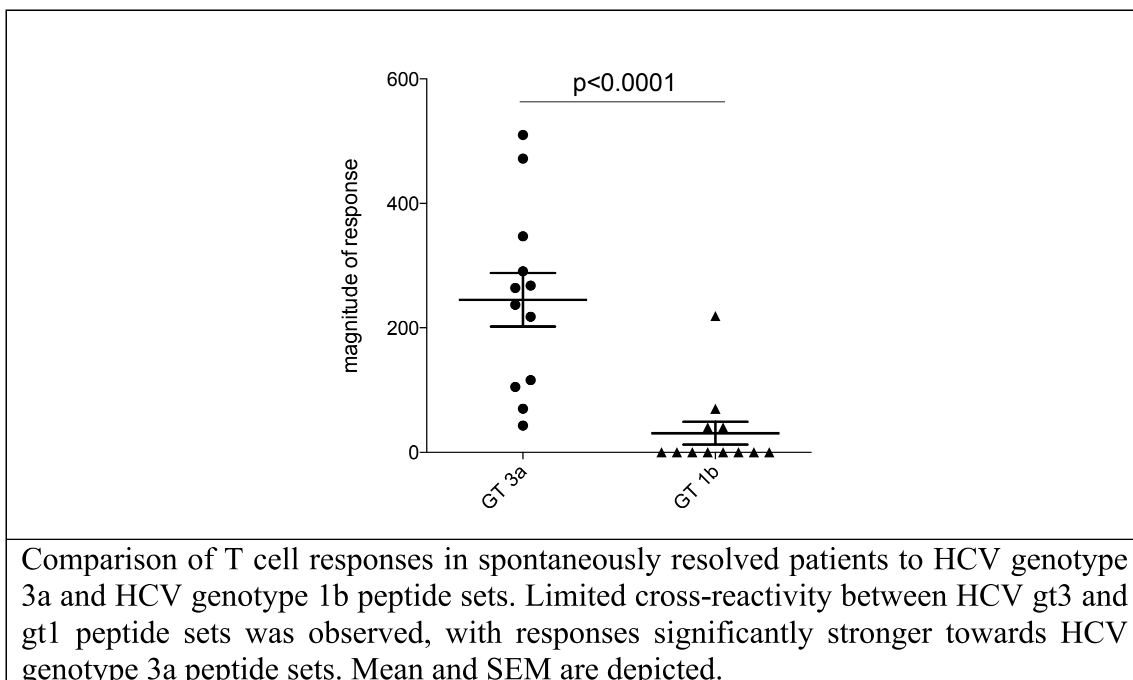
Figure 3-3: T cell responses in spontaneous resolvers to HCV genotype 1b specific peptides



T cell responses measured by ELISpot assays to a HCV genotype 1b (viral strain J4) specific peptide set. Responses were measured in patients with spontaneously resolved HCV infection. * not done.

When comparing the overall response against HCV gt1 and gt3a specific peptides in paired patient samples, stronger responses were seen against the gt3a peptide set (Figure 3-4).

Figure 3-4: Comparison of T cell responses in spontaneously resolved patients against HCV genotype 3a and 1b peptide sets



Comparison of T cell responses in spontaneously resolved patients to HCV genotype 3a and HCV genotype 1b peptide sets. Limited cross-reactivity between HCV gt3 and gt1 peptide sets was observed, with responses significantly stronger towards HCV genotype 3a peptide sets. Mean and SEM are depicted.

3.5.2 T cell responses in chronic HCV genotype 3a in infection

To analyse responses to HCV genotype 3a specific peptides in chronic HCV infection, 108 patients were recruited from the John Radcliffe Hospital in Oxford, and Plasma and PBMC were derived as described in the methods section (see 2.2.1 page 67, and 2.3.1 page 72). For all patients, the infecting HCV genotype was confirmed as subtype 3a. A table detailing patient characteristics (age, sex, IL28B genotype) and treatment details can be found in Table 3-2 (page 88).

T cell responses were measured by ELISpot assays using overlapping peptide pools spanning the full length of the HCV gt3 genome in 108 chronic HCV gt3a infected patients (Figure 3-5). No response was detected in 51 out of 108 patients (47%). A response to HCV gt3a peptides was detected in 57 patients (53%), with responses mainly targeting the core (39 out of 57, 68%) and NS3 helicase and protease region (26 out of 57, 46%). Responses were low in magnitude (94.34 ± 13.16), with only 17% of tested patients exhibiting overall T cell responses over 200 SFU/million cells in magnitude.

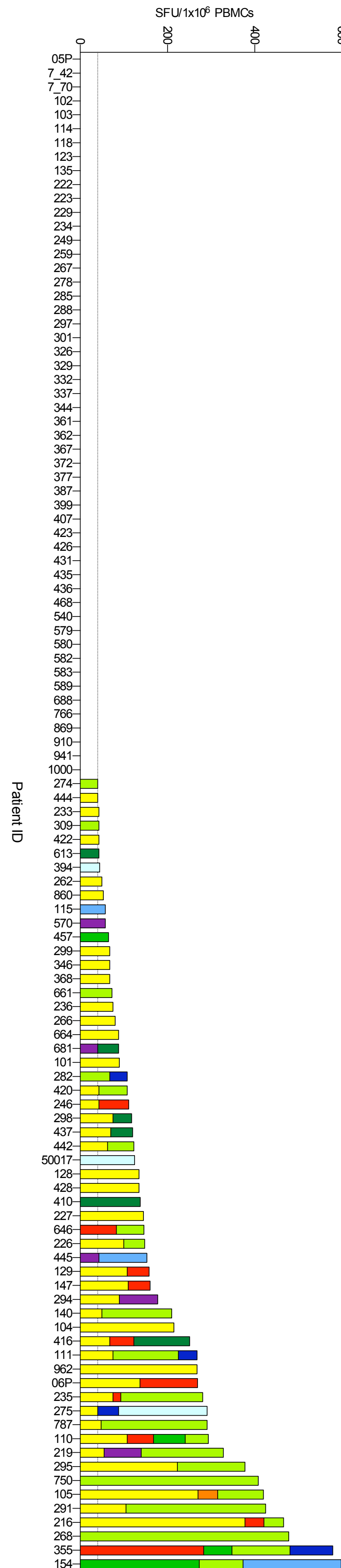
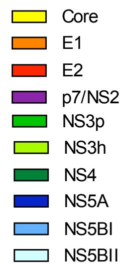


Figure 3-5: T cells responses against a HCV gt3a peptide set in chronically infected HCV gt3a patients

T cell responses detected in patients chronically infected with HCV gt3a, measured by IFNγ ELISpot using a HCV gt3a peptide set spanning the entire HCV genome. T cell responses over cut-off were detected in 56/108 patients. ELISpot assays on 12 patients were performed by Isla Humphreys.



HCV gt3 specific responses were compared to HCV gt1 responses in 44 gt1 infected patients. T cell responses were assessed using a HCV gt1b peptide set (strain J4b) in IFNγ ELISpot assays (Figure 3-6). 20 out of 44 patients did not form an IFNγ T cell response (45%). The HCV region most frequently targeted was core, with all but one responder targeting this region (23/24). This is in line with previously published evidence: T cell responses in chronic gt1 infection have been previously described as mainly targeting HCV core (Humphreys et al. 2012; V. M. Fleming et al. 2010).

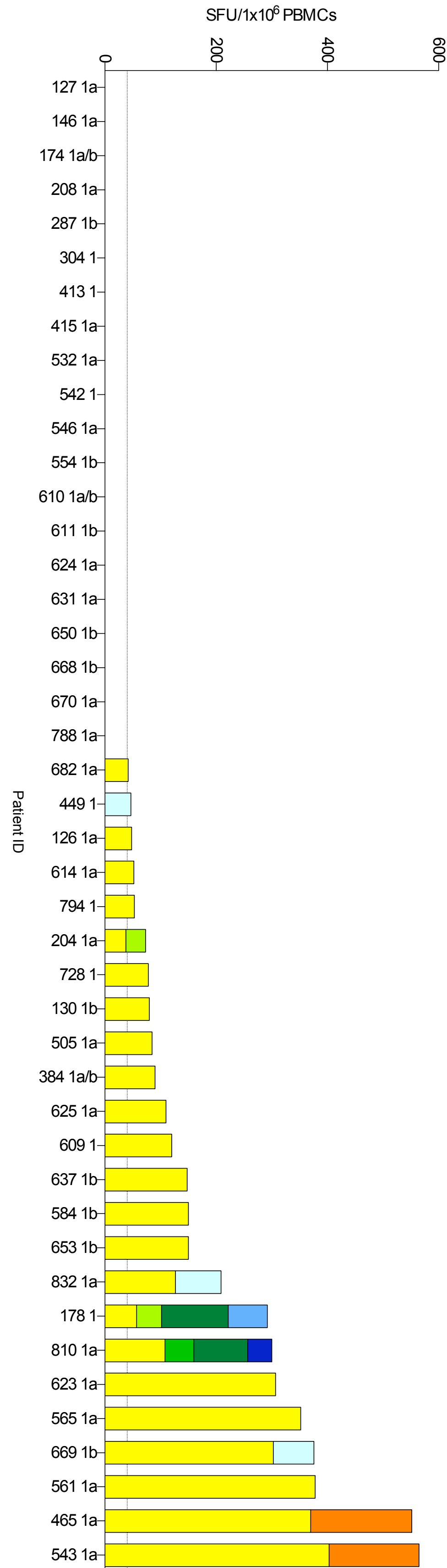
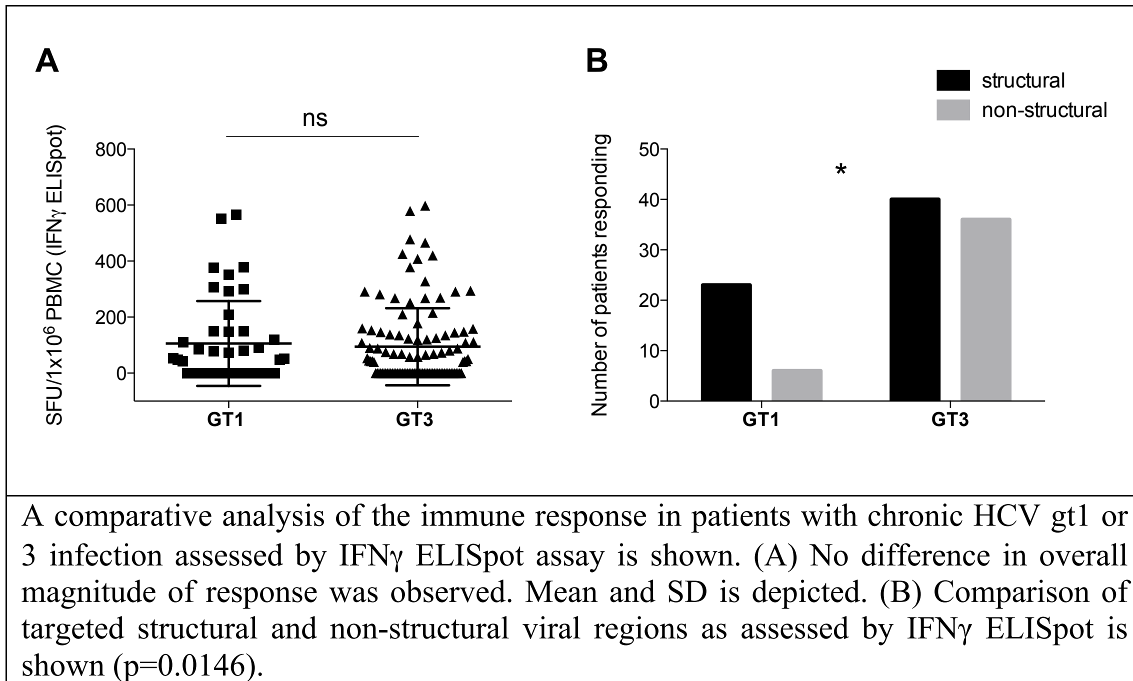


Figure 3-6: T cell responses against a HCV gt1b peptide set in chronically infected HCV gt1 patients.

T cell responses detected in patients chronically infected with HCV gt1, measured by IFNγ ELISpot using a HCV gt1b peptide set spanning the entire HCV genotype. T cell responses over cut-off were detected in 24/44 patients. ELISpot assays were performed by Anthony Brown.

When comparing the overall magnitude of response in HCV gt1 and gt3 infected patients, no difference was seen (Figure 3-7). However, when comparing targeted viral regions in HCV gt1 and gt3 infection, responses measured in HCV gt3a infection were significantly more aimed at non-structural regions (Fisher's exact test, p=0.0146).

Figure 3-7: Comparison of T cell responses to HCV viral regions in HCV gt1 and gt3 infection.



A comparative analysis of the immune response in patients with chronic HCV gt1 or 3 infection assessed by IFNγ ELISpot assay is shown. (A) No difference in overall magnitude of response was observed. Mean and SD is depicted. (B) Comparison of targeted structural and non-structural viral regions as assessed by IFNγ ELISpot is shown (p=0.0146).

3.5.3 T cell responses in acute HCV genotype 3a infection

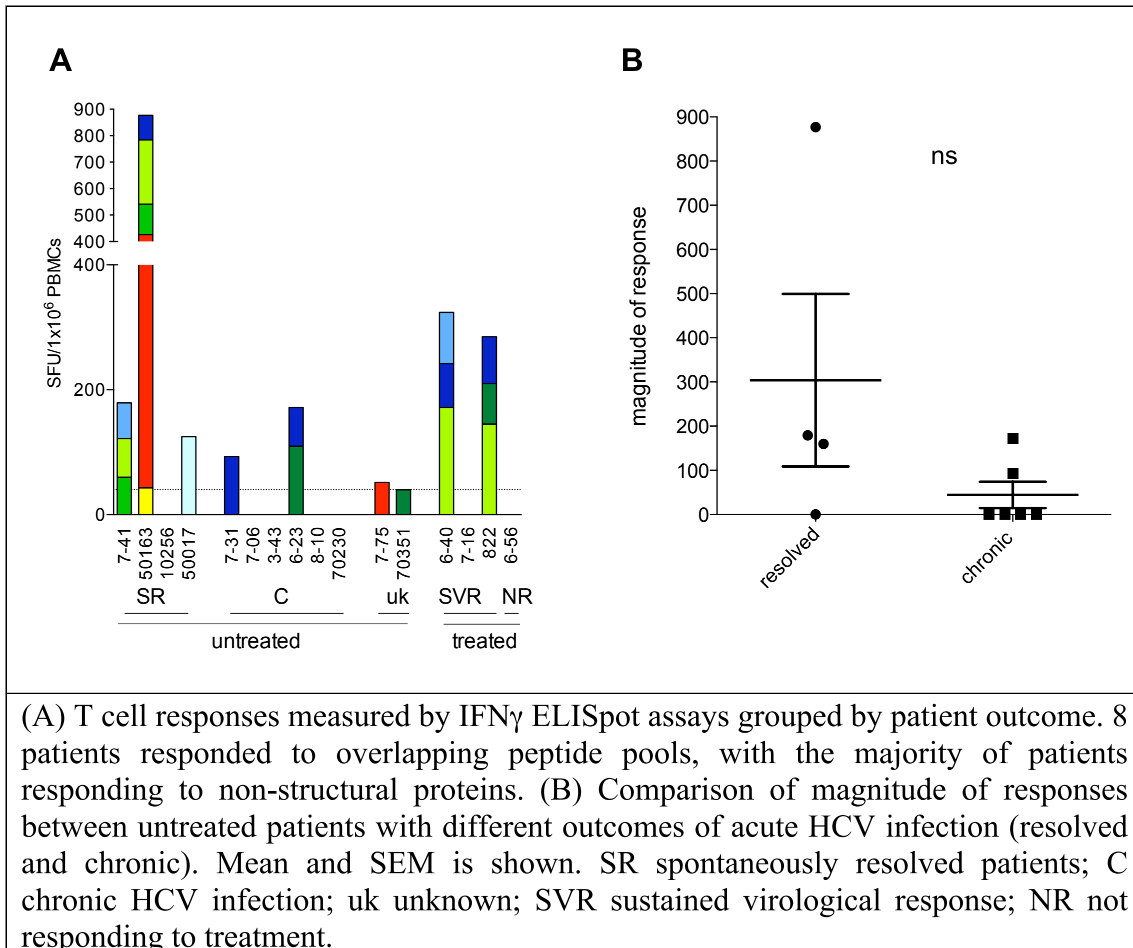
Immune responses in acute HCV gt3a infection have not been studied previously. Here, we had the exceptional opportunity to study an acute cohort of 16 samples, generously provided by Georg Lauer (Boston) and Andrea Cox (Baltimore). One patient was recruited at the John Radcliffe Hospital, Oxford (for patient details, see Table 3-3, page 90).

Of 16 patients studied, two patients were lost to follow up before outcome was determined, four cleared the infection spontaneously, four were treated during the acute phase with three patients clearing the infection under therapy and one patient not responding to treatment, and six patients developed chronic infection.

We tested T cell responses in all 16 patients acutely infected with HCV (Figure 3-8). Responses were detected in 9 out of 16 patients, with the majority of patients responding to HCV non-structural regions (Figure 3-8).

Magnitude of T cell responses was compared between patients not receiving treatment who either resolved HCV infection or developed chronic infection. Stronger responses were observed in patients resolving infection spontaneously, although this trend did not reach statistical significance (t test p=0.1318).

Figure 3-8: Magnitude of T cell responses did not differ in acutely infected patients who spontaneously cleared infection or developed chronic infection.

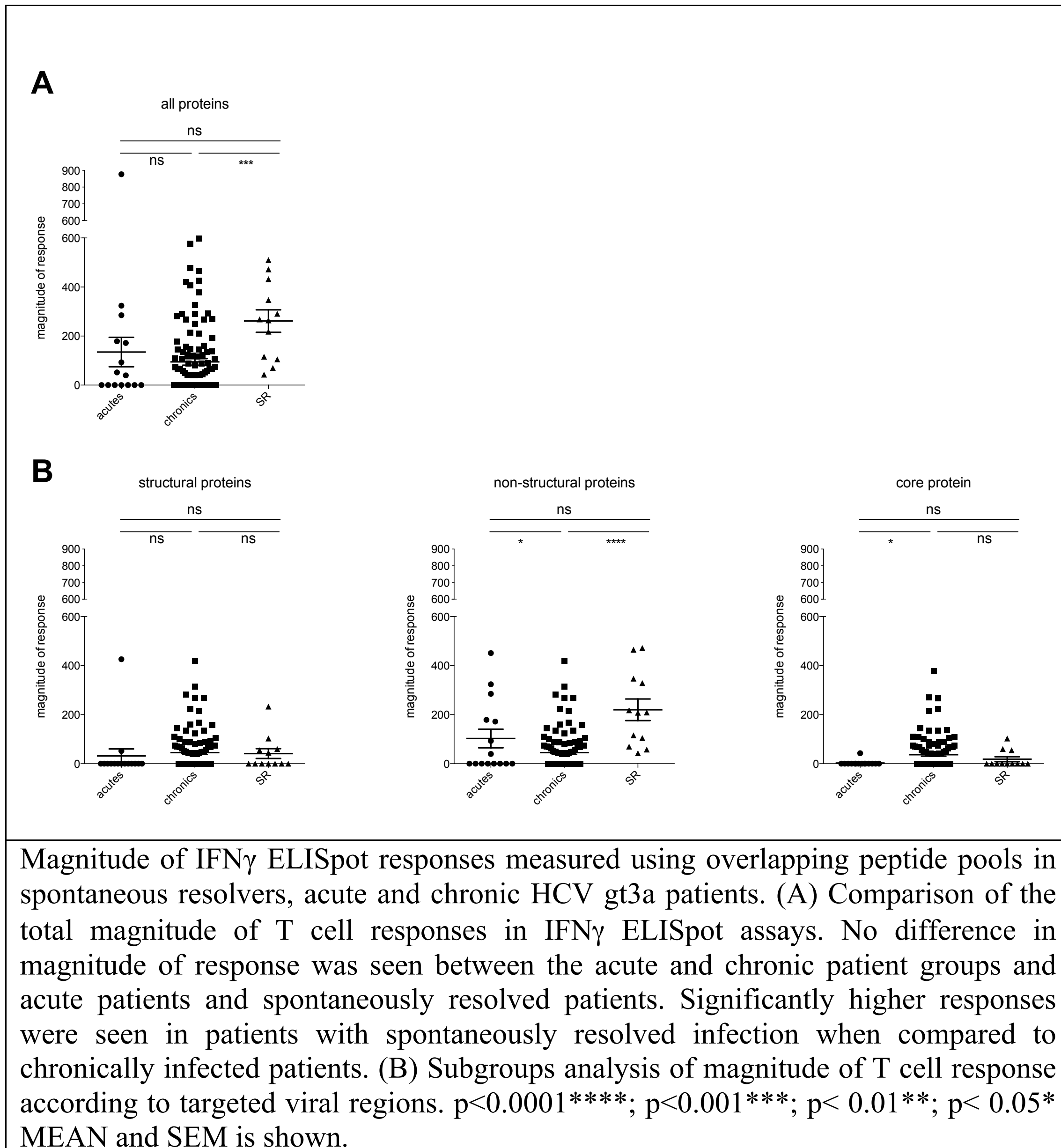


(A) T cell responses measured by IFNγ ELISpot assays grouped by patient outcome. 8 patients responded to overlapping peptide pools, with the majority of patients responding to non-structural proteins. (B) Comparison of magnitude of responses between untreated patients with different outcomes of acute HCV infection (resolved and chronic). Mean and SEM is shown. SR spontaneously resolved patients; C chronic HCV infection; uk unknown; SVR sustained virological response; NR not responding to treatment.

3.5.4 Comparison of magnitude of T cell responses and targeted viral regions using overlapping peptide pools in spontaneous resolvers, acutely and chronically infected patients

The magnitude of T cell responses measured by IFNγ ELISpot assay was compared in spontaneous resolvers, and acutely and chronically infected patients (Figure 3-9 A). No difference in magnitude of response was seen when comparing acutely infected patients to spontaneously resolved patients and the chronic patient group. In contrast, significantly higher responses were seen in patients with spontaneously resolved infection when compared to chronically infected patients (unpaired t test, $p<0.001$).

Figure 3-9: Magnitude of T cell responses in spontaneously resolved patients and acute chronic HCV genotype 3a infected patients



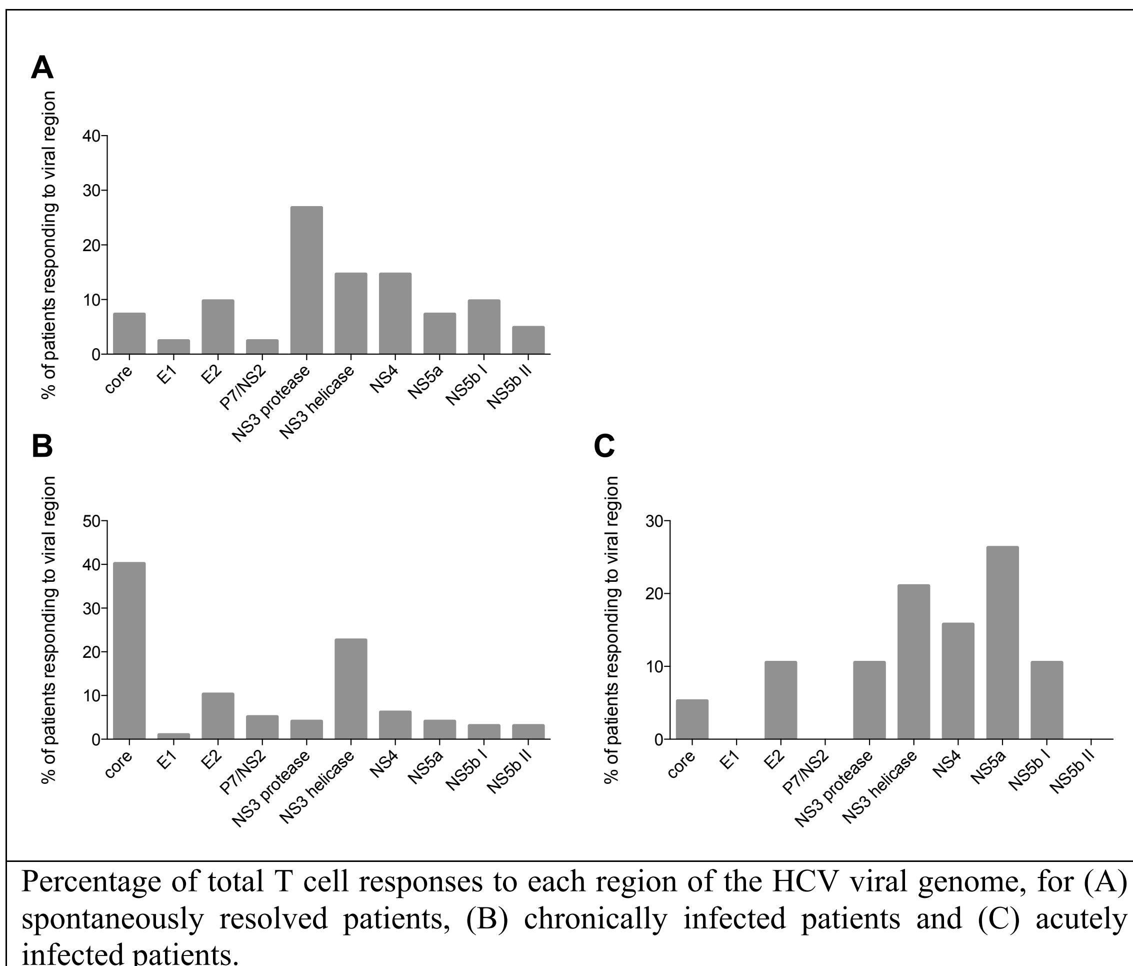
Magnitude of IFNγ ELISpot responses measured using overlapping peptide pools in spontaneous resolvers, acute and chronic HCV gt3a patients. (A) Comparison of the total magnitude of T cell responses in IFNγ ELISpot assays. No difference in magnitude of response was seen between the acute and chronic patient groups and acute patients and spontaneously resolved patients. Significantly higher responses were seen in patients with spontaneously resolved infection when compared to chronically infected patients. (B) Subgroups analysis of magnitude of T cell response according to targeted viral regions. p<0.0001****; p<0.001***; p< 0.01**; p< 0.05* MEAN and SEM is shown.

Subsequently, T cell responses to structural and non-structural regions of the virus were compared separately (Figure 3-9 B). There was no difference in the magnitude of response to structural proteins in all three patients groups. The magnitude of responses to non-structural proteins resembled the observations made when comparing the overall magnitude of T cell responses (Figure 3-9). Chronic patients showed low responses, with significantly higher responses in acutely infected patients (t test, p=0.0202), and spontaneously resolved patients (t test, p<0.0001). No difference in magnitude was observed between acutely infected patients and spontaneous resolvers. Of note, when comparing responses in the core region,

chronically infected patients showed significantly higher responses than acutely infected patients, but no difference between chronically infected and spontaneously resolved patients was observed.

These results are reflected when analysing the percentage of T cell responses mounted against different HCV viral regions in the different patient groups (Figure 3-10): In spontaneous resolved patients (A), the majority of T cell responses was aimed at the NS3 protease region, with both the NS3 helicase and NS4 being the next most commonly targeted regions. Contrary, T cells in acute gt3 infection (C) commonly target the NS5a, NS3 and NS4 region, most likely reflecting that some patients progress to spontaneous resolution. Regions targeted in chronic HCV gt3 infection (B) differ from those targeted in acute and spontaneously resolved infection: here, the regions most commonly targeted were core, and the NS3 helicase region.

Figure 3-10: T cell responses of spontaneously resolved, acutely and chronically infected HCV patients target different viral region



Percentage of total T cell responses to each region of the HCV viral genome, for (A) spontaneously resolved patients, (B) chronically infected patients and (C) acutely infected patients.

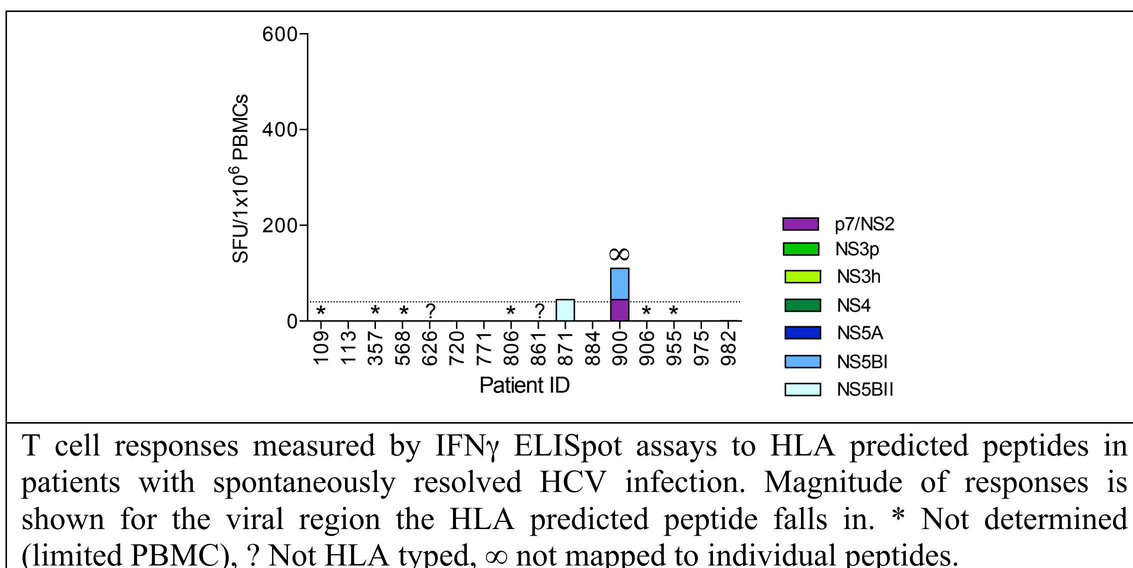
3.6 T cell responses to HLA predicted peptides

In addition to the experiments run with overlapping HCV genotype 3a peptides covering the whole viral genome (see 3.5 T cell responses to overlapping peptide pools, page 92), T cell responses in HCV gt3a patients were evaluated using HLA predicted peptides in spontaneously resolved patients, and patients acutely and chronically infected with HCV genotype 3a using IFNγ ELISpot assays. Peptides were matched to the patients HLA type, reducing the number of PBMCs used in each experiment. Since only a limited number of optimal peptides were used in each experiment, potential enhancement of the number and magnitude of detected immune responses was expected; due to optimal peptide length and limited potential competition for HLA binding sites.

3.6.1 Responses to HLA predicted peptides in spontaneous resolvers

To assess T cell responses to HLA predicted peptides in a cohort of patients who spontaneously resolved HCV infection, we first determined HLA types of patients with resolved infection (14/16 patients). Next, T cell responses were assessed using peptides matching the patients HLA type in IFNγ ELISpot assays. Due to limited availability of PBMCs in patients with spontaneously resolved infection, T cell responses to HLA predicted gt3 specific peptides were tested in only 8 out of 14 typed patients. Of these, T cell response to HLA predicted peptides were detected in two patients (Figure 3-11).

Figure 3-11: T cell responses in spontaneous resolvers to HLA predicted peptides



T cell responses measured by IFNγ ELISpot assays to HLA predicted peptides in patients with spontaneously resolved HCV infection. Magnitude of responses is shown for the viral region the HLA predicted peptide falls in. * Not determined (limited PBMC), ? Not HLA typed, ∞ not mapped to individual peptides.

3.6.2 T cell responses to HLA predicted peptides in chronically infected patients

HLA types of 64 chronically HCV gt3a infected patients were obtained. Next, HLA predicted peptides were selected according to the patients HLA type, and T cell responses measured using IFNγ ELISpot assays. Out of tested 64 patients, a T cell response to HLA predicted peptides was detected in 13 patients (20%) (Figure 3-12).

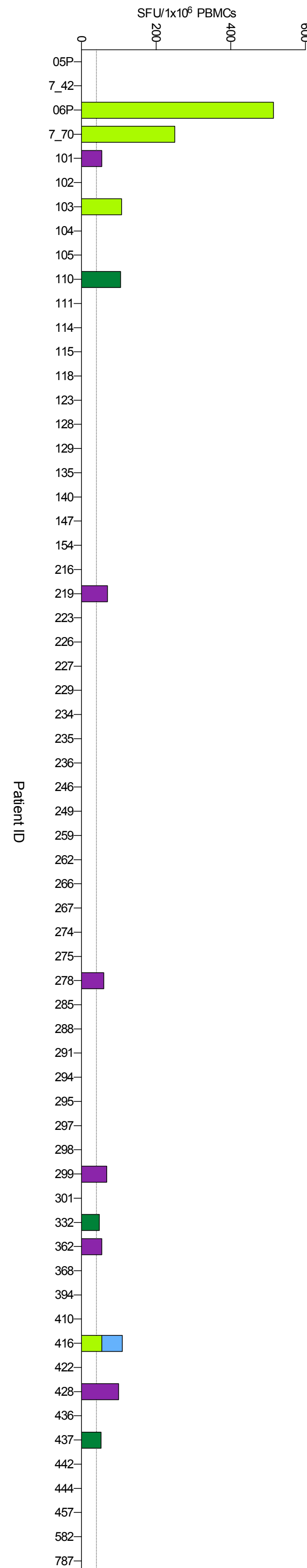


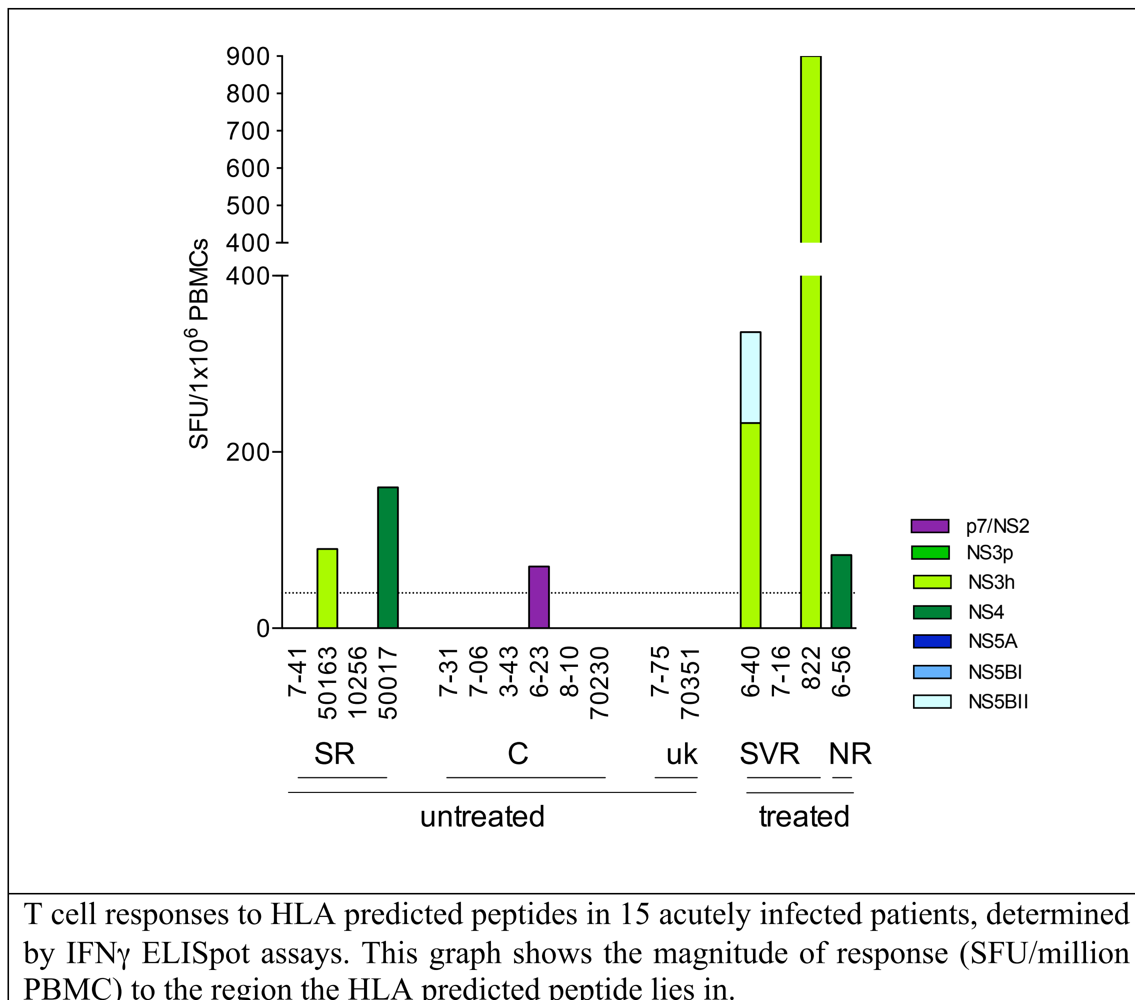
Figure 3-12: T cell responses in chronically infected patients to HLA predicted peptides

T cell responses to HLA predicted peptides were tested in 64 tested patients in IFNγ ELISpot assays. 13 patients responded to HLA predicted peptides. This graph shows the magnitude of response to the region the HLA predicted peptide lies in.

3.6.3 Responses to HLA predicted peptides in acutely infected patients

HLA types of all acutely infected patients were obtained, and T cell responses tested using IFNγ ELISpot assays using HLA predicted peptides matched to patients' HLA type. T cell responses were detected in 6 out of 16 tested patients (37.5%), with 3 patients responding to peptides from the NS3 region. Other responses were aimed at NS2, NS4 and NS5B viral regions.

Figure 3-13: T cell responses to HLA predicted peptides in acutely infected HCV gt3a patients.



T cell responses to HLA predicted peptides in 15 acutely infected patients, determined by IFNγ ELISpot assays. This graph shows the magnitude of response (SFU/million PBMC) to the region the HLA predicted peptide lies in.

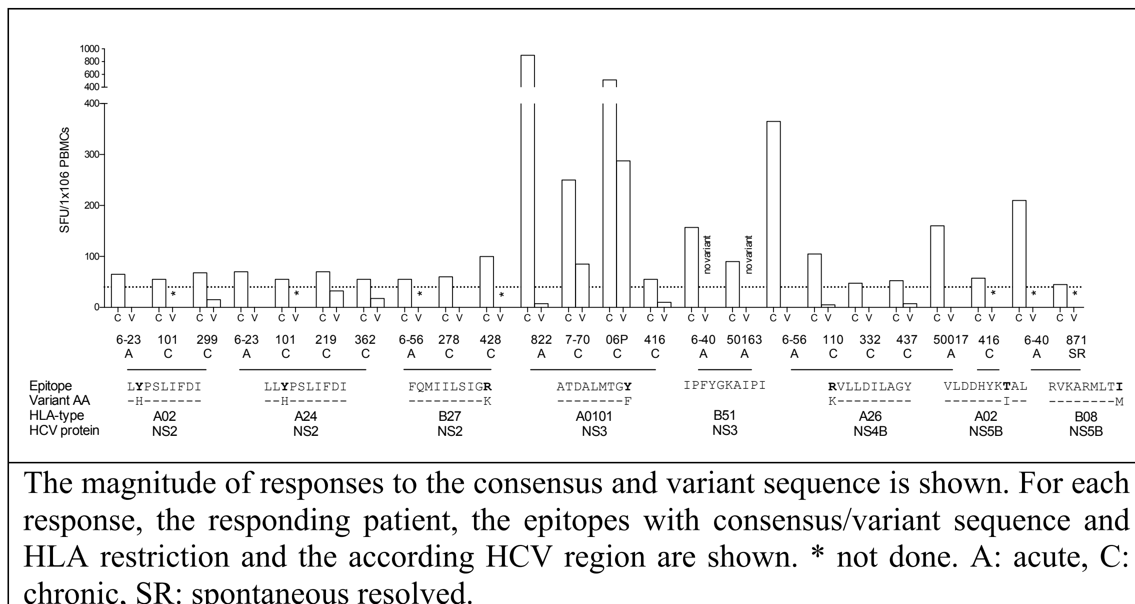
3.6.4 Responses to single peptides and viral variants

Next, we aimed to determine the overall number of peptides detected using HLA predicted peptides. T cell responses to single HLA predicted peptides in spontaneous resolvers, acute and chronic patients are detailed in Figure 3-14. Responses were detected to eight different peptides out of the whole peptide set composed of 65 peptides (all peptides are detailed under 2.4.2 HLA predicted peptides, page 78), with

one to five patients recognizing each peptide. Three peptides were detected in the HCV NS2 region (`LLYPSLIFDI/LYPSLIFDI` (A02/A24 restricted) and `FQMIILSIGR` (B27)), two peptides in the NS3 region (`ATDALMTGY` (A01), `IPFYGKAIPI` (B51)), one peptide in the NS4B region (`RVLLDILAGY` (A26)), and two peptides in the NS5B region (`VLDDHYKTAL` (A02) and `RVKARMLTI` (B08)).

Since the HLA predicted peptide set was based on sequence polymorphisms associated with patients HLA class-I types, we aimed to define the T cell response against the variant sequence originally associated with the restricting HLA type. If the variant described in the original publication was contained within the tested consensus peptide predicted by epitope prediction programs (some polymorphisms were flanking the predicted epitopes), the described peptide variant was tested in IFNγ ELISpot assays. In all patients tested, the response to the peptide variant was lower than the response to the consensus epitope (Figure 3-14).

Figure 3-14: Details of all responses to HLA predicted peptides.



The magnitude of responses to the consensus and variant sequence is shown. For each response, the responding patient, the epitopes with consensus/variant sequence and HLA restriction and the according HCV region are shown. * not done. A: acute, C: chronic, SR: spontaneous resolved.

In one patient (patient ID 900), responses were not mapped from pools to single peptides due to the limitation of PBMC (marked with ∞ in Figure 3-11, page 101). T cell responses were seen to two pools of A02 specific peptides containing NS2 peptides (`VMGGKYFQM, ILSIGRWFNT, LLYPSLIFDI, LLARGSRDGV`) and NS5B peptides (`VLDDHYKTAL, KAINQIRSV, SLSSKAINQI`). Commonly detected responses in other patients suggest that the T cell response may be targeting epitopes `LLYPSLIFDI` in the NS2 region and `VLDDHYKTAL` in the NS5B region.

3.7 Comparison of T cell responses to overlapping and HLA matched peptide sets

After testing patients with two independent peptide sets, I aimed to define whether T cell responses detected by HLA predicted peptides were also detected using overlapping peptides. Since HLA predicted peptides were only defined within the non-structural region; responses to overlapping peptide pools within the non-structural region only were used for this comparison. In Figure 3-15, a direct comparison of T cell responses detected using HLA predicted peptides and overlapping peptide pools in 8 spontaneous resolvers, 16 acutely and 64 chronically infected patients is depicted.

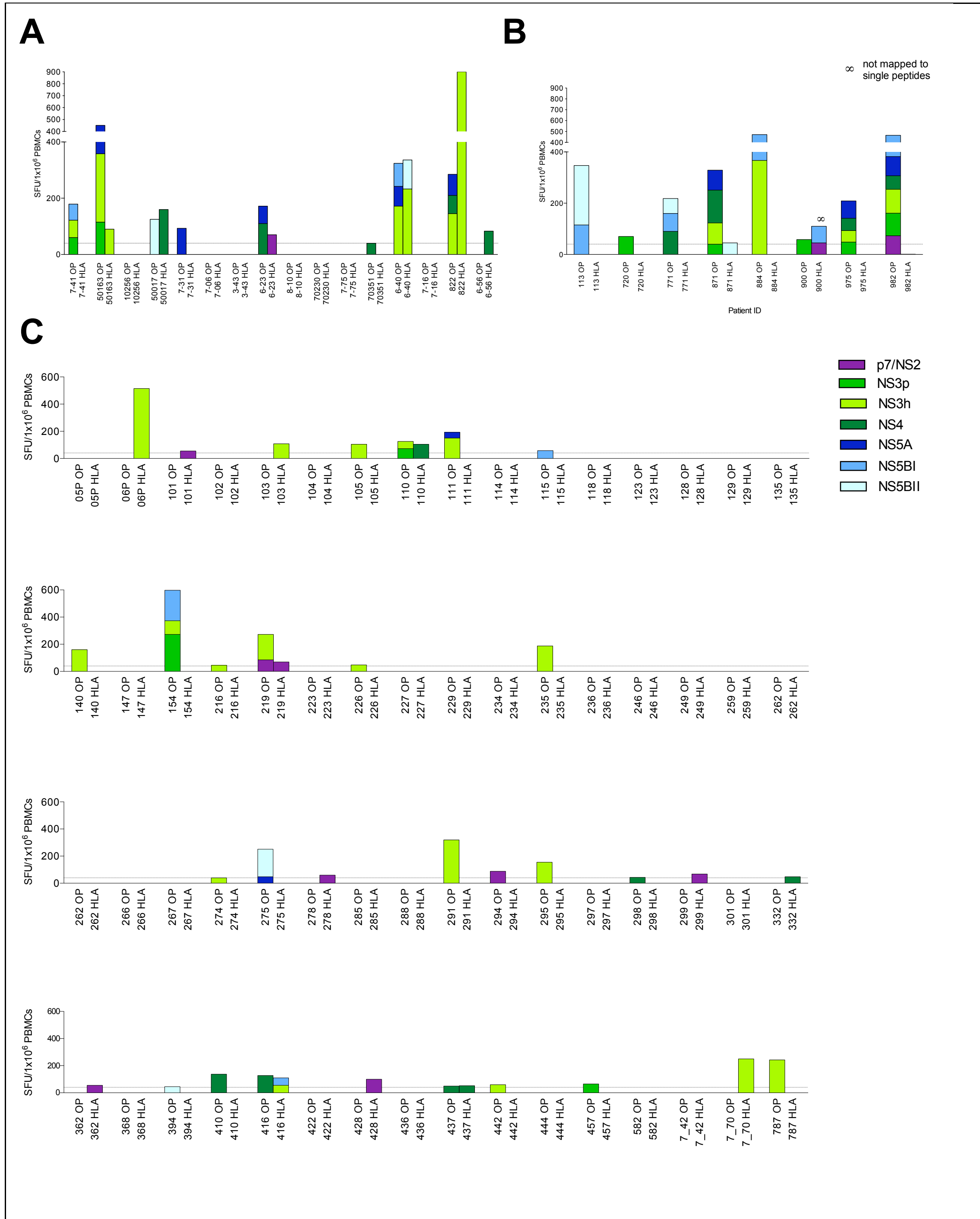
Overall, 44% (39 out of 88) of tested patients did not respond to either peptide set (Figure 3-16). The majority of all patients formed a T cell response (49/88, 56%); responses to overlapping pools were detected in 44% (39/88) of all patients, whereas responses to HLA predicted peptides were observed in 24% (21/88) of patients. 12% of patients responded to both peptide sets, but in only 6% (5/88) of patients the response was aimed at the same pool for overlapping peptide pools and HLA predicted peptides (Figure 3-16). 44% (39/89) of tested patients only responded to one of the two peptide sets, either HLA predicted peptides (10/88, 11%) or overlapping peptide pools (29/88, 33%).

Therefore, cross-reactivity between both peptide sets was low, with only 5 patients responding to the same peptide pools using both approaches. When considering different stages of infection (spontaneous resolvers, acute infection and chronic infection), and then comparing the overlap of responses, divergent results were observed for different patient groups (Figure 3-16, bottom).

All ***spontaneous resolved patients*** responded to either or both of the peptide sets. This is in line with described high T cell responses in patients who spontaneously clear infection (C. L. Day et al. 2002; Lauer et al. 2004; Schulze zur Wiesch et al. 2005). Of note, a response to HLA predicted peptides was only detected in two patients. This low response rate might be due to the method of HLA predicted peptide generation: viral sequence polymorphisms were associated with HLA types in a cohort of chronic HCV gt3a patients, where escape may have already occurred in commonly targeted T cell epitopes. In patients who clear the infection, a response to

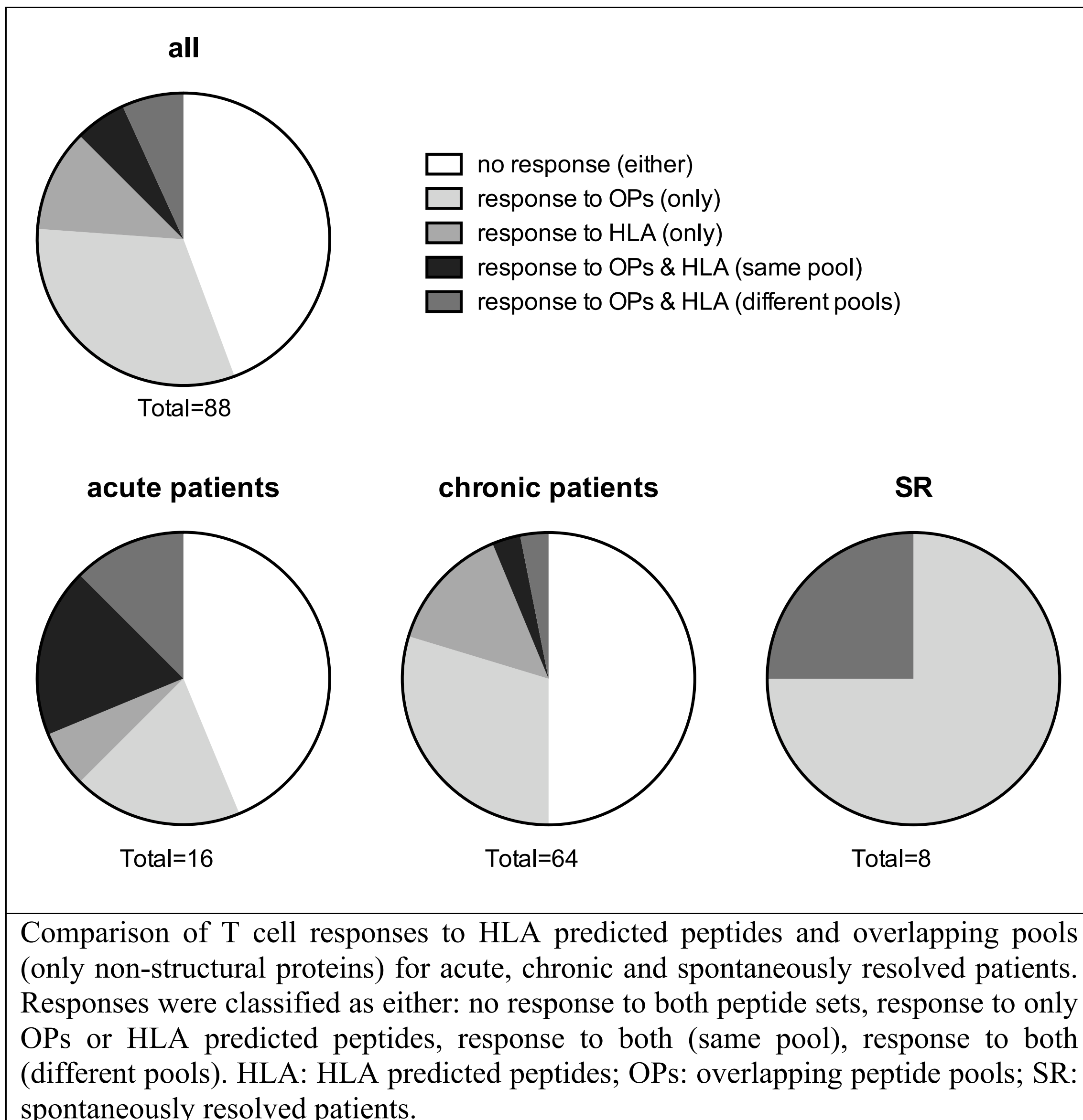
sites of mutation is not likely, since viral regions targeted in these patients do not tend to escape (Erickson et al. 2001; A. L. Cox, Mosbruger, Mao, et al. 2005).

Figure 3-15: Comparison of responses to overlapping pools and HLA predicted peptides in spontaneous resolvers and acutely and chronically infected patients.



T cell responses to HLA predicted peptides and overlapping peptide pools were measured in patients with (B) spontaneously resolved infection (n=8), (A) acutely (n=16) and (C) chronically infected (n=64) patients. For this comparison, responses to overlapping pools are compared to the HCV viral region in which the HLA predicted peptide falls (NS3p, NS3h, NS4, NS5a, NS5b1, NS5b2).

Figure 3-16: HCV specific T cell responses to HLA predicted peptides.



Comparison of T cell responses to HLA predicted peptides and overlapping pools (only non-structural proteins) for acute, chronic and spontaneously resolved patients. Responses were classified as either: no response to both peptide sets, response to only OPs or HLA predicted peptides, response to both (same pool), response to both (different pools). HLA: HLA predicted peptides; OPs: overlapping peptide pools; SR: spontaneously resolved patients.

The highest overlap between two analyses was found in ***acute patients***, where 3 out of 16 (19%) patients responded to matching pools from both sets. In two acute patients, the responses were mapped to single peptide level (for details on peptide mapping of overlapping peptide pools, see Table 3-6 and Table 3-7, page 113 and 114), confirming that the response was indeed formed to an identical epitope (patient 6-40 and 822, Table 3-4). The majority of acute patients in our cohort developed chronic infection; the comparably high level of cross-reactive responses between the two peptide sets in acute patients might be due to patients targeting immunodominant epitopes that subsequently mutate under T cell pressure. Due to the limited availability of PBMC, not all responses to overlapping peptide pools were mapped to

single peptide level. This was the case for acute patient 50163, and chronic patients 219 and 437. It is therefore possible, that, even though the same pools were targeted in both analyses and T cell specificity may be identical in both analyses, the response has not been included in Table 3-4.

In ***chronically infected patients***, overlap between the two peptide approaches was minimal, with only 2 patients (3%) responding to matching pools. However, about half of the patients (49%) did not form a detectable T cell response to either of the peptide sets. This might be due to an exhausted T cell response in chronic infection, with IFNγ production dropping below the level of assay detection.

Comparison of T cell targets detected by two peptide sets

Detection of the same CD8+ T cell epitopes by the two tested approaches was minimal (Table 3-4). Immune responses to the same epitope in both peptide sets were only detected in two acutely infected individuals at epitopes IPFYGKAIPI and ATDALMTGY. For peptide VLDDHYKTAL, a response to the equivalent overlapping peptide was detected, but in a different patient.

Table 3-4: Details of T cell responses to HLA predicted peptides in spontaneously resolved, acute and chronic patients.

Protein	Position	3a peptide sequence	HLA type	Patient ID	Patient HLA type						detected using OPs
NS2	886-896	LLYPSLIFDI/LYPSLIFDI	A02/A24	6-23	A*2402	A*0201	B*3502	B*4403	C*0401	C*1601	
			A24	219	A*1101	A*2404	B*18	B*3501	C*0401	C*1203	
			A24	362	A*24		B*3501	B*4403	C*0401	C*0409	
			A02/A24	101	A*0205	A*2402	B*1302	B*4901	C*0602	C*0701	
			A02	299	A*0101	A*0201	B*0702	B*4001	C*0304	C*0702	
	941-951	FQMIILSIGR	B27	6-56	A*0201	A*2601	B*2702	B*3801	C*1203		
				278	A*0201	A*1101	B*1801	B*2705	C*0102	C*1203	
				428	A*0201		B*5701	B*2702	C*0701	C*1501	
NS3	1379-1387	IPFYGKAIPI	B51	6-40	A*0101		B*0801	B*5101	C*0102	C*0701	yes (6-40)
				50163			B*4403	B*5101	C*1402	C*1601	
	1443-1451	ATDALMTGY	A01	822	A*0101		B*5201	B*5701	C*0602	C*1202	yes (822)
				7-70	A*0101	A*2402	B*0801	B*3906	C*0701	C*0702	
				06P	A*0101	A*0301	B*0702	B*0702	C*0702	C*0702	
				416	A*0101	A*0201	B*0801	B*5701	C*0602	C*0701	
NS4B	1853-1862	RVLLDILAGY	A26	6-56	A*0201	A*2601	B*2702	B*3801	C*120301		
				50017	A*0201	A*2601	B*3801	B*4402	C*0501	C*1203	
				110	A*2301*	A*2601*	B*3801	B*4901	C*0701*	C*1203	
				332	A*0101	A*2601	B*4501	B*0801	C*0602	C*0701	
				437	A*0201	A*2601	B*3801	B*2702	C*1203	C*0102	
NS5B	2490-2499	VLDDHYKTAL	A02	416	A*0101	A*0201	B*0801	B*5701	C*0602	C*0701	yes (6-40)
	2508-2516	RVKARMLTI	B08	6-40	A*0101		B*0801	B*5101	C*0102	C*0701	
				871	A*3201	A*0101	B*0801	B*4402	C*0501	C*0701	

HCV gt3a specific T cell responses to HLA predicted peptides are shown. For each peptide, the position within the HCV genome, the peptide sequence, restricting HLA type, responding patients and patients HLA types are shown. Peptides additionally detected with overlapping peptide pools are depicted. Colour coding: yellow (chronic patients), blue (acute patients), orange (spontaneous resolvers).

Several epitopes were only recognized using HLA predicted peptides ((L)LYPSLIFDI, FQMIILSIGR, ATDALMTGY, RVLLDILAGY, RVKARMLTI), but not overlapping peptide pools, suggesting that the exact peptide length matters for MHC presentation and T cell recognition of these specific epitopes.

Another possible reason that these peptides were not recognized using overlapping peptide pools, is that viral variants were included in the overlapping peptide set at HLA predicted peptide sites. To assess this hypothesis, viral sequence information of overlapping peptides was compared to peptide sequences of consensus HLA predicted peptides. In 7 out of 65 cases, sequence differences between the screening HLA predicted peptide (consensus) to the screening overlapping peptide were observed (Table 3-5). Only in two cases, differences were observed at the polymorphic sites (IPFYGKAIP**L**, EK**A**LGLLQR) originally described in Rauch *et al.*'s publication. In our hands, T cells responses measured by ELISpot assays were recognized only one peptide (IPFYGKAIP**L**) with existing sequence differences between the two peptide sets.

Table 3-5: Sequence difference between overlapping peptide pools and HLA predicted peptides.

HCV Protein	Linked HLA	HLA predicted peptide Sequence	Overlapping Peptide Sequence	Additional Mutations
NS2	B4402	ATLGAG**I**LVL	ATLGAG**IL**<u>A</u>LFGFFTL G**IL**<u>A</u>LFGFFTLSPWYKHW	V→A
	B44	AG**I**LVLFGFF		
	C03	I**L**VLFGFFTL		
NS3	B2705	IPFYGKAIP**L**	EGEIPFYGKAIP**I**AQLK	None
	A0201	**A**LLKGGRHLI	YGKAIPI**A**<u>Q</u>LKGGRHLIF **A**<u>Q</u>LKGGRHLIFCHSKKK	L→Q
	B5101	I**A**LLKGGRHL		
NS4	A68	EK**A**LGLLQR	HQFKEK**V**LGLLQRATQQQA	None
NS5a	A2601	ET**D**AELSVA	PLRAET**D**<u>D</u>AELS<u>A</u>AAECF	V→A Extra AA D

HLA predicted peptides with varying overlapping peptide sequence. HLA predicted peptides are specified, with the polymorphism linked to HLA type marked in bold. For the overlapping peptide, the polymorphic site is marked bold, with additional mutations and insertions underlined. AA, amino acid.

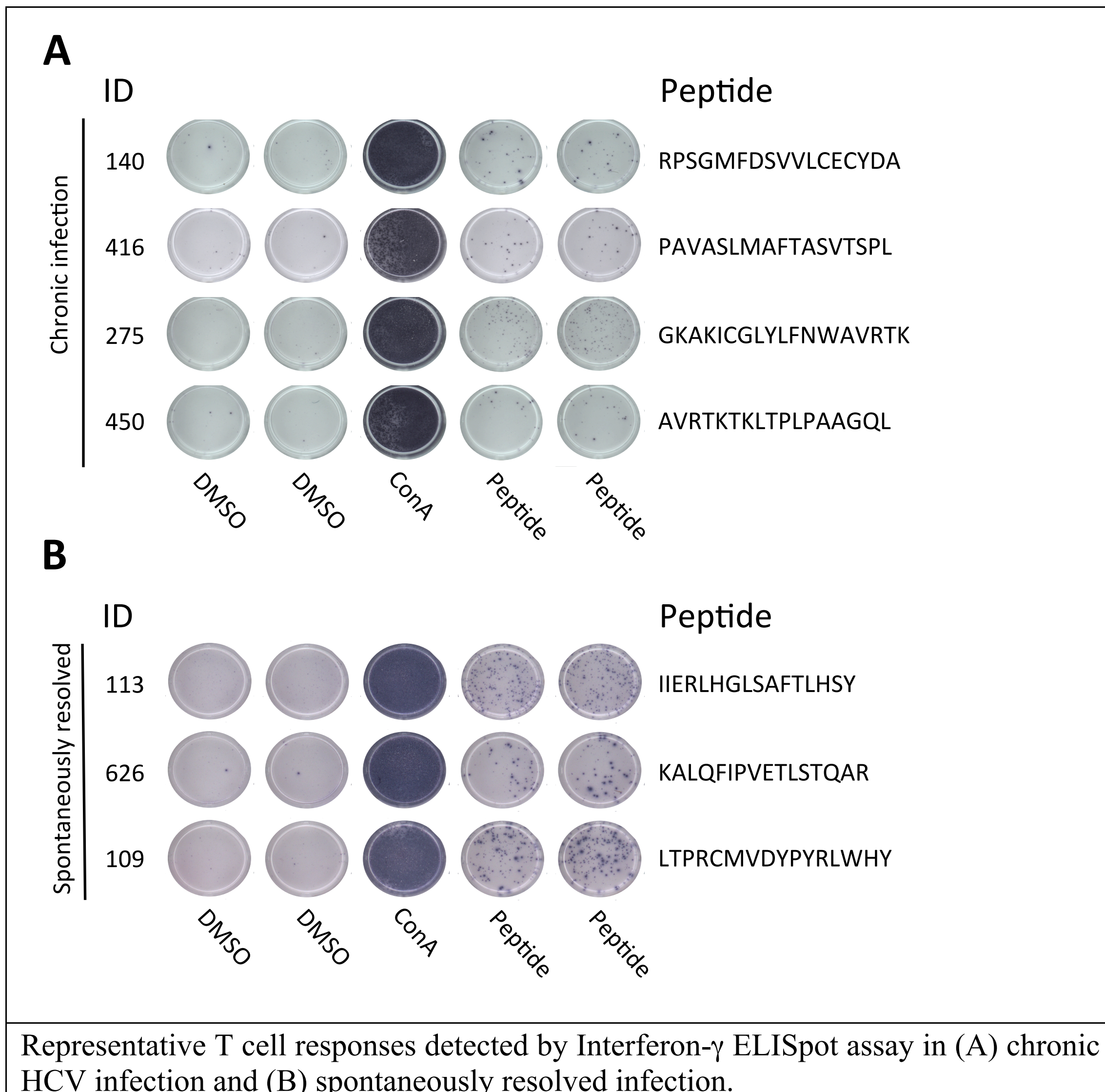
3.8 High resolution analysis of HCV genotype 3 specific T cell responses

Patient samples are often difficult to obtain and the number of PBMC generated from each sample is limited, so experimental strategies have to be optimized to increase the amount of information generated from one sample. To comprehensively assess T cell responses in HCV infection, a set of 450 peptides is necessary to span the whole genome (using peptides of 15 to 18 amino acids in length, overlapping by 11 amino acids). Since 450 individual peptides cannot be experimentally assessed separately in each patient, peptide-screening strategies have therefore been either based on matrixes combining different sets of peptides in each well, or peptide pools. In our experiments, we used the peptide pool approach, screening samples with 10 peptide pools, each containing up to 72 peptides, and in case of a detected T cell responses, subsequently mapping responses to sub-pool and single peptide level (see Figure 2-4, page 77).

3.8.1 Confirmation of T cell responses measured by IFNγ ELISpot assay

One issue with using overlapping pools and subsequently mapping responses to single peptides is that, even if the response to the overlapping pool was strong and easily detectable, the response to a single mapped peptide can be very low, and often range at the detection cut-off. Representative ELISpot responses to single peptides are depicted in Figure 3-17, with low responses (some just above detection cut-off) shown for chronically infected individuals, and stronger responses depicted for spontaneously resolved infection.

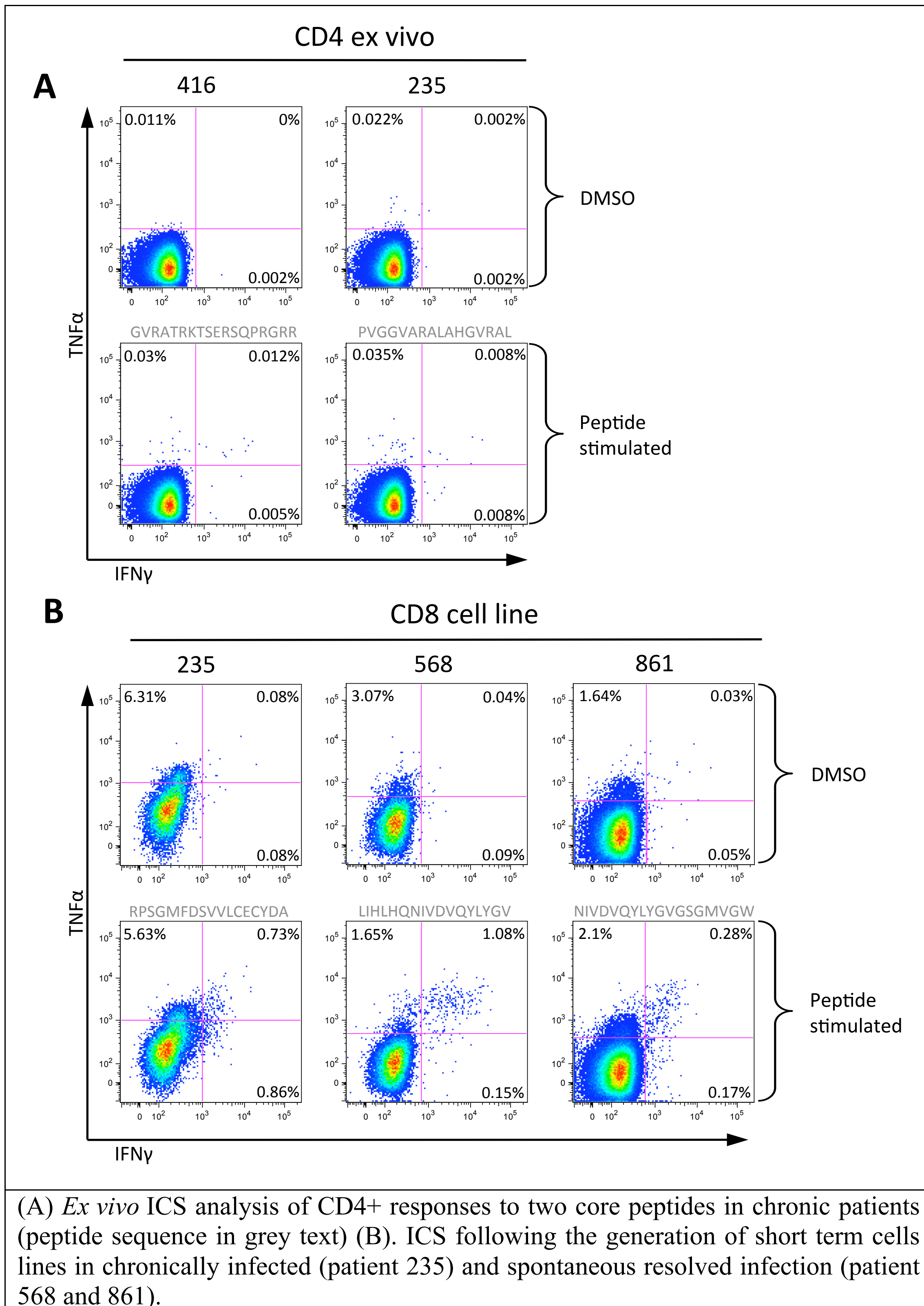
Figure 3-17: Representative T cell responses measured by IFNγ ELISpot assay.



Representative T cell responses detected by Interferon-γ ELISpot assay in (A) chronic HCV infection and (B) spontaneously resolved infection.

We therefore first confirmed that T cell responses measured at ELISpot detection level to single peptides were not measured false positives. To confirm responses using a different technique, we analysed responses using intracellular staining techniques (ICS) and flow cytometry analysis, both on patient samples *ex vivo* and T cell lines (Figure 3-18).

Figure 3-18: CD4+ and CD8+ specific T cell responses measured by ICS.



(A) *Ex vivo* ICS analysis of CD4+ responses to two core peptides in chronic patients (peptide sequence in grey text) (B). ICS following the generation of short term cells lines in chronically infected (patient 235) and spontaneous resolved infection (patient 568 and 861).

3.8.2 Individual peptides targeted by T cells in HCV gt3 infection

T cell responses to overlapping peptide pools were mapped to individual peptides using the peptide mapping strategy described in the methods section (Figure 2-4, page 77). In this thesis, individual overlapping peptides targeted by T cells will be termed as "T cell targets", whereas T cell targets identified using HLA predicted peptides are termed as "epitopes", since the optimal binding peptides have been defined for these using epitope prediction programs.

Responses to individual peptides in the structural HCV region are shown in Table 3-6. Eleven regions were targeted by T cells from spontaneously resolved, acutely and chronically infected HCV gt3 patients, six single peptides in the core region, and five in E2. No T cell responses to E1 were detected. If responses were found to two neighbouring peptides, they were grouped as one response in this table, assuming that it was likely that the response was targeting the overlapping region between peptides. It is possible though, that two separate epitopes are targeted.

Table 3-6: T cell responses to individual peptides in HCV structural regions mapped from overlapping peptide pools.

Protein	Position	3a peptide sequence	Patient ID	Patient HLA type						Patient HLA type		CD8/CD4
core	27-44	GGQIVGGVYVLPRRGPRL	362	A*2402		B*3501	B*4403	C*0401	C*0409	DRB1*0103	DRB1*0701	
	34-51	VYVLPRRGPRLGVRATRK	06P	A*0101	A*0301	B*0702	B*0702	C*0702	C*0702			
			884	A*3201	A*2501	B*5701	B*5101	C*0601	C*1502	DRB1*0404	DRB1*0701	
	66-83	PKARRSEGRSWAQPGYPW	105	A*1101	A*1101	B*3501	B*5101	C*0401	C*1402	DRB1*0101	DRB1*0701	
			128	A*0201	A*6801	B*0801	B*4402	C*0701	C*05	DRB1*0301	DRB1*0401	
			275	A*1101		B*0702(	B*4402	C*0501	C*0702	DRB1*0401	DRB1*0407	
			450									CD4 *
	73-90	GRSWAQPGYPWPLYGNEG	416	A*0101	A*0201	B*0801	B*5701	C*0602	C*0701	DRB1*0101	DBR1*0701	CD4 *
	130-147	FADLMGYIPLVGAPVGGV	140	A*2402	A*3002	B*0702	B*3501			DRB1*0101	DRB1*1501	
			105	A*1101	A*1101	B*3501	B*5101	C*0401	C*1402	DRB1*0101	DRB1*0701	
	143-158	PVGGVARALAHGVRAL	6-56	A*0201	A*2601	B*2702	B*3801	C*1203				
			110	A*2301	A*2601	B*3801	B*4901	C*0701	C*1203	DRB1*0801	DRB1*1101	CD4
			128	A*0201	A*6801	B*0801	B*4402	C*0701	C*05	DRB1*0301	DRB1*0401	CD4 *
			129	A*1101	A*7401	B*4403	B*38	C*04	C*0702	DRB1*1504	DRB1*0701	
			216	A*0201	A*6801	B*5101	B*5108	C*1502	C*1602	DRB1*0301	DRB1*1602	CD4 *
			219	A*1101	A*2402	B*18	B*3501	C*0401	C*1203	DRB1*0403	DRB1*0701	
			226	A*0201	A*3101	B*1801	B*55	C*0102	C*0701	DRB1*0404	DRB1*1104	CD4 *
			227	A*0201		B*1501	B*5701(	C*0401	C*0602	DRB1*0701	DRB1*0801	CD4 *
			235	A*0301	A*3002	B*0702	B*3501	C*0401	C*0702	DRB1*0101	DRB1*1501	
			331	ND								CD4
			422	A*02	A*03	B*44	B*35	C*0401		DRB1*0101	DRB1*1501	
			437	A*0201	A*2601	B*3801	B*2702	C*1203	C*0102	DRB1*0101		CD4
	148-165	ARALAHGVRALEDGINFA	962									CD8
E2	453-476	CPQRLSSCKPITFFRQGWGSLTDA	822	A*0101		B*5201	B*5701	C*0602	C*1202	DRB1*0701	DRB1*1502	CD4
		FRQGWGSLTDANITGPSD	06P	A*0101	A*0301	B*0702	B*0702	C*0702	C*0702			
	610-625	LTPRCMVDYPYRLWHY	06P	A*0101	A*0301	B*0702	B*0702	C*0702	C*0702			
			109	A*0201	A*0101	B*0801	B*4402	C*0501	C*0701	DRB1*0401	DRB1*0301	
			900	A*0201	A*6601	B*3502	B*1401	C*1701	C*0802	DRB1*1104	DRB1*0103	
	635-650	KVRMFVGGFEHRFTAA	7-75	A*0201	A*3101	B*0702	B*0702	C*0702	C*0702			
			129	A*1101	A*7401	B*4403	B*38	C*04	C*0702	DRB1*1504	DRB1*0701	
	696-712	LIHLHQNIVDVQYLYGV	568	A*02	A*3201	B*2707	B*1501	C*0602	C*0304	DRB1*1501	DRB1*1101	CD8 *
	702-719	NIVDVQYLYGVGSGMVGW	6-56	A*0201	A*2601	B*2702	B*3801	C*1203				
			861	ND								CD8 *

Mapped HCV gt3a specific T cell responses to single peptides in structural regions of the HCV genome are shown. For each peptide, the position within the HCV genome, the peptide sequence, responding patients and patients HLA types, as well as CD8/CD4 restriction is shown. * CD8/CD4 restriction determined by ICS; colour coding: yellow (chronic patients), blue (acute patients), orange (spontaneous resolvers).

T cell responses targeting individual peptides in HCV non-structural regions were measured by IFNγ ELISpot assay in spontaneous resolvers, acutely and chronically infected patients (Table 3-7).

Table 3-7: T cell responses to individual peptides in HCV non-structural regions mapped from overlapping peptide pools.

Protein	Position	3a peptide sequence	Patient ID	Patient HLA type						Patient HLA type		CD8/CD4
NS3	1198-1213	KALQFIPVETLSTQAR	626	ND								
	1246-1261	KVPAAYVAQGYNVLVL	806	A*0101	-	B*4002	B*1501	C*0202	C*0303	DRB1*0101	DRB1*0303	
	1264-1281	SVAATLGFGSFMSRAYGI	7-75	A*0201	A*3101	B*0702	B*0702	C*0702	C*0702			
			720	A*0301	A*0201	B*4402	B*3801	B0501	C*1203	DRB1*0401	DRBA*1601	
	1282-1298	DPNIRTGNRTVTTGAKL	109	A*0201	A*0101	B*0801	B*4402	C*0501	C*0701	DRB1*0401	DRB1*0301	
			900	A*0201	A*6601	B*3502	B*1401	C*1701	C*0802	DRB1*1104	DRB1*0103	
	1370-1387	EEVALGSEGEIPFYGKAI	6-40	A*0101		B*0801	B*5101	C*0102	C*0701			CD8
	1423-1440	AYYRGLDVSVIPTAGDVV	226	A*0201	A*3101	B*1801	B*55	C*0102	C*0701	DRB1*0404	DRB1*1104	
			884	A*3201	A*2501	B*5701	B*5101	C*0601	C*1502	DRB1*0404	DRB1*0701	CD4
	1437-1451	GDVVVCATDALMTGF	822	A*0101		B*5201	B*5701	C*0602	C*1202	DRB1*0701	DRB1*1502	
	1520-1537	RPSGMFDSVVLCECYDAGCSWYDL	105	A*1101	A*1101	B*3501	B*5101	C*0401	C*1402	DRB1*0101	DRB1*0701	
			111	A*0101	A*0201	B*35	B*14	C*0401	C*0802	DRB1*0101	DRB1*0701	
			140	A*2402	A*3002	B*0702	B*3501			DRB1*0101	DRB1*1501	CD8
			154	A*1101	A*1101	B*1517	B*3501	C*0401	C*0701	DRB1*0101	DRB1*1301	
			219	A*1101	A*2402	B*18	B*3501	C*0401	C*1203	DRB1*0403	DRB1*0701	
			235	A*0301	A*3002	B*0702	B*3501	C*0401	C*0702	DRB1*0101	DRB1*1501	CD8
			268									
			291	A*0101	A*2901	B*3501	B*4403C	C*0401	C*1601	DRB1*0103	DRB1*0701	
			295	A*0301	A*2402	B*3501	B*5501C	C*0301	C*0401	DRB1*0101	DRB1*0901	
			362	A*2402		B*3501	B*4403	C*0401	C*0409	DRB1*0103	DRB1*0701	CD8
			750									
			787	A*0201	A*1101	B*3501	B*4001	C*0401	C*0301	DRB1*0101	DRB1*1501	
			906	A*0301	A*2601	B*3501	B*2701	C*0401	C*0202	DRB1*0404	DRB1*1301	
			955	A*0201	A*1101	B*4402	B*3501	C*0401	C*0501	DRB1*0401	DRB1*1102	
	1547-1569	RAYLSTPGLPVCQDHLDFWESVF	219	A*1101	A*2402	B*18	B*3501	C*0401	C*1203	DRB1*0403	DRB1*0701	
NS4B	1792-1808	PAVASLMAFTASVTSPL	6-23	A*2402	A*0201	B*3502	B*4403	C*0401	C*1601			
			416	A*0101	A*0201	B*0801	B*5701	C*0602	C*0701	DRB1*0101	DRB1*0701	CD8
			437	A*0201	A*2601	B*3801	B*2702	C*1203	C*0102	DRB1*0103	DRB1*0101	CD8
	1805-1822	TSPLTTNQTMFFNILGGW	806	A*0101	-	B*4002	B*1501	C*0202	C*0303	DRB1*0101	DRB1*0303	
			906	A*0301	A*2601	B*3501	B*2701	C*0401	C*0202	DRB1*0404	DRB1*1301	
	1825-1842	THLAGPQSSSAFVVSGLA	437	A*0201	A*2601	B*3801	B*2702	C*1203	C*0102	DRB1*0103	DRB1*0101	N/A
	1917-1932	EGAVQWMNRLIAFASR	410	A*0101	A*3001	B*1302	B*4402	C*0602	C*0501	DRB1*0701		CD8
NS5A	2030-2047	GVMSTRCPCGASIAGHVK	129	A*1101	A*7401	B*4403	B*38	C*04	C*0702	DRB1*1504	DRB1*0701	CD8
	2119-2136	CPCQVPAAEFFTEVDGVR	6-23	A*2402	A*0201	B*3502	B*4403	C*0401	C*1601			
	2126-2141	AEFFTEVDGVRLHRYA	6-40	A*0101		B*0801	B*5101	C*0102	C*0701			CD8
			822	A*0101		B*5201	B*5701	C*0602	C*1202	DRB1*0701	DRB1*1502	
			871	A*3201	A*0101	B*0801	B*4402	C*0501	C*0701	DRB1*0103	DRB1*0101	
			975	A*3201	A*0301	B*1401	B*0702	C*0802	C*0702	DRB1*0101	DRB1*0701	
	2145-2162	KPLLRDEITFMVGLNSYA	7-31	A*0101	A*0301	B*0801	B*1801	C*0701	C*0701			
NS5B	2484-2499	TFDRLQVLDDHYKTAL	6-40	A*0101		B*0801	B*5101	C*0102	C*0701			CD8
NS5B	2548-2565	NQIRSVWEDLLEDTTTPI	113	A*0301	A*0101	B*4402	B*4001	C*0501	C*0304	DRB1*0701	DRB1*1454	CD4
	2603-2618	KRALYDVIQKLSIETM	884	A*3201	A*2501	B*5701	B*5101	C*0601	C*1502	DRB1*0404	DRB1*0701	CD4
	2844-2861	IMYAPTIWVRMVMMTHFF	50017	A*0201	A*2601	B*3801	B*4402	C*0501	C*1203			
	2893-2908	IIERLHGLSAFTLHSY	113	A*0301	A*0101	B*4402	B*4001	C*0501	C*0304	DRB1*0701	DRB1*1454	CD4
	2947-2964	GKAKICGLYLFNWAVRTK	275	A*1101		B*0702	B*4402	C*0501	C*0702	DRB1*0104	DRB1*0115	CD8
	2967-2976	KLTPLPAAGQL	450	ND								CD8

Mapped HCV gt3a specific T cell responses to single peptides in structural regions of the HCV genome are shown. For each peptide, the position within the HCV genome, the peptide sequence, responding patients and patients HLA types, as well as CD8/CD4 restriction is shown. * CD8/CD4 restriction determined by ICS; colour coding: yellow (chronic patients), blue (acute patients), orange (spontaneous resolvers).

T cell responses in **chronic patients** (Table 3-6 and Table 3-7) mainly targeted two peptides: a core peptide targeted by CD4+ T cells (PVGGVARALAHGVRAL), and two overlapping peptides targeted by CD8+ T cells in the NS3 region (RPSGMFDSVVLCECYDAGCSWYDL). Additionally to chronically infected patients targeting these individual peptides, the CD4 peptide was targeted by an acutely infected patient (6-56) later developing chronic infection and not responding to treatment, and the CD8 peptide was targeted by two patients (906, 955) with spontaneously resolved HCV infection. Further individual peptides targeted by chronically infected patients were located in the core region (PKARRSEGRSWAQPGYPW, CD4+ restricted; and FADLMGYIPLVGAPVGGV, restriction not defined) and NS4B region (PAVASLMAFTASVTSPL, CD8+ restricted, targeted by two chronic patients and one acute patient).

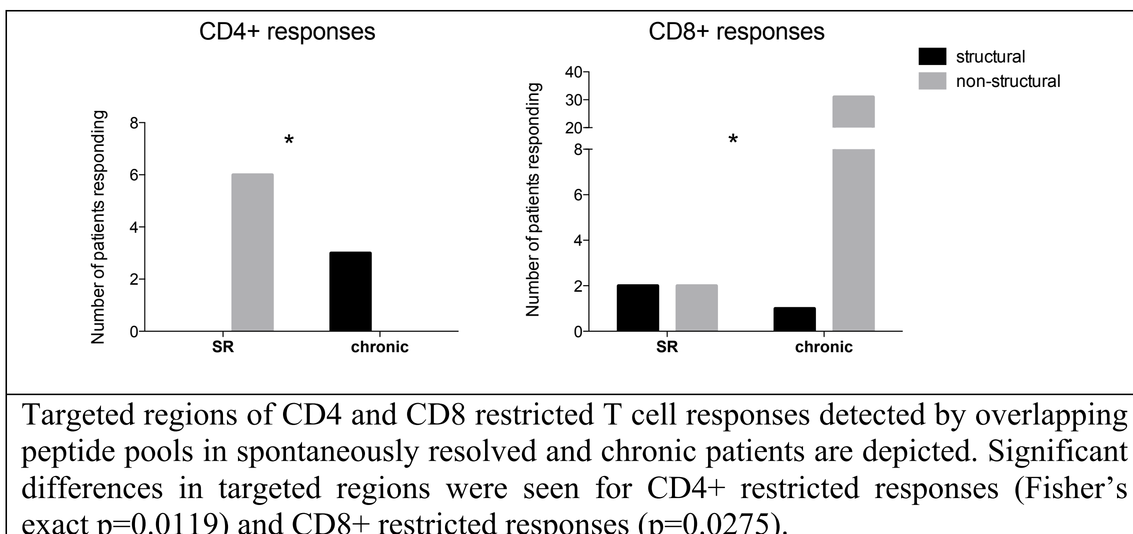
The pattern of T cell responses to single peptides in **acute patients** (Table 3-6 and Table 3-7) was similar to that seen in spontaneous resolvers: a wide range of individual peptides was targeted, with the majority of responses formed to non-structural proteins. In three cases, identical single peptides were targeted in acute patients and spontaneous resolvers (NIVDVQYLYGVGSGMVGW, E2, 6-56 and 861; SVAATLGFGSFMSRAYGI, NS3, 7-75 and 720; AEFFTEVDGVRLHRYA, NS5A, 6-40/882 and 871/975). This result is somewhat unexpected, since most of acutely infected patients did not clear infection spontaneously. Acute patient 6-56 did not respond to treatment, acute patients 6-40 and 822 both cleared the infection with IFNα/Ribavirin treatment, and patient 7-75 was lost to follow up (see Table 3-3, page 31 for details). It is possible, though, that responses to identical peptides are formed during acute infection in all patients, regardless of outcome, but are subsequently lost in chronic infection. This hypothesis would be in line with observations from a recent study observing initially high CD4+ responses in acute HCV infection regardless of HCV outcome (Schulze Zur Wiesch et al. 2012). Several T cell targets were also shared with chronically infected patients: the dominant core (PVGGVARALAHGVRAL) and the NS4B peptide (PAVASLMAFTASVTSPL) mentioned in the previous paragraph, and two E2 peptides (FRQGWGSLTDANITGPSD, KVRMFVGGFEHRFTAA) were targeted by both acutely and chronically infected patients.

Mapped T cell responses to single peptides in **spontaneous resolvers** (Table 3-6 and Table 3-7) were spread over all HCV structural and non-structural proteins, targeting

a wide range of peptides. In four cases, more than one patient with spontaneously resolved infection targeted identical individual peptides: peptide LTPRCMVDYPYRLWHY (E2, targeted in patients 109 and 900, and targeted by chronic patient 06P), peptide DPNIRTGNRTVTTGAKL (NS3, targeted in patients 109 and 900), peptide RPSGMFDSVVLCECYDAGCSWYDL (NS3, targeted in patients 906 and 955, also frequently targeted in chronic infection), and peptide TSPLTTNQTMFFNILGGW, (NS4B, in patients 806 and 906). Of note, patients 109 and 900 both targeted two identical individual peptides.

Next, CD4/CD8 restriction of T cell responses was determined wherever enough cells were available (Table 3-6 and Table 3-7). Restriction was either defined using CD8+ depletion IFNγ ELISpot assays (see Figure 2-2 for an example, page 74), or through intracellular staining (marked with a star in Table 3-6 and Table 3-7). The distribution of CD4+ and CD8+ responses varied significantly in different patient groups (Figure 3-19), with CD4+ T cells in spontaneous resolvers only targeting HCV non-structural proteins, whereas CD4+ cells in chronically infected patients only targeted structural proteins, mainly core. Contrary, CD8+ cells from chronic patients mainly targeted non-structural proteins (except for one CD8+ response in chronic patient 962 to core peptide ARALAHGVRALEDGINFA). CD8+ cells in spontaneous resolvers targeted both structural and non-structural regions equally.

Figure 3-19: CD4/CD8 responses in spontaneous resolvers and chronic patients.



Targeted regions of CD4 and CD8 restricted T cell responses detected by overlapping peptide pools in spontaneously resolved and chronic patients are depicted. Significant differences in targeted regions were seen for CD4+ restricted responses (Fisher's exact p=0.0119) and CD8+ restricted responses (p=0.0275).

These observations are in in line with previously published data suggesting that a strong CD4+ response is required for HCV clearance (Diepolder et al. 1995; K. M. Chang et al. 2001; Thimme et al. 2001; C. L. Day et al. 2002; Rehermann 2009). In

addition, a recent study suggests that CD4+ responses are primed early in acute infection, but disappear quickly in patients developing chronic infection (Schulze Zur Wiesch et al. 2012). The phenomenon of CD8+ T cell responses being readily detectable in chronically infected patients, and CD4 responses in spontaneous resolved infection, has also been observed previously (K. M. Chang et al. 2001).

Overall, single peptides targeted in spontaneous resolvers, acute and chronic patients did vary, with distinct preferential regions targeted in chronic patients (core, NS3) and acute patients/spontaneous resolvers (broad responses to all viral proteins). A striking feature was the differential distribution of CD4+/CD8+ responses in spontaneous resolvers and chronic patients. However, even though a differential CD4+/CD8+ distribution was observed, different patient groups did target identical individual peptides, suggesting that not the targeted epitope but the phenotype of the cells has a dominant influence on outcome of infection.

3.8.3 Viral sequence analysis at individual peptides targeted by T cells

It was shown that HLA class-I restricted T cell pressure promotes viral evolution by causing mutations within CD8 epitopes (Ruhl et al. 2011). To establish whether the detected T cell responses are targeting circulating virus, viral sequence analysis was attempted in all samples with a detectable T cell responses response to individual peptides and plasma available. Methods for RNA isolation, first and second round PCR and sequencing are described in detail in the methods section (2.2.1 to 2.2.8, pages 67 to 71). Results of viral sequence analysis are shown for structural and non-structural viral regions in Table 3-8 and Table 3-9, respectively.

In acutely and chronically HCV gt3 infected patients, 28 independent T cell responses to single peptides were detected in the structural regions, and 50 to peptides in non-structural regions. Sequence information was obtained in 26/28 cases (93%) for non-structural T cell targets, and 31/50 cases (62%) in non-structural regions.

Sequence polymorphisms were rarely observed at individual peptides targeted by CD4+ cells in the core region (Table 3-8). Of 21 patients sequenced at 5 individual peptides targeted by CD4+ T cells, a sequence polymorphism was only observed in one patient (5%, P**V**GGVARALAHGVRAL/**A**).

Sequence polymorphisms were more common within individual targeted peptides sequenced in the E2 region (2/6, 33%), where polymorphisms were seen in a chronically infected patient (06P) at two T cell targets, with two amino acid

differences observed at peptide F**R**QGWGSLTDANI**T**GPSD/**N,S** and one amino acid difference observed at peptide LTPRC**M**VDYPYRLWHY/**L**. However, no sequence polymorphisms were observed in three acutely infected patients targeting individual peptides within the E2 region.

Table 3-8: Viral consensus sequence analysis at T cell targets detected in HCV structural regions.

Protein		3a peptide sequence	Patient	Patient Viral Sequence at Pre-Treatment
core	27-44	GGQIVGGVYVLPRRGPRL	362	GGQIVGGVYVLPRRGPRL
	34-51	VYVLPRRGPRLGVRATRK	06P	VYVLPRRGPRLGVRATRK
	66-83	PKARRSEGRSWAQPGYPW	105	ND
			128	PKARRSEGRSWAQPGYPW
			275	PKARRSEGRSWAQPGYPW
			450	PKARRSEGRSWAQPGYPW
	73-90	GRSWAQPGYPWPLYGNEG	416	GRSWAQPGYPWPLYGNEG
	130-147	FADLMGYIPLVGAPVGGV	140	FADLMGYIPLVGAPVGGV
			105	ND
	143-158	PVGGVARALAHGVRAL	6-56	PVGGVARALAHGVRAL
			110	P**A**GGVARALAHGVRAL
			128	PVGGVARALAHGVRAL
			129	PVGGVARALAHGVRAL
			216	PVGGVARALAHGVRAL
			219	PVGGVARALAHGVRAL
			226	PVGGVARALAHGVRAL
			227	PVGGVARALAHGVRAL
			235	PVGGVARALAHGVRAL
			331	PVGGVARALAHGVRAL
			422	PVGGVARALAHGVRAL
			437	PVGGVARALAHGVRAL
	148-165	ARALAHGVRALEDGINFA	962	ARALAHGVRALEDGINFA
E2	453-476	CPQRLSSCKPITFFRQGWGSLTDA	822	CPQRLSSCKPITFFRQGWGSLTDA
		FRQGWGSLTDANITGPSD	06P	F**N**QGWGSLTDANI**S**GPSD
	610-625	LTPRCMVDYPYRLWHY	06P	LTPRC**L**VDYPYRLWHY
	635-650	KVRMFVGGFEHRFTAA	7-75	KVRMFVGGFEHRFTAA
			129	KVRMFVAGFEHRFTAA
	702-719	NIVDVQYLYGVGSGMVGW	6-56	NIVDVQYLYGVGSGMVGW

Viral sequence analysis at single overlapping peptides in patients with a detectable T cell response in the HCV structural region. The peptide sequence, amino acid position, patients with a detectable T cell response (colour coded: yellow-chronic patients, blue-acute patients), and the according circulating viral sequence are shown. Sequence polymorphisms in the circulating viral sequence compared to the peptide sequence are marked red. ND not determined (insufficient RNA, or plasma sample not available).

When analysing sequence polymorphisms within the non-structural regions, fairly conserved sequences were observed for the dominant CD8+ epitope contained in individual peptide RPSGMFDSVVLCECYDAGCSWYDL in the NS3 region, with only one out of seven sequenced patients showing a mutation (Table 3-9).

On the other hand, sequence polymorphisms were observed for the majority of sequenced patients at the CD8+ restricted epitope ATDALMTGY. Mutational escape at this epitopes has been described previously in the literature for HCV gt1 and 3 infected patients (Neumann-Haefelin et al. 2008). Of note, in the acute patient where a T cell response to this epitope was detected, the viral sequencing showed the wild type ATDALMTG**Y**, whereas the variant version ATDALMTG**F** was observed in all chronic patients (Table 3-9). Similarly, sequence polymorphisms within T cell epitopes were also observed at the newly described, GT3 specific CD8+ restricted epitopes RVLLDILAGY in two sequenced patients (**K**VLLDILAGY) and FQM**I**ILS**I**GR, with three diverse sequence polymorphisms seen (FQM**A**ILSIGR, FQMIILS**V**GR, and FQM**V**ILSIG**K**). Sequence polymorphisms in several individual peptides targeted by T cells were also commonly observed in the NS5A and NS5B region (Table 3-9).

Overall, sequence polymorphisms were rarely observed at CD4+ T cell targets detected in chronic infection, but commonly seen at CD8+ restricted targets detected using HLA predicted peptides. For further detailed analysis and assessment of experimental T cell cross-reactivity towards sequence variants, refer to Chapter 4.

Table 3-9: Viral sequence analysis at T cell targets detected in non-structural regions. (Table on the next page).

Viral sequence analysis at targeted viral regions in patients with a detectable T cell response to single overlapping peptides in the HCV non-structural region. The peptide sequence with the location of the individual peptide, the responding patients (colour coded: yellow-chronic patients, blue-acute patients), and the according circulating viral sequence are shown. Sequence polymorphisms in the circulating viral sequence compared to the peptide sequence are marked red. ND not determined (insufficient RNA, or plasma sample not available).

Protein		3a peptide sequence	Patient	Patient Viral Sequence at Pre-Treatment
NS2	886-896	LLYPSLIFDI	6-23	LLYPSLIFDI
		LYPSLIFDI	6-23	LYPSLIFDI
			101	LYPSLIFDI
			219	LYPSLIFDI
			299	LYPSLIFDI
			362	LY**S**SLIFDI
	941-951	FQMIiLSIGR	6-56	FQM**A**ILSIGR
			278	FQMIILG**V**GR
			428	FQM**V**ILSIG**K**
NS3	1264-1281	SVAATLGFGSFMSRAYGI	7-75	SVAATLGFGSFMS**H**AYGI
	1370-1387	EEVALGSEGEIPFYGKAI	6-40	ND
	1379-1387	IPFYGKAIPI	6-40	ND
			50163	ND
	1423-1440	AYYRGLDVSVIPTAGDVV	226	AYYRGLDVSVIPTAGDVV
	1436-1447	GDVVVCATDALMTGF	822	GDVVVCATDALMTG**Y**
	1442-1447	ATDALMTGY	822	ATDALMTG**Y**
			7-70	ATDALMTG**F**
			06P	ATDALMTG**F**
			416	ATDALMTG**F**
	1520-1537	RPSGMFDSVVLCECYDAGCSWYDL	105	ND
			111	RPSGMFDSVVLCECYDAGCSWYDL
			140	RPSGMFDSVVLCECYDAGCSWYDL
			154	RPSGMFDSVVLCECYDAGCSWYDL
			219	RPSGMFDSVVLCECYDAGCSWYDL
			235	RPSGMFDSVVLCECYDAGC**A**WYDL
			268	ND
			291	ND
			295	RPSGMFDSVVLCECYDAGCSWYDL
			362	RPSGMFDSVVLCECYDAGCSWYDL
			750	ND
			787	ND
	1547-1569	RAYLSTPGLPVCQDHLDFWESVF	219	RAYLSTPGLPVCQDHLDFWESVF
NS4B	1792-1808	PAVASLMAFTASVTSPL	6-23	ND
			416	PAVASLMAFTASVTSPL
			437	ND
	1825-1842	THLAGPQSSSAFVVSGLA	437	ND
	1853-1862	RVLLDILAGY	6-56	**K**VLLDILAGY
			50017	ND
			110	**K**VLLDILAGY
			332	ND
			437	ND
	1917-1932	EGAVQWMNRLIAFASR	410	EGAVQWMNRLIAFASR
NS5A	2030-2047	GVMSTRCPCGASIAGHVK	129	GVMSTRCPCGASI**T**GHVK
	2119-2136	CPCQVPAAEFFTEVDGVR	6-23	CPCQVPA**P**EFFTEVDGVR
	2126-2141	AEFFTEVDGVRLHRYA	6-40	ND
			822	ND
	2145-2162	KPLLRDEITFMVGLNSYA	7-31	ND
NS5B	2484-2499	TFDRLQVLDDHYKTAL	6-40	TFDRLQVLDDHYKTAL
	2490-2499	VLDDHYKTAL	416	ND
	2508-2516	RVKARMLTI	6-40	RVKARMLTI
	2844-2861	IMYAPTIWVRMVMMTHFF	50017	ND
	2947-2964	GKAKICGLYLFNWAVRTK	275	GKAKI**T**GLYLFNWAVRTK
	2967-2976	KLTPLPAAGQL	450	KLTPLPAAG**L**L

3.9 Comparison of T cell specificity in HCV gt1 and gt3 infection

Cross-reactivity between HCV gt1 and 3 T cell responses is an important consideration for T cell vaccine design. To date, it is not clear whether T cells primed in infection with different genotypes target identical viral regions. To ultimately design a single vaccine that covers different viral strains and genotypes, more information is needed on T cell immunity in natural infection in different genotypes.

To assess cross-reactivity between T cell immunity previously defined in the well-studied genotype HCV gt1 and immunity assessed in HCV gt3 in the Oxford cohort in this thesis, I analysed:

(1) Whether responses detected in the Oxford HCV gt3 cohort were previously described in the literature in HCV gt1 infection.

(2) Whether dominant HCV gt1 epitopes were targeted in the Oxford HCV gt3 cohort.

(3) HCV gt1 and 3 cross-reactivity on single peptide level for peptides targeted in HCV gt3 patients.

3.9.1 Were T cell targets detected in the Oxford HCV gt3 cohort previously described in the literature?

To assess whether T cell targets detected in HCV gt3a infection in the Oxford cohort were previously described in the literature, a comprehensive assessment of T cell targets published on the immune epitope database (IEDB) at sites of HCV gt3 specific immune response was performed. Methods for this analysis are described in detail in the methods section (2.6.1 Analysis of epitopes described in natural HCV infection (IEDB), page 81). In brief, T cell targets previously described for HCV gt1 and gt3 were obtained from the immune epitope database resource (IEDB, www.iedb.com). Crosschecking the primary publications for the described immune responses ensured quality control. Peptide overlap was defined as a sequence overlap of more than 3 amino acids. Table 3-10 shows the results of the comparative analysis between 16 CD8+ restricted T cell targets described in the Oxford cohort and previously published CD8+ epitopes.

Table 3-10: CD8+ restricted T cell targets previously described in the literature in the region of CD8+ restricted T cell targets defined in this study.

GT3 CD8+ epitopes (Oxford Study Cohort)													GT1 epitopes described in corresponding regions			
Position	3a peptide sequence	Viral protein	HLA	Patient ID	HLA type								Peptide (Literature)		HLA	First author
941-951	FQMIiLSIGR	NS2	B27	6-56	A*0201	A*2601	B*2702	B*3801	C*1203		#		no CD8 epitopes described			
				278	A*0201	A*1101	B*1801	B*2705	C*0102	C*1203	#					
				428	A*0201		B*5701	B*2702	C*0701	C*1501	#					
1917-1932	EGAVQWMNRLIAFASR	NS4B		410	A*0101	A*3001	B*1302	B*4402	C*0602	C*0501	#	*	no CD8 epitopes described			
2030-2047	GVMSTRCPCGASIAGHVK	NS5A		129	A*1101	A*7401	B*4403	B*38	C*04	C*0702	#	*	no CD8 epitopes described			
2484-2499	TFDRLQVLDDHYKTAL	NS5B		6-40	A*0101		B*0801	B*5101	C*0102	C*0701	#		no CD8 epitopes described			
2490-2499	VLDDHYKTAL		A02	416	A*0101	A*0201	B*0801	B*5701	C*0602	C*0701	#					
2508-2516	RVKARMLTI	NS5B	B08	6-40	A*0101		B*0801	B*5101	C*0102	C*0701	#		no CD8 epitopes described			
				871	A*3201	A*0101	B*0801	B*4402	C*0501	C*0701						
2967-2976	KLTPLPAAGQL	NS5B		450							#	*	no CD8 epitopes described			
1853-1862	RVLLDILAGY	NS4B	A26	6-56	A*0201	A*2601	B*2702	B*3801	C*120301		#		ILAGYGAGV		A2	M Battegay
				50017	A*0201	A*2601	B*3801	B*4402	C*0501	C*1203	#		ILAGYGAGV		A2	N H Gruener
				110	A*2301*	A*2601*	B*3801	B*4901	C*0701*	C*1203	#		ILAGYGAGV		A2	T Kuntzen
				332	A*0101	A*2601	B*4501	B*0801	C*0602	C*0701	#					
				437	A*0201	A*2601	B*3801	B*2702	C*1203	C*0102	#					
886-896	LLYPSLIFDI/LYPSLIFDI	NS2	A02/A24	6-23	A*2402	A*0201	B*3502	B*4403	C*0401	C*1601	#		**H**P**T**L**V**FDITK		Class I	A L Cox
				101	A*1101	A*2404	B*18	B*3501	C*0401	C*1203	#		**H**P**T**L**V**FDITKL		Class I	T Kuntzen
				219	A*24		B*3501	B*4403	C*0401	C*0409	#					
				299	A*0205	A*2402	B*1302	B*4901	C*0602	C*0701	#					
				101	A*0205	A*2402	B*1302	B*4901	C*0602	C*0701	#					
				362	A*0101	A*0201	B*0702	B*4001	C*0304	C*0702	#					
2947-2964	GKAKICGLYLFNWAVRTK	NS5B		275	A*1101		B*0702	B*4402	C*0501	C*0702	#	*	RGG**R**A**A**ICG**K**YLFNWAVR		Class I	C Neumann-Haefelin
													G**R**A**A**ICG**K**Y		B27	C Neumann-Haefelin
													G**R**A**A**ICG**K**YLFNWAV		Class I	P T F Kennedy
													KYLFNWAV**K**		A2	Z Guo
148-165	ARALAHGVRALEDGINFA	core		962							#		GVR**V**LEDG**V**		A2	H F Löhr
													R**V**LEDGVN**Y**		Class I	D D Anthony
													VLEDGVN**Y**ATGNLPG		Class I	K Sugimoto
1520-1537	RPSGMFDSVVLCECYDAGCSWYDL	NS3	B35/C04	105	A*1101	A*1101	B*3501	B*5101	C*0401	C*1402			MFDS**S**VLCECYDAGC		Class I	D Ciuffreda
				111	A*0101	A*0201	B*35	B*14	C*0401	C*0802						
				140	A*2402	A*3002	B*0702	B*3501			#	*				
				154	A*1101	A*1101	B*1517	B*3501	C*0401	C*0701						
				219	A*1101	A*2402	B*18	B*3501	C*0401	C*1203						
				235	A*0301	A*3002	B*0702	B*3501	C*0401	C*0702	#	*				
				268												
				291	A*0101	A*2901	B*3501	B*44030	C*0401	C*1601						
				295	A*0301	A*2402	B*3501	B*55010	C*0301	C*0401						
				362	A*2402		B*3501	B*4403	C*0401	C*0409	#	*				
				750												
				787	A*0201	A*1101	B*3501	B*4001	C*0401	C*0301						
				906	A*0301	A*2601	B*3501	B*2701	C*0401	C*0202						
				955	A*0201	A*1101	B*4402	B*3501	C*0401	C*0501						
2126-2141	AEFFTEVDGVRLHRYA	NS5A		6-40	A*0101		B*0801	B*5101	C*0102	C*0701	#		FFTE**L**DGVRLHRFAP		Class I	D Ciuffreda
				822	A*0101		B*5201	B*5701	C*0602	C*1202						
				871	A*3201	A*0101	B*0801	B*4402	C*0501	C*0701						
				975	A*3201	A*0301	B*1401	B*0702	C*0802	C*0702						
1792-1808	PAVASLMAFTASVTSPL	NS4B		6-23	A*2402	A*0201	B*3502	B*4403	C*0401	C*1601			SLMAFTA**A**V		A2	B Rehermann
				416	A*0101	A*0201	B*0801	B*5701	C*0602	C*0701	#	*	SLMAFTA**A**V		A2	K M Chang
				437	A*0201	A*2601	B*3801	B*2702	C*1203	C*0102	#	*	SLMAFTA**A**V		A2	N H Gruener
696-712	LIHLHQNIVDVQYLYGV	E2		568	A*02	A*3201	B*2707	B*1501	C*0602	C*0304	#	*	ALSTGLIHLHQNIVD		Class I	D Ciuffreda
702-719	NIVDVQYLYGVGSGMVGW			6-56	A*0201	A*2601	B*2702	B*3801	C*1203				LHQNIVDVQYLYGVG		Class I	D Ciuffreda
				861							#	*				
1379-1387	IPFYGKAIPI	NS3	B51	6-40	A*0101		B*0801	B*5101	C*0102	C*0701	#		IPFYGKAI	&	B51	S Giugliano
				50163			B*4403	B*5101	C*1402	C*1601	#		IPFYGKAIP**L**		B51	D Yerly
1436-1447	GDVVVCATDALMTGF	NS3		822	A*0101		B*5201	B*5701	C*0602	C*1202			ATDALMTGY		A1	G M Lauer, 2002
1442-1447	ATDALMTGY		A01	822	A*0101		B*5201	B*5701	C*0602	C*1202	#		**V**ATDALMTGYTGDFD		Class I	A M Werthheimer
				7-70	A*0101	A*2402	B*0801	B*3906	C*0701	C*0702	#		ATDALMTGF	&	A1	S Giugliano
				06P	A*0101	A*0301	B*0702	B*0702	C*0702	C*0702	#		ATDALMTGF		A1	T Kuntzen
				416	A*0101	A*0201	B*0801	B*5701	C*0602	C*0701	#		ATDALMTGY		A1	A L Cox
													ATDALMTGY		A1	E Barnes
													ATDALMTGY		A1	G M Lauer, 2004

HCV gt3 specific CD8+ T cell targets confirmed in this study, and corresponding CD8+ T cell targets described in the literature. T cell targets were classified by the likelihood of overlap between described and detected targets: no matching CD8+ T cell targets described (red); Targets described, but major gt1/gt3 sequence differences (light red); Targets described, minor gt1/gt3 sequence differences (light green); Matching T cell targets described, no sequence differences (green). The viral region, position within the viral genome and gt3 sequence is given. All patients targeting each individual peptide are specified (colour coding: orange-SR, blue-acute, yellow chronic), including patient's HLA class-I types where obtained. Sequences of CD8+ T cell targets published in the literature are given, including the restricting HLA type and first author of the publication. Differences to HCV gt3 sequences are marked in bold, overlap between T cell targets is underlined. # CD8 restriction for this patients experimentally defined, * T cell targets from the Oxford cohort previously described in (Humphreys et al. 2012). & T cell targets previously described for gt1 and 3.

(M Battegay et al. 1995; Rehermann et al. 1996; K. M. Chang et al. 1997; Löhr et al. 1999; Grüner et al. 2000; Wertheimer et al. 2003; Sugimoto et al. 2003; Lauer et al. 2004; A. L. Cox, Mosbruger, Lauer, et al. 2005; Kennedy et al. 2006; Kuntzen et al. 2007; Yerly, Heckerman, T. M. Allen, et al. 2008; Ciuffreda et al. 2008; Giugliano et al. 2009; Neumann-Haefelin et al. 2008; Barnes et al. 2012; Z. Guo et al. 2012)

For 6 gt3 specific T cell targets, no immune response to the equivalent HCV gt1 sequence was previously described in the literature (colour coded 'red' in Table 3-10). It is therefore likely, that FQMIILSIGR (NS2), EGAVQWMNRLIAFASR (NS4B), GVMSTRCPCGASIAGHVK (NS5A), VLDDHYKTAL, RVKARMLTI and KLTPLPAAGQL (NS5B) are newly described HCV gt3 specific CD8+ restricted T cell targets. HCV gt3a sequences were compared to reference gt1a and gt1b sequences (Table 3-11). At all sites but one (EGAVQWMNRLIAFASR), significant sequence differences were observed at multiple amino acid positions. This suggests that sequence differences between HCV gt1 and gt3 are the cause for differing MHC class-I presentation and CD8+ T cell recognition between the two genotypes.

Table 3-11: Sequence comparison for HCV gt3 and gt1 at newly described HCV gt3 CD8+ restricted T cell targets.

	Gt3 peptide sequence	Gt1a H77 reference	Gt1b J4B reference
NS2	FQMIILSIGR	**V**QM**A**I**IKL**G**A**	**V**QM**VFMKL**G**A**
NS4B	EGAVQWMNRLIAFASR	EGAVQWMNRLIAFASR	EGAVQWMNRLIAFASR
NS5A	GVMSTRCPCGASIAGHVK	G**I**M**H**TRC**H**CGA**E**I**T**GHVK	G**I**M**Q**T**T**CPCGA**Q**IAGHVK
NS5B	VLDDHYKTAL	VLD**S**HY**QDV**L	VLDDHY**RDV**L
NS5B	RVKARMLTI	**E**VKA**AASKV**	**T**VKA**KL**L**S**I
NS5B	KLTPLPAAGQL	KLTPI**A**AAG**R**L	KLTPIPAASQL

Sequence comparison for detected HCV gt3a T cell targets in the Oxford cohort with no T cell responses previously described in the literature in HCV gt1 infection. Sequence variations between HCV gt3 peptides and gt1a (H77) and gt1b (J4B) reference sequences are shown.

For three T cell targets described in the gt3 Oxford cohort, identical targeted regions were described in gt1 infection (ATDALMTGY and IPFYGKAIPI (NS3, A*0101 and B*51 restricted), and region LIHLHQNIVDVQYLYGVGSGMVGW (E2), marked in green in Table 3-10), suggesting an overlap of T cell responses between HCV gt1 and gt3 at these sites. Only a single mutation between HCV gt3a an gt1 was observed at epitope SLMAFTA**S**V (gt1: SLMAFTA**A**V (NS4B)), which was previously described as A*02 restricted in gt1 infection (Rehermann et al. 1996; K. M. Chang et al. 1997; Grüner et al. 2004). Further supporting the hypothesis that this epitope was targeted in both genotypes, it was observed that all patients targeting this epitope within the HCV gt3 cohort carried the A*02 allele.

For the remaining 6 T cell targets, potential overlap between gt1 and gt3 T cell immunity is less well defined (marked in light green or light red in Table 3-10): Single sequence polymorphisms in targeted regions between gt1 and gt3 suggest possible HCV gt1 and 3 cross-reactivity at two further T cell targets (NS3: RPSGMFDSVVLCECYDAGCSWYDL/ MFDS**S**VLCECYDAGC, NS5A: AEFFTEVDGVRLHRYA/ FFTE**L**DGVRLHRFAP); however, HLA restriction was not defined in the publication describing these targets in gt1 infection (Ciuffreda et al. 2008). Another potential overlap between HCV gt1 and gt3 was observed for HCV gt3 core target ARALAHGVRALEDGINFA, but two amino acid substitutions were observed compared to a A*02 restricted gt1 T cell targets in the same region G**V**RVLEDG**V** (Löhr et al. 1999; Anthony et al. 2002; Sugimoto et al. 2003); further complicating the comparison patient 962 from our cohort was not HLA typed.
For T cell targets LLYPSLIFDI (NS2), RVLLDILAGY (NS4B) and GKAKICGLYLFNWAVRTK (NS5B), substantial sequence differences or a sequence overlap of less than 5 amino acids were observed in described gt1 epitopes, suggesting that these T cell targets are not shared between genotypes.
Overall, out of 16 CD8+ restricted T cell targets defined in the Oxford gt3 cohort, matching gt1 responses were described in 3 cases (19%). Similar T cell targets with single amino acid substitutions were observed in 3 cases (19%), or 2 amino acid substitutions in one case (6%). However, for the majority of T cell targets no matching gt1 epitopes were described (9/16, 56%), with major sequence differences or small overlap in 3 cases (19%), or no described gt1 T cell targets in 6 cases (38%).

Next, 7 T cell targets from the Oxford cohort defined as CD4+ restricted were compared to gt1 targets previously described in the literature (Table 3-12). For 6/7 (86%) T cell targets described in the Oxford cohort, CD4+ epitopes in the corresponding region were described in the literature for HCV gt1 infection.
For one targeted region, CPQRLSSCKPITFFRQGWGSLTDANITGPSD (E2), detected in two patients in the Oxford cohort, no matching T cell response was described in the literature (epitope marked red in Table 3-12). This might be due to major sequence differences at 13 positions between the gt3 consensus and gt1a peptide (H77, CP**E**RL**A**SC**R**R**LT**D**F**A**QGWGPI**SY**AN**GS**G**LDE**) and 16 positions between gt3 and 1b (J4b, CP**E**R**MA**SC**R**PI**DW**F**A**QGWGPIT**YTKPNSSDQ**).

Since peptide KRALYDVIQKLSIETM (2603-2618, NS5B) described in the Oxford cohort and the published T cell targets overlap by 7 amino acids only, of which 4 amino acids differed in sequence, it is unlikely that a matching epitope between genotypes is contained in these peptides (T cell target marked light red in Table 3-12). For the 5 remaining CD4+ T cell targets, sequences were similar (marked in light green), with only 1-3 amino acids differing within the overlapping sequences for gt3 and gt1.

Table 3-12: CD4+ T cell targets previously described in the literature in the region of CD4+ T cell targets defined in HCV gt3 infection in this study.

GT3 CD4+ epitopes (Oxford Cohort)								GT1 epitopes described in corresponding regions		
Position	3a peptide sequence	Viral protein	Patient ID					Peptide (Literature)	HLA	First author
453-476	CPQRLSSCKPITFFRQGWGSLTDANITGPSD	E2	822	DRB1*0701	DRB1*1502	#		no CD4 epitopes described		
			06P							
2603-2618	KRALYDVIQKLSIETM	NS5B	884	DRB1*0404	DRB1*0701	#	*	KL**PLAV**MGSSYGFQYSPGQR	Class-II	J Schulze zur Wiesch, 2007
								KL**PLAV**MGSSYGFQYSPGQR	Class-II	C L Day
66-83	PKARRSEGRSWAQPGYPW	core	105	DRB1*0101	DRB1*0701			RRQPIPKARR**P**EGR**T**WAQPG	Class-II	J Schulze zur Wiesch, 2012
			128	DRB1*0301	DRB1*0401			RQPIPKVRR**P**EGR**T**	HLA-DR	J J Lasarte
			275	DRB1*0401	DRB1*0407			K**V**RR**P**EGR**T**WAQPG	HLA-DR	J J Lasarte
			450			#		PEGR**T**WAQPGYPWPLYGNEG	Class-II	J Schulze zur Wiesch, 2007
73-90	GRSWAQPGYPWPLYGNEG		416	DRB1*0101	DBR1*0701	#	*	PEGR**T**WAQPGYPWPLYGNEGCGW	Class-II	H F Löhr
								PEGR**T**WAQPGYPWP	HLA-DR	J J Lasarte
								PEGR**T**WAQPGYPWPL	Class-II	P T F Kennedy
143-158	PVGGVARALAHGVRAL	core	6-56					ADLMGYIPLVGAP**L**GG**A**ARA	Class-II	J Schulze zur Wiesch, 2007
			110	DRB1*0801	DRB1*1101	#	*	ADLMGYIPLVGAP**L**GG**A**AR	HLA-DR	F A Castelli
			128	DRB1*0301	DRB1*0401	#		ADLMGYIPLVGAP**L**GG**A**ARA	Class-II	J Schulze zur Wiesch, 2012
			129	DRB1*1504	DRB1*0701			LMGYIPLVGAP**L**GG**A**	Class-II	D Ciuffreda
			216	DRB1*0301	DRB1*1602	#		LVGAP**L**GG**A**ARAL	Class-II	H F Löhr
			219	DRB1*0403	DRB1*0701			LVGAP**L**GG**A**ARALAH		K Sugimoto
			226	DRB1*0404	DRB1*1104	#		GAP**L**GG**A**ARALAHGVR**V**LED	Class-II	J Schulze zur Wiesch, 2007
			227	DRB1*0701	DRB1*0801	#		GAP**L**GG**A**ARALAHGVR**V**LED	Class-II	A J MacDonald
			235	DRB1*0101	DRB1*1501			GAP**L**GG**A**ARALAHGVR**V**LED	Class-II	C L Day
			331			#	*	GAP**L**GG**A**ARALAHGVR**V**LED	Class-II	J Schulze zur Wiesch, 2012
			422	DRB1*0101	DRB1*1501			GG**A**ARALAHGVR**V**LE	Class-II	D Ciuffreda
			437	DRB1*0101		#	*	ALAHGVR**V**L	Class-II	H F Löhr
								LAHGVR**V**LEDGVNYATGNLP	Class-II	J Schulze zur Wiesch, 2012
1423-1440	AYYRGLDVSVIPTAGDVV	NS3	226	DRB1*0404	DRB1*1104			GINAVAYYRGLDVSVIPT**S**G	Class-II	J Schulze zur Wiesch, 2007
			884	DRB1*0404	DRB1*0701	#	*	GINAVAYYRGLDVSV	Class-II	A M Wertheimer
								GINAVAYYRGLDVSVIPT**S**G	Class-II	C L Day
								VAYYRGLDVSVIPT**S**	Class-II	A M Wertheimer
								LDVSVIPT**S**GDVVVVATDAL	Class-II	J Schulze zur Wiesch, 2007
								IPT**S**GDVVVVSTDALMTG	Class-II	N M Tabatabai
2548-2565	NQIRSVWEDLLEDTTTPI	NS5B	113	DRB1*0701	DRB1*1454	#	*	HARKAV**TH**I**N**SVWKDLLED**N**	Class-II	J Schulze zur Wiesch, 2012
								SVWKDLLED**NV**TPIDTTIMA	Class-II	C L Day
2893-2908	IIERLHGLSAFTLHSY	NS5B	113	DRB1*0701	DRB1*1454	#	*	PII**Q**RLHGLSAF**S**LHSYSPG	Class-II	J Schulze zur Wiesch, 2007
								PII**Q**RLHGLSAF**S**LHSYSPG	Class-II	J Schulze zur Wiesch, 2012

HCV gt3 specific CD4+ T cell targets confirmed in this study, and responses described in the literature as CD4+ restricted or without defined CD4/CD8 restriction. For each targeted region, the viral region, position within the viral genome and gt3 sequence is given. All patients targeting the viral region are specified (colour coding: orange-SR, blue-acute, yellow chronic), including patient's HLA class-II types where obtained. #CD4 restriction for this patients experimentally defined, *previously described in (Humphreys et al. 2012). Sequences of CD4+ T cell targets published in the literature are given, including the restricting HLA type and the first author of the publication. Differences to HCV gt3 sequences are marked in bold. (J J Lasarte et al. 1998; Lamonaca et al. 1999; Löhr et al. 1999; Tabatabai et al. 1999; C. L. Day et al. 2002; A. J. MacDonald et al. 2002; Sugimoto et al. 2003; Wertheimer et al. 2003; Kennedy et al. 2006; Castelli et al. 2007; Schulze zur Wiesch et al. 2007; Ciuffreda et al. 2008; Schulze Zur Wiesch et al. 2012) .

Taken together, these results suggest that CD4+ T cell target similar regions in HCV gt3 and gt1 infection, even though peptides may vary at 1-3 amino acids within

epitopes, which may be due to the more promiscuous nature of MHC class-II presentation (Siebenkotten 1998).

The picture is less clear for T cell targets defined in the Oxford cohort where the CD4/CD8 restriction was not defined. For these 14 individual peptides, T cell targets previously described in the literature were analysed, including targets with either defined CD4+/CD8+ restriction or without defined restriction (Table 3-13).

T cell targets without defined restriction detected in the Oxford HCV gt3 cohort were compared to published HCV gt1 CD8+ and CD4+ T cell targets separately, and grouped accordingly (Table 3-13). For depiction in Table 3-13, data was compressed and is shown without literature references.

I first analysed CD8+ restricted T cell targets previously described in the literature and compared them to T cell targets without defined restriction from the Oxford cohort (Table 3-13, middle column). If HLA restrictions of published T cell targets matched the patients HLA type (HLA types marked in blue), the HCV gt3 peptide was marked as “likely overlapping” (light green). If the patient did not have the restricting HLA allele (HLA types marked in grey), the HCV gt3 peptide was marked as “unlikely overlapping” (red). A similar classification for CD4+ restricted T cell targets was not attempted; Due to their promiscuous nature, CD4+ responses were classified as “likely overlapping” (light green) if a T cell target was described in the literature. Following from this comparison, HCV gt3 T cell targets were grouped into three categories:

(1) CD8+ T cell targets previously described, with matching HLA types; no CD4+ T cell targets described (top section).

(2) CD8+ and CD4+ T cell targets previously defined (middle section).

(3) CD4+ T cell targets previously described, no CD8+ T cell targets defined (bottom section).

Two T cell targets are likely to be CD8+ restricted. For seven individual peptides, both CD4 and CD8 T cell targets were described, however, the restricting class-I HLA type matched the patients HLA type, making a CD8+ restriction likely (marked in blue in Table 3-13). However, this is just an approximation: for a definite comparison, CD4/CD8 and ideally HLA restriction will have to be defined.

In summary, T cell responses to T cell targets without defined CD4/CD8 restriction targeted in our HCV gt3 cohort have been previously described in the literature.

Table 3-13: HCV gt1 T cell targets previously described in the literature in the region of HCV gt3 T cell targets with non-defined CD4/CD8 restriction.

GT3 responses restriction not defined (Oxford cohort)					CD8+ responses Literature			CD4+ responses Literature		
610-625	LTPRCMVDYPYRLWHY	E2	C SR	 *	+	CLVDYPYRL DYPYRLWHY	A2 Cw7		no CD4 epitopes described	
2119-2136	CPCQVPAAEFFTEVDGVR	NS5A	A		+	FFTELDGVRLHRFAP	CD8+		no CD4 epitopes described	
27-51	GGQIVGGVYVLPRRGPRLGVRATRK	core	C C SR	*	(+)	GQIVGGVYLL IVGGVYLLPR GVYLLPRRGPRLGVR NVKFPGGGQIVGGVYVLPRR	B60 A11 CD8+ ND	+	DVKFPGGGQIVGGVYLLPRR GGQIVGGVYLLPRRGPRLGVRATRK NVKFPGGGQIVGGVYVLPRR VGGVYLLPRRGPRLGVRATR GVYLLPRRGPRLGVR LLPRRGPRL RRGPRLGVRATRKTSERSQPRGRRQ GPRLGVRATRKTSERSQPRG	HLA-DRB1*15:01 HLA-DR ND CD4+ CD4+ CD4+ CD4+ CD4+
130-147	FADLMGYIPLVGAPVGGV	core	C		+	GFADLMGYIPL (A)DLMGYIPLV GFADLMGYI	A2 A2 A24	+	TLTCGFADLMGYIPL ADLMGYIPLVGAPLGGAAR DLMGYIPLV LMGYIPLVGAPLGGA LVGAPLGGAARAL GAPLGGAARALAHGVRVLED	CD4+ HLA-DR CD4+ CD4+ CD4+ CD4+
1198-1213	KALQFIPVETLSTQAR	NS3	SR	*	+	KAVDFIPVENLETTMRSP VDFVPVESM	CD8+ B40	+	RAAVCTRGVAKAVDFIPVEN AKAVDFIPVENLETTMRSPVFTD	CD4+ HLA-DR
1246-1261	KVPAAYVAQGYNVLVL	NS3	SR	*	+	GKSTKVPAAYAAQGYKVL	CD8+	+	GSGKSTKVPAAYAA GKSTKVPAAYAAQGYKVLV PAAYAAQGYKVLVLNPSVAA GYKVLVLNPSVAATLGFGAY	CD4+ CD4+ CD4+ CD4+
1264-1281	SVAATLGFGSFMSRAYGI	NS3	A SR		+	ATLGFGAYMSKA GAYMSKAHGV/GSFMSRAYGI TLGFGAYMSK/TLGFGSFMSR	A3 CD8+ A11	+	GYKVLVLNPSVAATLGFGAY VLVLNPSVAATLGFGAYM VAATLGFGAYMSKAHGVD TLGFGAYMSKAHGIDPNIRT	CD4+ CD4+ CD4+ CD4+
1282-1298	DPNIRTGNRTVTTGAKL	NS3	SR	*	+	GVDPNIRTGV/GIDPNIRTGN	A2	+	GVRTITTGSPITYSTYGKFL	CD4+
1547-1569	RAYLSTPGLPVCQDHLDFWESVF	NS3	C		+	EFWESVFTGLTHIDAHFL/ DFWESVFTGLTHIDAHFL	CD8+	+	TPAETTVRLRAYMNTPGLPV TTVRLRAYMNTPGLP LRAYMNTPGLPVCQD	CD4+ CD4+ CD4+
635-650	KVRMFVGGFEHRFTAA	E2	A C		(+)	RMYVGGVEHR	A3	+	TINYTIFKIRMYVGGVEHRL KVRMYVGGVEHRLEA MYVGGVEHRLEAACNWTRGE	CD4+ CD4+ CD4+
1805-1822	TSPLTTNQTMFFNILGGW	NS4B	SR	*	(+)	LTTSQTLLF LFFNILGGWV	B57 A2	+	VTSPLTTSQTLLFNILGGWV LTTSQTLLFNILGGWVAAQL LLFNILGGWVAAQLAAPGAA	CD4+ CD4+ CD4+
1825-1842	THLAGPQSSSAFVVSGLA	NS4B	C	*		no CD8+ epitopes described		+	AAQLAAPGAATAFVGAGLAG AAPGAATAFVGAGLAGAAIG	CD4+ CD4+
2844-2861	IMYAPTIWVRMVMMTHFF	NS5B	A		(+)	ARMILMTHF	B27	+	NIIMFAPTLWARMILMTHFF ARMILMTHFFSVLIARDQLE	CD4+ HLA-DR11
2145-2162	KPLLRDEITFMVGLNSYA	NS4B	A		(+)	LLREEVSFRV	A2		no CD4 epitopes described	

T cell targets from the Oxford HCV gt3 cohort with non-defined restriction are shown and were grouped into 4 categories: (1) T cell targets with no CD4+ , but matching CD8+ T cell responses described in the literature. (2) T cell targets with matching CD4+ and CD8+ responses described in the literature. (3) T cell targets with no CD8+, but matching CD4+ T cell responses described in the literature. (4) T cell targets with no matching CD4+ and CD8+ responses described. Colour coding; Red: no/unlikely overlap with described T cell targets/no T cell targets described. Green: possible overlap with described T cell targets. Blue: Literature HLA type matches patients HLA type. Grey: Literature HLA type does not match patients HLA type. Orange: Sequences of T cell targets that were described in HCV gt1/gt3 infection.

I conclude that for most genomic regions with T cell responses detected in the Oxford HCV gt3 cohort, T cell targets were previously described in the literature. However, only for the minority of T cell targets sequence identity was observed between HCV gt1 and 3. Observed overlap between HCV gt3 and gt1 was higher for CD4+ restricted than for CD8+ restricted T cell targets (71% vs. 38%).

3.9.2 Are CD8+ T cell epitopes dominant in HCV gt1 infection targeted in HCV gt3 infection?

Next, I aimed to determine whether CD8+ T cell epitopes commonly targeted in HCV gt1 infections were detected in the Oxford HCV gt3 cohort.

CD4+ and CD8+ T cell epitopes commonly targeted in HCV gt1 infection have been described in detail (Lauer et al. 2004; Kuntzen et al. 2007; Lauer et al. 2005; Schulze zur Wiesch et al. 2005; Schulze zur Wiesch et al. 2007; Lechner et al. 2000; Fitzmaurice et al. 2011). As we have shown in a comprehensive comparison of HCV gt3 T cell targets identified in the Oxford cohort with previously published T cell targets in HCV gt1 infection (3.9.1, page 121), T cell specificity varies considerably between these two genotypes. However, it is not clear whether epitopes commonly described in the literature for HCV gt1 are detected in HCV gt3 infection.

To perform a comparative analysis, we first obtained published information on CD8+ restricted epitopes targeted in HCV gt1 infection from the IEDB (detailed description in 2.6.1, page 81). Subsequently, all epitopes described in more than 5 publications after the exclusion of duplicates were defined as dominant. Results from this comparative analysis are shown in Table 3-14.

Table 3-14: Were dominant HCV gt1 CD8+ restricted epitopes detected in the Oxford cohort?

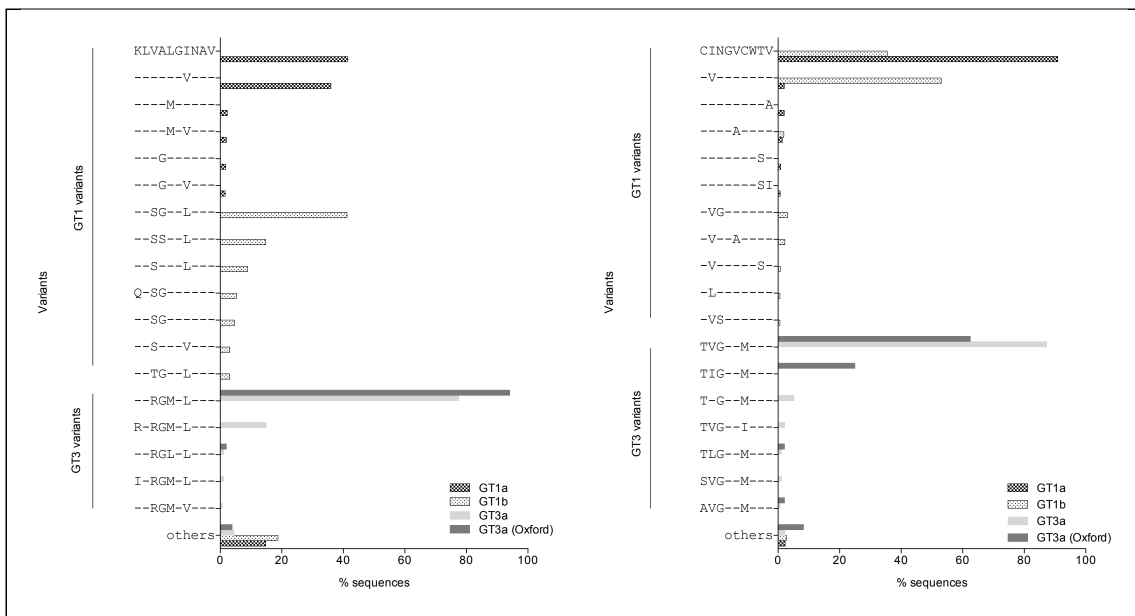
Epitope sequence	Viral region	start AA	end AA	HLA type	number of publications	targeted by T cells in Oxford gt3 cohort	Restriction of HCV gt3 epitopes
YLLPRRGPRL	core	35	44	A2	10	+	CD4
GPRLGFRAT	core	41	49	B7	5	+	CD4
NEGLGWTGW	core	87	95	B44	5		
ADLMGYIPLV	core	131	140	A2	11	+	CD4
LLALLSCLTV	core	178	187	A2	5		
SMVGNWAKV	E1	363	371	A2	5		
SLLAPGAKQNV	E2	401	411	A2	5		
NTRPPLGNW	E2	541	549	B57	5		
RDWAHNGL	NS2	957	964	B37	7		
CINGVCWTV	NS3	1073	1081	A2	31		
LLCPAGHAV	NS3	1169	1177	A2	5		
HPNIEEVAL	NS3	1359	1367	B35	7		
HSKKKCDEL	NS3	1395	1403	B8	8		
KLVALGINAV	NS3	1406	1415	A2	32		
ATDALMTGY	NS3	1436	1444	A1	7	+	
LLFNILGGWV	NS4b	1807	1816	A2	7		
VLSDFKTWL	NS4a	1987	1995	A2	7		
ALYDVVTKL	NS5b	2594	2602	A2	11		

Comparison of CD8+ restricted epitopes dominant in HCV gt1 infection (defined as described in 5 or more publications on the IEDB). Epitopes that were targeted in the Oxford HCV gt3 cohort are marked with +. Epitopes described as CD4+ restricted T cell targets in the Oxford gt3 cohort falling into regions of CD8+ restricted epitopes described as HCV gt1 epitopes in the literature are marked in grey.

18 HCV gt1 specific CD8+ epitopes from the immune epitope database were described in five or more publications. Of these published epitopes, most were located in the HCV core (n=5, 28%) and NS3 (n=6, 33%) region. The majority of dominant epitopes were restricted by HLA type A*02 (11/18, 61%). When comparing dominant CD8+ epitopes in HCV gt1 infection to T cell targets detected in the Oxford HCV gt3 cohort, HCV gt3 T cell responses were only found in the region of 4 HCV gt1 epitopes, of which three were CD4+ restricted. Therefore, the majority of CD8+ epitopes dominant in HCV gt1 infection (17/18, 94%) was not targeted in HCV gt3 infection, with only one dominant HCV gt1 epitope (ATDALMTGY) detected in the Oxford HCV gt3 cohort (1/18, 6%).

Possible reasons for the lack in cross-reactivity are sequence differences between HCV gt1 and gt3. To assess sequence variation at particularly common HCV gt1 CD8+ epitopes that were described in over thirty publications (`KLVALGINAV` and `CINGVCWTV`), sequences were further evaluated for HCV genotypes 1 and 3. Sequence material was obtained from the Los Alamos database and additional HCV gt3 sequences generated in-house included. Figure 3-20 shows circulating viral variants for HCV gt1 and gt3.

Figure 3-20: GT3 sequence variants of commonly targeted epitopes in HCV genotype 1 infection.



Comparison of viral sequences for dominant HCV gt1 epitopes for gt1 and gt3. Sequences for these epitopes were obtained from the Los Alamos database, with additional gt3 sequences from our laboratory.

HCV gt1 and gt3 sequences at dominant HCV gt1 epitopes vary at 4 sites: KLVALGINAV/KL**RGM**G**L**NAV and CINGVCWTV/**TVG**GV**M**WTV. Since already a single mutation can hamper peptide presentation by MHC complexes or recognition by T cells (Söderholm et al. 2006; Fytili et al. 2008), these mutations may cause the lack of cross-reactivity between the two genotypes. Indeed, for epitope CINGVCWTV, a loss in binding activity of the HCV gt3 version **TVG**GV**M**WTV to HLA A*02 was previously described, as well as an inability of HCV gt1 primed T cell to recognise the HCV gt3 epitope *in vitro* (Fytili et al. 2008).

3.9.3 Limited HCV gt1 and gt3 cross-reactivity on single peptide level

Next, we aimed to assess cross-reactivity between HCV gt3 and gt1 experimentally. Therefore, HCV gt3 patients with confirmed responses to the gt3 peptide set with PBMC available were tested with HCV gt1 peptide equivalents (Table 3-15). Limited cross-reactivity between gt3 and gt1 peptides was observed on single peptide level, with decreased or lost responses towards HCV gt1 peptides compared to responses observed against HCV gt3 peptides in all but one tested response.

Table 3-15: HCV gt3a and gt1 peptide cross-reactivity.

Protein	3a amino acid	3a peptide	gt3a SFU/ million PBMC	1a and/or 1b Equivalent Sequence		gt1 SFU/ million PBMC	3a/1 cross-reactivity
Core	73-90	GRSWAQPGYPWPLYGNEG	33	1a	GR**T**WAQPGYPWPLYGNEG	28	↓
				1b	GR**A**WAQPGYPWPLYGNEG**L**	23	↓
	143-158	PVGGVARALAHGVRAL	50	1a=1b	**GA**P**L**G**A**VARALAHGVR**V**L	23	↓
NS3h	1423-1440	AYYRGLDVSVIPTAGDVV	267	1b	AYYRGLDVSVIPT**I**GDVV	0	↓
	1513-1523	RPSGMFDSVVL	30	1a=1b	RPSGMFDS**S**VL	8	↓
NS4b	1791-1806	PAVASLMAFTASVTSPL	65	1a	PA**I**ASLMAFTA**A**VTSPL	0	-
				1b	PA**I**ASLMAFTAS**I**TSPL	0	-
NS5a	2029-2046	GVMSTRCPCGASIAGHVK	23	1a	**I**V**H**STR**H**PCG**E**S**T**AGHVK	30	↑
				1b	G**I**M**Q**T**T**CPCGA**Q**IAGHVK	13	↓
NS5b	2603-2618	KRALYDVIQKLSIETM	100	1b	K**M**ALYDV**VST**L**PQVV**M	8	↓
	2965-2975	AVRTKTKLTPLPAAGQL	33	1a	VRTK**L**KLTP**IA**AAG**R**LDL	0	-
				1b	AVRTK**L**KLTP**I**PAA**S**QL	5	↓

T cell responses were assessed in HCV gt3 patients using the positively identified HCV gt3a peptides and equivalent 1a and 1b peptides (IFNγ ELISpot assays). Amino acids differing between genotypes are shown in bold. ↑ cross-reactivity maintained, - loss of cross reactivity, ↓ Decrease in reactivity. This table was previously published in (Humphreys et al. 2012).

In conclusion, overlap in T cell specificity between HCV gt1 and 3 was limited when T cell targets detected in the Oxford HCV gt3 cohort were compared to HCV gt1 responses previously published in the literature. T cell targets overlapped at 6 out of

16 described CD8+ T cell targets (38%) between HCV gt3 and 1, however, identical sequences between genotypes were only observed at three CD8+ restricted targets. The majority of CD8+ restricted T cell targets described in HCV gt3 infection in the Oxford cohort were not previously described in HCV gt1 (56%). A higher amount of overlap between genotypes was observed for promiscuous CD4+ restricted T cell targets (71%).

When comparing epitopes dominant in HCV gt1 infection (defined as found in >5 publications in the literature) to HCV gt3 specific T cell targets defined in the Oxford HCV gt3 cohort, minimal overlap of just one epitope was observed.

Finally, T cell cross-reactivity between HCV genotype 1 and 3 was assessed experimentally. When comparing cross-reactivity of T cell responses primed in natural HCV gt3 infection towards HCV gt1a and gt1b peptides, a reduction of T cell responses was observed.

3.10 Associations of magnitude of HCV gt3 specific T cell responses with clinical predictors of spontaneous resolution of infection and treatment response

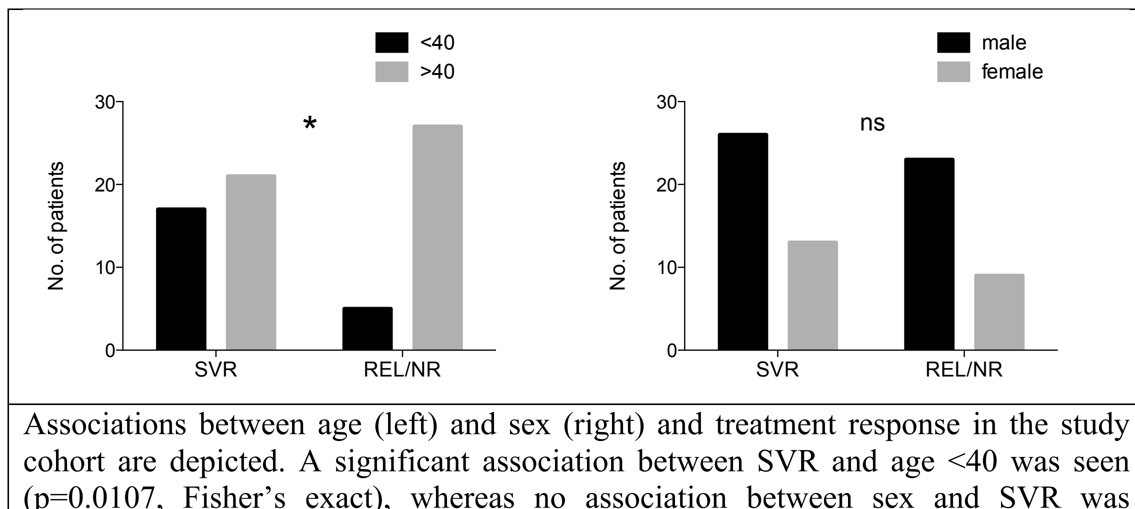
The magnitude of T cell responses influences the outcome of acute infection (Lauer et al. 2004). In the following section, I aimed to assess whether the magnitude of T cell responses in HCV gt3 infection prior to combination treatment with peg-IFNα/Ribavirin was associated with treatment outcome and known independent predictors of treatment outcome, like age, sex, pre-treatment viral load and IL28B genotype (Kau et al. 2008; Neukam et al. 2012).

3.10.1 Associations of treatment outcome with known predictors of treatment response

First, known predictors of treatment response were evaluated within our cohort. Of 108 chronic patients, 78 (72.2%) were treated with IFNα/Ribavirin therapy (Table 3-2, page 88). Treatment response to the first treatment administered was evaluated. 42 patients achieved SVR (53.8%), 25 patients relapsed, 6 patients did not respond and 3 patients discontinued treatment due to interferon side effects. Patients not responding to treatment and relapsed patients were grouped for the following analyses.

In the cohort analysed in this study (Table 3-2, page 88), a significant association between younger age and response to combination treatment (SVR, pegylated IFNα/Ribivirin) was observed, whereas no association between sex and response to treatment was seen (Figure 3-21).

Figure 3-21: Association between age and sex and treatment outcome with peg-IFNα and Ribavirin.

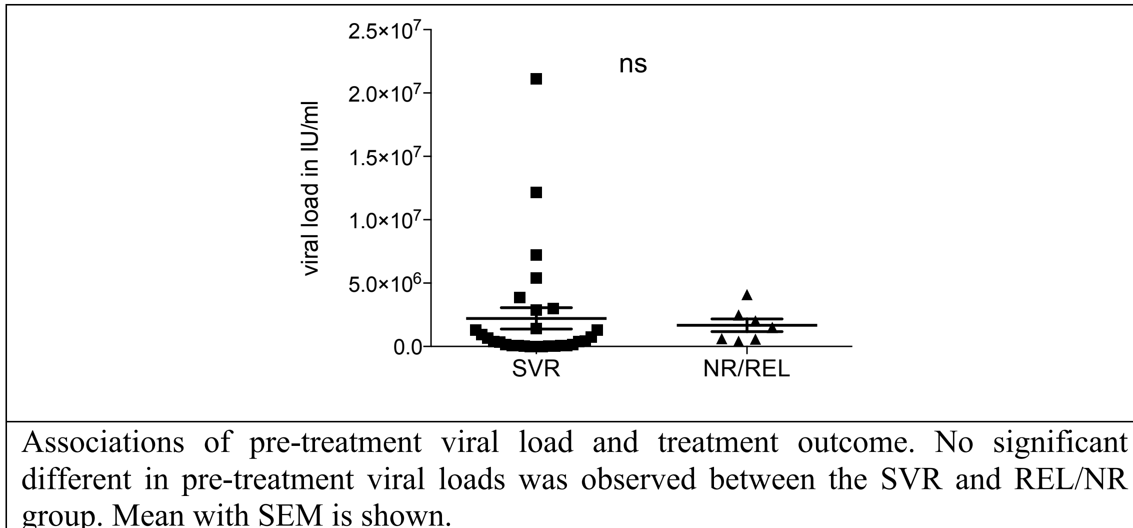


Associations between age (left) and sex (right) and treatment response in the study cohort are depicted. A significant association between SVR and age <40 was seen (p=0.0107, Fisher's exact), whereas no association between sex and SVR was

observed.

Pre-treatment viral loads were measured in 38 chronically HCV gt3a infected patients, who subsequently received combination therapy with peg-IFNα and Ribavirin. 29 of these patients resolved infection (SVR), 9 patients did not respond to treatment or had a viral relapse (NR/REL). No association between treatment response and pre-treatment viral load was observed (Figure 3-22).

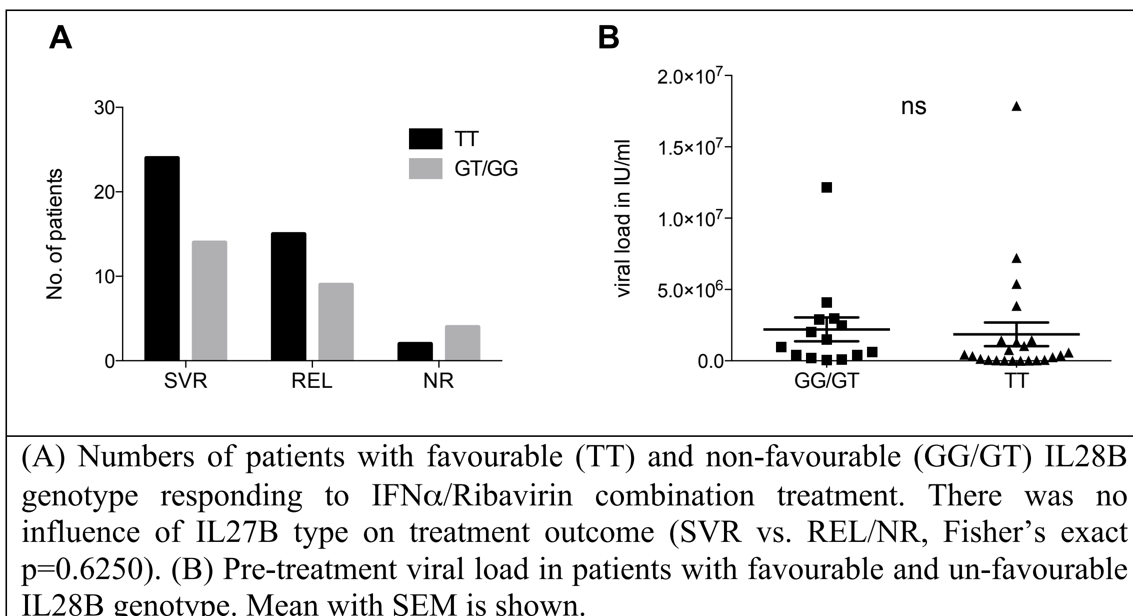
Figure 3-22: Pre-treatment viral load and treatment outcome.



Associations of pre-treatment viral load and treatment outcome. No significant different in pre-treatment viral loads was observed between the SVR and REL/NR group. Mean with SEM is shown.

Next, distributions of IL28B genotypes were analysed regarding HCV treatment outcome (Figure 3-23, A).

Figure 3-23: IL28B and influence on treatment response and baseline viral load.



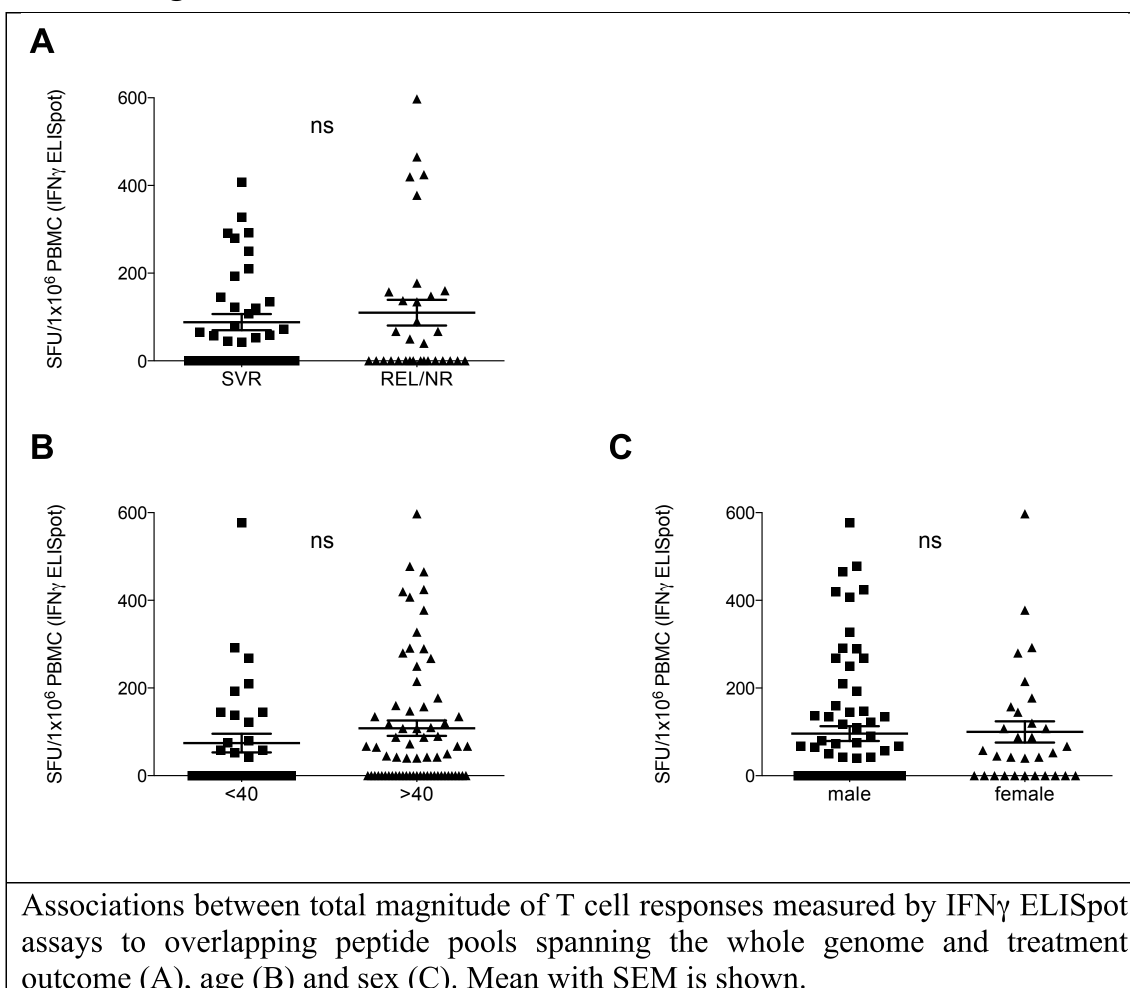
(A) Numbers of patients with favourable (TT) and non-favourable (GG/GT) IL28B genotype responding to IFNα/Ribavirin combination treatment. There was no influence of IL27B type on treatment outcome (SVR vs. REL/NR, Fisher's exact p=0.6250). (B) Pre-treatment viral load in patients with favourable and un-favourable IL28B genotype. Mean with SEM is shown.

There were no differences between groups achieving SVR and not responding to treatment (SVR vs. REL/NR, Fisher's exact p=0.6250). In the non-responder group, more patients had an unfavourable genotype (GT/GG) than in the other groups, however this difference did not reach statistical significance (Fisher's exact test, p=0.27, compared to patients achieving SVR). Additionally, there was no influence of IL28B genotype on pre-treatment viral load in the HCV gt3 chronic patients analysed in this study (Figure 3-23B).

3.10.2 Magnitude of T cell response

Subsequently, the magnitude of total T cell responses, measured by IFNγ ELISpot assays to overlapping peptides spanning the whole HCV genome, was associated with treatment response and known predictors of treatment outcome. Magnitude of T cell responses in chronic patients was not associated with treatment outcome, or with age and sex (Figure 3-24).

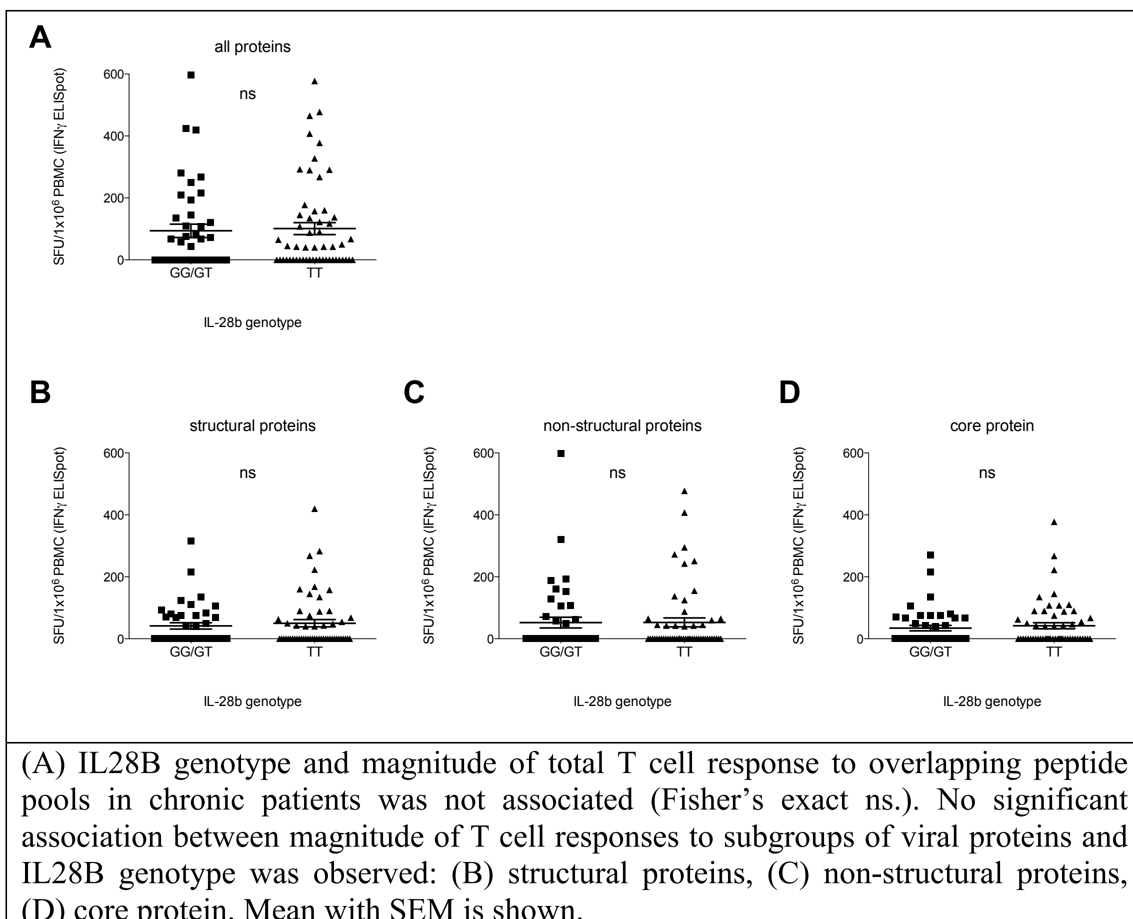
Figure 3-24: Associations of magnitude of T cell response and treatment outcome, age or sex.



Associations between total magnitude of T cell responses measured by IFNγ ELISpot assays to overlapping peptide pools spanning the whole genome and treatment outcome (A), age (B) and sex (C). Mean with SEM is shown.

Next, the total magnitude of IFNγ ELISpot responses was analysed regarding patients' IL28B genotype in 98 chronically infected patients (4 patients GG, 39 patients GT, 55 patients TT). There was no significant difference in magnitude of T cell response seen between both groups (Figure 3-25, A). In this work, we have observed significant differences in magnitude of T cell response between spontaneous resolvers, acute and chronic patients when comparing responses to different viral proteins separately (structural/non-structural proteins and HCV core), even if no significant difference was observed in total T cell response (Figure 3-9, page 99). However, when repeating this sub-division regarding patients' HLA types in chronically infected patients, no relation between T cell response and IL28B genotype was observed, even if different viral regions were considered (Figure 3-25, B-D).

Figure 3-25: Magnitude of T cell response and IL28B genotype.



(A) IL28B genotype and magnitude of total T cell response to overlapping peptide pools in chronic patients was not associated (Fisher's exact ns.). No significant association between magnitude of T cell responses to subgroups of viral proteins and IL28B genotype was observed: (B) structural proteins, (C) non-structural proteins, (D) core protein. Mean with SEM is shown.

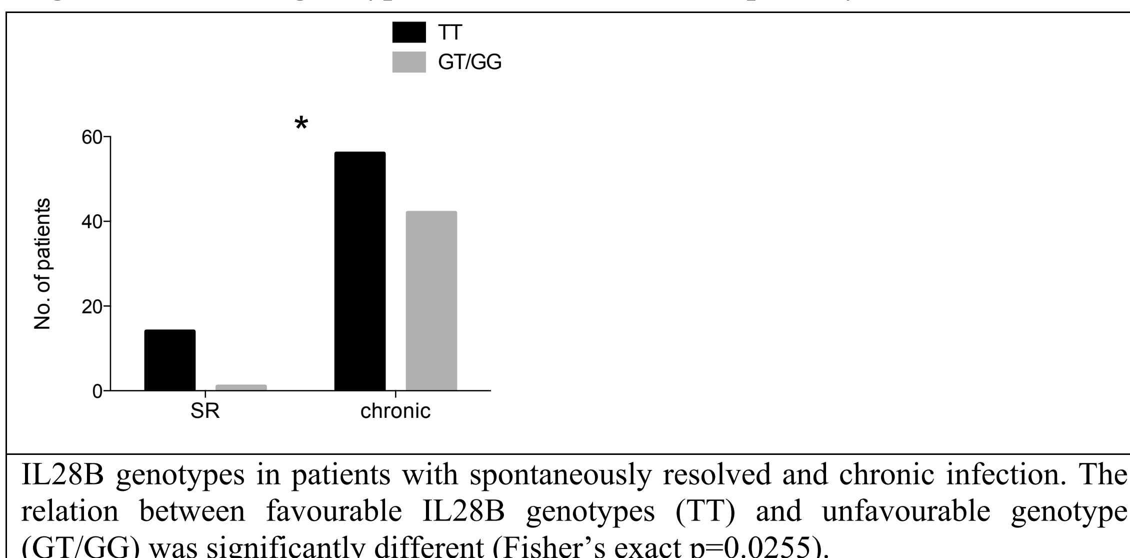
3.10.3 IL-28B status and outcome of primary infection

Finally, the influence of IL28B status on spontaneous resolution was assessed within our cohort. IL28B genotypes at SNP rs8099917 were determined in 15 out of 17 (88%) patients with spontaneously resolved infection, 98 out of 108 (91%) chronically infected patients, and 8 out of 16 (50%) patients acutely infected with HCV gt3 infection. At this SNP, the favorable gene is 'TT'. Polymorphisms 'GT' and 'GG' were grouped since associated with a non-SVR in HCV genotype 1 infection (Ge et al. 2009)

<u>IL28B type and spontaneous resolution of HCV infection</u>

When comparing IL28B genotype and clinical outcome, there was a significant difference between patients who spontaneously resolved infection and developed chronic infection; as expected, significantly more patients had the favorable genotype TT in the spontaneously resolved group than in the group that developed chronic HCV gt3a infection (Figure 3-26).

Figure 3-26: IL28B genotype and clinical outcome of primary infection.



IL28B genotypes in patients with spontaneously resolved and chronic infection. The relation between favourable IL28B genotypes (TT) and unfavourable genotype (GT/GG) was significantly different (Fisher's exact p=0.0255).

3.11 Discussion and conclusions

T cell responses are crucial in determining the outcome of acute HCV infection. Strong and multi-specific T cell responses are associated with spontaneous resolution of infection, whereas progression to chronic infection is characterized by weak and narrowly focused T cell responses. HCV has 7 major infecting viral subtypes, which vary in about 30-35% of nucleotides (Simmonds 2004). Adaptive immune responses have been studied comprehensively in HCV gt1 infection, but data on cross-reactivity of responses to other genotypes (Yi Wang et al. 2006; Fytili et al. 2008; L. Duan et al. 2011) and comprehensive analyses of T cell immunity in non-gt1 subtypes remain scarce (Giugliano et al. 2009; Humphreys et al. 2012).

Although numerous publications included patients infected with HCV gt3a (Schulze zur Wiesch et al. 2007), studies using specific gt3a cohorts and specific gt3 peptide sets are limited to one investigating T cell responses to the NS3 protein (Giugliano et al. 2009) and a recent study from our laboratory (Humphreys et al. 2012). To date, no comprehensive comparison of responses in HCV gt1 and gt3 infection using a subtype specific peptide set is published.

In this Chapter, I aimed to comprehensively assess T cell responses in HCV genotype 3 infection and compare them to previously studied T cell immunity in gt1 infection.

Here, a cohort of spontaneously resolved patients (n=16), and acutely (n=16) and chronically (n=108) HCV gt3 infected patients were assessed, allowing a comprehensive comparison to previously published T cell targets described for HCV gt1. T cell responses were tested by two different peptide sets: a set of overlapping peptides (15-18AA, overlapping by 11AA) spanning the entire HCV genome based on a HCV gt3 consensus sequence, and HLA predicted peptides with single peptides from the HCV non-structural regions. The two peptide sets were complementary and mostly detected different responses.

T cell responses detected by HCV gt3 overlapping peptides:

Multiple new T cell targets in HCV gt3 infection have been identified using a set of overlapping peptide pools as a screening methodology. Since this peptide set is covering the whole genome, we are confident that this description of HCV gt3 T cell immunity is based on a comprehensive assessment.

T cell responses in ***spontaneously resolved infection*** detected by overlapping peptide pools were broad and strong, targeting multiple regions all over the viral genome. HCV gt3a specific T cell responses in spontaneous resolvers were significantly higher than gt3a specific responses in chronic HCV infection (see 3.5.2, page 94), but not as high as responses in spontaneous resolvers described in the literature in gt1 infection (Lauer et al. 2004).

Spontaneously resolved patients were not antibody typed for infecting genotype. T cell responses in spontaneous resolvers were therefore tested to HCV gt3 and gt1 peptides, with significantly lower responses to the gt1 peptide set compared to gt3. Following from strong T cell responses against gt3 but not gt1 peptides, we deduct that most of spontaneously resolved patients are likely to have cleared an infection with the HCV gt3 subtype. This also seems a likely scenario since HCV gt3a infection is very common in our patient cohort and patient infected with gt3 are more likely to clear the infection spontaneously than those infected with HCV gt1 (M. Lehmann et al. 2004; Bortolotti et al. 2005).

In ***chronically HCV gt3 infected patients***, T cell responses were weak, with CD4+ T cells mainly targeting structural proteins, particularly core, and CD8+ T cell targeting non-structural proteins. Two frequently detected individual peptides were identified: a HCV core peptide targeted by CD4+ T cells ($core_{143}$, PVGGVARALAHGVRAL), and a HCV NS3 peptide ($NS3_{1520}$, RPSGMFDSVVLCECYDAGCSWYDL) targeted by CD8+ T cells. The equivalent gt1 peptide to HCV gt3 $core_{143}$ is known to be a DRB1*1101 restricted CD4 T cell target (Godkin et al. 2001; C. L. Day et al. 2002; A. J. MacDonald et al. 2002; Schulze zur Wiesch et al. 2005; Schulze Zur Wiesch et al. 2012). A response to the gt1 equivalent of peptide $NS3_{1520}$ (MFDS**S**VLCECYDAGC) was previously detected in HCV gt1 infection in a single patient (Ciuffreda et al. 2008).

Overall, T cells in chronically infected individuals targeted non-core regions more readily than in HCV gt1 infection (Semmo et al. 2005; G Harcourt et al. 2006; V. M. Fleming et al. 2010). It has been shown previously that the NS3 region is more readily targeted in HCV gt3 infection (Hultgren et al. 2004), an observation also made in this study. However, Hultgren *et al.* used thymidine proliferation assays stimulating T cell using whole NS3 protein, and it is therefore expected that mainly CD4+ responses are

detected. In contrast, majority of T cell responses targeting HCV non-structural proteins in chronically infected patients in this study were CD8+ restricted.

In ***acutely infected patients***, T cell responses of variable magnitude were detected. Magnitude of responses in this cohort was not linked to the outcome of infection during acute infection. However, a trend towards a stronger response in patients who spontaneously cleared than in patients developing chronic infection was observed when comparing the mean responses (352 SFU/10^6PBMC vs. 44 SFU/10^6PBMC). This is in line with published literature, where strong and broad responses have been described in patients clearing HCV infection spontaneously (C. L. Day et al. 2002; Lauer et al. 2004; A. L. Cox, Mosbruger, Lauer, et al. 2005). However, it has been reported recently that T cell responses in acutely infected patients are initially strong in both patients clearing the infection spontaneously and patients developing chronic infection (Schulze zur Wiesch et al. 2012). However, T responses disappear rapidly from the blood in patients developing chronic infection. The trend towards stronger responses in patients clearing infection in the cohort analysed in this study therefore might have been due to a selection bias. Since patients with resolving infection shown stronger clinical symptoms, they may have been diagnosed earlier, when T cell responses were still strong and broad.

When comparing responses detected by overlapping peptide pools between patient groups (SR, acute and chronically infected), a striking differential distribution of CD4+ and CD8+ restricted responses across the HCV viral genome was observed between chronically infected patients and those with spontaneous resolution of infection. CD4+ restricted T cells in chronic infection predominately targeted HCV structural regions, but mainly HCV non-structural regions in spontaneously resolved infection. In contrast, CD8+ cells in chronic infection mainly targeted HCV non-structural regions, whereas the both structural and non-structural regions were targeted in spontaneously resolved infection. This differential distribution of T cell responses aimed at structural *vs* non-structural HCV proteins in spontaneously resolved and chronic HCV infection has not been previously described in the literature. In some studies, though, a differential hierarchy in the role of $CD4^+$ and $CD8^+$ T cell responses was observed: In acute infection in chimpanzees, effective CD4+ responses were associated with clearance of HCV infection, whereas CD8+ T cells seem to evolve in the blood of the acutely infected patients regardless of outcome (Kaplan et al. 2007).

In this Chapter, HCV gt1 and gt3 specific T cell responses in chronic infection were compared: No difference in magnitude of responses was observed, however, when comparing targeted regions of the HCV viral genome, more responses to non-structural regions were observed in HCV gt3 infection compared to gt1. This variation in distribution of responses between HCV gt1 and other HCV genotypes was previously noted, with CD4+ T cells of HCV non-gt1 patients responding to HCV non-structural regions more commonly than HCV gt1 patients (Schulze zur Wiesch et al. 2007).

A potential reason for differential recognition of different viral regions between HCV gt3 and 1 may be that the HCV gt1b peptide set used was based on a single patient sequence strain (J4b), like most peptide sets previously used in the literature (Lauer et al. 2002; A. L. Cox, Mosbruger, Mao, et al. 2005; Schulze zur Wiesch et al. 2005). In contrast, the cohort adapted HCV gt3 peptide set was based on a full length consensus sequence derived from 15 Oxford patients sequences, an approach previously used for a comparative study of T cell responses to NS3 in gt1 and gt3 infection (Giugliano et al. 2009). I hypothesize that responses to conserved regions like HCV core will be readily detected using a peptide set based on a single unrelated HCV patient sequence (as used for HCV gt1), and less responses in variable regions. In contrast, a peptide set matched to the cohort consensus like the HCV gt3 overlapping peptide set might be more tailored to detect responses in variable HCV regions. To ensure that the Oxford HCV gt3 cohort did not represent a single outbreak leading to readily detectable CD8 responses in the HCV non-structural regions, viral diversity was assessed in the Oxford HCV gt3 cohort using full-length HCV sequences (Supplementary Figure 8-1, page 288), data published in (Humphreys et al. 2012)). Phylogenetic analysis showed significant diversity and a gt3 reference strain (accession number D28917) fell within the Oxford gt3 cluster.

Another potential reason that responses in HCV gt1 were restricted to HCV core may be that a subgroup of gt1 patients was infected with subtype 1a and therefore did not match the peptide strain used for screening (1b). However, when responses in a sub-group of 1a patients were assessed using 1a peptides (strain H77), no striking differences in the pools targeted were observed (Supplementary Figure 8-1, page 288, data published in (Humphreys et al. 2012)).

Potential problems with using overlapping peptide pools for the assessment of T cell targets are the large number of cells needed for initial screening and subsequent mapping experiments. Work intensive experiments with numerous mapping steps are needed to map responses to single epitopes. Using a peptide matrix would have been an experimental alternative, however, this approach requires a large number of cells for the initial experiment, potentially leaving less material for further functional analyses of responses. A second potential drawback with using a peptide set based on a consensus sequence is the possibility that the consensus sequence contains variants common in chronically infected patients, potentially limiting the detection of T cell responses primed by the wild type sequence.

T cell responses detected by HLA predicted peptides

In addition to overlapping peptides, a novel approach to identify gt3 specific T cell targets was used in this thesis: HLA predicted peptides.

The rationale for a second peptide screening approach was based on the considerations, that (i) T cell responses may be missed screening with overlapping peptides of 15-18AA length; (ii) Fewer PBMCs may be used in screening assays using a peptide set tailored to the tested patients' HLA type; (iii) Time efficient functional characterization using epitope specific tetramers may be achieved if a peptide is identified using an HLA predicted peptide, where the optimal binding peptide and presenting HLA type is known. Using overlapping peptide pools as a screening approach, work intensive experiments have to be performed to define T cell immunity to the single peptide, peptide optimals and restricting HLA types.

The HLA predicted peptide set was based on a sequence polymorphisms in a dataset of 136 HCV gt3 sequences; sequence polymorphisms were linked to patients HLA types to identify potential immune mediated escape mutations (Rauch et al. 2009). We subsequently evaluated these polymorphisms that were associated with patients' HLA types using online epitope prediction programs to identify putative T cell epitopes, for which peptides were manufactured and used in ELISpot assays matched to the patient's HLA types.

Different percentages of patients responded to HLA predicted peptides in spontaneously resolved patients, acutely and chronically infected patient groups. Even though broad and strong responses were detected in spontaneous resolvers using the overlapping peptide pools, only two patients (13%) mounted a response to HLA

predicted peptides. A potential reason for this comparative lack of responses is that the HLA predicted peptide set was based on sequence polymorphisms seen in a cohort of chronically infected gt3 patients: a T cell target observed in chronic patients with concomitant escape mutations is unlikely to be an epitope in spontaneously resolved patients whose T cell response cleared the infection. In contrast, responses detected by HLA predicted peptides in acutely and chronically gt3 infected patients were more abundant than in spontaneous resolved patients, detected in 33% and 22% of patients, respectively. A higher number of T cell responses detected in chronic patients compared to spontaneous resolvers were expected; since the peptide set was based on escape mutants within epitopes had been associated with HCV persistence previously (Bowen and Walker, 2005, Cox, Lauer, 2005, Tester 2005),

Overall, only a limited number of T cell targets were detected, with responses to HLA predicted peptides only seen to 8 out of a total of 65 peptides. The lack of responses might be due to either experimental limitations or weaknesses of the original peptide design that are discussed further in the following section.

***Experimental outline*:** Several potential reasons for the lack of detected responses are inherent to our study design: First, patient T cells may be sequestered in the liver, therefore not detectable in peripheral blood and not available for analysis (Schirren et al. 2000; Guidotti & Iannacone 2013). Second, weak *ex vivo* T cell responses under detection threshold may have been missed, potentially due to inefficient priming of T cell responses or exhausted T cells in chronic infection (Rehermann 2009; Velazquez et al. 2012). Third, limited numbers of HLA specific peptides may have been detected due to an insufficient number of HLA matched patients within our patient cohort, especially possible for uncommon HLA types.

In addition, potential weaknesses discussed by the authors in the original study describing sequence polymorphisms associated with patient HLA type may have contributed to limited peptide detection using HLA predictive peptides (Rauch et al. 2009): Due to an underrepresentation of rare HLA types associations of polymorphisms to patients' HLA types may have been missed. In addition, 40% of individuals enrolled in the analysis were co-infected with HIV, potentially limiting T cell pressure due to the lack of CD4+ help.

***Statistical methods*:** Other potential drawbacks relate to statistical methods initially used for the linkage of polymorphisms to HLA type (Rauch et al. 2009): HLA associations were assessed using the Fisher's exact test, and multiple statistical tests

were subsequently applied to ensure statistically significant associations (phylogenetic correction, false positive discovery rates/q-values).

Phylogenetic corrections: Polymorphisms across the viral genome can be due to neutral evolution or natural selection (such as T cell pressure). In a group with an overrepresented HLA type (phylogenetic relatedness) and infection with a common ancestor virus, associations between HLA type and mutations may reflect a founder effect rather than immune pressure exerted by the associated HLA allele (Rauch et al. 2009). To address this issue, Rauch *et al*. identified clusters of relatively homogenous sequences (L. Kaufman & Rousseeuw 1990), in which HLA distribution should be random. After adjusting for cluster strata, a Mantel-Haenszel test was performed. This procedure combines the associations between viral polymorphisms and HLA alleles within clusters of possible related sequences, which should further decrease the risk of confounding through overrepresentation of HLA alleles in phylogenetically related sequences (Rauch et al. 2009).

False discovery rates (q-value): In Rauch *et al*.'s association study, multiple different values were associated to define HLA associations with sequence mutations. To avoid false positive due to multiple testing problems, a q-value of $q \leq 0.2$ was applied (J. D. Storey & Tibshirani 2003). This indicates that up to 20% of the significant associations identified may be false positive; therefore peptides that are not T cell targets may have been included in the peptide set further contributing to the lack of T cell responses detected in this study.

HLA types (Low resolution of HLA typing and Linkage disequilibrium): Further potential problem originates a low resolution of HLA typing used in the Rauch *et al*.'s study, with HLA typing only performed to a 2 digit level (e.g. A*02). As shown for A*02 and B*27 subtypes, similar HLA types from the same supertype can vary substantially in their binding preferences (D. Barouch et al. 1995; Hülsmeyer et al. 2002). Already a single amino acid substitution within the MHC binding complex can alter binding properties substantially (J. M. Burrows et al. 2007). Therefore, HLA subtype specific mutations might have been missed due to a low resolution of HLA typing.

***Epitope prediction*:** Described sequence polymorphisms associated with patient HLA type (Rauch et al. 2009) were further assessed using online available epitope prediction servers SYFPEITHI and BIMAS to define putative epitopes at matching HLA types (Rammensee et al. 1999; K. C. Parker et al. 1994). Both epitope prediction

algorithms are based on matrixes and score residues based on previously published T cell epitopes and MHC ligands (Lundegaard et al. 2010). Recent publications (H. H. Lin et al. 2008) have shown that matrix based computing algorithms are inferior to those based on methods like artificial neural networks (ANNs) or support vector machines (SMVs), since they cannot take correlated effects into account (e.g. where the contribution to the binding affinity at amino acid depends on amino acids at other positions of the peptide (H. P. Adams & Koziol 1995; Gulukota et al. 1997; S Buus et al. 2003; M. Nielsen et al. 2003)). Therefore, through using BIMAS and Syfpeithi as prediction programs for our peptide set design, we potentially included false positives in our dataset. Usage of other prediction servers based on ANNs (such as NetCTL) may have enabled the discovery of additional T cell targets and a minimization of false positive predictions of T cell targets.

In addition to these considerations, MHC binding predictions are always only estimates of the *in vivo* scenario: often immunodominant peptides in natural infection have much lower prediction scores than other peptides, where no T cell response is detectable *in vivo*. Furthermore, even though the peptide binding step is the most selective step in peptide presentation, several programs have been incorporating predictions on other steps of the antigen processing and presentation pathway, such as proteasomal degradation and TAP binding (MAPP, netCLT), which may have enabled a more comprehensive estimation of peptide presentation *in vivo*.

In conclusion, peptide sets were complimentary and mostly detected distinct T cell targets. Overlapping peptide pools detected a greater number of new peptides and may be a more comprehensive method for screening for new HCV gt3 specific T cell targets. However, using a targeted approach like HLA predicted peptides enabled the discovery of novel epitopes using limited PBMC samples. This was possible through a reduction of screening peptides tailored to the patients HLA type.

Cross-reactivity of HCV gt1 and gt3 specific immune responses

In this Chapter, a comprehensive comparison of HCV gt3a specific T cell responses detected in the Oxford cohort to previously described responses in HCV gt1 infection was performed.

T cell targets between HCV gt1 and gt3 varied considerably. Overlap between HCV genotype 1 and 3 was particularly scarce for HCV gt3 CD8+ restricted T cell targets

detected in the Oxford gt3 cohort, with only detected 3 epitopes (19%) matching HCV gt1 T cell targets described in the literature. For all other HCV gt3 CD8+ restricted T cell targets, sequence differences in described regions were observed. For CD4+ restricted T cell targets, potential overlap between genotypes was higher than that observed for CD8+ targets, possibly due to the promiscuous binding properties of MHC class-II complexes (Siebenkotten et al. 1998). Overall, even though for most genomic regions targeted in the Oxford HCV gt3 cohort, HCV gt1 specific T cell targets were previously described in the literature, sequence identity between HCV gt1 and 3 was only observed for the minority of epitopes.

Next, I examined whether CD8+ restricted epitopes commonly targeted in gt1 infection were detected in our gt3 cohort. Minimal overlap was found between dominant HCV gt1 epitopes and T cell targets detected in the Oxford HCV gt3 cohort, with only common epitope ATDALMTGY detected in both HCV gt1 and gt3 patients. Intriguingly, dominant HCV gt1 A02 restricted epitopes (KLVALGINAV and CINGVCWTV) were not detected in a single HCV gt3 infected patient. Of note, it has been previously described that epitope specific T cells did not recognize the gt3 version of epitope CINGVCWTV (Fytili et al. 2008).

Most likely, distinct T cell specificity is due to substantial sequence differences between HCV genotypes 1 and 3 (overall about 24%), which may abrogate HLA binding at critical anchor residues within viral epitopes. Also, genotype-specific polymorphisms outside the anchor positions may impair the recognition of the HLA-epitope complex by $CD8^+$ T cells. Additionally, divergent sequences flanking immunogenic epitopes may have a substantial impact on epitope processing: proteasomal processing predictions by Rauch *et al*. have shown that about 20% of genotype-associated amino acid variants influence proteasomal cleavage of target peptides (Rauch et al. 2009). It is therefore possible that, even though an immune response may be directed at the same genomic region, different epitopes are targeted in HCV gt1 and 3 infection.

I propose that further studies of subtype specific immune responses are necessary for informed cross-reactive vaccine design. Especially in understudied HCV genotypes like 5 and 6, with consensus sequences differing from HCV genotype 1, major differences in T cell targets compared to published data are to be expected.

Associations of patient parameters and treatment response

Known predictors of treatment response were assessed in the Oxford HCV gt3 cohort, both independently and in relation to the magnitude of detected T cell responses.

First, associations between treatment response (SVR and REL/NE) and age and sex were analysed. Younger patients achieved SVR more readily than older patients, but no difference in treatment response was observed between male and female patients. Similar results were reported in the literature: A strong association between younger age and SVR has been observed in several studies (Poynard et al. 2000; Manns et al. 2001; Fried et al. 2002; M. L. Shiffman et al. 2007). Female sex has been associated with SVR (Poynard et al. 2000), however, this association was not repeatedly observed in several big peg-IFNα/Ribavirin trials (Manns et al. 2001; Fried et al. 2002).

No difference in pre-treatment viral loads in patients subsequently achieving SVR compared to non-responders/relapsers was observed in this study. In contrast, low pre-treatment viral loads have been associated with a higher SVR rate in HCV genotype 3 infected patients in the literature (Kau et al. 2008), with baseline viral loads of 400,000, 400,000–800,000 and 800,000 IU/ml achieving SVR rates of 81, 70, and 59%, respectively.

Subsequently, the impact of IL28B status on HCV treatment response was evaluated, with no differences in treatment response observed between the patients with favourable versus un-favourable IL28B genotype. This is in line with data from a study published on HCV gt3a infection by our research group recently, with a bigger cohort (n=158 patients) including the patient group analysed in this thesis (Bucci et al. 2013): the SVR rate was 61% in the group with the favourable allele TT, and 63% in the GT/GG group (p=0.1419). Comparably, the SVR rates in the sub-group analysed in this thesis were 63% (TT allele) and 54% in the GT/GG group. No difference was observed in pre-treatment viral loads between patients with the favourable and un-favourable IL28B allele. Contrary, in Bucci *et al.*'s study, where patients analysed here were included, a significantly higher viral baseline viral load was observed for patients with the favourable IL28B genotype TT (Bucci et al. 2013).

Association of magnitude of responses

Next, I aimed to assess whether the magnitude of T cell responses was associated treatment responses to combination therapy with pegylated-Interferonα/Ribavirin, and predictors of treatment responses, such as age, sex and IL28B phenotype. No significant associations between magnitude of T cell response measured by ELISpot assay and either of the described parameters was observed. It is possible that significant associations have been missed: Patient groups analysed here only form a small test group, potentially preventing significant associations (patients achieving SVR n=40; patients relapsing/not responding n=31).

There is no consensus regarding the role of HCV specific T cell responses in outcome of HCV treatment. Some studies have suggested that pre-treatment proliferative capacity and magnitude of ELISpot CD4+ and CD8+ T cell responses are important for successful treatment response (Pilli et al. 2007; Rosen et al. 2007). Others have shown that responses induced during therapy play a role in SVR (Cramp et al. 2000; Kamal et al. 2002), whereas no association was found in several studies (Barnes et al. 2002; Aberle et al. 2007). Some groups, including our own, who monitored magnitude of T cell responses longitudinally, have shown a decrease of T cell responses during treatment in acute (F. Rahman et al. 2004; Lauer et al. 2005) and chronic (Barnes, Gelderblom, et al. 2009; Humphreys et al. 2012) infection. In chronic patients, the loss of T cell responses was more pronounced in the SVR group, suggesting that T cell responses might not be required to establish SVR (Barnes, Gelderblom, et al. 2009).

In conclusion, strong and broad HCV gt3 specific T cell responses were found in spontaneously resolved infection, whereas responses in chronic infection were low and mainly targeted two dominant epitopes in the HCV core protein and NS3; however, the impact of HCV gt3 specific T cell responses on treatment outcome remains illusive. Importantly, HCV gt3 specific T cell specificity is markedly different from that previously described in HCV gt1 infection.

4 Characterization of dominant T cell responses in HCV genotype 3a infection

4.1 Abstract

Background: Conserved immunodominant epitopes may act as attractive targets in cross-reactive immunogen design, especially in a highly variable pathogen like Hepatitis C virus (HCV). In this Chapter, I aimed to define frequently detected T cell targets in HCV genotype (gt) 3 infection and assess when they emerge during the course of infection. Since the failure to contain the virus in chronic HCV infection has been associated with mutational escape within targeted T cell epitopes, sequence variability within and between HCV genotypes and T cell cross-reactivity against identified HCV gt1 and gt3 sequence variants was assessed. Furthermore, T cell function and phenotype at dominant epitopes was evaluated.

Methods: T cell responses were assessed in 140 patients with HCV gt3 infection using IFNγ ELISpot assays, and defined as frequent if detected in >4 patients within the Oxford gt3 cohort. Sequence heterogeneity at frequently detected T cell targets was evaluated in the Oxford cohort and at population level, and T cell cross-reactivity assessed experimentally using ELISpot analyses. T cell phenotypic profiles were evaluated using tetramers and intracellular staining assays, implementing a novel methodology for the detection of low level responses at detection threshold based on pre-incubation of PBMC at high cell concentrations.

Results: Seven frequently detected HCV gt3 specific responses were identified; two CD4+ T cell responses targeting HCV core and five CD8+ T cell responses targeting non-structural proteins. CD4+ T cell responses to HCV core emerge once chronic infection is established; responses were low in magnitude, but could be increased by short-term incubation at high cell concentrations. Frequently detected HCV gt3 T cell targets were variable within HCV gt3 and between HCV genotypes, with limited T cell cross-reactivity against identified sequence variants detected in IFNγ ELISpot assays. HCV gt3 specific CD4+ and CD8+ T cells were monofunctional and showed low expression of activation markers.

Conclusions: At T cell targets frequently detected in HCV gt3 specific infection sequence polymorphisms are common and T cell cross-reactivity against viral variants is limited. Therefore, HCV gt3 epitopes detected in natural infection may not be ideal targets for cross-reactive vaccine design.

4.2 Introduction

An ideal HCV immunogen would induce immune responses that are cross-reactive against viral variants of all HCV genotypes. Conserved epitopes detected in natural infection may act as attractive targets, especially in a variable pathogen like HCV.

Only few epitopes have been consistently linked to spontaneous resolution of HCV infection (A. Y. Kim et al. 2010; Neumann-Haefelin et al. 2010; Fitzmaurice et al. 2011; Y. Kim et al. 2012); the evidence for the majority of dominant T cell epitopes remains ambiguous, since identical epitopes may be targeted in both patients with spontaneous resolution of infection and those who proceed to chronic infection, as shown for epitope CVNGVCWTV ($NS3_{1073}$) in HCV gt1 infection (Lechner et al. 2000; Thimme et al. 2001; Grüner et al. 2000; Lauer et al. 2004). In addition, it has been shown recently that epitopes linked to protection in HCV gt1 infection are not conserved across HCV genotypes, and that epitope specific T cells are not cross-reactive against viral variants observed in other HCV genotypes (Skibbe et al. 2014).

Therefore, several considerations support the use of conserved epitopes previously described in natural infection in cross-reactive vaccine design: Using a previously detected epitope implies that the viral region has proven immunogenicity, and the epitopes' level of sequence conservation may indicate that the epitope is not subject to immune mediated escape and may be cross-reactive between genotypes.

Viral escape

The failure to contain HCV in chronic HCV infection is associated with mutational escape within targeted T cell epitopes (Erickson et al. 2001). In individuals of the same HLA type identical escape mutations can be observed, as it has been demonstrated for some HCV gt1 specific CD8+ restricted epitopes (Gaudieri et al. 2006; Neumann-Haefelin et al. 2006). However, functional constraints and fitness costs limit the amount of tolerated escape mutations, and reversion to the wild type sequence after viral transmission to an individual without the restricting HLA type has been observed in HIV (T. M. Allen et al. 2004; Friedrich et al. 2004) and HCV infection (Timm et al. 2004; S. C. Ray et al. 2005). In contrast, other mutations can be stable and persist after transmission, driving viral sequence evolution, as suggested for HCV non-structural proteins (A. L. Cox, Mosbruger, Mao, et al. 2005).

T cell cross-reactivity

Even if an immunodominant epitope is subject to mutational escape, T cells primed against wild type sequences may show some degree of cross-reactivity to other sequence variants. In the HCV literature, there is only limited evidence on the ability of epitope specific T cells to recognize sequence variants from other genotypes. Even though some T cell cross-genotype reactivity has been described (Giugliano et al. 2009; Humphreys et al. 2012; Skibbe et al. 2014), a systematic analysis considering all possible variants at dominant epitopes within and across genotypes has only been published for the immunodominant HCV gt1 epitope CVNGVCWTV ($NS3_{1073}$) (Fytili et al. 2008). For this epitope, T cell cross-reactivity of the response primed against the HCV gt1a variant against variants from HCV genotypes 1, 4-6 was detected, but responsiveness against HCV gt2 and gt3 variants was low. In addition, limited cross-reactivity within HCV gt1 peptide variants was observed at an immunodominant HLA A*02 restricted epitope in work by our research group (Kelly et al., manuscript in preparation). Depending on the analysed epitope, a single amino acid substitution may either completely prevent the recognition by wild-type primed CD8+ T cells or allow for some T cell cross-reactivity between variants (Neumann-Haefelin et al. 2008; Fytili et al. 2008; Söderholm et al. 2006). Thus, an aim of this Chapter was to study the sequence variability at dominant HCV gt3 epitopes and to assess the capacity of T cells specific to cross-react with identified HCV viral variants at these epitopes.

T cell function

T cell function is thought to play an important role in effective viral clearance. A series of studies in viral infection have shown that distinct functional signatures correlate with levels of viral replication and disease activity (reviewed by (Pantaleo & Harari 2006)). In particular, T cell responses linked to effective viral clearance or viral control in chronic viral infections like HCV and HIV infection show a high degree of polyfunctionality (Lauer et al. 2005; J. R. Almeida et al. 2007; Ciuffreda et al. 2008; Seder et al. 2008; I. G. Rodrigue-Gervais et al. 2010). Polyfunctionality of T cells has been defined as the ability to produce cytokines like IL-2 and IFNγ and to proliferate (Pantaleo & Harari 2006). So far, T cell functionality at dominant HCV gt3 specific epitopes has not been defined.

Aims of this Chapter

The aims for this Chapter were therefore (Figure 4-1):

(1) To identify T cell targets frequently detected in HCV gt3 infection.

HCV gt3 T cell specificity shows major differences to that previously described for HCV gt1. In Chapter 3 of this thesis, multiple T cell responses to HCV gt3 peptides were identified; in this Chapter, I aim to identify T cell targets frequently detected in HCV gt3 infection.

(2) To assess when frequently detected HCV gt3 specific CD4+ T cell responses evolve over the course of clinical infection.

To further define the role of frequently detected HCV gt3 specific CD4+ responses in HCV infection, I will analyse when responses occur during the course of infection, and whether they are abundant in acutely or chronically infected individuals, or spontaneously resolved patients.

(3) To define sequence variation within HCV gt3 and between HCV genotypes 1 and 3 at frequently detected T cell targets.

To assess whether frequently detected T cell targets defined in HCV gt3 infection may serve as targets for cross-reactive immunogen design, I will define the level of sequence conservation in frequently targeted regions. Identified sequence variants within HCV gt3 and between HCV genotypes will inform assessment of T cell cross-reactivity experiments between defined sequence variants.

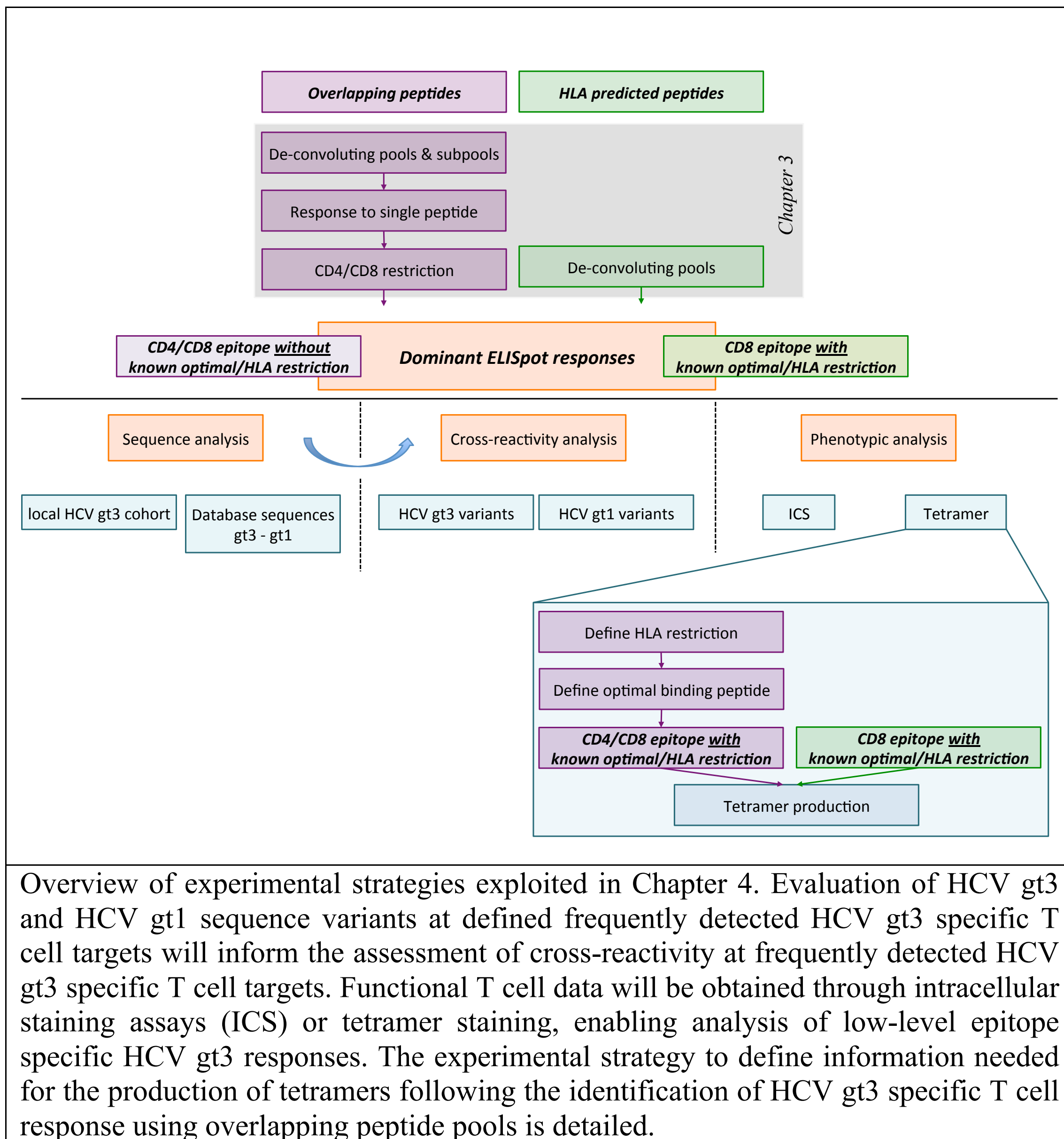
(4) To evaluate T cell cross-reactivity against HCV sequence variants.

Cross-reactivity of T cells primed in natural HCV gt3 infection against identified viral sequence variants will be evaluated in IFNγ ELISpot assays.

(5) Define the cellular phenotype and function of HCV gt3 specific responses.

Finally, the phenotype and function of T cells targeting frequently detected T cell targets will be evaluated using flow cytometry based techniques.

Figure 4-1: Experimental overview for Chapter 4.



Overview of experimental strategies exploited in Chapter 4. Evaluation of HCV gt3 and HCV gt1 sequence variants at defined frequently detected HCV gt3 specific T cell targets will inform the assessment of cross-reactivity at frequently detected HCV gt3 specific T cell targets. Functional T cell data will be obtained through intracellular staining assays (ICS) or tetramer staining, enabling analysis of low-level epitope specific HCV gt3 responses. The experimental strategy to define information needed for the production of tetramers following the identification of HCV gt3 specific T cell response using overlapping peptide pools is detailed.

4.3 Selection of frequently detected T cell responses in HCV gt3 infection in the Oxford HCV genotype 3 cohort.

T cell responses to 41 HCV gt3 individual peptides were described in this thesis (Table 3-6 and Table 3-7, page 113 and 114). Of these, 10 HCV gt3 specific T cell targets were located in the structural region and 31 in the non-structural region.

Frequently detected T cell targets were defined as individual peptides detected in at least four patients in the Oxford HCV gt3 cohort, and selected for further analysis. 7 T cell targets were defined as frequently detected, with 4 to 14 patients responding to each individual peptide (Table 4-1). Both CD4+ and CD8+ T cells targeted frequently

detected individual peptides. Two frequently detected peptides targeted by CD4+ T cells were located in the HCV core region, whereas the remaining 5 frequently detected peptides were targeted by CD8+ T cells and located in HCV non-structural regions. In this Chapter, CD4+ and CD8+ responses will be discussed separately.

Table 4-1: Frequently detected HCV gt3 specific T cell responses.

Frequently detected HCV gt3 specific T cell response						**Patients responding**			
Viral region	**AA position**	**Sequence**	**HLA**	**CD4/ CD8**	**Pept. set**	**S**	**C**	**A**	**Total**
Core	66-83	PKARRSEGRSWAQPGYPW		CD4	OP		5		**5**
	143-158	PVGGVARALAHGVRAL		CD4	OPs		11	1TxN	**12**
NS2	886-896	LLYPSLIFDI	A02	CD8	HLA		2	1AC	**3**
		LYPSLIFDI	A24	CD8			3	1AC	**4**
NS3	1443-1451	ATDALMTGY *	A01	CD8	HLA		3	1TxS	**4**
	1520-1537	RPSGMFDSVVLCECYDAGCSWYDL		CD8	OPs	2	12		**14**
NS4b	1853-1862	RVLLDILAGY	A26	CD8	HLA		3	1TxS 1SR	**5**
NS5a	2126-2141	AEFFTEVDGVRLHRYA		CD8	OPs	2		2TxS	**4**

T cell responses frequently detected in HCV gt3a infection are depicted. For each targeted individual peptide, the amino acid (AA) position, peptide sequence, restricting HLA type, CD4/CD8 restriction and detecting peptide sets are specified. The total number of patients responding to the peptide and their status of infection (S: spontaneously resolved, C: chronic, A: acute [AC: acute proceeding to chronic, TxS: treated achieving SVR, TxN: not responding to treatment, SR spontaneously resolved]) is detailed. An epitope previously described in HCV gt1 infection is marked with a star.

4.4 CD4+ restricted T cell responses to HCV core frequently detected in chronic genotype 3a infection

T cell responses in HCV chronically infected individuals have been described as highly attenuated, especially in peripheral blood (Lauer et al. 2004; Bowen & C M Walker 2005a; A. L. Cox, Mosbruger, Mao, et al. 2005). Proliferation assays and T cell lines in chronically infected patients have typically found limited CD4+ T cell responses to HCV structural and non-structural proteins (Diepolder et al. 1995). In contrast to this data, we and others have consistently detected robust, but low-level CD4+ restricted IFNγ T cell responses specific to HCV core in HCV gt1 infection (Semmo et al. 2005; Ruys et al. 2008; V. M. Fleming et al. 2010). These responses had low proliferative capacity and a loss of IL-2 production (Barnes et al. 2002; Semmo et al. 2005; G Harcourt et al. 2006; Ruys et al. 2008; V. M. Fleming et al. 2010).

In the Oxford HCV gt3 cohort, core specific CD4+ response were almost exclusively observed in chronically infected individuals, suggesting that these responses are not contributing to clearing HCV infection. However, there is scarce data suggesting that CD4+ responses may have a protective role against liver disease; it has been shown that rapid progression of liver disease in chronically HCV/HIV patients appeared to be most dramatic in persons with $CD4^{+}$ T cell counts of <200 cells/mm^3. Reconstitution of immunity through antiretroviral therapy leads to a decrease in the rate of progression of fibrosis and risk of clinical events due to liver disease (Qurishi et al. 2003). It has also been suggested that core specific CD4+ cell may form a regulatory T cell population (Langhans et al. 2010; Dominguez-Villar et al. 2012). However, the exact role of core specific CD4+ responses in HCV infection, particularly HCV gt3 infection, remains unclear.

In addition to these considerations, HCV core, as a region highly conserved between genotypes, would likely be included in a vaccine consisting of conserved regions and aiming to induce T cell responses cross-reactive between HCV genotypes. To compare core-specific T cell responses detected in natural infection to those induced by a potential vaccine, core-specific response will have to be further defined.

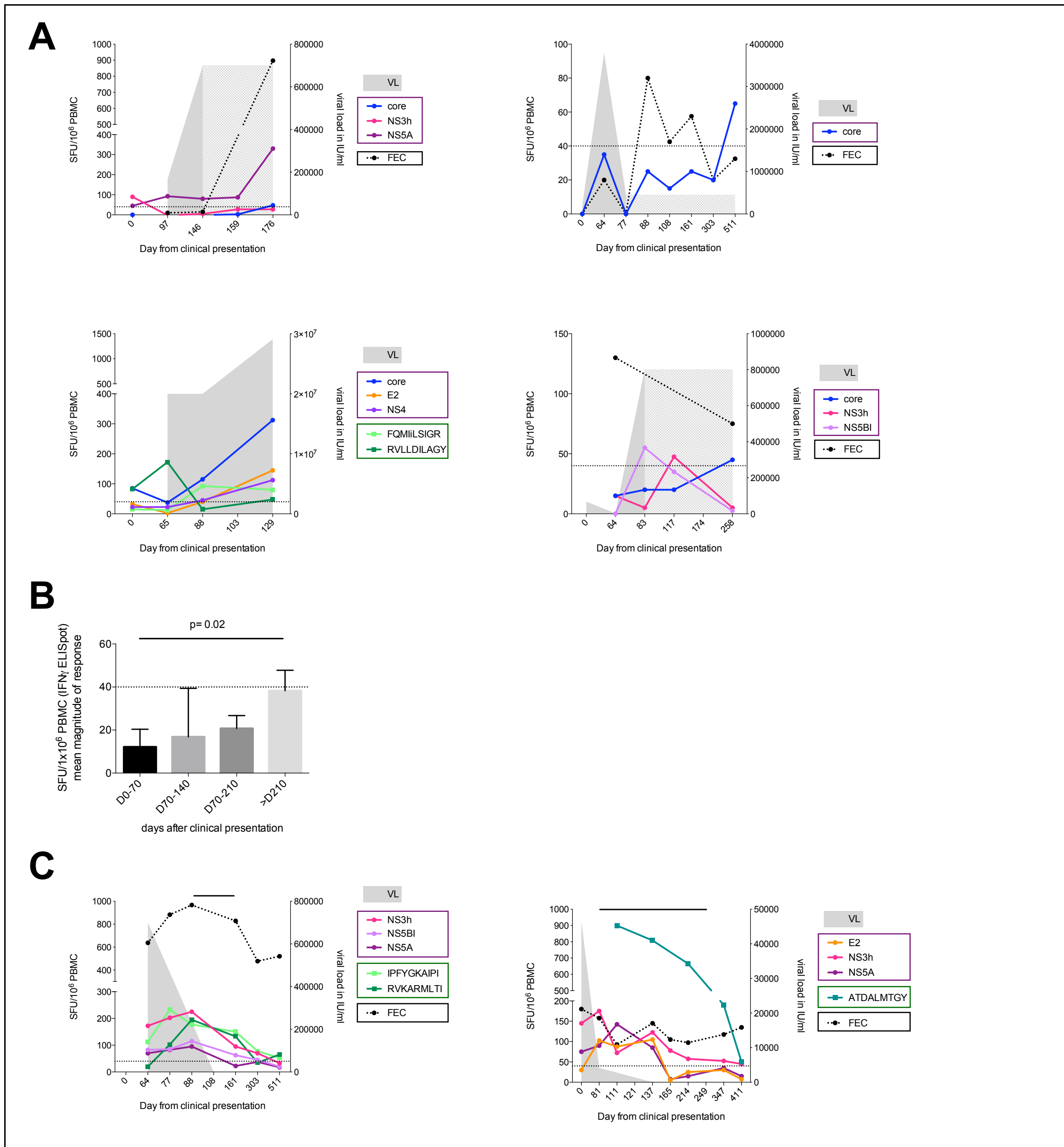
Therefore, core-specific T cell responses of low magnitude frequent in HCV gt3 infection were analysed further. Two frequently targeted CD4+ restricted T cell targets in the HCV core region have been identified: core 66-83 (`PKARRSEGRSWAQPGYPW`) and core 143-158 (`PVGGVARALAHGVRAL`), both detected in more than five chronically HCV gt3 infected patients. Since these responses were not observed in spontaneously resolved patients, we aimed to establish when responses arise through tracking responses longitudinally from acute infection into chronic disease. In addition, sequence variation within HCV gt3 and between HCV gt3 and gt1 infection, cross-reactivity between viral variants and the cellular phenotype at frequently detected CD4+ restricted T cell targets was assessed.

4.4.1 Core responses in HCV gt3 infection evolve over time in acutely infected patients developing chronic infection

To establish at which point during HCV infection core specific CD4+ responses emerge, we analysed T cell responses in acutely infected patients over time (Figure 4-2). For 4 acutely infected patients developing chronic infection, PBMC for multiple time points were available. In all four patients, responses to core over threshold (40

SFU/10^{6}PBMC) were observed, with core responses increasing over time. Responses to other HCV viral proteins did not consistently increase over time, further supporting the distinct role of HCV core (Figure 4-2 A).

Figure 4-2: Viral load and HCV specific T cell responses in acutely infected patients measured over time.



T cell responses monitored over the course of acute infection measured by IFNγ ELISpot assays. T cell responses to overlapping peptide pools and individual HLA predicted peptides are depicted, as well as HCV viral load. (A) T cell responses in four acutely HCV gt3 infected patients proceeding to chronic infection. Shaded areas represent HCV viral loads, light grey shadings were added when values were missing to represent that the patient developed chronic infection. (B) HCV specific core responses increase over the duration of acute infection. (C) T cell responses in acutely infected patients treated with peg-IFNα during the acute phase of infection subsequently achieving sustained virological response (SVR). VL - viral load; FEC - Flu, EBV and CMV.

When analysing HCV gt3 core specific T cell responses over time in all acutely infected patients developing chronic infection (n=6, Figure 4-2 B), a significant increase in magnitude of T cell response to HCV core measured by IFNγ ELISpot analysis was observed.

In contrast, core specific T cell responses were not detected in patients clearing the infection under combination therapy (pegylated Interferon-α/Ribavirin) during the acute phase (Figure 4-2 C). Of note, HCV specific T cell responses declined during treatment, a phenomenon observed previously by our research group and contributed to either the immunomodulatory effect of interferons or increased T cell homing to the liver (Humphreys et al. 2012). Both here and in our previous publication, a concomitant decrease of a pool of CD8+ specific flu, EBV and CMV specific epitopes was noted during treatment (Humphreys et al. 2012).

In summary, HCV core specific T cell responses in HCV gt3 infection evolve over the course of acute infection, and are readily detectable once chronic infection is established.

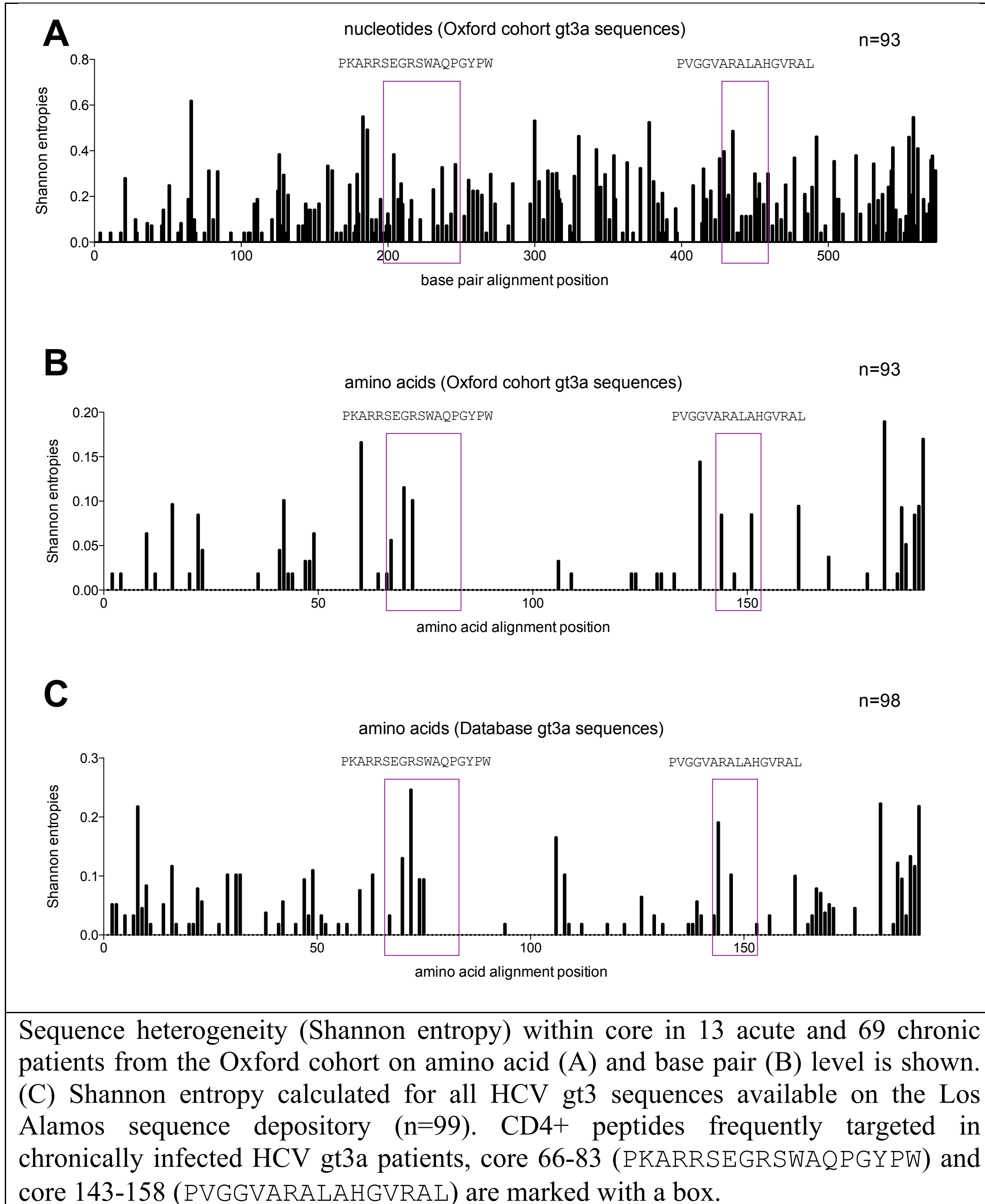
4.4.2 Sequence variability at CD4+ epitopes frequently targeted in HCV gt3 infection

Sequence heterogeneity within the HCV gt3 core region

Even though core specific CD4+ T cell responses were detected frequently in our chronic HCV gt3 cohort, they do not control virus in persistent infection. A possible reason for CD4+ T cell failure may be immune mediated escape within targeted epitopes. Immune mediated escape has been frequently observed in common CD8+ epitopes (Bowen & C M Walker 2005b; A. L. Cox, Mosbruger, Mao, et al. 2005), has been associated with viral persistence and also has an effect on T cell cytokine secretion (H. Wang & D D Eckels 1999; J. H. Wang et al. 2003). However, mutational escape at CD4+ restricted T cell epitopes in persistent infection has not yet been thoroughly analysed, as CD4+ T cell responses in chronic infection are often weak and difficult to detect. We therefore aimed to define sequence variants at two epitopes frequently targeted in chronic HCV gt3 infection, core 66-83 (PKARRSEGRSWAQPGYPW, 54% of mapped responses) and core 143-158 (PVGGVARALAHGVRAL, 23% of mapped responses).

HCV gt3 sequences of HCV core were obtained in 80 chronic and 13 acute patients from the Oxford cohort. Sequence diversity for all sequences was calculated using the Shannon entropy calculation (see 2.2.8, page 71), and is depicted in Figure 4-3.

Figure 4-3: Shannon entropy of core sequences.



Sequence heterogeneity (Shannon entropy) within core in 13 acute and 69 chronic patients from the Oxford cohort on amino acid (A) and base pair (B) level is shown. (C) Shannon entropy calculated for all HCV gt3 sequences available on the Los Alamos sequence depository (n=99). CD4+ peptides frequently targeted in chronically infected HCV gt3a patients, core 66-83 (PKARRSEGRSWAQPGYPW) and core 143-158 (PVGGVARALAHGVRAL) are marked with a box.

Next, variability at T cell targets (mean Shannon entropy of the targeted peptide) was compared to the overall variability of the HCV region (mean Shannon entropy for HCV core). Variability at nucleotide level within targeted regions was not higher than

in core regions not frequently targeted in HC gt3 infection (Figure 4-3 A). However, amino acid substitutions at 4 (PKARRSEGRSWAQPGYPW) and 3 (PVGGVARALAHGVRAL) positions were observed within the cohort at frequently targeted peptides (Figure 4-3 B). When analysing a bigger HCV gt3 sequence dataset obtained from the Los Alamos database (n=98), multiple amino acid substitutions were observed at the two frequently detected T cell targets, with peptide PKARRSEGRSWAQPGYPW (66-83) being one of the most variable regions within HCV core (Figure 4-3 C).

Sequence variants at frequently detected T cell targets with in the Oxford HCV gt3 cohort

Next, sequence variability was assessed at CD4+ T cell targets frequently detected within the Oxford HCV gt3 cohort (Figure 4-4). In addition, it was evaluated whether sequence polymorphisms occurred more commonly in patients with detectable a CD4+ T cell response to the according epitope in IFNγ ELISpot assays, which may suggest immune mediated escape.

For T cell target core 66-83 (PKARRSEGRSWAQPGYPW), sequence polymorphisms were observed at amino acid positions 1, 2, 5 and 7 (**PK**AR**R**S**E**GRSWAQPGYPW/ **1** L; **2** A/E/R; 5Q; **7** D/G) (Figure 4-4 A), with substitution (R→Q) PKAR**R**SEGRSWAQPGYPW being most abundant and detected in 11% of sequenced individuals. For T cell target core 143-158 (PVGGVARALAHGVRAL), sequence polymorphisms were observed at positions 2, 5, and 9 (P**V**GG**V**ARA**L**AHGVRAL/ 2**A**; 5**A**; 9**F**) (Figure 4-4 A). Polymorphisms were not more common in patients targeting the two frequently detected HCV core CD4+ T cell targets (marked in grey in Figure 4-4). This suggests rare immune mediated escape in these CD4+ restricted core epitope, a finding supported by previously reported experimental data observing rare mutational escape in CD4+ epitopes in humans (V. M. Fleming et al. 2010) and chimpanzees (Fuller et al. 2010).

Figure 4-4: Sequence information at frequently detected T cell targets `PKARRSEGRSWAQPGYPW` **and** `PVGGVARALAHGVRAL` **in the HCV core region in HCV gt3 patients with and without a T cell response.**

A

```
       PKARRSEGRSWAQPGYPW    262  ------------------    570   ------------------
128    ------------------    266  ------------------    579   ------------------
275    ------------------    267  ------------------    582   ------------------
416    ------------------    274  ----Q-------------    589   ------D-----------
450    ------------------    278  ------------------    661   ------------------
                             285  ------------------    688   ------------------
099    ------------------    288  ------------------    750   ------------------
101    ------------------    290  ------------------    787   ------------------
102    ------------------    294  ------------------    860   ------------------
110    ------------------    297  ------------------    869   ------G-----------
115    ------------------    298  ------------------    910   ------------------
118    ------------------    299  ------------------    962   ------------------
122    ------------------    301  ------------------    06P   ------------------
129    ------------------    309  ------------------
135    ------D-----------    326  ------------------    acute
140    ------------------    331  ------------------
143    ------------------    332  ------------------
147    ------------------    335  ------------------    7-42  ------------------
154    ------------------    346  ------------------    7-75  ------------------
216    -R----------------    362  ------D-----------    6-56  ------------------
218    ----Q-------------    387  ------------------    7-70  ------------------
219    ----Q-------------    394  ----Q-------------    3-43  ------------------
222    ----Q-------------    397  ------------------    6-23  ----Q-------------
223    ------------------    398  ------------------    6-40  ------------------
226    -E----------------    407  ----Q-------------    7-06  ------------------
227    ------------------    410  ----Q-------------    7-16  ------------------
229    ------------------    411  -A----------------    7-31  ------------------
230    ------------------    422  ------------------    7-41  ------------------
231    ------------------    425  ----Q-G-----------    8-10  ----Q-------------
233    ------------------    428  ------------------    822   ------G-----------
235    ------------------    437  L-----------------
236    ------------------    442  ------------------    EliSpot response to
246    ------------------    457  ------------------    PKARRSEGRSWAQPGYPW
259    ------------------    540  ------------------
```

B

```
       PVGGVARALAHGVRAL    262  ----------------    540   ----A-----------
110    -A--------------    266  ----------------    570   ----------------
128    ----------------    267  ----------------    579   ----------------
129    ----------------    274  ----------------    582   ----------------
216    ----------------    275  ----------------    589   ----------------
219    ----------------    278  ----------------    661   --------?-------
226    ----------------    285  ----------------    688   ----------------
227    ----------------    288  -A--------------    750   ----------------
235    ----------------    290  ----------------    787   ----------------
422    ----------------    294  ----------------    860   ----------------
437    ----------------    297  ----------------    869   ----------------
                           298  ----------------    910   ----------------
099    ----------------    299  ----------------    962   -A--------------
101    --------F-------    301  -A--------------    06P   ----------------
102    ----------------    309  ----------------
115    ----------------    326  --------F-------    acute
118    ----------------    331  ----------------
122    ----------------    332  --------F-------    7-42  ----------------
135    ----------------    335  ----------------    7-75  ----------------
140    ----------------    346  ----------------    6-56  ----------------
143    ----------------    362  ----------------    3-43  ----------------
147    ----------------    387  ----------------    6-23  -A--------------
154    ----------------    394  ----------------    6-40  -A--------------
218    ----------------    397  ----------------    7-06  ----------------
222    -A------F-------    398  ----------------    7-16  ----------------
223    ----------------    407  ----------------    7-31  ----------------
229    ----------------    410  ----------------    7-41  ----------------
230    ----------------    416  ----------------    7-70  ----------------
231    ----------------    425  ----------------    8-10  ----------------
234    ----------------    428  ----------------    822   ----------------
233    ----------------    442  -A--------------
236    ----------------    450  --------F-------    EliSpot responses to
246    ----------------    457  ----------------    PVGGVARALAHGVRAL
259    ----------------    411  ----------------
```

HCV core sequence analysis by PCR in 93 HCV gt3 infected patients from the Oxford cohort, for CD4+ T cell targets **(A)** core 66-83 (`PKARRSEGRSWAQPGYPW`) and **(B)** core 143-158 (`PVGGVARALAHGVRAL`), which were frequently targeted in chronically infected HCV gt3 infection. Sequences of chronic patients (n=80) are shown in columns 1 and 2, and acute patients (n=13) in column 3. The grey shading marks patients that responding to the according epitope in INFγ ELISpot assays.

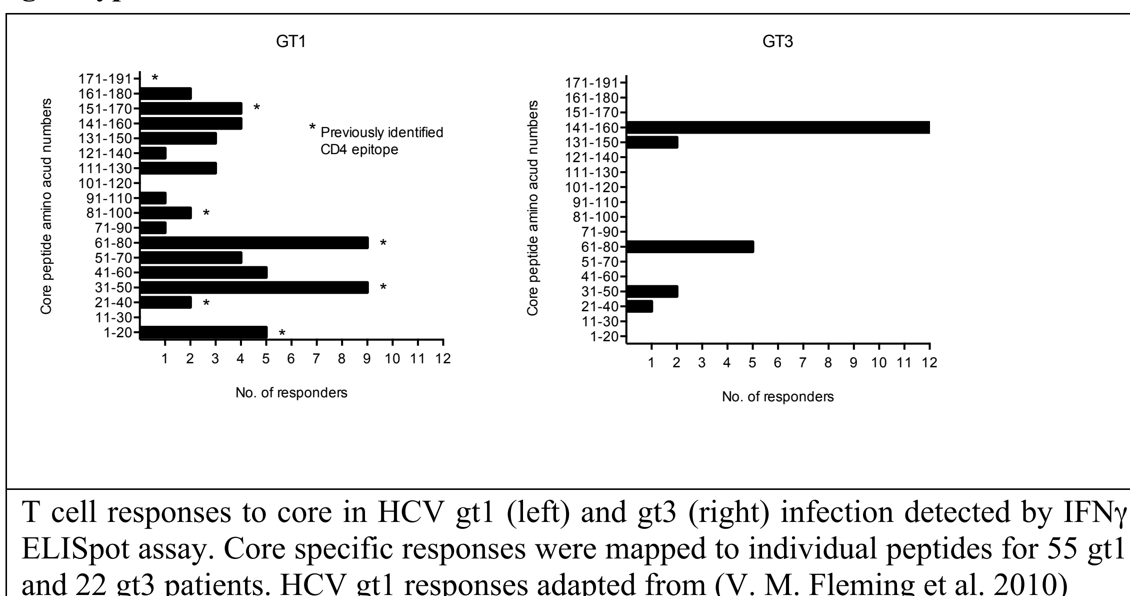
4.4.3 Comparison of CD4+ T cell specificity against HCV core in HCV gt1 and gt3 infection

HCV core is a comparatively conserved region within the HCV genome. To assess whether HCV gt3 and gt1 T cell specificity was distinct between HCV gt1 and gt3, HCV gt3 specific T cell responses described in this thesis were compared to previously published data in HCV gt1 infection (V. M. Fleming et al. 2010).

T cell specificity against HCV core in HCV gt1 and gt3 infection varies considerably (Figure 4-5): HCV gt1 core specific T cell responses are broad and target multiple peptides, whereas core specific T cell responses in HCV gt3 infection are more focused and frequently target peptides core 141-160 and 61-80.

T cell responses to core peptides that were frequently targeted in HCV gt3 infection have been previously described in HCV gt1 infection (Löhr et al. 1996; V. M. Fleming et al. 2010). In contrast, peptides frequently targeted in HCV gt1 infection were either rarely (core 31-59 (Lamonaca et al. 1999; A. J. MacDonald et al. 2002; V. M. Fleming et al. 2010)), or not at all detected in HCV gt3 infected patients (core 1-20 and 151-170 (A. J. MacDonald et al. 2002; V. M. Fleming et al. 2010)).

Figure 4-5: Comparison of CD4+ T cell responses to the core region in HCV genotype 1 and 3 infection.



T cell responses to core in HCV gt1 (left) and gt3 (right) infection detected by IFNγ ELISpot assay. Core specific responses were mapped to individual peptides for 55 gt1 and 22 gt3 patients. HCV gt1 responses adapted from (V. M. Fleming et al. 2010)

The distinct T cell specificity could be due to sequence differences between gt1 and 3 at targeted individual peptides. Indeed, a comparison of peptide sequences at frequently targeted gt1 and 3 peptides reveal differences at some targeted peptides (Table 4-2). However, even though no sequence differences were observed at amino

acid positions 31-50 and 41-60, no responses were detected in the HCV gt3 cohort. This could be due to a smaller sample size in the gt3 cohort, or competition at the presenting HLA type.

Table 4-2: HCV gt1 and 3 sequence comparison at commonly targeted core peptides.

AA positions	GT1 sequence	GT3 sequence	predominantely targed in GT
1 to 20	MSTNPKPQRK TKRNTNRRPQ	MST**L**PKPQRK TKRNT**I**RRPQ	1
31 to 50	VGGVYLLPRR GPRLGVRATR	VGGVY**V**LPRR GPRLGVRATR	1
41 to 60	GPRLGVRATR KTSERSQPRG	GPRLGVRATR KTSERSQPRG	1
61 to 80	RRQPIPKVRR PEGRTWAQPG	RRQPIPK**A**RR **S**EGR**S**WAQPG	1 and 3
141 to 160	GAPLGGAARA LAHGVRVLED	GAP**V**GG**V**ARA LAHGVR**A**LED	3

Sequence variations of HCV screening peptide sets for genotype 1 (used in Fleming et al. 2010) and GT3 (consensus sequence) are depicted. Peptides dominantly targeted in HCV gt3 infection are underlined. Sequence variation between HCV gt1 and gt3 peptides is marked in bold.

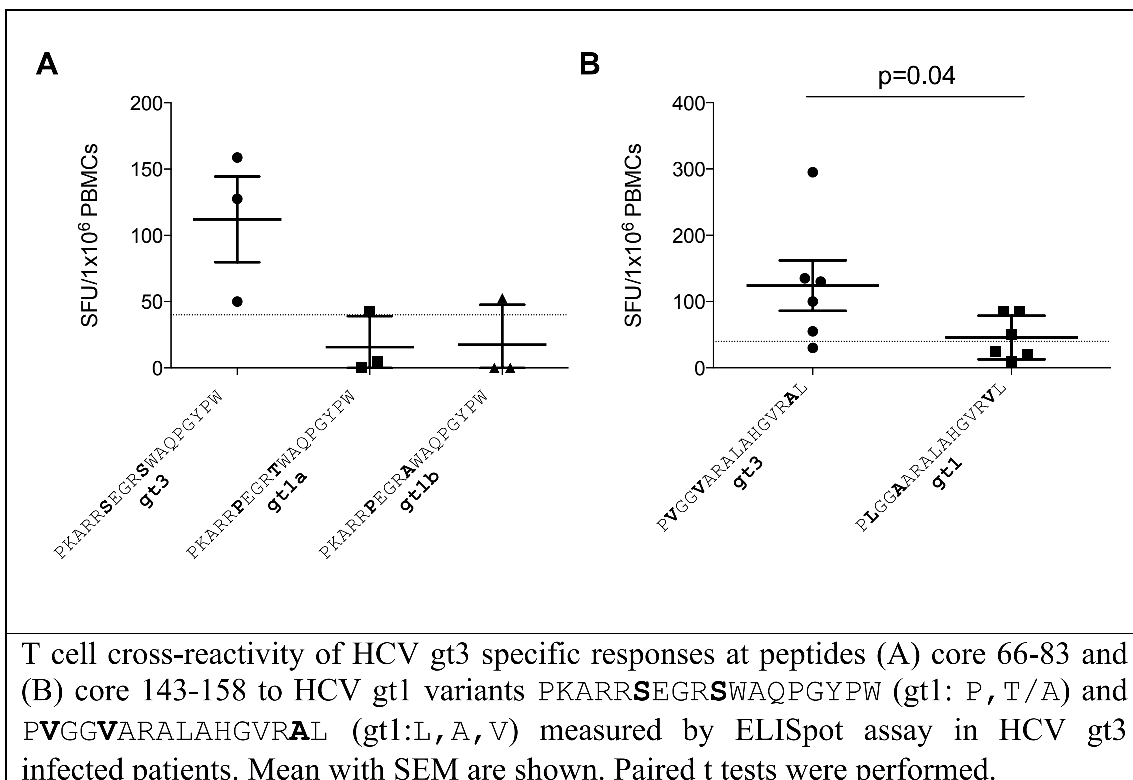
Of note, sites of inter- and intra-genotypic sequence variation were different at the frequently targeted gt3 peptide core 66-83 (PK**A**RRSEGR**S**WAQPGYPW vs. **PK**AR**R**S**E**GRSWAQPGYPW), whereas sites of variation within HCV gt3 and between HCV gt1 and gt3 were identical at HCV gt3 peptide core 143-158 (P**V**GG**V**ARALAHGVR**A**L).

4.4.4 Cross-reactivity of CD4+ T cells targeting HCV gt3 core peptides against HCV gt1 sequence variants

To assess whether CD4+ T cells primed in HCV gt3 infection targeting frequently detected HCV core T cell targets also target circulating HCV gt1 viral variants, T cell cross-reactivity against identified HCV gt3 and gt1 sequence variants was assessed by IFNγ ELISpot assays.

In HCV gt3 patients responding to the individual peptide core 66-83 (PK**A**RR**S**EGR**S**WAQPGYPW, gt3) in IFNγ ELISpot assays, reduced T cell responses were observed when stimulating with the gt1 variants PKARR**P**EGR**T**WAQPGYPW and PKARR**P**EGR**A**WAQPGYPW (Figure 4-6 A). Likewise, in patients targeting dominant T cell target core 143-158 (P**V**GG**V**ARALAHGVR**A**L), a significantly reduced T cell response against the gt1 peptide variant (P**L**GG**A**ARALAHGVR**V**L) was detected in IFNγ ELISpot assays (Figure 4-6B).

Figure 4-6: Experimental cross-reactivity against HCV gt3 and gt1 sequence variants at dominant CD4+ restricted HCV core T cell targets.



T cell cross-reactivity of HCV gt3 specific responses at peptides (A) core 66-83 and (B) core 143-158 to HCV gt1 variants PKARR**S**EGR**S**WAQPGYPW (gt1: P,T/A) and P**V**GG**V**ARALAHGVR**A**L (gt1:L,A,V) measured by ELISpot assay in HCV gt3 infected patients. Mean with SEM are shown. Paired t tests were performed.

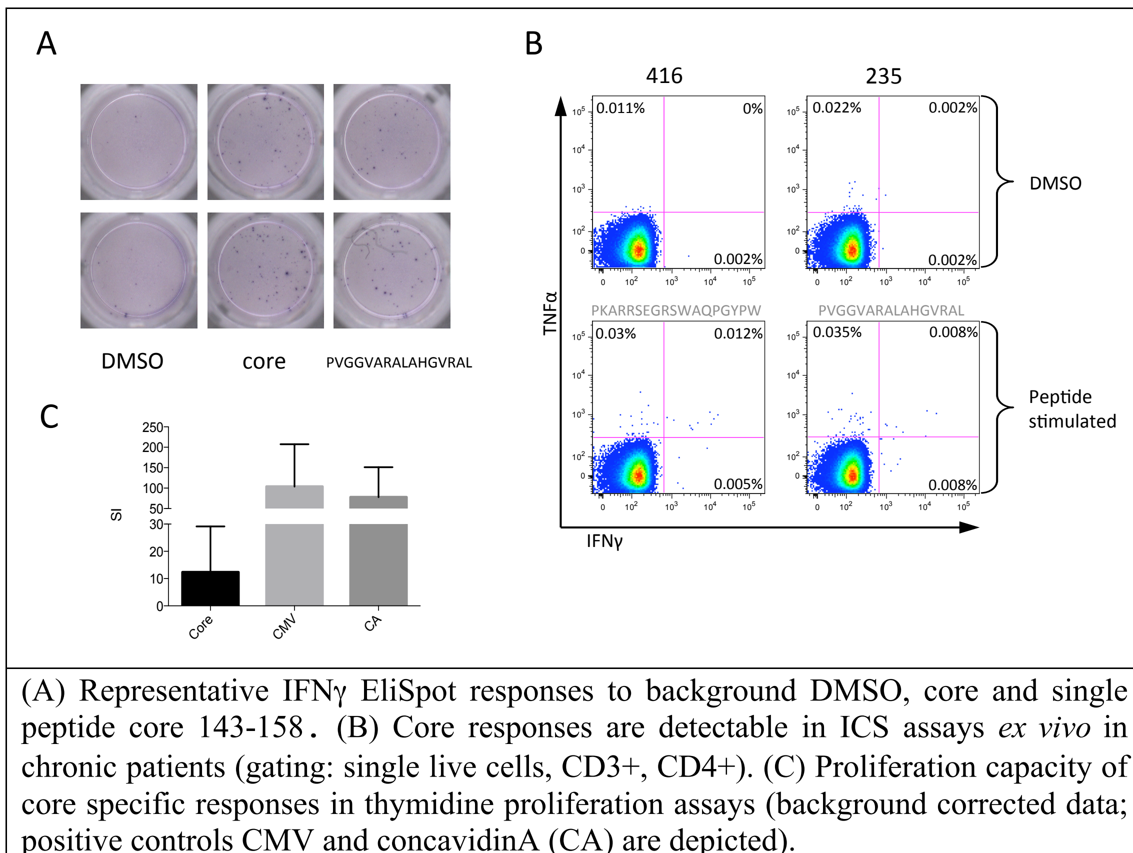
4.4.5 Cytokine profile of core specific CD4+ responses

4.4.5.1 Ex-vivo ICS responses – cytokine profile

Next, I analysed the functional profile of dominant core specific CD4+ responses in HCV gt3 infection, with the aim of further defining the role of these responses in HCV infection.

First, T cell responses were verified using IFNγ ELISpot assays and *ex vivo* intracellular cytokine staining (ICS) (Figure 4-7 A and B). Next, the proliferation capacity of HCV gt3 core specific T cell responses was assessed using ^{3}H incorporation assays as described previously (Missale et al. 1996) (Figure 4-7 C). Core specific responses showed only low proliferation capacity in comparison with CMV specific responses and the positive control concavidinA (n=12).

Figure 4-7: Verification of HCV gt3a specific core responses.

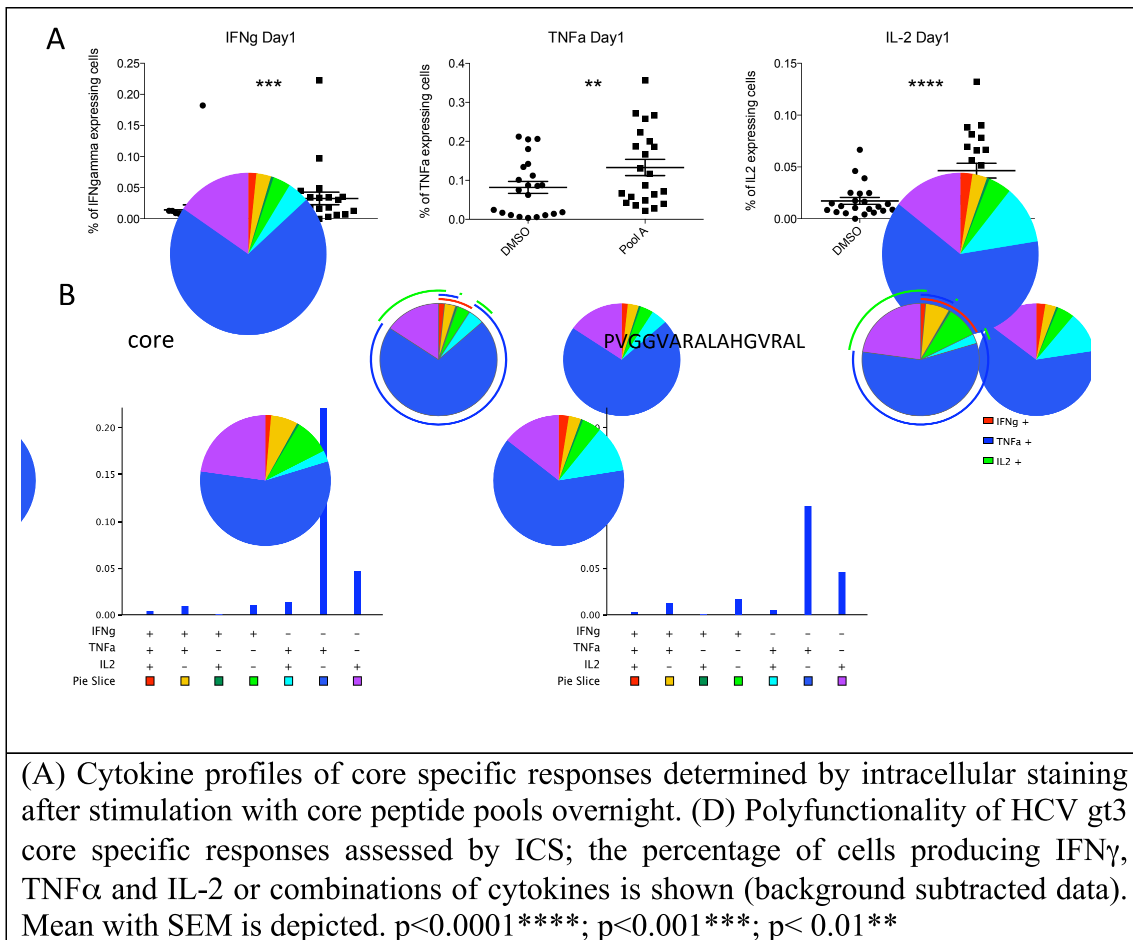


(A) Representative IFNγ EliSpot responses to background DMSO, core and single peptide core 143-158. (B) Core responses are detectable in ICS assays *ex vivo* in chronic patients (gating: single live cells, CD3+, CD4+). (C) Proliferation capacity of core specific responses in thymidine proliferation assays (background corrected data; positive controls CMV and concavidinA (CA) are depicted).

Subsequently, the functionality of the antiviral response was analysed using ICS for IFN-γ, TNFα (tumor necrosis factor–α), and IL-2 (interleukin-2). The cytokine profile

of core specific HCV genotype 3a responses was assessed in 22 chronically infected patients with confirmed ELISpot responses to core (Figure 4-8).

Figure 4-8: Cytokine profile of core specific HCV gt3a responses.



(A) Cytokine profiles of core specific responses determined by intracellular staining after stimulation with core peptide pools overnight. (D) Polyfunctionality of HCV gt3 core specific responses assessed by ICS; the percentage of cells producing IFNγ, TNFα and IL-2 or combinations of cytokines is shown (background subtracted data). Mean with SEM is depicted. $p<0.0001$****; $p<0.001$***; $p< 0.01$**

PBMC were stimulated overnight with core peptide pools and to peptide PVGGVARALAHGVRAL (AA position 143-158) in patients where a response to the single peptide was mapped. Intracellular staining assays were performed and cytokine secretion assessed using flow cytometry. After stimulation with core, T cells secreted significant amounts of IFNγ, TNFα and IL-2 in comparison to un-stimulated cells (Figure 4-8 A). Polyfunctionality of cytokine producing T cells was calculated (SPICE software, (Roederer et al. 2011); the majority of core specific T cells are monofunctional producing only one single cytokine, with TNFα being most abundant, followed by IL-2 and IFNγ (Figure 4-8 B). Analysed patients with a confirmed response to peptide core 143-158 (PVGGVARALAHGVRAL) showed a comparative phenotypical profile to core specific CD4+ T cell responses in ICS assays.

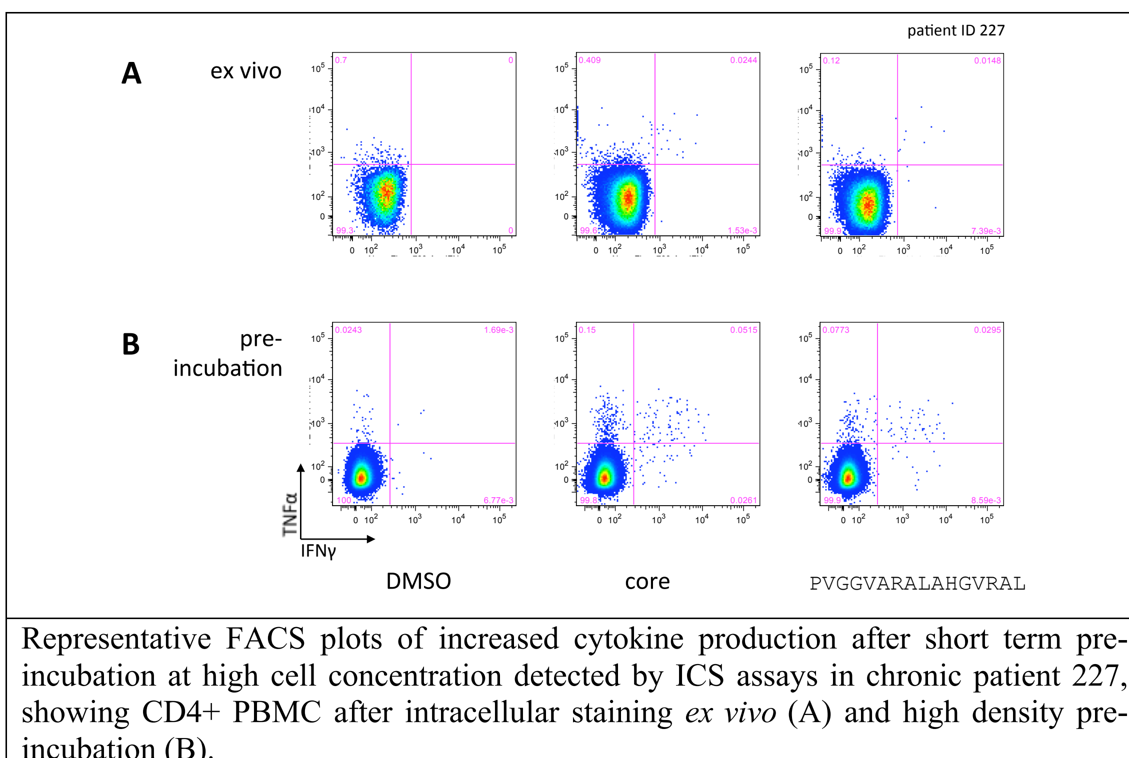
4.4.5.2 Short-term incubation of core specific CD4+ T cells at high cell concentrations to assessed cytokine production

HCV gt3 specific CD4+ T cell responses in chronic infection were often difficult to detect. Even in ELISpot assays, a very sensitive method, most responses range at the detection threshold, making the detection of responses and the mapping to single peptides challenging. Therefore, T cell responses to frequently detected HCV gt3 core peptides were often not detectable in ICS assays, a technique that was used for phenotypic characterization without a tetramer available to specifically mark and enrich cells at low frequency.

In a recent publication, it has been shown that pre-incubation of PBMC at high (over $1x10^6$cells/ml), but not at low concentrations increased the amount of cytokines released upon stimulation (Römer et al. 2011). The discovery was related to a clinical trial, where human volunteers received TGN1412, a humanized CD28-specific monoclonal antibody, and subsequently experienced a life-threatening cytokine release syndrome (Suntharalingam et al. 2006). Preclinical tests using human PBMCs had failed to announce the rapid release of toxic cytokines in response to this CD28 agonist (C.-H. Lin & Hünig 2003; Beyersdorf et al. 2008; Eastwood et al. 2010). However, a coincidental discovery led to a surprising observations: when cells were pre-incubated at cell high densities for 48 hours before performing the experiments, high cytokine releases were detected. This increase in reactivity is thought to be mediated through increased cell-cell contact, with T cells acquiring tissue-like properties during high-density culture. Moreover, other reports demonstrated a change of functional signatures in antigen specific T cell responses after overnight resting of PBMC (Kutscher et al. 2013)

Based on this evidence, high-density pre-culture at high cell concentrations was tested in HCV gt3 patients with a confirmed core specific T cell response, before assessing CD4+ T cells in ICS assays. T cells were incubated with HCV core peptides either *ex vivo*, or after pre-incubation at high cell concentrations. Representative FACS plots of HCV gt3 specific responses against gt3 core peptides *ex vivo* and after short-term incubation at high densities are shown in Figure 4-9.

Figure 4-9: Representative FACS plots for T cell responses ex-vivo and after short-term incubation.

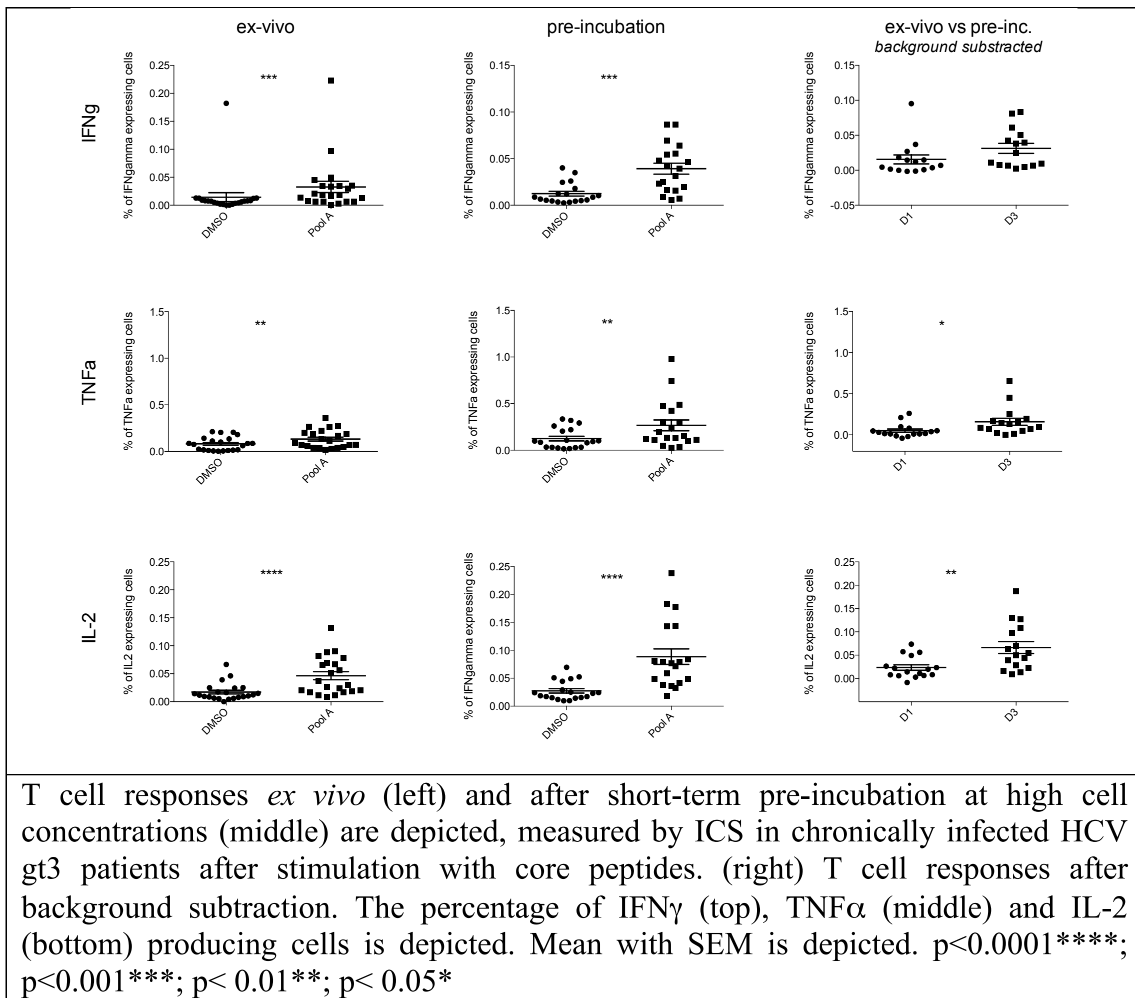


Representative FACS plots of increased cytokine production after short term pre-incubation at high cell concentration detected by ICS assays in chronic patient 227, showing CD4+ PBMC after intracellular staining *ex vivo* (A) and high density pre-incubation (B).

A significant increase of cytokine release in HCV gt3a core peptide stimulated samples after pre-incubation of T cells compared to *ex vivo* samples was observed (Figure 4-10). Background cytokine production was increased in samples after pre-incubation. However, when comparing IFNγ, TNFα and IL-2 release after background subtraction for T cell responses *ex vivo* and after high concentration pre-incubation, a significant increase in cytokine production was observed for TNFα and IL-2 in pre-incubated samples (Figure 4-10). The increase in cytokine production after pre-incubation at high cell concentrations was CD4+ specific; no equivalent effect was observed when testing HCV CD8+ T cells (data not shown).

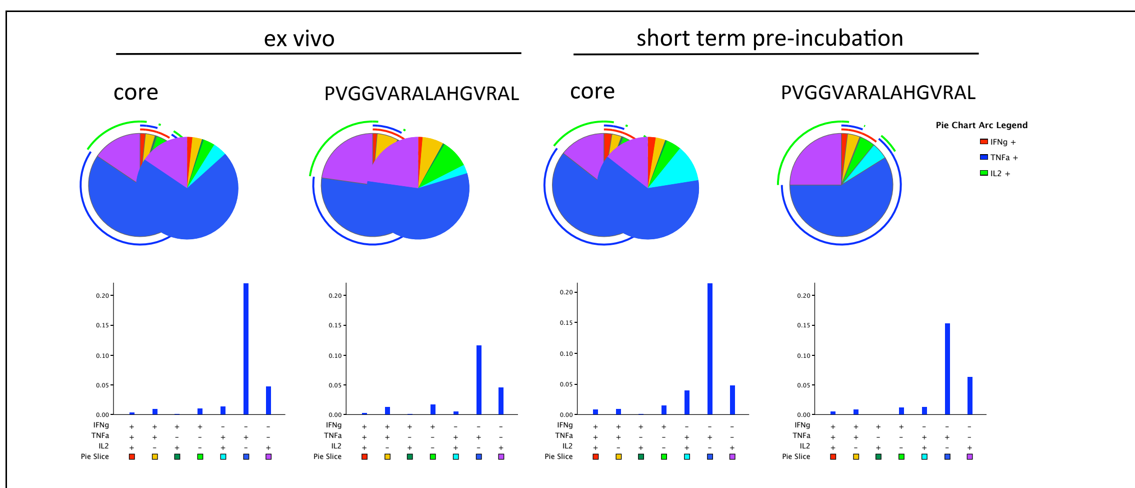
To assess whether core specific T cells after pre-incubation at high concentrations alter their phenotypic profile, polyfunctionality of CD4+ core specific T cells *ex vivo* and after pre-incubation was assessed. T cell responses were evaluated after stimulation with HCV gt3a core peptide pools and frequently detected peptide core 143-158 in patients responding core or the peptide in IFNγ ELISpot assays. No alteration in T cell polyfunctionality was observed after pre-incubation of T cells, with core specific T cells producing only single cytokines (mainly TNFα and IL-2) *ex vivo* and after pre-incubation (Figure 4-11).

Figure 4-10: Cytokine release of core specific T cells after short-term incubation.



T cell responses *ex vivo* (left) and after short-term pre-incubation at high cell concentrations (middle) are depicted, measured by ICS in chronically infected HCV gt3 patients after stimulation with core peptides. (right) T cell responses after background subtraction. The percentage of IFNγ (top), TNFα (middle) and IL-2 (bottom) producing cells is depicted. Mean with SEM is depicted. p<0.0001****; p<0.001***; p< 0.01**; p< 0.05*

Figure 4-11: Comparison of polyfunctionality of HCV gt3 core specific T cells after short-term incubation.



Cytokine profile of core and core 143-158 (PVGGVARALAHGVRAL) specific responses assessed *ex vivo* and after short-term incubation at high cell concentrations by ICS assays in chronically HCV gt3 infected patients.

Further work on dominant low level core specific CD4+ epitopes may be facilitated by the use of MHC class-II tetramers and include the assessment of cytokines and phenotypic markers related to regulatory and memory T cells. However, when analysing the class-II HLA types in patients targeting dominant CD4+ epitopes, no common HLA types were observed (data not shown), most likely due to the promiscuous nature of peptide binding properties of MHC class-II complexes with an identical peptide presented by distinct MHC class-II types.

In summary, sequence polymorphisms were rarely observed at two CD4+ T cell targets frequently detected in the HCV gt3 core region. However, even though HCV core is a comparably conserved viral region within the HCV genome, major differences in T cell specificity were observed between HCV genotypes 3 and 1. In addition, T cell responses primed in natural HCV gt3 infection showed limited cross-reactivity against HCV gt3 and gt1 sequence variants in IFNγ ELISpot assays.

4.5 Frequently detected CD8+ restricted T cell responses targeting HCV genotype 3 non-structural regions

Of seven frequently detected T cell targets defined in this study, five were CD8+ restricted and targeted HCV non-structural viral proteins (Table 4-3).

Table 4-3: Frequently detected CD8+ restricted T cell targets in HCV gt3 non-structural regions.

Frequently detected HCV gt3 specific T cell responses						Patients responding			
Viral region	**AA position**	**Sequence**	**HLA**	**Pept. set**	**Abbrev.**	**S**	**C**	**A**	**Total**
NS2	886-896	LLYPSLIFDI	A02	HLA	$NS2_{886}$		2	1AC	**3**
		LYPSLIFDI	A24				3	1AC	**4**
NS3	1443-1451	ATDALMTGY *	A01	HLA	$NS3_{1443}$		3	1TxS	**4**
	1520-1537	RPSGMFDSVVLCECYDAGC SWYDL		OPs	$NS3_{1520}$	2	12		**14**
NS4b	1853-1862	RVLLDILAGY	A26	HLA	$NS4b_{1853}$		3	1TxS 1SR	**5**
NS5a	2126-2141	AEFFTEVDGVRLHRYA		OPS	$NS5a_{2126}$	2		2TxS	**4**

CD8+ T cell responses frequently detected in HCV gt3 infection are depicted. For each targeted individual peptide, the amino acid (AA) position, peptide sequence, restricting HLA type and detecting peptide sets are specified. The number of patients responding to the peptide and their status of infection (T: total, C: chronic, A: acute [AC: acute proceeding to chronic, TxS: treated achieving SVR, SR spontaneously resolved], S: spontaneously resolved) is detailed. An epitope previously described in HCV gt1 infection is marked with a star.

Using HLA predicted peptides, three dominant CD8+ restricted T cell responses to HCV gt3 non-structural regions were identified (Table 4-3). Responses were located in the HCV viral regions NS2, NS3 and NS4b, were restricted by HLA type A*02/A*24, A*01 and A*26 and recognized by 3 (A*02)/ 4 (A*24), 3 (A*01) and 5 (A*26) HCV gt3a patients, respectively. The majority of patients recognizing HLA predicted peptides were chronically infected individuals.

Using overlapping peptides, two dominant T cell targets were defined in the Oxford HCV gt3 cohort (Table 4-3). The response to the HCV NS3 peptide RPSGMFDSVVLCECYDAGCSWYDL was mainly observed in chronically infected patients, and found in 14 out of 29 (48%) mapped responses to the HCV NS3 region (for details, see 3.8.2 Individual peptides targeted by T cells in HCV gt3 infection, page 113). Of 14 patients targeting this individual peptide, two patients spontaneously resolved infection, whereas 12 were chronically infected. A second T cell target located in the NS5a region (AEFFTEVDGVRLHRYA) was detected in two spontaneously resolved patients and two acutely infected patients who cleared the

infection under therapy. Both CD8+ restricted responses detected by overlapping peptides were of particular interest, since they were detected in patients spontaneously clearing HCV infection and therefore may serve as an interesting vaccine target, possibly in combination with other epitopes.

Similarly to the analysis of HCV gt3 core specific CD4+ responses, sequence polymorphisms at dominant CD8+ restricted T cell targets were analysed, both within HCV genotype 3 and across HCV genotypes. This was followed by an experimental assessment of T cell cross-reactivity against identified sequence variants in IFNγ ELISpot assays. Subsequently, T cell phenotype and function of CD8+ restricted T cell responses was evaluated.

Abbreviations as stated in Table 4-3 will be used interchangeably with peptide sequences: LLYPSLIFDI ($NS2_{886}$), ATDALMTGY ($NS3_{1443}$) and RVLLDILAGY ($NS4b_{1853}$) for T cell targets detected by HLA predicted peptides, and RPSGMFDSVVLCECYDAGCSWYDL ($NS3_{1520}$) and AEFFTEVDGVRLHRYA ($NS5a_{2126}$) for T cell targets detected by overlapping peptides.

4.5.1 Sequence variability at dominant CD8+ restricted T cell targets

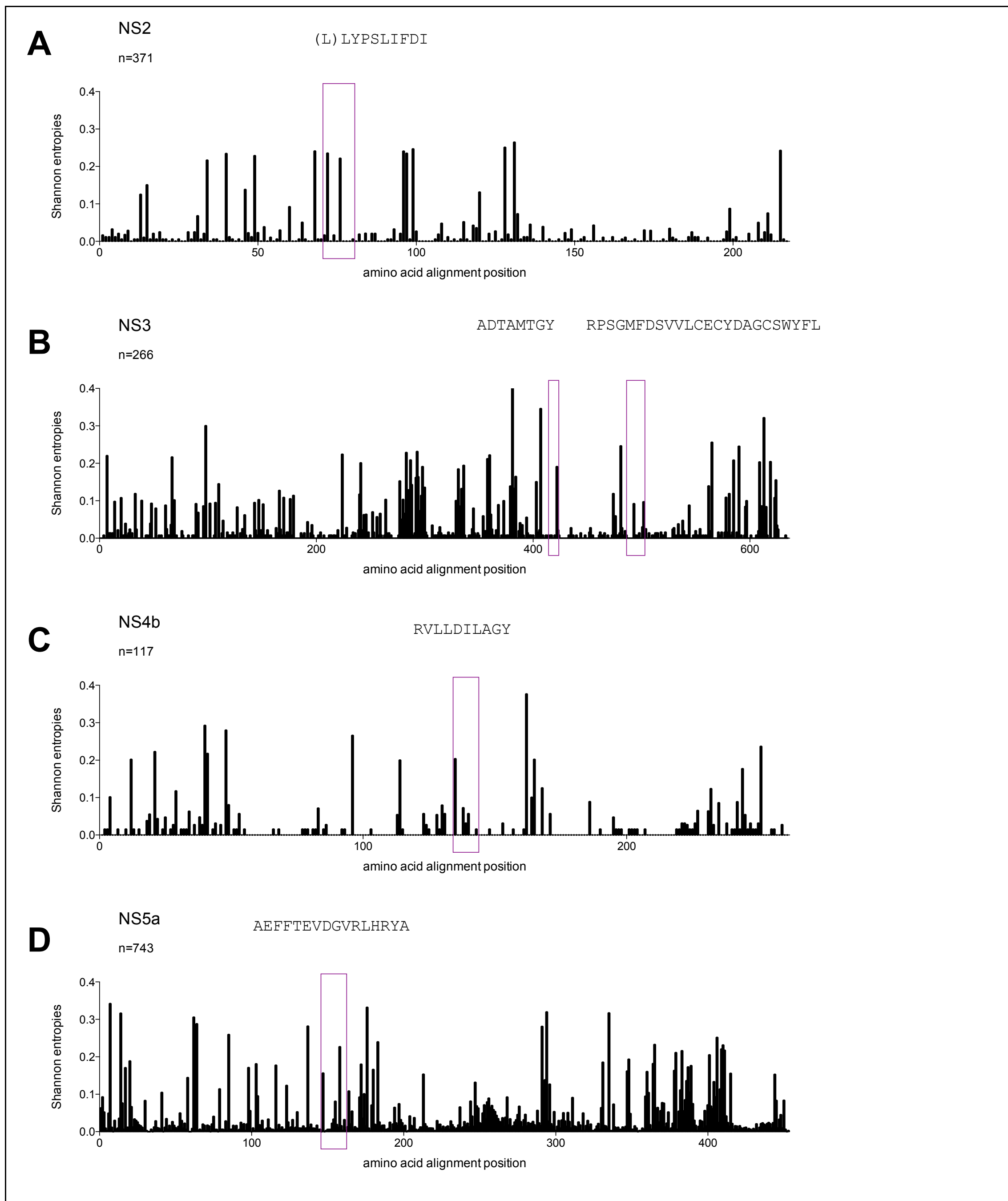
4.5.1.1 HCV gt3 sequence variability at dominant CD8+ T cell targets on population level

Initially, sequences of CD8+ T cell targets frequently detected in HCV gt3 infection were evaluated at population level. For this analysis, HCV gt3 sequences for regions containing dominantly frequently targeted peptides were obtained from the Los Alamos sequence depository and sequence heterogeneity was calculated using the Shannon entropy calculation (see 2.2.8, page 71) (Figure 4-12). Variability was assessed by comparing the mean entropy at each T cell target to mean entropy score of the analysed HCV region (NS2, NS3, NS4b and NS5, respectively), and subsequently defining the T cell target as “ conserved” or “variable”.

High sequence heterogeneity was observed at three dominant epitopes detected by HLA predicted peptides. For epitope $NS2_{886}$ restricted by HLA types A*02 and A*24 ((L)LYPSLIFDI, Figure 4-12 A), two dominant sequence polymorphisms were noted on population level: at position 3, amino acid tyrosine was replaced with histidine (Y→H), and at position 7, isoleucine was replaced with valine (I→V). With the amino acid substitution at position 3, binding properties are expected to change

considerably from a hydrophobic (Y) to a positively charged side chain. This polymorphism was also linked to HLA types A*02 and A*24 in the original publication linked HCV gt3 polymorphisms to HLA types (Rauch et al. 2009).

Figure 4-12: Sequence heterogeneity on population level at frequently detected CD8+ restricted T cell targets within HCV non-structural regions.



The sequence heterogeneity (Shannon entropy) is depicted for HCV viral regions where dominant CD8+ restricted epitopes were detected. Shannon entropy calculated for all HCV gt3 sequences available on the Los Alamos sequence depository for the according region: (A) NS2, (B) NS3, (C) NS4b and (D) NS5a. In each panel, the number of analysed sequences is depicted. Frequently detected T cell targets are marked with a box, with the sequence of each targeted peptide given.

At epitope $NS3_{1443}$ (ATDALMTGY), a single polymorphisms at position 9 in the otherwise conserved A*01 restricted epitope was observed, both within the Oxford gt3 cohort (data not shown) and on population level (Figure 4-12 B). Epitope ATDALMTGY is a well-known epitope in HCV gt1 and gt3 infection (A. L. Cox, Mosbruger, Lauer, et al. 2005; Neumann-Haefelin et al. 2008; Giugliano et al. 2009; Barnes et al. 2012). The sequence polymorphisms at position 9 (F) has been previously identified as immune mediated escape and emerges as consensus (negatope), consistent with the fact that HLA-A*0101 is a relatively high-frequency HLA allele (Neumann-Haefelin et al. 2008). This suggests that the escape mutation is of similar or improved replicative capacity in comparison to the wild type. In this cohort, a response to the variant sequence ATDALMTG**F** has been detected in all three chronically HCV gt3 infected patients mounting a response to the wild type (see Figure 3-14, page 106).

At epitope $NS4b_{1853}$ (RVLLDILAGY) (Figure 4-12 C), a common amino acid substitution from arginine to lysine [R→K] at position 1 was observed, a polymorphism previously linked to HLA type A*26 (Rauch et al. 2009).

Two CD8+ restricted T cell responses were frequently detected using overlapping peptide pools. Comparably low sequence heterogeneity was observed at targeted peptide $NS3_{1520}$ (Figure 4-12 B). However, an accumulation of non-synonymous changes was observed on base pair level in the Oxford cohort (Supplementary Figure 8-3 A, page 290); however, these changes were not reflected on amino acid level, suggesting either functional or structural constraints at this epitope (Supplementary Figure 8-3 B, page 290). When analysing the second T cell target frequently detected by overlapping peptides, $NS5a_{2126}$, sequence heterogeneity assessed by Shannon entropy showed two common polymorphisms within HCV gt3 at position 1 and 13 (**A**EFFTEVDGVR**L**HRYA/P,I) (Figure 4-12 D).

Overall, sequence polymorphisms were commonly observed at three epitopes frequently detected using the HLA predicted peptide set and at T cell target $NS5a_{2126}$ identified using overlapping peptide pools, whereas T cell target $NS3_{1520}$ was comparably conserved within HCV gt3. Since the HLA predicted peptide screening set was derived from sequence polymorphisms associated with HLA types (Rauch et al. 2009), common polymorphisms were to be expected at these epitopes within the Oxford HCV gt3 cohort.

4.5.1.2 Sequence variability at CD8+ T cell targets frequently detected within the Oxford HCV gt3 cohort, in context of T cell responses detected by IFNγ ELISpot and patients' HLA type

Subsequently, sequence polymorphisms at CD8+ epitopes frequently detected were evaluated within the Oxford HCV gt3 cohort (Figure 4-13 to Figure 4-16). Sequence information was obtained using capillary Sanger sequencing, and analysed regarding detected T cell responses in IFNγ ELISpot assays and patients' HLA type.

I initially compared whether sequence variants were more common in patients mounting an IFNγ ELISpot response than in those not responding to the peptide of interest (left side of Figure 4-13 to Figure 4-15). For epitopes detected by HLA predicted peptides, I subsequently assessed whether sequence variants were more common in patients with the restricting HLA type than in patients with other HLA types (right side of Figure 4-13 to Figure 4-15)

Epitope $NS2_{886}$ ((L)L**Y**PSLIFDI) was detected in five patients and restricted by HLA types A*02 and A*24. Of 42 HLA A*02 positive patients tested in IFNγ ELISpot assays, a positive T cell response was detected in three patients (7%) (Figure 4-13, left panel). In comparison, $NS2_{886}$ was more commonly targeted in A*24 positive patients; IFNγ secretion was detected in four out of ten A*26 positive patients (40%). In patients with confirmed T cell response, polymorphisms within the targeted epitope were observed in two out of five sequenced patients (40%).

Next, sequence polymorphisms were additionally analysed regarding patients' HLA types (Figure 4-13, right panel). Polymorphisms did not seem to cluster in A*02 and A*24 positive patients. In 12.5% (2/16) of A*02 positive patients polymorphisms were observed, compared to 33% (6/18) A*02 negative patients (p=0.2327, Fisher's exact). When comparing A*24 positive to A*24 negative patients, a higher percentage of A*24 positive individuals had sequence polymorphisms (3/7; 43%) compared to 18.5% (5/27) of A*24 negative individuals, however, this difference was not statistically significant (p=0.3153, Fisher's exact).

Within the Oxford gt3 cohort, only a single patient showed the escape mutation commonly observed at population level and linked to HLA A*02 and A*24 ((L)L**Y**PSLIFDI/**H**), however, this patients was A*02/A*24 negative and therefore not tested for the HLA predicted peptide. A mutation at position 7, that was common at population level (I→V), was observed in four patients in the Oxford cohort, but did

not seem to be more common in patients with HLA types A*02 and A*24, suggesting that this mutation is not caused by HLA restricted immune pressure.

Figure 4-13: Sequence information at dominant epitope NS2$_{886}$ (L)LYPSLIFDI in HCV gt3 patients with and without a T cell response.

EliSpot results

```
A02+                          A24+
.............................................................
responders                  : responders
                            :
      LLYPSLIFDI            :       LYPSLIFDI
6-23  ------V-EV            : 6-23  -----V-EV
101   ----------            : 101   ---------
299   ----------            : 219   ---------
                            : 362   --S------
.............................................................
non-responders              : non-responders
102   ----------            : 104   ---------
111   ----------            : 140   -----V---
118   ----------            : 295   Not seq
128   ----------            : 436   Not seq
135   ----------            : 457   Not seq
147   ----------            : 10256 Not seq
223   ----------            :
227   ----------            :
234   ------V---            :
259   ----------            :
274   ----------            :
278   ----------            :
103   Not seq               :
114   Not seq               :
216   Not seq               :
226   Not seq               :
246   Not seq               :
285   Not seq               :
394   Not seq               :
416   Not seq               :
422   Not seq               :
428   Not seq               :
436   Not seq               :
437   Not seq               :
457   Not seq               :
582   Not seq               :
720   Not seq               :
771   Not seq               :
787   Not seq               :
900   Not seq               :
982   Not seq               :
7-42  Not seq               :
50017Not seq                :
70230Not seq                :
7-41  Not seq               :
6-23  Not seq               :
6-56  Not seq               :
7-75  Not seq               :
7-06  Not seq               :
```

Patient HLA types

```
A02+                  A24+                  A02/A24-

      LLYPSLIFDI            LLYPSLIFDI            LLYPSLIFDI
6-23  ------V-EV      6-23  ------V-EV      110   ----------
101   ----------      101   ----------      115   ----T-----
102   ----------      104   ----------      129   ----------
111   ----------      110   ----------      143   ----------
118   ----------      140   ------V---      229   ----------
122   ----------      219   ----------      235   ------V---
128   ----------      362   ---S------      236   --H-------
135   ----------                            237   ----------
147   ----------                            275   ----------
223   ----------      A24-                  288   ----------
227   ----------                            294   ----------
234   ------V---      102   ----------      297   ----------
259   ----------      111   ----------      298   -F--------
274   ----------      115   ----T-----      301   ----------
278   ----------      118   ----------
299   ----------      122   ----------
                      128   ----------      not typed
                      129   ----------
A02-                  135   ----------      099   ----------
                      143   ----------      222   ----------
104   ----------      147   ----------      228   ----------
110   ----------      223   ----------      230   ----------
115   ----T-----      227   ----------      233   ----------
129   ----------      229   ----------      290   ----------
140   ------V---      234   ------V---      331   ----------
143   ----------      235   ------V---      387   ----------
219   ----------      236   --H-------
229   ----------      237   ----------
235   ------V---      259   ----------
236   --H-------      274   ----------
237   ----------      275   ----------
275   ----------      278   ----------
288   ----------      288   ----------
294   ----------      294   ----------
297   ----------      297   ----------
298   -F--------      298   -F--------
301   ----------      299   ----------
362   ---S------      301   ----------
```

In depth analysis of sequences at epitope NS2$_{886}$ (L)LYPSLIFDI. Left: ELISpot responses to NS2$_{886}$ were detected in 3/42 A*02 and 4/10 A*24 positive patients. Patients responding to NS2$_{886}$ in IFNγ ELISpot assays are marked in grey. Right: Patient sequences were grouped after patient HLA types A*02 and A*24. Not seq: sequence not determined.

T cell responses against epitope NS3$_{1443}$ (ATDALMTGY) were detected in 4 out of 26 HLA A*01 positive patients tested in IFNγ ELISpot assays (15%) (Figure 4-14, left). Common sequence polymorphisms at position 9 were observed in both patients with and without a detected T cell response. The sequence polymorphisms at position 9 (Y→ F) has been previously identified as immune mediated escape in chronically HCV gt1 infected patients (Neumann-Haefelin et al. 2008). Of patients with a detected T cell response in IFNγ ELISpot assays, three chronic patients carried the

escape version ATDALMTG**F**, whereas the acutely infected patient 822 did not carry the escape variant at the first analysed time point.

When analysing sequence polymorphisms in HLA A*01 positive and negative patients across the Oxford gt3 cohort, immune mediated escape was more common in A*01 positive patients (9/11, 82%) than in HLA A*01 negative patients (23/34, 68%), however, this difference was not statistically significant (p=0.4666, Fisher's exact, Figure 4-14, right panel).

Figure 4-14: Sequence information at dominant epitope NS3$_{1443}$ ATDALMTGY/F (A*01) in HCV gt3 patients with and without a T cell response.

```
ELISpot results

A01+

responders
-------------------------
        ATDALMTGY
822     ---------
416     --------F
7-70    --------F
06P     --------F

non-responders
-------------------------
111     --------F
229     --------F
274     --------F
297     --------F
299     ---------
7-31    --------F
7-70    --------F
06P     --------F
114     not seq
118     not seq
123     not seq
259     not seq
291     not seq
332     not seq
410     not seq
442     not seq
6-40    not seq
7-42    not seq
7-06    not seq
70351   not seq
103     not seq
871     not seq

109     not tested
806     not tested
```

```
Patient HLA types

A01+

        ATDALMTGY
822     ---------
111     --------F
229     --------F
274     --------F
297     --------F
299     ---------
416     --------F
7-31    --------F
7-70    --------F
06P     --------F
7-70    --------F

not HLA typed

099     --------F
218     ---------
222     --------F
228     --------F
230     --------F
233     --------F
290     --------F
331     ---------
387     --------F
```

```
A01-

        ATDALMTGY
7-41    ---------
7-16    --------F
7-75    --------F
6-23    --------F
101     --------F
102     ---------
104     ---------
110     --------F
115     ---------
122     --------F
128     --------F
129     --------F
135     ---------
140     --------F
143     --------F
147     --------F
154     --------F
219     ---------
223     --------F
226     --------F
227     ---------
234     ---------
235     --------F
236     --------F
275     ---------
278     --------F
288     ---------
294     ---------
295     --------F
297     --------F
298     --------F
301     --------F
362     --------F
6-56    --------F
```

In depth analysis of sequences at epitope NS3$_{1443}$ ATDALMTG**Y/**F. Left: ELISpot responses to NS3$_{1443}$ were detected in 4/26 A*01 positive patients. Patients responding to NS3$_{1443}$ in IFNγ ELISpot assays are marked in grey. Right: Patient sequences were grouped after patients' HLA type A*01. Not seq: sequence information not determined.

T cell responses against NS4b epitope RVLLDILAGY were detected in five out of eight HLA A*26 positive patients tested in IFNγ ELISpot assays (62.5%) (Figure 4-15, left). When comparing sequence variants in HLA A*26 positive patients versus HLA A*26 negative patients across the Oxford HCV gt3 cohort (Figure 4-15, right

panel), a significant accumulation of polymorphisms in HLA A*26 positive patients was observed (p=0.0166, Fisher's exact), with sequence variation observed in all A*26 positive patients (4/4/, 100%), compared to 6/21 (29%) of A*26 negative patients. The most common amino acid substitution observed was a R→K substitution at position 1, which was also commonly seen on population level, and initially linked to HLA type A26 in Rauch et al.'s study (Rauch et al. 2009).

Figure 4-15: Sequence information at dominant epitope $NS4b_{1853}$ RVLLDILAGY (A*26) in HCV gt3 patients with and without a T cell response.

```
ELISpot results                 | Patient HLA types
                                |
A26+                            | A26+                 A26-
                                |
responders                      | 6-56 K---------      102  ----------
--------------------------------| 110  K---------      128  ----------
      RVLLDILAGY                | 115  K---------      129  K---------
6-56  K---------                | 288  ---I------      135  ----------
110   K---------                |                      140  ----------
50017 not seq                   |                      147  ----------
332   not seq                   |                      219  ----------
437   not seq                   |                      227  ---I-V----
                                |                      229  K---------
non-responders                  |                      234  K---------
--------------------------------|                      235  ----------
115   K---------                |                      259  K---------
288   ---I------                | not typed            274  ----------
236   not seq                   |                      275  ----------
                                | 099  K---------      278  ----------
                                | 230  ----------      294  ----------
                                | 237  ----------      297  ----------
                                | 290  K---------      298  ----------
                                | 331  ----------      299  ----------
                                | 387 ----------       301  K---------
                                |                      362  ----------
```

In depth analysis of sequences at epitope $NS4b_{1853}$ **R**VLLDILAGY. Left: ELISpot responses to $NS4b_{1853}$ were detected in 5/8 tested A*26 positive patients. Patients responding to $NS4b_{1853}$ in IFNγ ELISpot assays are marked in grey. Right: Patient sequences were grouped after patients' HLA type A*26. not seq: sequence not determined.

In summary, sequence polymorphisms at dominant epitopes detected by HLA predicted peptides were common. Sequence polymorphisms clustered in patients carrying the associated HLA type at epitopes $NS4b_{1853}$ and $NS3_{1443}$. For targeted epitope $NS2_{886}$, polymorphisms did not seem to cluster in patients carrying HLA types A*02 and A*24 that were associated with sequence polymorphisms in the original HCV gt3 cohort (Rauch et al. 2009). However, it is possible that polymorphisms were missed due to an insufficient number of sequences obtained; a clearer picture might have been achieved with a fully sequenced cohort. The percentage of HLA positive patients mounting detectable T cell responses to the

according epitope was variable: the majority of A*26 positive patients recognized peptide NS4b$_{1853}$ (62.5%), whereas 15% of A*01 positive patients recognized epitope NS3$_{1443}$, and 40% of A*24 but only 7% of A*02 positive patients mounted a T cell response to epitope NS2$_{886}$, respectively. T cell responses may have been missed in other patients positive for the restricting HLA type, since responses in chronically infected patients are typically very low and may have been below detection threshold.

Next, sequence polymorphisms at CD8+ T cell targets frequently detected by overlapping peptides were evaluated (NS3$_{1520}$ (`RPSGMFDSVVLCECYDAGCSWYDL`) and NS5a$_{2126}$ (`AEFFTEVDGVRLHRYA`)) (Figure 4-16).

Sequences of T cell target NS3$_{1520}$ were obtained in 43 patients (chronic n=36, acute n=7, Figure 4-16 A). Only rare sequence polymorphisms were observed, which did not accumulate in patients with a detectable IFNγ ELISpot to this epitope (1/7 (14.3%) *vs* 4/43 (9.3%), p=0.5591, Fisher's exact).

Unfortunately, it was not possible to obtain any sequences of patients who responded to peptide NS5a$_{2126}$ in IFNγ ELISpot assays: two responders were spontaneously resolved patients without detectable RNA, and sequencing was not possible in two acutely infected patients clearing the infection under treatment. In the remaining patients without a detectable IFNγ ELISpot response, polymorphisms were common at positions 1 (alanine→ proline) and 13 (leucine to isoleucine) (Figure 4-16), both also observed on population level (Figure 4-12, page 171).

Overall, sequence polymorphisms within HCV gt3 were commonly observed at most frequently detected CD8+ T cell targets. This data suggests that sequence variation at frequently detected T cell targets is structurally and functionally tolerated and therefore circulate within the infected host.

Figure 4-16: Sequence information at frequently detected T cell targets NS3$_{1520}$ (`RPSGMFDSVVLCECYDAGCSWYDL`) and NS5a$_{2126}$ (`AEFFTEVDGVRLHRYA`) in HCV gt3 patients with and without a T cell response.

A

```
      RPSGMFDSVVLCECYDAGCSWYD       228   -----------------------
111   -----------------------       229   -----------------------
140   -----------------------       230   -----------------------
154   -----------------------       233   -----------------------
219   -----------------------       236   -------------------A---
235   -------------------A---       274   -----------------------
295   -----------------------       275   -----------------------
362   -----------------------       278   -----------------------
                                    288   -----------------------
                                    290   -----------------------
                                    294   -----------------------
099   K----------------------       297   -----------------------
101   -----------------------       298   -----------------------
102   -----------------------       299   -----------------------
104   -----------------------       301   -----------------------
110   -----------------------       331   -----------------------
115   -----------------------       387   -----------------------
122   -----------------------       416_  -----------------------
128   -----------------------       06P   -----------------------
129   -----------------A-----
135   -----------------------       acute patients
143   -----------------------       822   -----------------------
147   -----------------------       6-23  -----------------------
218   -----------------------       6-56  -----------------------
222   -----------------------       7-16  -----------------------
223   K----------------------       7-31  -----------------------
226   -----------------------       7-70  -----------------------
227   -----------------------       7-75  -----------------------
```

B

```
      AEFFTEVDGVRLHRYA       235   P---------------
099   ----------------       237   -----------I----
110   ----------------       274   ----------------
115   ----------------       275   P---------------
128   P---------------       288   ----------------
129   ----------------       290   ----------------
135   ----------------       294   ----------------
140   P---------------       297   ----------------
143   ----------------       298   P---------------
147   ----------------       299   ----------------
218   ----------------       301   ----------------
219   ----------------       331   -------------F--
223   -----------I----       387   ----------------
227   ----------------       411   P----------I----
228   ----------------       420   ----------------
229   -----------I----       429   P----------I----
230   ----------------       6-23  P---------------
233   -----------V----       7-31  P---------------
```

HCV gt3a patient sequences obtained by PCR at frequently detected CD8+ T cell targets detected by overlapping peptides. (A) Peptide NS3$_{1520}$ `RPSGMFDSVVLCECYDAGCSWYDL` and (B) Peptide `AEFFTEVDGVRLHRYA`. Patients responding to NS3$_{1520}$ in IFNγ ELISpot assays are marked in grey.

4.5.1.3 Sequence variability between HCV gt3 and gt1 at CD8+ T cell targets frequently detected in HCV gt3 infection

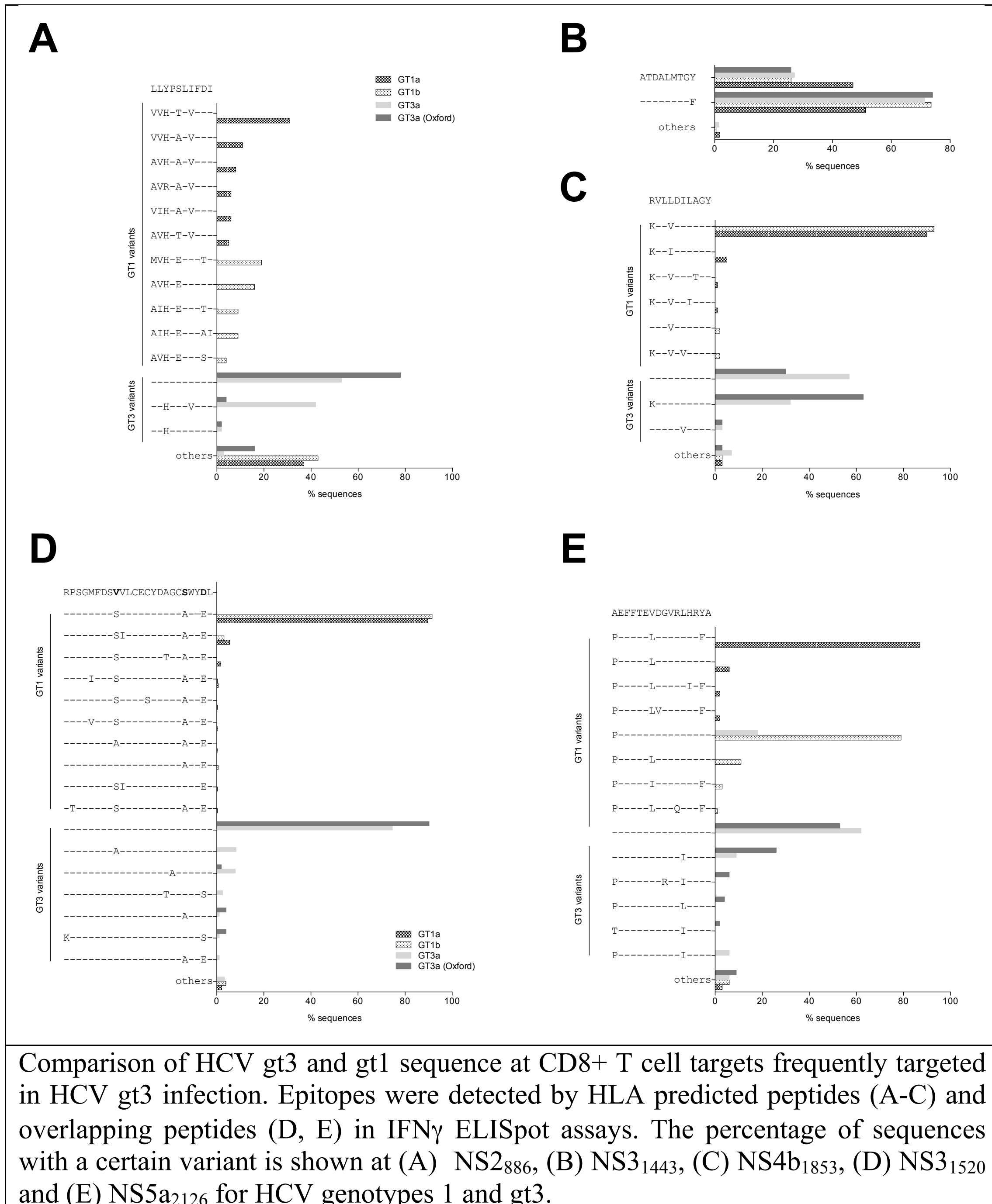
To exploit the potential usage of CD8+ T cell targets frequently detected in HCV gt3 infection for cross-genotypic vaccine design, sequence variability between HCV genotype 1 and 3 was assessed (Figure 4-17). Sequences for HCV gt1a, gt1b and gt3a sequences were obtained from the Los Alamos sequence database, with additional in-house sequences for HCV gt3a.

For epitope $NS2_{886}$ (L)LYPSLIFDI (Figure 4-17 A), major sequence variation within and between genotypes was observed, suggesting that this epitopes is indeed a HCV gt3 specific epitope. Epitope $NS4b_{1853}$ RVLLDILAGY (Figure 4-17 C) is variable between genotypes, but more conserved within subtypes 1a and 1b in comparison to the highly variable epitope $NS2_{886}$. In contrast, epitope $NS3_{1443}$ ATDAMTGY was highly conserved between HCV genotypes 1 and 3, with an identical escape variant in both genotypes (ATDAMTG**Y/F**) that was previously linked to immune pressure (Figure 4-17 B) (Neumann-Haefelin et al. 2008).

At T cell target $NS3_{1520}$ (RPSGMFDSVVLCECYDAGCSWYDL, Figure 4-17 D), sequence polymorphisms between HCV gt1 and gt3 were observed at three amino acid positions: at position 9 (S→V/A), position 20 (A→S) and position 23 (E→D/S/E) (Figure 4-17 D). However, $NS3_{1520}$ was highly conserved within HCV genotypes 1 and 3. In contrast, significant variability within and between HCV genotypes 1a, 1b and 3a was observed at T cell target $NS5a_{2126}$ (AEFFTEVDGVRLHRYA, Figure 4-17 E). Of note, the common HCV gt1b sequence variant **P**EFFTEVDGVRLHRYA was also observed in several HCV gt3 infected patients.

In summary, CD8+ T cell targets frequently detected in HCV gt3 infection were not conserved between HCV genotypes 1 and 3.

Figure 4-17: Comparison of HCV gt1 and gt3 sequences at frequently detected HCV gt3 specific CD8+ T cell targets.



Comparison of HCV gt3 and gt1 sequence at CD8+ T cell targets frequently targeted in HCV gt3 infection. Epitopes were detected by HLA predicted peptides (A-C) and overlapping peptides (D, E) in IFNγ ELISpot assays. The percentage of sequences with a certain variant is shown at (A) $NS2_{886}$, (B) $NS3_{1443}$, (C) $NS4b_{1853}$, (D) $NS3_{1520}$ and (E) $NS5a_{2126}$ for HCV genotypes 1 and gt3.

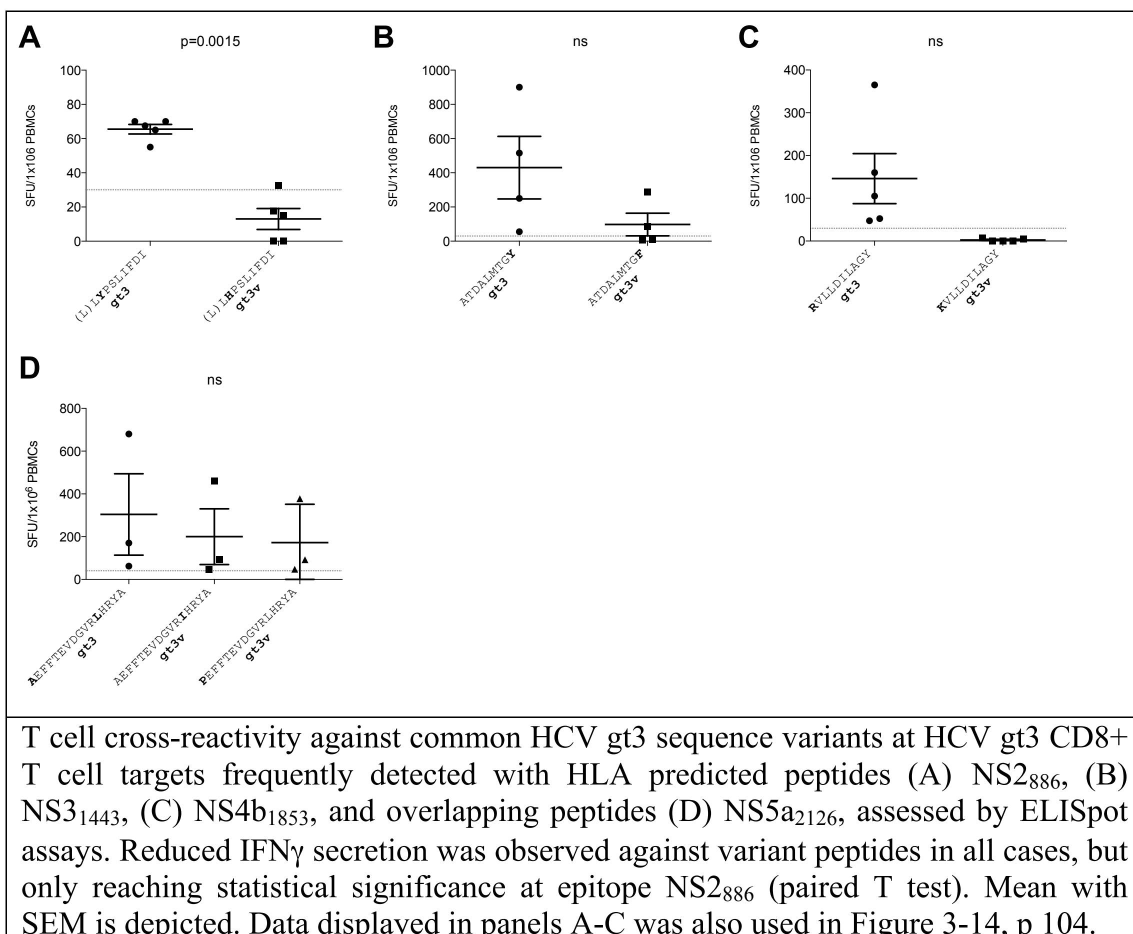
4.5.2 Cross-reactivity of T cell responses primed in HCV genotype 3 infection against common viral variants

To exploit the potential use of HCV gt3 specific CD8+ T cell targets in cross-reactive immunogen design, T cell cross-reactivity against identified HCV gt3 and gt1 viral variants was assessed in IFNγ ELISpot assays.

Cross-reactivity of T cells specific for frequently detected HCV gt3 T cell targets against common HCV gt3 sequence variants

Initially, the impact of identified HCV gt3 variants on T cell cross-reactivity was evaluated at three T cell targets detected by HLA predicted peptides and one detected by overlapping peptides (Figure 4-18). At epitopes frequently detected by HLA predicted peptides, reduced T cell responses were observed to the gt3 variant peptide in comparison to cytokine release when stimulated with the wild type peptide (Figure 4-18 A-C). Reduced, but not abrogated cross-reactivity was observed against identified gt3 sequence variants at peptide $NS5a_{2126}$ ((Figure 4-18 D).

Figure 4-18: T cell cross-reactivity against HCV gt3 variants at CD8+ T cell targets frequently detected in HCV gt3 infection.



T cell cross-reactivity against common HCV gt3 sequence variants at HCV gt3 CD8+ T cell targets frequently detected with HLA predicted peptides (A) $NS2_{886}$, (B) $NS3_{1443}$, (C) $NS4b_{1853}$, and overlapping peptides (D) $NS5a_{2126}$, assessed by ELISpot assays. Reduced IFNγ secretion was observed against variant peptides in all cases, but only reaching statistical significance at epitope $NS2_{886}$ (paired T test). Mean with SEM is depicted. Data displayed in panels A-C was also used in Figure 3-14, p 104.

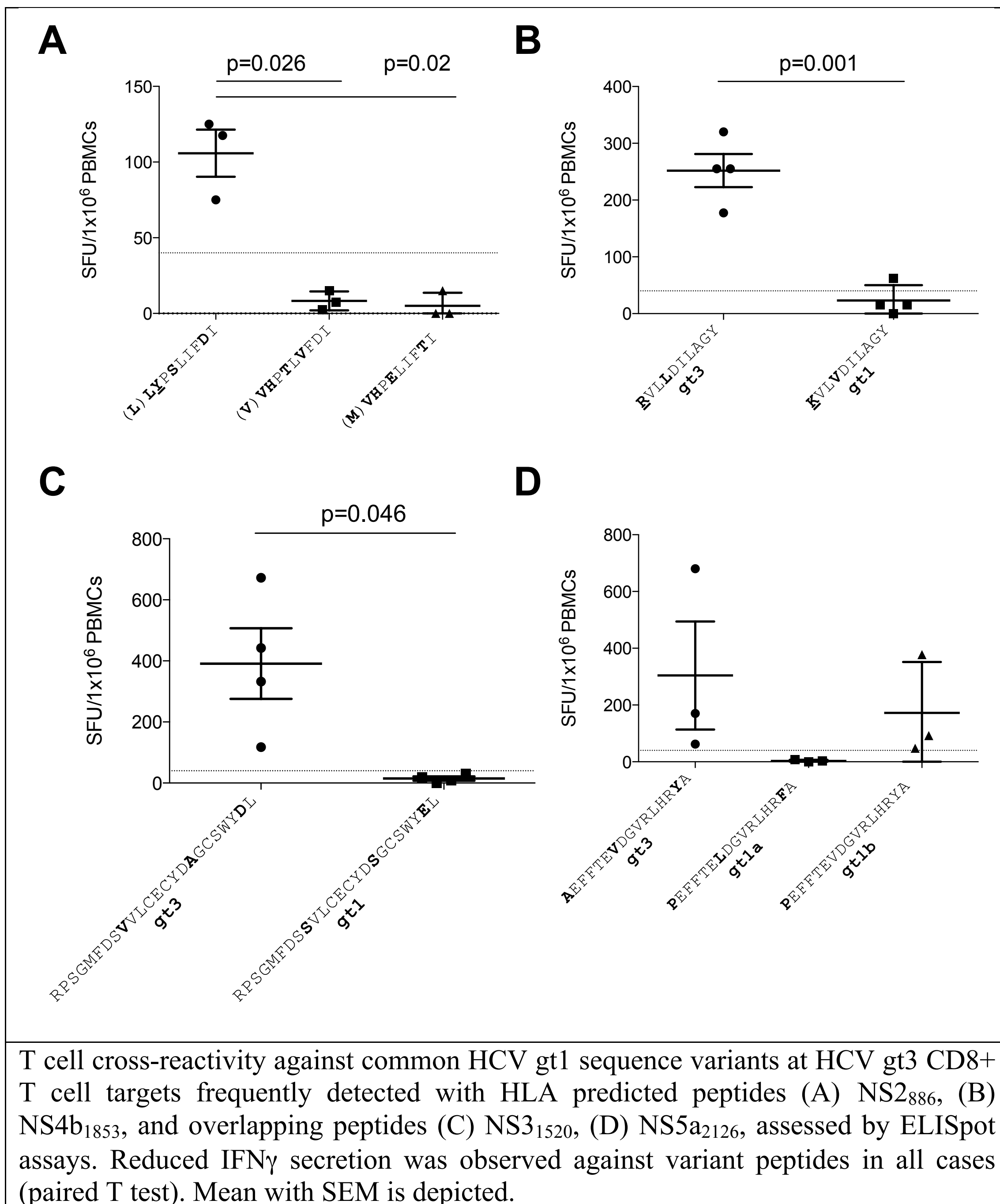
Cross-reactivity of T cells specific for frequently detected HCV gt3 T cell targets against common HCV gt1 sequence variants

Next, T cell cross-reactivity against common HCV gt1 variants was assessed at CD8+ T cell targets frequently detected in HCV gt3 infection (Figure 4-18). Limited T cell cross-reactivity was observed against identified HCV gt1 variants at epitopes (L)LYPSLIFDI and RVLLDILAGY (Figure 4-19 A, B). HCV gt1 variants at epitope ATDAMTGY did not differ from previously observed variants within HCV gt3 (ATDAMTG**Y**/**F**), where reduced IFNγ production has been observed in ELISpot assays when stimulated with variant ATDAMTG**F** (see Figure 4-18, page 181).

At two CD8+ T cell targets detected by overlapping peptides, limited cross-reactivity of T cells primed in HCV gt3 infection was observed against HCV gt1 variants (Figure 4-19 C, D). The HCV gt1 sequence variant of peptide $NS3_{1520}$ was not recognized by T cells primed in HCV gt3 infection (Figure 4-19 C). At peptide $NS5a_{2126}$, some T cell cross-reactivity was observed against the tested HCV gt1b variant (**P**EFFTEVDGVRLHRYA), which was also detected in HCV gt3 infection; however, IFNγ release was abrogated when stimulating with the identified gt1b variant (Figure 4-19 D).

Overall, limited cross-reactivity against sequence variants within HCV gt3 and between HCV genotypes was observed at CD8+ T cell targets frequently detected in HCV gt3 infection. These results strongly suggest that epitopes frequently detected in natural HCV gt3 infection may not serve as attractive targets in vaccine design aiming to induce cross-reactive responses.

Figure 4-19: T cell cross-reactivity against HCV gt1 sequence variants at frequently detected CD8+ T cell targets.



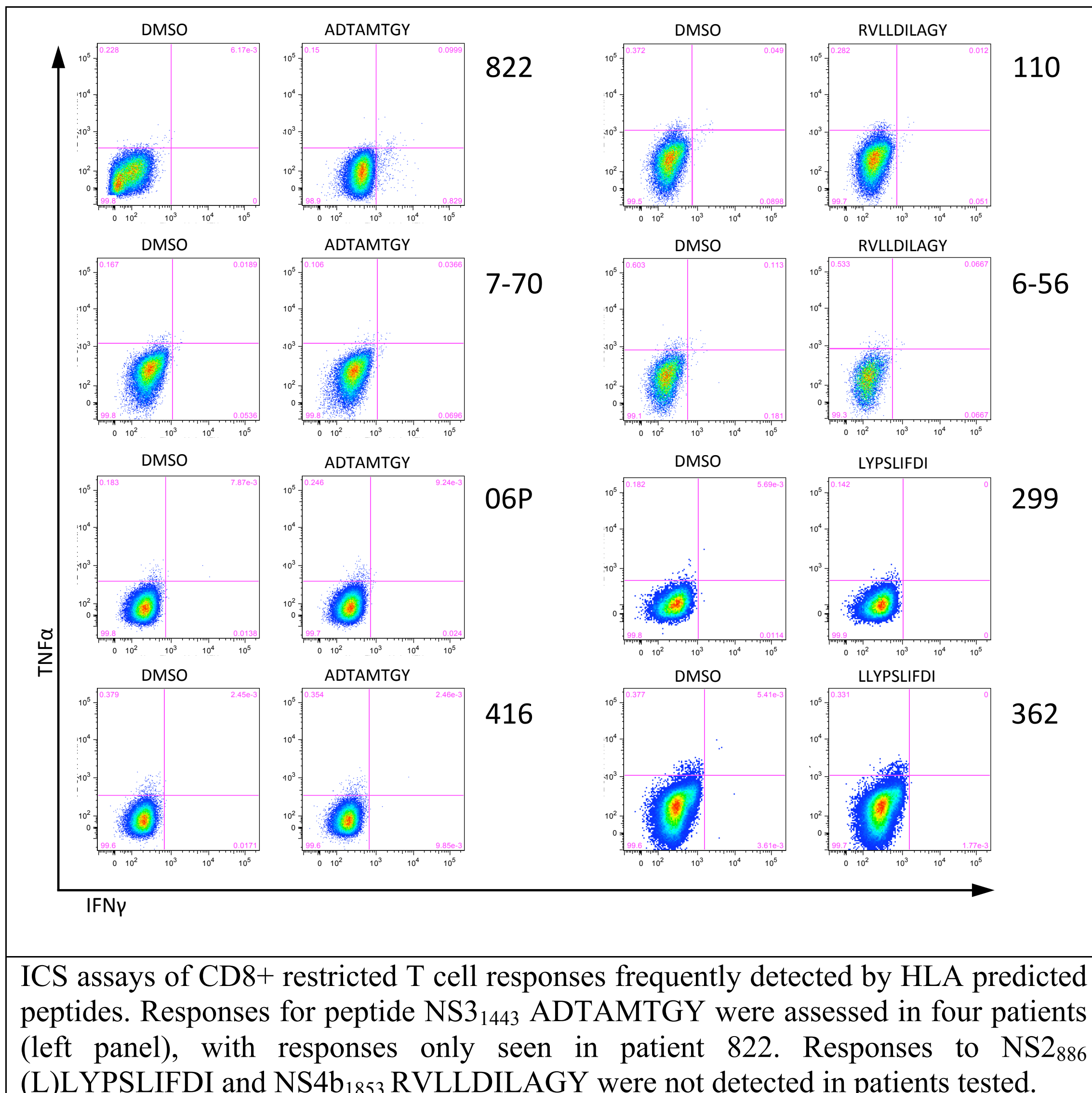
T cell cross-reactivity against common HCV gt1 sequence variants at HCV gt3 CD8+ T cell targets frequently detected with HLA predicted peptides (A) $NS2_{886}$, (B) $NS4b_{1853}$, and overlapping peptides (C) $NS3_{1520}$, (D) $NS5a_{2126}$, assessed by ELISpot assays. Reduced IFNγ secretion was observed against variant peptides in all cases (paired T test). Mean with SEM is depicted.

4.5.3 Function of frequently detected HCV gt3 specific CD8+ T cell responses

4.5.3.1 Phenotypic analysis of CD8+ responses frequently detected by HLA predicted peptides

Intracellular staining assays: To assess the phenotype and function of T cell responses frequently detected using HLA predicted peptides, patient PBMC were initially assessed using intracellular staining assays (ICS) in 8 patients with cells available (Figure 4-20). *Ex vivo* T cell responses using ICS assays after overnight peptide stimulation were only detected for acutely infected patient 822, with all other patients showing no or only very weak T cell responses upon peptide stimulation.

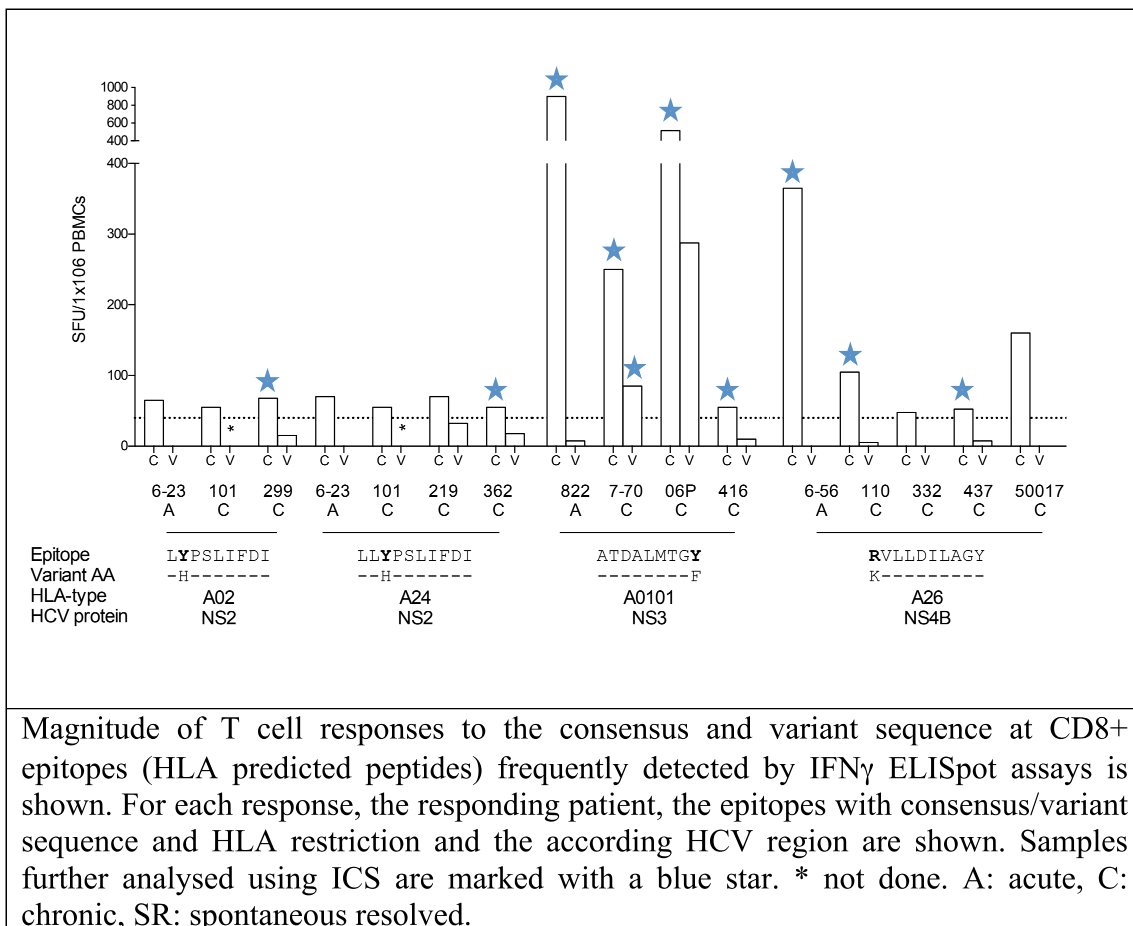
Figure 4-20: T cell responses to HLA predicted peptides frequently detected in IFNγ ELISpot assays assessed by intracellular staining assay



ICS assays of CD8+ restricted T cell responses frequently detected by HLA predicted peptides. Responses for peptide $NS3_{1443}$ ADTAMTGY were assessed in four patients (left panel), with responses only seen in patient 822. Responses to $NS2_{886}$ (L)LYPSLIFDI and $NS4b_{1853}$ RVLLDILAGY were not detected in patients tested.

The lack of detected responses may be due to lower sensitivity of ICS assays in comparison to IFNγ ELISpot assays. In our experience, CD8+ T cell responses less than 200 SFU/10^6 PBMC in ELISpot assays are not detectable by ICS. Indeed, most of the responses tested (marked with a blue star in Figure 4-21) were low in magnitude. Of note, patients 822 and 6-56 were the only acutely infected patients that were tested in ICS assays; none of the chronically infected patients tested exhibited a T cell response in ICS assays. T cell responses for acute patient 822 were IFNγ^{high}, TNFα^{+}, Mip1b^{+} and negative for IL-2 (data not shown).

Figure 4-21: Magnitude of T cell responses in IFNγ ELISpot to HLA predicted peptides, which were further tested in ICS assays.



Magnitude of T cell responses to the consensus and variant sequence at CD8+ epitopes (HLA predicted peptides) frequently detected by IFNγ ELISpot assays is shown. For each response, the responding patient, the epitopes with consensus/variant sequence and HLA restriction and the according HCV region are shown. Samples further analysed using ICS are marked with a blue star. * not done. A: acute, C: chronic, SR: spontaneous resolved.

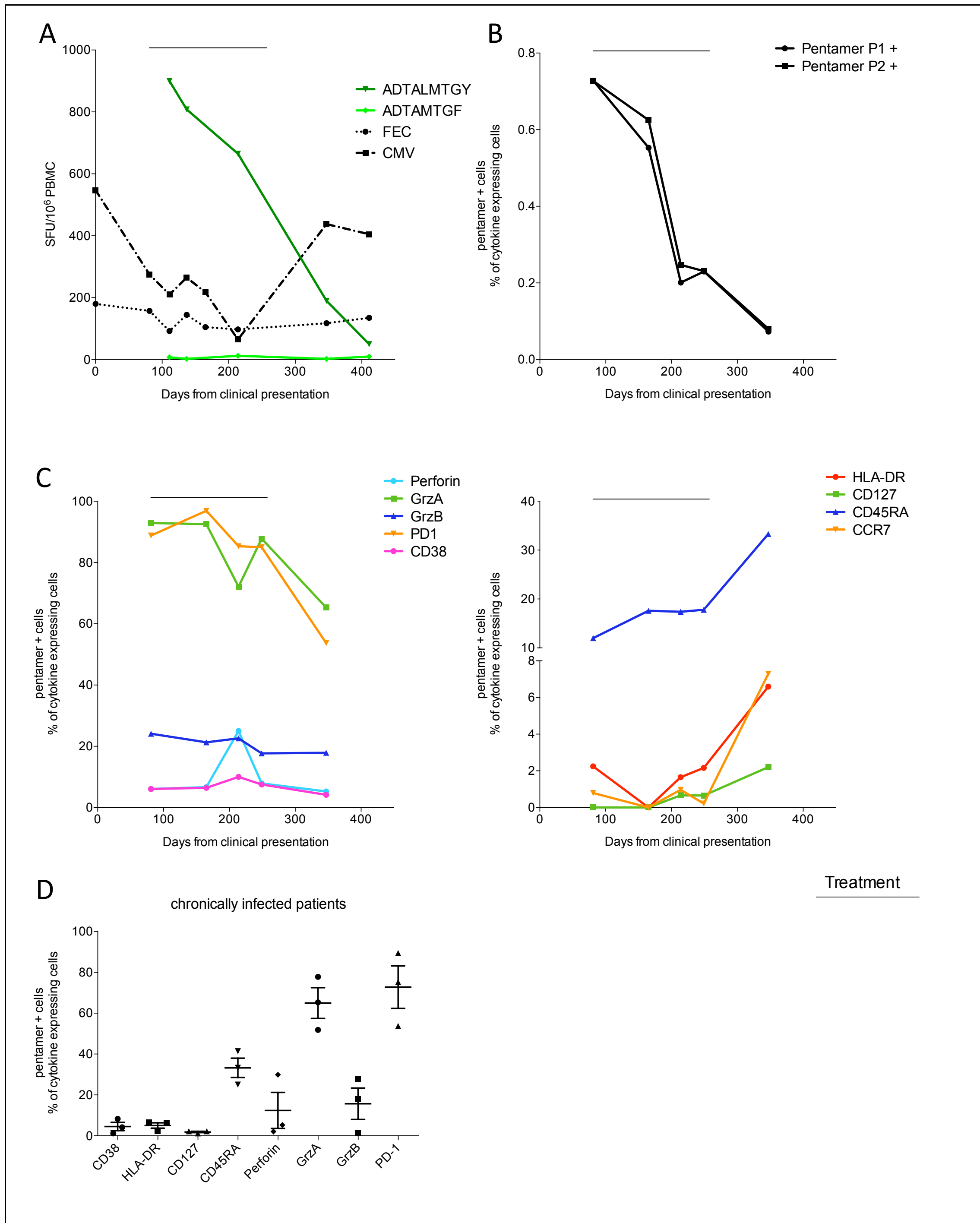
Taken together, the analysis of T cell responses by conventional intracellular staining techniques was complicated by the low magnitude of T cell responses in chronically infected patients, in which exhaustion and T cell dysfunction is common. To enable further analysis of frequently detected low level T cell responses, the use specific labelling techniques was required.

Tetramer staining assays: Next, epitope $NS3_{1443}$ ADTAMTGY specific T cell responses were assessed in patients with detected responses in IFNγ ELISPot assays (one acute patient (822) and three chronic patients (7-70, 06P, 416), Figure 4-22). In acute infection, T cell responses were primed to the wild type sequence ADTAMTGY, with no detectable response to the escape variant ADTAMTG**F** (Figure 4-22 A). Detected tetramer positive cells reflected the magnitude of T cell responses seen in ELISpot assays (Figure 4-22 B). Epitope $NS3_{1443}$ specific T cell responses decreased during treatment in the acute phase, a phenomenon previously described in chronic HCV infection (Golden-Mason et al. 2007; V. Kasprowicz et al. 2008; R. H. McMahan et al. 2010). Before treatment, HCV specific T cells were high in PD-1, a marker reflecting cell activation that has also been linked to exhaustion (Golden-Mason et al. 2007; V. Kasprowicz et al. 2008; McMahan et al. 2010), and also displayed high amounts of Granzyme A, a cytokine commonly expressed in memory cells. Intriguingly, Granzyme B expression, dependent on recent activation and associated with cytotoxic activity in the literature (Harari et al. 2009), was detected in only 30% of epitope specific T cells in acute infection. Epitope specific cells also showed low expression of activation markers HLA-DR and CD38, and of the cytotoxic cytokine Perforin. The percentage of cells expressing memory markers like CCR7 and CD127 increased over time during treatment in acute infection (Figure 4-22 C), as well as CD45RA expression (a marker of antigen exposure and expressed on terminally differentiated CD8+ T cells).

Epitope $NS3_{1443}$ specific cells in chronically infected patients displayed a similar phenotypes (Figure 4-22 D), with only few cells expressing activation markers like CD38 and HLA-DR, about 10% to 40% of cells expressing the differentiation marker CD45RA and cytotoxic cytokines like Perforin and Granzyme B, whereas high percentages of tetramer specific cells were positive for Granzyme A and PD-1. For details on analysed phenotypic markers refer to Methods, Table 2-3, page 75.

Overall, the T cell responses detected in ELISpot assays to HLA predicted peptides were low in magnitude, complicating phenotypic analysis by intracellular staining assays. Tetramers were used to further analyse low-level responses: $NS3_{1443}$ specific T cells in acute infection displayed low expression of activation markers and cytotoxic cytokines, with similar phenotypic profiles observed at HCV epitope specific cells in chronic infection.

Figure 4-22: Phenotypic profile of ADTAMTGY in acute and chronic HCV gt3 infection.



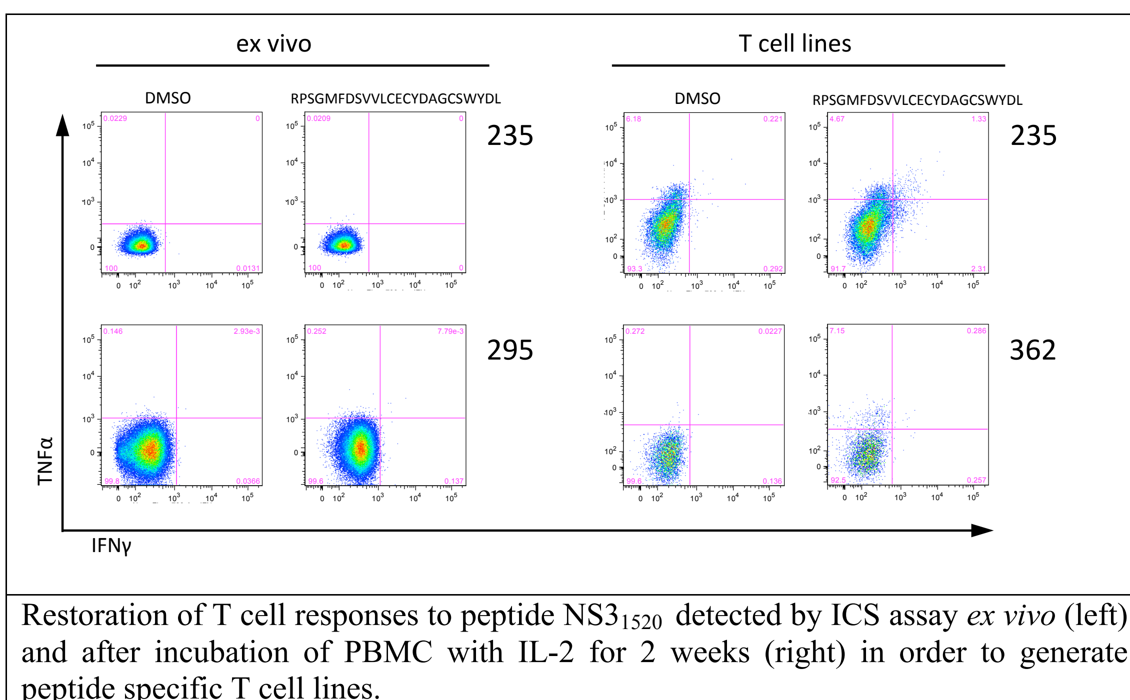
Phenotypic analysis of HCV gt3a specific T cell responses towards epitope NS3 1443 ADTALMTGY in acute and chronic infection. (A) In acute infection (pt 822), a strong T cell response was detected in IFNγ ELISpot assays to the priming epitope (ADTAMTG**Y**), but none to the escape variant (ADTAMTG**F**). The response decreased during treatment (black line). Responses to non-HCV responses CD8+ restricted (FEC: flu, EBV, CMV) and CD4+ restricted epitopes (CMV) are depicted. (B) Percentage of tetramer+ cells. (C) Phenotype of virus specific, shown as % of tetramer+ cells expressing different cytokines. (D) Phenotypic profile of tetramer+ cells in chronic infection does not vary from phenotype detected in acute infection.

4.5.3.2 Phenotypic characterization of a CD8+ restricted T cell response frequently detected using overlapping peptides

The T cell response to T cell target $NS3_{1520}$ RPSGMFDSVVLCECYDAGCSWYDL was particularly frequent response in the Oxford gt3 cohort, detected in 12 chronically infected patients as well as in two individuals spontaneously resolving HCV infection. I aimed to assess whether T cell phenotype and function at this T cell target is distinct in spontaneously resolved and chronically infected individuals.

ICS responses ex vivo T cell lines: Initially, $NS3_{1520}$ specific T cells were evaluated using intracellular staining assays (Figure 4-23) in chronically infected individuals. T cell were incubated with and without peptide $NS3_{1520}$ overnight, and cytokine production was assessed. However, even though strong T cell responses in chronically infected patients were detected using ELISpot assays (160 and 165 SFU/10^6PBMC in patients 235 and 295, respectively), I was not able to detect this CD8+ response *ex vivo* in ICS assays (Figure 4-23, left). However, responses in chronically infected patients were recovered after stimulation with IL-2 for 2 weeks (Figure 4-23, right). Ideally, $NS3_{1520}$ specific T cell phenotypes will be evaluated using tetramers, enabling a straight forward labelling technique and detection of responses not detectable by ICS.

Figure 4-23: Responses to peptide $NS3_{1520}$ were recovered by stimulation with IL-2.

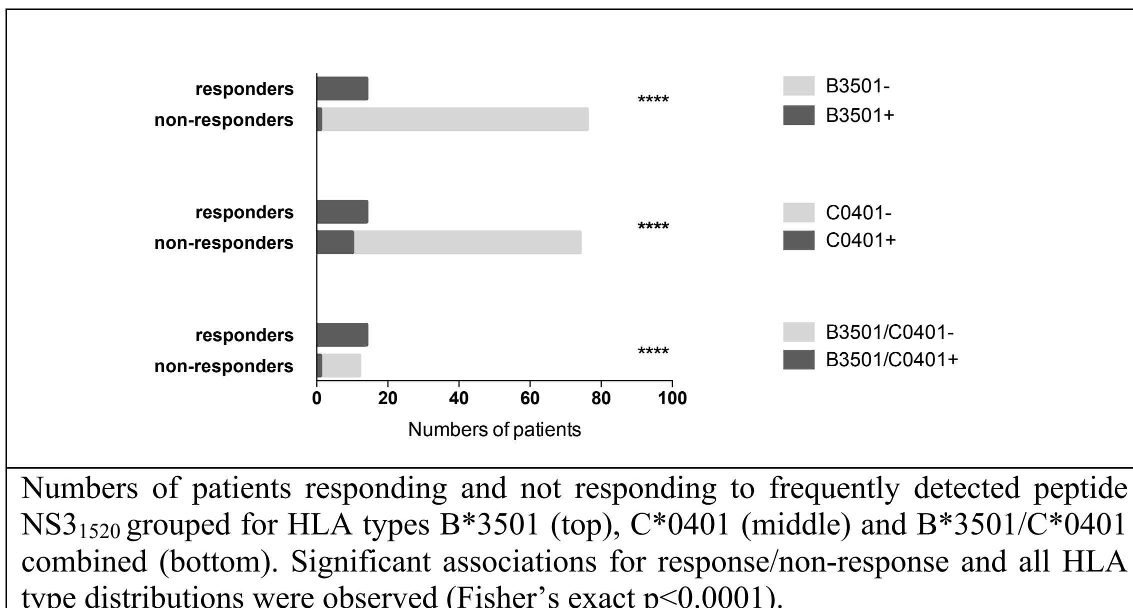


Restoration of T cell responses to peptide $NS3_{1520}$ detected by ICS assay *ex vivo* (left) and after incubation of PBMC with IL-2 for 2 weeks (right) in order to generate peptide specific T cell lines.

Tetramer development: Since a T cell response to the HCV gt3 specific T cell target $NS3_{1520}$ has not been described previously, no tetramer was available on the market. I therefore aimed to define HLA restriction at this frequently detected peptide, as well as the optimal binding sequence to enable tetramer production.

To define the restricting HLA type for the epitope contained in peptide $NS3_{1520}$, HLA type distributions of responding patients versus non-responding patients in the Oxford HCV gt3 cohort to peptide $NS3_{1520}$ were assessed (Figure 4-24). All patients responding to peptide $NS3_{1520}$ carried HLA alleles B*3501 and C*0401, but in one patient with positive for B*3501 and C*0401 no response was detected. Of note, even though a significant association between the HLA distribution was observed for HLA type C*0401 was noted, not all patients who only had the C*0401 allele, but not the B*3501 allele responded to peptide $NS3_{1520}$.

Figure 4-24: HLA type distributions in patients with and without T cell response to peptide $NS3_{1520}$.



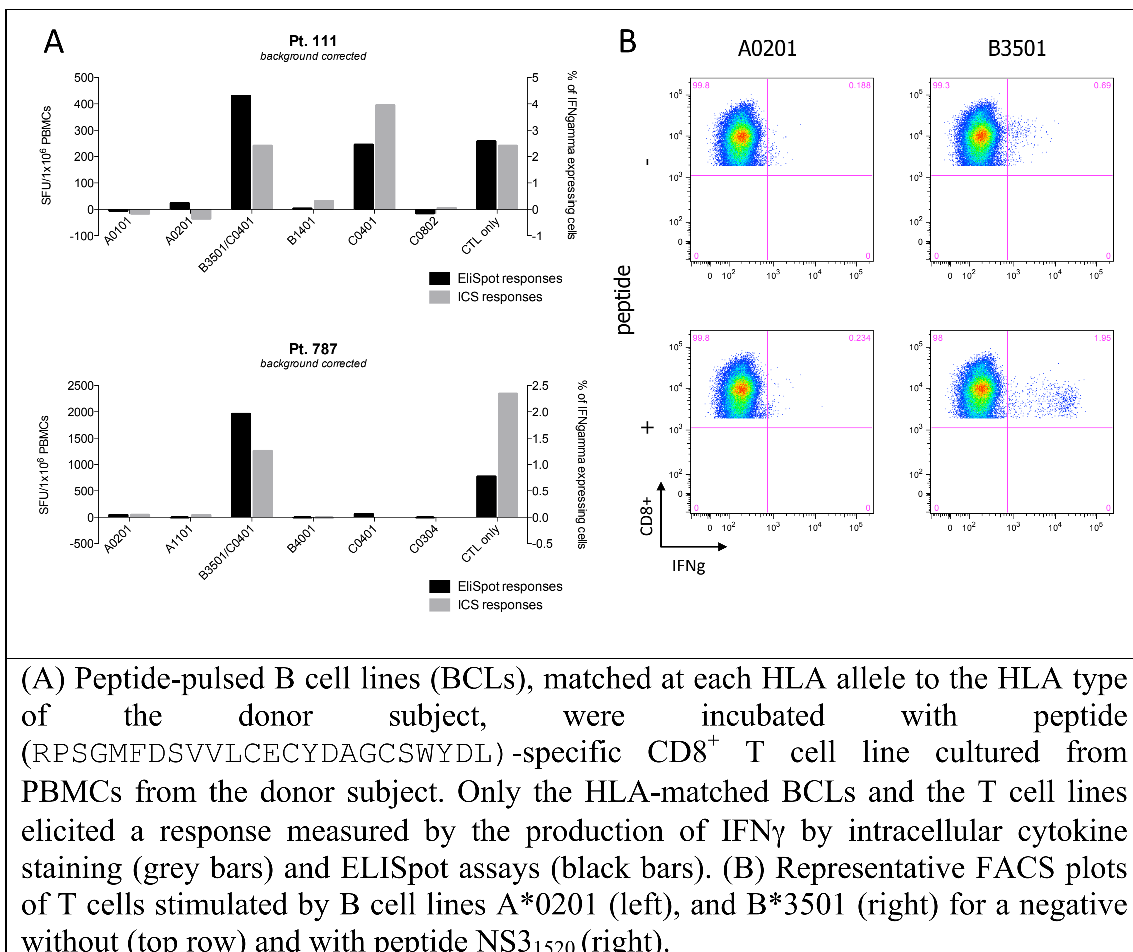
Numbers of patients responding and not responding to frequently detected peptide $NS3_{1520}$ grouped for HLA types B*3501 (top), C*0401 (middle) and B*3501/C*0401 combined (bottom). Significant associations for response/non-response and all HLA type distributions were observed (Fisher's exact $p<0.0001$).

It has been previously described that HLA alleles B*3501 and C*04 are in strong linkage disequilibrium. Linkage disequilibrium is the non-random association of alleles at two or more loci, which descend from single, ancestral chromosome. This phenomenon is common for HLA class I loci, particularly for HLA-B and -C, which are separated by only 100 kb on chromosome 6 (Reich et al. 2001; Miretti et al. 2005). Of note, the linked HLA types B*35 and C*04 have been linked to disease progression HIV and HCV infection: In HIV, the HLA class-I alleles B*35/C*04 were associated with rapid development of AIDS-defining conditions in Caucasians

(Carrington et al. 1999). In HCV infection, HLA type C*04 was associated with viral persistence in a study assessing HLA patterns in 231 individuals with spontaneously resolved and 444 matched individuals with persistent infection (Thio et al. 2002).

To confirm the restriction of the T cell responses against peptide $NS3_{1520}$ by HLA types B*3501 and C*0401, HLA restriction experiments were performed as described previously (Payne et al. 2010). In brief, peptide specific T cell lines from fresh cells from donors with previously confirmed responses to $NS3_{1520}$ were grown for 2 weeks under IL-2 stimulation. Subsequently, peptide was presented to patient T cell lines by B cell lines (BCLs) matched to the donors HLA type. Reactivity of T cells was assessed by measuring IFNγ production using flow cytometry and ELISpot assays. Background corrected results for IFNγ production measured by ELISpot assay and flow cytometry are depicted in Figure 4-25.

Figure 4-25: HLA restriction experiments at peptide $NS3_{1520}$.



(A) Peptide-pulsed B cell lines (BCLs), matched at each HLA allele to the HLA type of the donor subject, were incubated with peptide (RPSGMFDSVVLCECYDAGCSWYDL)-specific $CD8^+$ T cell line cultured from PBMCs from the donor subject. Only the HLA-matched BCLs and the T cell lines elicited a response measured by the production of IFNγ by intracellular cytokine staining (grey bars) and ELISpot assays (black bars). (B) Representative FACS plots of T cells stimulated by B cell lines A*0201 (left), and B*3501 (right) for a negative without (top row) and with peptide $NS3_{1520}$ (right).

HLA restriction was confirmed in two patients with available fresh PBMC (chronic patients 111 and 787). Since HLA-type B*3501 is in strong linkage disequilibrium

with C*04, we were not able to generate a B cell line for B*3501 alone. Strong IFNγ production was observed after peptide stimulation from B cell lines with B*3501/C*04 restriction in both tested patients. Only patient 111 showed a reaction to the peptide presented by C*04 alone, which was not observed in patient 787.

Next, I aimed to define the peptide with optimal binding properties within the overlapping peptide RPSGMFDSVVLCECYDAGCSWYDL (Figure 4-26).

Figure 4-26: Strategies of optimal mapping for overlapping peptide NS3$_{1520}$ **and representative ELISpot responses to different optimal mapping strategies.**

A **RPSGMFDSVVLCECYDAGCSWYDL**

1. overlapping peptide

RPSGMFDSVVLCECYDAGCSWYDL
RPSGMFDSVVLCECYDA P231
P232 DSVVLCECYDAGCSWYDL

2. Epitope prediction

B*3501 RPSGMFDSVVLCECYDAGCSWYDL
B*3501 RPSGMFDSVVLCECYDAGCSWYDL
C*0401 RPSGMFDSVVLCECYDAGCSWYDL

3. 11mers

RPSGMFDSVVLCECYDAGCSWYDL
...

4. Shortening from both sides

RPSGMFDSVVLCECYDAGCSWYDL

RPSGMFDSVVLCECYDA P231

P232 DSVVLCECYDAGCSWYDL

B

patient 295

patient 787

SFU/1x10^6 PBMC (IFNγ ELISpot)

0, 50, 100, 150, 200

G, P231/232, P231-1, P231-2, P231-3, P231-4, P231-5, P231-6, P231-7, P231-8, P231-9, P231-10, P231-11, P231-12, P231-13, P231-14

Strategies followed for optimal mapping at peptide NS3$_{1520}$. (A) First, the overlapping peptide was tested. Then, epitopes predicted as binders for B*3501 and C*0401 were assessed in ELISpot assays. Since no response was detected, overlapping 11mers were tested. Finally, both peptides with detected responses were shortened from both sides. (B) Representative results from ELISpot assays for optimal peptide mapping, here shown for testing overlapping 11mers.

Since this T cell target was identified using two neighboured overlapping peptides (RPSGMFDSVVLCECYDA and DSVVLCECYDAGCSWYDL) we first tested the overlapping 11mer as an optimal (DSVVLCECYDA) (Figure 4-26 A1). No response to this 11mer was observed. We subsequently followed the stragety outlined in Figure 4-26 (A): To optimize cell use in ELISpot experiments, an epitope prediction using freely available online tools (Syfpeithi, BIMAS and NetMHC) was run for the overlapping peptide, with 3 strong binders predicted (Figure 4-26 A2). However, no T cell responses were detected in IFNγ ELISpot assays. Next, we tested 14 consecutive peptides of 11 amino acids in length (Figure 4-26 A3); again, no T cell response was detected using this approach (representative responses in Figure 4-26 B).
B*3501 was reported to present untypically long peptide for HLA class-I complexes, with peptide binding length up to 14 amino acids described (Probst-Kepper et al. 2004). B*3501 has its anchor binding residues at position 2 (Proline, P) and position 9/10 (Tyrosine, Y) (Takamiya et al. 1994); however, was described as a relatively non-restrictive molecule. This is a feature it shares with other HLA alleles linked to HIV progression (Kosmrlj et al. 2010). To account for this B*3501 specific properties, we next shortened the presented overlapping peptide on both sides (Figure 4-26 A4). This strategy would also account for an unlikely but possible scenario, where two different peptides are contained in $NS3_{1520}$, both restricted by HLA types B*3501 or C*0401. However, using this approach I was not able to detect the optimal binding peptide length.
The failure to recognize T cell responses to optimals derived from overlapping peptides could be due to a contamination of the original screening peptide with a different peptide that was exclusively presented by HLA type B*3501. To test for contamination of used overlapping peptides, peptides RPSGMFDSVVLCECYDA and DSVVLCECYDAGCSWYDL were assessed by mass spectrometry. The sequences of the two peptides were confirmed, but no contamination was detected (data not shown). Taken together, we were unable to define the optimal binding peptide, and therefore were not able to produce a tetramer for this response.

In summary, the epitope contained in peptide $NS3_{1520}$ is a B*3501/C*0401 restricted. In chronically infected patients, cytokine release was not detected *ex vivo*, but recovered after *in vitro* IL-2 stimulation.

4.6 Discussion

In this thesis, I have identified multiple HCV gt3 specific T cell responses (Chapter 3). In Chapter 4, I further assessed whether T cell targets frequently detected in HCV gt3 infection would serve as potential targets in cross-reactive immunogen design. The role of HCV gt3 specific T cell targets frequently detected in acute, chronic and spontaneously resolved HCV gt3 infection was evaluated regarding; sequence variability at frequently detected T cell targets, cross-reactivity to identified sequence variants within HCV gt3 and between different genotypes and functionality of T cells.

Definition of T cell targets frequently detected in HCV gt3 infection: Seven frequently detected HCV gt3 T cell targets were identified, defined as those observed in more than four patients. Two common targets in the core region were CD4+ restricted (core 66-83, PKARRSEGRSWAQPGYPW and core 143-158, PVGGVARALAHGVRAL), whereas five T cell targets in HCV non-structural regions were targeted by CD8+ T cells, of which three were detected by HLA predicted peptides ($NS2_{886}$ LLYPSLIFDI, $NS3_{1443}$ ATDALMTGY and $NS4b_{1853}$ RVLLDILAGY), and two by overlapping peptides ($NS3_{1520}$ RPSGMFDSVVLCECYDAGCSWYDL and $NS5a_{2126}$ AEFFTEVDGVRLHRYA).

Due to the composition of the study cohort (108 chronic patients and only 16 spontaneously resolved patients), the majority of T cell responses defined as “frequently detected” were identified in chronically infected patients. Even though T cell responses detected in spontaneously resolved infection were significantly higher in magnitude than those detected in chronic infection, most have not been defined as “frequently detected”. In addition to the potential reason discussed above, spontaneously resolved patients targeted a wide range of different peptides, further decreasing the likelihood of being defined as “frequently detected” following the definition applied here. However, two frequently detected peptides were targeted by spontaneously resolved patients; $NS3_{1520}$ and $NS5a_{2126}$ are therefore of particular interest for potential vaccine design.

Sequence variability and T cell cross-reactivity within and between HCV genotypes: To assess whether identified frequently detected T cell targets could serve as attractive targets for cross-reactive immunogen design, sequence variability was assessed within HCV gt3 and across HCV genotypes.

CD4+ restricted responses to HCV core: In comparison to other HCV viral regions, core is comparably conserved (Contreras et al. 2002), potentially leading to a high level of cross-reactivity of T cells primed in HCV gt3 infection against viral variants within and between HCV genotypes. Indeed, for two frequently detected core specific CD4+ responses, sequence polymorphisms were rarely observed within HCV gt3 cohort sequences.

However, sequence variants were observed between HCV gt3 and gt1 at CD4+ T cell targets frequently detected in HCV gt3 infection. In addition, it was shown that T cell specificity is distinct in HCV genotype 1 and 3 infection (Semmo et al. 2005; G Harcourt et al. 2006; V. M. Fleming et al. 2010; Rohrbach et al. 2010): Compared to gt1 infection, where broad CD4+ core responses to multiple peptides have been described (V. M. Fleming et al. 2010), specificity of T cell responses in gt3 infection was narrow with mainly two frequently detected peptides targeted. Since core is a comparably conserved region of the HCV genome, this phenomenon can only be partly explained by sequence diversity between genotypes. An additional possible reason for distinct peptide presentation between genotypes may be binding interference of other competing peptides at presenting HLA class-II epitopes, which might skew peptide binding and presentation of core peptides. Of note, distinct response profiles between HCV gt1 and gt3 have been previously described in the literature in a cohort of HIV/HCV co-infected individuals (Rohrbach et al. 2010); high CD4+ T cell responses to core peptides were specifically observed in the subgroup of patients infected with HCV gt3, however, T cell specificity has not been defined in this publication.

Next, cross-reactivity of T cells specific for frequently detected CD4+ T cell targets against identified HCV sequence variants was tested in IFNγ ELISpot assays. Limited cross-reactivity was observed against common HCV gt1 sequence variants at peptides core 66-83 and core 143-158.

CD8+ restricted responses to non-structural viral regions: For four out of five frequently detected CD8+ T cell targets detected in HCV non-structural regions, high numbers of polymorphisms were observed within HCV gt3. However, one CD8+ T cell target, $NS3_{1520}$, was conserved within HCV gt3. Sequence analysis between HCV genotypes 3 and 1 at frequently detected CD8+ T cell targets demonstrated that four targets were not conserved between genotypes, including peptide $NS3_{1520}$ (highly

conserved *within* HCV gt3 and gt1). In contrast, the frequently detected epitope ATDALMTGY is conserved between HCV genotypes 1 and 3, with a common T cell mediated escape variant has been previously described in the literature (Neumann-Haefelin et al. 2008), which was also observed in this study.

CD8+ T cells specific for frequently detected HCV gt3 specific T cell targets showed limited cross-reactivity in IFNγ ELISpot assays against identified HCV gt3 and gt1 viral variants.

Taken together, high degrees of sequence variability at frequently detected HCV gt3 specific T cell targets have been observed, with limited T cell cross-reactivity against viral variants within gt3 and across genotypes. This was especially true for epitopes detected by HLA predicted peptides, which were based on escape variants within a large HCV gt3 sequence dataset, suggesting that they might be particularly prone to accumulation of sequence polymorphisms. These results are in line with limited cross-reactivity previously reported for T cells primed in HCV gt1 infection (Fytili et al. 2008; Skibbe et al. 2014), and will potentially limit the sensible use of frequently detected HCV gt3 specific T cell targets as vaccine targets priming cross-reactive immune responses.

Phenotypic analyses of T cell responses: Polyfunctional T cells have been defined as having the ability to secrete cytokines and mediate cytolysis and are highly predictive of protective immunity. However, in chronic viral infection, T cells often become exhausted in response to persistent antigen stimulation, upregulating multiple inhibitory receptors and get progressively monofunctional (Minton 2014). It has been recently described that exhausted CD4+ and CD8+T cells in chronic infection exhibit distinct profiles (Crawford et al. 2014). To further define the role of T cells specific for frequently detected HCV gt3 CD4+ and CD8+ T cell targets, function and phenotype was assessed.

Analysis of CD4+ T cell phenotype and function: The analysis of core-specific CD4+ restricted T cells was complicated by the low magnitude of responses just above IFNγ ELISpot detection threshold. Based on previous studies, where a substantial increase in T cell responses was described after short term pre-incubation of cells at high concentrations (Römer et al. 2011), a similar experiment for CD4+ restricted HCV core responses was attempted. Increased responses are thought to be mediated by increased cell-cell contact in high concentration cultures, thus reflecting

a cell phenotype similar to PBMC activation *in vivo.* Indeed, we observed an increase in percentage of cytokine secreting HCV core specific CD4+ cells after pre-incubation at high cell concentrations. However, in contrast to a previous study reporting increased cytokine production after pre-incubation and an altered cytokine profile (Kutscher et al. 2013), no difference in T cell polyfunctionality after short term incubation was observed in our hands. Core specific CD4+ T cell were poorly proliferative and mainly monofunctional. T cells were positive for TNFα only with some cells producing IL-2, which is in contrast to previously published data, where IL-2 was rarely observed in core specific HCV gt3 responses (Semmo et al. 2005). This data proves that short-term pre-incubation of T cells may serve as a technique to detect and further analyse low level CD4+ T cell responses just above detection threshold. This technique may be potentially useful for analysing other low level responses in cancer and other persistent infections like HIV and HCV where epitope specific tetramers are not available.

Analysis of CD8+ T cell phenotype and function: Similarly to analysed CD4+ responses, CD8+ restricted responses were weak and not detectable by ICS in the majority of samples, however, it was possible to recover CD8+ T cell responses by culturing cells with IL-2. Tetramer analysis showed that responses had low expression of activation markers and cytotoxic cytokines, suggesting that they are exhausted and do not exert immune pressure on the virus, similarly to dominant HCV gt1 responses previously described in the literature (Bengsch et al. 2010; Jo et al. 2012).

Further work: Future experiments could address cross-reactivity at frequently detected HCV gt3 specific T cell targets against HCV variants of other genotypes (2, 4, 5, 6) to enable an informed decision on whether these would indeed not serve as attractive targets in cross-reactive immunogen design.

The particular role of CD4+ restricted responses remains illusive and could be focus of future experiments. Evidence from the literature points to a T regulatory phenotype: it has been described previously that HCV core primes a regulatory T cell phenotype (Langhans et al. 2010). In addition, core-transduced CD4+ cells have been shown to exhibit a regulatory phenotype, with an increased basal secretion of the regulatory cytokine IL-10, a decreased IFNγ production and expression of regulatory cytokines, CTLA-4 and FoxP3, compared to non-transduced CD4+ cells (Dominguez-Villar et al. 2012; Fernandez-Ponce et al. 2014). Specifically, T cells

specific for the HCV gt1 version of core 143-158 have been shown to produce high amounts of IL-10, but not in chronically infected individuals (Godkin et al. 2001).
The increase in core specific CD4+ responses in late acute infection in individuals developing chronic infection is reminiscent of that previously observed in cytomegalovirus and Parvovirus 4 infection (Sylwester et al. 2005; R. Simmons et al. 2013). In these infections, some virus-specific T cells do not contract after the proliferative expansion during the acute phase, but instead are maintained and/or accumulate at high frequencies, a phenomenon termed memory inflation (O'Hara et al. 2012). It is therefore also possible that core specific CD4+ cells represent an effector memory population, a hypothesis requiring further evaluation.
A comprehensive analysis of phenotypic markers expressed on frequently detected CD4+ and CD8+ HCV gt3 specific T cells has not been assessed in this study; further experiments could define phenotypic markers greater detail. MHC class-I and -II tetramers specific would have facilitated a better analysis of cell phenotypes, and may be subject of further research effort.

In summary, T cell targets frequently detected in natural HCV gt3 infection show high levels of variability between genotypes, with limited cross-reactivity against sequence variants both within genotype 3 and across genotypes, suggesting that they might not serve as attractive targets for cross-reactive vaccine design. Further supporting this hypothesis, it was previously shown that cross-reactivity profiles at dominant gt1 epitopes did not vary between cells primed in natural infection and vaccine-induced T cells (Fytili et al. 2008); therefore, it is likely that the exclusive use of frequently detected HCV gt3 specific T cell targets in an immunogen aiming to prime cross-reactive responses will not be applicable. Leading from this data, it might be beneficial for the induction of cross-reactive immune responses by a T cell vaccine to exclude variable regions. Thereby, no T cell responses to immunodominant epitopes falling in variable regions would be induced, and subdominant responses against conserved regions potentially cross-reactive against multiple genotypes may be up-ranked.

5 Immunogen design

5.1 Abstract

Background: Currently, no vaccines exist for Hepatitis C virus (HCV), which is one of the most diverse viral pathogens known and exists as 6 major genotypes with ~80% sequence homology. We have previously shown that immune responses induced by a potent T cell vaccine containing HCV genotype-1b non-structural regions target epitopes dominant in natural infection. However, viral regions corresponding to these epitopes are highly variable at population level; *ex vivo* analysis of vaccine-induced T cells showed poor cross-reactive between variants. We therefore aimed to design a generic algorithmic approach generating cross-genotypic immunogens containing only conserved viral regions.

Methods: A computer algorithm defining sequence variability within sequence datasets was generated, which was applied to HCV sequence datasets obtained from standard database resources and generated in-house. To assess potential immunogenicity of conserved segments under a set threshold, we assessed T cell epitopes described in natural infection and additionally predicted epitopes using epitope prediction programs. Designed immunogens were blasted to exclude potential cross-reactivity with human material.

Results: Immunogens containing viral regions conserved between major HCV genotypes (for 1a/1b, 1/3a, 1-6) were designed. Patient sequences most similar to the segment consensus were selected for final immunogens, ensuring cross-genotypic reactivity. Conserved segments contain multiple epitopes described in natural infection, suggesting *in vivo* immunogenicity of designed constructs; additionally, strong binding peptides were predicted for conserved segments. Predicted artificial epitopes in junctional regions (~20%) were abrogated through the insertion of 2-6 amino acid linkers. To optimize peptide presentation, a TPA leader sequence was included in all immunogens.

Conclusions: This work paves the way for future studies of novel HCV immunogens based on conserved viral segments, and can be applied for other variable pathogens.

5.2 Rationale

5.2.1 Do we need a vaccine for HCV?

Hepatitis C virus infects 170 million people worldwide and is the leading cause of end-stage liver disease and transplantation in the UK (Health Protection Agency 2013). 25% of all HIV infected people are co-infected with HCV, and it is the leading cause of death in HIV+ people on HAART.

HCV treatment has become considerably more effective in the era of directly acting antivirals (DAA) (Schinazi et al. 2014). Although the DAAs are a big step forward in HCV therapy, treatment will not be available to many infected patients; the new drugs are very expensive and even health care systems in developed countries might not be able to afford treatment of all patients, a limitation that is even more pronounced in developing countries (Cammà et al. 2012; Cammà et al. 2013). In addition, most people infected with HCV worldwide are unaware that they carry the virus, preventing them from being treated in the first place. Compliance can be poor in patient groups where HCV incidence is high, such as IVDU. Furthermore, treatment success remains variable in patient groups difficult to treat: previous non-responders and relapsers on interferon, patients with advanced liver disease and decompensated cirrhosis, and patients not infected with HCV genotype 1 (gt1) (see Introduction, 1.3.2.2 Directly acting antivirals (DAA, page 31). Currently, side effects in treatment combinations still administered with interferon are considerable, but are expected to decrease with the new combinations of DAAs. Finally, successful treatment does not prevent re-infection: a problem particularly relevant in an IVDU population.

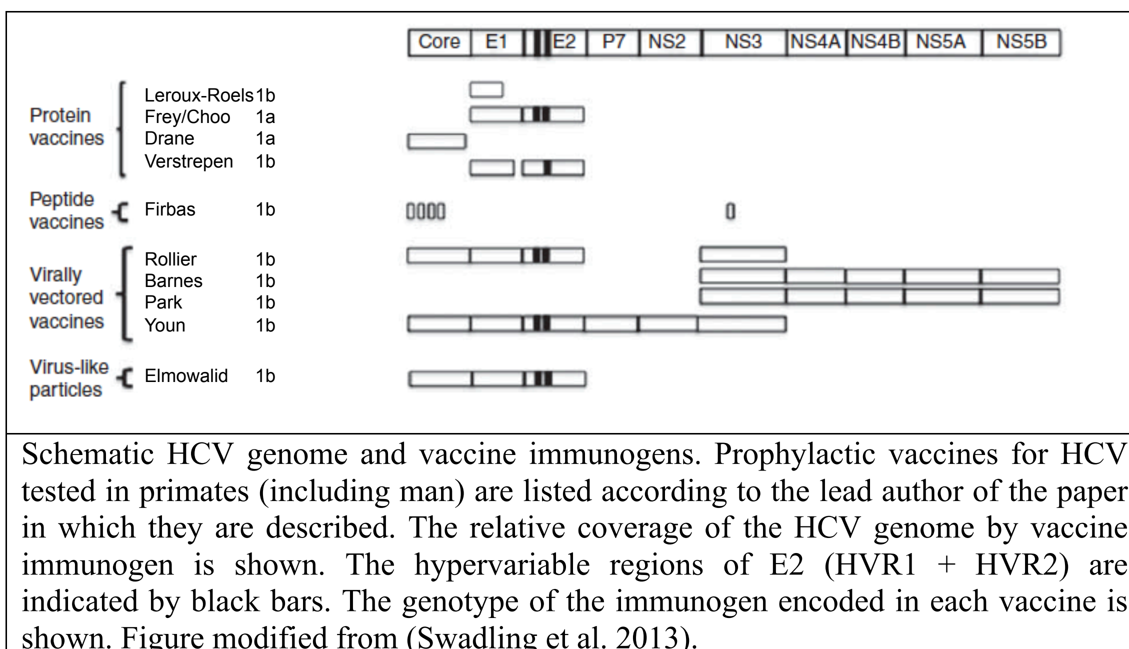
In conclusion, available treatment options do not ensure effective treatment, and most importantly, do not prevent further infections: in lessons learned with HIV infection, we know that even with various treatment options available for more than a decade, treatment does not reach about 7.5 million people infected with HIV (WHO report 2011). Therefore, a HCV vaccine is a pressing need.

5.2.2 Current HCV vaccine candidates

For HCV, several technologies have been investigated for the induction of specific antigens to induce protective immunological memory. Vaccine strategies have been recently reviewed by (Swadling et al. 2013) and (L. M. J. Law et al. 2013) and include (Figure 5-1):

1. Recombinant protein vaccines based on HCV structural regions: core (Drane et al. 2009), E1 (Leroux-Roels et al. 2004) , E1 and E2 (Verstrepen et al. 2011), E1/E2 combined (Choo et al. 1994; Frey et al. 2010);
2. Peptide vaccines, using HLA-A2 epitopes (Firbas et al. 2006; Firbas et al. 2010);
3. Virally vectored vaccines encoding non-structural regions (E Fattori et al. 2006; Folgori et al. 2006; S.-H. Park et al. 2012; Barnes et al. 2012), core-NS3 (Youn et al. 2008), or structural proteins core-E2 plus NS3 (Rollier et al. 2007);
4. Virus-like particles encoding structural proteins (Elmowalid et al. 2007)

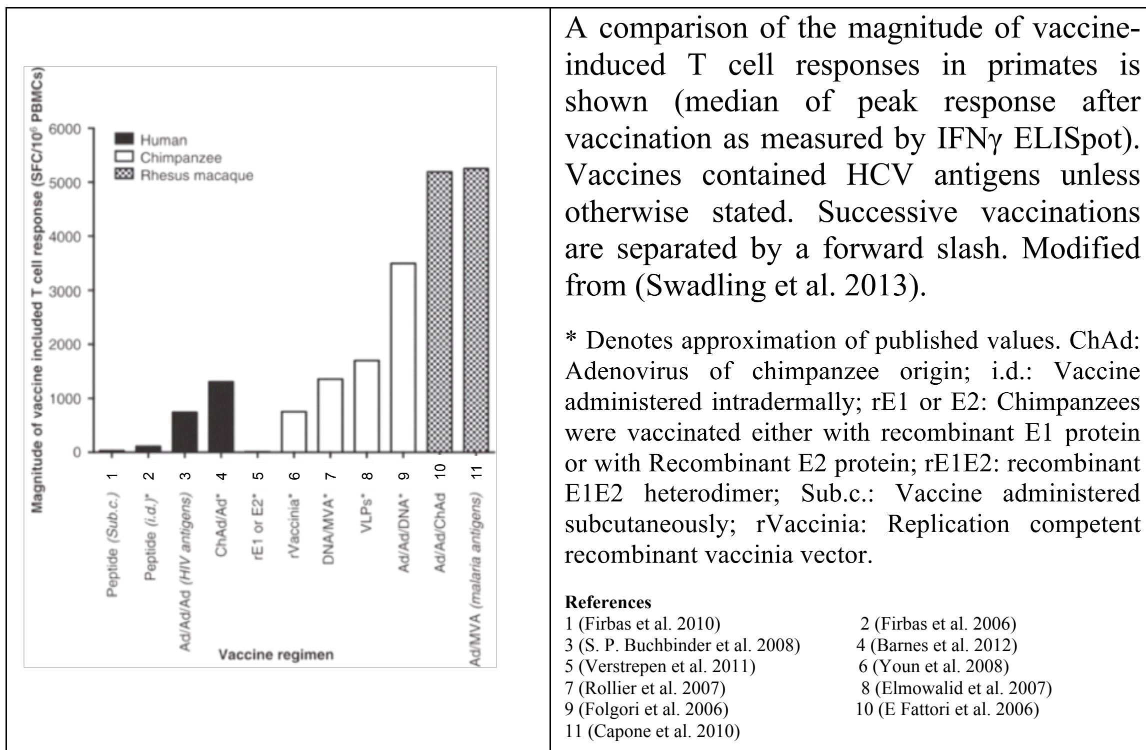
Figure 5-1: HCV vaccine candidates



Schematic HCV genome and vaccine immunogens. Prophylactic vaccines for HCV tested in primates (including man) are listed according to the lead author of the paper in which they are described. The relative coverage of the HCV genome by vaccine immunogen is shown. The hypervariable regions of E2 (HVR1 + HVR2) are indicated by black bars. The genotype of the immunogen encoded in each vaccine is shown. Figure modified from (Swadling et al. 2013).

When comparing the magnitude of IFNγ T cell responses, as a surrogate marker of immune response, a wide range of magnitudes have been observed using approaches discussed above (Figure 5-2): highest responses were seen against virally vectored vaccines in Rhesus macaques. In humans, highest responses were seen using combinations of different viral adenovectors (Barnes et al. 2012). Even higher responses ranging up to over 6000 SFU/million PBMC were seen using a combination of adenovectors and MVA in humans (manuscript in preparation, Swadling et al.).

Figure 5-2: Magnitude of T cell responses to different vaccine approaches.



A comparison of the magnitude of vaccine-induced T cell responses in primates is shown (median of peak response after vaccination as measured by IFNγ ELISpot). Vaccines contained HCV antigens unless otherwise stated. Successive vaccinations are separated by a forward slash. Modified from (Swadling et al. 2013).

* Denotes approximation of published values. ChAd: Adenovirus of chimpanzee origin; i.d.: Vaccine administered intradermally; rE1 or E2: Chimpanzees were vaccinated either with recombinant E1 protein or with Recombinant E2 protein; rE1E2: recombinant E1E2 heterodimer; Sub.c.: Vaccine administered subcutaneously; rVaccinia: Replication competent recombinant vaccinia vector.

References
1 (Firbas et al. 2010) 2 (Firbas et al. 2006)
3 (S. P. Buchbinder et al. 2008) 4 (Barnes et al. 2012)
5 (Verstrepen et al. 2011) 6 (Youn et al. 2008)
7 (Rollier et al. 2007) 8 (Elmowalid et al. 2007)
9 (Folgori et al. 2006) 10 (E Fattori et al. 2006)
11 (Capone et al. 2010)

Even though the magnitude of immune response represents a fundamental characteristic of T cell responses in clinical vaccine studies, it does not necessarily correlate with immune protection (Rollier et al. 2007). Other parameters like breadth of response, the profile of T cells including cytokine production, cytotoxicity, phenotype, and immune exhaustion are known to play an additional important role, but are difficult to compare between studies.

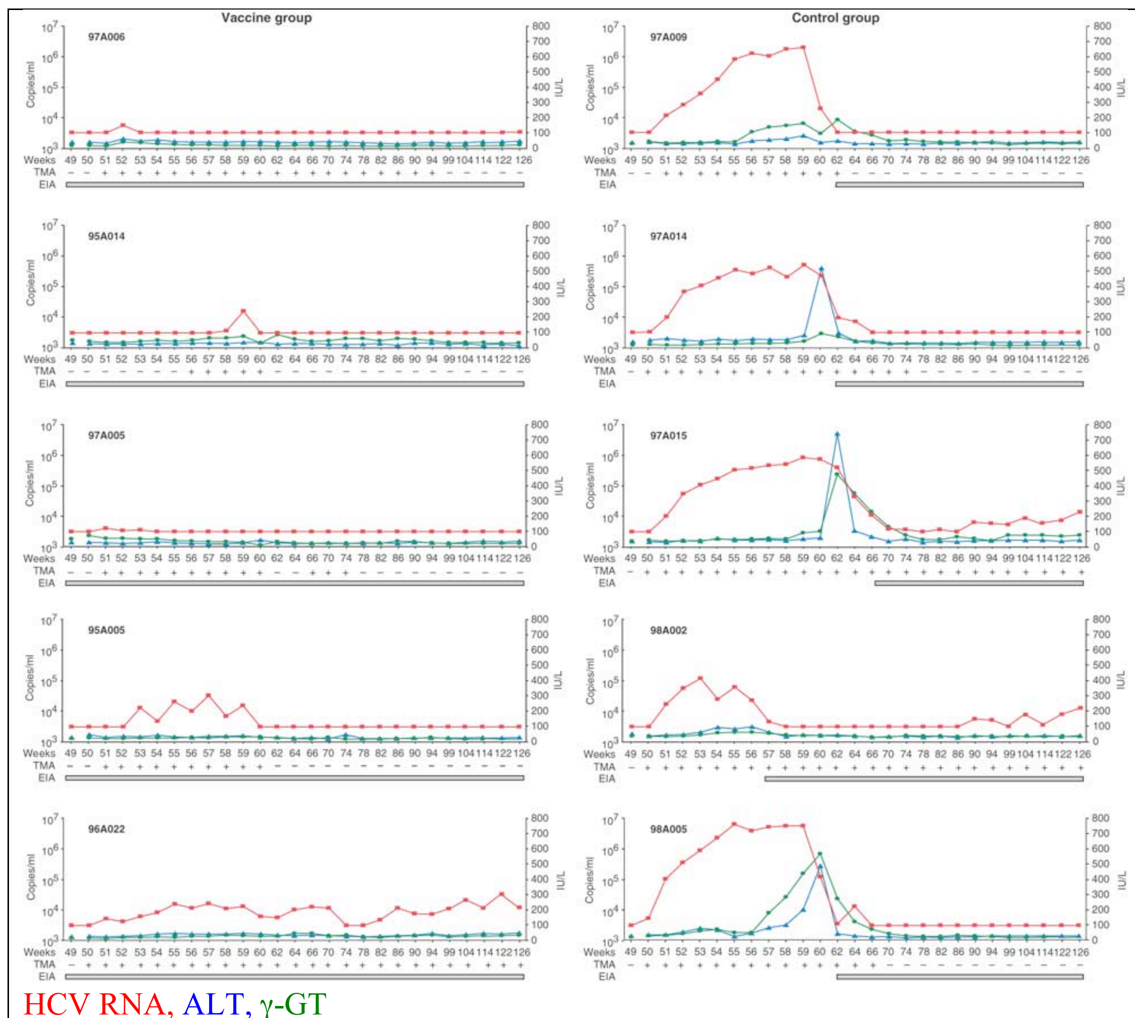
Of note, all strategies exploited to date are based on HCV gt1 immunogens, and cross-reactivity remains one of the challenges to be addressed.

5.2.3 Why develop a T cell vaccine for HCV?

There is plenty of evidence that the adaptive immune response plays a crucial role in HCV infection: We know that a robust and broad T cell response plays a crucial role in the clearance of acute infection (Lechner et al. 2000; Lauer et al. 2004; Lucas et al. 2007), and it has been shown that HLA types A03, B57 and B27 are associated with clearance of acute HCV infection (Neumann-Haefelin et al. 2010; Fitzmaurice et al. 2011). On the other hand, a down-regulation of T cell responses, T cell exhaustion, viral escape and suppression by regulatory T cells is associated with viral persistence (reviewed by (Neumann-Haefelin et al. 2007)). If $CD8^+$ specific T cell responses in chimpanzees that previously cleared HCV infection are blocked using specific

antibodies, re-infection with HCV resulted in prolonged viral replication (Shoukry et al. 2003). Finally, strong evidence for the importance of T cells for viral clearance originates from prophylactic vaccine data (Adeno/DNA) in a chimpanzee challenge model (Folgori et al. 2006). In this experiment, vaccinated animals developed no or very low viremia and no elevation of liver parameters, whereas all control animals had marked rises of liver parameters and substantial viremia (Figure 5-3).

Figure 5-3: HCV viral load and liver parameters in animals vaccinated with a prophylactic vaccine and control animals



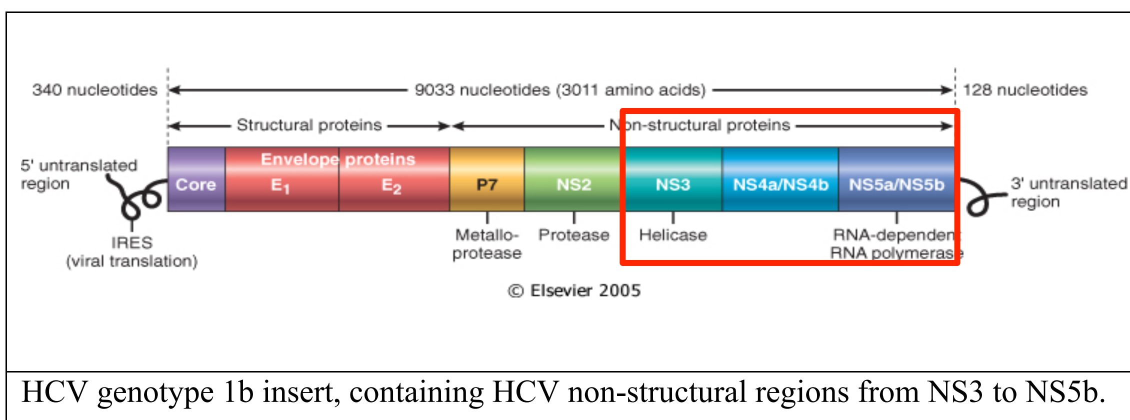
HCV RNA, ALT, γ-GT

The course of acute HCV infection (18 months follow up) is shown for animals vaccinated with a vaccine containing an HCV immunogen in an adenoviral vector (left panels) and control chimpanzees (right panels). Red lines represent the kinetics of HCV RNA measured in plasma samples using branched DNA assay, blue and green lines represent kinetics of ALT and GGT liver enzymes, respectively. Blue and green dots at T = 49 weeks represent the average ALT and GGT values calculated from 10 time points before challenge. Results of the qualitative HCV RNA assay (TMA; detection limit, 50 copies/ml) are reported as + or – at the bottom of each graph. Anti-HCV seroconversion is indicated by a grey bar across the bottom of each graph. Figure and legend modified from (Folgori et al. 2006).

This evidence has lead to the assessment of a HCV vaccine inducing T cell responses in humans.

In our research group, the clinical Phase-I trial using a HCV genotype 1b T cell vaccine in healthy volunteers and chronically HCV infected patients was started in 2011. The immunogenic insert contained HCV non-stuctural regions from NS3 to NS5b, including a mutation in the NS5b protein preventing replication (Figure 5-4) (Barnes et al. 2012).

Figure 5-4: HCV genotype 1b insert



HCV genotype 1b insert, containing HCV non-structural regions from NS3 to NS5b.

The HCV gt1b insert was incorporated in Adenoviral and MVA vectors, and given to healthy volunteers and patients in different regimens:

1. HCV 001 Adenoviral prime and boost in healthy volunteers (Barnes et al. 2012);
2. HCV 002 Adenoviral prime and boost in chronically infected individuals (therapeutic vaccine) with standard of care therapy (peg-INF-α/Ribavirin) (Manuscript in preparation, Kelly et al.);
3. HCV 003 Adenoviral prime and MVA boost in healthy volunteers and chronically infected patients (Manuscript in preparation, Swadling et al.).

Immune responses to the vaccine insert were assessed using ELISpot assays, and strong HCV genotype 1 responses were observed in healthy volunteers (Barnes, Folgori, et al. 2009; Barnes et al. 2012).

5.2.4 Cross-reactivity of HCV genotype 1b vaccine-induced responses

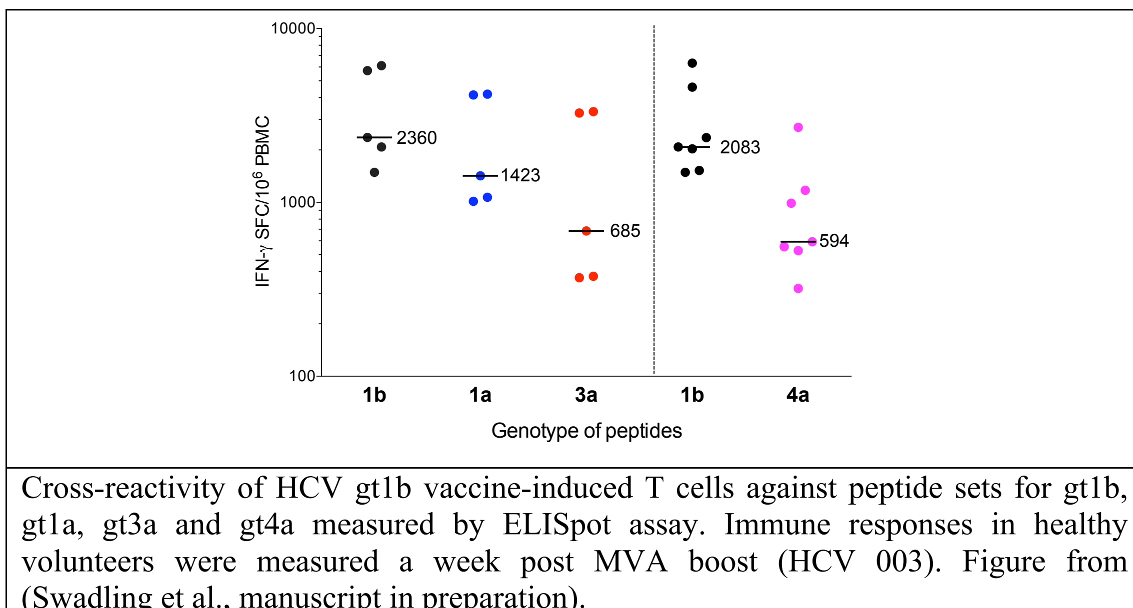
HCV is one of the most variable pathogens known, and exists in both multiple viral genotypes and quasi-species within each infected individual (see introduction; page 18, HCV genotypes and variability of the HCV genome). An effective HCV T cell

vaccine should induce T cells that target multiple viral strains in conserved regions, and therefore should be effective against multiple subtypes and quasi-species of the virus. The T cell vaccine described above was therefore tested against both peptide sets of different genotypes (HCV vaccine trial 001 and 003) and patient circulating viral sequences (HCV vaccine trial 002).

5.2.4.1 Genotype cross-reactivity of vaccine-induced responses

Cross-reactivity of vaccine-induced T cells in healthy volunteers (HCV 003) was assessed using peptide sets for genotype 1b (J4 strain), genotype 1a (H77 strain), genotype 3a (K3a strain), and genotype 4a (ED43 strain; EMBL accession Y11604) (Figure 5-5): In comparison to HCV gt1b specific responses, lower immune responses were detected to HCV gt1a, 3a and 4a peptide sets. (Swadling et al., manuscript in preparation). This decrease was most likely due to sequence differences in immunodominant epitopes in genotypes 1a, 3a and 4a compared to the priming genotype 1b sequence. This suggests limited cross-reactivity of HCV gt1b vaccine-primed T cells to other genotypes.

Figure 5-5: Cross-reactivity of immune responses against HCV genotype 1b, 1a, 3a and 4a peptide sets in patients vaccinated with a HCV genotype 1b insert.



Cross-reactivity of HCV gt1b vaccine-induced T cells against peptide sets for gt1b, gt1a, gt3a and gt4a measured by ELISpot assay. Immune responses in healthy volunteers were measured a week post MVA boost (HCV 003). Figure from (Swadling et al., manuscript in preparation).

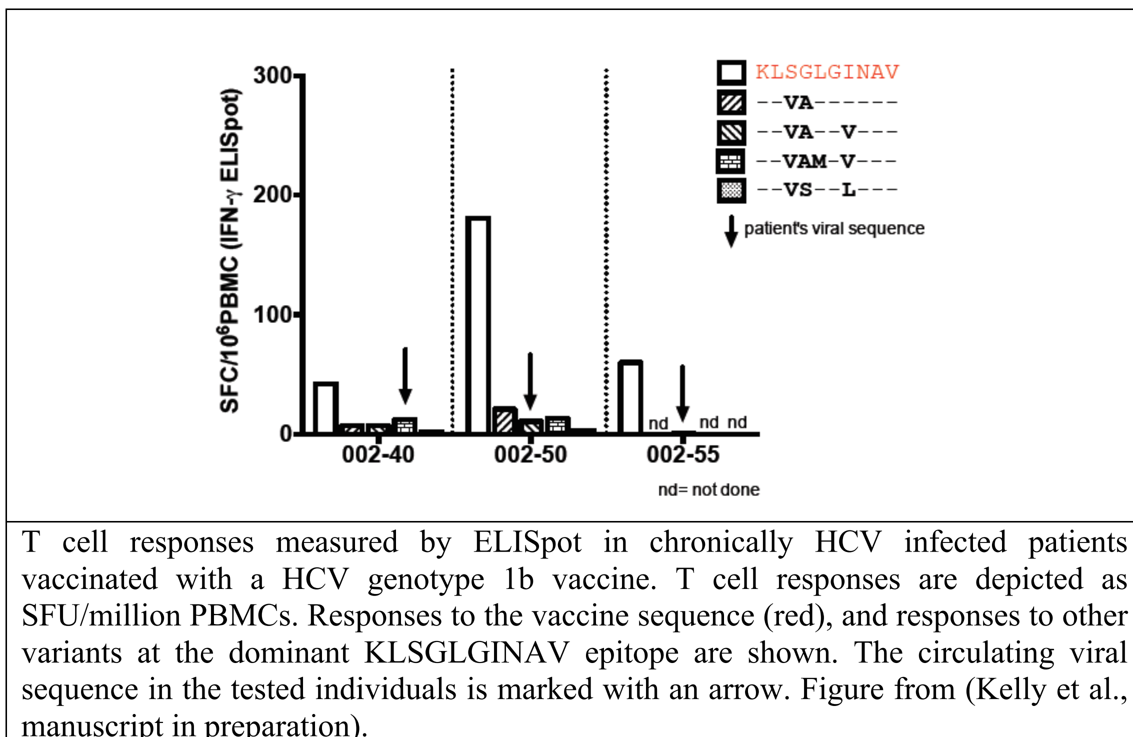
5.2.4.2 Cross-reactivity against quasi-species in a single host

HCV does not only exist in different genotypes, but also in different circulating viral quasi-species in each infected individual (S. Duffy et al. 2008). The abundance of multiple sequence variants especially at dominant epitopes, is due to a rapidly mutating virus. The HCV polymerase does not contain a proofreading function,

resulting in multiple viral copies, of which some are replication efficient. Additionally, T cell pressure towards immunodominant epitopes can result in immune mediated escape variants, further increasing viral variability. Therefore, even in HCV gt1b patients vaccinated with the therapeutic vaccine (HCV 002), T cells primed with a gt1b vaccine immunogen may not necessarily target circulating virus.

To test whether vaccine-induced immune responses in chronically infected patients target circulating virus, the HCV gt1b vaccine was administered in a prime/boost regimen to patients chronically infected with HCV genotype 1 (HCV 002). Some of the most dominant primed T cell responses were directed towards highly variable epitopes with little evidence of cross-reactivity between variants. Interferon (INF)-γ production of vaccine-induced T cells was assessed using ELISpot assays, measuring immune responses against the vaccine sequence and patient's circulating viral sequence at dominant epitopes (Figure 5-6).

Figure 5-6: Cross-reactivity of vaccine-induced T cell responses against circulating patient variants.



T cell responses measured by ELISpot in chronically HCV infected patients vaccinated with a HCV genotype 1b vaccine. T cell responses are depicted as SFU/million PBMCs. Responses to the vaccine sequence (red), and responses to other variants at the dominant KLSGLGINAV epitope are shown. The circulating viral sequence in the tested individuals is marked with an arrow. Figure from (Kelly et al., manuscript in preparation).

Vaccine-induced T cells showed a strong response against the epitope variant contained in the vaccine, but little or no IFN-γ production was observed against other common epitope variant peptides at this epitope. Of note, even the response against the epitope circulating in the patients was barely detectable, suggesting either a low T

cell response against the infecting virus or a low cross-reactivity of vaccine-induced T cells.

Overall, circulating viral sequences at dominant epitopes targeted in the HCV 002 trial were markedly different to the HCV 1b sequence used as a vaccine immunogen (Table 5-6). This reflects that the inserted HCV 1b strain did not contain epitope variants most abundant in the population.

Table 5-1: Comparison of vaccine epitope sequences at immunodominant epitopes with circulating viral sequences in patients.

T cell response detected to immunogen peptide	Circulating viral sequence in patient at baseline	Patient ID
KLSGLGINAV	KLVSLGLNAV	27
	KLVALGINAV	28
	KLVALGVNAV	50
	KLVALGVNAV	55
	KLVALGINAV	103
	KLVAMGVNAV	40
CVNGVCWTV	CINGVCWTC	27
	CINGVCWTC	55
ALYDVVSTL	ALYDVVSKL	27
HSKKKCDEL	HSKRKCDEK	45
SVVIVGRIIL	CVVIVGRVVL	27
QEFDEMEECASHLPY	REFDEMEECSQHLPY	27
APNYSRALWRVAAEE	APNYTFALWRVSAEE	39
VTLTHPITKYIMACM	IILTHPITKYIMACM	45
FQVGLNQYLVGSQLP	FRVGLHDYPVGSQLP	39

T cell responses detected in HCV infected individuals vaccinated with a HCV gt1b vaccine in comparison to circulating viral strain sequences at immunogenic epitopes. In all patients with detectable T cell responses, the sequence of the vaccine-induced T cell epitopes does not match the circulating viral sequence. Figure from: (Kelly et al., manuscript in preparation)

In summary, from clinical and experimental data evaluating the HCV gt1b vaccine, there are two lines of evidence supporting the need of a cross-reactive vaccine:

(1) Lack of genotype cross-reactivity by vaccine-induced T cells; and
(2) Lack of quasi-species cross-reactivity within an infected host by vaccine-induced T cells.
To overcome these limitations of a single genotype vaccine - inducing responses to immunodominant, but not cross-reactive epitopes - we therefore aimed to design a cross-reactive HCV vaccine.

5.2.5 Limited cross-reactivity of HCV gt3 and gt1 subtype specific T cell responses in natural infection

In previous Chapters of this thesis, I have demonstrated that T cell specificity is distinct between HCV subtypes 1 and 3 (Chapter 3). In addition, I have shown that there is limited experimental T cell cross-reactivity at T cell targets frequently detected in HCV gt3 infection, both between viral sequence variants within HCV gt3 and across HCV genotypes 1 and 3 (Chapter 4). This suggests that T cell targets primed in natural HCV infection, particularly those detected in chronically infected individuals, might not serve as attractive targets for cross-reactive vaccine design.

5.2.6 Rationale for a HCV vaccine based on conserved viral regions

Several options are possible to address the highly variable nature of HCV through immunogen design.

A. Use common sequence variants at immunodominant epitopes.

One option to generate vaccine-induced responses targeting a wide range of circulating strains is choosing a common sequence variant at known immunodominant epitopes. However, in a highly variable pathogen like HCV, it is difficult to choose a common sequence variant between different genotypes, since (i) common sequence variants vary massively between genotypes, as demonstrated for HCV gt1 and gt3 in Chapters 3 and 4 of this thesis; (ii) some immunodominant epitopes (such as NS3 1406 for HCV gt1) show major sequence variability even within a single genotype; and (iii) T cell specificity, and therefore immunodominant epitopes, are likely to be different in all HCV genotypes. This is to be confirmed for HCV genotypes 2, 4, 5 and 6, in which immunogenicity has not been analysed in detail, and in which different immunodominant epitopes might be defined. In addition, even if common sequence variants are defined for each immunogenic epitope, it is not guaranteed that they will be capable of priming immune responses *in vivo*.

B. Choice of a cross-reactive version at each immunodominant epitope.

This approach is based on evaluating different sequence variants for the capability to prime cross-reactive T cells. These variants may either occur in natural infection, or may be generated *in silico* using computer prediction algorithms, as previously obtained for HIV epitopes (W. Fischer et al. 2007). Capability of epitopes to prime T cell responses potentially cross-reactive between viral variants and genotypes may be assessed *in vivo* using vaccines, or, more applicable, in *in vitro* models using dendritic cells to present peptide variants and subsequently assessing cross-reactivity of primed T cells (see Kelly et al., manuscript in submission). However, this approach is very labour intensive, and it is not guaranteed that a cross-reactive version exists at each immunodominant epitope.

C. Use of ancestral sequences

This option was previously evaluated in HCV infection. A vaccine design based on ancestral sequences was evaluated using this rational approach, with the aim of maximizing immunogenicity and minimizing genetic distance across circulating strains (Burke et al. 2012). Three different approaches were compared: a combination of epitopes with predicted tight MHC binding, a consensus sequence and a representative ancestral sequence derived using Bayesian phylogenetic tools; $CD8^+$ T cells expanded with representative ancestral sequence recognized diverse circulating HCV strains better than T cells expanded with consensus sequence or naturally occurring sequence variants (Burke et al. 2012).

D. Exclusion of variable HCV epitopes

Another option is to exclude variable epitopes from the design, and only choose highly conserved epitopes or conserved viral regions for immunogen design.

Based on the evidence discussed above, we hypothesized that T cell vaccines targeting the most conserved regions of the HCV proteome, which are common to most subtypes and bear fitness costs when mutated, will generate effector T cells that efficiently recognise and kill virus-infected cells early enough after transmission to potentially impact on HCV replication and will do so more efficiently than T cell vaccines using the whole viral genome as a vaccine insert. However, a common concern with using conserved regions in immunogen design is that they may not be capable of priming potent T cell responses based on poor HLA avidity, or insufficient processing and presentation. I hypothesize that even if these regions are sub-dominant in priming immune responses in natural infection, they may be capable of priming

strong cross-reactive T cell responses once epitopes immunodominant in natural infection are removed from the immunogen, an effect previously observed for a conserved HIV immunogen (E.-J. Im et al. 2011).

For other viruses, especially for HIV, some of these approaches for cross-reactive vaccine have been exploited: Approaches aiming to generate cross-reactive HIV vaccines were based on an choosing the most conserved sequence variant at each amino acid, termed "mosaic T cell approach" (W. Fischer et al. 2007; Thurmond et al. 2008; Stephenson et al. 2012), vaccines based on HIV Gag conserved regions (Kulkarni et al. 2014), conserved "promiscuous" CD4+ T cell epitopes (Ribeiro et al. 2010; D. S. Rosa et al. 2011; R. R. Almeida et al. 2012), conserved B cell epitopes, or on viral regions conserved between genotypes.
A vaccine consisting of conserved regions of HIV (HIV_{cons}) designed by the Hanke group at the Jenner Institute (Oxford University) was pre-clinically tested in mice (Létourneau et al. 2007; Ondondo et al. 2013), macaques (SIV.consv, (Koopman et al. 2013)) and a recent Phase I clinical trial (Borthwick et al. 2013). The vaccine immunogen was designed using conserved regions between major HIV-1 clades A, B, C, and D, containing a consensus sequence between the four clades. The included segments alternate between the different clades to ensure equal coverage. In mice and macaques, significant immune responses to the immunogen were seen in mice using different vaccine vectors and administration methods (Koopman et al. 2013; Rosario et al. 2012). In humans, the vaccine was administered using DNA, simian adenovirus and MVA vector strategies, with encouraging results: high levels of induced effector T cells that recognised virus-infected autologous $CD4^+$ cells and inhibited HIV replication (Borthwick et al. 2013), with effector T cells typically subdominant in natural infection mediating the virus inhibition.
A cross-reactive immunogen design approach similar to the approach exploited by Tom Hanke's groups has also been attempted for Dengue virus, and the immunogenicity of this construct was subsequently tested in mice (Khan et al. 2006; Khan et al. 2008).
In addition, an epitope based vaccine design approach aiming to generate broadly neutralizing antibodies was described for respiratory syncytial virus (RSV), which was based on a computational method to design scaffold proteins with full backbone flexibility (Correia et al. 2014).

5.2.7 Aims of this work

Based on this evidence, we aimed to design a cross-reactive vaccine for HCV based on conserved viral regions. Ideally, a single vaccine would prevent infection with HCV genotypes 1 to 6, ensuring that the vaccine could be used in multiple continents. HCV genotype 7 was not included in the vaccine design, since only one single sequence was described for this subtype to date (D. B. Smith et al. 2013). In the UK, both HCV genotype 1 and genotype 3 are common infecting strains, and therefore a preventative vaccine would protect against both subtypes.

The aims of this Chapter were to:

(1) Define sequence regions conserved between HCV genotypes
(2) Select conserved HCV sequence segments for conserved immunogen design
(3) Design HCV immunogens for genotypes 1a/b, 1/3a and 1-6
(4) Evaluate potential immunogenicity of conserved regions and junction areas through
    a. Computational approaches using epitope prediction servers
    b. Assessment of previously described epitopes in natural infection falling into conserved regions
(5) Optimize designed immunogens for:
    a. Potential artificial, non-HCV epitopes in junctional regions

5.3 Programming approach for conserved immunogen design

To design a vaccine consisting of only conserved segments, it was necessary to define conserved segments within a sequence dataset, and select appropriate conserved segments for subsequent immunogen design. Work described in this section (5.3) was done in collaboration with Jose Lourenco (Zoology, Oxford).

5.3.1 Sequence selection for HCV genotypes 1 to 6

First, a sequence dataset had to be selected. For this, sequences from the HCV LOS ALAMOS HCV sequence database and additional in-house sequences for HCV genotype 3a were used. To ensure equal representation of differing strains, for each genotype sequences were selected from different research groups and countries. To design 3 different immunogens for (1) HCV gt1, (2) HCV gt1 and 3 and (3) HCV gt1-6, we generated 3 different sequence datasets for further analysis (Table 5-2). 36 sequences for each included subtype (HCV genotypes 1 to 6) were selected, considering different geographic regions, different research groups and genotype subtypes. Due to the lack of a sufficient number of genotype 5 sequences, two different analyses had to be run: one only using the 3 available full length genotype 5 sequence, with a resulting under-representation of genotype 5 in the resulting dataset, and a second analysis using the 3 available sequence sets multiplied by 12, resulting in an equal number of input sequences, but in a lower variability.

Table 5-2: Sequence selection for immunogen design, overview.

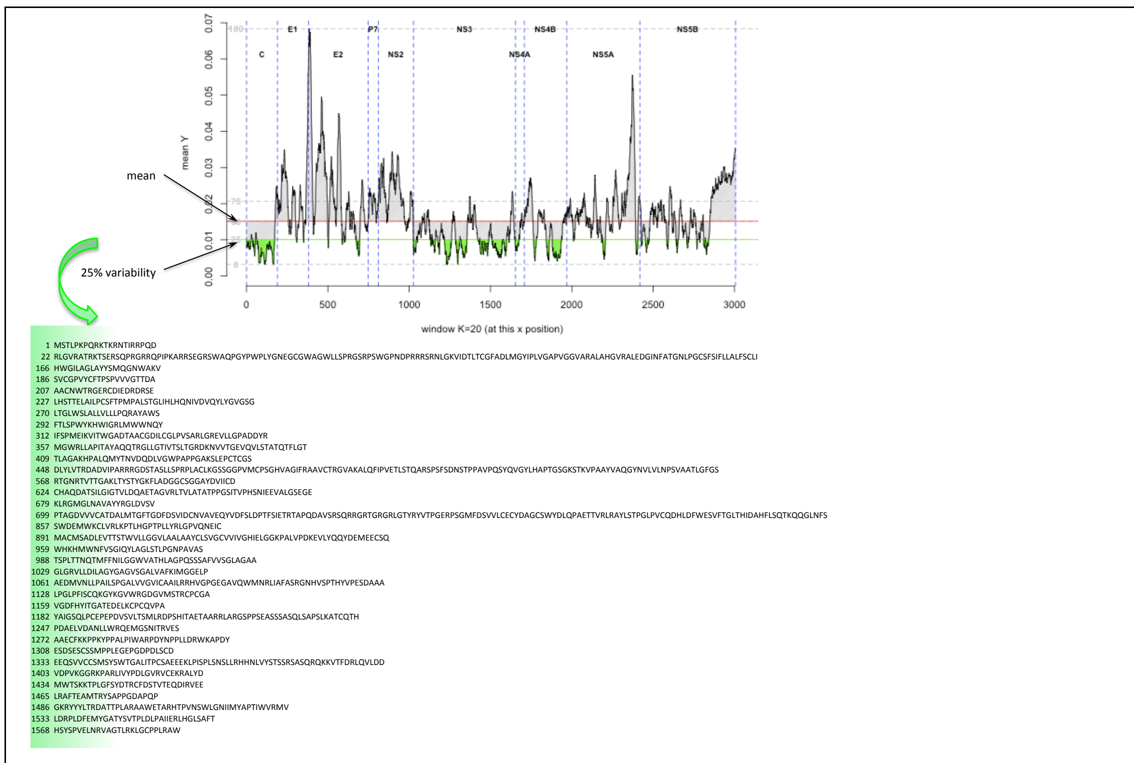
Immunogen	Included subtypes	Number of sequences
HCV gt1	1a	48
	1b	48
HCV gt1/3a	1a/1b	36
	3a	36
HCV gt1-6	1a/1b	36
	2a/2b	36
	3a	36
	4a/4other	36
	5	3*12
	6a/6other	36

Sequence datasets selected for immunogens for HCV gt1, HCV gt1/3a, HCV gt1-6, including equal numbers of sequences for each subtype included in the according immunogen design. Since only 3 full-length sequences were available for HCV gt5, this sequence material was multiplied to ensure equal sequence representation for this subtype in the HCV gt 1-6 dataset.

5.3.2 Calculating sequence diversity within a sequence dataset

Sequence diversity was calculated using pairwise comparisons within a certain window size. To design the immunogens containing conserved regions, a window size of 20 AA was used, aiming to capture both CD4 and CD8 epitopes. Most of these epitopes vary between 8-12 AA in length for CD8+ T cell epitopes, and 11-16 AA for CD4 cells. Subsequently, amino acids within certain window sizes are compared using pairwise comparisons, and a mean variability threshold for each amino acid is calculated (Figure 5-7). To define conserved segments, a cut off at 25% of the variability within each dataset was chosen (marked green in Figure 5-7). An example set of sequences of conserved segments is shown underneath the variability plot. Between 25 and 30 conserved segments were defined for the different analyses: 30 for the GT1a/1b dataset, 27 for the GT1/3a and 25 for the GT1-6 dataset.

Figure 5-7: Sequence diversity of the entire HCV genome between HCV genotypes 1 and 3.



Calculated sequence diversity for an example sequence dataset (HCV gt1/3a, containing 72 sequences) using a window size of 20, shown for the full HCV genome. For conserved vaccine design, segments with a variability <25% (lowest quartile), marked in green, were defined as conserved and selected for immunogen design and are depicted as consensus sequences.

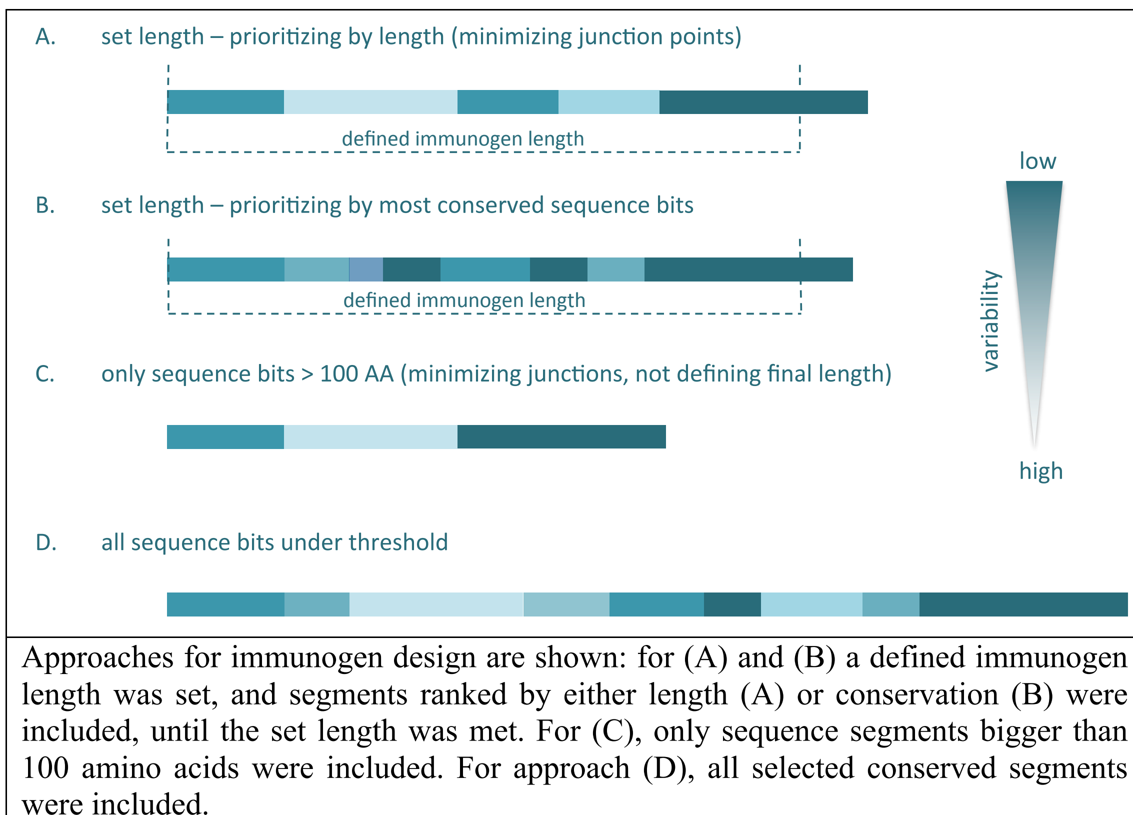
5.3.3 Approaches for immunogen design

The final strategy for immunogen design was based on several key considerations:

1. By joining previously non-related sequences segments, artificial (non-HCV) epitopes might be generated.
2. Highly conserved segments might be very short, which would cause multiple junction sites if all segments were included.
3. Long immunogens are often difficult to manufacture, and therefore, segment selection might be constrained by the total immunogen length.

Therefore, four approaches (A to D) for immunogen design were considered (Figure 5-8). Selection of conserved segments to be included in the final immunogen were prioritized by either (A) individual segment length; (B) most conserved segments; (C) including only sequence segments over a certain length; or (D) all segments under the set threshold.

Figure 5-8: Approaches for immunogen design



Approaches for immunogen design are shown: for (A) and (B) a defined immunogen length was set, and segments ranked by either length (A) or conservation (B) were included, until the set length was met. For (C), only sequence segments bigger than 100 amino acids were included. For approach (D), all selected conserved segments were included.

Considerations for favouring either of the immunogen approaches are detailed in Table 5-3.

Table 5-3: Potential advantages/ disadvantages for different immunogen designs

A	+	• Minimizing junction sides, less artificial epitopes • Shorter immunogens: easier to produce
	-	• Not all segments included: important immunogenic epitopes might be missing
B	+	• Highly conserved segments included: generated immune responses most likely cross-reactive
	-	• Segments tend to be very short: many junctions and thereby many artificial epitopes
C	+	• Minimizing junction sides, less artificial epitopes
	-	• Comparably short immunogens, potentially restricting the number of presented epitopes
D	+	• All selected sequence segments included: maximisation of immunogenic regions and thereby potential epitopes
	-	• Long immunogens: difficult to manufacture • Multiple junction sides, many artificial epitopes

To address considerations detailed in Table 5-3, a short pilot epitope prediction experiment was run for all approaches with datasets for HCV gt1a, gt1b and 3a using the online prediction server NetMHC (for details on epitope prediction, refer to: 5.5.2.3 and 5.5.3, page 236 and 238). Immunogens were assessed for immunogen length, potential immunogenicity and possible artificial, non-HCV epitopes at junction sites between segments. Results for this analysis are shown in Figure 5-9.

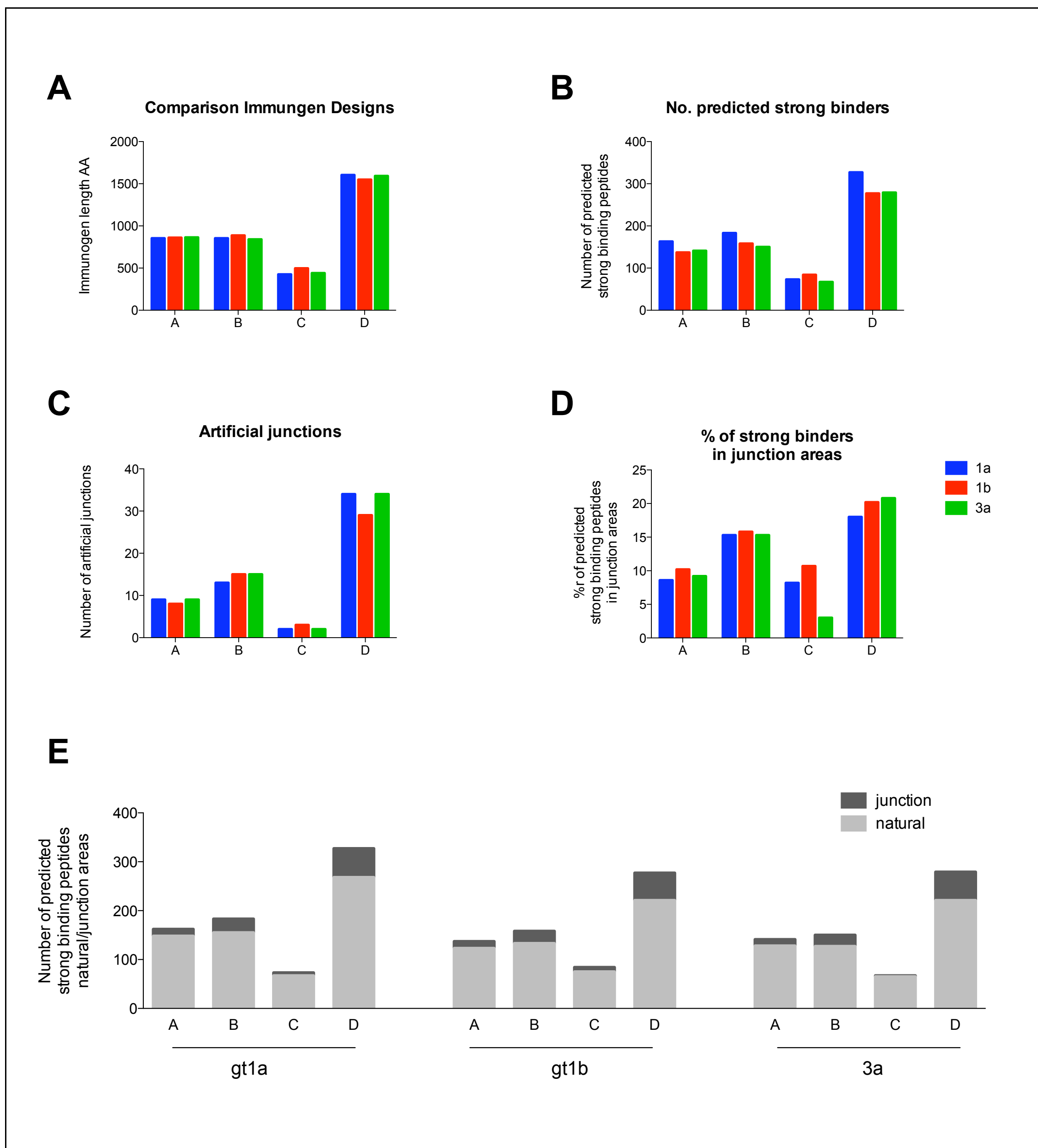
Based on this data, two immunogen design approaches were excluded from further analysis:

Approach B was excluded because of a high percentage of artificial epitopes in comparison to Approach A, most likely caused by a high proportion of segment junction sites;

Approach C was excluded because of the potential shortage of immunogenic material.

Based on these considerations, we decided to only include approach A and D in further analyses and the final immunogen design.

Figure 5-9: Pilot epitope prediction analysis for immunogens for HCV gt1a, 1b and 3a designed by approaches A to D.



Results of online epitope prediction pilot experiment using the NetMHC prediction server: Immunogen design approaches were compared regarding (A) immunogen length, (B) potential immunogenicity assessed by number of predicted strong binders (NetMHC score <0.5), (C) number of artificial junctions in each immunogen, (D) the percentage of strong binders falling into junction areas potentially forming artificial epitopes, and (E) number of predicted strong binders in segment areas forming natural HCV epitopes, and strong binders in junction areas forming potential artificial, non-HCV epitopes.

5.3.4 Comparison of conserved sequence segments for HCV gt1, HCV gt1/3a and HCV gt1-6 datasets

Subsequently, conserved segments for all three analyses (HCV gt1, HCV gt1/3a and HCV gt1-6) were compared (Figure 5-10). Significant overlap between conserved segments was observed; 1090 of all amino acids positions were shared between the three analyses, with an overlap of 1090/1544 (71%) for the HCV gt1 immunogen, 1090/1444 (75%) for HCV gt1/3a and 1377 (79%) for HCV gt1-6, respectively.

Figure 5-10: Comparison of conserved segments for HCV gt1, HCV gt1/3a and HCV gt1-6 datasets (Figure next pages)

Consensus sequences of conserved segments for all three analyses (HCV gt1, HCV gt1/3a, HCV gt1-6) are shown. Non-conserved regions are depicted as – for each amino acid. HCV viral protein regions are colour coded. Bold sequences represent segments that were included in Immunogen approach A (see 5.3.3 Approaches for immunogen design, page 213).

```
GT1       MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARA
GT1and3   MSTLPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRSWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPNDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGVARA
GT1-6     ----------TKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRSWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPNDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPVGGVARA
Core

GT1       LAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPASA------------------------------------------------------------------------------------YVGDLCGSVFLVSQLFTFSPR----
GT1and3   LAHGVRALEDGVNFATGNLPGCSFSIFLLALLSCLTHPAS-----------------------------------------------------------------------VDLLVGAATMCSALYVGDLCG------------------
GT1-6     LAHGVRVLEDGINYATGNLPGCSFSIFLLALLSCLTVPAS--------------------------------------------------------------------------------------------------------------
Core/E1

GT1       --------YPGHITGHRMAWDMMMNWSP--------------------------------------------------------------------------------------------------------------------------
GT1and3   ---------------------------------------------------HWGVLAGLAYYSMQGNWAKV-------------------------------------------------------------------------------
GT1-6     ---CNCSIYPGHITGHRMAWDMMMNWSPTT--------------------------------------------------------------------------------------NGSWHINRTALNCNDSLNTGFI------------
E1/E2

GT1       -----------------------------------------------------VCGPVYCFTPSPVVVGTTDR------------TDVFLLNNTRPPLGNWFGCTWM------------------------------CPTDCFRKHPEAT
GT1and3   ----------------------------------------------------SVCGPVYCFTPSPVVVGTTDR-----------------------------------------------------------------------------
GT1-6     ----------------------------------------------------SVCGPVYCFTPSPVVVGTTDR----------------------------------------------------------------CPTDCFRKHPEAT
E2

GT1       YSRCGSGPWLTPRCLVDYPYRLW---------------------------------------DRDRSELSPLLLSTTEWQVLPCSFTTLPALSTGLIHLHQNIVDVQYLYGVGSSV-------------------ARVCACLWMMLLIAQ
GT1and3   --RCGSGPWLTPRCLVDYPYRLWHYPCT--------------------AACNWTRGERCDIEDRDRSEL------TTELAILPCSFTPLPALSTGLIHLHQNIVDVQYLYGVGSGM----------------------------------
GT1-6     YTKCGSGPWLTPRCLVDYPYRLWHYPCTVNF----------------------------------------LLHSTTEWAILPCSFTTLPALSTGLIHLHQNIVDVQYLYGVGS------------------------------------
E2

GT1       AEAALENLV---------------------------------------------------------------------------------------------------------------------------------------------
GT1and3   ------------------------------------------------------------------------------------------------------------------------------------------------------
GT1-6     ------------------------------------------------------------------------------------------------------------------------------------------------------
E2/[illegible]/NS2

GT1       -------------------------------------------------------------------------------------------------------------------------------QGWRLLAPITAYAQQTRGLLGCI
GT1and3   -------------------------------------------------------------------------------------------------------------------------------MGWRLLAPITAYAQQTRGLLGCI
GT1-6     -------------------------------------------------------------------------------------------------------------------------------KGWRLLAPITAYAQQTRGLLGTI
NS2/[illegible]

GT1       ITSLTGRDKNQVEGEVQIVST-QSFLATCINGVCWTVYHGAG------------------------------------PCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAV
GT1and3   VTSLTGRDKNVVEGEVQVLSTA--------------------------------------------------------PCTCGSSDLYLVTRHADVIPARRRGDSRASLLSPRPLAYLKGSSGGPLLCPSGHVAGIFRAAVCTRGVAKAL
GT1-6     VTSLTGRDKN--------------------------------------------------------------------PCTCGSSDLYLVTRHADVIPARRRGDSRASLLSPRPISTLKGSSGGPVLCPSGHVVGIFRAAVCTRGVAKAL
NS3

GT1       DFIPVESLETTMRSPVFTDNSSPPAVPQTFQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGV-----TGVRTITTGAPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIGTVLDQAETAGARL
GT1and3   QFIPVETL----RSPSFSDNSTPPAVPQSYQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKA--------TGNRTITTGAKITYSTYGKFLADGGCSGGAYDIIICDECHSQDATSILGIGTVLDQAETAGARL
GT1-6     DFIPVESLE--MRSPVFTDNSTPPAVPQTYQVGHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAYGI----RTGVRTVTTGAPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATTILGIGTVLDQAETAGVRL
NS3

GT1       VVLATATPPGSVTVPHPNIEEV-----------------------------------------------NAVAYYRGLDVSVIPTSGDVVVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTLPQDAVSRTQRRGRTGR
GT1and3   TVLATATPPGSITVPHSNIEEVAL--------------------------------------------------------SVIPTSGDVVVCATDALMTGFTGDFDSVIDCNTCVEQTVDFSLDPTFSIETTTAPQDAVSRSQRRGRTGR
GT1-6     VVLATATPPG-------------------------------IKGGRHLIFCHSKKKCDELA--------NAVAYYRGLDVSVIPTSGDVVVCATDALMTGFTGDFDSVIDCNVAVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGR
NS3

GT1       GR---YRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYLNTPGLPVCQDHLEFWEGVFTGLTHIDAHFLSQTKQAGENFPYLVAYQATVCARAQAPPPSWDQMWKCLIRLKP--------------------------K
GT1and3   GRLGTYRFVTPGERPSGMFDSVVLCECYDAGCAWYELQPAETTVRLRAYLNTPGLPVCQDHLEFWESVFTGLTHIDAHFLSQTKQQGLNFPYLVAYQATVCARAQAPPPSWDEMWKCLVRLKPTLHGPTPLLYRLGAVQN----------
GT1-6     GRLGIYRYVSPGERPSGMFDSVVLCECYDAGCAWYELTPAETTVRLRAYLNTPGLPVCQDHLEFWEGVFTGLTHIDAHFLSQTKQGGENFPYLVAYQATVCARAKAPPPSWDTMW-----------------------------THPITK
NS3
```

```
GT1            YIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRIVLSG---------------------------------------------------------------------FWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAF
GT1and3        --MACMSADLEVTTSTWVLLGGVLAALAAYCLSTGCVVIVGH--------------------------------------------------------------------------FWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAF
GT1-6          YIMACMSADLEVVTSTWVLVGGVLAALAAYCLSVGCVVIVG---------------------------------------------------------------------------FWAKHMWNFISGIQYLAGLSTLPGNPAVASLMAF
NS3/NS4a/NS4b

GT1            TAAVTSPLTTQ-----------------------------------GSIGLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAAR
GT1and3        TA-------------------------------------------IGSIGLGKVLVDILAGYGAGVAGALVAFKIMGGELPSTEDMVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAAR
GT1-6          TAAVTSPL----------------------------------GAAVGSIGLGKVLVDILAGYGAGVAGALVAFKIMSGEKPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAAR
NS4b

GT1            VTQILS-------------------------SWLRDIWDWICEVLSDFKTWLK-------------------------------------------------------------------------------------------------
GT1and3        VTAILSSLT---------------------------------------------------------------------------------------------------------------------------------------------
GT1-6          V-----------------------------------------------------------------------------------------------------------------------------------------------------
NS4b/NS5a

GT1            ------------------CPCQVPSPEFFTELDGVRLHR-------------------------------------------------------RRLARGSPPSLASSSASQLSAPSLKATCTTNH------------------------
GT1and3        ------------------CPCQVPAPEFFTEVDGVRLHR-----------------------IGSQLPCEPEPDVSVLTSML---------TAARRLARGSPPSEASSSASQLSAPSLKATCQTHH------------------------
GT1-6          ----------------------------------------------------------------------------------------TAETAARRLARGSPPSLASSSASQLSAPSLKATCT---------------------------
NS5a

GT1            ------------------------------------------------------------------------------------------------------------------------------------------------SDAESY
GT1and3        -----------------------------------------------------------------------------------------------------------------------------------------------ESDSESC
GT1-6          ---------------------------------------------------------------------------------------------------------------------------------------------------ESY
NS5a

GT1            SSMPPLEGEPGDPDLSDGSWSTVSSEE-----------------------------------------------------LSNSLLRHHNMVYATTSRSA---------------------------------------------LTPPH
GT1and3        SSMPPLEGEPGDPDLSDGSWSTVSD------------------------QDVVCCSMSYSWTGALITPCAAEEEKLPINPLSNSLLRHHNLVYSTSSRSASQRQKKVTFDRLQVLDDHYK------------------------------
GT1-6          SSMPPLEGEPGDPDLEFEQVESQ-----------------------------VCCSMSYSWTGALITPCAAEEEKLPINPLSNSLLRHHNLVYSTTSRSASLRQKKVTFDR---------------------------------------
NS5a/NS5b

GT1            SAKSKFGYGAKDVRCH-------------------------------------------KPARLIVFPDLGVRVCEKMALYDVV----------------------------------PMGFSYDTRCFDSTVTESDIRTE---------
GT1and3        -----------------------------------------------------PEKGGRKPARLIVYPDLGVRVCEKMALYDV-----------------------------WTSKKTPMGFSYDTRCFDSTVTEQDIRVEEE-------
GT1-6          -----------------------------------------------------PSKGGRKPARLIVYPDLGVRVCEKRALYDV--------------------------------KKTPMGFSYDTRCFDSTVTERDIRTE---------
NS5b

GT1            ---------SLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKASAACRAAKL--CTMLVCGDDLVVICESAGTQEDA-SLRAFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTTPLARA
GT1and3        ----------------------------CGYRRCRASGVLPTSCGNTLTCYIKATAAC-------------------------------ALRAFTEAMTRYSAPPGDAPQP--------------------KGKRVYYLTRDPTTPLARA
GT1-6          ----------------------------CGYRRCRASGVLTTSMGNTLTCYIKALAA--------------------------------------EAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHD--------------TPLARA
NS5b

GT1            AWETARHTPVNSWLGNIIMFAPTLWARMIL-----------------------------EPLDLPQIIQRLHGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRA-----GGRAAICGKYLFNWAVRTKLKLTPIPAA
GT1and3        AWETARHTPVNSWLGNIIMYAPTIWVRMVL----------------------YGATYSVEPLDLPAIIERLHGLSAFSLHSYSPGEINRVAGCLRKLGCPPLRAWRHR---VRAKLLSQGGRAAICGKYLFNWAVRTK---------
GT1-6          AWETARHTPVNSWLGNIIMYAPTIWVRMVLMTHFFSILQ-------------YGATYSVTPLDLPAIIQRLHGLSAFSLHSYS--ELNRVAACLRKLGVPPLRAWRHRARAVRAKLIAQGGRAAICGKYLFNWAV------------
NS5b
```

5.4 Sequence selection for final immunogens

5.4.1 Comparison of within-genotype consensus sequences to all-genotype consensus

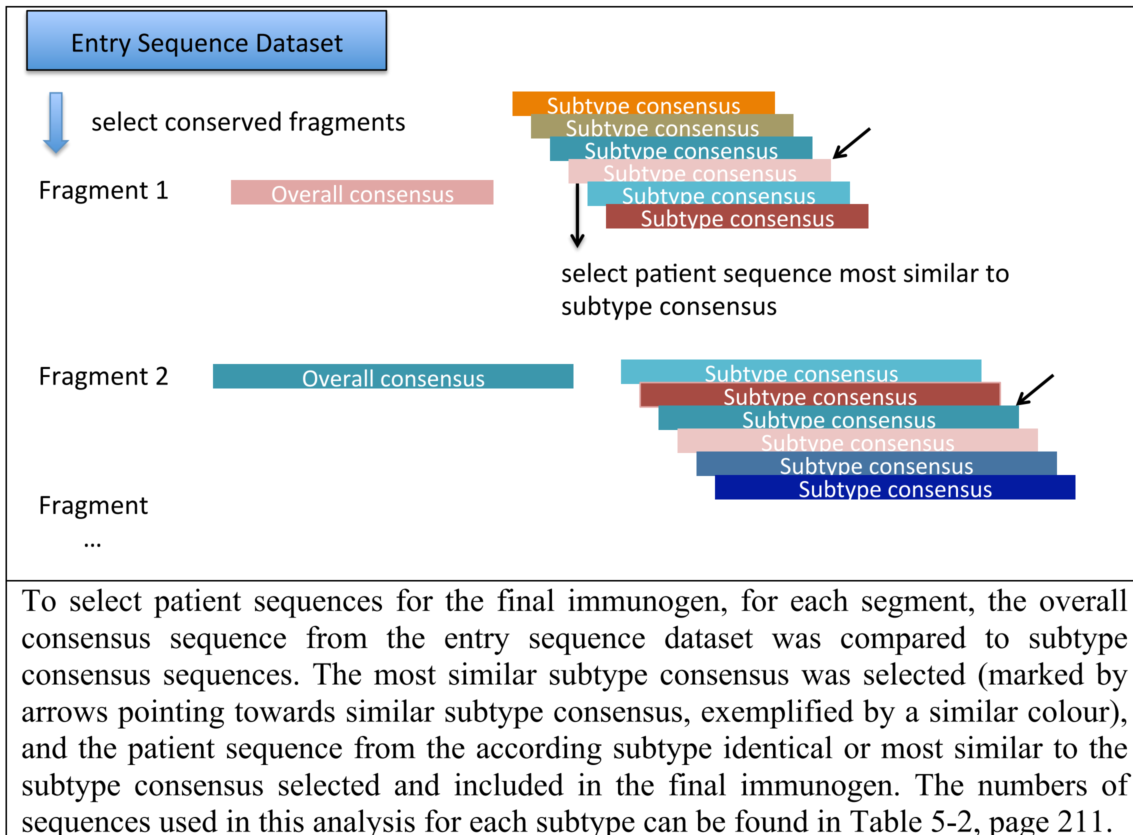
Consensus sequences of conserved segments generated across different HCV genotypes might show major differences to real patient sequences. Subsequently, an immunogen built from consensus sequences might generate immune responses to artificial, non-HCV epitopes.

We therefore decided to only include real patient sequences in our final immunogens that were optimally matched for identity to the "overall consensus" of the conserved segment. The term "overall consensus" was defined as a consensus generated from all sequences included for each immunogen design (for gt1, gt1/3a and gt1-6), whereas the term "subtype consensus" is used for a consensus sequence generated from sequences of single genotype (for example, for only genotype 3 sequences in the gt1-6 analysis).

To choose patient sequences that were included in the final immunogen a selection approach was followed, in which the similarity between overall consensus and subtype consensus sequences of the included genotypes was analysed for each conserved segment (Figure 5-11).

The subtype consensus with the highest sequence similarity to the overall consensus was selected. Next, all patient sequences of the according subtype were compared to the subtype consensus, and an identical or almost identical sequence was selected for the immunogen.

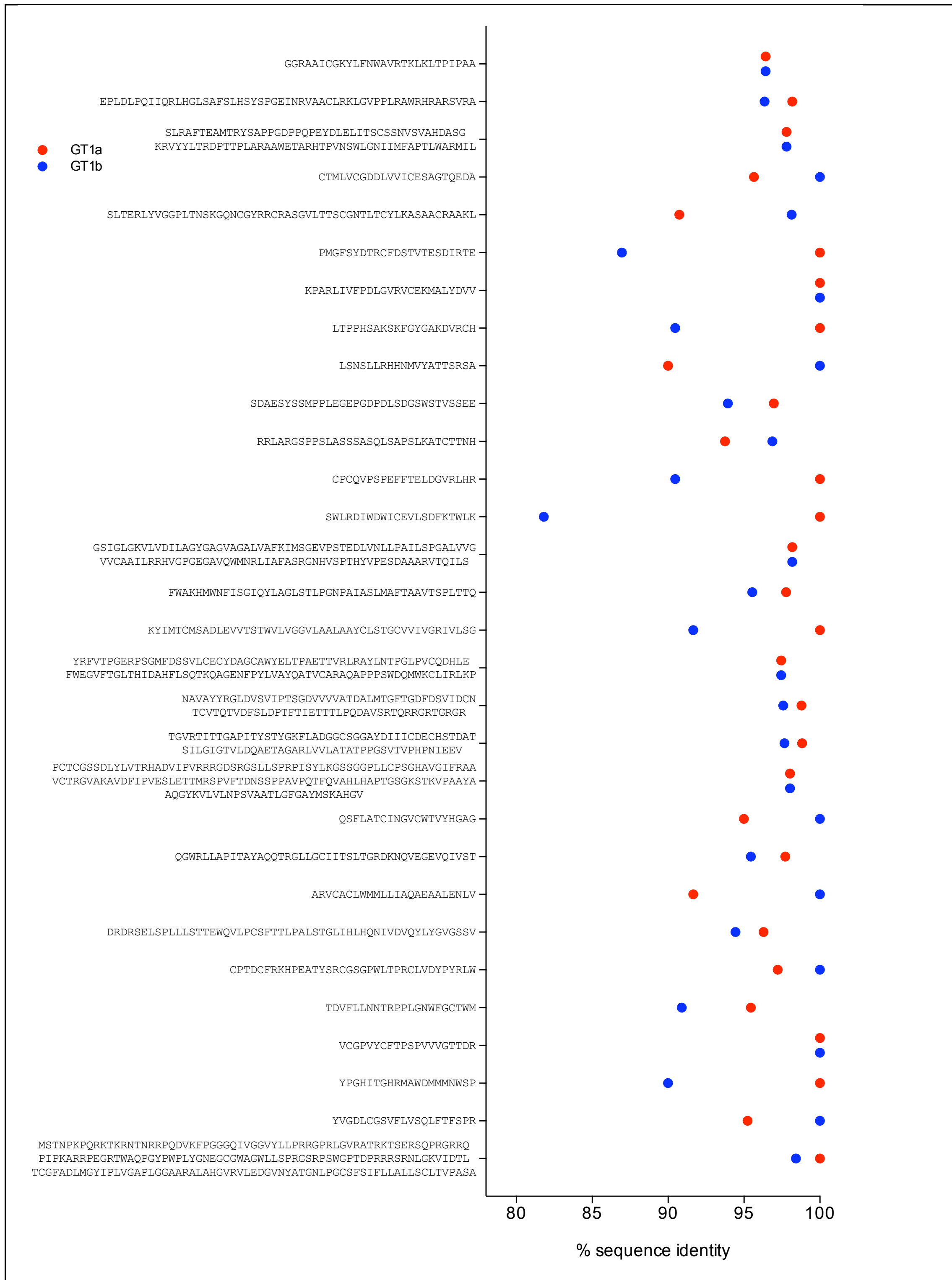
Figure 5-11: Selection process for sequences to be included in the final immunogen



To select patient sequences for the final immunogen, for each segment, the overall consensus sequence from the entry sequence dataset was compared to subtype consensus sequences. The most similar subtype consensus was selected (marked by arrows pointing towards similar subtype consensus, exemplified by a similar colour), and the patient sequence from the according subtype identical or most similar to the subtype consensus selected and included in the final immunogen. The numbers of sequences used in this analysis for each subtype can be found in Table 5-2, page 211.

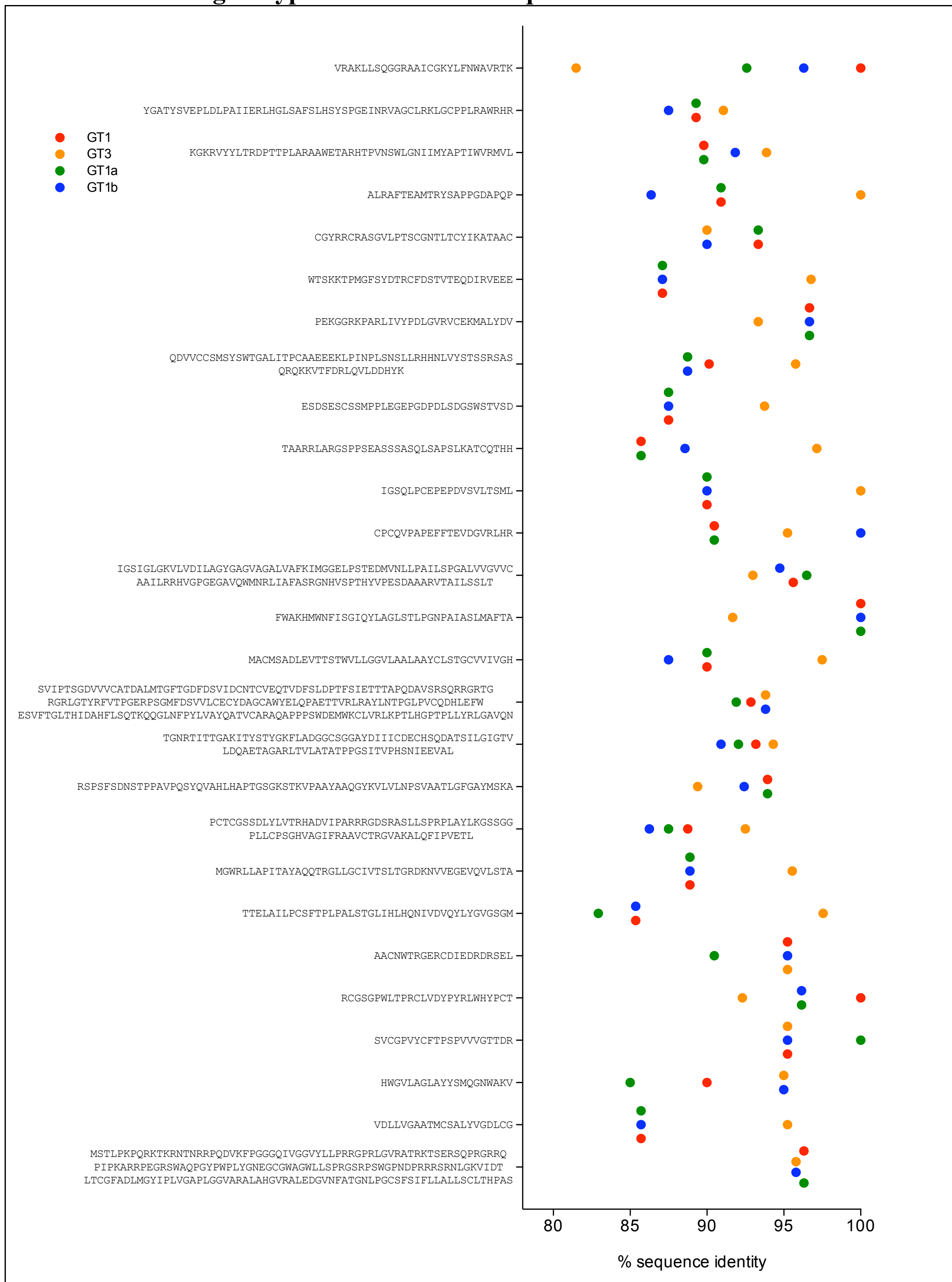
Results of this analysis are shown in Figure 5-12 for the HCV genotype 1 immunogen, Figure 5-13 for the HCV genotype 1 and 3 immunogen, and Figure 5-14 for the HCV genotype 1-6 immunogen.

Figure 5-12: Comparison of HCV genotype 1a and 1b consensus sequences to the overall HCV genotype 1 consensus sequence



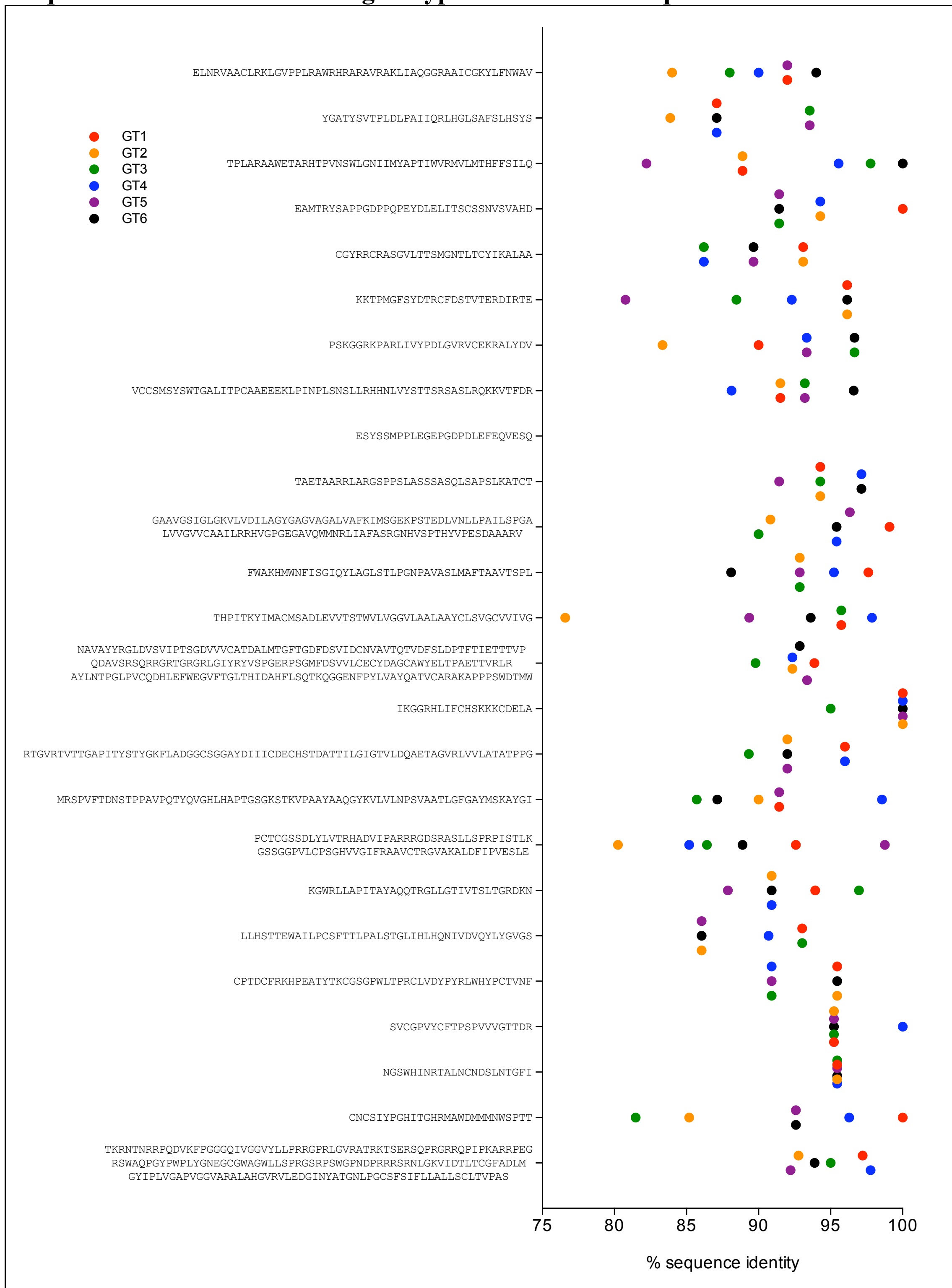
Results of the comparison of HCV gt1a and gt1b consensus sequences to the HCV gt1 consensus are shown. Consensus sequences for conserved segments are shown on the left, with core segments at the bottom of the figure and NS5b segments at the top. For each conserved segment, the percentage of sequence identity with the overall consensus is shown: in red for the HCV gt1a consensus and blue for HCV gt1b consensus.

Figure 5-13: Comparison of HCV genotype 1a, 1b and 3a consensus sequences to the overall HCV genotype 1/3a consensus sequence



Results of the comparison of HCV gt1 and gt3a consensus sequences to the HCV gt1/3a consensus are shown. Consensus sequences for subtypes 1a and 1b were analysed separately. Consensus sequences for conserved segments are shown on the left, with core segments at the bottom of the figure and NS5b segments at the top. For each conserved segment, the percentage of sequence identity with the overall consensus is shown.

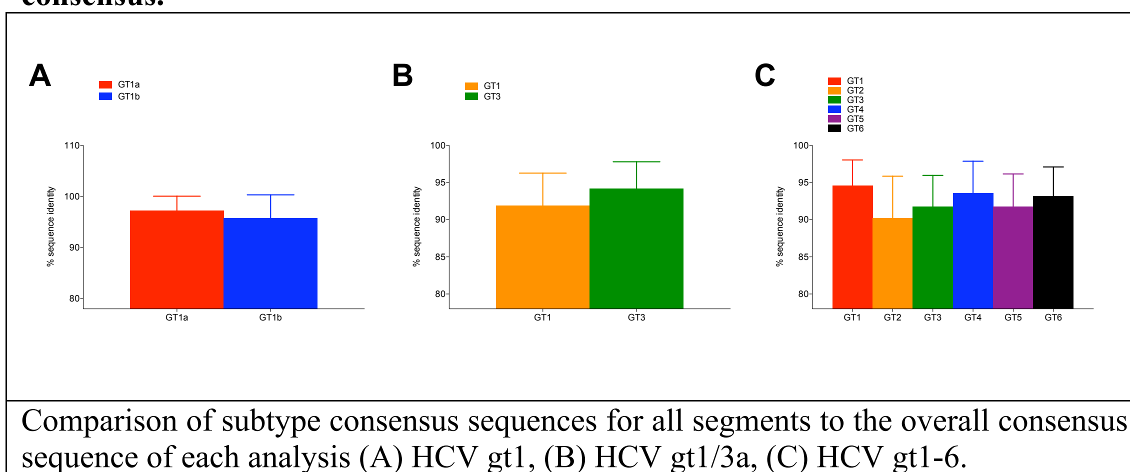
Figure 5-14: Comparison of HCV genotype 1, 2, 3, 4, 5 and 6 consensus sequences to the overall HCV genotype 1-6 consensus sequence



Results of the comparison of HCV gt1-6 subtype consensus sequences to the HCV gt1-6 consensus are shown. Consensus sequences for conserved segments are shown on the left, with core segments at the bottom of the figure and NS5b segments at the top. For each conserved segment, the percentage of sequence identity with the overall consensus is shown.

Next, the percentage of sequence identity of the subtype consensus sequences for each segment was compared to the overall consensus at each segment. This analysis was done using input datasets for HCV gt1, HCV gt1/3a, and HCV gt1-6 (Figure 5-15). When similarity of subtype consensus to overall consensus at each segment was summed, consensus sequences for HCV gt1a are most similar to the overall HCV gt1 consensus (A), HCV gt3a consensus most similar to the HCV gt1/3a consensus (B), and HCV gt1 consensus most similar to the HCV gt1-6 consensus (C).

Figure 5-15: Comparison of subtype consensus (all segments) to overall consensus.



Comparison of subtype consensus sequences for all segments to the overall consensus sequence of each analysis (A) HCV gt1, (B) HCV gt1/3a, (C) HCV gt1-6.

5.4.2 Selection of patient sequences for final immunogen constructs

Patient sequences for the final immunogen have been selected by the following process:

(1) Comparison of overall consensus to all subtype consensuses (5.4.1) to select the most similar.

(2) Comparison of patient sequences of the selected subtype to the subtype consensus. If no 100% match was found, patient sequences of other subtypes were examined and selected, if a higher match to the overall consensus was found.

Selected patient sequences are depicted in Table 5-4 for the immunogen designs ((A) HCV gt1, (B) HCV gt1/3a, (C) HCV gt1-6)). The match percentage between the patient viral sequence and the overall consensus was > 90% for all conserved segments. Segment 17 of the HCV gt1-6 analysis was not included, since this segment was selected as conserved, but sequence information was only available in 2 subtypes (gt4 and gt6), where this region was an insertion.

Table 5-4: Patient sequences selected for final HCV gt1 immunogen.

Patient sequences selected for the final immunogen are shown for each conserved segment defined for immunogen designs for (A) HCV gt1, (B) HCV gt1/3a, (C) HCV gt1-6). For each conserved segment, the segment identifier, starting position and HCV viral region are shown. The subtype with the consensus sequence most closely related to the overall consensus is given, as well as the selected patient sequence identifier, the match percentage between patient sequence and overall consensus and the included patient sequence. If the segment was included in the shorter immunogen (Approach A), the segment is marked in the column labelled "A".

A HCV genotype 1 immunogen

segment	starting position	region	selected subtype	selected pt.sequence	match % (pt/overall)	subtype consensus sequence	A
1	1	core	GT1a	1a.EU862823.DE.HCV-1a/DE/BID-V24/2003._	100	MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERS QPRGRRQPIPKARRPEGRTWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDP RRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGN LPGCSFSIFLLALLSCLTVPASA	x
2	276	E1	GT1b	1b.EU256059.US.HCV-1b/US/BID-V	100	YVGDLCGSVFLVSQLFTFSPR	
3	309	E1	GT1a	1a.EU862823.DE.HCV-1a_DE_BID-V24_2003._	100	YPGHITGHRMAWDMMMNWSP	
4	502	E2	GT1a/GT1b	1b.EU256061.US.HCV-1b_US_BID-V364_2006._	100	VCGPVYCFTPSPVVVGTTDR	
5	534	E2	GT1a	1a.EU862841.US.HCV-1a_US_BID-V48_2003._	100	TDVFLLNNTRPPLGNWFGCTWM	
6	581	E2	GT1b	1a.EU362899.US.1013q.1013	100	CPTDCFRKHPEATYSRCGSGPWLTPRCLVDYPYRLW	
7	656	E2	GT1a	1a.EU687193.DE.HCV-1a_DE_BID-V1562_2005._	98	DRDRSELSPLLLSTTQWQVLPCSFTTLPALSTGLIHLHQNIVDVQYLYGVGSSV	x
8	729	E2/p7	GT1b	1b.AB049088.JP.HCVT094._	100	ARVCACLWMMLLIAQAEAALENLV	
9	1021	NS3	GT1a	1a.EU255975.US.HCV-1a_US_BID-V183_1994._	98	KGWRLLAPITAYAQQTRGLLGCIITSLTGRDKNQVEGEVQIVST	
10	1067	NS3	GT1a	1b.EU482859.CH.HCV-1b_CH_BID-V272_2003._	100	QSFLATCINGVCWTVYHGAG	
11	1122	NS3	both	1b.JN120912.FR.Lex._	99	PCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPLLCPSGHAV GIFRAAVCTRGVAKAVDFIPVESMETTMRSPVFTDNSSPPAVPQTFQVAHLHAPTG SGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAHGV	x
12	1280	NS3	GT1a	1a.EU862836.CH.HCV-1a_CH_BID-V244_2003._	99	TGVRTITTGSPITYSTYGKFLADGGCSGGAYDIIICDECHSTDATSILGIGTVLDQ AETAGARLVVLATATPPGSVTVPHPNIEEV	x
13	1413	NS3	GT1a	1a.EU529681.CH.HCV-1a_CH_BID-V251_2005._	100	NAVAYYRGLDVSVIPTSGDVVVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDP TFTIETTTLPQDAVSRTQRRGRTGRGR	x
14	1499	NS3	both	1b.EU256059.US.HCV-1b_US_BID-V220_2005._	99	YRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETTVRLRAYLNTPGLPVCQDH LEFWEGVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPPSWDQMWKC LIRLKP	x
15	1643	NS3/NS4a	GT1a	1a.EU155233.US.HCV-1a_US_BID-V445_2006._	100	KYIMTCMSADLEVVTSTWVLVGGVLAALAAYCLSTGCVVIVGRIVLSG	x
16	1760	NS4b	GT1a	1b.EU155220.US.HCV-1b_US_BID-V147_2004._	98	FWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTASVTSPLTTQ	
17	1840	NS4b	both	1b.AF165051.JP.MD4-0.MD-4	100	GSIGLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVG VVCAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTQILS	x
18	1975	NS5a	GT1a	1a.EF032886.BR.BR601.BR601	100	SWLRDIWDWICEVLSDFKTWLK	
19	2112	NS5a	GT1a	1a.EU362876.US.1013_FU24.1013	100	CPCQVPSPEFFTELDGVRLHR	
20	2188	NS5a	GT1b	1b.AF165051.JP.MD4-0.MD-4	100	RRLARGSPPSLASSSASQLSAPSLKATCTTNH	
21	2380	NS5a	GT1a	1a.DQ838745.US.LIV23-11.LIV23	97	SDAESYSSMPPLEGEPGDPDLSDGSWSTVSSEA	
22	2445	NS5b	GT1b	1b.U45476.DE.HD-1._	100	LSNSLLRHHNMVYATTSRSA	
23	2511	NS5b	GT1a	1a.AF011753.US.H77-H21.Patient	100	LTPPHSAKSKFGYGAKDVRCH	
24	2575	NS5b	both	1b.EU256059.US.HCV-1b_US_BID-V220_2005._	100	KPARLIVFPDLGVRVCEKMALYDVV	
25	2634	NS5b	GT1a	1a.EU155233.US.HCV-1a_US_BID-V445_2006._	100	PMGFSYDTRCFDSTVTESDIRTE	
26	2675	NS5b	GT1b	1b.AB154179.IE.No--4.P2	100	SLTERLYVGGPLTNSKGQNCGYRRCRASGVLTTSCGNTLTCYLKASAACRAAKL	x
27	2731	NS5b	GT1b	1b.D50480.JP.HCV-K1-R1.Patient	100	CTMLVCGDDLVVICESAGTQEDA	
28	2755	NS5b	both	1b.AB049088.JP.HCVT094._	99	SLRAFTEAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHDASGKRVYYLTRDPTT PLARAAWETARHTPVNSWLGNIIMYAPTLWARMIL	x
29	2875	NS5b	GT1a	1a.EU529681.CH.HCV-1a_CH_BID-V251_2005._	100	EPLDLPQIIQRLHGLSAFSLHSYSPGEINRVAACLRKLGVPPLRAWRHRARSVRA	x
30	2935	NS5b	both	1b.AB442220.JP.KAH5._	100	GGRAAICGKYLFNWAVRTKLKLTPIPAA	

B HCV genotype 1/3a immunogen

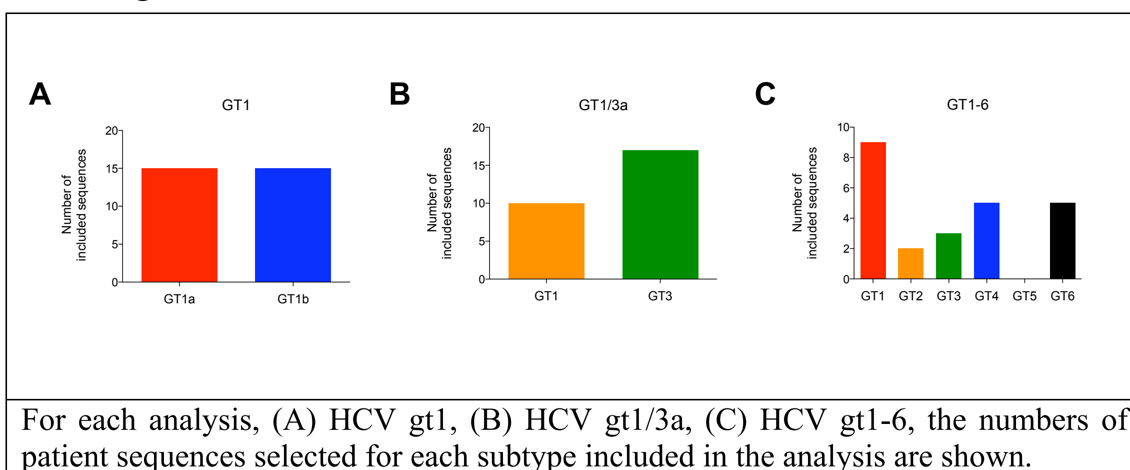
segment	starting position	region	selected subtype	selected pt.sequence	match % (pt/overall)	subtype consensus sequence	A
1	1	core	GT1	1a.EF032886.BR.BR601.BR601	97	MSTNPKPQRKTKRNTNRRPQDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRSWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPTDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPLGGAARALAHGVRVLEDGVNYATGNLPGCSFSIFLLALLSCLTVPAS	x
2	262	E1	GT3	3a.GQ356206.GB.Patient	95	VDLLVGAATMCSALYVGDMCG	
3	352	E1	GT3	3a.GQ356202.GB.Patient	100	HWGVLAGLAYYSMQGNWAKV	
4	503	E2	GT1	1a.D10749.JP.HC-J1.HC-J1	100	SVCGPVYCFTPSPVVVGTTDR	
5	603	E2	GT1	1a.EU862841.US.HCV-1a_US_BID-V	100	RCGSGPWLTPRCLVDYPYRLWHYPCT	
6	649	E2	both	3a.GQ356201.GB.Patient	100	AACNWTRGERCDIEDRDRSEL	
7	676	E2	GT3	3a.JQ717257.IN.ILBSRAS6._	100	TTELAILPCSFTPLPALSTGLIHLHQNIVDVQYLYGVGSGM	x
8	1028	NS2/3	GT3	3a.GQ356204.GB.Patient	96	MGWRLLAPITAYAQQTRGLLGTIVTSLTGRDKNVVTGEVQVLSTA	x
9	1129	NS3	GT3	3a.GQ356200.GB.Patient	93	PCTCGSADLYLVTRDADVIPARRRGDSTASLLSPRPLACLKGSSGGPVMCPSGHVAGIFRAAVCTRGVAKALQFIPVETL	x
10	1213	NS3	GT1	1a.EU256072.CH.HCV-1a_CH_BID-V	97	RSPVFSDNSSPPAVPQSYQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKA	x
11	1287	NS3	GT3	3a.GQ356215.GB.Patient	95	TGNRTITTGAKLTYSTYGKFLADGGCSGGAYDVIICDECHAQDATSILGIGTVLDQAETAGVRLTVLATATPPGSITVPHSNIEEVAL	x
12	1431	NS3	GT3	3a.JN714194.IN.RASILBS2-SR-PO.	95	SVIPTAGDVVVCATDALMTGFTGDFDSVIDCNVAVEQYVDFSLDPTFSIETRTAPQDAVSRSQRRGRTGRGRLGTYRYVGPGERPSGMFDSVVLCECYDAGCAWYELQPAETTVRLRAYLSTPGLPVCQDHLDFWESVFTGLTHIDAHFLSQTKQQGLNFPYLTAYQATVCARAQAPPPSWDEMWKCLVRLKPTLHGPTPLLYRLGPVQN	x
13	1653	NS3/NS4a	GT3	3a.GQ356200.GB.Patient	98	MACMSADLEVTTSTWVLLGGVLAALAAYCLSVGCVVIVGH	
14	1767	NS4b	GT1	1a.EF407449.US.2027.2027	100	FWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTA	
15	1846	NS4b	GT1	1a.EU529681.CH.HCV-1a_CH_BID-V	96	IGSVGLGKVLVDILAGYGAGVAGALVAFKIMSGEIPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARVTAILSSLT	x
16	2119	NS5a	both	1b.EU256088.US.HCV-1b_US_BID-V	100	CPCQVPAPEFFTEVDGVRLHR	
17	2163	NS5a	GT3	3a.GQ356206.GB.Patient	100	IGSQLPCEPEPDVSVLTSML	
18	2192	NS5a	GT3	3a.GQ356201.GB.Patient	97	TAARRLARGSPPSEASSSASQLSAPSLKATCQTHR	
19	2391	NS5a	GT3	3a.GQ356209.GB.Patient	94	ESDSESCSSMPPLEGEPGDPDLSCDSWSTVSD	
20	2427	NS5a/NS5b	GT3	3a.GQ356214.GB.Patient	97	QSVVCCSMSYSWTGALITPCSAEEEKLPINPLSNSLLRHHNLVYSTSSRSASQRQKKVTFDRLQVLDDHYK	x
21	2581	NS5b	GT1	1a.AB520610._.HCV-RMT._	100	PEKGGRKPARLIVYPDLGVRVCEKMALYDV	
22	2640	NS5b	GT3	3a.GQ356212.GB.Patient	100	WTSKKTPMGFSYDTRCFDSTVTEQDIRVEEE	
23	2706	NS5b	GT1	1a.AF011753.US.H77-H21.Patient	93	CGYRRCRASGVLTTSCGNTLTCYIKARAAC	
24	2767	NS5b	GT3	3a.GQ356203.GB.Patient	100	ALRAFTEAMTRYSAPPGDAPQP	
25	2809	NS5b	GT3	3a.JQ717257.IN.ILBSRAS6._	96	KGKRYYYLTRDPTTPLARAAWETARHTPVNSWLGNIIMYAPTIWVRMVM	x
26	2880	NS5b	GT3	3a.GQ356210.GB.Patient	91	YGATYSVTPLDLPAIIERLHGLSAFTLHSYSPVELNRVAGTLRKLGCPPLRAWRHR	x
27	2939	NS5b	GT1	1b.AB442220.JP.KAH5._	100	VRAKLLSQGGRAAICGKYLFNWAVRTK	

C HCV genotype 1 – 6 immunogen

segment	starting position	region	selected subtype	selected pt.sequence	match % (pt/overall)	subtype consensus sequence	A (1000
1	11	core	GT4	4Q_RWAN_239836682_FJ462434_QC2	99	TKRNTNRRPMDVKFPGGGQIVGGVYLLPRRGPRLGVRATRKTSERSQPRGRRQPIPKARRPEGRSWAQPGYPWPLYGNEGCGWAGWLLSPRGSRPSWGPNDPRRRSRNLGKVIDTLTCGFADLMGYIPLVGAPVGGVARALAHGVRALEDGINYATGNLPGCSFSIFLLALLSCLTVPAS	x
2	304	E1	GT1	1a.AY615798.US.HCV-TWB-1._	100	CNCSIYPGHITGHRMAWDMMMNWSPTT	
3	417	E2	either	2b.JP._.AB661431._	100	NGSWHINRTALNCNDSLNTGFI	
4	503	E2	GT4	4a.EG.2006.Eg10._	100	SVCGPVYCFTPSPVVVGTTDR	
5	588	E2	GT1/GT2/G	1b.AB049088.JP.HCVT094._	100	CPTDCFRKHPEATYTKCGSGPWLTPRCLVDYPYRLWHYPCTVNF	x
6	672	E2	GT1/GT3	1b.AY587844.RU.N589.589	95	LLLSTTEWQILPCSFTTLPALSTGLIHLHQNIVDVQYLYGVGS	x
7	1028	NS2/NS3	GT3	3a.D17763._.NZL1.NZL1	97	MGWRLLAPITAYAQQTRGLLGTIVTSLTGRDKN	
8	1129	NS3	GT1	1a.EU529681.CH.HCV-1a_CH_BID-V	94	PCTCGSSDLYLVTRHADVIPVRRRGDSRGSLLSPRPISYLKGSSGGPVLCPSGHAVGIFRAAVCTRGVAKAVDFIPVESLE	x
9	1212	NS3	GT4	4a.EG.2006.Eg4._	99	MRSPVFTDNSTPPAVPQTYQVAHLHAPTGSGKSTKVPAAYAAQGYKVLVLNPSVAATLGFGAYMSKAYGI	x
10	1286	NS3	GT1/GT4	4a.EG.2006.Eg12._	96	RSGVRTITTGAPITYSTYGKFLADGGCSGGAYDIIICDECHSTDSTTILGIGTVLDQAETAGVRLVVLATATPPG	x
11	1392	NS3	all except	1a.AY615798.US.HCV-TWB-1._	100	IKGGRHLIFCHSKKKCDELA	
12	1420	NS3	GT1	1b.EU155357.CH.HCV-1b_CH_BID-V	94	NAVAYYRGLDVSVIPTSGDVVVVATDALMTGFTGDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRGRRGIYRFVTPGERPSGMFDSSVLCECYDAGCAWYELTPAETSVRLRAYLNTPGLPVCQDHLEFWEGVFTGLTHIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPPSWDQMW	x
13	1645	NS3/NS4a	GT4	4a.EG.2006.Eg4._	98	THPITKYIMACMSADLEVVTSTWVLVGGVLAALAAYCLSVGSVVIVG	x
14	1767	NS4b	GT1	1a.EF407449.US.2027.2027	98	FWAKHMWNFISGIQYLAGLSTLPGNPAIASLMAFTAAVTSPL	x
15	1843	NS4b	GT1	1a.DQ430811.US.TN168-1FL.TN168	98	GAAVGSVGLGKVLVDILAGYGAGVAGALVAFKIMSGEVPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGPGEGAVQWMNRLIAFASRGNHVSPTHYVPESDAAARV	x
16	2189	NS5a	GT4/GT6	6d.VN._.VN235.VN235	100	TAETAARRLARGSPPSLASSSASQLSAPSLKATCT	
17	2398	NS5a		***segment not included***		ESYSSMPPLEGEPGDPDLEFEQVESQ	
18	2453	NS5a/NS5b	GT6	6a.HK._.6a66._	97	VCCSMSYSWTGALITPCAAEEEKLPINPLSNSLIRHHNMVYSTTSRSASLRQKKVTFDR	x
19	2604	NS5b	GT3/GT6	3a.JQ717260.IN.ILBSRAS9._	100	PSKGGRKPARLIVYPDLGVRVCEKRALYDV	
20	2666	NS5b	GT1/GT6	6h.VN._.VN004.VN004	100	KKTPMGFSYDTRCFDSTVTERDIRTE	
21	2729	NS5b	GT1/GT2	2a._._.NDM228._	97	CGYRRCRASGVLTTSMGNTITCYIKALAA	
22	2796	NS5b	GT1	1a.EU256072.CH.HCV-1a_CH_BID-V	100	EAMTRYSAPPGDPPQPEYDLELITSCSSNVSVAHD	
23	2845	NS5b	GT6	6a.HK._.6a61._	100	TPLARAAWETARHTPVNSWLGNIIMYAPTIWVRMVLMTHFFSILQ	x
24	2903	NS5b	GT3/GT5	3a.GQ356206.GB.Patient	94	YGATYSVTPLDLPAIIERLHGLSAFTLHSYS	
25	2936	NS5b	GT6	6a.HK._.6a33._	96	ELNRVGACLRKLGVPPLRAWRHRARAVRAKLIAQGGKAAICGKYLFNWAV	x

In Figure 5-16, the numbers of sequences selected from each subtype are shown. Of note, the numbers of selected patient sequences do not entirely reflect the percentage of homology depicted in Figure 5-15: for example, no sequences were selected from HCV gt5, and proportionally less sequence from HCV genotype 2. This is based on cases where two subtypes were similar to the overall consensus (e.g. segment 24 for immunogen design HCV gt1-6, Table 5-4, page 225), and a patient sequence from one subtype most similar to the overall consensus (in the example from subtype 3 and not subtype 5) was included in the final immunogen.

Figure 5-16: Number of patient sequences of each subtype included in the final immunogen.



For each analysis, (A) HCV gt1, (B) HCV gt1/3a, (C) HCV gt1-6, the numbers of patient sequences selected for each subtype included in the analysis are shown.

5.5 Evaluation of potential immunogenicity of designed immunogens

In general, immune responses in every infected individual focus on a small number of immunodominant epitopes (Yewdell 2006). Since the immune response mounted by each individual is limited and mainly focused on a few immunodominant regions, issues have been raised regarding the immunogenicity of vaccines containing only conserved regions, thereby potentially missing many immunodominant epitopes.

On the contrary, there is evidence that conserved regions with subdominant epitopes are capable of inducing robust immune responses if dominant epitopes were absent. In a recent study, it was shown that sequential inactivation of dominant HIV epitopes up-ranks the remaining sub-dominant epitopes (E.-J. Im et al. 2011).

Since we only included conserved regions in the immunogens, there is a chance that only regions with poor immunogenicity are selected. However, by removing

immunodominant T cell epitopes, the immunogenicity of sub-dominant HCV epitopes might be improved as has been observed in HIV(E.-J. Im et al. 2011).

To evaluate potential immunogenicity of the designed immunogens, we took two different approaches: (1) Epitopes described in natural HCV gt1 and gt3 infection were analysed regarding their position in conserved or non-conserved segments, and (2) Designed immunogens were evaluated using HLA epitope prediction programs. Subsequently, the number of potential artificial, non-HCV MHC class-I epitopes caused by joining conserved segments was assessed.

5.5.1 Epitopes in conserved regions described in natural infection

We evaluated conserved segments selected for the immunogen with regards to previously described epitopes in the literature. To do this, we selected all T cell epitope entries on the immune epitope database (http://www.iedb.org, (Vita et al. 2010) search date 19/02/2014) for HCV genotype 1 and 3. Subsequently, we checked retrieved data for the following parameters and excluded all publications where requirements were not met:

1. Only human data;
2. Results from T cell assays (ELISpot, Thymidine proliferation, ICS, FACS etc.), but no prediction data was to be included;
3. Duplicates of the same epitopes in the same publication were excluded.

Epitope datasets were then analysed regarding their position in defined HCV conserved/non-conserved regions.

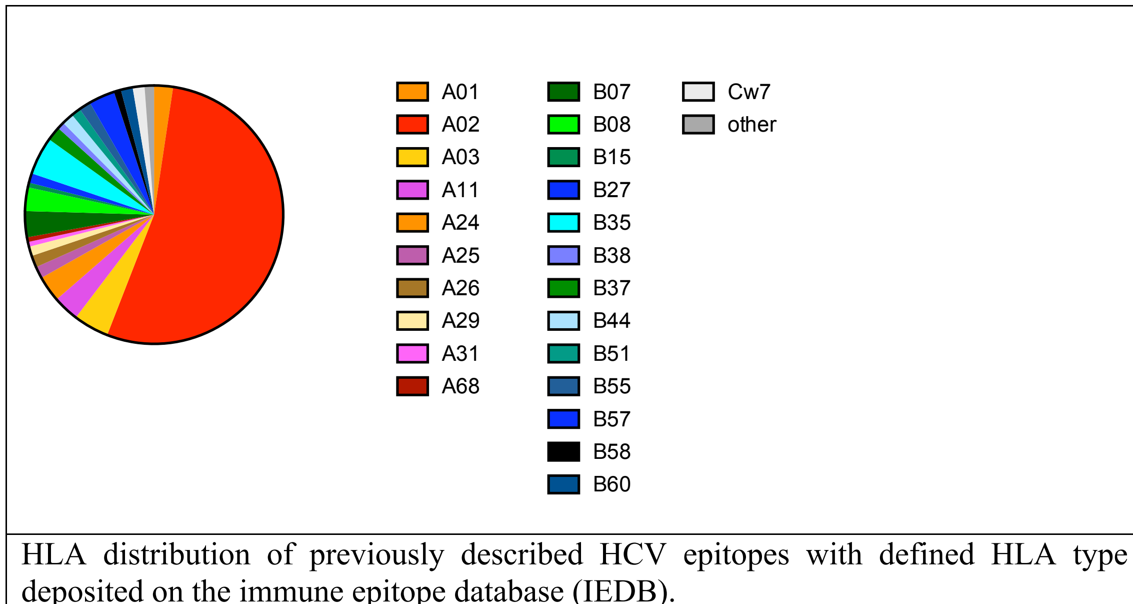
5.5.1.1 Do CD8+ epitopes described in natural HCV gt1 infection fall into conserved regions?

CD8+ restricted epitopes previously described in natural HCV gt1 infection and deposited on the IEDB were included in the analysis. After exclusion of duplicates, 454 CD8 HCV gt1 epitopes described in natural infection were included in the analysis.

For 160 out of 454 deposited CD8 restricted T cell epitopes, the HLA class-I restriction was not defined. The HLA type distribution of HCV gt1 specific CD8 epitopes is depicted in Figure 5-17. CD8+ restricted epitopes were mainly A2 restricted (181 described epitopes), which is most likely due to an experimental and publication bias. For HLA type A2, several well-characterized tetramers are available,

which have been used in multiple studies, such as CINGVCWTV and KLVALGINAV.

Figure 5-17: HLA type distribution of published epitopes as derived from the IEDB.

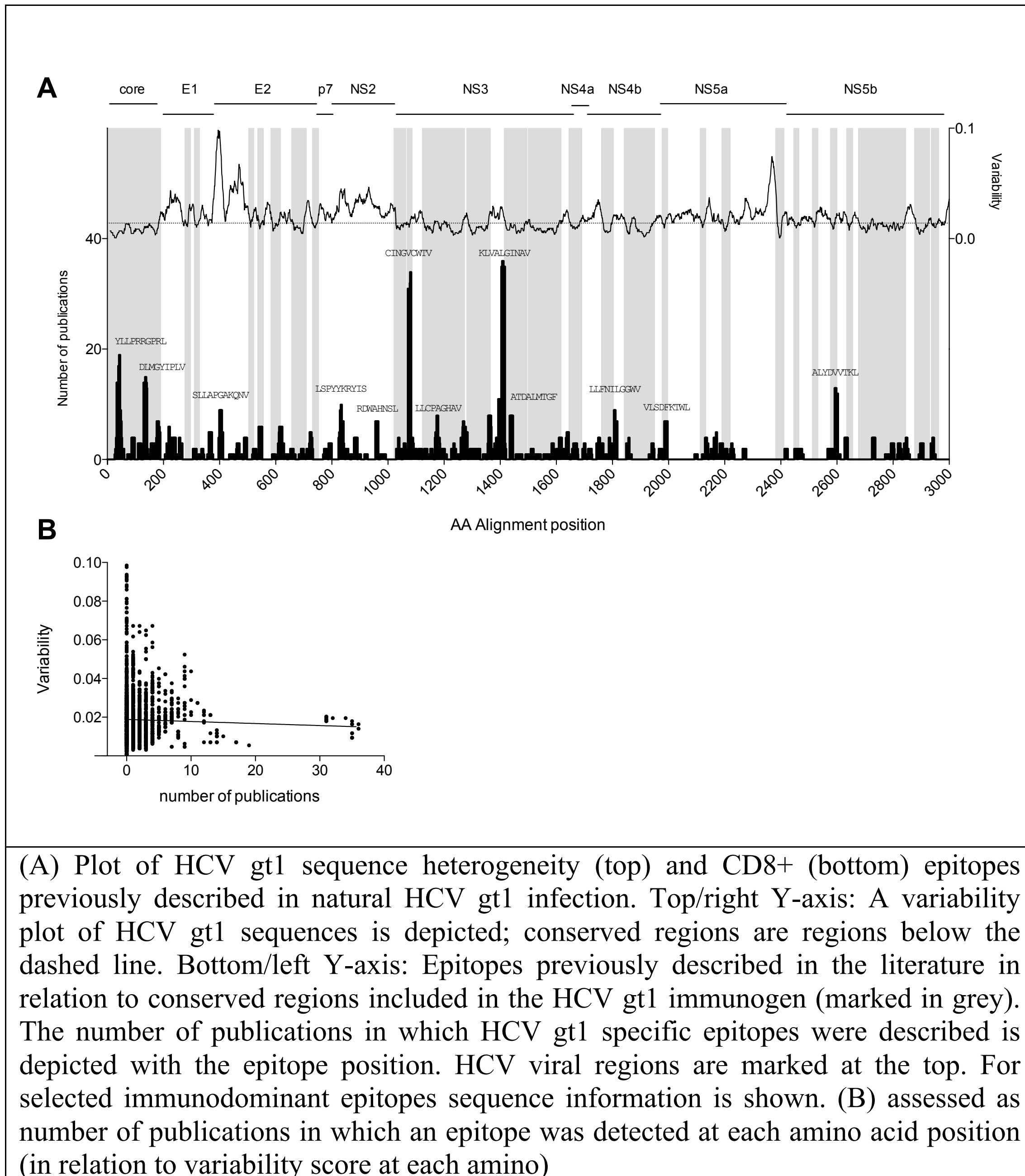


HLA distribution of previously described HCV epitopes with defined HLA type deposited on the immune epitope database (IEDB).

To define whether conserved regions were capable of priming responses in natural HCV infection, previously described CD8+ restricted epitopes were mapped on defined conserved regions (Figure 5-18 A). It has to be noted that the number of publications, used here as an approximation for immunodominance, may relate to real immunodominance or HLA dominance.

It was shown that conserved regions (as defined for the HCV gt1 immunogen defined in this thesis) are not immunologically inert in natural infection. However, the most immunodominant epitopes outside of HCV core fall into variable regions. Overall, there was no clear relation between regions commonly targeted by immune responses and sequence variability observed, represented by an association plot of immunogenicity (assessed as number of publications in which an epitope was detected at each amino acid position) in relation to variability score at each amino acid (Figure 5-18 B).

Figure 5-18: Distribution of epitopes previously described in natural HCV gt1 infection in relation to HCV gt1 sequence variability



(A) Plot of HCV gt1 sequence heterogeneity (top) and CD8+ (bottom) epitopes previously described in natural HCV gt1 infection. Top/right Y-axis: A variability plot of HCV gt1 sequences is depicted; conserved regions are regions below the dashed line. Bottom/left Y-axis: Epitopes previously described in the literature in relation to conserved regions included in the HCV gt1 immunogen (marked in grey). The number of publications in which HCV gt1 specific epitopes were described is depicted with the epitope position. HCV viral regions are marked at the top. For selected immunodominant epitopes sequence information is shown. (B) assessed as number of publications in which an epitope was detected at each amino acid position (in relation to variability score at each amino)

5.5.1.2 Are HCV gt1 epitopes previously linked to protection included in the immunogens?

Next, I assessed whether epitopes previously linked to spontaneous resolution of natural HCV infection in the literature were included in the designed immunogens. Four epitopes were defined as linked to protection in HCV gt1 infection (see Table 5-5, page 231), which were mapped on selected conserved viral regions. Only one epitope (NTRPPLGNW) was contained in full in the HCV gt1 construct. For the other

three epitopes, only fragments of the epitopes were defined as "conserved" within gt1 sequences. Epitopes KSKKTPMGF and ARMILMTHF were contained in the HCV gt1/3a and HCV gt1-6 construct, respectively. However, for these epitopes, the sequence version included in the final immunogen varied from the sequence of the epitope described as protective in HCV gt1 infection by one (**T**SKKTPMGF) or two (**V**RM**V**LMTHF) amino acids (Table 5-5). In addition, it has been shown previously that the protective effect of HLA-B27 is limited to HCV genotype 1 infection and does not expand to the genotypic sequence variant of HCV gt3 (Neumann-Haefelin et al. 2010).

Table 5-5: Are epitopes linked to protection in HCV genotype 1 included in the conserved segments.

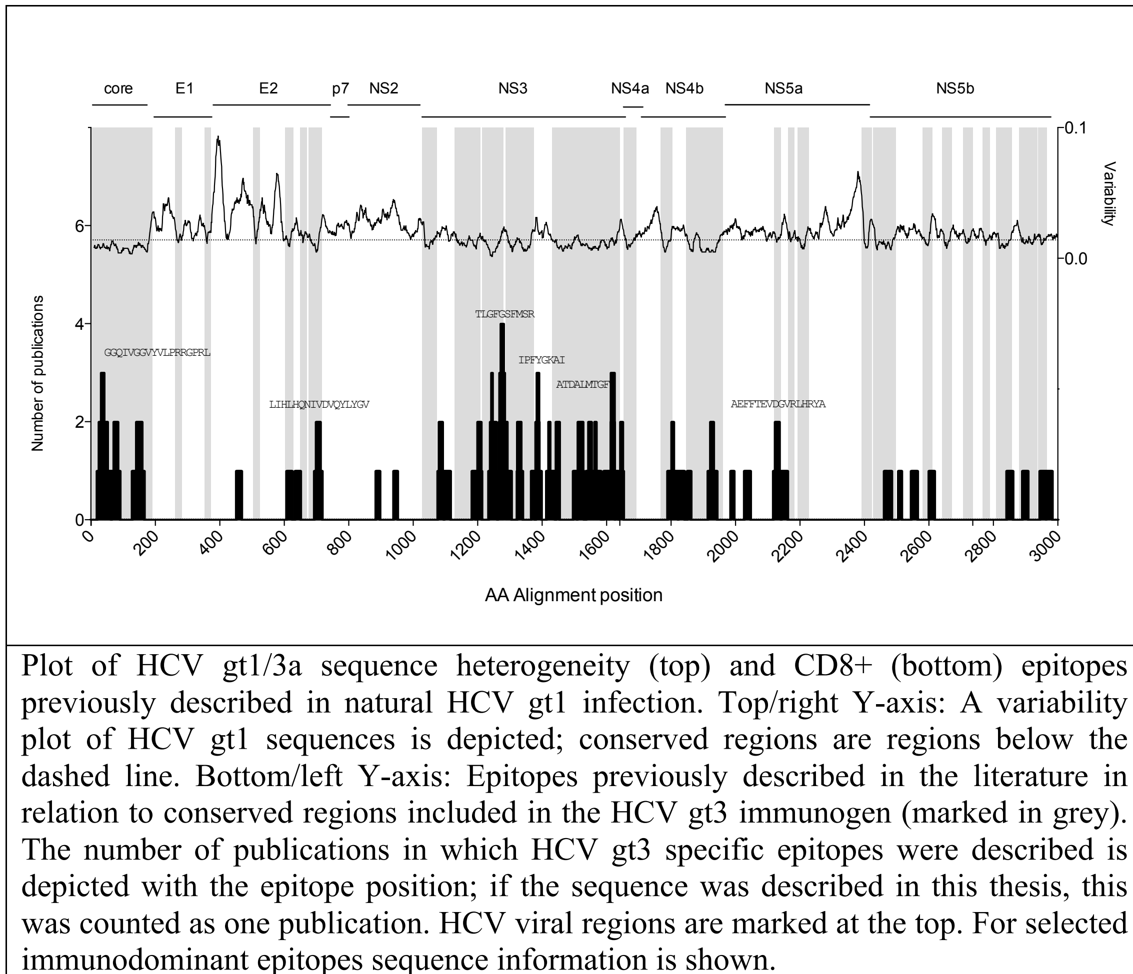
Linked HLA	Epitope	AA	Reference	construct	included?	sequence version included in immunogen
B*57	NTRPPLGNW escape: NTRPP**X**GNW	541	(A. Y. Kim et al. 2010)	gt1	yes	NTRPPLGNW
				gt1/3a	no	
				gt1-6	no	
A*03	TVYHGAGTK escape: TVYHGAG**XX**	1080 - 1088	(Fitzmaurice et al. 2011)	gt1	partially	TVYHGAG
				gt1/3a	no	
				gt1-6	no	
B*57	KSKKTPMGF escape: **XX**KK**X**PMGF	2629	(A. Y. Kim et al. 2010)	gt1	partially	PMGF
				gt1/3a	yes	**T**SKKTPMGF
				gt1-6	partially	KKTPMGF
B*27	ARMILMTHF escape: **X**RMIL**X**THF	2841 - 2849	(Neumann-Haefelin et al. 2008)	gt1	partially	ARMIL
				gt1/3a	partially	**V**RM**V**L
				gt1-6	yes	**V**RM**V**LMTHF

5.5.1.3 Epitopes described in HCV gt3 infection

Subsequently, it was assessed whether epitopes detected in natural HCV gt3 infection were located in conserved or variable HCV gt3 regions. Due to the limited number of epitopes described in HCV gt3 infection, and the fact that for the majority of epitopes a CD4/CD8 restriction was not described, both CD4 and CD8 epitopes and as well as epitopes without restriction were included in the analysis. HCV gt3 specific epitopes deposited on the immune epitope database included epitopes from four publications (Giugliano et al. 2009; H. Qureshi et al. 2011; Ruhl et al. 2012; Humphreys et al. 2012). In addition, epitopes described in Chapter 3 and 4 of this thesis were mapped, counting detection in this thesis as the equivalent to one publication (Figure 5-19). Similarly to epitopes described in HCV gt1 infection, epitopes described in HCV gt3 infection fell into conserved and non-conserved regions. Of note, the dominant

epitope AEFFTEVDGVRLHRYA that was found in 2 spontaneously resolved patients in this study (Chapter 4) fell into a conserved region and was included in the immunogen. So far, no epitopes in HCV gt3 infection were clearly linked to spontaneous resolution of HCV infection, therefore no analysis on potential protective epitopes has been attempted.

Figure 5-19: Distribution of epitopes previously described in natural HCV gt3 infection in relation to HCV gt1/3a sequence variability



Plot of HCV gt1/3a sequence heterogeneity (top) and CD8+ (bottom) epitopes previously described in natural HCV gt1 infection. Top/right Y-axis: A variability plot of HCV gt1 sequences is depicted; conserved regions are regions below the dashed line. Bottom/left Y-axis: Epitopes previously described in the literature in relation to conserved regions included in the HCV gt3 immunogen (marked in grey). The number of publications in which HCV gt3 specific epitopes were described is depicted with the epitope position; if the sequence was described in this thesis, this was counted as one publication. HCV viral regions are marked at the top. For selected immunodominant epitopes sequence information is shown.

In summary, based on this data from natural HCV gt1 and gt3 infection, we hypothesize that designed immunogens will not be immunogenically inert, but may be capable of priming immune responses to previously subdominant regions.

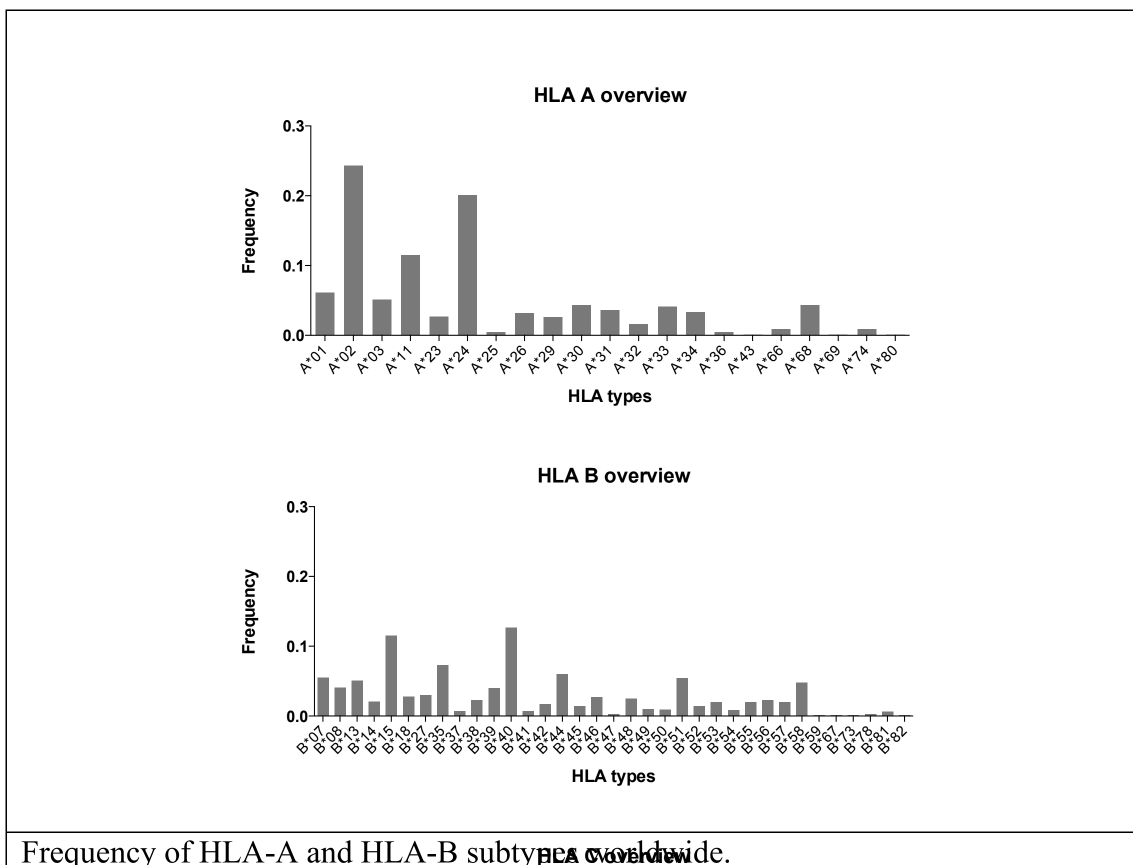
5.5.2 Evaluation of immunogenicity using epitope prediction programs

Next, we aimed to evaluate potential immunogenicity by MHC class-I epitope prediction using freely available online prediction programs. First, we evaluated common HLA types, to ensure that epitope prediction was attempted for HLA types abundant in the population. Next, we chose three online epitope prediction servers from the wide range of freely available servers based on literature evidence. Subsequently, selected conserved viral regions were evaluated for potential immunogenicity using these three servers. Particular attention was paid to immunogenicity in junctional regions between conserved segments

5.5.2.1 Evaluation of HLA type distribution

To ensure that common HLA types were included in the epitope prediction analysis (see 5.5.2.3, page 236), worldwide HLA type distributions were evaluated using the NCBI HLA resource (http://www.ncbi.nlm.nih.gov/projects/gv/mhc/ihwg.cgi, accessed in February 2013). Distributions for HLA-A and B types were assessed for different regions and populations combined (Figure 5-20).

Figure 5-20: Frequency of HLA types worldwide.



Frequency of HLA-A and HLA-B subtypes worldwide.

Of the HLA types listed in Figure 5-20, the most common HLA types were subsequently evaluated for potential MHC class-I epitopes using freely available online prediction programs.

5.5.2.2 Choosing an epitope prediction server

Available online epitope prediction servers offer predictions for different steps in peptide processing and binding to MHC class-I and II complexes. Options allow predictions for proteasomal degradation, TAP processing, MHC binding and MHC peptide presentation. Since peptide predictions are complex, we tried to keep the analysis to a well-defined step: the peptide binding stage.

Several computational methods for the assessment of T cell epitopes have been developed, using a variety of statistical and machine learning approaches. However, it gets increasingly more complicated to choose the best server for the desired analysis, since there is no guarantee that servers produce good quality predictions (Brusic et al. 2004). Therefore, an epitope prediction server was chosen based on previously published evidence. Lin et al. compared 30 combinations of servers and computational methods for MHC class-I epitope prediction (Table 5-6) (H. H. Lin et al. 2008).

Table 5-6: Evaluated MHC class-I epitope prediction servers and computational methods

Server	URL	Method	Reference
BIMAS	http://www-bimas.cit.nih.gov/molbio/hla_bind/	Matrix	(K. C. Parker et al. 1994)
HLA Ligand	http://hlaligand.ouhsc.edu/prediction.htm	Matrix	(M Sathiamurthy et al. 2003)
IEDB (ANN)	http://tools.immuneepitope.org/analyze/html/mhc_binding.html	ANN	(M. Nielsen et al. 2003)
IEDB (ARB)	http://tools.immuneepitope.org/analyze/html/mhc_binding.html	Matrix	(Bui et al. 2005)
IEDB (SMM)	http://tools.immuneepitope.org/analyze/html/mhc_binding.html	Matrix	(B. Peters & Sette 2005)
MAPPP (Bimas)	http://www.mpiib-berlin.mpg.de/MAPPP/binding.html	Matrix	(Hakenberg et al. 2003)
MAPPP (SYFPEITHI)	http://www.mpiib-berlin.mpg.de/MAPPP/binding.html	Matrix	(Hakenberg et al. 2003)
MHC Binder Prediction	http://www.vaccinedesign.com/	Matrix	-
MHC-BPS	http://bidd.cz3.nus.edu.sg/mhc/	SVM	(Cui et al. 2006)
MHC-I (Multiple matrix)	http://atom.research.microsoft.com/hlabinding/hlabinding.aspx	SBM	(Jojic et al. 2006)
MHC-I (Single matrix)	http://atom.research.microsoft.com/hlabinding/hlabinding.aspx	SBM	(Jojic et al. 2006)
MHCPred (Amino Acids)	http://www.jenner.ac.uk/MHCPred/	PLS	(Guan et al. 2006)
MHCPred (Interactions)	http://www.jenner.ac.uk/MHCPred/	PLS	(Guan et al. 2006)
MULTIPRED (ANN)	http://antigen.i2r.a-star.edu.sg/multipred1/	ANN	(G. L. Zhang et al. 2005)
MULTIPRED (HMM)	http://antigen.i2r.a-star.edu.sg/multipred1/	HMM	(G. L. Zhang et al. 2005)
MULTIPRED (SVM)	http://antigen.i2r.a-star.edu.sg/multipred1/	SVM	(Bhasin & Raghava 2007)
NetMHC (ANN)	http://www.cbs.dtu.dk/services/NetMHC/	ANN	(S Buus et al. 2003)
NetMHC (Weight Matrix)	http://www.cbs.dtu.dk/services/NetMHC/	Matrix	(M. Nielsen et al. 2004)
nHLAPred (ANNPred)	http://www.imtech.res.in/raghava/nhlapred/neural.html	ANN	(Bhasin & Raghava 2007)
nHLAPred (ComPred)	http://www.imtech.res.in/raghava/nhlapred/comp.html	ANN/Matrix	(Bhasin & Raghava 2007)
PepDist	http://www.pepdist.cs.huji.ac.il/	DF	(Wan et al. 2006)
PeptideCheck	http://www.peptidecheck.org/	Matrix	(Reche et al. 2002)
Predep	http://margalit.huji.ac.il/Teppred/mhc-bind/index.html	SBM	(Wan et al. 2006)
ProPred1	http://www.imtech.res.in/raghava/propred1	Matrix	(H. Singh & Raghava 2003)
Rankpep	http://bio.dfci.harvard.edu/Tools/rankpep.html	Matrix	(Reche et al. 2002)
SMM	http://zlab.bu.edu/SMM/	Matrix	(B. Peters et al. 2003)
SVMHC (MHCPEP)	http://www.sbc.su.se/~pierre/svmhc/new.cgi	SVM	(Dönnes & Kohlbacher 2006)
SVMHC (SYFPEITHI)	http://www.sbc.su.se/~pierre/svmhc/new.cgi	SVM	(Dönnes & Kohlbacher 2006)
SVRMHC	http://SVRMHC.umn.edu/SVRMHCdb	SVM	(Wan et al. 2006)
SYFPEITHI	http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm	Matrix	(Rammensee et al. 1999)

Servers and computational methods evaluated in (H. H. Lin et al. 2008). ANN artificial neural networks, SBM, structure based model, PLS partial least square, DF distant function

The performance of different servers and computational methods was assessed using an experimental dataset from a full-overlapping binding study of 9-mer peptides to seven HLA-I molecules (Peters & Sette 2005). Peptides were derived from a tumor antigen and a fragment of a viral antigen, and assessed against HLA types. Prediction results of the different servers were normalized and then compared regarding different HLA types. Overall, HLA types A*0201, A*0301, A*1101, B*0702, B*0801, B*1501 had excellent prediction accuracy, and B*2402 moderate classification accuracy. According to this study, NetMHC artificial neural networks (ANN computing method, at http://www.cbs.dtu.dk/services/NetMHC) is likely to produce best prediction results, being ranked as the best overall peptide predictor, showing best performance for thresholds that optimize the selection of T cell epitopes, and thresholds that do not allow false positives (H. H. Lin et al. 2008).

Prediction capabilities of available MHC class-II servers do not match MHC class-I servers (H. H. Lin et al. 2008). 21 available MHC class-II prediction servers were evaluated (Table 5-7), comparing prediction results to experimental data from four independent studies involving 721 peptide binding assays to seven common HLA-DR molecules: DRB1*0101, 0301, 0401, 0701, 1101, 1301, and 1501. NetMHCIIpan was the best individual predictor, closely followed by PROPRED, IEDB (Consensus), and MULTIPRED (SVM). Overall, the authors conclude that available servers offer only limited prediction accuracy and further development of class-II predictors is needed.

Table 5-7: Evaluated MHC class-II epitope prediction servers and computational methods

Server	URL	Method	Reference
HLA-DR4Pred (ANN)	http://www.imtech.res.in/raghava/hladr4pred/index.html	ANN	(Bhasin & Raghava 2007)
HLA-DR4Pred (SVM)	http://www.imtech.res.in/raghava/hladr4pred/index.html	SVM	(Bhasin & Raghava 2007)
IEDB (ARB)	http://tools.immuneepitope.org/analyze/html/mhc_II_binding.html	Matrix	(Bui et al. 2005)
IEDB (SMM)	http://tools.immuneepitope.org/analyze/html/mhc_II_binding.html	Matrix	(M. Nielsen et al. 2007)
IEDB (Saturniolo)	http://tools.immuneepitope.org/analyze/html/mhc_II_binding.html	Matrix	(Sturniolo et al. 1999)
IEDB (Consensus)	http://tools.immuneepitope.org/analyze/html/mhc_II_binding.html	Matrix	
MHC Binder Prediction	http://www.vaccinedesign.com	Matrix	
MHC2Pred	http://www.imtech.res.in/raghava/mhc2pred	SVM	
MHC-BPS	http://bidd.cz3.nus.edu.sg/mhc	SVM	(Cui et al. 2006)
MHCPred	http://www.jenner.ac.uk/MHCPred	PLS	(Guan et al. 2006)
Multipred1 (ANN)	http://antigen.i2r.a-star.edu.sg/multipred1	ANN	(G. L. Zhang et al. 2005)
Multipred1 (HMM)	http://antigen.i2r.a-star.edu.sg/multipred1	HMM	(G. L. Zhang et al. 2005)
Multipred1 (SVM)	http://antigen.i2r.a-star.edu.sg/multipred1	SVM	(G. L. Zhang et al. 2007)
NetMHCII	http://www.cbs.dtu.dk/services/NetMHCII	Matrix	(M. Nielsen et al. 2007)
NetMHCIIpan	http://www.cbs.dtu.dk/services/NetMHCIIpan	ANN	(M. Nielsen et al. 2008)
PeptideCheck	http://www.peptidecheck.org	Matrix	(DeLuca et al. 2007)
ProPred	http://www.imtech.res.in/raghava/propred	Matrix	(H. Singh & Raghava 2003)
Rankpep	http://bio.dfci.harvard.edu/Tools/rankpep.html	Matrix	(Reche et al. 2002)
SVMHC	http://www-bs.informatik.uni-tuebingen.de/SVMHC/index_html	Matrix	(Dönnes & Kohlbacher 2006)
SVRMHC	http://www-bs.informatik.uni-tuebingen.de/SVMHC/index_html	SVM	(Wan et al. 2006)
SYFPEITHI	http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm	Matrix	(Rammensee et al. 1999)

Servers and computational methods evaluated in (H. H. Lin et al. 2008). ANN artificial neural networks, SBM, structure based model, PLS partial least square, DF distant function

Hence, we chose to analyse our immunogens for HLA class-I types A and B with 3 different, freely available online prediction programs:

(1) NetMHC (ANN)/ NetMHCIIpan at http://www.cbs.dtu.dk/services/NetMHC, which ranked highest in Lin et al., 2008,

(2) Syfpeithi (http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm), and

(3) BIMAS, (http://www-bimas.cit.nih.gov/molbio/hla_bind/).

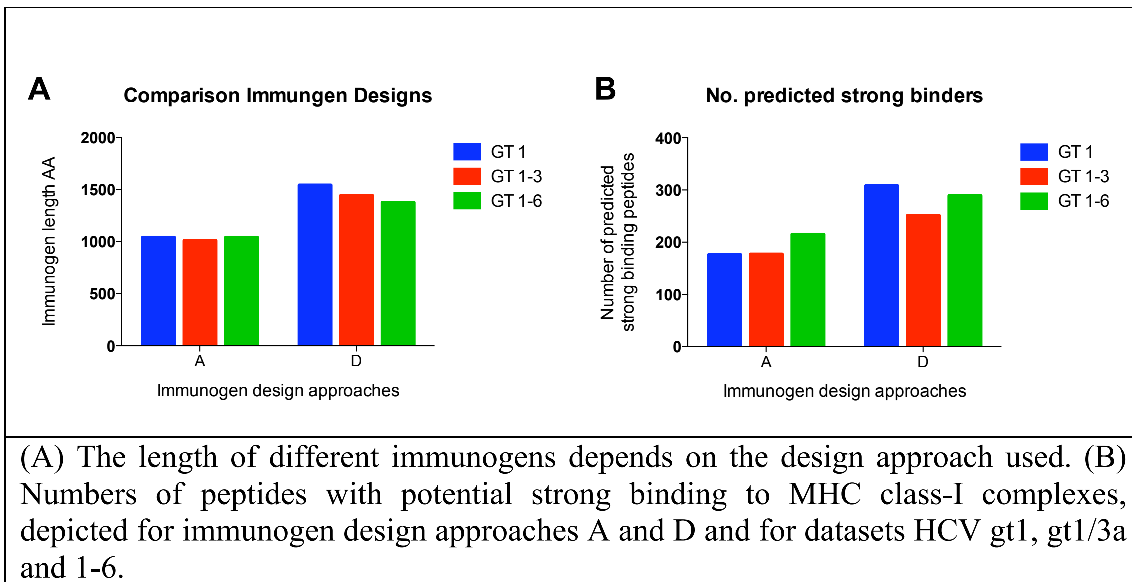
Syfpeithi and BIMAS servers were included in the analysis since these servers have been commonly used in our department previously. Based on the restricted accuracy of MHC class-II prediction, potential CD4+ restricted epitopes were not evaluated in the designed immunogens.

5.5.2.3 Are the designed immunogens immunogenic? - MHC class-I epitope prediction using the NetMHC server

To assess the potential immunogenicity of the designed immunogens, we ran an epitope prediction analysis for full-length immunogens for HCV gt1, gt1/3a and gt1-6 (approaches A and D). The following settings were selected for the NetMHC (ANN) analysis: HLA supertype representatives available on the NetMHC server were selected for the analysis (HLA A*0101, *0201, *0301, *2402, *2601, HLA B*0702, *0801, *1501, *2705, *3901, *4001, and *5801). For these HLA types, peptides of 8 to 11 amino acids in length were predicted, with a threshold of a predicted IC50 of 50 for a strong binder and a score <0.5. We used the 2.8 server version available in November 2013.

Short and long versions of the designed immunogens for each sequence dataset (HCV gt1a/1b, HCV gt1/3a, HCV gt1-6) had significant immunogenicity, with more than 190 strong binders predicted for the shorter immunogen length (approach A), and more than 250 strong binders predicted for the long immunogen versions (approach D) (Figure 5-21 B). The number of predicted strong binders was related to the length of the immunogen (Figure 5-21 A).

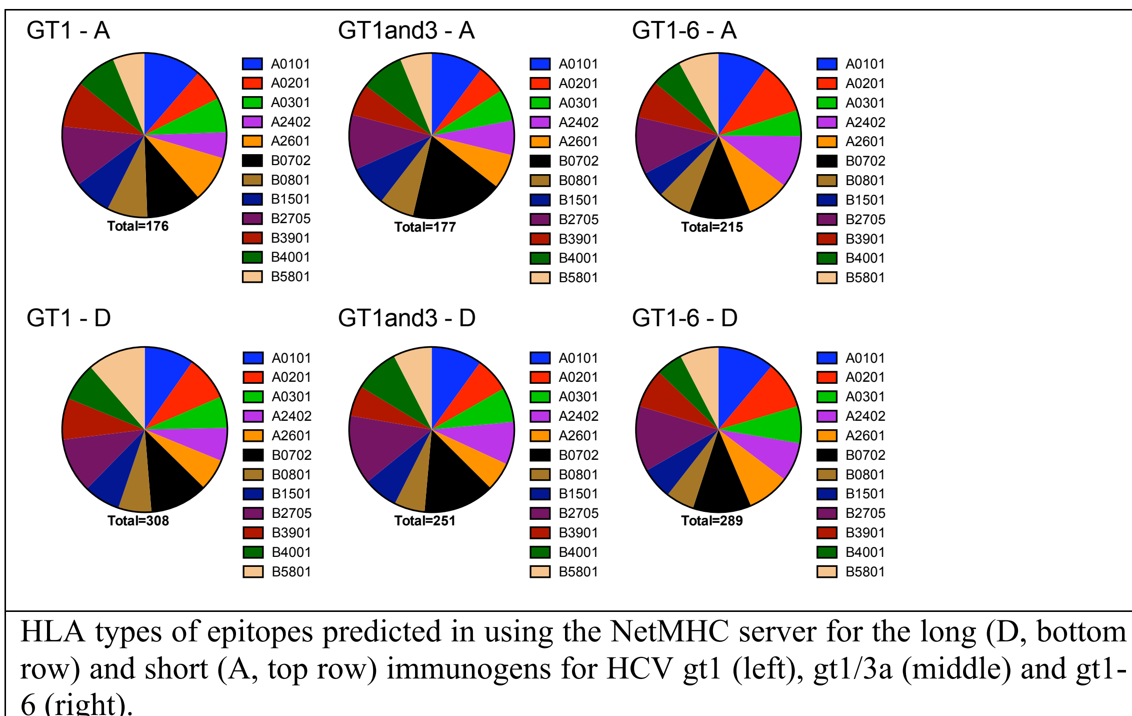
Figure 5-21: Numbers of predicted strong binders to MHC class-I epitopes.



(A) The length of different immunogens depends on the design approach used. (B) Numbers of peptides with potential strong binding to MHC class-I complexes, depicted for immunogen design approaches A and D and for datasets HCV gt1, gt1/3a and 1-6.

To ensure that no major HLA type was underrepresented by reducing the immunogen length from an all segments approach (analysis D) to an immunogen with a restricted number of segments (analysis A), HLA type distributions for predicted MHC class-I epitopes were compared between different immunogen approaches. No major change in HLA type distributions of predicted strong binders was observed when shortening the long version (version D) of the designed immunogens (Figure 5-22).

Figure 5-22: HLA type distributions of predicted epitopes in different analyses and immunogens.



HLA types of epitopes predicted in using the NetMHC server for the long (D, bottom row) and short (A, top row) immunogens for HCV gt1 (left), gt1/3a (middle) and gt1-6 (right).

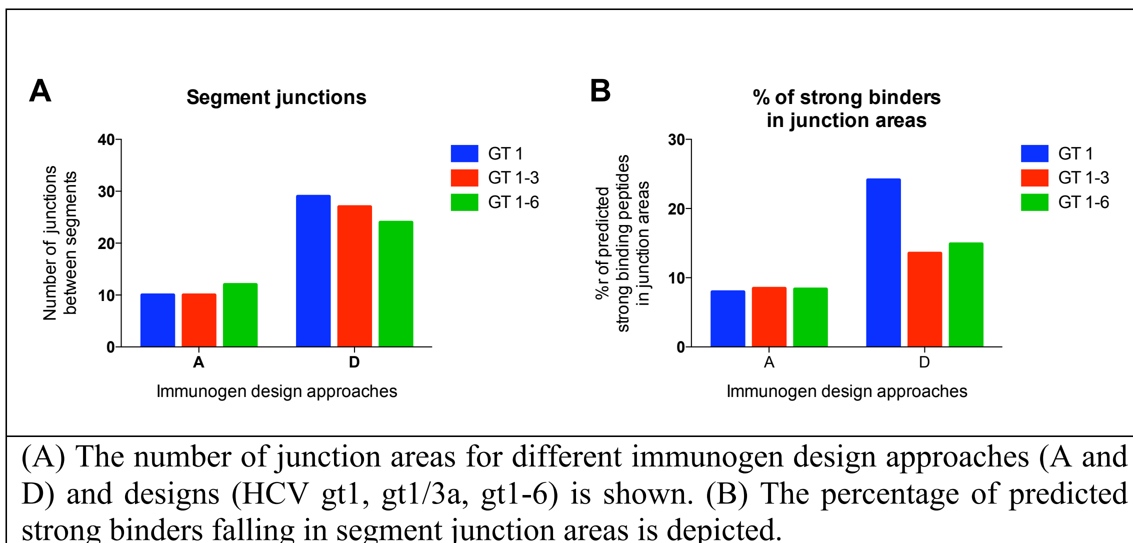
Based on this epitope prediction analysis, I conclude that the designed conserved immunogens have the potential of inducing potent immune responses to subdominant epitopes to a range of different HLA types.

5.5.3 Putative artificial MHC class-I epitopes in junction areas

If conserved segments are joined together in a single, open reading frame immunogen, immune responses to segment junctions will target artificial, non-HCV epitopes. To assess putative artificial epitopes, junction regions between conserved segments were defined using different online prediction servers. Junctional regions were defined as 10 amino acids on each side of each conserved segment.

Initially, junction regions were assessed for artificial, non-HCV epitopes using the NetMHC prediction server (Figure 5-23). Depending on the immunogen design approach, about 8-25% of all predicted strong binding epitopes fall in junction regions. The number of segment junctions is directly related to the amount of putative artificial epitopes formed.

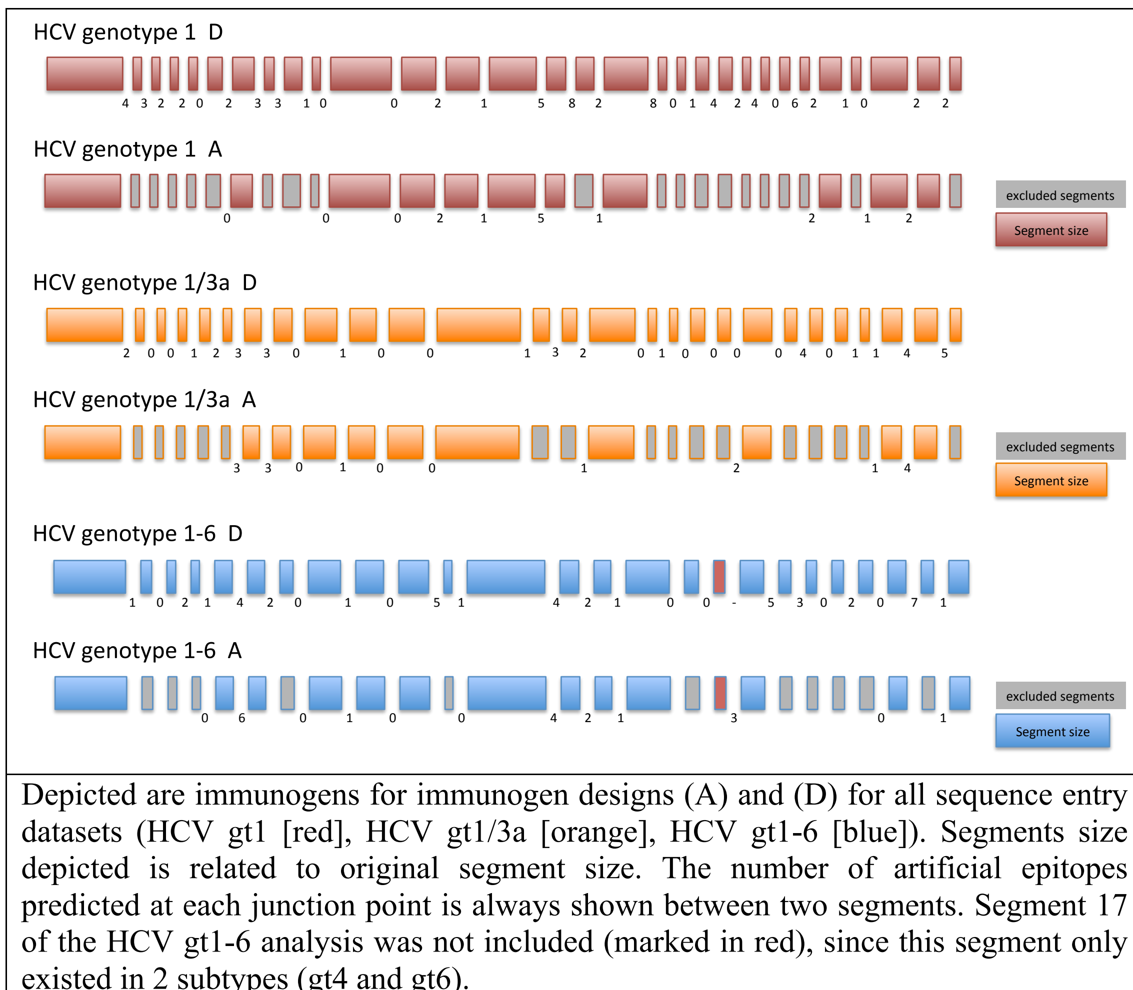
Figure 5-23: Potential artificial epitopes generated at segment junctions



(A) The number of junction areas for different immunogen design approaches (A and D) and designs (HCV gt1, gt1/3a, gt1-6) is shown. (B) The percentage of predicted strong binders falling in segment junction areas is depicted.

A detailed analysis showing numbers of predicted epitopes for each junction region of the different immunogens is depicted in Figure 5-24.

Figure 5-24: Detailed numbers of potential artificial epitopes in junction regions for all immunogen designs.



Depicted are immunogens for immunogen designs (A) and (D) for all sequence entry datasets (HCV gt1 [red], HCV gt1/3a [orange], HCV gt1-6 [blue]). Segments size depicted is related to original segment size. The number of artificial epitopes predicted at each junction point is always shown between two segments. Segment 17 of the HCV gt1-6 analysis was not included (marked in red), since this segment only existed in 2 subtypes (gt4 and gt6).

Since epitope prediction analyses will never fully reflect the immune response *in vivo*, we aimed to optimize prediction results by including and comparing predictions from different servers. Ideally, this should enable the exclusion of false positive predictions without a loss of prediction sensitivity. Three different prediction servers were used to predict artificial epitopes:

(1) NetMHC server (ANN, http://www.cbs.dtu.dk/services/NetMHC)

(2) Syfpeithi server (http://www.syfpeithi.de)

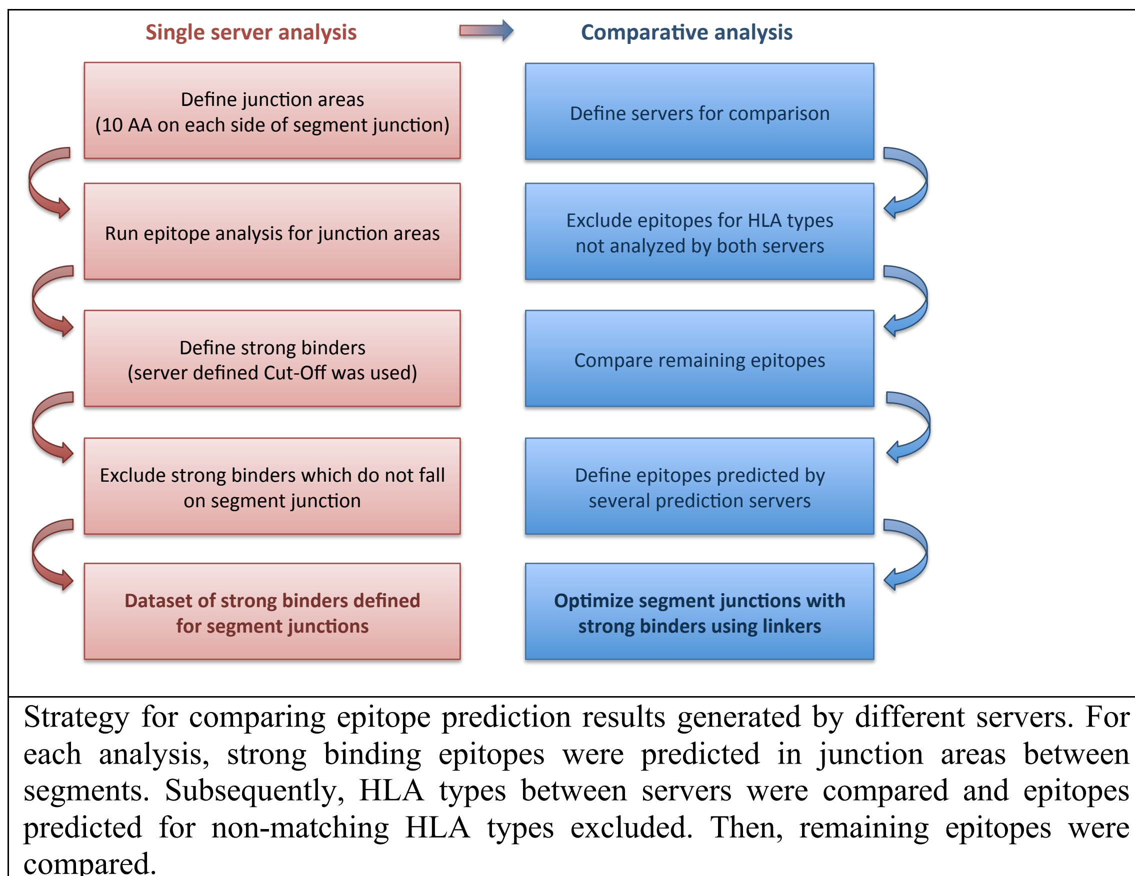
(3) BIMAS server (http://www-bimas.cit.nih.gov/molbio/hla_bind/)

The selection was based on a meta-analysis of available prediction servers (NetMHC), as well as on the previous usage of prediction servers in our laboratory (Syfpeithi/BIMAS).

The following strategy for the comparison of predicted epitopes was used (Figure 5-25); first, junction regions were defined, selecting 10 amino acids on each side of

the segment junction for each junction of designed immunogens (gt1, gt1/3, gt1-6). Then, the epitopes prediction analyses were run for all available MHC class-I types for each of the three servers, and strong binders defined following the recommendations on each prediction server (NetMHC score <0.5, Syfpeithi score >20, BIMAS score >100). Strong binders not falling exactly on the junction (e.g. a predicted peptide of 8AA in length, starting at the first amino acid of the defined junction region) were excluded from the analysis. To then compare the datasets generated by different epitope prediction servers, we excluded epitopes predicted for HLA types that were not shared between servers. If different peptide lengths for the same HLA type were predicted, these were not excluded, since we reasoned that interesting information was contained, even if the exact match between peptides was not met: these epitopes were classified as "similar" in our analysis. Remaining epitopes were compared, and junction regions with strong binders adapted using linkers between segments (see 5.5.3.1, page 242).

Figure 5-25: Strategy for comparing different epitope prediction servers



Strategy for comparing epitope prediction results generated by different servers. For each analysis, strong binding epitopes were predicted in junction areas between segments. Subsequently, HLA types between servers were compared and epitopes predicted for non-matching HLA types excluded. Then, remaining epitopes were compared.

A detailed description of the analysis and results can be found in the appendix (see 8.2 Comparative analysis of CD8+ restricted epitopes predicted by three different

online epitope prediction servers, page 291 following). In brief, the overlap in prediction of strong binders by the used epitope prediction servers was minimal.
When comparing the NetMHC to Syfpeithi prediction, of 169 artificial strong binders predicted by NetMHC (70 epitopes) and Syfpeithi (99 epitopes) servers, 24 epitopes were identical, reflecting an overlap of 34% (NetMHC) and 24% (Syfpeithi) (Supplementary Table 8-3, page 293). All other predicted strong binders were either predicted for different HLA types or lengths. If predicted epitopes differed in length by one or two amino acids, but were predicted for identical HLA types, they were classified as “similar”: this applied to 8 more epitopes, increasing the overlap from 34% to 46% (NetMHC) and 24% to 32%, respectively.
Next, I compared prediction results of NetMHC and BIMAS servers: of 107 predicted artificial strong binders [46 (NetMHC), 61 (BIMAS)], only 5 predicted epitopes were identical, estimating at 11% overlap of epitopes predicted by the NetMHC server and 8% overlap for the BIMAS analysis (Supplementary Table 8-6, page 299). All other predicted strong binders were either predicted for different HLA types or lengths. 5 more epitopes were classified as “similar”, increasing the match percentage from 11% to 22% (NetMHC) and 8% to 16% (BIMAS).
191 epitopes were predicted in junction regions between segments by Syfpeithi and BIMAS servers [96 (Syfpeithi) and 95 (BIMAS)], of which 24 were identical for both prediction analyses, which estimates at 25% for both analyses (Table 8-9, page 302). A further 2 epitopes were classified as “similar”, increasing the percentage of matching epitopes to 26%.
Overall, overlap percentage ranged from 8% to 34% (or 16% and 46%, if similar epitopes were considered), with the greatest overlap between the Syfpeithi and NetMHC server, and a low overlap for predicted epitopes between NetMHC and BIMAS. In conclusion, there is a significant mismatch between strong binders predicted for identical HLA types by different servers. Overall, the NetMHC server predicted the smallest number of epitopes, implying a higher rate of false positives using both of the other servers. This is in line with published evidence by Lin et al. (H. H. Lin et al. 2008). The comparative analysis implies that there is more overlap of predicted epitopes between NetMHC and Syfpeithi than between NetMHC and BIMAS, with 21 (+6 similar; NetMHC/Syfpeithi) versus 5 (+5 similar; NetMHC/BIMAS) identical epitopes predicted. This is potentially caused by an overrepresentation of HLA B*27 and B*51 epitopes in the BIMAS prediction.

5.5.3.1 Avoiding potential artificial epitope: Usage of linkers

Linkers between two joined segments are used in immunogen design to avoid the generation of artificial antigens. Commonly used linkers include various glycine/proline or glycine/serine combinations. Examples for these linkers used in previous immunogen designs include GGGPGGG in the heterotypic Influenza A NP+M1 vaccine (T. K. Berthoud et al. 2011), GGGSGGG in a recent Malaria immunogen (Biswas et al. 2011) or GS, used in the ME-TRAP malaria vaccine (S. C. Gilbert et al. 1997). Other possibilities include lysine linkers (A. Yano et al. 2013), sulfhydryl linkers potentially augmenting T cell immune responses (Timmerman 2009), glutaraldehyde linkers (Hurvitz & Timmerman 2005) and furin sensitive linkers (J. Lu et al. 2004). However, not all vaccines containing several joint segments include linkers at junction sides, the HIV immunogen HIV_{cons} designed by Tom Hanke's group (Létourneau et al. 2007) does not contain any linkers. Here, about 20% of the generated immune response in mice was aimed against junctional regions (correspondence with Tomas Hanke).

Since the conserved immunogens contained multiple different segments, and according to epitope prediction programs about 20% of potential epitopes were lying in junction areas, we chose to include linkers in junction areas where strong binders were predicted in epitope prediction analyses by at least 2 different servers. Linkers were designed to abrogate artificial binders. However, linker design was complicated by divergent prediction results obtained through epitope prediction servers. Therefore, linker design was only attempted in cases where epitopes were predicted by at least two independent epitope prediction servers. After the insertion of linkers at junctions with strong binding artificial epitopes, regions were re-assessed for strong binders using online prediction servers. If strong binders were not abolished by the usage of linkers, original junction sites without linkers were used. Since multiple identical linkers included in a construct might cause dimerization during cellular immunogen expression and therefore complicate production, different linkers were designed to avoid dimerization.

In the following section, results for the HCV gt1 immunogen (design D) are discussed as an example. For a detailed depiction of results of this analysis, refer to the appendix (Table 8-10, page 305). For this immunogen, 11 of the 29 junction sites were not modified: no epitopes were predicted at two junctions (5, 18), and no overlap between prediction servers was seen at 4 junction sites (3, 7, 10, 27), and for

the remaining 5 junction sites (4, 8, 9, 17, 20), inserted linkers did not modify strong binders, or created other putative epitopes. An example of modified junction regions with different linkers is shown in Table 5-8.

Table 5-8: Inserted linker regions for HCV gt1 immunogen (approach D)

Junction	Amino acid sequence	Suggested linkers at junctions
0	SQEIHARFRRSTNPKPQRKT	SQEIHARFRRkgggpgggkSTNPKPQRKT
1	LSCLTVPASAYVGDLCGSVF	LSCLTVPASAggsgYVGDLCGSVF
2	VSQLFTFSPRYPGHITGHRM	VSQLFTFSPRpYPGHITGHRM
3	AWDMMMNWSPVCGPVYCFTP	
4	SPVVVGTTDRTDVFLLNNTR	
5	LGNWFGCTWMCPTDCFRKHP	
6	CLVDYPYRLWDRDRSELSPL	CLVDYPYRLWgDRDRSELSPL
7	QYLYGVGSSVARVCACLWMM	
8	QAEAALENLVKGWRLLAPIT	
9	VEGEVQIVSTQSFLATCING	
10	VCWTVYHGAGPCTCGSSDLY	
11	GAYMSKAHGVTGVRTITTGS	GAYMSKAHGVsgTGVRTITTGS
12	TVPHPNIEEVNAVAYYRGLD	TVPHPNIEEVgpgNAVAYYRGLD
13	QRRGRTGRGRYRFVTPGERP	QRRGRTGRGRpgggsgggYRFVTPGERP
14	MWKCLIRLKPKYIMTCMSAD	MWKCLIRLKPggkgpggKYIMTCMSAD
15	VIVGRIVLSGFWAKHMWNFI	VIVGRIVLSGgpgFWAKHMWNFI
16	ASVTSPLTTQGSIGLGKVLV	ASVTSPLTTQpGSIGLGKVLV
17	AAARVTQILSSWLRDIWDWI	
18	VLSDFKTWLKCPCQVPSPEF	
19	TELDGVRLHRRRLARGSPPS	TELDGVRLHRkkgpgsgpgpRRLARGSPPS
20	SLKATCTTNHSDAESYSSMP	
21	GSWSTVSSEALSNSLLRHHN	GSWSTVSSEAgsgsLSNSLLRHHN
22	MVYATTSRSALTPPHSAKSK	MVYATTSRSAgpLTPPHSAKSK
23	GYGAKDVRCHKPARLIVFPD	GYGAKDVRCHsgsgggsKPARLIVFPD
24	CEKMALYDVVPMGFSYDTRC	CEKMALYDVVggPMGFSYDTRC
25	TVTESDIRTESLTERLYVGG	TVTESDIRTEggsgggSLTERLYVGG
26	ASAACRAAKLCTMLVCGDDL	ASAACRAAKLggsggCTMLVCGDDL
27	CESAGTQEDASLRAFTEAMT	
28	APTLWARMILEPLDLPQIIQ	APTLWARMILggsggEPLDLPQIIQ
29	WRHRARSVRAGGRAAICGKY	WRHRARSVRAsGGRAAICGKY

Modified junction regions between conserved segments by inserting a linker. Junction regions are numbered (junction 0 is between TPA leader sequence and conserved segment 1). On the left, designed linker sequences are depicted in blue letters.

In the following, divergent results were obtained using the different online epitope prediction programs used:

NetMHC: For the first prediction analysis in junction areas, 70 artificial strong binders were predicted. 23 (33%) of these were not modified, since linker insertion at the binding site did not achieve a binding score reduction, or epitopes were not predicted by two independent prediction servers (junction 4, 7, 8, 9, 10, 17 and 27). The number of remaining epitopes (47) was reduced by 57% to 20 (Supplementary Table 8-10, page 305). Of these 20 epitopes, 2 were new predicted strong binders (defined as new predicted HLA type, or entirely different sequence to previously predicted epitopes, marked in purple). 15 epitopes were modified epitopes (defined as modified predicted strong binders, predicted for the same HLA type, partly overlapping with previously predicted epitope but now including linker regions,

marked light red) predicted containing linker amino acids. For 3 epitopes, a stronger binding score was calculated after modification (marked in red).

Syfpeithi: For the first prediction analysis in junction areas, 148 artificial strong binders were predicted. 35 (24%) of these were not modified (junction 4, 7, 8, 9, 10, 17 and 27). The number of remaining epitopes (113) was reduced by 24% to 86 (Supplementary Table 8-10, page 305). Of these 86 epitopes, 46 were new predicted strong binders (defined as new predicted HLA type, or entirely different sequence to previously predicted epitopes, marked in purple). 38 epitopes were modified epitopes (defined as modified predicted strong binders, predicted for the same HLA type, partly overlapping with previously predicted epitopes but now including linker regions, marked light red) predicted containing linker amino acids. For 5 epitopes, a stronger binding score was calculated after modification (marked in red).

BIMAS: For the first prediction analysis in junction areas, 103 artificial strong binders were predicted. 14 (14%) of these were not modified (junction 4, 7, 8, 9, 10, 17 and 27). The number of remaining epitopes (89) was reduced by 66% to 30 (Supplementary Table 8-10, page 305). Of these 30 epitopes, 4 were new predicted strong binders (defined as new predicted HLA type, or entirely different sequence to previously predicted epitopes, marked in purple). 22 epitopes were modified epitopes (defined as modified predicted strong binders, predicted for the same HLA type, partly overlapping with previously predicted epitope but now including linker regions, marked light red) predicted containing linker amino acids. For 4 epitopes, a stronger binding score was calculated after modification (marked in red).

Overall, numbers of modified predicted strong binders were reduced by 49%. An example for removal of strong binders within several junction regions, evaluated by NetMHC, Syfpeithi and BIMAS, is shown in Figure 5-26. 24% of predicted epitopes were not modified, since predictions did not overlap between servers or modifications did not results in a decreased binding score.

Figure 5-26: Examples for abrogated predicted strong binders by insertion of linkers

				score without linker				score with linker			
ID	Junction sequence	Linker	HLA type	peptide	NetMHC	Syfpeithi	BIMAS	peptide	NetMHC	Syfpeithi	BIMAS
24	CEKMALYDVVPMGFSYDTRC	GG	A*2402	LYDVVPMGF	0.4	22	140	LYDVVGGPMGF	0.8	11	0.5
25	TVTESDIRTESLTERLYVGG	GGSGG	A*01	RTESLTERLY	0.12	30	112.5	RTEGGSGGSLTERLY	2	17	2.25
26	ASAACRAAKLCTMLVCGDDL	GGSGG	B*2705	CRAAKLCTM	0.5	21	600	CRAAKLGGSGGCTM	5	12	27

Binding scores before and after insertion on linkers are shown at 2 predicted epitopes with strong binding predictions. Strong binders were defined as: NetMHC <0.5, Syfpeithi >20, BIMAS >100. In these cases, strong binders were removed by the insertion of linkers.

5.6 Adaptations of the designed immunogens

5.6.1 TPA leader sequence and invariant chain

There is evidence that the tissue plasminogen activator (tPA) leader sequence encodes a strong signal peptide for protein secretion, thus mediating the secretion of antigen into extracellular space (Weinberger et al. 2013). Secreted antigen can be taken up and processed by APC, which then promotes the presentation of antigenic material onto MHC class-II complexes, resulting in presentation of antigenic peptides on MHC-II. Peptides will also be presented on MHC class-I molecules by a mechanism called cross-priming (see 1.5.1 Antigenic peptide presentation - MHC class-I and -II processing, page 44).

It was shown in several studies that the inclusion of a TPA leader increased the expression of viral proteins such as Rotavirus VR4 protein (Choi et al. 1998) and the SIV envelope protein. In addition, in a comparison of leader sequences, the TPA leader induced highest levels of HIV-gp120 secretion in a baculovirus system (Golden et al. 1998). Based on this evidence, we included a TPA leader sequence (MDAMKRGLCCVLLLCGAVFVSPSQEIHARFRR) in the conserved immunogens. Newly created junction sides between the TPA leader and the first segment (junction 0) were evaluated for putative epitopes using NetMHC, Syfpeithi and BIMAS servers, and modified accordingly in the case of predicted strong binders in junction areas.

5.6.2 BLAST assessment for selected immunogens

Since we generated artificial junction sites between conserved segments, there was a possibility of creating vaccine-induced epitopes with potential cross-reactivity reactivity to human epitopes. To evaluate junction sites and to exclude potential

epitopes, we performed a BLAST assessment for all junction regions (including TPA leader sequence – first HCV segment) of conserved immunogens for HCV gt1, HCV gt1/3a and HCV gt1-6. Linker regions, defined as 10 amino acids on both sides of the junction, plus the linker sequence, were blasted using the ncbi blast website (http://blast.ncbi.nlm.nih.gov). Since $CD8^+$ T cell epitopes are typically 8 to 11 amino acids in length, we aimed to modify junction sequences with a homology to human proteins of more than 7 amino acids.

All junction regions (after linker insertion, or without linker if none was included) had a sequence homology to human proteins lower than 7 amino acids. Therefore, no additional adjustments of junctional sequence material based on BLAST results were made.

5.6.3 Safety

In the immunogen design approach described here we have identified conserved regions that have been spliced together, in some cases with linker regions, to form a single open reading frame. None of the constructs include the complete coding sequence for any structural or non-structural protein in the HCV genome. All seven constructs encode one chimeric protein, and we do not expect this protein to possess any of the functions of native HCV antigens *in vivo* in mice or humans. In our design approach we have taken the additional precaution of modifying linker sequences to avoid the creation of strong T cell epitopes. We additionally blasted junction regions against the human genome to check for homology against human proteins (see 5.6.2). Consequently we have minimized the risk of any auto-reactive immune responses.

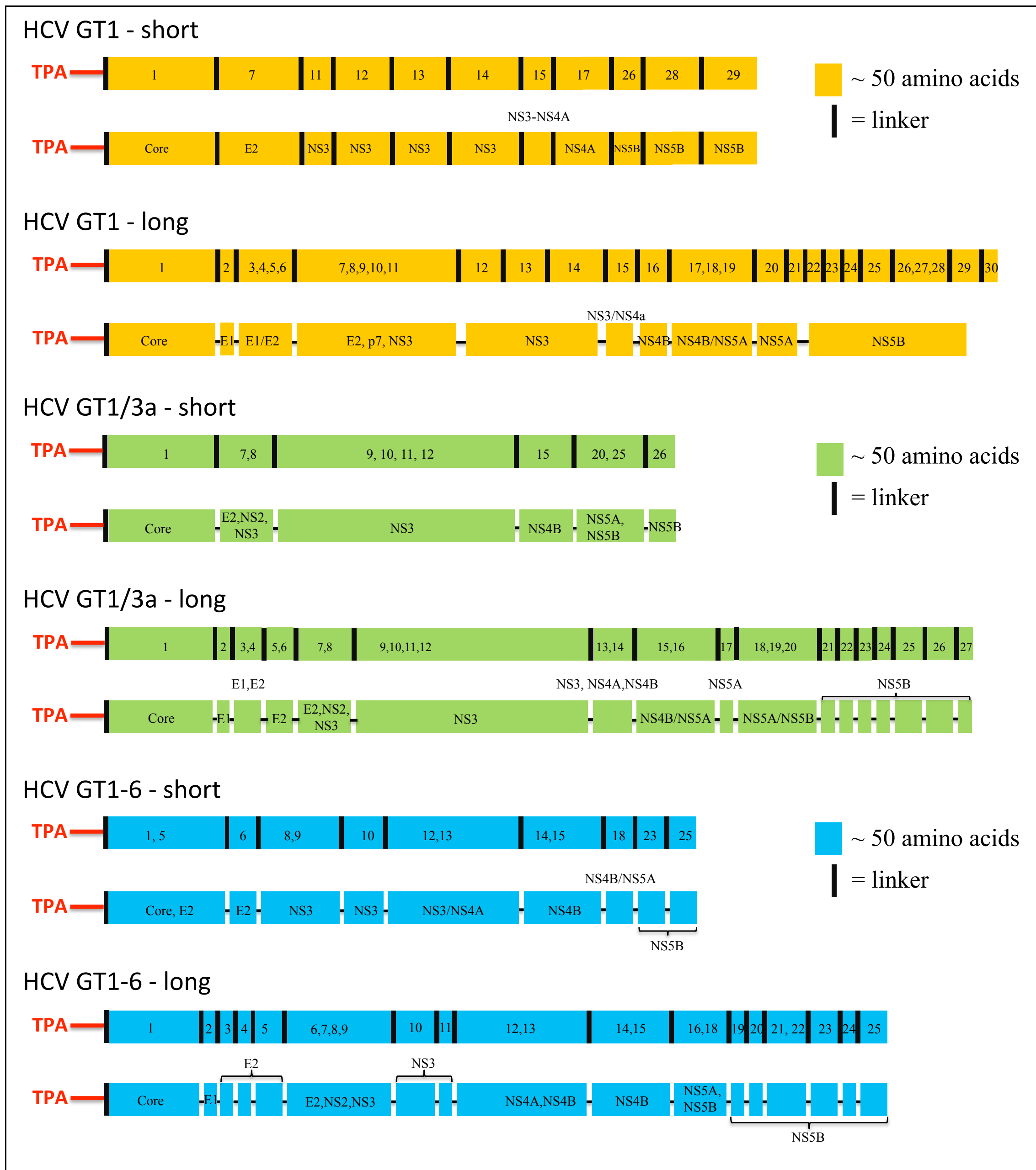
The strategy employed here for HCV is similar to the HIV_{cons} vaccines currently being evaluated by the Hanke group at the Jenner Institute for HIV-1. Based on these studies (Létourneau et al. 2007; Koopman et al. 2013; Borthwick et al. 2013), the risks of oncogenicity, toxicity and allergenicity can be presumed to be negligible.

5.7 Summary of results and discussion

5.7.1 Final immunogens

To address the diversity of HCV, we have designed novel T cell immunogens by assembling the most conserved regions of the HCV proteome. Schematics of all final immunogens are shown in Figure 5-27.

Figure 5-27: Schematics of final immunogens



Schematics of final immunogens for HCV gt1, gt1/3a and gt1-6 including linkers and TPA leader sequence in long (approach D) and short (approach A) versions are shown. The length of the depicted immunogens represents the length of the defined constructs. Additionally, the viral regions from which segments originate are shown.

The segments in each construct represent conserved sequences within Genotype 1, between Genotypes 1 and 3, or Genotypes 1-6. Two of the constructs possess conserved regions of HCV Genotype 1 (HCV gt1 long and short constructs), two constructs contain conserved regions of Genotypes 1 and 3 (HCV gt1/3a long and short constructs), and the final 3 constructs encode conserved sequences from Genotypes 1-6 (HCV gt1-6, 2 long and 1 short construct). Conserved segments were joined to constructs of a final length of 832 AA (gt1), 819 AA (gt1/3a), 820 AA for the short constructs, and 1543 AA (gt1), 1443 AA (gt1/3a) or 1376 AA (gt1-6) for the long constructs.

The designed HCV immunogens are novel and differ from previously described HCV vaccine inserts. They are assembled from protein regions rather than epitopes (Firbas et al. 2006; Firbas et al. 2010), enabling a broader coverage of epitopes presented by multiple HLA types rather than epitopes based vaccines. In addition, they combine patient sequences from different strains similar to the overall consensus of each segment, ensuring possible priming of broad, cross-reactive immune responses.

Immunogens for all genotypes have been evaluated for potential immunogenicity using two independent approaches: first, epitopes previously described in natural infection mapped onto conserved regions, second, putative T cell epitopes were evaluated using three freely available online epitope prediction programs. Subsequently, linkers were included in places where strong artificial, non-HCV epitopes were predicated in junction areas between segments. For one construct, we did not include linkers between conserved segments (HCV gt1-6, long), to evaluate immunogenicity of junction regions *in vivo*.

Final adaptations of designed immunogens included the insertion of a TPA leader sequence to improve MHC class-II presentation. Epitopes were blasted to exclude potential cross-reactivity to human epitopes.

5.7.2 Discussion of results

Designing an effective HCV vaccine is challenging: a vaccine should induce immune responses which ideally prevent, but at least control infection, even though immune responses in natural HCV infection often fail to achieve this.

One of the key obstacles in HCV immunogen design is the extensive variability of the virus (D. B. Smith et al. 2013). This genetic diversity makes it difficult to design one

single vaccine that covers all different strains and subtypes of the virus (Houghton & Abrignani 2005) .

Initial effort in HCV vaccine design has focused on sterilizing immunity using an E1/E2 heterodimer dimer (Verstrepen et al. 2011). Although neutralizing antibodies can play a role in HCV clearance (Osburn et al. 2009; Dowd et al. 2009; Pestka et al. 2007), specific envelope sequences that drive neutralizing AB generation remain unknown. Strategies that use only a handful of epitopes or maximize recognition of several epitope variants (Firbas et al. 2006; Firbas et al. 2010) may optimize specific recognition of a single epitope, but most likely at the expense of other regions needed for immune mediated control. Other strategies included the selection of a single viral strain, as in the vaccine previously trialled in our laboratory (Barnes et al. 2012). However, single sequence selection is always arbitrary, since a single strain cannot resemble the majority of circulating viral strains in a population. This is reflected by limited cross-reactivity of vaccine-induced responses seen between genotypes and circulating viral strains (Swadling et al., manuscript in preparation; Kelly et al., manuscript in preparation).
In the past, HCV research has often benefitted from results of studies on other variable viruses, particularly in HIV research. The HIV vaccine design field has been extraordinarily busy over the last years, approaching the challenge of designing immunogens for this highly variable pathogen in a multitude of ways. Attempted approaches included poly-epitope vaccines, such as HIVA and EP HIV-1090, which have exhibited only poor immunogenicity in clinical trials. Other strategies focused on central vaccines based on HIV ancestral sequences and polyvalent mosaic vaccines, with more successful results in preclinical studies (reviewed by (B. T. Korber et al. 2009)). A strategy similar to the one described in this thesis has been attempted by Thomas Hanke's research group, with promising results in preclinical and clinical trials. Nonetheless, a study by a different group in rhesus monkeys assessing the immune response elicited in response to vaccination with a full-length HIV vaccine in comparison to a conserved-region-only vaccine saw greater magnitude using the full-length construct compared to the conserved-region-only construct (Stephenson et al. 2012). However, comparable breadth of responses was observed, showing no clear benefit of the full-length construct over the conserved-regions-only immunogen. T cells primed by conserved immunogens are expected to be highly cross-reactive and

target a multitude of circulating viral strains. In addition, T cells targeting conserved viral regions might be particularly beneficial, since immune mediated mutations in these regions may cause high fitness costs to the virus.

Here, I describe the design of a vaccine based on conserved sequences of HCV subtypes 1 to 6. To target common viral strains in different regions, immunogens were designed for three different genotype selections: HCV genotype 1a/1b, HCV genotype 1/3a and HCV genotype 1-6. The rationale for this selection of sequences was that HCV gt1 is the most common genotype world wide, and prevalent in the US, Europe, Russia, China, Australia and South America. An immunogen for HCV genotypes 1 and 3a was designed to cover the most common subtypes in the UK; and an immunogen based on conserved regions between HCV genotypes 1 to 6 would ideally protect against strains worldwide.

Sequence selection for immunogen design: The input datasets were generated from sequence material deposited on the Los Alamos database, including additional sequence material derived in-house for HCV gt3. Only full-length sequences deposited were selected to generate the different sequence datasets. Sequences were checked manually for human origin, and to ensure high diversity and minimize bias sequences from different research groups and areas were selected for each genotype.

However, since only full-length sequences were included in the dataset, sequence selection per se was biased: only research groups with sufficient funds and according research interest will deposit full-length sequences. Furthermore, numbers of full-length sequences for the different genotypes found on the databases vary massively; only 2 full-length sequences were reported for HCV genotype 5, whereas 762 full-length sequences were reported for only subtype 1a. We tried to account for this bias by weighing subtypes equally; the same numbers of sequences were included in the entry sequence dataset for each genotype. However, since we did not want to limit numbers of sequences for each genotype to 3, and, at the same time, wanted to weigh subtype 5 equally to other genotypes, so that sequences for subtype 5 were multiplied to match the other genotypes. Included genotype subtypes were chosen according to common circulating viral strains, for example subtype 3a for HCV genotype 3.

A different possible approach would have been a use of all sequences deposited on the database, including sequence segments: this would have increased diversity, but not decreased publication bias.

Patient sequences vs. consensus sequences: In contrast to the HIV_{cons} immunogen (Létourneau et al. 2007), patient sequences rather than consensus sequences have been used in the HCV designed immunogens. The rationale for using patient sequences was based on the observation that a consensus sequence derived from up to 6 HCV genotypes may hold artificial epitopes. In addition, escape mutations can become the dominant sequence for epitopes where the restricting HLA type is common, as it was demonstrated for a common HCV HLA A*01 epitope (Neumann-Haefelin et al. 2008). To prevent these limitations, patient sequences with a high similarity to the consensus sequence were selected. This does prevent the generation of artificial epitopes, however, escape mutations due to immune pressure in the individual patient might be present.

To select patient sequences for the immunogens, a separate analysis was performed for each conserved segment. First, the subtype most similar to the overall consensus was defined, and then a patient sequence most similar to the overall consensus was selected. Overall, selected patient sequences were very similar to the overall consensus, with all sequences showing over 90% homology with the overall consensus (99.6% for gt1, 97.4% for gt1/3a and 94% for gt1-6 constructs). HCV genotype 1 sequences were most commonly included.

Instead of patient sequences, it would have been possible to include ancestral sequences, as described for HCV in a recent publication by Andrea Cox's group (Munshaw et al. 2012; Burke et al. 2012). To re-construct ancestral sequences, phylogenetic calculations build genetic trees where recent, host-specific changes are placed close to the tips, whereas sequences deeper in the tree reflect early shared ancestors (Bhattacharya et al. 2007). Although the ancestral sequence is similar to currently circulating viral strains, they may not reflect patient-specific escape mutations. Cox and colleagues describe a better ability to expand functional and cross-reactive T cell responses to known HCV epitopes using the representative phylogenetic HCV 1a ancestral sequence compared to a gt1a consensus sequence. In addition, the ancestral sequence contained the greatest number of optimal epitopes (Burke et al. 2012). However, T cell reactivity was assessed in a surrogate model of expansion of pre-existing T cell responses in chronically infected individuals; ideally, vaccination of naïve hosts with the consensus and ancestral sequence would have been performed.

Potentially limited immunogenicity: A second potential weakness of designed immunogens is that the immunogenicity of conserved regions might be limited. Reasons for relative conservation of these regions might be related to structural and functional constraints, but could also reflect a lack of immunological pressure, e.g. absence of epitopes that are presented and select for immune mediated mutations. I have addressed the problem of a potential lack of immunogenicity using two approaches:

First, epitopes described in natural infection were mapped on the HCV variability curve, to analyse whether they fall into variable regions excluded from the immunogens, or conserved regions included in the immunogens. Conserved regions are well populated with known, but less dominant, HCV CD4 and CD8+ restricted epitopes. For HCV genotype 1 epitopes, similar amounts of epitopes in conserved and non-conserved regions were observed, proving that non-conserved regions indeed have the potential of priming immune responses. A comprehensive analysis of HCV epitopes described in natural infection based on compiled data from the IEDB found that epitopes frequently recognized by both T cells and B cells correlated with low variability at sequence level (Y. Kim et al. 2012). This observation was not replicated in this thesis. This might be due to the fact that in Kim et al.'s work, all published epitope data was used for the analysis, whereas this study crosschecked IEDB data with original publications, and non-human data. Inaccurately included positives and duplicate epitopes within the same publication were excluded from the analysis. Furthermore, the entropy analysis in Kim et al. included sequence data across all six genotypes, subsequently comparing variability to T cell data obtained in genotype 1, whereas we compared epitope data to only gt1 sequence variability.

Second, I performed an epitope prediction analysis to estimate potential immunogenicity of designed constructs. High numbers of strong binders were predicted for conserved regions using the online epitope prediction server NetMHC, suggesting that conserved regions indeed possess the ability to induce potent immune responses. However, this analysis was only performed for MHC class-I complexes; leaving the potential generation of CD4+ restricted epitopes unaddressed.

Leading from these results, we expect conserved regions to be immunogenic, even though this has to be proven *in vivo* by trials in mice and humans. Supporting this assumption, strong immune responses to conserved regions immune-subdominant in natural infection were detected using a similar conserved HIV construct in mice,

macaques and humans (Létourneau et al. 2007; Rosario et al. 2012; Borthwick et al. 2013). In experiments in mice, it was also shown that serially deleting immunodominant epitopes tripled the frequencies of immune responses to previously subdominant epitopes (Létourneau et al. 2007).

However, algorithmic epitope prediction has multiple potential pitfalls, since the "real life" differs from computer predictions, which can always only re-model a small part of the complex path of epitope presentation. About 90% of immunodominance can be explained by the finding that only ~1% of peptides bind to MHC class-I complexes with sufficient stability to be presented (Yewdell 2006). Therefore, the epitope prediction analysis in this work was limited to the peptide binding step. However, once the threshold for peptide binding to the MHC complex is met, other factors like the location within the polypeptide, proteasomal degradation, peptide abundance, and thymic education and T cell pre-cursor frequencies seem to play a dominant role in determining immunodominance (Yewdell 2006; Assarsson et al. 2007). In addition, only poor correlation between strong peptide binders and recognition of T cell responses was observed in previous *in vivo* experiments (Burke et al. 2012).

Epitopes linked to protection: Four epitopes linked to protection were described in HCV infection (A. Y. Kim et al. 2010; Neumann-Haefelin et al. 2010; Fitzmaurice et al. 2011), however, to date no epitopes have been linked to protection in HCV gt3 infection. To define whether epitopes linked to protection were included in the conserved immunogens, protective epitopes were mapped on selected conserved segments. In each immunogen (gt1, gt1/3a, gt1-6), only a single epitope was fully included (gt1: NTRPPLGNW, gt1/3: **T**SKKTPMGF, gt1-6: **V**RM**V**LMTHF). However, the sequence for the two epitopes included in the gt1/3 (**T**SKKTPMGF) and gt1-6 (**V**RM**V**LMTHF) construct were different to epitopes sequences described as linked to protection (KSKKTPMGF, ARMILMTHF). Since it was recently reported that no cross-reactivity was observed between the HCV gt1 and gt3 version of epitope ARMILMTHF (Skibbe et al. 2014), only limited protective benefit is expected at this epitope when priming with the gt3 version **V**RM**V**LMTHF. Theoretically, it would have been possible to extend the conserved set to include epitopes that were previously described in acute infection. We decided against this approach, to keep the immunogen design stringent and limited to conserved regions defined initially.

Artificial non-HCV epitopes in junction areas –linker design: The combination of multiple conserved segments in one open reading frame immunogen potentially leads to the formation of artificial, non-HCV epitopes. To avoid these potentially immunodominant targets, linkers were designed at junction sites where strong binders were predicted. However, linker design was complicated by divergent results obtained by different peptide prediction servers used. Therefore, linker design was only attempted in regions where strong binders were predicted by two independent prediction algorithms, In addition, where designed linkers did not abrogate predicted strong binders, no linkers were inserted in the final immunogen. However, this approach is error prone: a potential strong binder only predicted by a single prediction server will not be modified. In addition, prediction data will not reflect *in vivo* immunodominance, and unnecessary linkers may be included, potentially altering the peptide presentation and immune responses towards HCV peptide material.

In summary, 6 immunogens based on regions conserved between viral genotypes (for 1a/1b, 1/3a, 1-6) have been designed. Patient sequences most similar to the segment consensus were selected for final immunogens, ensuring cross-genotypic reactivity. Conserved segments contain multiple epitopes described in natural infection, suggesting *in vivo* immunogenicity of designed constructs; additionally, strong binding peptides were predicted for conserved segments. Predicted artificial epitopes in junctional regions (~20%) were abrogated through the insertion of 2-6 amino acid linkers. To optimize peptide presentation, a TPA leader sequence was included in all immunogens.

5.7.3 Cross-reactive HCV vaccine – the future

In this thesis, 7 HCV immunogens based on conserved viral regions were defined (HCV_{cons}). To progress to clinical use, immunogens will have to be further assessed in animal models and Phase 1 to 3 clinical studies. However, it will not be feasible to assess seven different immunogens in human Phase I to III studies. Therefore, selection criteria will based on experimental studies and immunological results will be necessary.

Selection of a delivery modus – viral vectors

Viral vaccine vectors are known to induce strong and broad T cell responses in healthy volunteers (Draper & Heeney 2010). Viral vectors based on chimpanzee adenoviral strains have been successful in priming T cell responses, with the

additional advantage that pre-existing immunity to chimpanzee strains is lower than that to rare human subtypes (Dicks et al. 2012; Antrobus et al. 2014), whereas Modified Vaccinia Ankara (MVA) as a viral vector has been successful in boosting T cell responses primed by adenoviral vectors (McConkey et al. 2003; Gherardi et al. 2003; Webster et al. 2005). An HCV non-structural insert (HCV-NSmut) in a prime-boost regimens using chimpanzee adenoviruses and MVA induced stronger immune responses than combinations of human and chimpanzee adenoviruses (Barnes et al. 2012); Swadling et al, in submission).

For immunogens designed in this thesis, a replication-defective E1/E3 deleted chimpanzee adenovirus vector (ChAdOx1) will be used (Dicks et al. 2012). ChAdOx1 including HCV_{cons} was manufactured by the Viral Vector Core Facility (Jenner Institute, University of Oxford, Oxford). In this vector, the expression of the HCV_{cons} immunogen is driven by the human cytomegalovirus immediate early promoter, and the conserved immunogens were inserted at the E1 locus of the ChAdOx1 genome.

To generate MVA-HCV_{cons} vectors, the expression cassette containing the conserved immunogen HCV_{cons} was subcloned into the MVA shuttle vector pMCA-GFP-TD flanked by TKL (thymidine kinase gene left region) and TKR (thymidine kinase gene rightregion) generating the transfer vector pMVA-GFP-TD-HCV_{cons}. pMVA-GFP-TD-HCV_{cons} drives the antigen expression using the vaccinia P7.5 early/late promoter, and expression of green fluorescent protein (GFP) using the fowlpox late promotor, FP4b. The production of recombinant MVA-HCV_{cons} was based on in vitro recombination between the MVA-Red genome and homologous sequence (TKL and TKR) within the transfer vector pMVA-GFP-TD-HCV_{cons}.

Assessment of HCV immunogens in animal models- basic immunogenicity

The designed conserved HCV immunogens are chimaeric and not natural proteins, and concatenating conserved fragments may impact the processing of intact epitopes that are embedded within the fragments. It is therefore important to demonstrate that the immunogens are capable of inducing T cells in pre-clinical animal model.

However, lack of animal models remains a major challenge in HCV vaccine development. The only animal model for HCV are chimpanzees, with other primate models like macaques and rhesus monkeys being unsusceptible to HCV (J. Bukh et al. 2001). However, the use of chimpanzees in HCV research is banned in Europe and has recently been limited in the USA (National Academy of Sciences 2011). We will therefore resort to testing our immunogens in inbred and outbred mice, aiming to

prove basic immunogenicity, even though results will not represent epitope presentation in humans. Mice with human adaptive immune functions (Pajot et al. 2004) may fill this gap, however, will only present epitopes specific for transgenic HLA alleles. Immunogenicity will be compared to results obtained using the existing HCV vaccine with an HCV-gt1b-NSmut insert. An immunogen that generates large numbers of CD8+ and CD4+ T cell responses with a broad range of presented epitopes will be preferentially selected over one with limited epitope presentation at low magnitude.

Alternative assessments of immunogenicity

A potential alternative may be the evaluation of immunogenicity in an *in vitro* system. Even though possible models have limitations inherent to any *in vitro* system, this could enable an assessment of epitopes presented by several dominant HLA types, rather than the preclinical assessment of only one single HLA in humanized mice. One possible experimental scenario would be the transfection of dendritic cell lines expressing defined, dominant MHC complexes with the designed immunogens, and subsequently eluting and analysing presented peptides by mass spectrometry. If facilitated as a high-throughput method, this approach could allow routine analyses of vaccine constructs prior to the investigation in animal and human models. The breadth of epitope presentation defined using this approach could serve as an additional selection parameter when ranking conserved immunogens for clinical studies.

Selection of a single conserved immunogen for clinical studies

An ideal HCV immunogen should be highly immunogenic, covering as many genotypes as possible and inducing maximal number of cross-reactive T cell responses. If, due to financial considerations, only a small number of immunogen can be tested in clinical studies, pre-clinical parameters will have to serve for the selection of a single immunogen.

An obvious selection criterion is the genotype coverage of the conserved immunogens. If sufficient immunogenicity can be proven in pre-clinical studies, HCV_{cons} immunogens covering HCV genotypes 1 to 6 or HCV genotypes 1 and 3 will be chosen for further studies. In addition, it has been shown in HIV infection that full-length HIV immunogens induce greater magnitude of T cell responses and comparable breadth in comparison to shorter conserved immunogens (Stephenson et al. 2012). Therefore, the longer HCV_{cons} immunogen version will be favoured over the shorter one, if manufacturing can be accomplished.

Phase I studies - safety

Phase I studies will enrol a small number of individuals at low risk for HCV infection. The primary outcome measured is safety, and adverse events will be recorded in all vaccinated individuals. As a secondary outcome, vaccine immunogenicity will be monitored. Despite the uncertainty what constitutes an effective HCV immune response, the breadth and magnitude of HCV specific CD4+ and CD8+ responses will be assessed. As previously shown for the HCV-NSmut vaccine, which was administered using viral vectors (AdCh3, AdHu6, MVA), the magnitude and breadth of adaptive immune responses targeting the immunogen and anti-vector immunity will be assessed for the HCV_{cons} immunogens (Folgori et al. 2006; Barnes et al. 2012). Phase I trials may also gather data on vaccine doses and vaccine schedules, parameters currently being assessed in our laboratory using the HCV-NSmut vaccine. It is likely that results will influence the study outline for the HCV_{cons} immunogens.

Phase II/III studies – efficacy and further data on safely and immunogenicity

Phase II studies will enrol larger numbers of individuals and aim to yield further data on safety and immune responses elicited by the vaccine. Phase III studies are designed to test whether a vaccine provides protection against HCV infection and can include individuals at greater risk of infection. HCV infection rates in vaccinated individuals are to be compared to infection rates in placebo vaccinated individuals, in a double-blinded study layout. A realistic goal for vaccine efficiency should be the prevention of HCV chronic infection.

Unlike in malaria (Webster et al. 2005), it is currently not ethical to perform HCV challenge studies to vaccine efficacy. However, given the astonishing efficiency of new HCV drugs with treatment cure rates over 95%, the perception on whether it is ethical to infect a patient with HCV for trial purposes might undergo a substantial change over the coming years.

The current state of the art is to evaluate the efficacy of HCV vaccines in prospective studies performed in at risk groups, where HCV prevalence and incidence rates of the cohort are known. A current Phase II study in intravenous drug using cohorts in San Francisco/Baltimore assesses HCV infection rates in 300 individuals, either vaccinated with the HCV genotype 1b vaccine HCV-NSmut (n=150) or placebo vaccinated (n=150) (ClinicalTrials.gov 2014). In these cohorts, about 90% of patients are infected with HCV genotype 1, and 10% with HCV genotype 3. Results of this study are expected in 216. Potential suitable cohorts in the UK include cohorts of

intravenous drug users (IVDU) and men who have sex with men (MSM). The advantage of MSM cohorts is the high compliance, however, epidemic rates are though to be peaking and new infections are declining. Well-defined IVDU cohorts in the UK exist in Glasgow and Plymouth, and the mix of infecting genotypes (50% HCV gt3, 50% HCV gt1) would make them suitable for testing a vaccine aiming to induce T cell responses cross-reactive between genotypes. To date, no HCV vaccine has reached phase III development (L. M. J. Law et al. 2013).

Clearly, there are major challenges to take until a new vaccine reaches clinical use. But, despite the new HCV drugs, with global prevalence rates at 170 million, an efficient vaccine inducing T cell responses cross-reactive against several HCV genotypes will be necessary to combat HCV infection worldwide.

6 Concluding remarks

HCV infection constitutes a major health burden, leading to chronic infection with liver fibrosis and hepatocellular cancer. Major advances in HCV treatment have been achieved over the recent years; however, the absence of HCV screening programs, high treatment costs and limited access in low and middle-income countries with high HCV prevalence rates remain important challenges. The administration of an effective HCV vaccine will undoubtedly be key to broadly prevent new infections, especially in high-risk groups like intravenous drug users, and crucial for HCV eradication. However, the significant genetic diversity is a major challenge in the development of a viral vaccine.

This thesis aimed to design a HCV vaccine immunogen priming T cell responses active against all major viral genotypes.

To define the need for a cross-reactive vaccine, I initially assessed whether T cell specificity differs between HCV subtypes, and whether immunodominant responses are cross-reactive. So far, extensive research has been published on HCV gt1 specific immunity, but data on T cell immunity towards other genotypes remains scarce. I therefore comprehensively evaluated T cell specificity against HCV gt3, which is now the major infecting genotype in the UK, and cross-reactivity between HCV subtypes 1 and 3. T cell specificity was markedly different between HCV gt1 and gt3, and limited T cell cross-reactivity was observed at immunodominant epitopes. Leading from this data, a vaccine immunogen was designed based entirely on regions conserved between HCV genotypes.

Comprehensive assessment of HCV gt3 specific T cell responses

To assess T cell responses in HCV gt3 infection, 140 patients (acute (n=16), chronic (n=108)) infected with HCV and patients who spontaneously resolved infection (n=16 patients) were screened using IFNγ ELISpot assays. T cell responses were assessed using an HCV gt3 specific peptide set (15-18aa in length, overlapping by 11aa) based on a cohort consensus sequence spanning the whole viral genome.

Similarly to T cell responses previously described in HCV genotype 1 infection (C. L. Day et al. 2002; Lauer et al. 2004; A. L. Cox, Mosbruger, Lauer, et al. 2005), strong and broad T cell responses were detected in spontaneously resolved infection. In contrast, T cells in chronically infected individuals were weak and mainly targeting

two epitopes in the core (CD4+ restricted) and NS3 (CD8+ restricted) region. In acute HCV infection, patients who subsequently cleared infection had higher T cell responses than those who developed chronic infection, however, this difference was not statistically significant. No association was found between magnitude of T cell response and clinical outcome after therapy or between the magnitude of T cell response and other known predictors of treatment response.

Distribution of CD4 and CD8 responses targeting the different HCV regions is distinct in spontaneous resolvers and chronically infected patients: in spontaneously resolved patients, non-structural regions were mainly targeted by CD4 cells, whereas CD8+ restricted cells targeted both HCV structural and non-structural regions. In contrast, CD4+ responses in chronic patients were only formed against structural proteins and were very low in magnitude, and CD8+ responses mainly targeted HCV non-structural proteins. This result suggests that particularly CD4+ specific responses primed in acute infection targeting HCV non-structural proteins play a role in resolving the infection, as it has been previously suggested in chimpanzee studies (Kaplan et al. 2007).

In addition to the overlapping peptide set used, T cell responses were assessed using an HLA predicted peptide set based on previously described sequence polymorphisms linked patients' HLA types (Rauch et al. 2009). The polymorphisms were evaluated using online epitope prediction tools, and subsequently tested in IFNγ ELISpot assays matched to the individuals HLA type. The rationale for the assessment of T cell responses using a second peptide set was based on several considerations; (1) HCV specific T cell responses are only in the minority of chronically infected patients. It is known that T responses can be missed using overlapping peptides, but detected using the optimal peptide length (Draenert et al. 2004) (2) using an overlapping peptide, binding optimals, CD4/CD8 and HLA restriction have to be defined after detecting a T cell response in ELISpot assays to enable further detailed phenotypic analysis of a T cell response. Since this information is pre-defined when using an HLA predicted peptide, subsequent experiments are less laborious. The two peptide sets used identified distinct epitopes, underlining the relevance of testing for HCV specific responses using optimal HLA matched peptides. Using the HLA predicted peptide set, T cell responses were mainly identified in chronically infected patients with low level responses, but not in spontaneously resolved patients. One possible reason for this

observation is that spontaneous resolution is associated with T cell responses that target HCV virus that is not subject to viral escape, whereas the HLA predicted peptide set was determined through polymorphisms induced by T cell pressure.

HCV gt3 specific T cell responses are markedly different to those detected in HCV gt1 infection

HCV genotypes vary at 20-30% on sequence level (Simmonds et al. 2005). I hypothesized that this variation would lead to a distinct T cell specificity in HCV genotypes. I therefore aimed to assess whether T cell specificity varies between HCV gt1 and gt3 infection.

To compare HCV gt3 and HCV gt1 specific immunity, genotype specific T cell responses were assessed in chronically infected HCV gt1 patients. A comparison of magnitude of T cell responses between chronic HCV gt1 and gt3 showed no difference. However, the distribution of responses to different HCV viral regions was distinct. In HCV gt3 infected individual, non-structural regions were more commonly targeted than in HCV gt1 infection. Similar results have been observed in a study comparing T cell responses in HCV gt1 and non-gt1 patients using a HCV gt1 peptide set (Schulze zur Wiesch et al. 2007).

Next, I determined whether T cell specificity varies between genotypes 1 and 3. 42 HCV gt3 specific epitopes were identified using both peptide sets in IFNγ ELISpot assays. These HCV gt3 specific epitopes were compared to previously described HCV gt1 epitopes in the literature. HCV gt1 epitopes were obtained from the immune epitope database (IEDB, www.iedb.org). T cell specificity was markedly different between HCV gt3 and gt1 infection, supporting the observation that there are major differences in sequence polymorphisms linked to immune mediated pressure between the two genotypes (Rauch et al. 2009). Inter-genotypic differences were especially observed for CD8+ restricted epitopes, whereas a higher overlap in targeted regions was seen for CD4+ restricted epitopes. When comparing epitopes frequently targeted in HCV genotype 1 infection (defined as those deposited 5 or more times on the IEDB) to those detected in HCV gt3 infection in this study, only a single epitope of those dominant in HCV gt1 infection was also detected in this study (ATDALMTGY), which has been previously described as cross-reactive between HCV genotypes.

Major differences between HCV gt1 and gt3 T cell specificity suggest that T cell epitopes primed in natural HCV infection are not those that one would seek to induce

if the aim was to generate a T cell vaccine that was cross-reactive between HCV gt1 and gt3. To further assess this hypothesis, cross-reactivity at epitopes dominant in HCV gt3 infection was evaluated experimentally.

Definition of dominant HCV gt3 specific epitopes

Even though T cell targets in natural infection vary between HCV gt1 and gt3, it is possible that T cells primed with one genotype are cross-reactive against sequence variants from other genotypes. Cross-reactive epitopes might serve as attractive targets for a vaccine covering different viral genotypes. I therefore aimed to assess cross-reactivity at dominant T cell epitopes primed in natural HCV gt3 infection.

Initially, I defined dominant HCV gt3 epitopes as those targeted in more than 4 patients from our cohort. Following this definition, seven epitopes were selected: two dominant CD4+ restricted epitopes in the HCV core region and five dominant CD8+ restricted epitopes within the HCV non-structural regions. Of five CD8+ restricted epitopes, three were detected using gt3 specific HLA predicted peptides and two were detected using overlapping peptide pools. Most immunodominant epitopes were detected in chronic patients, which is likely due to the overrepresentation of chronic patients within the Oxford HCV gt3 cohort in comparison to acute and spontaneously resolved patients.

High sequence variability and limited T cell cross-reactivity was observed within HCV gt3 and between gt3 and gt1 sequence variants

Sequence variability at immunodominant epitopes was initially assessed within HCV gt3; with low observed variability at epitopes targeted by CD4+ T cells, and high variability at CD8+ epitopes detected using HLA predicted peptides. T cell responses against HCV gt3 epitope variant sequences were reduced at all tested sites. Subsequently, sequence variability at frequently targeted epitopes was assessed between HCV gt3 and gt1. High variability was observed at all immunodominant epitopes but one (ATDALMTGY), which was described as conserved between viral genotype previously (Fytili et al. 2008). Next, T cell cross-reactivity between HCV gt1 and gt3 variants was assessed for defined immunodominant epitopes: Limited cross-reactivity of responses primed in natural HCV genotype 3 infection was observed to HCV gt1 sequence variants. Furthermore, phenotypic analysis of immunodominant epitopes revealed limited polyfunctionality of T cells specific for immunodominant epitopes, particularly in chronically infected individuals. This data suggests that responses dominantly targeted in natural infection might not serve as

attractive targets for vaccine aiming to induce responses cross-reactive against several HCV genotypes.

Rationale for a cross-reactive vaccine

HCV is one of the most variable pathogens known, constituting a major challenge for cross-reactive immunogen design. We have previously shown that T cells induced by a potent HCV gt1 vaccine immunogen are poorly cross-reactive between HCV viral genotypes and circulating HCV gt1 variants (Swadling et al. manuscript in preparation, Kelly et al., manuscript in preparation). In addition to this data, I have demonstrated in Chapters 3 and 4 of this thesis that T cell specificity varies between HCV genotypes 1 and 3 and cells primed in natural HCV gt3 infection are poorly cross-reactive against HCV gt1 variants. Therefore, we reasoned that epitopes immunodominant in natural infection might not serve as attractive targets for cross-reactive immunogen design. Based on this experimental data, we aimed to design a vaccine entirely based on viral regions conserved between HCV subtypes.

Design of a HCV vaccine based on conserved viral regions

Immunogens based on conserved viral regions between HCV subtypes 1a/1b, genotypes 1/3a and genotypes 1-6 were designed. Conserved viral segments were determined using a computer algorithm defining variability using pairwise comparisons and selecting conserved segments within the lowest variability quartile. Conserved segments were then joined to form long and short immunogen versions based on length of conserved segments. Patient sequences with high similarity to the overall consensus were included in final immunogen design. Junction regions between conserved segments were optimized to avoid non-HCV epitopes based on prediction data. Finally, a TPA leader sequence was inserted to increase MHC class-II presentation and immunogens were blasted to estimate similarity to human sequences.

Evaluation of potential immunogenicity of HCV vaccine immunogens based on conserved viral regions

A concern when using conserved regions for vaccine immunogen design is the possibility of choosing regions of poor immunogenicity. This might be based on poor proteasomal processing of conserved regions, as well as limited HLA avidity and peptide presentation. Additionally, they might present epitopes for which there is a reduced T cell precursor frequency. To address these potential limitations, we aimed to assess the designed immunogens for their potential immunogenicity.

Ideally, immunogenicity would be evaluated *in vivo*, using either humanized mouse models or in humans. However, humanized mouse models are limited to one single HLA type, and direct assessment in humans is not practicable at this stage of development. We therefore aimed to estimate potential immunogenicity using two surrogate approaches; (i) the evaluation of epitopes detected in natural HCV infection falling into defined conserved regions, and (ii) computer algorithm based epitope prediction. However, these results will not fully represent immunogenicity *in vivo*; computer predictions reliably only utilize peptide binding affinity, whereas only approximate predictions are possible for proteasomal peptide cleavage, TAP processing and endosomal transport. Therefore, predictions in this work only accounted for peptide binding affinity.

I have shown that (i) Epitopes detected in natural infection were distributed across conserved and non-conserved segments. Highly immunodominant HCV gt1 epitopes like CINGVCWTV and KLVALGINAV fell into variable regions.

(ii) Strong binding epitopes were predicted within conserved regions, suggesting that these viral regions are not immunologically inert.

Even though conserved viral regions may be less immunodominant than variable viral regions in natural infection, we hypothesize that potent cross-reactive immune responses will be primed using conserved viral regions if variable viral regions that harbour dominant epitopes are removed from the vaccine. This hypothesis will have to be assessed experimentally. The sub-dominance might be advantageous, since T cell responses elicited towards conserved regions will be different to those mainly detected in natural infection, which fail to control the virus in the majority of infected individuals. Supporting this hypothesis, serial up-ranking of sub-dominant epitopes has been demonstrated previously for an HIV immunogen based on conserved regions (E.-J. Im et al. 2011).

Limitations of this work:

This study assessed cross-reactivity of immune responses in natural HCV gt3 and compared them to HCV gt1 to inform rational cross-reactive vaccine design.

Several limitations are inherent to the study: Firstly, peripheral blood mononuclear cells were obtained to analyse T cell responses in HCV gt3 infection. This represents a common method to assess immunity in HCV infection, since there are ethical issues about obtaining T cells from liver biopsies – a procedure associated with patient

morbidity and mortality. Therefore, relevant cell populations may have been missed, as they might be homing to the liver during infection. However, it has been shown that peripheral blood mononuclear cells reflect the properties of liver resident T cells, albeit at lower frequencies (M. Mueller et al. 2010). Secondly, subtype specific immunity was only assessed in one single HCV genotype, which was compared to the commonly analysed HCV subtype 1. The choice of HCV gt3 was based on the fact that this genotype represents the second most common genotype globally. Other HCV genotypes (2, 4, 5, and 6) were not assessed. Evaluation of these might have facilitated the discovery of immunodominant, broadly cross-reactive epitopes, which could have served as attractive targets for cross-reactive vaccine design. Thirdly, the majority of recruited patients were chronically infected individuals, which biased the definition of immunodominant epitopes that were further analysed. If a larger number of patients with spontaneously resolved infection had been included, a different selection of epitopes would have been defined as immunodominant. This might have led to the definition of highly cross-reactive epitopes conserved between all genotypes that would have constituted attractive vaccine targets.

Viral sequence information was obtained using capillary Sanger sequencing to evaluate variability at dominant epitopes. Better coverage, more depth and the possibility to analyse viral quasi-species in a single host might have been achieved by deep sequencing using techniques like RNAseq (Batty et al. 2013). In addition, complete sequence datasets at immunodominant epitopes might have enabled the proof of immune mediated escape. Further work could include analysis of sequences of regions flanking epitopes, since it has been shown previously that polymorphisms in these regions alter the T cell repertoire in the case of a CD4+ influenza epitope (Cole et al. 2012), which might have implications on sequence versions included in future immunogens.

Although a range of different assays to assess T cell function was performed in this study, the use of HCV gt3 specific tetramers would have enhanced this analysis. Further analyses of CD4+ and CD8+ restricted epitopes would have enabled closer definition of T cell function, such as regulatory properties or exhaustion profiles. In addition, more detailed and extensive analysis may have been achieved using the recent CYTOF technology (cytometry by time-of-flight) (Newell et al. 2012).

I aimed to generate a cross-reactive vaccine using a stringent rational approach based entirely on HCV sequence regions conserved between HCV viral genotypes. However, we used previously deposited sequence information for vaccine design that may have been subject to selection bias. We tried to account for this bias by carefully choosing sequences from different research groups and weighting full-length sequence information from different genotypes equally. A different option for sequence selection would have been the inclusion of all HCV sequences from sequence depository. Even so, this approach would have increased the publication bias and secondary measures would have been necessary to down-rank over-represented sequence information, especially for HCV gt1.

To assess for potential strong binders, immunogens were assessed using freely available online prediction algorithms. However, epitope prediction analyses using computing models can always only estimate the complex cellular process of proteasomal processing, peptide binding and presentation. Therefore, adaptations to vaccine design at junction sites between conserved regions based on epitope prediction data to avoid strong binders are always approximations, and potential immunodominant, non-HCV epitopes cannot be excluded definitively.

Further applications of this work:

A prophylactic vaccine remains a highly desirable goal to prevent the global 3-4 million new HCV infections every year. In this study, immunogens aiming to prime cross-reactive T cell responses have been designed based on conserved viral regions. Designed immunogens will be assessed for immunogenicity in mice, before being trialled in humans.

7 References

Abbvie press releases, 2014. AbbVie Completes Largest Phase III Program of an All-Oral, Interferon-Free Therapy for the Treatment of Hepatitis C Genotype 1 - Jan 31, 2014. Available at: http://abbvie.mediaroom.com/2014-01-31-AbbVie-Completes-Largest-Phase-III-Program-of-an-All-Oral-Interferon-Free-Therapy-for-the-Treatment-of-Hepatitis-C-Genotype-1 [Accessed April 1, 2014].

Abbvie press releases, 2013. AbbVie Releases First of Six Phase III Results from Investigational All-Oral, Interferon-Free, 12-week Regimen, Showing 96 Percent SVR12 in Genotype 1 Hepatitis C Patients New to Therapy - Nov 18, 2013. Available at: http://abbvie.mediaroom.com/2013-11-18-AbbVie-Releases-First-of-Six-Phase-III-Results-from-Investigational-All-Oral-Interferon-Free-12-week-Regimen-Showing-96-Percent-SVR12-in-Genotype-1-Hepatitis-C-Patients-New-to-Therapy [Accessed April 1, 2014].

Abel, M. et al., 2006. Intrahepatic virus-specific IL-10-producing CD8 T cells prevent liver damage during chronic hepatitis C virus infection. *Hepatology*, 44(6), pp.1607–1616.

Aberle, J.H. et al., 2007. CD4+ T cell responses in patients with chronic hepatitis C undergoing peginterferon/ribavirin therapy correlate with faster, but not sustained, viral clearance. *The Journal of infectious diseases*, 195(9), pp.1315–1319.

Abid, K. et al., 2005. An in vitro model of hepatitis C virus genotype 3a-associated triglycerides accumulation. *Journal of Hepatology*, 42(5), pp.744–751.

Accapezzato, D. et al., 2004. Hepatic expansion of a virus-specific regulatory CD8+ T cell population in chronic hepatitis C virus infection. *Journal of Clinical Investigation*, 113(7), pp.963–972.

Adams, H.P. & Koziol, J.A., 1995. Prediction of binding to MHC class I molecules. *Journal of immunological methods*, 185(2), pp.181–190.

Adams, N.J. et al., 1997. Complete Coding Sequence of Hepatitis C Virus Genotype 6a. *Biochemical and Biophysical Research Communications*, 234(2), pp.393–396.

Adinolfi, L.E. et al., 2001. Steatosis accelerates the progression of liver damage of chronic hepatitis C patients and correlates with specific HCV genotype and visceral obesity. *Hepatology (Baltimore, Md.)*, 33(6), pp.1358–1364.

Afdhal, N.H. et al., 2011. Hepatitis C pharmacogenetics: state of the art in 2010. *Hepatology (Baltimore, Md.)*, 53(1), pp.336–345.

Aghemo, A. et al., 2009. The pattern of pegylated interferon-alpha2b and ribavirin treatment failure in cirrhotic patients depends on hepatitis C virus genotype. *Antiviral therapy*, 14(4), pp.577–584.

Agnello, V. et al., 1999. Hepatitis C virus and other flaviviridae viruses enter cells via low density lipoprotein receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 96(22), pp.12766–12771.

Allen, T.M. et al., 2004. Selection, transmission, and reversion of an antigen-processing cytotoxic T-lymphocyte escape mutation in human immunodeficiency virus type 1 infection. *Journal of virology*, 78(13), pp.7069–7078.

Almeida, J.R. et al., 2007. Superior control of HIV-1 replication by CD8+ T cells is reflected by their avidity, polyfunctionality, and clonal turnover. *The Journal of experimental medicine*, 204(10), pp.2473–2485.

Almeida, R.R. et al., 2012. Broad and cross-clade CD4+ T-cell responses elicited by a DNA vaccine encoding highly conserved and promiscuous HIV-1 M-group consensus peptides. *PloS one*, 7(9), p.e45267.

Alric, L. et al., 2000. Study of host- and virus-related factors associated with spontaneous hepatitis C virus clearance. *Tissue antigens*, 56(2), pp.154–158.

Angelico, M. et al., 1997. Chronic liver disease in the Alexandria governorate, Egypt: contribution of schistosomiasis and hepatitis virus infections. *Journal of Hepatology*, 26(2), pp.236–243.

Ank, N. et al., 2008. An important role for type III interferon (IFN-lambda/IL-28) in TLR-induced antiviral activity. *Journal of immunology (Baltimore, Md.: 1950)*, 180(4), pp.2474–2485.

Ank, N. et al., 2006. Lambda interferon (IFN-lambda), a type III IFN, is induced by viruses and IFNs and displays potent antiviral activity against select virus infections in vivo. *Journal of Virology*, 80(9), pp.4501–4509.

Ank, N. & Paludan, S.R., 2009. Type III IFNs: new layers of complexity in innate antiviral immunity. *BioFactors (Oxford, England)*, 35(1), pp.82–87.

Anthony, D.D. et al., 2002. Comprehensive determinant mapping of the hepatitis C-specific CD8 cell repertoire reveals unpredicted immune hierarchy. *Clinical immunology (Orlando, Fla.)*, 103(3 Pt 1), pp.264–276.

Antrobus, R.D. et al., 2014. Clinical assessment of a novel recombinant simian adenovirus ChAdOx1 as a vectored vaccine expressing conserved Influenza A antigens. *Molecular Therapy: The Journal of the American Society of Gene Therapy*, 22(3), pp.668–674.

Appel, N. et al., 2006. From structure to function: new insights into hepatitis C virus RNA replication. *The Journal of biological chemistry*, 281(15), pp.9833–9836.

Arnaud, N. et al., 2010. Hepatitis C virus controls interferon production through PKR activation. *PloS One*, 5(5), p.e10575.

Arnaud, N. et al., 2011. Hepatitis C virus reveals a novel early control in acute immune response. *PLoS Pathogens*, 7(10), p.e1002289.

Assarsson, E. et al., 2007. A quantitative analysis of the variables affecting the repertoire of T cell specificities recognized after vaccinia virus infection. *Journal of immunology (Baltimore, Md.: 1950)*, 178(12), pp.7890–7901.

Bacon, B.R. et al., 2011. Boceprevir for Previously Treated Chronic HCV Genotype 1 Infection. *New England Journal of Medicine*, 364(13), pp.1207–1217.

Barnes, E., Gelderblom, H.C., et al., 2009. Cellular immune responses during high-dose interferon-alpha induction therapy for hepatitis C virus infection. *The Journal of Infectious Diseases*, 199(6), pp.819–828.

Barnes, E. et al., 2012. Novel adenovirus-based vaccines induce broad and sustained T cell responses to HCV in man. *Science translational medicine*, 4(115), p.115ra1.

Barnes, E., Folgori, A., et al., 2009. Phase I trial of a highly immunogenic T-cell vaccine for hepatitis C virus based on novel adenoviral vectors from rare serotypes. *Hepatology*, 50, Supp.4, p.A198.

Barnes, E. et al., 2002. The dynamics of T-lymphocyte responses during combination therapy for chronic hepatitis C virus infection. *Hepatology (Baltimore, Md.)*, 36(3), pp.743–754.

Barouch, D. et al., 1995. HLA-A2 subtypes are functionally distinct in peptide binding and presentation. *The Journal of experimental medicine*, 182(6), pp.1847–1856.

Barrett, S., Ryan, E. & Crowe, J., 1999. Association of the HLA-DRB1*01 allele with spontaneous viral clearance in an Irish cohort infected with hepatitis C virus via contaminated anti-D immunoglobulin. *Journal of Hepatology*, 30(6), pp.979–983.

Barth, H. et al., 2003. Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *The Journal of biological chemistry*, 278(42), pp.41003–41012.

Bassendine, M.F. et al., 2013. Lipids and HCV. *Seminars in immunopathology*, 35(1), pp.87–100.

Batty, E.M. et al., 2013. A Modified RNA-Seq Approach for Whole Genome Sequencing of RNA Viruses from Faecal and Blood Samples. *PLoS ONE*, 8(6), p.e66129.

Bellecave, P. et al., 2010. Cleavage of mitochondrial antiviral signaling protein in the liver of patients with chronic hepatitis C correlates with a reduced activation of the endogenous interferon system. *Hepatology (Baltimore, Md.)*, 51(4), pp.1127–1136.

Bengsch, B. et al., 2007. Analysis of CD127 and KLRG1 Expression on Hepatitis C Virus-Specific CD8+ T Cells Reveals the Existence of Different Memory T-Cell Subsets in the Peripheral Blood and Liver. *Journal of Virology*, 81(2), pp.945–953.

Bengsch, B. et al., 2010. Coexpression of PD-1, 2B4, CD160 and KLRG1 on exhausted HCV-specific CD8+ T cells is linked to antigen recognition and T cell differentiation. *PLoS Pathogens*, 6(6), p.e1000947.

Berthoud, T.K. et al., 2011. Potent CD8+ T-Cell Immunogenicity in Humans of a Novel Heterosubtypic Influenza A Vaccine, MVA−NP+M1. *Clinical Infectious Diseases*, 52(1), pp.1–7.

Beyersdorf, N. et al., 2008. Protection from graft-versus-host disease with a novel B7 binding site–specific mouse anti–mouse CD28 monoclonal antibody. *Blood*, 112(10), pp.4328–4336.

Bhasin, M. & Raghava, G.P.S., 2007. A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes. *Journal of biosciences*, 32(1), pp.31–42.

Bhattacharya, T. et al., 2007. Founder effects in the assessment of HIV polymorphisms and HLA allele associations. *Science (New York, N.Y.)*, 315(5818), pp.1583–1586.

Biggins, S.W. et al., 2012. Projected future increase in aging hepatitis C virus-infected liver transplant candidates: a potential effect of hepatocellular carcinoma. *Liver transplantation: official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society*, 18(12), pp.1471–1478.

Biron, C., 1999. Initial and innate responses to viral infections — pattern setting in immunity or disease. *Current Opinion in Microbiology*, 2(4), pp.374–381.

Biswas, S. et al., 2011. Transgene Optimization, Immunogenicity and In Vitro Efficacy of Viral Vectored Vaccines Expressing Two Alleles of Plasmodium falciparum AMA1. *PLoS ONE*, 6(6), p.e20977.

Bjorkman, P.J. et al., 1987. Structure of the human class I histocompatibility antigen, HLA-A2. *Nature*, 329(6139), pp.506–512.

Bluestone, J.A. & Abbas, A.K., 2003. Natural versus adaptive regulatory T cells. *Nature reviews. Immunology*, 3(3), pp.253–257.

BMS press release, 2013. Efficacy and Safety Study of Pegylated Interferon Lambda-1a With Ribavirin and Daclatasvir, to Treat naïve Subjects With Chronic HCV Genotypes 1, 2, 3, and 4 Who Are Co-infected With HIV - Full Text View - ClinicalTrials.gov. Available at: http://clinicaltrials.gov/show/NCT01866930 [Accessed April 10, 2014].

Bochud, P.-Y. et al., 2009. Genotype 3 is associated with accelerated fibrosis progression in chronic hepatitis C. *Journal of Hepatology*, 51(4), pp.655–666.

Bode, J.G. et al., 2003. IFN-alpha antagonistic activity of HCV core protein involves induction of suppressor of cytokine signaling-3. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, 17(3), pp.488–490.

Boettler, T. et al., 2005. T cells with a CD4+CD25+ regulatory phenotype suppress in vitro proliferation of virus-specific CD8+ T cells during chronic hepatitis C virus infection. *Journal of Virology*, 79(12), pp.7860–7867.

Bolacchi, F. et al., 2006. Increased hepatitis C virus (HCV)-specific CD4+CD25+ regulatory T lymphocytes and reduced HCV-specific CD4+ T cell response in HCV-infected patients with normal versus abnormal alanine aminotransferase levels. *Clinical and experimental immunology*, 144(2), pp.188–196.

Bolitho, P. et al., 2007. Apoptosis induced by the lymphocyte effector molecule perforin. *Current opinion in immunology*, 19(3), pp.339–347.

Le Bon, A. et al., 2003. Cross-priming of CD8+ T cells stimulated by virus-induced type I interferon. *Nature immunology*, 4(10), pp.1009–1015.

Le Bon, A. et al., 2006. Direct stimulation of T cells by type I IFN enhances the CD8+ T cell response during cross-priming. *Journal of immunology (Baltimore, Md.: 1950)*, 176(8), pp.4682–4689.

Booth, D.R. & George, Jacob, 2013. Loss of function of the new interferon IFN-λ4 may confer protection from hepatitis C. *Nature Genetics*, 45(2), pp.119–120.

Booth, J.C. et al., 1998. Comparison of the rate of sequence variation in the hypervariable region of E2/NS1 region of hepatitis C virus in normal and hypogammaglobulinemic patients. *Hepatology (Baltimore, Md.)*, 27(1), pp.223–227.

Borthwick, N. et al., 2013. Vaccine-elicited Human T Cells Recognizing Conserved Protein Regions Inhibit HIV-1. *Molecular therapy: the journal of the American Society of Gene Therapy*.

Bortolotti, F. et al., 2005. Hepatitis C virus (HCV) genotypes in 373 Italian children with HCV infection: changing distribution and correlation with clinical features and outcome. *Gut*, 54(6), pp.852–857.

Boulant, S. et al., 2006. Structural determinants that target the hepatitis C virus core protein to lipid droplets. *The Journal of biological chemistry*, 281(31), pp.22236–22247.

Bowen, D.G. et al., 2008. Variable Patterns of Programmed Death-1 Expression on Fully Functional Memory T Cells after Spontaneous Resolution of Hepatitis C Virus Infection. *Journal of Virology*, 82(10), pp.5109–5114.

Bowen, D.G. & Walker, C M, 2005a. Adaptive immune responses in acute and chronic hepatitis C virus infection. *Nature*, 436(7053), pp.946–952.

Bowen, D.G. & Walker, C M, 2005b. Mutational escape from CD8+ T cell immunity: HCV evolution, from chimpanzees to man. *The Journal of Experimental Medicine*, 201(11), pp.1709–1714.

Boyman, O. et al., 2007. Cytokines and T-cell homeostasis. *Current Opinion in Immunology*, 19(3), pp.320–326.

Brenndörfer, E.D. et al., 2009. Nonstructural 3/4A protease of hepatitis C virus activates epithelial growth factor-induced signal transduction by cleavage of the T-cell protein tyrosine phosphatase. *Hepatology (Baltimore, Md.)*, 49(6), pp.1810–1820.

Brooks, D.G. et al., 2005. Intrinsic functional dysregulation of CD4 T cells occurs rapidly following persistent viral infection. *Journal of virology*, 79(16), pp.10514–10527.

Brown, J.H. et al., 1993. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature*, 364(6432), pp.33–39.

Brusic, V., Bajic, V.B. & Petrovsky, N., 2004. Computational methods for prediction of T-cell epitopes--a framework for modelling, testing, and applications. *Methods (San Diego, Calif.)*, 34(4), pp.436–443.

Bucci, C. et al., 2013. “Favourable” IL28B polymorphisms are associated with a marked increase in baseline viral load in hepatitis C virus subtype-3a infection and do not predict sustained virological response after 24 weeks of therapy. *The Journal of general virology*.

Bui, H.-H. et al., 2005. Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics*, 57(5), pp.304–314.

Bukh, J. et al., 2001. Failure to infect rhesus monkeys with hepatitis C virus strains of genotypes 1a, 2a or 3a. *Journal of Viral Hepatitis*, 8(3), pp.228–231.

Bumgardner, G.L. et al., 1998. A functional model of hepatocyte transplantation for in vivo immunologic studies. *Transplantation*, 65(1), pp.53–61.

Bumgardner, G.L. & Orosz, C.G., 2000. Unusual patterns of alloimmunity evoked by allogeneic liver parenchymal cells. *Immunological reviews*, 174, pp.260–279.

Burbelo, P.D. et al., 2012. Serology-enabled discovery of genetically diverse hepaciviruses in a new host. *Journal of virology*, 86(11), pp.6171–6178.

Burke, K.P. et al., 2012. Immunogenicity and cross-reactivity of a representative ancestral sequence in hepatitis C virus infection. *Journal of immunology (Baltimore, Md.: 1950)*, 188(10), pp.5177–5188.

Burrows, J.M. et al., 2007. The impact of HLA-B micropolymorphism outside primary peptide anchor pockets on the CTL response to CMV. *European Journal of Immunology*, 37(4), pp.946–953.

Busch, M.P. et al., 2006. Correlates of hepatitis C virus (HCV) RNA negativity among HCV-seropositive blood donors. *Transfusion*, 46(3), pp.469–475.

Buus, S et al., 2003. Sensitive quantitative predictions of peptide-MHC binding by a “Query by Committee” artificial neural network approach. *Tissue antigens*, 62(5), pp.378–384.

Cabrera, R. et al., 2004. An immunomodulatory role for CD4(+)CD25(+) regulatory T lymphocytes in hepatitis C virus infection. *Hepatology (Baltimore, Md.)*, 40(5), pp.1062–1071.

Calne, R.Y. et al., 1969. Induction of immunological tolerance by porcine liver allografts. *Nature*, 223(5205), pp.472–476.

Calne, R.Y. et al., 1967. Prolonged survival of liver transplants in the pig. *British medical journal*, 4(5580), pp.645–648.

Cammà, C. et al., 2013. Cost-effectiveness of boceprevir or telaprevir for previously treated patients with genotype 1 chronic hepatitis C. *Journal of hepatology*, 59(4), pp.658–666.

Cammà, C. et al., 2012. Cost-effectiveness of boceprevir or telaprevir for untreated patients with genotype 1 chronic hepatitis C. *Hepatology (Baltimore, Md.)*, 56(3), pp.850–860.

Canadian Paediatric Society, 2008. Vertical transmission of the hepatitis C virus: Current knowledge and issues. *Paediatrics & Child Health*, 13(6), pp.529–534.

Candotti, D. et al., 2003. Frequent recovery and broad genotype 2 diversity characterize hepatitis C virus infection in Ghana, West Africa. *Journal of virology*, 77(14), pp.7914–7923.

Capone, S. et al., 2010. Immune responses against a liver-stage malaria antigen induced by simian adenoviral vector AdCh63 and MVA prime-boost immunisation in non-human primates. *Vaccine*, 29(2), pp.256–265.

Cardozo, C. & Kohanski, R.A., 1998. Altered properties of the branched chain amino acid-preferring activity contribute to increased cleavages after branched chain residues by the “immunoproteasome”. *The Journal of biological chemistry*, 273(27), pp.16764–16770.

Carrington, M. et al., 1999. HLA and HIV-1: Heterozygote Advantage and B*35-Cw*04 Disadvantage. *Science*, 283(5408), pp.1748–1752.

Casino, C. et al., 1999. Variation of hepatitis C virus following serial transmission: multiple mechanisms of diversification of the hypervariable region and evidence for convergent genome evolution. *The Journal of general virology*, 80 (Pt 3), pp.717–725.

Castelli, F.A. et al., 2007. Differential capacity of T cell priming in naive donors of promiscuous CD4+ T cell epitopes of HCV NS3 and Core

proteins. *European journal of immunology*, 37(6), pp.1513–1523.

Chang, K.M. et al., 2001. Differential CD4(+) and CD8(+) T-cell responsiveness in hepatitis C virus infection. *Hepatology (Baltimore, Md.)*, 33(1), pp.267–276.

Chang, K.M. et al., 1997. Immunological significance of cytotoxic T lymphocyte epitope variants in patients chronically infected by the hepatitis C virus. *The Journal of clinical investigation*, 100(9), pp.2376–2385.

Chayama, K. et al., 2012. Dual therapy with the nonstructural protein 5A inhibitor, daclatasvir, and the nonstructural protein 3 protease inhibitor, asunaprevir, in hepatitis C virus genotype 1b–infected null responders. *Hepatology*, 55(3), pp.742–748.

Cheent, K. & Khakoo, S.I., 2009. Natural killer cells: integrating diversity with function. *Immunology*, 126(4), pp.449–457.

Chicz, R.M. et al., 1992. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature*, 358(6389), pp.764–768.

Cho, S.S. et al., 1996. Activation of STAT4 by IL-12 and IFN-alpha: evidence for the involvement of ligand-induced tyrosine and serine phosphorylation. *Journal of immunology (Baltimore, Md.: 1950)*, 157(11), pp.4781–4789.

Choi, A.H. et al., 1998. Particle-bombardment-mediated DNA vaccination with rotavirus VP4 or VP7 induces high levels of serum rotavirus IgG but fails to protect mice against challenge. *Virology*, 250(1), pp.230–240.

Choo, Q.L. et al., 1994. Vaccination of chimpanzees against infection by the hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, 91(4), pp.1294–1298.

Christie, J.M. et al., 1997. Clinical outcome of hypogammaglobulinaemic patients following outbreak of acute hepatitis C: 2 year follow up. *Clinical and experimental immunology*, 110(1), pp.4–8.

Ciuffreda, D. et al., 2008. Polyfunctional HCV-specific T-cell responses are associated with effective control of HCV replication. *European journal of immunology*, 38(10), pp.2665–2677.

Clark, P.J. et al., 2012. Hepatitis C virus selectively perturbs the distal cholesterol synthesis pathway in a genotype-specific manner. *Hepatology (Baltimore, Md.)*, 56(1), pp.49–56.

Clemens, M.J., 2005. Translational control in virus-infected cells: models for cellular stress responses. *Seminars in Cell & Developmental Biology*, 16(1), pp.13–20.

ClinicalTrials.gov, 2014. Staged Phase I/II Hepatitis C Prophylactic Vaccine - NCT01436357. *ClinicalTrials.gov - NCT01436357*. Available at: http://clinicaltrials.gov/show/NCT01436357 [Accessed August 18, 2014].

Coburn, C.A. et al., 2013. Discovery of MK-8742: An HCV NS5A Inhibitor with Broad Genotype Activity. *ChemMedChem*, 8(12), pp.1930–1940.

Coccia, E.M. et al., 2004. Viral infection and Toll-like receptor agonists induce a differential expression of type I and lambda interferons in human plasmacytoid and monocyte-derived dendritic cells. *European journal of immunology*, 34(3), pp.796–805.

Cole, D.K. et al., 2012. Modification of the carboxy-terminal flanking region of a universal influenza epitope alters $CD4^{+}$ T-cell repertoire selection. *Nature communications*, 3, p.665.

Coller, K.E. et al., 2012. Molecular determinants and dynamics of hepatitis C virus secretion. *PLoS pathogens*, 8(1), p.e1002466.

Colonno, R. et al., 2011. 1200 CHARACTERIZATION AND IDENTIFICATION OF PPI-437, PPI-668 AND PPI-833 AS POTENT AND SELECTIVE HCV NS5A INHIBITORS WITH ACTIVITY AGAINST ALL HCV GENOTYPES. *Journal of Hepatology*, 54, Supplement 1, p.S474.

Contreras, A.M. et al., 2002. Viral RNA Mutations Are Region Specific and Increased by Ribavirin in a Full-Length Hepatitis C Virus Replication System. *Journal of Virology*, 76(17), pp.8505–8517.

Cooper, S. et al., 1999. Analysis of a successful immune response against hepatitis C virus. *Immunity*, 10(4), pp.439–449.

Corey, K.E. et al., 2009. Hepatitis C virus infection and its clearance alter circulating lipids: Implications for long-term follow-up. *Hepatology*, 50(4), pp.1030–1037.

Correia, B.E. et al., 2014. Proof of principle for epitope-focused vaccine design. *Nature*, advance online publication. Available at: http://www.nature.com/nature/journal/vaop/ncurrent/full/nature12966.html?WT.ec_id=NATURE-20140206 [Accessed February 12, 2014].

Cox, A.L., Mosbruger, T., Mao, Q., et al., 2005. Cellular immune selection with hepatitis C virus persistence in humans. *The Journal of experimental medicine*, 201(11), pp.1741–1752.

Cox, A.L., Mosbruger, T., Lauer, G.M., et al., 2005. Comprehensive analyses of CD8+ T cell responses during longitudinal study of acute human hepatitis C. *Hepatology*, 42(1), pp.104–112.

Cramp, M.E. et al., 2000. Hepatitis C virus–specific T-cell reactivity during interferon and ribavirin treatment in chronic hepatitis C. *Gastroenterology*, 118(2), pp.346–355.

Crawford, A. et al., 2014. Molecular and Transcriptional Basis of CD4+ T Cell Dysfunction during Chronic Infection. *Immunity*, 40(2), pp.289–302.

Crispe, I.N., 2009. The liver as a lymphoid organ. *Annual review of immunology*, 27, pp.147–163.

Crotta, S. et al., 2002. Inhibition of natural killer cells through engagement of CD81 by the major hepatitis C virus envelope protein. *The Journal of experimental medicine*, 195(1), pp.35–41.

Cui, J. et al., 2006. MHC-BPS: MHC-binder prediction server for identifying peptides of flexible lengths from sequence-derived physicochemical properties. *Immunogenetics*, 58(8), pp.607–613.

Dalgard, O. et al., 2008. Pegylated interferon alfa and ribavirin for 14 versus 24 weeks in patients with hepatitis C virus genotype 2 or 3 and rapid virological response. *Hepatology*, 47(1), pp.35–42.

Dalgard, O. et al., 2004. Treatment with pegylated interferon and ribavarin in HCV infection with genotype 2 or 3 for 14 weeks: A pilot study. *Hepatology*, 40(6), pp.1260–1265.

Davis, G.L. et al., 2003. Projecting future complications of chronic hepatitis C in the United States. *Liver Transplantation*, 9(4), pp.331–338.

Day, C.L. et al., 2002. Broad Specificity of Virus-Specific CD4+ T-Helper-Cell Responses in Resolved Hepatitis C Virus Infection. *J. Virol.*, 76(24), pp.12584–12595.

Dazert, E. et al., 2009. Loss of viral fitness and cross-recognition by CD8+ T cells limit HCV escape from a protective HLA-B27-restricted human immune response. *The Journal of clinical investigation*, 119(2), pp.376–386.

DeGoey, D.A. et al., 2014. Discovery of ABT-267, a Pan-Genotypic Inhibitor of HCV NS5A. *Journal of Medicinal Chemistry*, 57(5), pp.2047–2057.

DeLuca, D.S., Khattab, B. & Blasczyk, R., 2007. A modular concept of HLA for comprehensive peptide binding prediction. *Immunogenetics*, 59(1), pp.25–35.

Diamond, M.S. et al., 2011. Type I interferon is selectively required by dendritic cells for immune rejection of tumors. *The Journal of Experimental Medicine*, 208(10), pp.1989–2003.

Dicks, M.D.J. et al., 2012. A novel chimpanzee adenovirus vector with low human seroprevalence: improved systems for vector derivation and comparative immunogenicity. *PloS One*, 7(7), p.e40385.

Diepolder, H.M. et al., 1995. Possible mechanism involving T-lymphocyte response to non-structural protein 3 in viral clearance in acute hepatitis C virus infection. *Lancet*, 346(8981), pp.1006–1007.

Dill, M.T. et al., 2011. Interferon-induced gene expression is a stronger predictor of treatment response than IL28B genotype in patients with hepatitis C. *Gastroenterology*, 140(3), pp.1021–1031.

Dittmann, S. et al., 1991. Long-term persistence of hepatitis C virus antibodies in a single source outbreak. *Journal of hepatology*, 13(3), pp.323–327.

Doherty, D.G. & O'Farrelly, C., 2000. Innate and adaptive lymphoid cells in the human liver. *Immunological reviews*, 174, pp.5–20.

Dominguez-Villar, M. et al., 2012. Up-regulation of FOXP3 and induction of suppressive function in CD4+ Jurkat T-cells expressing hepatitis C virus core protein. *Clinical science (London, England: 1979)*, 123(1), pp.15–27.

Donnelly, R.P. & Kotenko, S.V., 2010. Interferon-lambda: a new addition to an old family. *Journal of interferon & cytokine research: the official journal of the International Society for Interferon and Cytokine Research*, 30(8), pp.555–564.

Dönnes, P. & Kohlbacher, O., 2006. SVMHC: a server for prediction of MHC-binding peptides. *Nucleic acids research*, 34(Web Server issue), pp.W194–197.

Dore, G.J. et al., 2013. 1418 Daclatasvir combined with PegInterferon alpha-2a and Ribavirin for 12 or 15 weeks in patients with HCV genotype 2 or 3 infection: COMMAND GT2/3 STUDY. *Journal of hepatology*, 58, pp.S570–S571.

Dowd, K.A. et al., 2009. Selection pressure from neutralizing antibodies drives sequence evolution during acute infection with hepatitis C virus. *Gastroenterology*, 136(7), pp.2377–2386.

Doyle, S.E. et al., 2006. Interleukin-29 uses a type 1 interferon-like program to promote antiviral responses in human hepatocytes. *Hepatology (Baltimore, Md.)*, 44(4), pp.896–906.

Draenert, R. et al., 2004. Impact of intrapeptide epitope location on CD8 T cell recognition: implications for design of overlapping peptide panels. *AIDS (London, England)*, 18(6), pp.871–876.

Dragan, A.I. et al., 2007. Mechanisms of activation of interferon regulator factor 3: the role of C-terminal domain phosphorylation in IRF-3 dimerization and DNA binding. *Nucleic Acids Research*, 35(11), pp.3525–3534.

Drane, D. et al., 2009. Priming of CD4+ and CD8+ T cell responses using a HCV core ISCOMATRIX vaccine: a phase I study in healthy volunteers. *Human vaccines*, 5(3), pp.151–157.

Draper, S.J. & Heeney, J.L., 2010. Viruses as vaccine vectors for infectious diseases and cancer. *Nature Reviews Microbiology*, 8(1), pp.62–73.

Drexler, J.F. et al., 2013. Evidence for Novel Hepaciviruses in Rodents. *PLoS Pathog*, 9(6), p.e1003438.

Duan, L. et al., 2011. Prediction and identification-based prediction of Chinese hepatitis C viral-specific cytotoxic T lymphocyte epitopes. *Journal of medical virology*, 83(8), pp.1315–1320.

Duffy, S., Shackelton, L.A. & Holmes, Edward C., 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nature Reviews Genetics*, 9(4), pp.267–276.

Duggal, P. et al., 2013. Genome-wide association study of spontaneous resolution of hepatitis C virus infection: data from multiple cohorts. *Annals of internal medicine*, 158(4), pp.235–245.

Dunham, R.M. et al., 2013. Hepatic stellate cells preferentially induce Foxp3+ regulatory T cells by production of retinoic acid. *Journal of immunology (Baltimore, Md.: 1950)*, 190(5), pp.2009–2016.

Duong, F.H.T. et al., 2004. Hepatitis C virus inhibits interferon signaling through up-regulation of protein phosphatase 2A. *Gastroenterology*, 126(1), pp.263–277.

Eastwood, D. et al., 2010. Monoclonal antibody TGN1412 trial failure explained by species differences in CD28 expression on CD4+ effector memory T-cells. *British Journal of Pharmacology*, 161(3), pp.512–526.

Egger, D. et al., 2002. Expression of hepatitis C virus proteins induces distinct membrane alterations including a candidate viral replication complex. *Journal of virology*, 76(12), pp.5974–5984.

El-Serag, H.B. et al., 2002. Extrahepatic manifestations of hepatitis C among United States male veterans. *Hepatology (Baltimore, Md.)*, 36(6), pp.1439–1445.

Elmowalid, G.A. et al., 2007. Immunization with hepatitis C virus-like particles results in control of hepatitis C virus infection in chimpanzees. *Proceedings of the National Academy of Sciences of the United States of America*, 104(20), pp.8427–8432.

Erickson, A.L. et al., 2001. The Outcome of Hepatitis C Virus Infection Is Predicted by Escape Mutations in Epitopes Targeted by Cytotoxic T Lymphocytes. *Immunity*, 15(6), pp.883–895.

European Paediatric Hepatitis C Virus Network, 2005. A significant sex--but not elective cesarean section--effect on mother-to-child transmission of hepatitis C virus infection. *The Journal of infectious diseases*, 192(11), pp.1872–1879.

Evans, M.J. et al., 2007. Claudin-1 is a hepatitis C virus co-receptor required for a late step in entry. *Nature*, 446(7137), pp.801–805.

Eyster, M.E., Sanders, J. & Goedert, J J, 2004. Viral clearance occurs very early during the natural resolution of hepatitis C virus infection in persons with haemophilia. *Haemophilia: the official journal of the World Federation of Hemophilia*, 10(1), pp.75–80.

Falk, K. et al., 1991. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature*, 351(6324), pp.290–296.

Farci, P. et al., 2000. The outcome of acute hepatitis C predicted by the evolution of the viral quasispecies. *Science (New York, N.Y.)*, 288(5464), pp.339–344.

Farmery, M.R. et al., 2000. The role of ERp57 in disulfide bond formation during the assembly of major histocompatibility complex class I in a synchronized semipermeabilized cell translation system. *The Journal of biological chemistry*, 275(20), pp.14933–14938.

Farquhar, M.J. et al., 2012. Hepatitis C virus induces CD81 and claudin-1 endocytosis. *Journal of virology*, 86(8), pp.4305–4316.

Fattori, E et al., 2006. Efficient immunization of rhesus macaques with an HCV candidate vaccine by heterologous priming-boosting with novel adenoviral vectors based on different serotypes. *Gene therapy*, 13(14), pp.1088–1096.

Fernandez-Ponce, C. et al., 2014. CD4+ Primary T Cells Expressing HCV-Core Protein Upregulate Foxp3 and IL-10, Suppressing CD4 and CD8 T Cells. *PLoS ONE*, 9(1). Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3896374/ [Accessed March 13, 2014].

Ferri, S. et al., 2011. Redistribution of regulatory T-cells across the evolving stages of chronic hepatitis C. *Digestive and liver disease: official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver*, 43(10), pp.807–813.

Firbas, C. et al., 2006. Immunogenicity and safety of a novel therapeutic hepatitis C virus (HCV) peptide vaccine: a randomized, placebo controlled trial for dose optimization in 128 healthy subjects. *Vaccine*, 24(20), pp.4343–4353.

Firbas, C. et al., 2010. Immunogenicity and safety of different injection routes and schedules of IC41, a Hepatitis C virus (HCV) peptide vaccine. *Vaccine*, 28(12), pp.2397–2407.

Fischer, W. et al., 2007. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nature Medicine*, 13(1), pp.100–106.

Fitzmaurice, K. et al., 2011. Molecular footprints reveal the impact of the protective HLA-A*03 allele in hepatitis C virus infection. *Gut*, 60(11), pp.1563–1571.

Fleming, V.M. et al., 2010. Virological footprint of CD4+ T-cell responses during chronic hepatitis C virus infection. *The Journal of General Virology*, 91(Pt 6), pp.1396–1406.

Folgori, A. et al., 2006. A T-cell HCV vaccine eliciting effective immunity against heterologous virus challenge in chimpanzees. *Nature medicine*, 12(2), pp.190–197.

Foster, G.R. et al., 2011. Telaprevir alone or with peginterferon and ribavirin reduces HCV RNA in patients with chronic genotype 2 but not genotype 3 infections. *Gastroenterology*, 141(3), pp.881–889.e1.

Foy, E. et al., 2005. Control of antiviral defenses through hepatitis C virus disruption of retinoic acid-inducible gene-I signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8), pp.2986–2991.

Foy, E. et al., 2003. Regulation of Interferon Regulatory Factor-3 by the Hepatitis C Virus Serine Protease. *Science*, 300(5622), pp.1145–1148.

Frank, C. et al., 2000. The role of parenteral antischistosomal therapy in the spread of hepatitis C virus in Egypt. *The Lancet*, 355(9207), pp.887–891.

Frey, S.E. et al., 2010. Safety and immunogenicity of HCV E1E2 vaccine adjuvanted with MF59 administered to healthy adults. *Vaccine*, 28(38), pp.6367–6373.

Friborg, J. et al., 2013. Combinations of lambda interferon with direct-acting antiviral agents are highly efficient in suppressing hepatitis C virus replication. *Antimicrobial agents and chemotherapy*, 57(3), pp.1312–1322.

Fried, M.W. et al., 2002. Peginterferon alfa-2a plus ribavirin for chronic hepatitis C virus infection. *The New England Journal of Medicine*, 347(13), pp.975–982.

Friedrich, T.C. et al., 2004. Reversion of CTL escape-variant immunodeficiency viruses in vivo. *Nature medicine*, 10(3), pp.275–281.

Fuller, M.J. et al., 2010. Selection-Driven Immune Escape Is Not a Significant Factor in the Failure of CD4 T Cell Responses in Persistent Hepatitis C Virus Infection. *Hepatology (Baltimore, Md.)*, 51(2), p.378.

Fytili, P. et al., 2008. Cross-genotype-reactivity of the immunodominant HCV CD8 T-cell epitope NS3-1073. *Vaccine*, 26(31), pp.3818–3826.

Gale, M. & Foy, E., 2005. Evasion of intracellular host defence by hepatitis C virus. *Nature*, 436(7053), pp.939–945.

Gale, M.J., Jr et al., 1997. Evidence that hepatitis C virus resistance to interferon is mediated through repression of the PKR protein kinase by the nonstructural 5A protein. *Virology*, 230(2), pp.217–227.

Gane, E.J. et al., 2013. Nucleotide polymerase inhibitor sofosbuvir plus ribavirin for hepatitis C. *The New England journal of medicine*, 368(1), pp.34–44.

Gane, E.J. et al., 2010. Oral combination therapy with a nucleoside polymerase inhibitor (RG7128) and danoprevir for chronic hepatitis C genotype 1 infection (INFORM-1): a randomised, double-blind, placebo-controlled, dose-escalation trial. *Lancet*, 376(9751), pp.1467–1475.

Gane, E.J. et al., 2014. SOFOSBUVIR/LEDIPASVIR FIXED DOSE COMBINATION IS SAFE AND EFFECTIVE IN DIFFICULT-TO-TREAT POPULATIONS INCLUDING GENOTYPE-3 PATIENTS, DECOMPENSATED GENOTYPE-1 PATIENTS, AND GENOTYPE-1 PATIENTS WITH PRIOR SOFOSBUVIR TREATMENT EXPERIENCE. , (Journal of Hepatology 2014 vol. 60 | S1–S22). Available at:

http://www.journal-of-hepatology.eu/ [Accessed April 10, 2014].

Gao, M. et al., 2010. Chemical genetics strategy identifies an HCV NS5A inhibitor with a potent clinical effect. *Nature*, 465(7294), pp.96–100.

Gastaminza, P. et al., 2008. Cellular determinants of hepatitis C virus assembly, maturation, degradation, and secretion. *Journal of virology*, 82(5), pp.2120–2129.

Gattoni, A. et al., 2006. Interferon-gamma: biologic functions and HCV therapy (type I/II) (1 of 2 parts). *La Clinica terapeutica*, 157(4), pp.377–386.

Gaudieri, S. et al., 2006. Evidence of viral adaptation to HLA class I-restricted immune pressure in chronic hepatitis C virus infection. *Journal of Virology*, 80(22), pp.11094–11104.

Gaudieri, S. et al., 2009. Hepatitis C virus drug resistance and immune-driven adaptations: relevance to new antiviral therapy. *Hepatology (Baltimore, Md.)*, 49(4), pp.1069–1082.

Ge, D. et al., 2009. Genetic variation in IL28B predicts hepatitis C treatment-induced viral clearance. *Nature*, 461(7262), pp.399–401.

Gerlach, T. et al., 2003. Acute hepatitis C: high rate of both spontaneous and treatment-induced viral clearance. *Gastroenterology*, 125(1), pp.80–88.

Gerlach, T. et al., 1999. Recurrence of hepatitis C virus after loss of virus-specific CD4(+) T-cell response in acute hepatitis C. *Gastroenterology*, 117(4), pp.933–941.

Germer, J.J. et al., 1999. Determination of Hepatitis C Virus Genotype by Direct Sequence Analysis of Products Generated with the Amplicor HCV Test. *Journal of Clinical Microbiology*, 37(8), pp.2625–2630.

Gherardi, M.M. et al., 2003. Prime-boost immunization schedules based on influenza virus and vaccinia virus vectors potentiate cellular immune responses against human immunodeficiency virus Env protein systemically and in the genitorectal draining lymph nodes. *Journal of Virology*, 77(12), pp.7048–7057.

Ghiringhelli, F. et al., 2006. The role of regulatory T cells in the control of natural killer cells: relevance during tumor progression. *Immunological reviews*, 214, pp.229–238.

Gilbert, S.C. et al., 1997. A protein particle vaccine containing multiple malaria epitopes. *Nature Biotechnology*, 15(12), pp.1280–1284.

Gilead press releases, 2014a. European Commission Grants Marketing Authorization for Gilead's Sovaldi® (Sofosbuvir) for the Treatment of Chronic Hepatitis C Infection in Genotypes 1-6. Available at: http://investors.gilead.com/phoenix.zhtml?c=69964&p=irol-newsArticle&ID=1891746 [Accessed April 1, 2014].

Gilead press releases, 2014b. European Medicines Agency Validates Gilead's Marketing Application for Ledipasvir/Sofosbuvir Fixed-Dose Combination Tablet for Genotype 1 Chronic Hepatitis C Infection | Gilead. Available at: http://www.gilead.com/news/press-releases/2014/3/european-medicines-agency-validates-gileads-marketing-application-for-ledipasvirsofosbuvir-fixeddose-combination-tablet-for-genotype-1-chronic-hepatitis-c-infection [Accessed April 1, 2014].

Gilead press releases, 2013. Gilead Announces SVR12 Rates From Three Phase 3 Studies Evaluating a Once-Daily Fixed-Dose Combination of Sofosbuvir and Ledipasvir for Genotype 1 Hepatitis C Patients | Gilead. Available at: http://www.gilead.com/news/press-releases/2013/12/gilead-announces-svr12-rates-from-three-phase-3-studies-evaluating-a-oncedaily-fixeddose-combination-of-sofosbuvir-and-ledipasvir-for-genotype-1-hepatitis-c-patients?mode=print [Accessed April 1, 2014].

Giugliano, S. et al., 2009. Degree of cross-genotype reactivity of hepatitis C virus-specific CD8+ T cells directed against NS3. *Hepatology (Baltimore, Md.)*, 50(3), pp.707–716.

Godkin, A. et al., 2001. Characterization of novel HLA-DR11-restricted HCV epitopes reveals both qualitative and quantitative differences in HCV-specific CD4+ T cell responses in chronically infected and non-viremic patients. *European journal of immunology*, 31(5), pp.1438–1446.

Golden, A. et al., 1998. Effect of promoters and signal sequences on the production of secreted HIV-1 gp120 protein in the baculovirus system. *Protein expression and purification*, 14(1), pp.8–12.

Golden-Mason, L. et al., 2007. Upregulation of PD-1 expression on circulating and intrahepatic hepatitis C virus-specific CD8+ T cells associated with reversible immune dysfunction. *Journal of Virology*, 81(17), pp.9249–9258.

Golden-Mason, L. & Rosen, H.R., 2006. Natural killer cells: primary target for hepatitis C virus immune evasion strategies? *Liver transplantation: official publication of the American Association for the Study of Liver Diseases and the International Liver Transplantation Society*, 12(3), pp.363–372.

Goossens, N. & Negro, F., 2013. Is the genotype 3 of the hepatitis C virus the new villain? *Hepatology*, p.n/a–n/a.

Grakoui, A. et al., 2003. HCV Persistence and Immune Evasion in the Absence of Memory T Cell Help. *Science*, 302(5645), pp.659–662.

Griffin, T.A. et al., 1998. Immunoproteasome assembly: cooperative incorporation of interferon gamma (IFN-gamma)-inducible subunits. *The Journal of experimental medicine*, 187(1), pp.97–104.

Grüner, N.H. et al., 2000. Association of hepatitis C virus-specific CD8+ T cells with viral clearance in acute hepatitis C. *The Journal of Infectious Diseases*, 181(5), pp.1528–1536.

Guan, P. et al., 2006. MHCPred 2.0: an updated quantitative T-cell epitope prediction server. *Applied bioinformatics*, 5(1), pp.55–61.

Guedj, J. et al., 2013. Modeling shows that the NS5A inhibitor daclatasvir has two modes of action and yields a shorter estimate of the hepatitis C virus half-life. *Proceedings of the National Academy of Sciences*, 110(10), pp.3991–3996.

Guidotti, L.G. et al., 1999. Noncytopathic Clearance of Lymphocytic Choriomeningitis Virus from the Hepatocyte. *The Journal of Experimental Medicine*, 189(10), pp.1555–1564.

Guidotti, L.G. & Iannacone, M., 2013. Effector CD8 T cell trafficking within the liver. *Molecular immunology*, 55(1), pp.94–99.

Gulukota, K. et al., 1997. Two complementary methods for predicting peptides binding major histocompatibility complex molecules. *Journal of*

molecular biology, 267(5), pp.1258–1267.

Habib, M. et al., 2001. Hepatitis C virus infection in a community in the Nile Delta: Risk factors for seropositivity. *Hepatology*, 33(1), pp.248–253.

Von Hahn, T. et al., 2007. Hepatitis C virus continuously escapes from neutralizing antibody and T-cell responses during chronic infection in vivo. *Gastroenterology*, 132(2), pp.667–678.

Hajarizadeh, B., Grebely, J. & Dore, Gregory J, 2013. Epidemiology and natural history of HCV infection. *Nature reviews. Gastroenterology & hepatology*, 10(9), pp.553–562.

Hakenberg, J. et al., 2003. MAPPP: MHC class I antigenic peptide processing prediction. *Applied bioinformatics*, 2(3), pp.155–158.

Haller, O., Staeheli, P. & Kochs, G., 2007. Interferon-induced Mx proteins in antiviral host defense. *Biochimie*, 89(6-7), pp.812–818.

Harari, A. et al., 2009. Distinct Profiles of Cytotoxic Granules in Memory CD8 T Cells Correlate with Function, Differentiation Stage, and Antigen Exposure. *Journal of Virology*, 83(7), pp.2862–2871.

Harcourt, G et al., 2006. Diminished frequency of hepatitis C virus specific interferon gamma secreting CD4+ T cells in human immunodeficiency virus/hepatitis C virus coinfected patients. *Gut*, 55(10), pp.1484–1487.

Harris, R.J. et al., 2012. Hepatitis C prevalence in England remains low and varies by ethnicity: an updated evidence synthesis. *European journal of public health*, 22(2), pp.187–192.

Harrison, S.A. et al., 2010. Serum cholesterol and statin use predict virological response to peginterferon and ribavirin therapy. *Hepatology*, 52(3), pp.864–874.

Health Protection Agency, 2013. Hepatitis C in the UK: 2013 report. Available at: http://www.hpa.org.uk/Publications/InfectiousDiseases/BloodBorneInfections/HepatitisCInTheUK/1307HepatitisCintheUK2013report/ [Accessed March 10, 2014].

Health Protection Agency, 2012. Hepatitis C in the UK: annual report 2012. *London: Health Protection Agency Centre for Infections*, pp.1–79.

Heeg, M.H.J. et al., 2009. FOXP3 expression in hepatitis C virus-specific CD4+ T cells during acute hepatitis C. *Gastroenterology*, 137(4), pp.1280–1288.e1–6.

Helle, F. et al., 2007. The neutralizing activity of anti-hepatitis C virus antibodies is modulated by specific glycans on the E2 envelope protein. *Journal of virology*, 81(15), pp.8101–8111.

Hezode, C. et al., 2009. Telaprevir and Peginterferon with or without Ribavirin for Chronic HCV Infection. *N Engl J Med*, 360(18), pp.1839–1850.

Hissar, S.S. et al., 2006. Hepatitis C virus genotype 3 predominates in North and Central India and is associated with significant histopathologic liver disease. *Journal of Medical Virology*, 78(4), pp.452–458.

Hofer, H. et al., 2003. Spontaneous viral clearance in patients with acute hepatitis C can be predicted by repeated measurements of serum viral load. *Hepatology*, 37(1), pp.60–64.

Hoffman, B. & Liu, Q., 2011. Hepatitis C viral protein translation: mechanisms and implications in developing antivirals. *Liver international: official journal of the International Association for the Study of the Liver*, 31(10), pp.1449–1467.

Höhne, M., Schreier, E. & Roggendorf, M, 1994. Sequence variability in the env-coding region of hepatitis C virus isolated from patients infected during a single source outbreak. *Archives of virology*, 137(1-2), pp.25–34.

Honda, K. et al., 2005. IRF-7 is the master regulator of type-I interferon-dependent immune responses. *Nature*, 434(7034), pp.772–777.

Honda, M. et al., 2010. Hepatic ISG Expression Is Associated With Genetic Variation in Interleukin 28B and the Outcome of IFN Therapy for Chronic Hepatitis C. *Gastroenterology*, 139(2), pp.499–509.

Hong, X. et al., 2005. Human leukocyte antigen class II DQB1*0301, DRB1*1101 alleles and spontaneous clearance of hepatitis C virus infection: a meta-analysis. *World journal of gastroenterology: WJG*, 11(46), pp.7302–7307.

Hopkins, S. et al., 2010. SCY-635, a Novel Nonimmunosuppressive Analog of Cyclosporine That Exhibits Potent Inhibition of Hepatitis C Virus RNA Replication In Vitro. *Antimicrobial Agents and Chemotherapy*, 54(2), pp.660–672.

Houghton, M. & Abrignani, S., 2005. Prospects for a vaccine against the hepatitis C virus. *Nature*, 436(7053), pp.961–966.

Huang, H. et al., 2007. Hepatitis C virus production by human hepatocytes dependent on assembly and secretion of very low-density lipoproteins. *Proceedings of the National Academy of Sciences of the United States of America*, 104(14), pp.5848–5853.

Huang, M. et al., 2011. 1212 ACH-2928: A NOVEL HIGHLY POTENT HCV NS5A INHIBITOR WITH FAVORABLE PRECLINICAL CHARACTERISTICS. *Journal of Hepatology*, 54, Supplement 1, p.S479.

Huber, J.P. & Farrar, J.D., 2011. Regulation of effector and memory T-cell functions by type I interferon. *Immunology*, 132(4), pp.466–474.

Hui, J.M. et al., 2002. Genotype-specific mechanisms for hepatic steatosis in chronic hepatitis C infection. *Journal of Gastroenterology and Hepatology*, 17(8), pp.873–881.

Hülsmeyer, M. et al., 2002. HLA-B27 Subtypes Differentially Associated with Disease Exhibit Subtle Structural Alterations. *Journal of Biological Chemistry*, 277(49), pp.47844–47853.

Hultgren, C. et al., 2004. Evidence for a relation between the viral load and genotype and hepatitis C virus-specific T cell responses. *Journal of Hepatology*, 40(6), pp.971–978.

Humphreys, I.S. et al., 2009. Full-length characterization of hepatitis C virus subtype 3a reveals novel hypervariable regions under positive selection during acute infection. *Journal of Virology*, 83(22), pp.11456–11466.

Humphreys, I.S. et al., 2012. HCV genotype-3a T cell immunity: specificity, function and impact of therapy. *Gut*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22337948 [Accessed May 17, 2012].

Hurvitz, S.A. & Timmerman, J.M., 2005. Current status of therapeutic vaccines for non-Hodgkin's lymphoma. *Current opinion in oncology*, 17(5), pp.432–440.

Im, E.-J. et al., 2011. Protective efficacy of serially up-ranked subdominant CD8+ T cell epitopes against virus challenges. *PLoS pathogens*, 7(5), p.e1002041.

Indolfi, G. et al., 2006. Alanine transaminase levels in the year before pregnancy predict the risk of hepatitis C virus vertical transmission. *Journal of medical virology*, 78(7), pp.911–914.

Inoue, G. et al., 2000. Hepatitis C virus clearance is prominent in women in an endemic area. *Journal of gastroenterology and hepatology*, 15(9), pp.1054–1058.

Jackel-Cram, C. et al., 2010. Hepatitis C Virus Genotype-3a Core Protein Enhances Sterol Regulatory Element Binding Protein-1 Activity through the PI3K-Akt-2 Pathway. *The Journal of General Virology*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20130133 [Accessed February 5, 2010].

Jacobson, I.M. et al., 2013. Sofosbuvir for Hepatitis C Genotype 2 or 3 in Patients without Treatment Options. *New England Journal of Medicine*, 368(20), pp.1867–1877.

Jacobson, I.M. et al., 2011. Telaprevir for Previously Untreated Chronic Hepatitis C Virus Infection. *New England Journal of Medicine*, 364(25), pp.2405–2416.

Janeway, C.A. et al., 2001. Immunobiology. Available at: http://www.ncbi.nlm.nih.gov/books/NBK10757/ [Accessed March 29, 2014].

Janssen, H.L.A. et al., 2013. Treatment of HCV Infection by Targeting MicroRNA. *New England Journal of Medicine*, 368(18), pp.1685–1694.

Jardetzky, T.S. et al., 1994. Three-dimensional structure of a human class II histocompatibility molecule complexed with superantigen. *Nature*, 368(6473), pp.711–718.

Jia, Z. et al., 2012. Test of IL28B Polymorphisms in Chronic Hepatitis C Patients Treated with PegIFN and Ribavirin Depends on HCV Genotypes: Results from a Meta-Analysis. *PLoS ONE*, 7(9), p.e45698.

Jiang, Z. et al., 2013. Hepatic stellate cells promote immunotolerance following orthotopic liver transplantation in rats via induction of T cell apoptosis and regulation of Th2/Th3-like cell cytokine production. *Experimental and Therapeutic Medicine*, 5(1), pp.165–169.

Jin, D.-Y., 2007. Molecular pathogenesis of hepatitis C virus-associated hepatocellular carcinoma. *Frontiers in Bioscience: A Journal and Virtual Library*, 12, pp.222–233.

Jinushi, M. et al., 2004. Negative regulation of NK cell activities by inhibitory receptor CD94/NKG2A leads to altered NK cell-induced modulation of dendritic cell functions in chronic hepatitis C virus infection. *Journal of immunology (Baltimore, Md.: 1950)*, 173(10), pp.6072–6081.

Jo, J. et al., 2009. Analysis of CD8+ T-cell-mediated inhibition of hepatitis C virus replication using a novel immunological model. *Gastroenterology*, 136(4), pp.1391–1401.

Jo, J. et al., 2012. Low perforin expression of early differentiated HCV-specific CD8+ T cells limits their hepatotoxic potential. *Journal of hepatology*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22425625 [Accessed June 18, 2012].

Jojic, N. et al., 2006. Learning MHC I--peptide binding. *Bioinformatics (Oxford, England)*, 22(14), pp.e227–235.

Jouan, L. et al., 2010. Distinct antiviral signaling pathways in primary human hepatocytes and their differential disruption by HCV NS3 protease. *Journal of Hepatology*, 52(2), pp.167–175.

Kamal, S.M. et al., 2002. Peginterferon alone or with ribavirin enhances HCV-specific CD4+ T-helper 1 responses in patients with chronic hepatitis C. *Gastroenterology*, 123(4), pp.1070–1083.

Kang, Y.-H. et al., 2012. CD161+CD4+ T cells are enriched in the liver during chronic hepatitis and associated with co-secretion of IL-22 and IFN-? *Frontiers in Immunology*, 3. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3502006/ [Accessed March 13, 2014].

Kaplan, D.E. et al., 2007. Discordant Role of CD4 T-Cell Response Relative to Neutralizing Antibody and CD8 T-Cell Responses in Acute Hepatitis C. *Gastroenterology*, 132(2), pp.654–666.

Kapoor, A. et al., 2011. Characterization of a canine homolog of hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, 108(28), pp.11608–11613.

Kapoor, A. et al., 2013. Identification of rodent homologs of hepatitis C virus and pegiviruses. *mBio*, 4(2), pp.e00216–00213.

Kasprowicz, V. et al., 2008. High Level of PD-1 Expression on Hepatitis C Virus (HCV)-Specific CD8+ and CD4+ T Cells during Acute HCV Infection, Irrespective of Clinical Outcome. *Journal of Virology*, 82(6), pp.3154–3160.

Kau, A., Vermehren, J. & Sarrazin, C., 2008. Treatment predictors of a sustained virologic response in hepatitis B and C. *Journal of hepatology*, 49(4), pp.634–651.

Kaufman, L. & Rousseeuw, P.J., 1990. Finding groups in data: an introduction to cluster analysis. Available at: http://www.bibsonomy.org/bibtex/25873cc08151ea18ae8cf9569f4f9bf34/weeeee [Accessed March 4, 2014].

Keating, S et al., 2005. Hepatitis C viral clearance in an intravenous drug-using cohort in the Dublin area. *Irish journal of medical science*, 174(1), pp.37–41.

Kennedy, P.T.F. et al., 2006. The influence of T cell cross-reactivity on HCV-peptide specific human T cell response. *Hepatology (Baltimore, Md.)*, 43(3), pp.602–611.

Kenny-Walsh, E., 1999. Clinical outcomes after hepatitis C infection from contaminated anti-D immune globulin. Irish Hepatology Research Group. *The New England journal of medicine*, 340(16), pp.1228–1233.

Khakoo, S.I. et al., 2004. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science (New York, N.Y.)*, 305(5685), pp.872–874.

Khan, A.M. et al., 2006. A systematic bioinformatics approach for selection of epitope-based vaccine targets. *Cellular immunology*, 244(2), pp.141–147.

Khan, A.M. et al., 2008. Conservation and variability of dengue virus proteins: implications for vaccine design. *PLoS neglected tropical diseases*,

2(8), p.e272.

Kim, A.Y. et al., 2010. Spontaneous control of HCV is associated with the expression of HLA-B*57 and preservation of targeted epitopes. *Gastroenterology*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20875418 [Accessed November 8, 2010].

Kim, Y. et al., 2012. A Meta-Analysis of the Existing Knowledge of Immunoreactivity against Hepatitis C Virus (HCV). *PLoS ONE*, 7(5), p.e38028.

Knolle, P.A. & Gerken, G, 2000. Local control of the immune response in the liver. *Immunological reviews*, 174, pp.21–34.

Knolle, P.A. & Thimme, R., 2014. Hepatic Immune Regulation and Its Involvement in Viral Hepatitis Infection. *Gastroenterology*, 146(5), pp.1193–1207.

Kolumam, G.A. et al., 2005. Type I interferons act directly on CD8 T cells to allow clonal expansion and memory formation in response to viral infection. *The Journal of Experimental Medicine*, 202(5), pp.637–650.

Koopman, G. et al., 2013. DNA/long peptide vaccination against conserved regions of SIV induces partial protection against SIVmac251 challenge. *AIDS (London, England)*.

Korber, B.T., Letvin, N.L. & Haynes, B.F., 2009. T-Cell Vaccine Strategies for Human Immunodeficiency Virus, the Virus with a Thousand Faces. *Journal of Virology*, 83(17), pp.8300–8314.

Kosmrlj, A. et al., 2010. Effects of thymic selection of the T-cell repertoire on HLA class I-associated control of HIV infection. *Nature*, 465(7296), pp.350–354.

Kotenko, S.V. et al., 2003. IFN-[lambda]s mediate antiviral protection through a distinct class II cytokine receptor complex. *Nat Immunol*, 4(1), pp.69–77.

Kulkarni, V. et al., 2014. Altered Response Hierarchy and Increased T-Cell Breadth upon HIV-1 Conserved Element DNA Vaccination in Macaques. *PLoS ONE*, 9(1), p.e86254.

Kuniholm, M.H. et al., 2010. Specific human leukocyte antigen class I and II alleles associated with hepatitis C virus viremia. *Hepatology*, 51(5), pp.1514–1522.

Kuntzen, T. et al., 2008. Naturally occurring dominant resistance mutations to HCV protease and polymerase inhibitors in treatment-naïve patients. *Hepatology (Baltimore, Md.)*, 48(6), pp.1769–1778.

Kuntzen, T. et al., 2007. Viral Sequence Evolution in Acute Hepatitis C Virus Infection. *J. Virol.*, 81(21), pp.11658–11668.

Kurosaki, M et al., 1994. Evolution and selection of hepatitis C virus variants in patients with chronic hepatitis C. *Virology*, 205(1), pp.161–169.

Kutscher, S. et al., 2013. Overnight resting of PBMC changes functional signatures of antigen specific T- cell responses: impact for immune monitoring within clinical trials. *PloS one*, 8(10), p.e76215.

Lagging, M. et al., 2008. Randomized comparison of 12 or 24 weeks of peginterferon α-2a and ribavirin in chronic hepatitis C virus genotype 2/3 infection. *Hepatology*, 47(6), pp.1837–1845.

Lamonaca, V. et al., 1999. Conserved hepatitis C virus sequences are highly immunogenic for CD4(+) T cells: implications for vaccine development. *Hepatology (Baltimore, Md.)*, 30(4), pp.1088–1098.

Lanford, R.E. et al., 2010. Therapeutic Silencing of MicroRNA-122 in Primates with Chronic Hepatitis C Virus Infection. *Science*, 327(5962), pp.198–201.

Langhans, B. et al., 2010. Core-specific adaptive regulatory T-cells in different outcomes of hepatitis C. *Clinical science (London, England: 1979)*, 119(2), pp.97–109.

Larghi, A. et al., 2002. Outcome of an outbreak of acute hepatitis C among healthy volunteers participating in pharmacokinetics studies. *Hepatology (Baltimore, Md.)*, 36(4 Pt 1), pp.993–1000.

Larrey, D. et al., 2012. Rapid and strong antiviral activity of the non-nucleosidic NS5B polymerase inhibitor BI 207127 in combination with peginterferon alfa 2a and ribavirin. *Journal of hepatology*, 57(1), pp.39–46.

Lasarte, J J et al., 1998. Cellular immunity to hepatitis C virus core protein and the response to interferon in patients with chronic hepatitis C. *Hepatology (Baltimore, Md.)*, 28(3), pp.815–822.

Laskus, T. et al., 2004. Analysis of hepatitis C virus quasispecies transmission and evolution in patients infected through blood transfusion. *Gastroenterology*, 127(3), pp.764–776.

Lauer, G.M. et al., 2002. Comprehensive Analysis of CD8+-T-Cell Responses against Hepatitis C Virus Reveals Multiple Unpredicted Specificities. *J. Virol.*, 76(12), pp.6104–6113.

Lauer, G.M. et al., 2005. Full-Breadth Analysis of CD8+ T-Cell Responses in Acute Hepatitis C Virus Infection and Early Therapy. *Journal of Virology*, 79(20), pp.12979–12988.

Lauer, G.M. et al., 2004. High resolution analysis of cellular immune responses in resolved and persistent hepatitis C virus infection. *Gastroenterology*, 127(3), pp.924–936.

Lauer, G.M. & Walker, B D, 2001. Hepatitis C Virus Infection. *New England Journal of Medicine*, 345(1), pp.41–52.

Law, L.M.J. et al., 2013. Progress towards a hepatitis C virus vaccine. *Emerging Microbes & Infections*, 2(11), p.e79.

Lawitz, E. et al., 2013. Sofosbuvir for Previously Untreated Chronic Hepatitis C Infection. *New England Journal of Medicine*, 368(20), pp.1878–1887.

Lechner, F. et al., 2000. Analysis of successful immune responses in persons infected with hepatitis C virus. *The Journal of Experimental Medicine*, 191(9), pp.1499–1512.

Lehmann, M. et al., 2004. High rate of spontaneous clearance of acute hepatitis C virus genotype 3 infection. *Journal of Medical Virology*, 73(3), pp.387–391.

Lemay, K.L. et al., 2013. A hepatitis C virus NS5A phosphorylation site that regulates RNA replication. *Journal of virology*, 87(2), pp.1255–1260.

Lenschow, D.J. et al., 2005. Identification of Interferon-Stimulated Gene 15 as an Antiviral Molecule during Sindbis Virus Infection In Vivo. *Journal of Virology*, 79(22), pp.13974–13983.

Leroux-Roels, G. et al., 2004. A candidate vaccine based on the hepatitis C E1 protein: tolerability and immunogenicity in healthy volunteers. *Vaccine*, 22(23-24), pp.3080–3086.

Létourneau, S. et al., 2007. Design and pre-clinical evaluation of a universal HIV-1 vaccine. *PloS one*, 2(10), p.e984.

Li, K. et al., 2005. Immune evasion by hepatitis C virus NS3/4A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF. *Proceedings of the National Academy of Sciences of the United States of America*, 102(8), pp.2992–2997.

Li, X.-D. et al., 2005. Hepatitis C virus protease NS3/4A cleaves mitochondrial antiviral signaling protein off the mitochondria to evade innate immunity. *Proceedings of the National Academy of Sciences of the United States of America*, 102(49), pp.17717–17722.

Lim, Y.-S. & Hwang, S.B., 2011. Hepatitis C virus NS5A protein interacts with phosphatidylinositol 4-kinase type IIIalpha and regulates viral propagation. *The Journal of biological chemistry*, 286(13), pp.11290–11298.

Lin, C.-H. & Hünig, T., 2003. Efficient expansion of regulatory T cells in vitro and in vivo with a CD28 superagonist. *European journal of immunology*, 33(3), pp.626–638.

Lin, H.H. et al., 2008. Evaluation of MHC class I peptide binding prediction servers: applications for vaccine research. *BMC immunology*, 9, p.8.

Lin, W. et al., 2006. Hepatitis C virus core protein blocks interferon signaling by interaction with the STAT1 SH2 domain. *Journal of virology*, 80(18), pp.9226–9235.

Lindh, M. et al., 2011. Interleukin 28B Gene Variation at rs12979860 Determines Early Viral Kinetics During Treatment in Patients Carrying Genotypes 2 or 3 of Hepatitis C Virus. *Journal of Infectious Diseases*, 203(12), pp.1748–1752.

Link, J.O. et al., 2014. Discovery of Ledipasvir (GS-5885): A Potent, Once-Daily Oral NS5A Inhibitor for the Treatment of Hepatitis C Virus Infection. *Journal of Medicinal Chemistry*, 57(5), pp.2033–2046.

Liu, Lin et al., 2010. Acceleration of Hepatitis C Virus Envelope Evolution in Humans Is Consistent with Progressive Humoral Immune Selection during the Transition from Acute to Chronic Infection. *Journal of Virology*, 84(10), pp.5067–5077.

Liu, Z.X. et al., 2000. NK cells cause liver injury and facilitate the induction of T cell-mediated immunity to a viral liver infection. *Journal of immunology (Baltimore, Md.: 1950)*, 164(12), pp.6480–6486.

Logvinoff, C. et al., 2004. Neutralizing antibody response during acute and chronic hepatitis C virus infection. *Proceedings of the National Academy of Sciences of the United States of America*, 101(27), pp.10149–10154.

Löhr, H.F. et al., 1996. Liver-infiltrating and circulating CD4+ T cells in chronic hepatitis C: immunodominant epitopes, HLA-restriction and functional significance. *Liver*, 16(3), pp.174–182.

Löhr, H.F. et al., 1999. The viral clearance in interferon-treated chronic hepatitis C is associated with increased cytotoxic T cell frequencies. *Journal of hepatology*, 31(3), pp.407–415.

Lok, A.S. et al., 2012. Preliminary Study of Two Antiviral Agents for Hepatitis C Genotype 1. *New England Journal of Medicine*, 366(3), pp.216–224.

Lorenz, I.C. et al., 2006. Structure of the catalytic domain of the hepatitis C virus NS2-3 protease. *Nature*, 442(7104), pp.831–835.

Lotteau, V. et al., 1990. Intracellular transport of class II MHC molecules directed by invariant chain. *Nature*, 348(6302), pp.600–605.

Lu, J. et al., 2004. Multiepitope Trojan antigen peptide vaccines for the induction of antitumor CTL and Th immune responses. *Journal of immunology (Baltimore, Md.: 1950)*, 172(7), pp.4575–4582.

Lucas, M. et al., 2007. Tracking virus-specific CD4+ T cells during and after acute hepatitis C virus infection. *PloS One*, 2(7), p.e649.

Lundegaard, C. et al., 2010. State of the art and challenges in sequence based T-cell epitope prediction. *Immunome Research*, 6(Suppl 2), p.S3.

MacDonald, A.J. et al., 2002. CD4 T Helper Type 1 and Regulatory T Cells Induced against the Same Epitopes on the Core Protein in Hepatitis C Virus-Infected Persons. *Journal of Infectious Diseases*, 185(6), pp.720–727.

Madden, D.R. et al., 1992. The three-dimensional structure of HLA-B27 at 2.1 A resolution suggests a general mechanism for tight peptide binding to MHC. *Cell*, 70(6), pp.1035–1048.

Mahnke, K. & Enk, A.H., 2005. Dendritic cells: key cells for the induction of regulatory T cells? *Current topics in microbiology and immunology*, 293, pp.133–150.

Mangia, A., Thompson, A.J., et al., 2010. An IL28B polymorphism determines treatment response of hepatitis C virus genotype 2 or 3 patients who do not achieve a rapid virologic response. *Gastroenterology*, 139(3), pp.821–827, 827.e1.

Mangia, A., Bandiera, F., et al., 2010. Individualized treatment with combination of Peg-interferon alpha 2b and ribavirin in patients infected with HCV genotype 3. *Journal of Hepatology*, 53(6), pp.1000–1005.

Mangia, A., Mottola, L. & Piazzolla, V., 2013. Update on the Treatment of Patients With Non–Genotype 1 Hepatitis C Virus Infection. *Clinical Infectious Diseases*, 56(9), pp.1294–1300.

Manigold, T. & Racanelli, V., 2007. T-cell regulation by CD4 regulatory T cells during hepatitis B and C virus infections: facts and controversies. *The Lancet infectious diseases*, 7(12), pp.804–813.

Manns, M.P. et al., 2001. Peginterferon alfa-2b plus ribavirin compared with interferon alfa-2b plus ribavirin for initial treatment of chronic hepatitis C: a randomised trial. *Lancet*, 358(9286), pp.958–965.

Mao, Q. et al., 2001. Human Immunodeficiency Virus Seroconversion and Evolution of the Hepatitis C Virus Quasispecies. *Journal of Virology*, 75(7), pp.3259–3267.

Marcello, T. et al., 2006. Interferons alpha and lambda inhibit hepatitis C virus replication with distinct signal transduction and gene regulation

kinetics. *Gastroenterology*, 131(6), pp.1887–1898.

Martell, M. et al., 1992. Hepatitis C virus (HCV) circulates as a population of different but closely related genomes: quasispecies nature of HCV genome distribution. *Journal of virology*, 66(5), pp.3225–3229.

Martens, S. & Howard, J., 2006. The Interferon-Inducible GTPases. *Annual Review of Cell and Developmental Biology*, 22(1), pp.559–589.

Matera, G. et al., 2002. Changes in the prevalence of hepatitis C virus (HCV) genotype 4 in Calabria, Southern Italy. *Diagnostic Microbiology and Infectious Disease*, 42(3), pp.169–173.

Matikainen, S et al., 2001. IFN-alpha and IL-18 synergistically enhance IFN-gamma production in human NK cells: differential regulation of Stat4 activation and IFN-gamma gene expression by IFN-alpha and IL-12. *European journal of immunology*, 31(7), pp.2236–2245.

McCarthy, J.J. et al., 2010. Replicated Association Between an IL28B Gene Variant and a Sustained Response to Pegylated Interferon and Ribavirin. *Gastroenterology*, 138(7), pp.2307–2314.

McConkey, S.J. et al., 2003. Enhanced T-cell immunogenicity of plasmid DNA vaccines boosted by recombinant modified vaccinia virus Ankara in humans. *Nature Medicine*, 9(6), pp.729–735.

McFarland, A.P. et al., 2014. The favorable IFNL3 genotype escapes mRNA decay mediated by AU-rich elements and hepatitis C virus-induced microRNAs. *Nature Immunology*, 15(1), pp.72–79.

McHutchison, J.G. et al., 2010. Telaprevir for previously treated chronic HCV infection. *The New England Journal of Medicine*, 362(14), pp.1292–1303.

McKiernan, S. et al., 2004. Distinct MHC class I and II alleles are associated with hepatitis C viral clearance, originating from a single source. *Hepatology*, 40(1), pp.108–114.

McLauchlan, J. et al., 2002. Intramembrane proteolysis promotes trafficking of hepatitis C virus core protein to lipid droplets. *The EMBO journal*, 21(15), pp.3980–3988.

McMahan, R.H. et al., 2010. Tim-3 expression on PD-1+ HCV-specific human CTLs is associated with viral persistence, and its blockade restores hepatocyte-directed in vitro cytotoxicity. *The Journal of clinical investigation*, 120(12), pp.4546–4557.

Mecenate, F. et al., 2010. Short versus standard treatment with pegylated interferon alfa-2A plus ribavirin in patients with hepatitis C virus genotype 2 or 3: the cleo trial. *BMC gastroenterology*, 10, p.21.

Mellor, J. et al., 1996. Survey of type 6 group variants of hepatitis C virus in Southeast Asia by using a core-based genotyping assay. *Journal of clinical microbiology*, 34(2), pp.417–423.

Membreno, F.E. & Lawitz, E., 2011. The HCV NS5B nucleoside and non-nucleoside inhibitors. *Clinics in liver disease*, 15(3), pp.611–626.

Mengshol, J.A. et al., 2010. A Crucial Role for Kupffer Cell-Derived Galectin-9 in Regulation of T Cell Immunity in Hepatitis C Infection. *PLoS ONE*, 5(3), p.e9504.

Messick, K. et al., 2001. Hepatitis C viral clearance and antibody reactivity patterns in persons with haemophilia and other congenital bleeding disorders. *Haemophilia*, 7(6), pp.568–574.

Micallef, J.M., Kaldor, J.M. & Dore, G J, 2006. Spontaneous viral clearance following acute hepatitis C infection: a systematic review of longitudinal studies. *Journal of viral hepatitis*, 13(1), pp.34–41.

Michaux, A. et al., 2014. A spliced antigenic peptide comprising a single spliced amino acid is produced in the proteasome by reverse splicing of a longer peptide fragment followed by trimming. *Journal of immunology (Baltimore, Md.: 1950)*, 192(4), pp.1962–1971.

Minton, K., 2014. HIV: Mechanisms of T cell polyfunctionality. *Nature Reviews Immunology*, 14(1), pp.7–7.

Miretti, M.M. et al., 2005. A High-Resolution Linkage-Disequilibrium Map of the Human Major Histocompatibility Complex and First Generation of Tag Single-Nucleotide Polymorphisms. *American Journal of Human Genetics*, 76(4), pp.634–646.

Mitchell, A.E., Colvin, H.M. & Palmer Beasley, R., 2010. Institute of Medicine recommendations for the prevention and control of hepatitis B and C. *Hepatology*, 51(3), pp.729–733.

Miyanari, Y. et al., 2007. The lipid droplet is an important organelle for hepatitis C virus production. *Nature Cell Biology*, 9(9), pp.1089–1097.

Moghaddam, A. et al., 2011. IL28B genetic variation and treatment response in patients with hepatitis C virus genotype 3 infection. *Hepatology*, 53(3), pp.746–754.

Moghaddam, A., Reinton, N & Dalgard, O., 2006. A rapid real-time PCR assay for determination of hepatitis C virus genotypes 1, 2 and 3a. *Journal of viral hepatitis*, 13(4), pp.222–229.

Mohd Hanafiah, K. et al., 2013. Global epidemiology of hepatitis C virus infection: new estimates of age-specific antibody to HCV seroprevalence. *Hepatology (Baltimore, Md.)*, 57(4), pp.1333–1342.

Montes-Cano, M.A. et al., 2010. Interleukin-28B genetic variants and hepatitis virus infection by different viral genotypes. *Hepatology (Baltimore, Md.)*, 52(1), pp.33–37.

Montoya, M. et al., 2002. Type I interferons produced by dendritic cells promote their phenotypic and functional activation. *Blood*, 99(9), pp.3263–3271.

Moorman, J.P. et al., 2012. Tim-3 pathway controls regulatory and effector T cell balance during hepatitis C virus infection. *Journal of immunology (Baltimore, Md.: 1950)*, 189(2), pp.755–766.

Morrice, N.A. & Powis, S.J., 1998. A role for the thiol-dependent reductase ERp57 in the assembly of MHC class I molecules. *Current biology: CB*, 8(12), pp.713–716.

Mosbruger, T. et al., 2010. Large-Scale Candidate Gene Analysis of Spontaneous Clearance of Hepatitis C Virus. *Journal of Infectious Diseases*, 201(9), pp.1371–1380.

Moschen, A.R. et al., 2008. Interferon-alpha controls IL-17 expression in vitro and in vivo. *Immunobiology*, 213(9-10), pp.779–787.

Mueller, M. et al., 2010. Virus-specific CD4+ T cell responses in chronic HCV infection in blood and liver identified by antigen-specific upregulation of CD154. *Journal of hepatology*, 52(6), pp.800–811.

Muir, A.J. et al., 2010. Phase 1b study of pegylated interferon lambda 1 with or without ribavirin in patients with chronic genotype 1 hepatitis C virus infection. *Hepatology (Baltimore, Md.)*, 52(3), pp.822–832.

Munshaw, S. et al., 2012. Computational reconstruction of Bole1a, a representative synthetic hepatitis C virus subtype 1a genome. *Journal of virology*, 86(10), pp.5915–5921.

Naggie, S. et al., 2012. Dysregulation of innate immunity in hepatitis C virus genotype 1 IL28B-unfavorable genotype patients: Impaired viral kinetics and therapeutic response. *Hepatology*, 56(2), pp.444–454.

Nakamoto, N. et al., 2008. Functional restoration of HCV-specific CD8 T-cells by PD1 blockade is defined by PD1 expression and compartmentalization. *Gastroenterology*, 134(7), pp.1927–1937.e2.

Nakamoto, N. et al., 2009. Synergistic Reversal of Intrahepatic HCV-Specific CD8 T Cell Exhaustion by Combined PD-1/CTLA-4 Blockade. *PLoS Pathogens*, 5(2). Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2642724/ [Accessed March 13, 2014].

Nascimbeni, M. et al., 2003. Kinetics of CD4+ and CD8+ Memory T-Cell Responses during Hepatitis C Virus Rechallenge of Previously Recovered Chimpanzees. *Journal of Virology*, 77(8), pp.4781–4793.

National Academy of Sciences, 2011. IOM Report Recommends Stringent Limits On Use Of Chimpanzees in Biomedical and Behavioral Research. Available at: http://iom.edu/Reports/2011/Chimpanzees-in-Biomedical-and-Behavioral-Research-Assessing-the-Necessity/Press-Release.aspx [Accessed August 21, 2014].

Nattermann, J. et al., 2005. The HLA-A2 restricted T cell epitope HCV core 35-44 stabilizes HLA-E expression and inhibits cytolysis mediated by natural killer cells. *The American journal of pathology*, 166(2), pp.443–453.

Neefjes, J. et al., 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Reviews Immunology*, 11(12), pp.823–836.

Negro, F. & Alberti, A., 2011. The global health burden of hepatitis C virus infection. *Liver international: official journal of the International Association for the Study of the Liver*, 31 Suppl 2, pp.1–3.

Nettles, R.E. et al., 2011. Multiple ascending dose study of BMS-790052, a nonstructural protein 5A replication complex inhibitor, in patients infected with hepatitis C virus genotype 1. *Hepatology*, 54(6), pp.1956–1965.

Neukam, K. et al., 2012. Prediction of response to pegylated interferon plus ribavirin in HIV/hepatitis C virus (HCV)-coinfected patients using HCV genotype, IL28B variations, and HCV-RNA load. *Journal of hepatology*, 56(4), pp.788–794.

Neumann, A.U. et al., 1998. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science (New York, N.Y.)*, 282(5386), pp.103–107.

Neumann-Haefelin, C. et al., 2008. Analysis of the evolutionary forces in an immunodominant CD8 epitope in hepatitis C virus at a population level. *Journal of Virology*, 82(7), pp.3438–3451.

Neumann-Haefelin, C. et al., 2006. Dominant influence of an HLA-B27 restricted CD8+ T cell response in mediating HCV clearance and evolution. *Hepatology (Baltimore, Md.)*, 43(3), pp.563–572.

Neumann-Haefelin, C. et al., 2007. Host and viral factors contributing to CD8+ T cell failure in hepatitis C virus infection. *World Journal of Gastroenterology: WJG*, 13(36), pp.4839–4847.

Neumann-Haefelin, C. et al., 2010. Protective effect of human leukocyte antigen B27 in hepatitis C virus infection requires the presence of a genotype-specific immunodominant CD8+ T-cell epitope. *Hepatology (Baltimore, Md.)*, 51(1), pp.54–62.

Newell, E.W. et al., 2012. Cytometry by time-of-flight shows combinatorial cytokine expression and virus-specific cell niches within a continuum of CD8+ T cell phenotypes. *Immunity*, 36(1), pp.142–152.

NICE Guidelines, 2014. Hepatitis C - peginterferon alfa and ribavirin. *NICE Hepatitis C - peginterferon alfa and ribavirin (TA200)*. Available at: http://www.nice.org.uk/ [Accessed March 31, 2014].

NICE guidlines for Boceprevir and Telaprevir, 2013. NICE publishes final guidance on two new drugs for chronic hepatitis C. *NICE*. Available at: http://www.nice.org.uk/ [Accessed January 16, 2014].

Nielsen, M. et al., 2004. Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach. *Bioinformatics (Oxford, England)*, 20(9), pp.1388–1397.

Nielsen, M. et al., 2008. Quantitative predictions of peptide binding to any HLA-DR molecule of known sequence: NetMHCIIpan. *PLoS computational biology*, 4(7), p.e1000107.

Nielsen, M. et al., 2003. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein science: a publication of the Protein Society*, 12(5), pp.1007–1017.

Nielsen, M., Lundegaard, C. & Lund, O., 2007. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC bioinformatics*, 8, p.238.

Nielsen, S. et al., 2006. Association between hepatitis C virus and very-low-density lipoprotein (VLDL)/LDL analyzed in iodixanol density gradients. *Journal of virology*, 80(5), pp.2418–2428.

Njouom, R. et al., 2003. High rate of hepatitis C virus infection and predominance of genotype 4 among elderly inhabitants of a remote village of the rain forest of South Cameroon. *Journal of Medical Virology*, 71(2), pp.219–225.

Nkontchou, G. et al., 2011. HCV genotype 3 is associated with a higher hepatocellular carcinoma incidence in patients with ongoing viral C cirrhosis. *Journal of Viral Hepatitis*, 18(10), pp.e516–e522.

O'Hara, G.A. et al., 2012. Memory T cell inflation: understanding cause and effect. *Trends in Immunology*, 33(2), pp.84–90.

Ondondo, B. et al., 2013. Absence of systemic toxicity changes following intramuscular administration of novel pSG2.HIVconsv DNA, ChAdV63.HIVconsv and MVA.HIVconsv vaccines to BALB/c mice. *Vaccine*, 31(47), pp.5594–5601.

Oni, A.Ø. & Harrison, T.J., 1996. Genotypes of hepatitis C virus in Nigeria. *Journal of Medical Virology*, 49(3), pp.178–186.

Osburn, W.O. et al., 2009. Spontaneous Control of Primary Hepatitis C Virus Infection and Immunity Against Persistent Reinfection. *Gastroenterology*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19782080 [Accessed January 5, 2010].

Oxenius, A., Zinkernagel, R.M. & Hengartner, H., 1998. Comparison of activation versus induction of unresponsiveness of virus-specific CD4+ and CD8+ T cells upon acute versus persistent viral infection. *Immunity*, 9(4), pp.449–457.

Pajot, A. et al., 2004. A mouse model of human adaptive immune functions: HLA-A2.1-/HLA-DR1-transgenic H-2 class I-/class II-knockout mice. *European Journal of Immunology*, 34(11), pp.3060–3069.

Pantaleo, G. & Harari, A., 2006. Functional signatures in antiviral T-cell immunity for monitoring virus-associated diseases. *Nature Reviews Immunology*, 6(5), pp.417–423.

Park, S.-H. et al., 2012. Successful vaccination induces multifunctional memory T-cell precursors associated with early control of hepatitis C virus. *Gastroenterology*, 143(4), pp.1048–1060.e4.

Parker, K.C., Bednarek, M.A. & Coligan, J.E., 1994. Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *Journal of immunology (Baltimore, Md.: 1950)*, 152(1), pp.163–175.

Pawlotsky, J.-M. et al., 2012. 1405 ALISPORIVIR PLUS RIBAVIRIN IS HIGHLY EFFECTIVE AS INTERFERON-FREE OR INTERFERON-ADD-ON REGIMEN IN PREVIOUSLY UNTREATED HCV-GT2 OR GT3 PATIENTS: SVR12 RESULTS FROM VITAL-1 PHASE 2B STUDY. *Journal of Hepatology*, 56, p.S553.

Pawlotsky, J.-M., 2014. New Hepatitis C Therapies: The Toolbox, Strategies, And Challenges. *Gastroenterology*.

Payne, R.P. et al., 2010. Efficacious early antiviral activity of HIV Gag- and Pol-specific HLA-B 2705-restricted CD8+ T cells. *Journal of virology*, 84(20), pp.10543–10557.

Pazienza, V. et al., 2010. Hepatitis C virus core protein genotype 3a increases SOCS-7 expression through PPAR-{gamma} in Huh-7 cells. *The Journal of General Virology*, 91(Pt 7), pp.1678–1686.

Peh, C.A. et al., 2000. Distinct functions of tapasin revealed by polymorphism in MHC class I peptide loading. *Journal of immunology (Baltimore, Md.: 1950)*, 164(1), pp.292–299.

Penna, A. et al., 2007. Dysfunction and functional restoration of HCV-specific CD8 responses in chronic hepatitis C virus infection. *Hepatology (Baltimore, Md.)*, 45(3), pp.588–601.

Perlemuter, G. et al., 2002. Hepatitis C virus core protein inhibits microsomal triglyceride transfer protein activity and very low density lipoprotein secretion: a model of viral-related steatosis. *The FASEB Journal*, 16(2), pp.185–194.

Perz, J.F. et al., 2006. The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide. *Journal of hepatology*, 45(4), pp.529–538.

Pestka, J.M. et al., 2007. Rapid induction of virus-neutralizing antibodies and viral clearance in a single-source outbreak of hepatitis C. *Proceedings of the National Academy of Sciences of the United States of America*, 104(14), pp.6025–6030.

Peters, B. et al., 2003. Examining the independent binding assumption for binding of peptide epitopes to MHC-I molecules. *Bioinformatics (Oxford, England)*, 19(14), pp.1765–1772.

Peters, B. & Sette, A., 2005. Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC bioinformatics*, 6, p.132.

Pflugheber, J. et al., 2002. Regulation of PKR and IRF-1 during hepatitis C virus RNA replication. *Proceedings of the National Academy of Sciences of the United States of America*, 99(7), pp.4650–4655.

Piasecki, B.A. et al., 2004. Influence of alcohol use, race, and viral coinfections on spontaneous HCV clearance in a US veteran population. *Hepatology (Baltimore, Md.)*, 40(4), pp.892–899.

Pileri, P et al., 1998. Binding of hepatitis C virus to CD81. *Science (New York, N.Y.)*, 282(5390), pp.938–941.

Pillai, V. et al., 2007. Transient regulatory T-cells: A state attained by all activated human T-cells. *Clinical Immunology*, 123(1), pp.18–29.

Pilli, M. et al., 2007. HCV-specific T-cell response in relation to viral kinetics and treatment outcome (DITTO-HCV project). *Gastroenterology*, 133(4), pp.1132–1143.

Ploss, A. et al., 2009. Human occludin is a hepatitis C virus entry factor required for infection of mouse cells. *Nature*, 457(7231), pp.882–886.

Poli, F. et al., 1998. A retrospective evaluation of HLA-A, B and -DRB1 matching in liver transplantation. *Transplant international: official journal of the European Society for Organ Transplantation*, 11 Suppl 1, pp.S347–349.

Polyak, S J et al., 2001. Hepatitis C virus nonstructural 5A protein induces interleukin-8, leading to partial inhibition of the interferon-induced antiviral response. *Journal of virology*, 75(13), pp.6095–6106.

Poordad, F. et al., 2011. Boceprevir for Untreated Chronic HCV Genotype 1 Infection. *New England Journal of Medicine*, 364(13), pp.1195–1206.

Power, J.P. et al., 1995. Molecular epidemiology of an outbreak of infection with hepatitis C virus in recipients of anti-D immunoglobulin. *Lancet*, 345(8959), pp.1211–1213.

Poynard, T. et al., 2000. Is an “à la carte” combination interferon alfa-2b plus ribavirin regimen possible for the first line treatment in patients with chronic hepatitis C? *Hepatology*, 31(1), pp.211–218.

Precopio, M.L. et al., 2008. Optimizing Peptide Matrices For Identifying T Cell Antigens,. *Cytometry. Part A : the journal of the International Society for Analytical Cytology*, 73(11), pp.1071–1078.

Prentoe, J. et al., 2011. Hypervariable region 1 differentially impacts viability of hepatitis C virus strains of genotypes 1 to 6 and impairs virus

neutralization. *Journal of virology*, 85(5), pp.2224–2234.

Probst, A. et al., 2011. Role of Hepatitis C virus genotype 3 in liver fibrosis progression – a systematic review and meta‐analysis. *Journal of Viral Hepatitis*, 18(11), pp.745–759.

Probst-Kepper, M. et al., 2004. Conformational Restraints and Flexibility of 14-Meric Peptides in Complex with HLA-B*3501. *The Journal of Immunology*, 173(9), pp.5610–5616.

Prokunina-Olsson, L. et al., 2013. A variant upstream of IFNL3 (IL28B) creating a new interferon gene IFNL4 is associated with impaired clearance of hepatitis C virus. *Nature Genetics*, 45(2), pp.164–171.

Protzer, U., Maini, M.K. & Knolle, P.A., 2012. Living in the liver: hepatic infections. *Nature Reviews Immunology*, 12(3), pp.201–213.

Puig-Basagoiti, F. et al., 2001. Influence of the genetic heterogeneity of the ISDR and PePHD regions of hepatitis C virus on the response to interferon therapy in chronic hepatitis C. *Journal of Medical Virology*, 65(1), pp.35–44.

Pybus, O.G. et al., 2009. Genetic history of hepatitis C virus in East Asia. *Journal of virology*, 83(2), pp.1071–1082.

Pybus, O.G. et al., 2003. The Epidemiology and Iatrogenic Transmission of Hepatitis C Virus in Egypt: A Bayesian Coalescent Approach. *Molecular Biology and Evolution*, 20(3), pp.381–387.

Quarato, G. et al., 2012. The cyclophilin inhibitor alisporivir prevents hepatitis C virus-mediated mitochondrial dysfunction. *Hepatology (Baltimore, Md.)*, 55(5), pp.1333–1343.

Qureshi, H. et al., 2011. Identification of immunogenic regions within the alternative reading frame protein of hepatitis C virus (genotype 3). *European journal of clinical microbiology & infectious diseases: official publication of the European Society of Clinical Microbiology*, 30(9), pp.1075–1083.

Qurishi, N. et al., 2003. Effect of antiretroviral therapy on liver-related mortality in patients with HIV and hepatitis C virus coinfection. *The Lancet*, 362(9397), pp.1708–1713.

Radziewicz, H. et al., 2008. Impaired Hepatitis C Virus (HCV)-Specific Effector CD8+ T Cells Undergo Massive Apoptosis in the Peripheral Blood during Acute HCV Infection and in the Liver during the Chronic Phase of Infection. *Journal of Virology*, 82(20), pp.9808 –9822.

Radziewicz, H. et al., 2007. Liver-infiltrating lymphocytes in chronic human hepatitis C virus infection display an exhausted phenotype with high levels of PD-1 and low levels of CD127 expression. *Journal of virology*, 81(6), pp.2545–2553.

Rahman, F. et al., 2004. Effects of antiviral therapy on the cellular immune response in acute hepatitis C. *Hepatology (Baltimore, Md.)*, 40(1), pp.87–97.

Rammensee, H. et al., 1999. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, 50(3-4), pp.213–219.

Rammensee, H., Falk, K. & Rötzschke, O., 1993. MHC molecules as peptide receptors. *Current Opinion in Immunology*, 5(1), pp.35–44.

Ramratnam, B. et al., 1999. Rapid production and clearance of HIV-1 and hepatitis C virus assessed by large volume plasma apheresis. *Lancet*, 354(9192), pp.1782–1785.

Randall, R.E. & Goodbourn, S., 2008. Interferons and viruses: an interplay between induction, signalling, antiviral responses and virus countermeasures. *Journal of General Virology*, 89(1), pp.1–47.

Rauch, A. et al., 2009. Divergent adaptation of hepatitis C virus genotypes 1 and 3 to human leukocyte antigen-restricted immune pressure. *Hepatology (Baltimore, Md.)*, 50(4), pp.1017–1029.

Rauch, A. et al., 2010. Genetic variation in IL28B is associated with chronic hepatitis C and treatment failure: a genome-wide association study. *Gastroenterology*, 138(4), pp.1338–1345, 1345.e1–7.

Ray, S.C. et al., 1999. Acute Hepatitis C Virus Structural Gene Sequences as Predictors of Persistent Viremia: Hypervariable Region 1 as a Decoy. *Journal of Virology*, 73(4), pp.2938–2946.

Ray, S.C. et al., 2005. Divergent and convergent evolution after a common-source outbreak of hepatitis C virus. *The Journal of experimental medicine*, 201(11), pp.1753–1759.

Rebsamen, M. et al., 2008. The antiviral adaptor proteins Cardif and Trif are processed and inactivated by caspases. *Cell Death Differ*, 15(11), pp.1804–1811.

Reche, P.A., Glutting, J.-P. & Reinherz, E.L., 2002. Prediction of MHC class I binding peptides using profile motifs. *Human immunology*, 63(9), pp.701–709.

Rehermann, B., 2009. Hepatitis C virus versus innate and adaptive immune responses: a tale of coevolution and coexistence. *The Journal of clinical investigation*, 119(7), pp.1745–1754.

Reich, D.E. et al., 2001. Linkage disequilibrium in the human genome. *Nature*, 411(6834), pp.199–204.

Reits, E.A. et al., 2000. The major substrates for TAP in vivo are derived from newly synthesized proteins. *Nature*, 404(6779), pp.774–778.

Rembeck, K. et al., 2012. Impact of IL28B-Related Single Nucleotide Polymorphisms on Liver Histopathology in Chronic Hepatitis C Genotype 2 and 3. *PLoS ONE*, 7(1), p.e29370.

Ribeiro, S.P. et al., 2010. A vaccine encoding conserved promiscuous HIV CD4 epitopes induces broad T cell responses in mice transgenic to multiple common HLA class II molecules. *PloS one*, 5(6), p.e11072.

Roche, P.A. & Cresswell, P., 1990. Invariant chain association with HLA-DR molecules inhibits immunogenic peptide binding. *Nature*, 345(6276), pp.615–618.

Rock, K.L., Rothstein, L. & Benacerraf, B., 1992. Analysis of the association of peptides of optimal length to class I molecules on the surface of cells. *Proceedings of the National Academy of Sciences of the United States of America*, 89(19), pp.8918–8922.

Rodrigue-Gervais, I.G. et al., 2010. Dendritic cell inhibition is connected to exhaustion of CD8+ T cell polyfunctionality during chronic hepatitis C virus infection. *Journal of immunology (Baltimore, Md.: 1950)*, 184(6), pp.3134–3144.

Roederer, M., Nozzi, J.L. & Nason, M.C., 2011. SPICE: exploration and analysis of post-cytometric complex multivariate datasets. *Cytometry. Part A: the journal of the International Society for Analytical Cytology*, 79(2), pp.167–174.

Rohrbach, J. et al., 2010. Cellular immune responses to HCV core increase and HCV RNA levels decrease during successful antiretroviral therapy. *Gut*, 59(9), pp.1252–1258.

Rollier, C.S. et al., 2007. Vaccine-induced early control of hepatitis C virus infection in chimpanzees fails to impact on hepatic PD-1 and chronicity. *Hepatology (Baltimore, Md.)*, 45(3), pp.602–613.

Römer, P.S. et al., 2011. Preculture of PBMC at high cell density increases sensitivity of T-cell responses, revealing cytokine release by CD28 superagonist TGN1412. *Blood*. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21931118 [Accessed December 9, 2011].

Romero-Brey, I. et al., 2012. Three-dimensional architecture and biogenesis of membrane structures associated with hepatitis C virus replication. *PLoS pathogens*, 8(12), p.e1003056.

Rosa, D.S. et al., 2011. A DNA vaccine encoding multiple HIV CD4 epitopes elicits vigorous polyfunctional, long-lived CD4+ and CD8+ T cell responses. *PloS one*, 6(2), p.e16921.

Rosario, M. et al., 2012. Prime-boost regimens with adjuvanted synthetic long peptides elicit T cells and antibodies to conserved regions of HIV-1 in macaques. *AIDS (London, England)*, 26(3), pp.275–284.

Rosen, H.R., 2013. Emerging concepts in immunity to hepatitis C virus infection. *The Journal of Clinical Investigation*, 123(10), pp.4121–4130.

Rosen, H.R. et al., 2007. Selective decrease in hepatitis C virus-specific immunity among African Americans and outcome of antiviral therapy. *Hepatology (Baltimore, Md.)*, 46(2), pp.350–358.

Ruhl, M. et al., 2011. CD8+ T-cell response promotes evolution of hepatitis C virus nonstructural proteins. *Gastroenterology*, 140(7), pp.2064–2073.

Ruhl, M. et al., 2012. Escape from a dominant HLA-B*15-restricted CD8+ T cell response against hepatitis C virus requires compensatory mutations outside the epitope. *Journal of virology*, 86(2), pp.991–1000.

Rushbrook, S.M. et al., 2005. Regulatory T cells suppress in vitro proliferation of virus-specific CD8+ T cells during persistent hepatitis C virus infection. *Journal of virology*, 79(12), pp.7852–7859.

Ruys, T.A. et al., 2008. HCV-specific T-cell responses in injecting drug users: evidence for previous exposure to HCV and a role for CD4+ T cells focussing on nonstructural proteins in viral clearance. *Journal of viral hepatitis*, 15(6), pp.409–420.

Sadasivan, B. et al., 1996. Roles for calreticulin and a novel glycoprotein, tapasin, in the interaction of MHC class I molecules with TAP. *Immunity*, 5(2), pp.103–114.

Saito, T. et al., 2008. Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. *Nature*, 454(7203), pp.523–527.

Sánchez-Quijano, A. et al., 1997. Unexpected high prevalence of hepatitis C virus genotype 4 in Southern Spain. *Journal of Hepatology*, 27(1), pp.25–29.

Sarasin-Filipowicz, M. et al., 2009. Alpha interferon induces long-lasting refractoriness of JAK-STAT signaling in the mouse liver through induction of USP18/UBP43. *Molecular and cellular biology*, 29(17), pp.4841–4851.

Sarasin-Filipowicz, M. et al., 2008. Interferon signaling and treatment outcome in chronic hepatitis C. *Proceedings of the National Academy of Sciences of the United States of America*, 105(19), pp.7034–7039.

Sarrazin, C. et al., 2011. Importance of IL28B gene polymorphisms in hepatitis C virus genotype 2 and 3 infected patients. *Journal of Hepatology*, 54(3), pp.415–421.

Sarrazin, C. et al., 2000. Mutations within the E2 and NS5A protein in patients infected with hepatitis C virus type 3a and correlation with treatment response. *Hepatology (Baltimore, Md.)*, 31(6), pp.1360–1370.

Sathiamurthy, M et al., 2003. Population of the HLA ligand database. *Tissue antigens*, 61(1), pp.12–19.

Scheel, T.K.H. & Rice, C.M., 2013. Understanding the hepatitis C virus life cycle paves the way for highly effective therapies. *Nature medicine*, 19(7), pp.837–849.

Schiavoni, G., Mattei, F. & Gabriele, L., 2013. Type I Interferons as Stimulators of DC-Mediated Cross-Priming: Impact on Anti-Tumor Response. *Frontiers in Immunology*, 4. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3872318/ [Accessed March 23, 2014].

Schinazi, R. et al., 2014. HCV direct-acting antiviral agents: the best interferon-free combinations. *Liver international: official journal of the International Association for the Study of the Liver*, 34 Suppl 1, pp.69–78.

Schirren, C.A. et al., 2000. Liver-derived hepatitis C virus (HCV)-specific CD4(+) T cells recognize multiple HCV epitopes and produce interferon gamma. *Hepatology (Baltimore, Md.)*, 32(3), pp.597–603.

Schoenborn, J.R. & Wilson, C.B., 2007. Regulation of Interferon‐γ During Innate and Adaptive Immune Responses. In Frederick W. Alt, ed. *Advances in Immunology*. Academic Press, pp. 41–101. Available at: http://www.sciencedirect.com/science/article/pii/S0065277607960022 [Accessed March 23, 2014].

Schubert, U. et al., 2000. Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*, 404(6779), pp.770–774.

Schulze zur Wiesch, J. et al., 2005. Broad Repertoire of the CD4+ Th Cell Response in Spontaneously Controlled Hepatitis C Virus Infection Includes Dominant and Highly Promiscuous Epitopes. *The Journal of Immunology*, 175(6), pp.3603 –3613.

Schulze zur Wiesch, J. et al., 2007. Immunologic evidence for lack of heterologous protection following resolution of HCV in patients with non–genotype 1 infection. *Blood*, 110(5), pp.1559 –1569.

Schulze Zur Wiesch, J. et al., 2012. Broadly directed virus-specific CD4+ T cell responses are primed during acute hepatitis C infection, but rapidly disappear from human blood with viral persistence. *The Journal of experimental medicine*, 209(1), pp.61–75.

Scott, J.D. et al., 2006. High Rate of Spontaneous Negativity for Hepatitis C Virus RNA after Establishment of Chronic Infection in Alaska Natives. *Clinical Infectious Diseases*, 42(7), pp.945–952.

Seder, R.A., Darrah, P.A. & Roederer, M., 2008. T-cell quality in memory and protection: implications for vaccine design. *Nature reviews. Immunology*, 8(4), pp.247–258.

Seeff, L.B., 2002. Natural history of chronic hepatitis C. *Hepatology (Baltimore, Md.)*, 36(5 Suppl 1), pp.35–46.

Semmo, N. et al., 2005. Preferential loss of IL-2-secreting CD4+ T helper cells in chronic HCV infection. *Hepatology (Baltimore, Md.)*, 41(5), pp.1019–1028.

Serfaty, L. et al., 2001. Hepatitis C virus induced hypobetalipoproteinemia: a possible mechanism for steatosis in chronic hepatitis C. *Journal of Hepatology*, 34(3), pp.428–434.

Sharma, N.R. et al., 2011. Hepatitis C virus is primed by CD81 protein for low pH-dependent fusion. *The Journal of biological chemistry*, 286(35), pp.30361–30376.

Shepard, C.W., Finelli, L. & Alter, M.J., 2005. Global epidemiology of hepatitis C virus infection. *The Lancet Infectious Diseases*, 5(9), pp.558–567.

Sheppard, P. et al., 2003. IL-28, IL-29 and their class II cytokine receptor IL-28R. *Nat Immunol*, 4(1), pp.63–68.

Shevach, E.M. et al., 2006. The lifestyle of naturally occurring CD4+ CD25+ Foxp3+ regulatory T cells. *Immunological reviews*, 212, pp.60–73.

Shi, S.T. et al., 2002. Hepatitis C Virus NS5A Colocalizes with the Core Protein on Lipid Droplets and Interacts with Apolipoproteins. *Virology*, 292(2), pp.198–210.

Shiffman, M.L. et al., 2007. Peginterferon Alfa-2a and Ribavirin for 16 or 24 Weeks in HCV Genotype 2 or 3. *New England Journal of Medicine*, 357(2), pp.124–134.

Shoukry, N.H. et al., 2003. Memory CD8+ T cells are required for protection from persistent hepatitis C virus infection. *The Journal of Experimental Medicine*, 197(12), pp.1645–1655.

Siebenkotten, I.M., Carstens, C. & Koch, N., 1998. Identification of a Sequence That Mediates Promiscuous Binding of Invariant Chain to MHC Class II Allotypes. *The Journal of Immunology*, 160(7), pp.3355–3362.

Silverman, R.H., 2007. Viral encounters with 2',5'-oligoadenylate synthetase and RNase L during the interferon antiviral response. *Journal of virology*, 81(23), pp.12720–12729.

Simmonds, P. et al., 2005. Consensus proposals for a unified system of nomenclature of hepatitis C virus genotypes. *Hepatology (Baltimore, Md.)*, 42(4), pp.962–973.

Simmonds, P. et al., 1996. Evolutionary analysis of variants of hepatitis C virus found in South-East Asia: comparison with classifications based upon sequence similarity. *Journal of General Virology*, 77(12), pp.3013–3024.

Simmonds, P., 2004. Genetic diversity and evolution of hepatitis C virus--15 years on. *The Journal of general virology*, 85(Pt 11), pp.3173–3188.

Simmonds, P., 2013. The origin of hepatitis C virus. *Current topics in microbiology and immunology*, 369, pp.1–15.

Simmons, R. et al., 2013. Evolution of CD8+ T Cell Responses after Acute PARV4 Infection. *Journal of Virology*, 87(6), pp.3087–3096.

Simons, J.N. et al., 1995. Isolation of novel virus-like sequences associated with human hepatitis. *Nature Medicine*, 1(6), pp.564–569.

Singh, H. & Raghava, G.P.S., 2003. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics (Oxford, England)*, 19(8), pp.1009–1014.

Skibbe, K. et al., 2014. P223 HLA-B*27 IS PROTECTIVE AGAINST HCV GENOTYPE 1 AND 3 AND ASSOCIATED WITH TARGETING OF DISTINCT GENOTYPE-SPECIFIC CD8+ T CELL EPITOPES. *Journal of Hepatology*, 60(1), p.S140.

Smith, D.B. et al., 2013. Expanded classification of hepatitis C Virus into 7 genotypes and 67 Subtypes: updated criteria and assignment web resource. *Hepatology (Baltimore, Md.)*.

Smuts, H.E. & Kannemeyer, J., 1995. Genotyping of hepatitis C virus in South Africa. *Journal of clinical microbiology*, 33(6), pp.1679–1681.

Söderholm, J. et al., 2006. Relation between viral fitness and immune escape within the hepatitis C virus protease. *Gut*, 55(2), pp.266–274.

Sommereyns, C. et al., 2008. IFN-lambda (IFN-lambda) is expressed in a tissue-dependent fashion and primarily acts on epithelial cells in vivo. *PLoS pathogens*, 4(3), p.e1000017.

Sourisseau, M. et al., 2013. Temporal analysis of hepatitis C virus cell entry with occludin directed blocking antibodies. *PLoS pathogens*, 9(3), p.e1003244.

Stapleton, J.T. et al., 2011. The GB viruses: a review and proposed classification of GBV-A, GBV-C (HGV), and GBV-D in genus Pegivirus within the family Flaviviridae. *The Journal of general virology*, 92(Pt 2), pp.233–246.

Stegmann, K.A. et al., 2010. Interferon-α–Induced TRAIL on Natural Killer Cells Is Associated With Control of Hepatitis C Virus Infection. *Gastroenterology*, 138(5), pp.1885–1897.e10.

Steinhauer, D.A., Domingo, E. & Holland, J.J., 1992. Lack of evidence for proofreading mechanisms associated with an RNA virus polymerase. *Gene*, 122(2), pp.281–288.

Stephenson, K.E. et al., 2012. Full-length HIV-1 immunogens induce greater magnitude and comparable breadth of T lymphocyte responses to conserved HIV-1 regions compared with conserved-region-only HIV-1 immunogens in rhesus monkeys. *Journal of virology*, 86(21), pp.11434–11440.

Stone, A.E.L. et al., 2013. Hepatitis C virus pathogen associated molecular pattern (PAMP) triggers production of lambda-interferons by human plasmacytoid dendritic cells. *PLoS pathogens*, 9(4), p.e1003316.

Storey, J.D. & Tibshirani, R., 2003. Statistical significance for genomewide studies. *Proceedings of the National Academy of Sciences*, 100(16), pp.9440–9445.

Strengell, M., Julkunen, Ilkka & Matikainen, Sampsa, 2004. IFN-alpha regulates IL-21 and IL-21R expression in human NK and T cells. *Journal of leukocyte biology*, 76(2), pp.416–422.

Sturniolo, T. et al., 1999. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nature biotechnology*, 17(6), pp.555–561.

Sugimoto, K. et al., 2003. Suppression of HCV-specific T cells without differential hierarchy demonstrated ex vivo in persistent HCV infection. *Hepatology (Baltimore, Md.)*, 38(6), pp.1437–1448.

Sulkowski, M. et al., 2014. Daclatasvir plus Sofosbuvir for Previously Treated or Untreated Chronic HCV Infection. *New England Journal of Medicine*, 370(3), pp.211–221.

Sumida, K. et al., 2013. Characteristics of splenic CD8+T cell exhaustion in patients with hepatitis C. *Clinical & Experimental Immunology*, 174(1), pp.172–178.

Summa, V. et al., 2012. MK-5172, a Selective Inhibitor of Hepatitis C Virus NS3/4a Protease with Broad Activity across Genotypes and Resistant Variants. *Antimicrobial Agents and Chemotherapy*, 56(8), pp.4161–4167.

Suntharalingam, G. et al., 2006. Cytokine storm in a phase 1 trial of the anti-CD28 monoclonal antibody TGN1412. *The New England journal of medicine*, 355(10), pp.1018–1028.

Suppiah, V. et al., 2009. IL28B is associated with response to chronic hepatitis C interferon-α and ribavirin therapy. *Nature Genetics*, 41(10), pp.1100–1104.

Swadling, L., Klenerman, P. & Barnes, E., 2013. Ever closer to a prophylactic vaccine for HCV. *Expert opinion on biological therapy*, 13(8), pp.1109–1124.

Sylwester, A.W. et al., 2005. Broadly targeted human cytomegalovirus-specific CD4+ and CD8+ T cells dominate the memory compartments of exposed subjects. *The Journal of Experimental Medicine*, 202(5), pp.673–685.

Szabo, S.J. et al., 2002. Distinct effects of T-bet in TH1 lineage commitment and IFN-gamma production in CD4 and CD8 T cells. *Science (New York, N.Y.)*, 295(5553), pp.338–342.

Tabatabai, N.M. et al., 1999. Functionally distinct T-cell epitopes within the hepatitis C virus non-structural 3 protein. *Human immunology*, 60(2), pp.105–115.

Takaki, A. et al., 2000. Cellular immune responses persist and humoral responses decrease two decades after recovery from a single-source outbreak of hepatitis C. *Nature Medicine*, 6(5), pp.578–582.

Takamiya, Y. et al., 1994. HLA-B*3501-peptide interactions: role of anchor residues of peptides in their binding to HLA-B*3501 molecules. *International immunology*, 6(2), pp.255–261.

Tanaka, K., 1994. Role of proteasomes modified by interferon-gamma in antigen processing. *Journal of leukocyte biology*, 56(5), pp.571–575.

Tanaka, Y. et al., 2004. Exponential Spread of Hepatitis C Virus Genotype 4a in Egypt. *Journal of Molecular Evolution*, 58(2), pp.191–195.

Tanaka, Y. et al., 2009. Genome-wide association of IL28B with response to pegylated interferon-[alpha] and ribavirin therapy for chronic hepatitis C. *Nat Genet*, 41(10), pp.1105–1109.

Tapper, E.B. & Afdhal, N.H., 2013. Is 3 the new 1: perspectives on virology, natural history and treatment for hepatitis C genotype 3. *Journal of viral hepatitis*, 20(10), pp.669–677.

Tay, C.H. & Welsh, R.M., 1997. Distinct organ-dependent mechanisms for the control of murine cytomegalovirus infection by natural killer cells. *Journal of virology*, 71(1), pp.267–275.

Tester, I. et al., 2005. Immune evasion versus recovery after acute hepatitis C virus infection from a shared source. *The Journal of Experimental Medicine*, 201(11), pp.1725 –1731.

Teyton, L. & Peterson, P.A., 1992. Invariant chain--a regulator of antigen presentation. *Trends in cell biology*, 2(2), pp.52–56.

Thimme, R. et al., 2001. Determinants of Viral Clearance and Persistence during Acute Hepatitis C Virus Infection. *The Journal of Experimental Medicine*, 194(10), pp.1395 –1406.

Thimme, R., Binder, M. & Bartenschlager, Ralf, 2012. Failure of innate and adaptive immune responses in controlling hepatitis C virus infection. *FEMS microbiology reviews*, 36(3), pp.663–683.

Thio, C.L. et al., 2002. HLA-Cw*04 and Hepatitis C Virus Persistence. *Journal of Virology*, 76(10), pp.4792–4797.

Thomas, D L et al., 2009. Genetic variation in IL28B and spontaneous clearance of hepatitis C virus. *Nature*, 461(7265), pp.798–801.

Thomas, E. et al., 2012. HCV infection induces a unique hepatic innate immune response associated with robust production of type III interferons. *Gastroenterology*, 142(4), pp.978–988.

Thomson, E.C., Smith, J.A. & Klenerman, P., 2011. The natural history of early hepatitis C virus evolution; lessons from a global outbreak in human immunodeficiency virus-1-infected individuals. *The Journal of general virology*, 92(Pt 10), pp.2227–2236.

Thurmond, J. et al., 2008. Web-based design and evaluation of T-cell vaccine candidates. *Bioinformatics*, 24(14), pp.1639–1640.

Thursz, M. et al., 1999. Influence of MHC class II genotype on outcome of infection with hepatitis C virus. The HENCORE group. Hepatitis C European Network for Cooperative Research. *Lancet*, 354(9196), pp.2119–2124.

Tiegs, G. & Lohse, A.W., 2010. Immune tolerance: what is unique about the liver. *Journal of autoimmunity*, 34(1), pp.1–6.

Timm, J. et al., 2004. CD8 Epitope Escape and Reversion in Acute HCV Infection. *The Journal of Experimental Medicine*, 200(12), pp.1593–1604.

Timmerman, J.M., 2009. Carrier protein conjugate vaccines: the “missing link” to improved antibody and CTL responses? *Human vaccines*, 5(3), pp.181–183.

Timpe, J.M. et al., 2008. Hepatitis C virus cell-cell transmission in hepatoma cells in the presence of neutralizing antibodies. *Hepatology (Baltimore, Md.)*, 47(1), pp.17–24.

Tindle, R.W. & Frazer, I.H., 1994. Immune response to human papillomaviruses and the prospects for human papillomavirus-specific immunisation. *Current topics in microbiology and immunology*, 186, pp.217–253.

Tseng, C.-T.K. & Klimpel, G.R., 2002. Binding of the hepatitis C virus envelope protein E2 to CD81 inhibits natural killer cell functions. *The Journal of experimental medicine*, 195(1), pp.43–49.

Urban, T.J. et al., 2010. IL28B genotype is associated with differential expression of intrahepatic interferon-stimulated genes in patients with chronic hepatitis C. *Hepatology (Baltimore, Md.)*, 52(6), pp.1888–1896.

Urbani, S., Amadei, B., Fisicaro, P., et al., 2006. Outcome of acute hepatitis C is related to virus‐specific CD4 function and maturation of antiviral memory CD8 responses. *Hepatology*, 44(1), pp.126–139.

Urbani, S., Amadei, B., Tola, D., et al., 2006. PD-1 expression in acute hepatitis C virus (HCV) infection is associated with HCV-specific CD8 exhaustion. *Journal of virology*, 80(22), pp.11398–11403.

Vali, B. et al., 2010. HCV-specific T cells in HCV/HIV co-infection show elevated frequencies of dual Tim-3/PD-1 expression that correlate with liver disease progression. *European Journal of Immunology*, 40(9), pp.2493–2505.

Velazquez, V.M. et al., 2012. Hepatic enrichment and activation of myeloid dendritic cells during chronic hepatitis C virus infection. *Hepatology (Baltimore, Md.)*, 56(6), pp.2071–2081.

Verstrepen, B.E. et al., 2011. Clearance of genotype 1b hepatitis C virus in chimpanzees in the presence of vaccine-induced E1-neutralizing antibodies. *The Journal of infectious diseases*, 204(6), pp.837–844.

Vigneron, N. et al., 2004. An antigenic peptide produced by peptide splicing in the proteasome. *Science (New York, N.Y.)*, 304(5670), pp.587–590.

Virgin, H.W., Wherry, E.J. & Ahmed, R., 2009. Redefining chronic viral infection. *Cell*, 138(1), pp.30–50.

Vita, R. et al., 2010. The immune epitope database 2.0. *Nucleic acids research*, 38(Database issue), pp.D854–862.

Von Wagner, M. et al., 2005. Peginterferon-α-2a (40KD) and Ribavirin for 16 or 24 Weeks in Patients With Genotype 2 or 3 Chronic Hepatitis C. *Gastroenterology*, 129(2), pp.522–527.

Walker, J. et al., 2014. Preclinical Characterization of GSK2336805, a Novel Inhibitor of Hepatitis C Virus Replication That Selects for Resistance in NS5A. *Antimicrobial Agents and Chemotherapy*, 58(1), pp.38–47.

Walker, M.R. et al., 2003. Induction of FoxP3 and acquisition of T regulatory activity by stimulated human CD4+CD25- T cells. *Journal of Clinical Investigation*, 112(9), pp.1437–1443.

Wan, J. et al., 2006. SVRMHC prediction server for MHC-binding peptides. *BMC bioinformatics*, 7, p.463.

Wang, H. & Eckels, D D, 1999. Mutations in immunodominant T cell epitopes derived from the nonstructural 3 protein of hepatitis C virus have the potential for generating escape variants that may have important consequences for T cell recognition. *Journal of immunology (Baltimore, Md.: 1950)*, 162(7), pp.4177–4183.

Wang, J. et al., 2007. Transient expression of FOXP3 in human activated nonregulatory CD4+ T cells. *European journal of immunology*, 37(1), pp.129–138.

Wang, J.H., Layden, Thomas J & Eckels, David D, 2003. Modulation of the peripheral T-Cell response by CD4 mutants of hepatitis C virus: transition from a Th1 to a Th2 response. *Human immunology*, 64(7), pp.662–673.

Wang, Yi et al., 2006. New epitope peptides derived from hepatitis C virus (HCV) 2a which have the capacity to induce cytotoxic T lymphocytes in HLA-A2+ HCV-infected patients. *Microbiology and immunology*, 50(11), pp.857–865.

Ward, S.M. et al., 2007. Quantification and localisation of FOXP3+ T lymphocytes and relation to hepatic inflammation during chronic HCV infection. *Journal of Hepatology*, 47(3), pp.316–324.

Wartelle-Bladou, C. et al., 2012. Hepatitis C therapy in non-genotype 1 patients: the near future. *Journal of Viral Hepatitis*, 19(8), pp.525–536.

Wawrzynowicz-Syczewska, M. et al., 2004. Natural history of acute symptomatic hepatitis type C. *Infection*, 32(3), pp.138–143.

Webster, D. et al., 2005. Enhanced T cell-mediated protection against malaria in human challenges by using the recombinant poxviruses FP9 and modified vaccinia virus Ankara. *Proceedings of the National Academy of Sciences of the United States of America*, 102(13), pp.4836–4841.

De Weerd, N.A. & Nguyen, T., 2012. The interferons and their receptors--distribution and regulation. *Immunology and cell biology*, 90(5), pp.483–491.

Weinberger, E.E. et al., 2013. The influence of antigen targeting to sub-cellular compartments on the anti-allergic potential of a DNA vaccine. *Vaccine*, 31(51), pp.6113–6121.

Wendt, A. et al., 2014. Chronic hepatitis C: future treatment. *Clinical Pharmacology : Advances and Applications*, 6, pp.1–17.

Wertheimer, A.M. et al., 2003. Novel CD4+ and CD8+ T‐cell determinants within the NS3 protein in subjects with spontaneously resolved HCV infection. *Hepatology*, 37(3), pp.577–589.

Westin, J. et al., 2002. Steatosis accelerates fibrosis development over time in hepatitis C virus genotype 3 infected patients. *Journal of Hepatology*, 37(6), pp.837–842.

Wherry, E.J., 2011. T cell exhaustion. *Nature Immunology*, 12(6), pp.492–499.

Wherry, E.J. & Ahmed, R., 2004. Memory CD8 T-cell differentiation during viral infection. *Journal of virology*, 78(11), pp.5535–5545.

WHO, H., 2014. WHO | Hepatitis C. *WHO*. Available at: http://www.who.int/mediacentre/factsheets/fs164/en/ [Accessed January 16, 2014].

Witte, K. et al., 2009. Despite IFN-[lambda] receptor expression, blood immune cells, but not keratinocytes or melanocytes, have an impaired response to type III interferons: implications for therapeutic applications of these cytokines. *Genes Immun*, 10(8), pp.702–714.

Witte, K. et al., 2010. IL-28A, IL-28B, and IL-29: promising cytokines with type I interferon-like properties. *Cytokine & growth factor reviews*, 21(4), pp.237–251.

Wölfl, M. et al., 2008. Hepatitis C virus immune escape via exploitation of a hole in the T cell repertoire. *Journal of Immunology (Baltimore, Md.: 1950)*, 181(9), pp.6435–6446.

Wolk, K. et al., 2008. Maturing dendritic cells are an important source of IL-29 and IL-20 that may cooperatively increase the innate immunity of keratinocytes. *Journal of leukocyte biology*, 83(5), pp.1181–1193.

Xu, L.-Z. et al., 1994. Hepatitis C virus genotype 4 is highly prevalent in central Africa (Gabon). *Journal of General Virology*, 75(9), pp.2393–2398.

Yamakawa, Y. et al., 1996. Higher elimination rate of hepatitis C virus among women. *Journal of Viral Hepatitis*, 3(6), pp.317–321.

Yanagi, M. et al., 1997. Transcripts from a single full-length cDNA clone of hepatitis C virus are infectious when directly transfected into the liver of a chimpanzee. *Proceedings of the National Academy of Sciences of the United States of America*, 94(16), pp.8738–8743.

Yang, G. et al., 2012. 845 PRECLINICAL CHARACTERISTICS OF ACH-3102: A NOVEL HCV NS5A INHIBITOR WITH IMPROVED POTENCY AGAINST GENOTYPE-1A VIRUS AND VARIANTS RESISTANT TO 1ST GENERATION OF NS5A INHIBITORS. Available at: http://www.sciencedirect.com/science/article/pii/S0168827812608575 [Accessed April 1, 2014].

Yang, Y. et al., 1995. Upregulation of class I major histocompatibility complex antigens by interferon gamma is necessary for T-cell-mediated elimination of recombinant adenovirus-infected hepatocytes in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, 92(16), pp.7257–7261.

Yano, A. et al., 2013. A novel method for enhancement of peptide vaccination utilizing T-cell epitopes from conventional vaccines. *Vaccine*, 31(11), pp.1510–1515.

Ydreborg, M. et al., 2013. Impact of IL28B-Related Single Nucleotide Polymorphisms on Liver Transient Elastography in Chronic Hepatitis C Infection. *PLoS ONE*, 8(11), p.e80172.

Yerly, D. et al., 2008. Design, expression, and processing of epitomized hepatitis C virus-encoded CTL epitopes. *Journal of immunology (Baltimore, Md.: 1950)*, 181(9), pp.6361–6370.

Yewdell, J.W., 2006. Confronting complexity: real-world immunodominance in antiviral CD8+ T cell responses. *Immunity*, 25(4), pp.533–543.

Yin, Z. et al., 2012. Type III IFNs are produced by and stimulate human plasmacytoid dendritic cells. *Journal of immunology (Baltimore, Md.: 1950)*, 189(6), pp.2735–2745.

Yokozaki, S. et al., 2011. Mutations in two PKR-binding domains in chronic hepatitis C of genotype 3a and correlation with viral loads and interferon responsiveness. *Journal of Medical Virology*, 83(10), pp.1727–1732.

Youn, J.-W. et al., 2008. Evidence for protection against chronic hepatitis C virus infection in chimpanzees by immunization with replicating recombinant vaccinia virus. *Journal of virology*, 82(21), pp.10896–10905.

Zeuzem, S. et al., 2011. Efficacy of the protease inhibitor BI 201335, polymerase inhibitor BI 207127, and ribavirin in patients with chronic HCV infection. *Gastroenterology*, 141(6), pp.2047–2055; quiz e14.

Zeuzem, S. et al., 1998. Hepatitis C virus dynamics in vivo: Effect of ribavirin and interferon alfa on viral turnover. *Hepatology*, 28(1), pp.245–252.

Zeuzem, S. et al., 2004. Peginterferon alfa-2b plus ribavirin for treatment of chronic hepatitis C in previously untreated patients infected with HCV genotypes 2 or 3. *Journal of Hepatology*, 40(6), pp.993–999.

Zeuzem, S. et al., 2012. PEGINTERFERON LAMBDA-1a (LAMBDA) COMPARED TO PEGINTERFERON ALFA-2A (ALFA) IN TREATMENT-NAÏVE PATIENTS WITH HCV GENOTYPES (G) 2 or 3: FIRST SVR24 RESULTS FROM EMERGE PHASE IIB. Available at: http://mobile.ilcapp.eu/EASL_161/poster_23758/program.aspx [Accessed April 10, 2014].

Zhang, G.L. et al., 2005. MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides. *Nucleic acids research*, 33(Web Server issue), pp.W172–179.

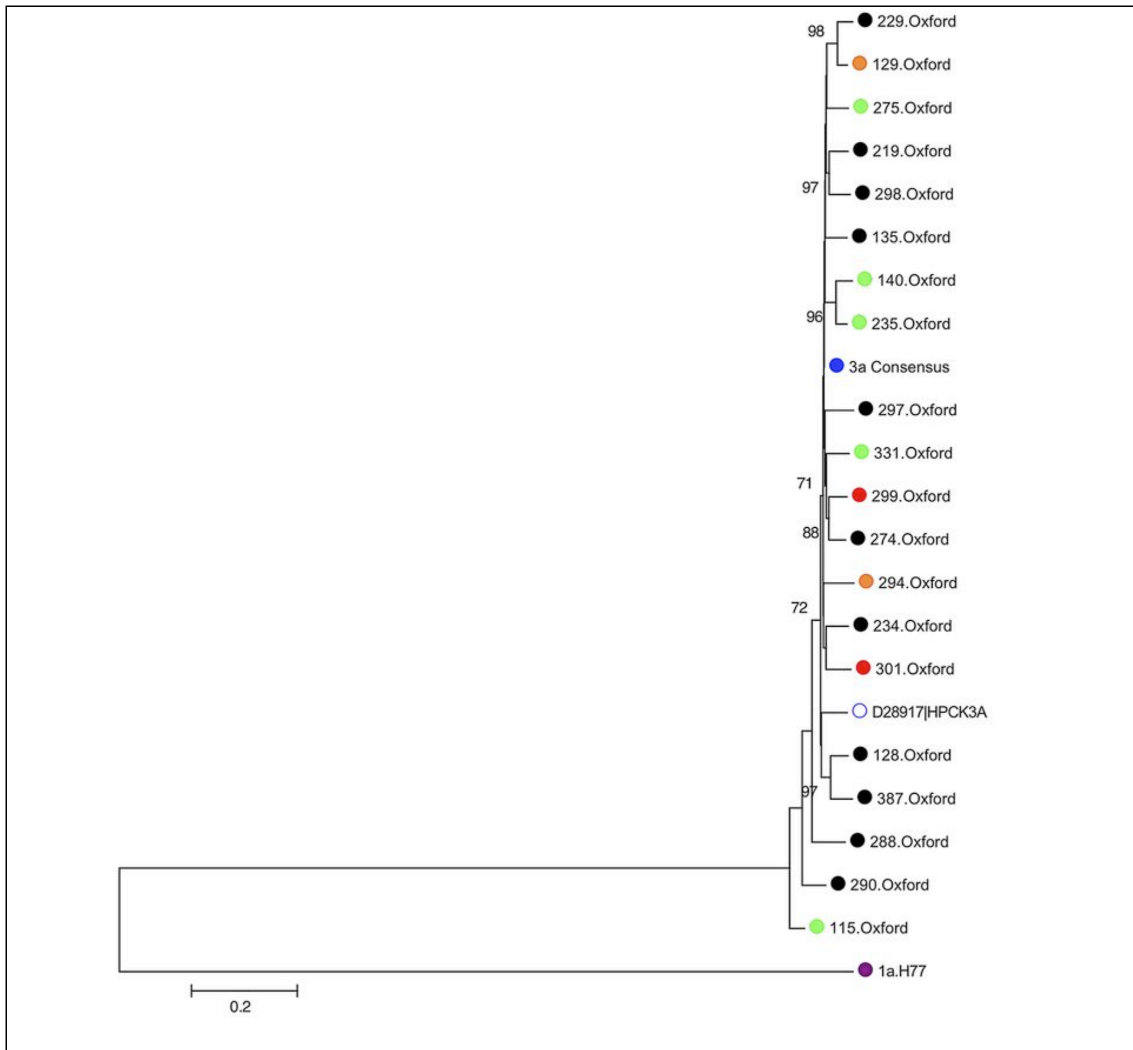
Zhang, G.L. et al., 2007. Prediction of supertype-specific HLA class I binding peptides using support vector machines. *Journal of immunological methods*, 320(1-2), pp.143–154.

Zhao, D.-M. et al., 2006. Activated CD4+CD25+ T cells selectively kill B lymphocytes. *Blood*, 107(10), pp.3925–3932.

8 Appendix

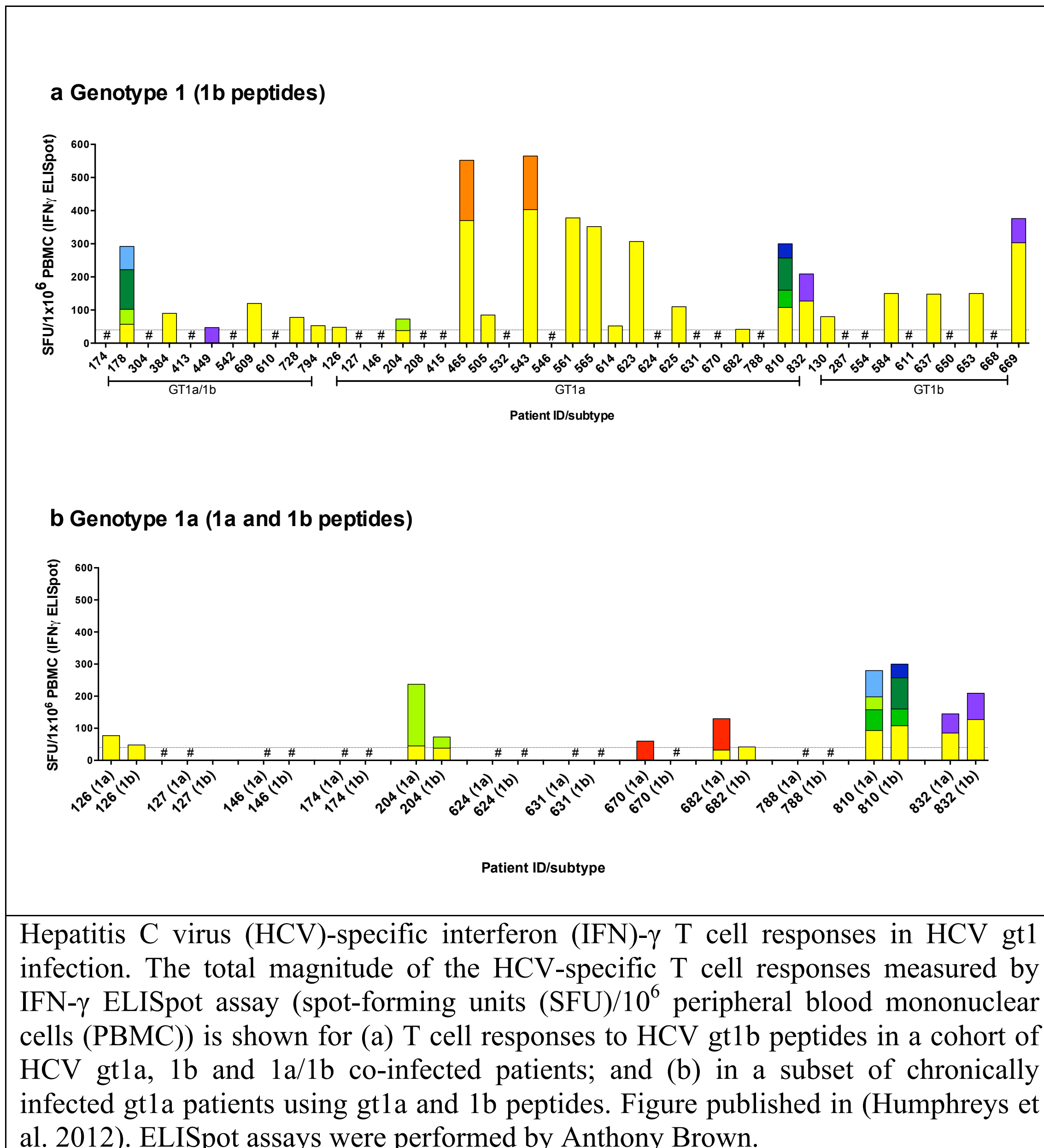
8.1 Supplementary figures

Figure 8-1: Sequence diversity of full-length genotype-3a sequences.



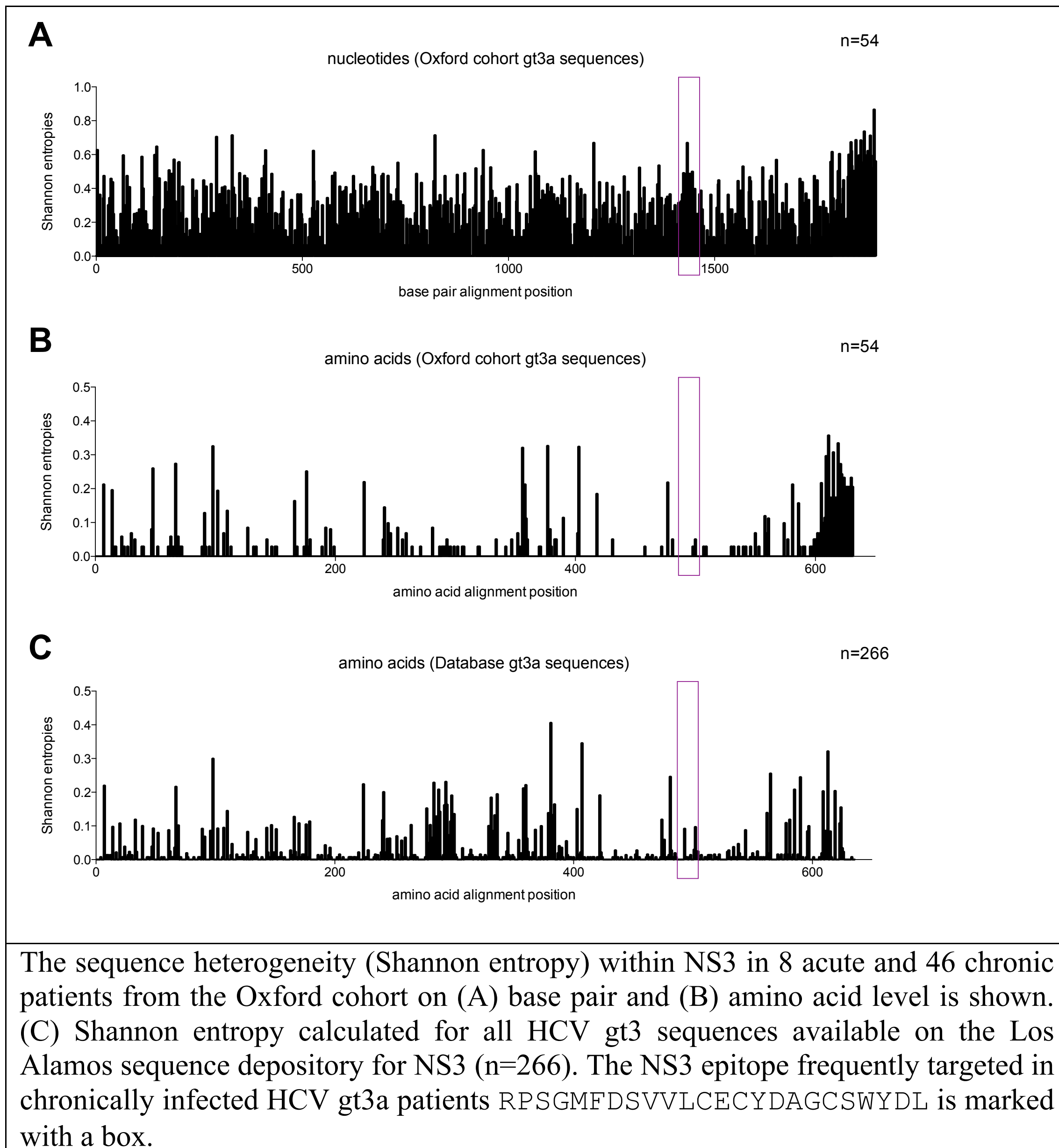
Neighbour-joining tree of full-length nucleotide sequences from 20 genotype-3a infected chronic patients, including eight patients followed longitudinally through combination therapy (green = sustained virological response, orange = relapse, red = non-responders, genotype-3a consensus sequence = blue). Also included are the genotype-3a peptide consensus sequence, a genotype-3a reference sequence (accession number D28917), together with H77 genotype-1a nucleotide sequence (accession number AF009606) used as an outgroup. Bootstrap scores >70% are shown. Data for this Figure was generated by Isla Humphreys.

Figure 8-2: T cell responses in HCV gt1 infection measured by IFNγ ELISPot assays using HCV subtype specific peptide sets.



Hepatitis C virus (HCV)-specific interferon (IFN)-γ T cell responses in HCV gt1 infection. The total magnitude of the HCV-specific T cell responses measured by IFN-γ ELISpot assay (spot-forming units (SFU)/10^6 peripheral blood mononuclear cells (PBMC)) is shown for (a) T cell responses to HCV gt1b peptides in a cohort of HCV gt1a, 1b and 1a/1b co-infected patients; and (b) in a subset of chronically infected gt1a patients using gt1a and 1b peptides. Figure published in (Humphreys et al. 2012). ELISpot assays were performed by Anthony Brown.

Figure 8-3: Sequence heterogeneity (Shannon entropy) of NS3 sequences



The sequence heterogeneity (Shannon entropy) within NS3 in 8 acute and 46 chronic patients from the Oxford cohort on (A) base pair and (B) amino acid level is shown. (C) Shannon entropy calculated for all HCV gt3 sequences available on the Los Alamos sequence depository for NS3 (n=266). The NS3 epitope frequently targeted in chronically infected HCV gt3a patients RPSGMFDSVVLCECYDAGCSWYDL is marked with a box.

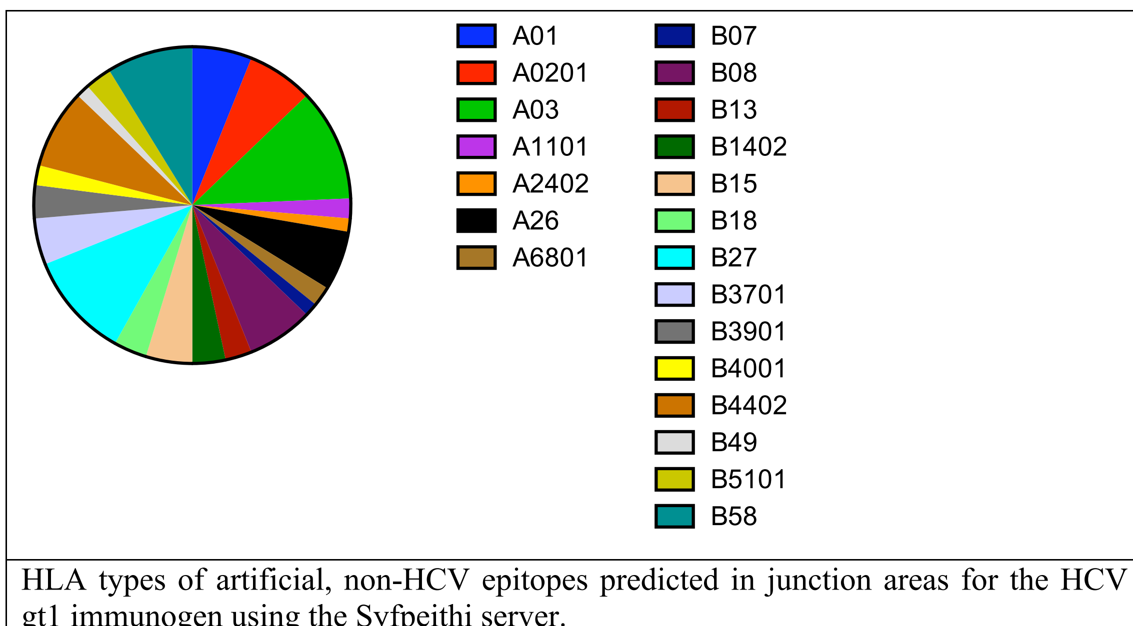
8.2 Comparative analysis of CD8+ restricted epitopes predicted by three different online epitope prediction servers in junction regions between conserved epitopes

MHC class-I epitopes in junction areas between conserved segments were predicted using three different online prediction servers: NetMHC, Syfpeithi and BIMAS. Junction regions were defined as 10 amino acids on each site of the junction between conserved segments. For all analyses, strong binders (defined according to each server's cut-off) were initially predicted for all available HLA types of each server. Predicted epitopes not falling on segment junction sites were excluded (e.g. 8mers and 9mers at each side of the junction). Subsequently, matching HLA types overlapping between servers that were to be compared were selected. As an example for the analysis, results for the HCV gt1 immunogen approach D (long construct) are discussed.

8.2.1 NetMHC – Syfpeithi

Using the Syfpeithi server, 205 strong binders were predicted in junction areas of the HCV gt1 immunogen (approach D) for all available HLA types of the server. Of these, 57 (28%) were excluded since they did not overlap with junction areas (e.g. 8mers and 9mers on each side of the junction). The HLA type distribution of predicted artificial epitopes predicted by the Syfpeithi server is shown in Figure 8-4.

Figure 8-4: HLA type distribution of predicted strong binders forming artificial epitopes in junction areas using the Syfpeithi prediction server.



HLA types of artificial, non-HCV epitopes predicted in junction areas for the HCV gt1 immunogen using the Syfpeithi server.

To compare MHC class-I prediction results of Syfpeithi and NetMHC servers, HLA types only available on either Syfpeithi or NetMHC prediction servers were excluded from the comparison. HLA types available on both servers and therefore included in the comparative analyses are marked in green in Table 8-1.

Table 8-1: MHC class-I types assessed for epitope prediction using the Syfpeithi prediction server.

NetMHC		Syfpeithi	
HLA-A	mers	HLA-A	mers
A*0101	8	A*01	8
A*0101	9	A*01	9
A*0101	10	A*01	10
A*0101	11	A*01	11
A*0201	8		
A*0201	9	A*0201	9
A*0201	10	A*0201	10
A*0201	11		
		A*0202	9
		A*0202	10
		A*0203	8
		A*0203	9
		A*0203	10
		A*0204	9
		A*0205	10
		A*0206	9
		A*0207	9
A*0301	8		
A*0301	9	A*03	9
A*0301	10	A*03	10
A*0301	11		
		A*1101	9
		A*1101	10
		A*1101	11
		A*23	9
A*2402	8		
A*2402	9	A*2402	9
A*2402	10	A*2402	10
A*2402	11		
		A*25	9
A*2601	8		
A*2601	9	A*26	9
A*2601	10	A*26	10
A*2601	11		
		A*2602	9
		A*29	9
		A*3101	9
		A*6801	9
		A*6801	10
		A*6801	11
		A*6801	15
		A*6802	9

NetMHC		Syfpeithi	
HLA-B	mers	HLA-B	mers
B*0702	8		
B*0702	9	B*0702	9
B*0702	10	B*0702	10
B*0702	11		
B*0801	8	B*08	8
B*0801	9	B*08	9
B*0801	10		
B*0801	11		
		B*13	9
		B*13	10
		B*1402	8
		B*1402	9
B*1501	8		
B*1501	9	B*1501	9
B*1501	10	B*1501	10
B*1501	11		
		B*1509	9
		B*1510	9
		B*1516	9
		B*18	8
		B*18	9
		B*1801	9
		B*2702	9
		B*2704	9
B*2705	8		
B*2705	9	B*2705	9
B*2705	10	B*2705	10
B*2705	11		
		B*2706	9
		B*2709	9
		B*3501	9
		B*3501	10
		B*37	8
		B*37	9
		B*3801	8
		B*3801	9
		B*3801	10
B*3901	8		
B*3901	9	B*3901	9
B*3901	10	B*3901	10
B*3901	11		
		B*3902	9

NetMHC		Syfpeithi	
HLA-B	mers	HLA-B	mers
		B*3902	10
B*4001	8		
B*4001	9	B*4001	9
B*4001	10	B*4001	10
B*4001	11		
		B*4002	9
		B*4002	10
		B*41	9
		B*4101	9
		B*4101	10
		B*4105	9
		B*44	9
		B*4402	9
		B*4402	10
		B*4402	11
		B*45	9
		B*4501	9
		B*4501	10
		B*4701	9
		B*4701	10
		B*4901	9
		B*4901	10
		B*5001	9
		B*5001	10
		B*51	10
		B*5101	8
		B*5101	9
		B*5301	9
		B*5301	10
		B*57	9
		B*5701	9
		B*5701	10
B*5801	8		
B*5801	9	B*5801	9
B*5801	10		
B*5801	11		
		B*5802	9
		B*5802	10

MHC class-I types available for prediction analysis on the Syfpeithi server. The HLA types which were analysed using the NetMHC server and then compared to the Syfpeithi results are marked in green.

Of 148 artificial epitopes predicted using the Syfpeithi server, 49 (33%) were excluded because of an HLA mismatch between Syfpeithi and NetMHC servers (Table 8-2). Of the strong binders predicted using the NetMHC server, none were excluded due to HLA mismatch.

Table 8-2: Predicted strong binders in junction areas excluded from the comparative analysis, because of non-matching HLA types in the comparative analysis between NetMHC and Syfpeithi servers

HLA type	mers	Junction	peptide	Syfpeithi score
A*1101	9	15	IVLSGFWAK	23
A*1101	10	16	TTQGSIGLGK	25
A*1101	9	27	GTQEDASLR	24
A*6801	9	12	EVNAVAYYR	22
A*6801	9	23	DVRCHKPAR	21
A*6801	10	25	DIRTESLTER	23
B*13	10	6	RLWDRDRSEL	21
B*13	10	9	VQIVSTQSFL	21
B*13	9	9	TQSFLATCI	22
B*13	10	16	TQGSIGLGKV	22
B*1402	9	4	DRTDVFLLN	21
B*1402	8	4	DRTDVFLL	23
B*1402	8	14	IRLKPKYI	23
B*1402	8	19	VRLHRRRL	25
B*1402	8	28	MILEPLDL	26
B*1516	9	1	LTVPASAYV	28
B*1516	9	3	WSPVCGPVY	22
B*18	9	12	IEEVNAVAY	22
B*18	8	12	EEVNAVAY	26
B*18	9	25	TESLTERLY	21
B*18	8	25	TESLTERL	21
B*18	9	27	QEDASLRAF	22
B*2709	9	14	IRLKPKYIM	22
B*2709	9	23	VRCHKPARL	21

HLA type	mers	Junction	peptide	Syfpeithi score
B*37	9	4	TDRTDVFLL	21
B*37	8	4	TDRTDVFL	21
B*37	8	6	WDRDRSEL	21
B*37	9	21	SEALSNSLL	25
B*37	8	21	SEALSNSL	24
B*37	8	24	YDVVPMGF	24
B*37	8	25	TESLTERL	28
B*4402	11	8	LENLVKGWRLL	23
B*4402	10	8	LENLVKGWRL	21
B*4402	11	9	GEVQIVSTQSF	21
B*4402	10	12	IEEVNAVAYY	21
B*4402	9	12	EEVNAVAYY	24
B*4402	9	12	IEEVNAVAY	25
B*4402	9	21	SEALSNSLL	23
B*4402	10	25	TESDIRTESL	23
B*4402	9	25	TESLTERLY	24
B*4402	9	27	QEDASLRAF	25
B*4402	10	28	LEPLDLPQII	21
B*4402	9	28	LEPLDLPQI	22
B*4901	10	8	LENLVKGWRL	22
B*4901	10	25	TESLTERLYV	22
B*5101	8	2	SPRYPGHI	24
B*5101	9	11	HGVTGVRTI	22
B*5101	8	11	KAHGVTGV	22
B*5101	9	16	SPLTTQGSI	24

Strong binders predicted by the Syfpeithi server that were excluded from the comparative analysis between Syfpeithi and NetMHC servers due to non-matching HLA types between the two servers.

Next, the remaining 169 artificial strong binders predicted by both NetMHC (70 epitopes) and Syfpeithi (99 epitopes) servers were compared (Table 8-3): 24 epitopes predicted by both servers were identical (marked in red), reflecting an overlap of 34% (NetMHC) and 24% (Syfpeithi). All other predicted strong binders were either predicted for different HLA types or lengths. If the HLA type matched for two epitopes, and they only differed in one or two amino acids in length, they were classified as “similar”: this applied to 8 more epitopes predicted by Syfpeithi and NetMHC, marked in grey in Table 8-3, increasing the overlap from 34% to 46% (NetMHC) and 24% to 32%, respectively.

Table 8-3: Comparison of MHC class-I prediction results using NetMHC and Syfpeithi servers (table on the next pages).

Comparison of epitope prediction results for NetMHC and Syfpeithi servers. Epitopes predicted by both servers are marked in red. Similar epitopes (epitopes of different length predicted for the same restricting HLA type) were marked in grey.

Junction ID	Junction sequence	HLA type	Syfpeithi	NetMHC	
1	LSCLTVPASAYVGDLCGSVF	A*0101		LTVPASAY	
		A*0201	CLTVPASAYV		
		A*03	CLTVPASAY		
		A*2601		LTVPASAY	
		B*0702	VPASAYVGDL	VPASAYVGDL	
		B*1501	CLTVPASAY		similar
		B*1501		LTVPASAY	similar
2	VSQLFTFSPRYPGHITGHRM	A*01	VSQLFTFSPRY	VSQLFTFSPRY	
		B*0702	SPRYPGHIT	SPRYPGHIT	similar
		B*0702		SPRYPGHI	similar
		B*08	SPRYPGHI		
		B*1501	QLFTFSPRY		
		B*2705	PRYPGHITGH		
3	AWDMMMNWSPVCGPVYCFTP	A*0201		MMNWSPVCGPV	
		A*26	PVCGPVYCF		
		B*0702		GPVYCFTPSPV	
4	SPVVVGTTDRTDVFLLNNTR	A*0101		TTDRTDVF	similar
		A*0101		TTDRTDVFLL	similar
		A*26	TTDRTDVFLL		
5	LGNWFGCTWMCPTDCFRKHP	none			
6	CLVDYPYRLWDRDRSELSPL	A*0201	RLWDRDRSEL		
		B*08	WDRDRSEL		
		B*2705		YRLWDRDRSEL	
		B*3901		YRLWDRDRSEL	
7	QYLYGVGSSVARVCACLWMM	A*0201	GVGSSVARV		
		A*03	YLYGVGSSVA		
		B*08	VARVCACL		
		B*5801		SSVARVCACLW	similar
		B*5801		SVARVCACLW	similar
		B*5801		VARVCACLW	similar
8	QAEAALENLVKGWRLLAPIT	A*0201	NLVKGWRLL		
		A*03	AEAALENLVK		
		B*08	ENLVKGWRL		
		B*4001	LENLVKGWRL	LENLVQGWRL	similar
		B*4001	LENLVQGWRLL	LENLVQGWRLL	similar
		B*58	AALENLVKGW	AALENLVQGW	
9	VEGEVQIVSTQSFLATCING	A*26	EVQIVSTQSF		similar
		A*26	EVQIVSTQS		similar
		B*1501	VQIVSTQSF	VQIVSTQSF	
10	VCWTVYHGAGPCTCGSSDLY	none			
11	GAYMSKAHGVTGVRTITTGS	A*0201	SKAHGVTGV		
12	TVPHPNIEEVNAVAYYRGLD	A*01	NIEEVNAVAYY	NIEEVNAVAYY	similar
		A*01	NIEEVNAVAY		similar
		A*01	IEEVNAVAYY		similar
		A*01	IEEVNAVAY		similar
		A*03	NIEEVNAVAY		
		A*26	EEVNAVAYY	EVNAVAYY	
13	QRRGRTGRGRYRFVTPGERP	B*2705	RRGRTGRGRY		
		B*2705	GRTGRGRYRF		similar
		B*2705	GRTGRGRYR		similar
		B*2705		GRYRFVTPGER	similar
		B*58	RTGRGRYRF		
14	MWKCLIRLKPKYIMTCMSAD	A*03	RLKPKYIMTC		
		A*03	CLIRLKPKY		
		B*0702		KPKYIMTCM	
		B*08	LIRLKPKYI	LIRLKPKYI	similar
		B*08	LIRLKPKY		similar
		B*0801		LIRLKPKYIM	similar
		B*08	RLKPKYIM	RLKPKYIM	similar1
		B*08	RLKPKYIMT		similar1
		B*0801		RLKPKYIMTCM	similar1
		B*1501	CLIRLKPKY		
		B*2705	IRLKPKYIM		
15	VIVGRIVLSGFWAKHMWNFI	A*03	IVGRIVLSGF		
		A*03	RIVLSGFWAK	RIVLSGFWAK	similar
		A*03	IVLSGFWAKH		similar
		A*03	IVLSGFWAK		similar
		A*0301		VLSGFWAK	similar
		A*2402		GFWAKHMWNF	
		A*26	IVGRIVLSGF		
		B*2705		GRIVLSGFWAK	
		B*58	LSGFWAKHM		similar1
		B*5801		LSGFWAKHMW	similar1
		B*5801		SGFWAKHMW	similar1
		B*5801		IVGRIVLSGFW	similar
		B*5801		RIVLSGFW	similar

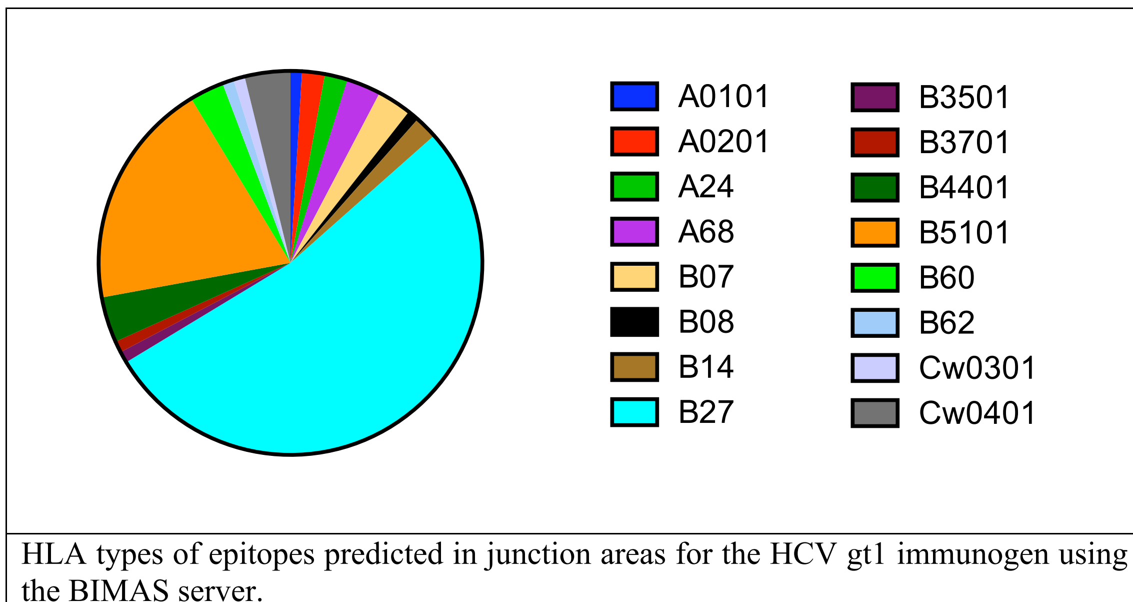
Junction ID	Junction sequence	HLA type	Syfpeithi	NetMHC	
16	ASVTSPLTTQGSIGLGKVLV	B*0702		SPLTTQGSI	similar
		B*0702		SPLTTQGSIGL	similar
17	AAARVTQILSSWLRDIWDWI	A*0201	QILSSWLRDI		similar
		A*0201	ILSSWLRDI		simialr
		B*2705		ARVTQILSSWL	
		B*58	SSWLRDIWDW	SSWLRDIWDW	
		B*58	RVTQILSSW	RVTQILSSW	
		B*58	LSSWLRDIW	LSSWLRDIW	
		B*5801		AARVTQILSSW	similar1
		B*5801		VTQILSSW	similar1
		B*5801		LSSWLRDIWDW	similar2
		B*5801		SSWLRDIW	similar2
18	VLSDFKTWLKCPCQVPSPEF none				
19	TELDGVRLHRRRLARGSPPS	A*03	RLHRRRLAR		
		B*2705	VRLHRRRLAR		similar
		B*2705		VRLHRRRL	similar
20	SLKATCTTNHSDAESYSSMP	A*01	CTTNHSDAESY	CTTNHSDAESY	
		A*01	TTNHSDAESY	TTNHSDAESY	
		A*0101		HSDAESYSSM	
		B*3901		NHSDAESYSSM	
21	GSWSTVSSEALSNSLLRHHN	A*03	ALSNSLLRH		
		B*4001	SEALSNSLL	SEALSNSLL	similar
		B*4001		SEALSNSL	similar
		B*58	VSSEALSNSL		similar1
		B*58	SSEALSNSLL		similar1
		B*58	SSEALSNSL		similar1
22	MVYATTSRSALTPPHSAKSK	A*03	ALTPPHSAK	ALTPPHSAK	similar
		A*0301		RSALTPPHSAK	similar
		A*2402	VYATTSRSAL		
		B*0702		MVYATTSRSAL	similar1
		B*0702		YATTSRSAL	similar1
23	GYGAKDVRCHKPARLIVFPD	A*26	DVRCHKPARL		
		B*08	VRCHKPARL		
		B*2705	VRCHKPARL		
		B*3901	CHKPARLIV		? similar
		B*3901	VRCHKPARL		? similar
24	CEKMALYDVVPMGFSYDTRC	A*01	LYDVVPMGFSY		
		A*0201	KMALYDVVPM		similar
		A*0201		ALYDVVPM	similar
		A*03	ALYDVVPMGF		
		A*03	DVVPMGFSY		
		A*2402	LYDVVPMGF	LYDVVPMGF	
		A*26	DVVPMGFSYD		
		A*26	DVVPMGFSY	DVVPMGFSY	
		B*1501	ALYDVVPMGF		similar1
		B*1501		KMALYDVVPM	similar1
		B*3901		EKMALYDVVPM	
		B*5801		MALYDVVPMGF	
25	TVTESDIRTESLTERLYVGG	A*01	RTESLTERLY	RTESLTERLY	
		B*2705	IRTESLTERL		similar
		B*2705	IRTESLTER		similar
		B*3901	IRTESLTERL		
		B*4001	TESDIRTESL	TESDIRTESL	
		B*58	RTESLTERLY		similar1
		B*58	RTESLTERL		similar1
26	ASAACRAAKLCTMLVCGDDL	B*08	AAKLCTML		
		B*2705	CRAAKLCTML		similar
		B*2705	CRAAKLCTM	CRAAKLCTM	similar
		B*3901	CRAAKLCTML		
		B*58	RAAKLCTML		
27	CESAGTQEDASLRAFTEAMT	B*58	ASLRAFTEAM		
28	APTLWARMILEPLDLPQIIQ	A*0201	ILEPLDLPQI		
		A*0201	RMILEPLDL		
		B*0801		WARMILEPL	
		B*2705	ARMILEPLDL		similar
		B*2705		ARMILEPL	similar
		B*3901	ARMILEPLDL		
29	WRHRARSVRAGGRAAICGKY	A*03	SVRAGGRAAI		similar
		A*03	SVRAGGRAA		similar
		B*0702		SVRAGGRAAI	similar1
		B*0702		SVRAGGRAA	similar1
		B*2705	VRAGGRAAI		
		B*2705	ARSVRAGGR		

8.2.2 NetMHC - BIMAS

Next, an epitope prediction analysis was performed for junction areas of the HCV gt1 immunogen (approach D) using the BIMAS epitope prediction server. The search was run for all available HLA types and including all available peptide lengths (8mers, 9mers and 10mers). Epitopes with a BIMAS score >100 were defined as strong binding epitopes.

159 strong binders in junction areas (10 amino acids on each side of the junction area) were predicted for the HCV gt1 D immunogen, of which 103 were artificial, non-HCV epitopes. Of note, HLA type distribution for the predicted epitopes in junction areas using the BIMAS server was skewed towards B*2705 and B*5101 in comparison to the other servers (Figure 8-5).

Figure 8-5: HLA type distribution of predicted strong binders forming artificial epitopes in junction areas using the BIMAS prediction server.



HLA types of epitopes predicted in junction areas for the HCV gt1 immunogen using the BIMAS server.

To compare results from the BIMAS analysis with those from the NetMHC analysis, I first determined HLA types shared by both servers (marked in blue in Table 8-4).

Table 8-4: MHC class-I types assessed for epitope prediction using the BIMAS prediction server.

NetMHC		BIMAS		NetMHC		BIMAS		NetMHC		BIMAS	
HLA-A	mers	HLA-A	mers	HLA-B	mers	HLA-B	mers	HLA-B	mers	HLA-B	mers
A*0101	8			B*0702	8			B*3901	9		
A*0101	9	A*01	9	B*0702	9	B*07	9	B*3901	10		
A*0101	10	A*01	10	B*0702	10	**B*07**	10	B*3901	11		
A*0101	11			B*0702	11			B*4001	8		
A*0201	8			B*0801	8			B*4001	9	B*40	9
A*0201	9	A*0201	9	B*0801	9	B*08	9	B*4001	10	B*40	10
A*0201	10	A*0201	10	B*0801	10	B*08	10	B*4001	11		
A*0201	11			B*0801	11					B*4403	9
A*0301	8					B*14	9			B*4403	10
A*0301	9	A*03	9			B*14	10			B*5101	8
A*0301	10	A*03	10	B*1501	8					B*5101	9
A*0301	11			B*1501	9	B*15	9			B*5101	10
		A*1101	9	B*1501	10	B*15	10			B*5102	9
		A*1101	10	B*1501	11					B*5102	10
A*2402	8					B*2702	9			B*5103	9
A*2402	9	A*24	9			B*2702	10			B*5103	10
A*2402	10	A*24	10	B*2705	8	B*2705	8			B*5201	8
A*2402	11			B*2705	9	B*2705	9			B*5201	9
A*2601	8			B*2705	10	B*2705	10			B*5201	10
A*2601	9			B*2705	11			B*5801	8		
A*2601	10					B*3501	8	B*5801	9		
A*2601	11					B*3501	9	B*5801	10		
		A*3101	9			B*3501	10	B*5801	11		
		A*3101	10			B*3701	9			B*60	9
		A*3302	9			B*3701	10			B*60	10
		A*3302	10			B*3801	9			B*61	9
		A*68	9			B*3801	10			B*61	10
		A*68	10	B*3901	8						

MHC class-I types avaiable for prediction analysis on the BIMAS server. The HLA types which were analysed using the NetMHC server and then compared to the BIMAS results are marked in blue.

Epitopes predicted for these HLA types were included in the comparative analysis. Of 103 artificial strong binders predicted by the BIMAS server, 42 (41%) were excluded from the BIMAS prediction dataset since HLA types did not match those used in the NetMHC analysis (Table 8-5, left). Likewise, of artificial 70 epitopes predicted for the NetMHC analysis, 24 were excluded (Table 8-5, right).

Table 8-5: Excluded predicted strong binders in junction areas in comparative analysis, because of non-matching HLA types in comparative analysis between NetMHC and BIMAS servers.

HLA type	mers	Junction	peptide	BIMAS score
A*68	9	12	EVNAVAYYR	600
A*68	9	15	IVLSGFWAK	240
A*68	9	23	DVRCHKPAR	600
B*14	10	19	DGVRLHRRRL	135
B*14	9	19	VRLHRRRLA	250
B*2702	10	13	RRGRTGRGRY	600
B*2702	10	13	GRTGRGRYRF	200
B*2702	9	15	GRIVLSGFW	100
B*2702	10	17	ARVTQILSSW	100
B*3501	9	14	KPKYIMTCM	240
B*3701	9	4	TDRTDVFLL	200
B*4403	10	12	IEEVNAVAYY	540
B*4403	9	12	IEEVNAVAY	180
B*4403	9	12	EEVNAVAYY	1080
B*4403	9	25	TESLTERLY	120
B*5101	10	1	VPASAYVGDL	130
B*5101	8	2	SPRYPGHI	400
B*5101	10	10	GPCTCGSSDL	110
B*5101	9	11	HGVTGVRTI	106.48
B*5101	10	12	HPNIEEVNAV	220
B*5101	9	16	SPLTTQGSI	484
B*5101	8	21	EALSNSLL	100
B*5101	9	26	AAKLCTMLV	100
B*5101	8	29	RAGGRAAI	110
B*5102	10	1	VPASAYVGDL	100
B*5102	10	7	YGVGSSVARV	528
B*5102	10	10	GPCTCGSSDL	121
B*5102	9	11	HGVTGVRTI	351.384
B*5102	10	12	HPNIEEVNAV	200
B*5102	9	16	SPLTTQGSI	1597.2
B*5102	10	19	DGVRLHRRRL	145.2
B*5102	9	22	YATTSRSAL	110
B*5102	10	26	RAAKLCTMLV	121
B*5103	10	26	RAAKLCTMLV	121
B*5103	9	26	AAKLCTMLV	133.1
B*60	9	8	LENLVKGWRL	176
B*60	9	21	SEALSNSLL	320
B*60	9	25	TESDIRTESL	320
Cw*0401	10	1	VPASAYVGDL	160
Cw*0401	10	15	GFWAKHMWNF	132
Cw*0401	10	22	VYATTSRSAL	200
Cw*0401	9	24	LYDVVPMGF	150

HLA type	mers	Junction	peptide	NetMHC score
A*2601	8	1	LTVPASAY	0.3
B*1501	8	1	LTVPASAY	0.5
B*3901	11	6	YRLWDRDRSEL	0.25
B*5801	9	7	VARVCACLW	0.15
B*5801	10	7	SVARVCACLW	0.25
B*5801	11	7	SSVARVCACLW	0.075
B*5801	10	8	AALENLVQGW	0.175
B*1501	9	9	VQIVSTQSF	0.025
A*2601	8	12	EVNAVAYY	0.05
B*5801	8	15	RIVLSGFW	0.5
B*5801	9	15	SGFWAKHMW	0.5
B*5801	10	15	LSGFWAKHMW	0.15
B*5801	11	15	IVGRIVLSGFW	0.5
B*5801	8	17	VTQILSSW	0.15
B*5801	8	17	SSWLRDIW	0.175
B*5801	9	17	RVTQILSSW	0.125
B*5801	9	17	LSSWLRDIW	0.075
B*5801	10	17	SSWLRDIWDW	0.2
B*5801	11	17	AARVTQILSSW	0.4
B*5801	11	17	LSSWLRDIWDW	0.075
B*3901	11	20	NHSDAESYSSM	0.2
A*2601	9	24	DVVPMGFSY	0.1
B*3901	11	24	EKMALYDVVPM	0.5
B*5801	11	24	MALYDVVPMGF	0.4

Strong binders predicted by the BIMAS server (left) and the NetMHC server (right) that were excluded from the comparative analysis between BIMAS and NetMHC servers due to non-matching HLA types between the two servers.

Next, I compared the remaining 107 artificial strong binders [46 (NetMHC), 61 (BIMAS)] predicted for matching HLA types between the NetMHC and BIMAS servers (Table 8-6): only 5 epitopes predicted by both servers were identical (marked in red), estimating at about 11% of epitopes for the NetMHC analysis and 8% for the BIMAS analysis. All other predicted strong binders were either predicted for different HLA types or lengths. If the HLA type matched for two epitopes, and they only differed in one or two amino acids in length, they were classified as “similar”. 5 more epitopes predicted by BIMAS and NetMHC servers only varying in length were classified as “similar” and are marked in grey in Table 8-6, increasing the match percentage from 11% to 22% (NetMHC) and 8% to 16% (BIMAS), respectively.

Table 8-6: Comparison of MHC class-I prediction results using NetMHC and BIMAS servers.

Comparison of epitope prediction results for BIMAS and NetMHC servers. Epitopes predicted by both servers are marked in red. Similar epitopes (epitopes of different length predicted for the same restricting HLA type) were marked in grey.

junction ID	Junction sequence	HLA* type	BIMAS	NetMHC	
1	LSCLTVPASAYVGDLCGSVF	A*0101		LTVPASAY	
		A*0204	CLTVPASAYV		
		B*0702		VPASAYVGDL	
2	VSQLFTFSPRYPGHITGHRM	A*0101		VSQLFTFSPRY	
		B*0702		SPRYPGHI	similar
		B*0702		SPRYPGHIT	similar
		B*15	SQLFTFSPRY		
		B*2705	SQLFTFSPRY		
		B*2705	PRYPGHITGH		similar1
		B*2705	PRYPGHIT		similar1
3	AWDMMMNWSPVCGPVYCFTP	A*0201		MMNWSPVCGPV	
		B*0702		GPVYCFTPSPV	
4	SPVVVGTTDRTDVFLLNNTR	A*0101		TTDRTDVF	similar
		A*0101		TTDRTDVFLL	similar
		B*2705	DRTDVFLL		
5	LGNWFGCTWMCPTDCFRKHP	none			
6	CLVDYPYRLWDRDRSELSPL	A*0203	RLWDRDRSEL		
		B*2705		YRLWDRDRSEL	similar
		B*2705	RLWDRDRSEL		similar
		B*2705	YRLWDRDR		similar1
		B*2705	YRLWDRDRS		similar1
7	QYLYGVGSSVARVCACLWMM	none			
8	QAEAALENLVKGWRLLAPIT	B*4001		LENLVQGWRL	similar
		B*4001		LENLVQGWRLL	similar
9	VEGEVQIVSTQSFLATCING	B*2705	VQIVSTQSFL		similar
		B*2705	VQIVSTQSF		similar
10	VCWTVYHGAGPCTCGSSDLY	none			
11	GAYMSKAHGVTGVRTITTGS	none			
12	TVPHPNIEEVNAVAYYRGLD	A*0101		NIEEVNAVAYY	
13	QRRGRTGRGRYRFVTPGERP	B*2705		GRYRFVTPGER	similar
		B*2705	GRYRFVTPGE		similar
		B*2705	GRYRFVTPG		similar
		B*2705	GRYRFVTP		similar
		B*2705	RRGRTGRGRY		similar1
		B*2705	GRTGRGRYRF		similar1
		B*2705	GRTGRGRYR		similar1
		B*2705	GRTGRGRY		similar1
		B*2705	GRGRYRFV		similar2
		B*2705	GRGRYRFVT		similar2
14	MWKCLIRLKPKYIMTCMSAD	B*0702		KPKYIMTCM	
		B*0801		RLKPKYIM	similar
		B*0801		RLKPKYIMTCM	similar
		B*0801		LIRLKPKYI	similar1
		B*0801		LIRLKPKYIM	similar1
		B*2705	IRLKPKYI		similar2
		B*2705	IRLKPKYIM		similar2
		B*2705	IRLKPKYIMT		similar2
15	VIVGRIVLSGFWAKHMWNFI	A*0301		VLSGFWAK	similar
		A*0301		RIVLSGFWAK	similar
		A*2402		GFWAKHMWNF	
		B*2705		GRIVLSGFWAK	similar1
		B*2705	GRIVLSGF		similar1
		B*2705	GRIVLSGFWA		similar1
		B*2705	GRIVLSGFW		similar1

junction ID	Junction sequence	HLA* type	BIMAS	NetMHC	
16	ASVTSPLTTQGSIGLGKVLV	B*0702		SPLTTQGSI	similar
		B*0702		SPLTTQGSIGL	similar
		B*2705	TQGSIGLGK		
17	AAARVTQILSSWLRDIWDWI	B*2705		ARVTQILSSWL	similar
		B*2705	ARVTQILSSW		similar
		B*2705	ARVTQILSS		similar
		B*2705	TQILSSWL		similar1
		B*2705	TQILSSWLR		similar1
18	VLSDFKTWLKCPCQVPSPEF	none			
19	TELDGVRLHRRRLARGSPPS	B*07	GVRLHRRRL		
		B*2705	VRLHRRRLAR		similar
		B*2705	VRLHRRRL	VRLHRRRL	similar
		B*2705	RRRLARGS		
20	SLKATCTTNHSDAESYSSMP	A*0101		TTNHSDAESY	similar
		A*0101		CTTNHSDAESY	similar
		A*0101		HSDAESYSSM	
21	GSWSTVSSEALSNSLLRHHN	B*4001		SEALSNSL	similar
		B*4001		SEALSNSLL	similar
22	MVYATTSRSALTPPHSAKSK	A*0301		ALTPPHSAK	similar
		A*0301		RSALTPPHSAK	similar
		A*24	VYATTSRSAL		
		B*0702		YATTSRSAL	similar1
		B*0702		MVYATTSRSAL	similar1
		B*2705	SRSALTPPHS		similar2
		B*2705	SRSALTPPH		similar2
23	GYGAKDVRCHKPARLIVFPD	B*07	DVRCHKPARL		
		B*08	DVRCHKPARL		
		B*2705	VRCHKPARL		similar
		B*2705	VRCHKPAR		similar
		B*2705	VRCHKPARLI		similar
24	CEKMALYDVVPMGFSYDTRC	A*0201		ALYDVVPM	
		A*24	LYDVVPMGF	LYDVVPMGF	
		B*1501		KMALYDVVPM	
25	TVTESDIRTESLTERLYVGG	A*01	RTESLTERLY	RTESLTERLY	
		B*2705	IRTESLTERL		similar
		B*2705	IRTESLTER		similar
		B*4001		TESDIRTESL	
26	ASAACRAAKLCTMLVCGDDL	B*2705	CRAAKLCTML		similar
		B*2705	CRAAKLCTM	CRAAKLCTM	similar
		B*2705	CRAAKLCT		similar
27	CESAGTQEDASLRAFTEAMT	B*2705	TQEDASLR		
28	APTLWARMILEPLDLPQIIQ	B*07	WARMILEPL		
		B*0801		WARMILEPL	
		B*2705	ARMILEPL	ARMILEPL	similar
		B*2705	ARMILEPLDL		similar
		B*2705	RMILEPLDL		similar
29	WRHRARSVRAGGRAAICGKY	B*0702		SVRAGGRAA	
		B*0702		SVRAGGRAAI	
		B*2705	ARSVRAGGR		similar
		B*2705	ARSVRAGGRA		similar
		B*2705	VRAGGRAAI		similar1
		B*2705	VRAGGRAAIC		similar1
		B*2705	VRAGGRAA		similar1

8.2.3 BIMAS – Syfpeithi

To compare results from the BIMAS analysis with those from the Syfpeithi analysis, I first determined HLA types shared by both servers (marked in orange in Table 8-7).

Table 8-7: MHC class-I types shared between Syfpeithi and BIMAS servers.

Syfpeithi HLA-A	mers	BIMAS HLA-A	mers
A*01	8		
A*01	9	A*01	9
A*01	10	A*01	10
A*01	11		
A*0201	9	A*0201	9
A*0201	10	A*0201	10
A*0202	9		
A*0202	10		
A*0203	8		
A*0203	9		
A*0203	10		
A*0204	9		
A*0205	10		
A*0206	9		
A*0207	9		
A*03	9	A*03	9
A*03	10	A*03	10
A*1101	9	A*1101	9
A*1101	10	A*1101	10
A*1101	11		
A*23	9		
A*2402	9	A*24	9
A*2402	10	A*24	10
A*25	9		
A*26	9		
A*26	10		
A*2602	9		
A*29	9		
A*3101	9	A*3101	9
		A*3101	10
		A*3302	9
		A*3302	10
A*6801	9	A*68	9
A*6801	10	A*68	10
A*6801	11		
A*6801	15		
A*6802	9		

Syfpeithi HLA-B	mers	BIMAS HLA-B	mers
B*0702	9	B*07	9
B*0702	10	**B*07**	10
B*08	8		
B*08	9	B*08	9
		B*08	10
B*13	9		
B*13	10		
B*1402	8	B*14	9
B*1402	9	B*14	10
B*1501	9	B*15	9
B*1501	10	B*15	10
B*1509	9		
B*1510	9		
B*1516	9		
B*18	8		
B*18	9		
B*1801	9		
B*2702	9	B*2702	9
		B*2702	10
B*2704	9		
		B*2705	8
B*2705	9	B*2705	9
B*2705	10	B*2705	10
B*2706	9		
B*2709	9		
		B*3501	8
B*3501	9	B*3501	9
B*3501	10	B*3501	10
B*37	8		
B*37	9	B*3701	9
		B*3701	10
B*3801	8		
B*3801	9	B*3801	9
B*3801	10	B*3801	10
B*3901	9		
B*3901	10		
B*3902	9		
B*3902	10		
B*4001	9	B*40	9
B*4001	10	B*40	10

Syfpeithi HLA-B	mers	BIMAS HLA-B	mers
B*4002	9		
B*4002	10		
B*41	9		
B*4101	9		
B*4101	10		
B*4105	9		
B*44	9		
B*4402	9		
B*4402	10		
B*4402	11		
		B*4403	9
		B*4403	10
B*45	9		
B*4501	9		
B*4501	10		
B*4701	9		
B*4701	10		
B*4901	9		
B*4901	10		
B*5001	9		
B*5001	10		
B*51	10		
B*5101	8	B*5101	8
B*5101	9	B*5101	9
		B*5101	10
		B*5102	9
		B*5102	10
		B*5103	9
		B*5103	10
		B*5201	8
		B*5201	9
		B*5201	10
B*5301	9		
B*5301	10		
B*57	9		
B*5701	9		
B*5701	10		
B*5801	9		
B*5802	9		
B*5802	10		
		B*60	9
		B*60	10
		B*61	9
		B*61	10

MHC class-I types avaiable for prediction analysis on the Syfpeithi and BIMAS server. The HLA types which were shared between the two servers and for which predicted epitopes were compared, are marked in orange.

Of 148 artificial strong binders predicted by the Syfpeithi server, 52 (35%) were excluded from the Syfpeithi prediction dataset since HLA types did not match those used in the BIMAS analysis (Table 8-8, left). Likewise, of artificial 103 epitopes predicted for the BIMAS analysis, 8 (8%) were excluded (Table 8-8, bottom).

Table 8-8: Excluded predicted strong binders in junction areas in comparative analysis, because of non-matching HLA types in comparative analysis between Syfpeithi and BIMAS servers.

HLA type	mers	Junction	Predicted peptide	Syfpeithi score
A*26	9	3	PVCGPVYCF	22
A*26	10	4	TTDRTDVFLL	21
A*26	10	9	EVQIVSTQSF	29
A*26	9	9	EVQIVSTQS	21
A*26	9	12	EEVNAVAYY	28
A*26	10	15	IVGRIVLSGF	23
A*26	10	23	DVRCHKPARL	26
A*26	10	24	DVVPMGFSYD	24
A*26	9	24	DVVPMGFSY	35
B*13	10	6	RLWDRDRSEL	21
B*13	10	9	VQIVSTQSFL	21
B*13	9	9	TQSFLATCI	22
B*13	10	16	TQGSIGLGKV	22
B*1516	9	1	LTVPASAYV	28
B*1516	9	3	WSPVCGPVY	22
B*18	9	12	IEEVNAVAY	22
B*18	8	12	EEVNAVAY	26
B*18	9	25	TESLTERLY	21
B*18	8	25	TESLTERL	21
B*18	9	27	QEDASLRAF	22
B*3901	9	23	CHKPARLIV	21
B*3901	9	23	VRCHKPARL	22
B*3901	10	25	IRTESLTERL	21
B*3901	10	26	CRAAKLCTML	21
B*3901	10	28	ARMILEPLDL	26
B*4402	11	8	LENLVKGWRLL	23
B*4402	10	8	LENLVKGWRL	21
B*4402	11	9	GEVQIVSTQSF	21
B*4402	10	12	IEEVNAVAYY	21
B*4402	9	12	EEVNAVAYY	24
B*4402	9	12	IEEVNAVAY	25
B*4402	9	21	SEALSNSLL	23
B*4402	10	25	TESDIRTESL	23
B*4402	9	25	TESLTERLY	24
B*4402	9	27	QEDASLRAF	25
B*4402	10	28	LEPLDLPQII	21
B*4402	9	28	LEPLDLPQI	22
B*4901	10	8	LENLVKGWRL	22
B*4901	10	25	TESLTERLYV	22
B*5802	10	8	AALENLVKGW	21
B*5802	9	13	RTGRGRYRF	22
B*5802	9	15	LSGFWAKHM	21
B*5802	10	17	SSWLRDIWDW	21
B*5802	9	17	RVTQILSSW	21
B*5802	9	17	LSSWLRDIW	21
B*5802	10	21	VSSEALSNSL	21
B*5802	10	21	SSEALSNSLL	21
B*5802	9	21	SSEALSNSL	21
B*5802	10	25	RTESLTERLY	27
B*5802	9	25	RTESLTERL	23
B*5802	9	26	RAAKLCTML	23
B*5802	10	27	ASLRAFTEAM	21

HLA type	mers	Junction	Predicted peptide	BIMAS score
B*15	10	2	SQLFTFSPRY	160
B*4403	10	12	IEEVNAVAYY	540
B*4403	9	12	IEEVNAVAY	180
B*4403	9	12	EEVNAVAYY	1080
B*4403	9	25	TESLTERLY	120
B*60	9	8	LENLVKGWRL	176
B*60	9	21	SEALSNSLL	320
B*60	9	25	TESDIRTESL	320

Strong binders predicted by the Syfpeithi server (left) and the BIMAS server (right) that were excluded from the comparative analysis between Syfpeithi and BIMAS servers due to non-matching HLA types between the two servers.

Of the remaining 191 [96 (Syfpeithi) and 95 (BIMAS)], epitopes, 24 were identical for both prediction analyses, which estimates at 25% for both analyses (Table 8-9). A further 2 epitopes were classified as “similar”, increasing the percentage of matching epitopes to 26%.

Table 8-9: Comparison of MHC class-I prediction results using Syfpeithi and BIMAS servers (table on next pages).

Comparison of epitope prediction results for Syfpeithi and BIMAS servers. Epitopes predicted by both servers are marked in red. Similar epitopes (epitopes of different length predicted for the same restricting HLA type) were marked in grey.

Junction	Sequence junction	HLA type	Sypfeithi Score	BIMAS score	
1	LSCLTVPASAYVGDLCGSVF	A*0201	CLTVPASAYV	CLTVPASAYV	
		A*03	CLTVPASAY		
		B*0702	VPASAYVGDL		
		B*1501	CLTVPASAY		
		B*5101		VPASAYVGDL	similar
		B*5102		VPASAYVGDL	similar
		Cw*0401		VPASAYVGDL	
2	VSQLFTFSPRYPGHITGHRM	A*01	VSQLFTFSPRY		
		B*0702	SPRYPGHIT		
		B*08	SPRYPGHI		
		B*1501	QLFTFSPRY		
		B*2705	PRYPGHITGH	PRYPGHITGH	similar
		B*2705		PRYPGHIT	similar
		B*2705		SQLFTFSPRY	
		B*5101	SPRYPGHI	SPRYPGHI	
3	AWDMMMNWSPVCGPVYCFTP none				
4	SPVVVGTTDRTDVFLLNNTR	B*1402	DRTDVFLLN		similar
		B*1402	DRTDVFLL		similar
		B*2705		DRTDVFLL	
		B*37	TDRTDVFLL	TDRTDVFLL	similar1
		B*37	TDRTDVFL		similar1
5	LGNWFGCTWMCPTDCFRKHP	none			
6	CLVDYPYRLWDRDRSELSPL	A*0201	RLWDRDRSEL	RLWDRDRSEL	
		B*08	WDRDRSEL		
		B*2705		YRLWDRDR	similar
		B*2705		YRLWDRDRS	similar
		B*2705		RLWDRDRSEL	similar
		B*37	WDRDRSEL		
7	QYLYGVGSSVARVCACLWMM	A*0201	GVGSSVARV		
		A*03	YLYGVGSSVA		
		B*08	VARVCACL		
		B*5102		YGVGSSVARV	
8	QAEAALENLVKGWRLLAPIT	A*0201	NLVKGWRLL		
		A*03	AEAALENLVK		
		B*08	ENLVKGWRL		
		B*4001	LENLVKGWRL		
9	VEGEVQIVSTQSFLATCING	B*1501	VQIVSTQSF		
		B*2705		VQIVSTQSFL	similar
		B*2705		VQIVSTQSF	similar
10	VCWTVYHGAGPCTCGSSDLY	B*5101		GPCTCGSSDL	
		B*5102		GPCTCGSSDL	
11	GAYMSKAHGVTGVRTITTGS	A*0201	SKAHGVTGV		
		B*5101	HGVTGVRTI	HGVTGVRTI	similar
		B*5101	KAHGVTGV		
		B*5102		HGVTGVRTI	similar
12	TVPHPNIEEVNAVAYYRGLD	A*01	NIEEVNAVAYY		similar
		A*01	NIEEVNAVAY		similar
		A*01	IEEVNAVAYY		similar1
		A*01	IEEVNAVAY		similar1
		A*03	NIEEVNAVAY		
		A*68	EVNAVAYYR	EVNAVAYYR	
		B*5101		HPNIEEVNAV	similar2
		B*5102		HPNIEEVNAV	similar2
13	QRRGRTGRGRYRFVTPGERP	B*2702		RRGRTGRGRY	similar
		B*2702		GRTGRGRYRF	similar
		B*2705	RRGRTGRGRY	RRGRTGRGRY	similar
		B*2705	GRTGRGRYRF	GRTGRGRYRF	similar
		B*2705	GRTGRGRYR	GRTGRGRYR	similar
		B*2705		GRTGRGRY	similar
		B*2705		GRGRYRFV	similar1
		B*2705		GRGRYRFVT	similar1
		B*2705		GRYRFVTPG	similar2
		B*2705		GRYRFVTPGE	similar2
		B*2705		GRYRFVTP	similar2
14	MWKCLIRLKPKYIMTCMSAD	A*03	RLKPKYIMTC		
		A*03	CLIRLKPKY		
		B*08	RLKPKYIMT		similar
		B*08	RLKPKYIM		similar
		B*08	LIRLKPKYI		similar1
		B*08	LIRLKPKY		similar1
		B*1402	IRLKPKYI		
		B*1501	CLIRLKPKY		
		B*2705	IRLKPKYIM	IRLKPKYIM	similar2
		B*2705		IRLKPKYI	similar2
		B*2705		IRLKPKYIMT	similar2
		B*2709	IRLKPKYIM		
		B*3501		KPKYIMTCM	
15	VIVGRIVLSGFWAKHMWNFI	A*03	IVLSGFWAKH		
		A*03	IVGRIVLSGF		
		A*03	RIVLSGFWAK		
		A*03	IVLSGFWAK		
		A*1101	IVLSGFWAK		
		A*68		IVLSGFWAK	
		B*2702		GRIVLSGFW	
		B*2705		GRIVLSGF	
		B*2705		GRIVLSGFWA	
		B*2705		GRIVLSGFW	
		Cw*0401		GFWAKHMWNF	

Junction	Sequence junction	HLA type	Syfpeithi Score	BIMAS score	
16	ASVTSPLTTQGSIGLGKVLV	A*1101	TTQGSIGLGK		
		B*2705		TQGSIGLGK	
		B*5101	SPLTTQGSI	SPLTTQGSI	similar
		B*5102		SPLTTQGSI	similar1
17	AAARVTQILSSWLRDIWDWI	A*0201	QILSSWLRDI		similar
		A*0201	ILSSWLRDI		similar
		B*2702		ARVTQILSSW	similar1
		B*2705		ARVTQILSSW	similar1
		B*2705		ARVTQILSS	similar1
		B*2705		TQILSSWL	similar2
		B*2705		TQILSSWLR	similar2
18	VLSDFKTWLKCPCQVPSPEF	none			
19	TELDGVRLHRRRLARGSPPS	A*03	RLHRRRLAR		
		B*07		GVRLHRRRL	
		B*14		DGVRLHRRRL	similar
		B*14		VRLHRRRLA	similar
		B*1402	VRLHRRRL		similar
		B*2705	VRLHRRRLAR	VRLHRRRLAR	similar1
		B*2705		VRLHRRRL	similar1
		B*2705		RRRLARGS	
		B*5102		DGVRLHRRRL	
20	SLKATCTTNHSDAESYSSMP	A*01	CTTNHSDAESY		similar
		A*01	TTNHSDAESY		similar
21	GSWSTVSSEALSNSLLRHHN	A*03	ALSNSLLRH		
		B*37	SEALSNSLL		similar
		B*37	SEALSNSL		similar
		B*4001	SEALSNSLL		
		B*5101		EALSNSLL	
22	MVYATTSRSALTPPHSAKSK	A*03	ALTPPHSAK		
		A*24	VYATTSRSAL	VYATTSRSAL	
		B*2705		SRSALTPPHS	similar
		B*2705		SRSALTPPH	similar
		B*5102		YATTSRSAL	
		Cw*0401		VYATTSRSAL	
23	GYGAKDVRCHKPARLIVFPD	A*68	DVRCHKPAR	DVRCHKPAR	
		B*07		DVRCHKPARL	
		B*08	VRCHKPARL		similar
		B*08		DVRCHKPARL	similar
		B*2705	VRCHKPARL	VRCHKPARL	similar1
		B*2705		VRCHKPAR	similar1
		B*2705		VRCHKPARLI	similar1
		B*2709	VRCHKPARL		
24	CEKMALYDVVPMGFSYDTRC	A*01	LYDVVPMGFSY		
		A*0201	KMALYDVVPM		
		A*03	ALYDVVPMGF		
		A*03	DVVPMGFSY		
		A*24	LYDVVPMGF	LYDVVPMGF	
		B*1501	ALYDVVPMGF		
		B*37	YDVVPMGF		
		Cw*0401		LYDVVPMGF	
25	TVTESDIRTESLTERLYVGG	A*01	RTESLTERLY	RTESLTERLY	
		A*6801	DIRTESLTER		
		B*2705	IRTESLTERL	IRTESLTERL	
		B*2705	IRTESLTER	IRTESLTER	
		B*37	TESLTERL		
		B*4001	TESDIRTESL		
26	ASAACRAAKLCTMLVCGDDL	B*08	AAKLCTML		
		B*2705	CRAAKLCTML	CRAAKLCTML	similar
		B*2705	CRAAKLCTM	CRAAKLCTM	similar
		B*2705		CRAAKLCT	similar
		B*5101		AAKLCTMLV	similar1
		B*5102		RAAKLCTMLV	similar1
		B*5103		RAAKLCTMLV	similar1
		B*5103		AAKLCTMLV	similar1
27	CESAGTQEDASLRAFTEAMT	A*1101	GTQEDASLR		
		B*2705		TQEDASLR	
28	APTLWARMILEPLDLPQIIQ	A*0201	ILEPLDLPQI		
		A*0201	RMILEPLDL		
		B*07		WARMILEPL	
		B*1402	MILEPLDL		
		B*2705	ARMILEPLDL	ARMILEPLDL	similar
		B*2705		ARMILEPL	similar
		B*2705		RMILEPLDL	similar
29	WRHRARSVRAGGRAAICGKY	A*03	SVRAGGRAAI		similar
		A*03	SVRAGGRAA		similar
		B*2705	VRAGGRAAI	VRAGGRAAI	
		B*2705	ARSVRAGGR	ARSVRAGGR	
		B*2705		ARSVRAGGRA	similar1
		B*2705		VRAGGRAAIC	similar1
		B*2705		VRAGGRAA	similar1
		B*5101		RAGGRAAI	

8.3 Linker design at junction regions between conserved segments to avoid predicted strong binders

Table 8-10: Prediction of artificial epitopes in junction regions by three different servers and modification through linkers (table depicted on the next 3 pages).

A prediction analysis was run using three independent epitope prediction servers before (left) and after (right) linker insertion at junction sites. Epitope sequence, predicted HLA type and prediction scores are shown are shown before and after linker insertion. Results for the different servers are depicted both before and after linker insertion, for NetMHC (left column), Syfpeithi (middle column) and BIMAS (right column. Scores of <0.5 (NetMHC), >20 (Syfpeithi) and >100 (BIMAS) were defined as strong binding peptides. Junction regions that were not modified are marked in light blue. Predicted epitopes are colour coded: Putative epitopes included in final immunogen are marked in light red (light red bar for junction regions that were not modified), newly predicted epitopes after linker insertion in purple and epitopes with a higher prediction score after linker insertion in bright red.

ID	Junction sequence	HLA	NetMHC	Score	Sylfpeithi	Score	BIMAS	Score	Linkers	HLA	NetMHC	Score	Syfpeithi	Score	BIMAS	Score
1	LSCLTVPASAYVGDLCGSVF	A*0101	LTVPASAY	0.4					LSCLTVPASAggsgYVGDLCGSVF	A*01:01	ASA**GGSG**Y	0.5				
	LSCLTVPASAYVGDLCGSVF	A*0201			CLTVPASAYV	21	CLTVPASAYV	383								
		A*03			CLTVPASAY	23										
		A*2601	LTVPASAY	0.3						A*26:01	TVPASA**GGSG**Y	0.3				
		B*0702	VPASAYVGDL	0.5	VPASAYVGDL	22				B*07:02	VPASA**GGSG**YV	0.5				
		B*1501	LTVPASAY	0.5												
		B*1501			CLTVPASAY	21										
		B*1516			LTVPASAYV	28				B*1516			ASA**GGSG**YV	22		
										B*3501			VPASA**GGSG**Y	20		
										B*5301			VPASA**GGSG**Y	21		
		B*5101					VPASAYVGDL	130		B*5101					SA**GGSG**YV	121
		B*5102					VPASAYVGDL	100								
		Cw*0401					VPASAYVGDL	160								
2	VSQLFTFSPRYPGHITGHRM	A*01	VSQLFTFSPRY	0.18	VSQLFTFSPRY	21			VSQLFTFSPRpYPGHITGHRM	A*01:01	FTFS**P**RPY	0.4				
										A*26:01	FTFS**P**RPY	0.25				
		B*0702	SPRYPGHI	0.15						B*07:02	SPR**P**YPGHI	0.3				
		B*0702	SPRYPGHIT	0.25	SPRYPGHIT	22				B*07:02	SPR**P**YPGHIT	0.5				
										B*07:02	R**P**YPGHIT	0.5				
		B*08			SPRYPGHI	21				B*08			S**P**RPYPGHI	20		
	VSQLFTFSPRYPGHITGHRM	B*15					SQLFTFSPRY	160	VSQLFTFSPRpYPGHITGHRM							
		B*1501			QLFTFSPRY	21				B*1501			QLFTF**S**PRPY	22		
										B*15:01	SQLFTFS**P**RPY	0.4				
		B*2705			PRYPGHITGH	26	PRYPGHITGH	100		B*2705			R**P**YPGHITGH	20		
		B*2705					SQLFTFSPRY	100								
		B*2705					PRYPGHIT	100								
		B*5101			SPRYPGHI	24	SPRYPGHI	400		B*5101			SP**R**PYPGHI	23	SPR**P**YPGHI	400
		B*5101								B*5102					SPR**P**YPGHI	200
										B*5802			FS**P**RPYPGHI	20		
3	AWDMMMNWSPVCGPVYCFTP	A*0201	MMNWSPVCGPV	0.4												
	AWDMMMNWSPVCGPVYCFTP	A*26			PVCGPVYCF	22										
		B*0702	GPVYCFTPSPV	0.4												
		B*1516			WSPVCGPVY	22										
4	SPVVVGTTDRTDVFLLNNTR	A*0101	TTDRTDVF	0.4												
		A*0101	TTDRTDVFLL	0.4												
	SPVVVGTTDRTDVFLLNNTR	A*26			TTDRTDVFLL	21										
		B*1402			DRTDVFLLN	21										
		B*1402			DRTDVFLL	23										
	SPVVVGTTDRTDVFLLNNTR	B*2705					DRTDVFLL	200								
		B*37			TDRTDVFLL	21	TDRTDVFLL	200								
		B*37			TDRTDVFL	21										
5	LGNWFGCTWMCPTDCFRKHP	none														
6	CLVDYPYRLWDRDRSELSPL	A*0201			RLWDRDRSEL	24	RLWDRDRSEL	727	CLVDYPYRLWgDRDRSELSPL							
	CLVDYPYRLWDRDRSELSPL	A*0203														
		B*08			WDRDRSEL	24				B*08			**G**DRDRSEL	25		
		B*13			RLWDRDRSEL	21										
	CLVDYPYRLWDRDRSELSPL	B*2705	YRLWDRDRSEL	0.5												
		B*2705					YRLWDRDR	1000		B*2705					YRLW**G**DRDR	300
		B*2705					YRLWDRDRS	200		B*2705					YRLW**G**DRDRS	200
		B*2705					RLWDRDRSEL	450		B*2705					RLW**G**DRDR	225
										B*2705			YRLW**G**DRDR	25		
		B*37			WDRDRSEL	21				B*37			**G**DRDRSEL	21		
		B*3901	YRLWDRDRSEL	0.25												
7	QYLYGVGSSVARVCACLWMM	A*0201			GVGSSVARV	21										
		A*03			YLYGVGSSVA	23										
		B*08			VARVCACL	22										
	QYLYGVGSSVARVCACLWMM	B*5102					YGVGSSVARV	528								
	QYLYGVGSSVARVCACLWMM	B*5801	SSVARVCACLW	0.08												
		B*5801	SVARVCACLW	0.25												
		B*5801	VARVCACLW	0.15												
8	QAEAALENLVKGWRLLAPIT	A*0201			NLVKGWRLL	22										
		A*03			AEAALENLVK	23										
		B*08			ENLVKGWRL	23										
	QAEAALENLVKGWRLLAPIT	B*4001	LENLVQGWRL	0.5	LENLVKGWRL	22										
		B*4001	LENLVQGWRLL	0.3	LENLVKGWRLL	23										
		B*4402			LENLVKGWRL	21										
		B*4901			LENLVKGWRL	22										
		B*58	AALENLVKGW	0.18	AALENLVKGW	21										
	QAEAALENLVKGWRLLAPIT	B*60					LENLVKGWRL	176								
9	VEGEVQIVSTQSFLATCING	A*26			EVQIVSTQSF	29										
		A*26			EVQIVSTQS	21										
		B*13			VQIVSTQSFL	21										
		B*13			TQSFLATCI	22										
	VEGEVQIVSTQSFLATCING	B*1501	VQIVSTQSF	0.03	VQIVSTQSF	22										
	VEGEVQIVSTQSFLATCING	B*2705					VQIVSTQSFL	200								
		B*2705					VQIVSTQSF	100								
		B*4402			GEVQIVSTQSF	21										
10	VCWTVYHGAGPCTCGSSDLY	B*5101					GPCTCGSSDL	110								
		B*5102					GPCTCGSSDL	121								
11	GAYMSKAHGVTGVRTITTGS	A*0201			SKAHGVTGV	21			GAYMSKAHGVsgTGVRTITTGS	A*0201			GV**SG**TGVRTI	21		
										B*1516			V**SG**TGVRTI	25		
		B*5101			HGVTGVRTI	22	HGVTGVRTI	106		B*5101			**SG**TGVRTI	21	KAHGV**SG**TGV	110
		B*5101			KAHGVTGV	22										
		B*5102					HGVTGVRTI	351		B*5802			V**SG**TGVRTI	27	KAHGV**SG**TGV	110
										B*5802			GV**SG**TGVRTI	23		
12	TVPHPNIEEVNAVAYYRGLD	A*01	NIEEVNAVAYY	0.4	NIEEVNAVAYY	23			TVPHPNIEEVgpgNAVAYYRGLD							
		A*01			NIEEVNAVAY	23										
		A*01			IEEVNAVAYY	24										
		A*01			IEEVNAVAY	23										
		A*03			NIEEVNAVAY	23				A*03			EV**GPG**NAVAY	27		
										A*03			EV**GPG**NAVA	22		
		A*26	EVNAVAYY	0.05	EEVNAVAYY	28				A*26:01	EV**GPG**NAVAYY	0.12	EV**GPG**NAVAY	32		
										A*26:01	EV**GPG**NAVAY	0.3				
										A*26			EV**GPG**NAVA	21		
		A*6801			EVNAVAYYR	22	EVNAVAYYR	600								
										B*1501			EV**GPG**NAVAY	23		
		B*18			IEEVNAVAY	22										
		B*18			EEVNAVAY	26										
										B*3501			**GPG**NAVAYY	20		
										B*4101			IEEV**GPG**NA	21		
										B*4101			EEV**GPG**NAVA	22		
		B*4402			IEEVNAVAYY	21				B*4402			EEV**GPG**NAVAY	27		
		B*4402			EEVNAVAYY	24										
		B*4402			IEEVNAVAY	25										
		B*4403					IEEVNAVAYY	540								
		B*4403					IEEVNAVAY	180								
		B*4403					EEVNAVAYY	1080								
										B*4501			IEEV**GPG**NA	21		
										B*4501			EEV**GPG**NAVA	24		
										B*4901			IEEV**GPG**NAV	20		
										B*5001			IEEV**GPG**NA	20		
										B*5001			EEV**GPG**NAVA	21		
		B*5101					HPNIEEVNAV	220								
		B*5102					HPNIEEVNAV	200								
										B*5301			**GPG**NAVAYY	20		

13	QRRGRTGRGRYRFVTPGERP							
	QRRGRTGRGRYRFVTPGERP	B*2702					RRGRTGRGRY	600
		B*2702					GRTGRGRYRF	200
	QRRGRTGRGRYRFVTPGERP	B*2705	GRYRFVTPGER	0.2				
		B*2705			RRGRTGRGRY	24	RRGRTGRGRY	3000
		B*2705			GRTGRGRYRF	28	GRTGRGRYRF	1000
		B*2705			GRTGRGRYR	25	GRTGRGRYR	300
		B*2705					GRTGRGRY	300
		B*2705					GRGRYRFV	600
		B*2705					GRGRYRFVT	200
		B*2705					GRYRFVTPG	100
		B*2705					GRYRFVTPGE	100
		B*2705					GRYRFVTP	100
		B*5802			RTGRGRYRF	22		
14	MWKCLIRLKPKYIMTCMSAD	A*03			RLKPKYIMTC	23		
		A*03			CLIRLKPKY	22		
	MWKCLIRLKPKYIMTCMSAD	B*0702	KPKYIMTCM	0.5				
		B*08	LIRLKPKYI	0.5	LIRLKPKYI	28		
		B*08	RLKPKYIM	0.3	RLKPKYIM	26		
		B*08			RLKPKYIMT	26		
		B*08			LIRLKPKY	21		
		B*0801	LIRLKPKYIM	0.4				
		B*0801	RLKPKYIMTCM	0.4				
		B*1402			IRLKPKYI	23		
		B*1501			CLIRLKPKY	22		
		B*2705			IRLKPKYIM	26	IRLKPKYIM	600
	MWKCLIRLKPKYIMTCMSAD	B*2705					IRLKPKYI	600
		B*2705					IRLKPKYIMT	200
		B*2709			IRLKPKYIM	22		
		B*3501					KPKYIMTCM	240
15	VIVGRIVLSGFWAKHMWNFI							
		A*03	RIVLSGFWAK	0.4	RIVLSGFWAK	24		
	VIVGRIVLSGFWAKHMWNFI	A*03			IVLSGFWAKH	21		
		A*03			IVGRIVLSGF	22		
		A*03			IVLSGFWAK	27		
		A*0301	VLSGFWAK	0.18				
		A*1101			IVLSGFWAK	23		
	VIVGRIVLSGFWAKHMWNFI	A*2402	GFWAKHMWNF	0.4				
		A*26			IVGRIVLSGF	23		
		A*68					IVLSGFWAK	240
		B*2702					GRIVLSGFW	100
		B*2705	GRIVLSGFWAK	0.18				
		B*2705					GRIVLSGF	1000
		B*2705					GRIVLSGFWA	200
		B*2705					GRIVLSGFW	200
		B*5801	IVGRIVLSGFW	0.5				
		B*5801	LSGFWAKHMW	0.15				
		B*5801	RIVLSGFW	0.5				
		B*5801	SGFWAKHMW	0.5				
		B*5802			LSGFWAKHM	21		
		Cw*0401					GFWAKHMWNF	132
16	ASVTSPLTTQGSIGLGKVLV	A*1101			TTQGSIGLGK	25		
	ASVTSPLTTQGSIGLGKVLV	B*0702	SPLTTQGSI	0.3				
		B*0702	SPLTTQGSIGL	0.3				
		B*13			TQGSIGLGKV	22		
	ASVTSPLTTQGSIGLGKVLV	B*2705					TQGSIGLGK	200
		B*5101			SPLTTQGSI	24	SPLTTQGSI	484
		B*5102					SPLTTQGSI	1597
17	AAARVTQILSSWLRDIWDWI	A*0201			QILSSWLRDI	22		
		A*0201			ILSSWLRDI	24		
	AAARVTQILSSWLRDIWDWI	B*2702					ARVTQILSSW	100
		B*2705	ARVTQILSSWL	0.5				
		B*2705					ARVTQILSSW	200
		B*2705					ARVTQILSS	200
		B*2705					TQILSSWL	200
		B*2705					TQILSSWLR	100
		B*58	LSSWLRDIW	0.08	LSSWLRDIW	21		
		B*58	RVTQILSSW	0.13	RVTQILSSW	21		
		B*58	SSWLRDIWDW	0.2	SSWLRDIWDW	21		
	AAARVTQILSSWLRDIWDWI	B*5801	AARVTQILSSW	0.4				
		B*5801	LSSWLRDIWDW	0.08				
		B*5801	SSWLRDIW	0.18				
		B*5801	VTQILSSW	0.15				
18	VLSDFKTWLKCPCQVPSPEF	none						
19	TELDGVRLHRRRLARGSPPS	A*03			RLHRRRLAR	28		
	TELDGVRLHRRRLARGSPPS	B*07					GVRLHRRRL	300
		B*14					DGVRLHRRRL	135
		B*14					VRLHRRRLA	250
		B*1402			VRLHRRRL	25		
	TELDGVRLHRRRLARGSPPS	B*2705	VRLHRRRL	0.5			VRLHRRRL	600
		B*2705			VRLHRRRLAR	23	VRLHRRRLAR	1000
		B*2705					RRRLARGS	600
		B*5102					DGVRLHRRRL	145
20	SLKATCTTNHSDAESYSSMP	A*01	CTTNHSDAESY	0.15	CTTNHSDAESY	21		
		A*01	TTNHSDAESY	0.25	TTNHSDAESY	21		
		A*0101	HSDAESYSSM	0.5				
		B*3901	NHSDAESYSSM	0.2				
21	GSWSTVSSEALSNSLLRHHN	A*03			ALSNSLLRH	21		
		B*37			SEALSNSLL	25		
		B*37			SEALSNSL	24		
	GSWSTVSSEALSNSLLRHHN	B*4001	SEALSNSL	0.4				
		B*4001	SEALSNSLL	0.4	SEALSNSLL	23		
		B*4402			SEALSNSLL	23		
	GSWSTVSSEALSNSLLRHHN	B*5101					EALSNSLL	100
		B*5802			VSSEALSNSL	21		
		B*5802			SSEALSNSLL	21		
		B*5802			SSEALSNSL	21		
		B*60					SEALSNSLL	320
22	MVYATTSRSALTPPHSAKSK	A*03	ALTPPHSAK	0.25	ALTPPHSAK	30		
		A*0301	RSALTPPHSAK	0.5				
	MVYATTSRSALTPPHSAKSK	A*24			VYATTSRSAL	22	VYATTSRSAL	200
		B*0702	MVYATTSRSAL	0.3				
		B*0702	YATTSRSAL	0.4				
		B*2705					SRSALTPPHS	200
		B*2705					SRSALTPPH	200
		B*5102					YATTSRSAL	110
		Cw*0401					VYATTSRSAL	200

13	QRRGRTGRGRpgggsgggYRFVTPGERP	B*1516			**GSGGG**YRFV	22		
		B*2705			**GGSGGG**YRF	20		
		B*2705					GRGR**PGGGS**	200
		B*3501			R**PGGGSGGG**Y	20		
		B*5301			R**PGGGSGGG**Y	20		
14	MWKCLIRLKPggkgpggKYIMTCMSAD	A*03			RLKP**GGKGP**	23		
		A*03			CLIRLKP**GGK**	28		
		A*03			LIRLKP**GGK**	25		
		B*08			LIRLKP**GGK**	21		
		B*08			LIRLKP**GG**	21		
		B*2705					IRLKP**GGK**	2000
		B*5101					**GGKGPGG**KYI	106
15	VIVGRIVLSGgpgFWAKHMWNFI	A*0201					VLSG**GPG**FWA	269
		A*03			RIVLSG**GPG**F	20		
		A*03			IVLSG**GPG**F	23		
		A*1101			LSG**GPG**FWAK	22		
		A*24:02	**G**FWAKHMWNF	0.4				
		A*26			IVLS**GGP**GF	20		
		B*5301			**G**PGFWAKHMW	20		
		B*58:01	IVLSGG**GPG**FW	0.4				
		B*58:01	LSGG**GPG**FW	0.17				
		Cw*0401					**G**FWAKHMWNF	132
16	ASVTSPLTTQpGSIGLGKVLV							
		B*07:02	SPLTTQ**P**GSI	0.5				
		B*2705					TQ**P**GSIGLGK	200
		B*2705					TQ**P**GSIGL	200
		B*5101					SPLTTQ**P**GSI	400
		B*5102					SPLTTQ**P**GSI	1452
		B*5102					Q**P**GSIGLGKV	440
		B*5101					Q**P**GSIGLGKV	242
19	TELDGVRLHRkkgpgsgpgpRRLARGSPPS	A*03			RLHR**KKGPGS**	20		
		A*03			ELDGVRLHR**K**	22		
		B*2705					HR**KKGPGS**	200
		B*5802			**GSGPGP**RRL	25		
21	GSWSTVSSEAgsgsLSNSLLRHHN							
		A*1101			**GS**LSNSLLR	29		
		B*37			SEA**GSGS**L	20		
		B*40:01	SEA**GSGS**L	0.15				
		B*5802			VSSEA**GSGS**L	21		
		B*5802			SSEA**GSGS**L	20		
		B*5802			**GSGS**LSNSLL	21		
		B*5802			**GSGS**LSNSL	21		
22	MVYATTSRSAgpLTPPHSAKSK	A*03			**P**LTPPHSAK	27		
		B*2705					SRSA**GP**LT	200

No.	Sequence	HLA	Peptide	Score	Peptide	Score	Peptide	Score
23	GYGAKDVRCHKPARLIVFPD							
		A*26			DVRCHKPARL	26		
		A*6801			DVRCHKPAR	21	DVRCHKPAR	600
		B*07					DVRCHKPARL	200
		B*08			VRCHKPARL	21		
		B*08					DVRCHKPARL	160
		B*2705			VRCHKPARL	24	VRCHKPARL	2000
		B*2705					VRCHKPAR	1000
		B*2705					VRCHKPARLI	180
		B*2709			VRCHKPARL	21		
		B*3901			CHKPARLIV	21		
		B*3901			VRCHKPARL	22		
24	CEKMALYDVVPMGFSYDTRC	A*01			LYDVVPMGFSY	26		
	CEKMALYDVVPMGFSYDTRC	A*0201	ALYDVVPM	0.4				
		A*0201			KMALYDVVPM	21		
		A*03			ALYDVVPMGF	24		
		A*03			DVVPMGFSY	24		
	CEKMALYDVVPMGFSYDTRC	A*24	LYDVVPMGF	0.4	LYDVVPMGF	22	LYDVVPMGF	140
		A*26	DVVPMGFSY	0.1	DVVPMGFSYD	24		
		A*26			DVVPMGFSY	35		
		B*1501	KMALYDVVPM	0.3				
		B*1501			ALYDVVPMGF	24		
		B*37			YDVVPMGF	24		
		B*3901	EKMALYDVVPM	0.5				
		B*5801	MALYDVVPMGF	0.4				
		Cw*0401					LYDVVPMGF	150
25	TVTESDIRTESLTERLYVGG	A*01	RTESLTERLY	0.13	RTESLTERLY	30	RTESLTERLY	113
		A*6801			DIRTESLTER	23		
		B*18			TESLTERLY	21		
		B*18			TESLTERL	21		
		B*2705			IRTESLTERL	24	IRTESLTERL	2000
		B*2705			IRTESLTER	27	IRTESLTER	1000
		B*37			TESLTERL	28		
		B*3901			IRTESLTERL	21		
		B*4001	TESDIRTESL	0.4	TESDIRTESL	23		
		B*4402			TESDIRTESL	23		
		B*4402			TESLTERLY	24		
		B*4403					TESLTERLY	120
		B*4901			TESLTERLYV	22		
		B*5802			RTESLTERLY	27		
		B*5802			RTESLTERL	23		
		B*60					TESDIRTESL	320
26	ASAACRAAKLCTMLVCGDDL							
		B*08			AAKLCTML	25		
	ASAACRAAKLCTMLVCGDDL	B*2705	CRAAKLCTM	0.5	CRAAKLCTM	21	CRAAKLCTM	600
		B*2705			CRAAKLCTML	22	CRAAKLCTML	2000
		B*2705					CRAAKLCT	200
		B*3901			CRAAKLCTML	21		
		B*5101					AAKLCTMLV	100
		B*5102					RAAKLCTMLV	121
		B*5103					RAAKLCTMLV	121
		B*5103					AAKLCTMLV	133
		B*5802			RAAKLCTML	23		
27	CESAGTQEDASLRAFTEAMT	A*1101			GTQEDASLR	24		
		B*18			QEDASLRAF	22		
	CESAGTQEDASLRAFTEAMT	B*2705					TQEDASLR	100
		B*4402			QEDASLRAF	25		
		B*5802			ASLRAFTEAM	21		
28	APTLWARMILEPLDLPQIIQ	A*0201			ILEPLDLPQI	26		
		A*0201			RMILEPLDL	21		
	APTLWARMILEPLDLPQIIQ	B*07					WARMILEPL	120
	APTLWARMILEPLDLPQIIQ	B*0801	WARMILEPL	0.5				
		B*1402			MILEPLDL	26		
		B*2705	ARMILEPL	0.25			ARMILEPL	2000
		B*2705			ARMILEPLDL	22	ARMILEPLDL	2000
		B*2705					RMILEPLDL	150
		B*3901			ARMILEPLDL	26		
		B*4402			LEPLDLPQII	21		
		B*4402			LEPLDLPQI	22		
29	WRHRARSVRAGGRAAICGKY	A*03			SVRAGGRAAI	21		
		A*03			SVRAGGRAA	22		
	WRHRARSVRAGGRAAICGKY	B*0702	SVRAGGRAA	0.4				
		B*0702	SVRAGGRAAI	0.15				
		B*2705			VRAGGRAAI	21	VRAGGRAAI	600
		B*2705			ARSVRAGGR	22	ARSVRAGGR	1000
		B*2705					ARSVRAGGRA	200
		B*2705					VRAGGRAAIC	200
		B*2705					VRAGGRAA	200
		B*5101					RAGGRAAI	110

No.	Sequence with linker	HLA	Peptide	Score	Peptide	Score	Peptide	Score
23	GYGAKDVRCHsgsgggsKPARLIVFPD	A*1101			H**SGSGGGS**K	21		
		A*1101			**GSGGG**SKPAR	22		
		A*26			DVRCH**SGSGG**	20		
		A*26			DVRCH**SGSG**	20		
		B*08			GG**GS**KPARL	21		
		B*1516			**GS**KPARLIV	27		
		B*5802			**GS**KPARLIVF	22		
24	CEKMALYDVVggPMGFSYDTRC							
		A*03			ALYDVV**GG**PM	22		
		A*03			ALYDVV**GG**P	20		
		A*03			DVV**GG**PMGF	20		
		A*03			VV**GG**PMGFSY	23		
		A*26	DVV**GG**PMGFSY	0.05	DVV**GG**PMGFS	23		
		A*26	DVV**GG**PMGF	0.1	DVV**GG**PMGF	30		
		A*26			VV**GG**PMGFSY	22		
		B*1501			VV**GG**PMGFSY	20		
		B*37			YDVV**GG**PM	21		
		Cw*0401					LYDVV**GG**PM	150
25	TVTESDIRTEggsgggSLTERLYVGG							
		A*1101			**GSGGG**SLTER	23		
		A*26			E**GGSGGG**SL	22		
		B*1516			**G**SLTERLYV	20		
		B*4001			TE**GGSGGG**SL	22		
		B*4402			TE**GGSGGG**SL	21		
		B*60					TE**GGSGGG**SL	160
26	ASAACRAAKLggsggCTMLVCGDDL	A*03			KL**GGSGG**CTM	22		
		B*1516			**GSGG**CTMLV	21		
		B*2705					CRAAKL**GGS**	200
27								
28	APTLWARMILggsggEPLDLPQIIQ	A*0201			IL**GGSGG**EPL	24		
		B*1402			**SGG**EPLDL	22		
		B*2705					ARMIL**GGS**	200
		B*4402			**G**EPLDLPQII	21		
		B*4402			**G**EPLDLPQI	22		
		B*4901			**G**EPLDLPQI	20		
		B*5802			**GSGG**EPLDL	21		
29	WRHRARSVRAsGGRAAICGKY	A*03			SVRA**S**GGRAA	20		
		A*03			SVRA**S**GGRA	20		
		B*07:02	SVRA**S**GGRAAI	0.25				
		B*07:02	RA**S**GGRAAI	0.4				
		B*2705					VRA**S**GGRAAI	600
		B*2705			ARSVRA**S**GGR	22	ARSVRA**S**GGR	1000
		B*5101			RA**S**GGRAAI	21	RA**S**GGRAAI	121
		B*5102					RA**S**GGRAAI	242
		B*5103					RA**S**GGRAAI	121
		B*5802			RA**S**GGRAAI	22		

- junction regions not modified
- all epitopes after linker insertion
- higher prediction scores after linker insertion
- new epitope after linker infertion