Structural Characterization of the anti-HIV-1 Broadly Neutralizing Monoclonal Antibody 2F5

by

Jean-Philippe Julien

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Department of Biochemistry
University of Toronto

© Copyright by Jean-Philippe Julien 2010
Abstract

Human immunodeficiency virus type 1 (HIV-1), the pathogen responsible for the onset of acquired immuno-deficiency syndrome (AIDS) in humans has reached pandemic proportions. To this day, no cure is available for infection with this virus and the only treatment option for this chronic infection is the life-long adherence to anti-retroviral therapy. Efforts in the quest to control the worldwide AIDS pandemic include the search for an effective anti-HIV-1 vaccine. Providing hope in this endeavor are a few monoclonal antibodies possessing broad neutralizing characteristics (bnmAbs) that have been isolated from the sera of rare patients that have a delayed progression to AIDS. In this thesis, one of these bnmAbs, 2F5 is extensively characterized at the atomic level to better understand its binding and neutralization mechanism. In total, 27 crystal structures of the 2F5 Fab’ in complex with various peptides representing its linear gp41 membrane proximal external region (MPER) epitope are presented. Furthermore, expression of the 2F5 Fab in a bacterial system allowed to design mutants of the 2F5 Fab and therefore investigate the implication of specific domains of 2F5 in mediating binding and neutralization. Atomic level characterization of this immune complex revealed a somewhat promiscuous recognition of 2F5 for its $^{664}$DKW$^{666}$ epitope as long as the following characteristics were conserved: the aspartate’s negative charge, the hydrophobic alkyl-π stacking arrangement
between the β-turn lysine and tryptophan, and the positive charge of the former. Moreover, it was demonstrated that 2F5 has an elongated and flexible complementary determining region 3 loop of the heavy chain (CDR H3), which is required for neutralization and is involved in secondary binding interactions other than to its core linear epitope. These contributions will significantly help in guiding the structure-based design of an HIV-1 vaccine looking to elicit 2F5-like antibody responses.
Acknowledgments

This thesis represents the culmination of a sum of events and people without which and whom the exact circumstances of this work would surely not have been possible.

First and foremost, I would like to recognize everyone who encouraged me to develop a sense of unbound curiosity from a very young age until the present. Particularly, my mother, my father and my grandparents have always surrounded me with great patience, belief, resources and wisdom, which laid the grounds for cultivating in me a deep passion for discovery.

To my dearest life partner, Rie, who I met during the course of my Ph.D. degree, I am thankful to you for continuously sharing the challenges and celebrations related to my work, for always supporting my dreams and for helping me keep my feet grounded in love, balance and life-long objectives.

To my colleagues and friends, especially Steve, Annie and Dave, I want to thank you for making coming to work everyday so enjoyable. It has been a true pleasure to work together with you and spend time outside of work enjoying the many pleasures of life. Particularly, I want to thank Steve for his generosity and patience in teaching, Dave for his sharing in and out of sciences, as well as Annie for her enthusiasm and help in bringing to life the projects that we created together.

In so many respects, I owe my development as a scientific thinker to my supervisor, Dr Emil Pai, and for that I am immensely thankful. By entrusting me with various responsibilities throughout the course of my stay in his laboratory, he significantly contributed to fostering many facets of my personality. From our numerous conversations, for which his door was always open, one of the most valuable teachings that I take away is the belief that structural biology is one of the cornerstones of understanding diseases and strategizing for designing treatment options.

In addition, I want to acknowledge both members of my supervisory committee, Dr David Isenman and Dr David Rose for always guiding my scientific path with wisdom and supporting me in various endeavours. Also, I would like to mention the enormous input of Dr José Luis Nieva in helping me to develop a comprehensive understanding of the field of HIV vaccine research, as well as thank him for his support and friendly supervision in the later stages of this thesis.
Finally, I would like to recognize the various institutions that contributed to my education, for believing that supporting me could have a beneficial impact on the immediate community and society as a whole. Particularly, I am thankful to individuals and ideals of the following organizations: École Secondaire Jean-Paul II, Lester B. Pearson College of the Pacific, McGill University, Canada Millennium Scholarship Foundation, Canadian Merit Scholarship Foundation, the University of Toronto, le Fonds québécois de la recherche sur la nature et les technologies and Canadian Institutes for Health Research.

-Je dédie cette thèse à tous ceux que l’expérience humaine a injustement fait souffrir et qui ont à travers le temps su trouver refuge dans la conscience et l’amour.
# Table of Contents

Acknowledgments .......................................................................................................................... iv  

Table of Contents ........................................................................................................................... vi  

List of Tables .................................................................................................................................. x  

List of Figures ................................................................................................................................ xi  

List of Appendices ....................................................................................................................... xiii  

List of Abbreviations ................................................................................................................... xiv  

Chapter 1 General Introduction ...................................................................................................... 1  

1 The road towards a Human Immunodeficiency Virus vaccine .................................................... 1  

1.1 The HIV pandemic .............................................................................................................. 1  

1.1.1 Epidemiology ................................................................................................................ 1  

1.1.2 Origin ............................................................................................................................. 1  

1.1.3 Pathogenesis .................................................................................................................. 2  

1.2 Advancements in the fight against AIDS ............................................................................ 6  

1.2.1 Antiretroviral therapy ................................................................................................. 6  

1.2.2 The difficulty in finding a cure for AIDS ................................................................. 7  

1.2.3 Prevention strategies ................................................................................................. 7  

1.3 Strategies for an HIV vaccine ............................................................................................. 9  

1.3.1 Results of past HIV vaccine trials ............................................................................ 9  

1.3.2 Strategies guiding HIV vaccine design ..................................................................... 12  

1.4 Broadly neutralizing antibodies against HIV-1 .................................................................. 15  

1.5 Broadly neutralizing monoclonal antibody 2F5 .................................................................. 18  

1.6 Thesis rationale ................................................................................................................... 19  

Chapter 2 Understanding the breadth of 2F5 HIV-1 neutralization ............................................. 21  

2 Crystallographic definition of the epitope promiscuity recognized by 2F5 ............................. 22
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>22</td>
</tr>
<tr>
<td>2.2</td>
<td>Materials and Methods</td>
<td>22</td>
</tr>
<tr>
<td>2.2.1</td>
<td>gp41 MPER sequence alignment analysis</td>
<td>22</td>
</tr>
<tr>
<td>2.2.2</td>
<td>2F5 Fab’ production and crystal complex formation</td>
<td>23</td>
</tr>
<tr>
<td>2.2.3</td>
<td>X-Ray diffraction analysis</td>
<td>26</td>
</tr>
<tr>
<td>2.3</td>
<td>Results</td>
<td>26</td>
</tr>
<tr>
<td>2.3.1</td>
<td>gp41 MPER variation</td>
<td>26</td>
</tr>
<tr>
<td>2.3.2</td>
<td>X-Ray diffraction experiments</td>
<td>29</td>
</tr>
<tr>
<td>2.4</td>
<td>Discussion</td>
<td>41</td>
</tr>
<tr>
<td>3</td>
<td>Epitope conformation, antigen-recognition loop mobility and anion binding site.</td>
<td>48</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>48</td>
</tr>
<tr>
<td>3.2</td>
<td>Materials and Methods</td>
<td>50</td>
</tr>
<tr>
<td>3.3</td>
<td>Results</td>
<td>53</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Structure of the ligand-free 2F5 Fab’</td>
<td>53</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Crystals of 2F5 Fab’-gp41 peptide complexes at low and high ionic strength.</td>
<td>53</td>
</tr>
<tr>
<td>3.3.3</td>
<td>The extended CDR H3 loop of 2F5 in the presence of peptide antigen</td>
<td>54</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Residues located at the C-terminus of the ELDKWA epitope core sequence</td>
<td>56</td>
</tr>
<tr>
<td>3.3.5</td>
<td>gp41 FP influences the conformation of an epitope peptide bound to the 2F5 paratope</td>
<td>61</td>
</tr>
<tr>
<td>3.3.6</td>
<td>The effect of the gp41 FP segment is specific</td>
<td>62</td>
</tr>
<tr>
<td>3.3.7</td>
<td>An anion-binding site at the base of CDR H3</td>
<td>65</td>
</tr>
<tr>
<td>3.4</td>
<td>Discussion</td>
<td>65</td>
</tr>
<tr>
<td>4</td>
<td>Ablation of the CDR H3 apex of the anti-HIV-1 broadly neutralizing monoclonal antibody 2F5 abrogates neutralizing capacity without affecting core epitope binding.</td>
<td>74</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>74</td>
</tr>
</tbody>
</table>
4.2 Materials and Methods...................................................................................................... 75
  4.2.1 Materials ............................................................................................................... 75
  4.2.2 Expression and purification of 2F5 Fab ................................................................. 76
  4.2.3 Expression and purification of non-glycosylated gp41 construct ......................... 77
  4.2.4 Production of vesicles ........................................................................................... 79
  4.2.5 Isothermal titration calorimetry ............................................................................ 79
  4.2.6 Cell-cell fusion and neutralization assays ............................................................. 79
  4.2.7 ELISA binding assays ........................................................................................... 80
4.3 Results....................................................................................................................... 82
  4.3.1 Supply of recombinant 2F5 Fab ............................................................................ 82
  4.3.2 Thermodynamics of 2F5ep binding to 2F5 Fab constructs ................................. 84
  4.3.3 Neutralization assays ............................................................................................ 85
  4.3.4 Binding affinities of recombinant 2F5 Fab and its mutants determined by ELISA ................................................................................................................... 86
  4.3.5 Relative binding affinity of recombinant 2F5 Fab constructs to different epitopes as determined by competitive ELISA..................................................... 89
  4.3.6 2F5 Fab interaction with membrane components ................................................. 91
4.4 Discussion .................................................................................................................. 92
Chapter 5 General Discussion.............................................................................................. 98
5 Towards an HIV vaccine capable of eliciting a broadly neutralizing antibody response .... 98
  5.1 Requirements for 2F5 recognition and neutralization .............................................. 99
    5.1.1 The rational design of a small immunogen capable of eliciting 2F5-like antibodies ................................................................................................................... 99
    5.1.2 Understanding the 2F5 epitope type-I β-turn in a biological context .................. 100
    5.1.3 An immunogen looking to elicit 2F5-like antibodies will need to incorporate other components in addition to its core epitope ........................................... 103
  5.2 The gp41 MPER as a target for HIV vaccine design ............................................ 106
    5.2.1 Roles of the gp41 MPER .................................................................................... 106
5.2.2 Immunogenicity of the MPER during natural infection ......................................... 107
5.2.3 bnmAbs targeting the gp41 MPER .................................................................... 108
5.2.4 Challenges associated with eliciting anti-MPER neutralizing antibodies .......... 110
5.3 Other anti-HIV-1 bnmAbs and Env vulnerability ...................................................... 113
5.4 Difficulties in eliciting bnmAbs against HIV-1 .......................................................... 116
5.5 Future Directions ........................................................................................................ 119
  5.5.1 Structural characterization of HIV-1 Env .............................................................. 119
  5.5.2 Conclusions ......................................................................................................... 121
References ..................................................................................................................... 122
Appendix 1 .................................................................................................................... 157
Copyright Acknowledgements ...................................................................................... 164
List of Tables

Table 2.1 List of peptide sequences (with associated MPER residue variation) that were used for co-crystallization experiments with 2F5 Fab’. Also listed are the X-ray diffraction data collection and refinement statistics for each 2F5 Fab’-peptide complex that formed crystals (with PDB identifier). ............................................................................................................................................... 25

Table 2.2 Natural sequence variation in HIV MPER and 2F5 neutralization (reported by Binley et al., J Virol 2004) across different clades. ................................................................................................................. 27

Table 3.1 Statistics for data collection and model refinement. Values in parentheses represent those in the highest resolution bin. ............................................................................................................. 52

Table 3.2 Contact Surface Areas by Residue on Peptide Epitopes and 2F5 Fab’. ......................... 62

Table 3.3 RMS deviations of gp41 epitope peptide residues for all structures presented in this work as well as for structure PDB ID 1TJH (Ofek et al. (2004)). ......................................................... 67

Table 4.1 Thermodynamic parameters of 2F5ep peptide (656NEQELLELDKWASLWN671) association to wild type 2F5 Fab and CDR H3 apex mutants. .............................................................. 85
List of Figures

Figure 1.1 Cryo-electron tomography of HIV-1 ................................................................. 3

Figure 1.2 Stages of HIV cellular entry and targets for broadly neutralizing MAbs .......... 4

Figure 1.3 Schematic diagram of the HIV life cycle showing the conventional and some new anti-retroviral drug targets ................................................................. 5

Figure 2.1 Enzymatic digestion of 2F5 IgG into antigen binding fragments using various enzymes .................................................................................................................. 24

Figure 2.2 Overlay of all MPER representing peptides (except for constrained peptides) characterized in complex with 2F5 Fab’ reported in this chapter ................................................. 30

Figure 2.3 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions at position 664 of the MPER ............................................................ 31

Figure 2.4 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions at position 665 of the MPER ............................................................ 33

Figure 2.5 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions at position 666 of the MPER ............................................................ 36

Figure 2.6 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions outside the DKW core, namely positions 662, 667 and 668 of the MPER 38

Figure 2.7 Crystal structures of 2F5 Fab’ in complex with peptides harboring β-turn stabilizing linkages between positions 663 and 667 of the MPER 40

Figure 2.8 Representation of spatial and chemical requirements for MPER residues for 2F5 core recognition as deduced from the present complex crystal structures .................. 45

Figure 3.1 Overall arrangement of the 2F5 Fab’ fragment bound to its extended gp41 peptide epitope as seen in the crystal structure of the Hyb3K peptide ........................................ 49
Figure 3.2 Visualization of the effect of crystal packing on the conformation of the 2F5 Fab’ CDR H3 loop in two different crystal systems. ............................................................... 56

Figure 3.3 The 2F5 Fab’ paratope in complex with a variety of gp41 epitope peptides in different crystal forms.................................................................................................................................. 60

Figure 3.4 Superposition of the gp41 peptide epitope conformations found in eight different crystal structures. ......................................................................................................................... 66

Figure 3.5 Model representation of the interaction between bnmAb 2F5 and HIV-1 gp41. ........ 71

Figure 4.1 Expression and purification of recombinant 2F5 Fab. ................................................ 77

Figure 4.2 Expression and purification of recombinant gp41 ectodomain. ............................... 78

Figure 4.3 Recombinant 2F5 Fab constructs. ........................................................................... 83

Figure 4.4 Isothermal titration calorimetry. Isotherms of 2F5ep(
\textsuperscript{656}NEQELLELDKWASLW\textsubscript{671}) binding to (A) wild-type 2F5 Fab and (B) 2F5 Fab delta CDR H3. ......................................................................................................................... 84

Figure 4.5 Neutralization and cell-cell fusion inhibition assays with the different 2F5 Fab constructs. ................................................................................................................................. 86

Figure 4.6 Direct and competitive ELISA assays measuring the binding of recombinant 2F5 Fab constructs to two gp41 constructs. .......................................................................................... 87

Figure 4.7 Competitive ELISA binding assays of the 2F5 Fab constructs with various epitopes. ........................................................................................................................................ 90

Figure 4.8 Direct and competitive ELISA assays measuring the binding of recombinant 2F5 Fab constructs to membrane components. .......................................................................................... 92
List of Appendices

Appendix 1: Attempt at crystallizing the 2F5 Fab’ in complex with the gp41 ectodomain in a post-fusion six-helix bundle conformation ................................................................. 157
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADCC</td>
<td>Antibody-dependant cellular cytotoxicity</td>
</tr>
<tr>
<td>AGM</td>
<td>African Green monkey</td>
</tr>
<tr>
<td>Ahx</td>
<td>6-amino-hexanoic acid</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired immunodeficiency syndrome</td>
</tr>
<tr>
<td>APOBEC3</td>
<td>Apoliprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G</td>
</tr>
<tr>
<td>BAFF</td>
<td>Human B-cell-activation factor</td>
</tr>
<tr>
<td>BCR</td>
<td>B-cell receptor</td>
</tr>
<tr>
<td>bnmAb</td>
<td>Broadly neutralizing monoclonal antibody</td>
</tr>
<tr>
<td>CCR5</td>
<td>CC chemokine receptor type 5</td>
</tr>
<tr>
<td>CD4</td>
<td>Cluster of differentiation 4</td>
</tr>
<tr>
<td>CD8</td>
<td>Cluster of differentiation 8</td>
</tr>
<tr>
<td>CDR</td>
<td>Complementarity determining region</td>
</tr>
<tr>
<td>CDR H3</td>
<td>Complementarity determining region 3 of the heavy chain</td>
</tr>
<tr>
<td>CHO</td>
<td>Chinese hamster ovary</td>
</tr>
<tr>
<td>Chol</td>
<td>Cholesterol</td>
</tr>
<tr>
<td>CXCR4</td>
<td>CXC chemokine receptor type 4</td>
</tr>
<tr>
<td>Dap</td>
<td>L-α-diaminopropionic acid</td>
</tr>
<tr>
<td>DC-SIGN</td>
<td>Dendritic cell-specific intracellular adhesion molecule-3-grabbing nonintegrin</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>DEAE</td>
<td>Diethylaminoethyl cellulose</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>Env</td>
<td>Envelope</td>
</tr>
<tr>
<td>EPR</td>
<td>Electron paramagnetic resonance</td>
</tr>
<tr>
<td>Fab</td>
<td>Fragment, antigen binding</td>
</tr>
<tr>
<td>Fc</td>
<td>Fragment, crystallizable</td>
</tr>
<tr>
<td>FP</td>
<td>Fusion peptide</td>
</tr>
<tr>
<td>FWR</td>
<td>Framework region</td>
</tr>
<tr>
<td>GALT</td>
<td>Gut-associated lymphoid tissue</td>
</tr>
<tr>
<td>GFP</td>
<td>Green fluorescence protein</td>
</tr>
<tr>
<td>gp120</td>
<td>Glycoprotein 120</td>
</tr>
<tr>
<td>gp160</td>
<td>Glycoprotein 160</td>
</tr>
<tr>
<td>gp41</td>
<td>Glycoprotein 41</td>
</tr>
<tr>
<td>HAART</td>
<td>Highly active antiretroviral therapy</td>
</tr>
<tr>
<td>HEK</td>
<td>Human embryonic kidney</td>
</tr>
<tr>
<td>HIV-1</td>
<td>Human immunodeficiency virus type 1</td>
</tr>
<tr>
<td>HIV-2</td>
<td>Human immunodeficiency virus type 2</td>
</tr>
<tr>
<td>HLA</td>
<td>Human leukocyte antigen</td>
</tr>
<tr>
<td>Acronym</td>
<td>Definition</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>HR1</td>
<td>Heptad repeat 1</td>
</tr>
<tr>
<td>HR2</td>
<td>Heptad repeat 2</td>
</tr>
<tr>
<td>IAVI</td>
<td>International AIDS Vaccine Initiative</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>ITC</td>
<td>Isothermal titration calorimetry</td>
</tr>
<tr>
<td>KIR</td>
<td>Killer cell IgG-like receptor</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani broth</td>
</tr>
<tr>
<td>LTNP</td>
<td>Long-term non-progressor</td>
</tr>
<tr>
<td>LUV</td>
<td>Large unilamellar vesicles</td>
</tr>
<tr>
<td>MAb</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MPER</td>
<td>Membrane proximal external region</td>
</tr>
<tr>
<td>NIAID</td>
<td>National Institute of Allergy and Infectious Diseases</td>
</tr>
<tr>
<td>NIH</td>
<td>National Institutes of Health</td>
</tr>
<tr>
<td>NK</td>
<td>Natural Killer</td>
</tr>
<tr>
<td>Nle</td>
<td>Norleucine</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NNRTI</td>
<td>Nonnucleoside reverse transcriptase inhibitor</td>
</tr>
<tr>
<td>Nrg</td>
<td>Nitroarginine</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>NRTI</td>
<td>Nucleoside reverse transcriptase inhibitor</td>
</tr>
<tr>
<td>NSLS</td>
<td>National Synchrotron Light Source</td>
</tr>
<tr>
<td>Orn</td>
<td>Ornithine</td>
</tr>
<tr>
<td>PA</td>
<td>1-Palmitoyl-2-hydroxy-(sn)-glycero-3-phosphate</td>
</tr>
<tr>
<td>Paf</td>
<td>(p)-Aminophenylalanine</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBMC</td>
<td>Peripheral blood mononuclear cell</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein Data Bank</td>
</tr>
<tr>
<td>PEPFAR</td>
<td>United States President’s Emergency Plan for AIDS Relief</td>
</tr>
<tr>
<td>PI</td>
<td>Protease inhibitor</td>
</tr>
<tr>
<td>pNPP</td>
<td>Alkaline phosphatase yellow</td>
</tr>
<tr>
<td>POPC</td>
<td>1-Palmitoyl-2-oleoylphosphatidylcholine</td>
</tr>
<tr>
<td>PT</td>
<td>Pig Tail monkey</td>
</tr>
<tr>
<td>rAD</td>
<td>Recombinant adenoviral vector</td>
</tr>
<tr>
<td>RM</td>
<td>Rhesus Macaque</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>Reverse transcriptase</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SIV</td>
<td>Simian immunodeficiency virus</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>SM</td>
<td>Sooty Mangabey</td>
</tr>
<tr>
<td>SPR</td>
<td>Surface plasmon resonance</td>
</tr>
<tr>
<td>Tat</td>
<td>Trans-activator of transcription protein</td>
</tr>
<tr>
<td>TBS</td>
<td>Tris-buffer saline</td>
</tr>
<tr>
<td>TCR</td>
<td>T cell receptor</td>
</tr>
<tr>
<td>Th17</td>
<td>T helper 17 cell</td>
</tr>
<tr>
<td>TM</td>
<td>Transmembrane domain</td>
</tr>
<tr>
<td>TMZ</td>
<td>Temozolomide</td>
</tr>
<tr>
<td>Treg</td>
<td>Regulatory T cell</td>
</tr>
<tr>
<td>TRIM5a</td>
<td>Antiviral restriction factor tripartite interaction motif 5a</td>
</tr>
<tr>
<td>UNAIDS</td>
<td>United Nations Joint Programme on HIV/AIDS</td>
</tr>
</tbody>
</table>
1 The road towards a Human Immunodeficiency Virus vaccine

1.1 The HIV pandemic

1.1.1 Epidemiology

In 2008, the Nobel Prize of Medicine was awarded to Françoise Barré-Sinoussi and Luc Montagnier, from France for the discovery of the Human Immunodeficiency Virus (HIV), 25 years earlier (20, 53). In addition to the pioneering work performed by Barré-Sinoussi and Montagnier, this Nobel Prize of Medicine served to recognize the global impact of the problems associated with HIV infections and the public health and scientific challenges that lie ahead in fighting this devastating epidemic.

As of today, approximately 25 million individuals have died from causes related to AIDS (273). An estimated 33 million were living with HIV in 2007, including 2.7 million new infections for that year alone (273). Although HIV/AIDS is of pandemic proportions, it can in fact be considered as two epidemics: one is in sub-Saharan Africa, where several countries have more than 2% of their population infected with HIV and going as high as 26% in Swaziland; the other is a “concentrated” epidemic in the rest of the world where it occurs mainly in high-risk groups. On a worldwide basis, heterosexual transmission represents the most important means of HIV acquisition. It therefore becomes clear that the people most at risk of becoming HIV positive and dying of AIDS are people who are sexually active with multiple partners, predominantly aged 15-30. In Southern Africa today, because of HIV/AIDS, life expectancy of youth 15-30 has dropped to 49 years old (273).

1.1.2 Origin

There is very little doubt among the scientific community that HIV-1 originated from the transmission of its simian equivalent (SIVcpz), from chimpanzees to humans. Where, when and how, however, remains an area of debate. The kick start of the epidemic is thought to have occurred in the large population centre of Kinshasa in the Democratic Republic of the Congo.
The most ancient isolated HIV samples go back to 1959 and 1960 (287, 303). From these samples, evolutionary and statistical analyses have demonstrated that the disease grew exponentially from one or at most a few individuals to the epidemic that in 2007 was estimated to have infected more than 55 million people (248, 287). The simplest explanation for how SIV<sub>cpz</sub> jumped to humans and became HIV-1 is through constant exposure of humans to the blood of chimpanzees butchered locally for bushmeat. Indeed, in a recent study, researchers found that 7 of 12 bushmeat samples being sold in West African markets were infected with SIV (12).

1.1.3 Pathogenesis

HIV is a *lentivirus* from the Retroviridae family. Its spherical shape of about 120 nm diameter makes it unusually large for a retrovirus (100). Its two copies of single stranded RNA (encoding nine genes (*gag, pol, env, tat, rev, nef, vif, vpr* and *vpu*) and 19 proteins) are encapsulated by a conical-shaped capsid made of the viral protein p24 (3). Surrounding the capsid is a matrix composed of the viral protein p17, which ensures the integrity of the virus. Finally, as seen in Fig. 1.1, this matrix is surrounded by the viral membrane, where many irregularly distributed copies of the Env protein are embedded in addition to proteins from the host cell membrane acquired during the initial budding of HIV from the host cell (3). The transmembrane glycoprotein Env is composed of two non-covalently associated trimers, gp120 and gp41, which are responsible for the attachment and fusion of the virus to CD4+ cells, the first step in initiating the infectious cycle (50).
For viral entry, HIV targets cells that possess on their surface the CD4 receptor, a protein found on cells of the immune system to assist during T cell receptor (TCR) activation. Membrane fusion mediated by HIV Env protein is a complex process that is only partially understood. To this day, the most accepted view of membrane fusion consists of a two-step model for receptor engagement that involves initial interaction between gp120 and the CD4 receptor (Fig. 1.2a), followed by conformational changes that allow interaction of the gp120-CD4 complex with a co-receptor on the cell surface (generally CCR5 or CXCR4) (Fig. 1.2b) (93). Binding of gp120 to a co-receptor triggers conformational changes in the gp41 subunit leading to formation of a pre-fusion intermediate termed the extended coil-coil (Fig. 1.2c) (173). Insertion of the N-terminal peptide of gp41, which occurs in the extended coil-coiled state, causes local defects and destabilization of the apposing membrane (93, 173). Following extended coiled-coil but prior to membrane fusion, a five-helix bundle is hypothetically formed which is responsible for the initial induction of membranes coming into close proximity of one another (93, 173). In addition to bundle formation, mutational analysis has shown the importance of a flexible hinge region
between the transmembrane (TM) anchor and the bundle to allow the viral and target membranes to be pulled together to close proximity (93, 190). Incorporation of the final C-terminus helix into the bundle provides the energy to bring the fusion peptides and viral anchors closer together, thus initiating membrane merging leading to the fusion of the membranes (Fig. 1.2d) (93). Env conformational rearrangement depends on a well-coordinated action of a bipartite system: raft organization appears to be crucial for HIV-1 entry in cells expressing low levels of receptors, probably through their concentration in distinct domains (173). In addition, it has been shown that at least three CD4 molecules and four to six CCR5 are required to form the fusion pore (147, 214). Although our understanding of HIV-1 mediated fusion has allowed us to establish a working model for the processes involved, many steps remain to be further elucidated.

Figure 1.2 Stages of HIV cellular entry and targets for broadly neutralizing MAbs. Each stage of viral entry is represented schematically. Numbers represent sites recognized by bnmAbs and are discussed in the text. In c), the blue arrow shows the C-terminal helix of gp41, whereas the orange arrow highlights the N-terminal peptide of gp41 (also termed Fusion Peptide). Figure reproduced with permission from Willey, S et al., TMBI, 2008 (285).

Upon entry in the cell cytoplasm, the virus initiates the disassembly of its core. First, its genomic RNA is converted to a DNA copy (a provirus) by the now uncoated HIV reverse transcriptase (RT) (Fig. 1.3). The provirus is then transported into the nucleus where it is integrated into host chromosomes by several viral and cellular proteins (125, 250). Transcription of provirus initially
depends solely on the cell’s RNA polymerase II machinery, and subsequently also on the HIV trans-activator of transcription (Tat), which enhances HIV transcription by several hundred fold (24). Then, together with cellular elements, HIV Rev mediates the export of transcribed viral RNAs to the cytoplasm and their post-translational regulation into spliced and unspliced RNA (161). The viral unspliced RNA, which constitutes the genomic RNA, forms a dimer based on structured RNA motifs and is thereafter integrated into a newly assembled viral particle, a process mediated by HIV Gag (43, 299). The end of the HIV life cycle is marked by the budding of the virus, which transits via the cell’s late endosomal pathway through multi-vesicular bodies after final assembly (205).

![Figure 1.3 Schematic diagram of the HIV life cycle showing the conventional and some new anti-retroviral drug targets.](image)

Figure 1.3 Schematic diagram of the HIV life cycle showing the conventional and some new anti-retroviral drug targets. Figure reproduced with permission from Bhattacharya, S et al., J Infect, 2009 (25).

In addition to infected blood, other fluids that contain high enough viral loads to cause initial infection in an individual include semen, vaginal fluid, and mother's milk. Mother-to-child transmission is also observed frequently during childbirth (15-35%) (268). The probability of HIV transmission per sexual act of vaginal intercourse, or infectivity, is relatively low and has been estimated to vary from approximately 0.0001 to 0.0014 (106). In the case of primary
infection through the sexual mucosa, HIV initially infects CD4+ T cells, macrophages and dendritic cells, which then migrate to the lymph nodes to activate an immune response (86). However, high HIV replication at these sites for the first 4-8 days depletes CD4+ T cells extensively, which leads to early and irreversible damage to the immune system during the acute phase of HIV-1 infection. The organ that is primarily affected is the gut or gut-associated lymphoid tissue (GALT), the human body’s largest lymphoid organ (86). The end of the acute phase of HIV-1 infection is marked by a sharp fall in virus in the peripheral blood (viral load), an event commonly attributed to the appearance of HIV-specific CD8+ cytotoxic T cells (13).

After the reduction in peak virus, the viral load settles down to a set point, a quantity predictive of the rate of progression to AIDS (181). Indeed, the higher the amount of virus in the peripheral blood at this stage, the faster the progression towards immunodeficiency tends to be. Following acute infection is a prolonged clinically asymptomatic period, where HIV continues to actively replicate and slowly deplete CD4+ T cells in the lymphoid organs. However, a significant regeneration of CD4+ T cells combined with the mounting of both cellular and humoral immune responses by the host leads to a constant immune activation to fight HIV infection. For several years, the host and pathogen are entangled in a constant battle where HIV generates its own substrate for replication through the induction of immune activation (77). It takes on average nine years for a person infected with HIV to develop AIDS, with five years being considered rapid and remaining well for more than 15 years being termed a case of long-term non-progression (86). Inevitably, however, CD4+ T cell counts will fall below 200 cells \( \mu l^{-1} \) at which point opportunistic infections such as tuberculosis take advantage of the mere absence of an immune system to infect the host and eventually cause his/her death.

1.2 Advancements in the fight against AIDS

1.2.1 Antiretroviral therapy

There is no doubt that the most significant progress made over the last two decades of scientific research in the battle against HIV/AIDS has been the development of effective antiretroviral drugs that allow for the chronic suppression of HIV replication. Highly active antiretroviral therapy (HAART) is at the centre of treatment strategies for patients infected with HIV to an extent that in almost all developed countries, HIV can be considered a chronic manageable condition. Antiretroviral drugs can be classified according to their viral targets: Nucleoside
Reverse Transcriptase Inhibitors (NRTIs), Nonnucleoside Reverse Transcriptase Inhibitors (NNRTIs), Protease Inhibitors (PIs), Fusion Inhibitors, Entry Inhibitors - CCR5 co-receptor antagonist and HIV integrase strand transfer inhibitors (Fig 1.3). A total of 32 approved drugs are currently available for the treatment of HIV/AIDS (2). Simplicity of dosing is increasingly becoming commonplace, such that patients can often rely on a course of therapy of one pill per day, usually consisting of a co-formulation that includes two NRTIs and one NNRTI (277). On average, drug regimens allow AIDS patients to live an additional 14 years (277).

1.2.2 The difficulty in finding a cure for AIDS

No cure is currently available to eradicate an HIV infection. Chronic suppressive therapy by HAART is the only method available to contain persistent infection in reservoirs and prevent viral replication that would otherwise speed up the progression towards AIDS. Indeed, discontinuation of HAART results in an immediate and sharp rebound in viral load, suggesting that there is a latent reservoir of HIV provirus in the infected individual that antiretroviral drugs cannot reach, even after prolonged treatment (225). The best characterized latent HIV reservoir are infected resting memory CD4+ T cells, although other drug-insensitive reservoirs such as the brain, macrophages and haematopoietic stem cells have also been hypothesized to harbour latent HIV provirus (225). Many efforts are currently invested in developing novel drugs combined with specific therapeutic regimens with the objective of strategically eliminating the source of latent HIV provirus (225). Succeeding with this goal would possibly allow moving away from the requirement of lifelong adherence to HAART therapy and surely bring us a step closer to finding a cure for HIV/AIDS.

1.2.3 Prevention strategies

In 2007, the HIV war chest of the global community to help the less fortunate combat HIV/AIDS totaled $10 billion (56). With no available cure for HIV/AIDS, a major fraction of this money served to provide anti-retroviral drugs to people who need them. Between 2003 and 2005, the number of people in developing countries receiving anti-retroviral drugs increased by an average of half-a-million each year, such that antiretroviral therapy coverage rose from 7% in 2003 to 42% in 2008 (1). Although these numbers show significant progress towards universal treatment for everyone infected with HIV, performing the harsh mathematics of this epidemic quickly shows that treatment without prevention is simply unsustainable. Even when assuming the
lowest possible prices for drugs, the annual cost of getting treatment to everyone in the world who is HIV positive would be more than $13 billion a year, every year (97). This figure doesn’t include any other expenses, such as testing, counseling, training of medical staff, research, prevention or orphan care. Moreover, this figure assumes no increase in the number of people living with HIV. Yet, according to a 2008 estimate, for each new person who started treatment, nearly three new people became infected with HIV (56). Clearly, in an era where no cure for HIV/AIDS exists, alongside antiretroviral treatment, prevention strategies need to be at the centre of the fight to conquer this devastating pathogen.

To this day, the most effective prevention strategy for the fight against HIV/AIDS remains the ABC program, for Abstain, Be faithful, use Condoms. This sensitization program was widely put in place over the last two decades by some African countries, and particularly encouraged by the United Nations Joint Programme on HIV/AIDS (UNAIDS) and the United States President's Emergency Plan for AIDS Relief (PEPFAR). One of the most famous success stories related to this prevention campaign is that of Uganda, where HIV prevalence reached a peak of 15% in 1993 to then steadily decline to 5% by 2001 (254). Over this period, it was found that fewer Ugandans were having sex at young ages, levels of monogamy increased and condom use rose steeply among unmarried sexually active men and women (254).

Another prevention approach in the fight against HIV/AIDS is male circumcision. Indeed, recent studies found that male circumcision may help to reduce female-to-male transmission of HIV by as much as 60%, in addition to reducing female genital symptoms and vaginal infections significantly (15, 104, 105). It is thought that the elimination of the foreskin diminishes the likelihood of transmission because this tissue is rich in Langerhans cells, an important entry point for HIV into the body (277). Furthermore, because the preputial mucosa of the foreskin is a significant source of viral shedding, male circumcision might protect women from the risk of HIV infection by lowering infectivity and by decreasing the chances of acquisition of HIV cofactors such as other sexually transmitted infections.

Public health programs rooted in epidemiologic observations and scientific understanding such as the ones mentioned above should be expanded as widely as possible in areas of high-risk of contracting HIV. However, behavioral, cultural, political and religious barriers often prevent such programs from becoming common practice. In light of the difficulties of establishing such
programs on a worldwide scale, it is clear that the most effective prevention method would be an effective HIV vaccine.

1.3 Strategies for an HIV vaccine

Since the discovery of a vaccine for smallpox by Jenner in 1796, the scientific community has successfully developed vaccines against a myriad of other pathogens: diphtheria, cholera, bubonic plague, tetanus, pertussis, pneumococcus, hepatitis A, hepatitis B, varicella, typhoid, measles, mumps, yellow fever, rubella, polio, rabies, human papillomavirus and influenza (11, 216). For most of these cases, the component conferring protection to the host through vaccination consists of a killed micro-organism (influenza, cholera, bubonic plague, polio and hepatitis A), a live-attenuated micro-organism (yellow fever, measles, rubella, mumps and typhoid) or a protein subunit of the micro-organism (hepatitis B, human papillomavirus and influenza).

1.3.1 Results of past HIV vaccine trials

1.3.1.1 Whole-killed and live-attenuated HIV vaccine

Whole-killed virus vaccines have been very successful in history to immunize patients against viral infections (249). In the case of HIV however, this vaccination method has proven ineffective mainly because the methods used to kill HIV seem to alter the native conformation of the virus, particularly that of the viral envelope (78). Hence this vaccine has been shown to be non-antigenic and incapable of generating broadly neutralizing antibody responses or HIV-specific cytotoxic T cell responses (19). Furthermore, there exist significant safety considerations that such a vaccine could be dangerous in the case that residual infectious virus remains present due to incomplete inactivation.

The utility of a live-attenuated vaccine for controlling the HIV epidemic has been long debated, as it would represent a significant tradeoff between being effective in protecting against infections and not being completely safe as the attenuated strain could cause AIDS in some vaccinated individuals. Indeed, one of the differences between HIV and other pathogens for which a vaccine has been designed resides in the fact that the human immune system is not capable of clearing an HIV infection on its own, even if it started from an attenuated virus. Use
of live attenuated SIV as vaccines in monkeys had initially provided near-complete protection from challenge with the same SIV virus strain, which created a lot of excitement at the turn of the millennium as it was a proof of principle for the feasibility of such an HIV vaccine (144). However, subsequent studies with heterologous SIV challenge did indeed cause immunodeficiency in monkeys over time, illustrating the additional difficulty of conferring protection against a highly variable virus (16). Moreover, an attenuated HIV virus with a \textit{nef} gene deletion that was originally thought to be a good candidate for a live-attenuated HIV vaccine was found to also lead to AIDS in a group of individuals infected with this specific HIV strain in the Sydney Blood Bank Cohort (284). Since these discoveries, the idea of a live-attenuated HIV vaccine has significantly faded and other vaccine strategies of higher safety are now at the forefront of research efforts.

1.3.1.2 Protein subunit HIV vaccine

In 1999, efforts of 15 years of dedicated scientific research by Genentech and Vaxgen costing over $300 million dollars resulted in the first efficacy trial of a candidate HIV-1 vaccine. In a phase III clinical trial conducted in Bangkok, Thailand, a protein subunit vaccine candidate consisting of a mixture of two gp120 surface viral protein subunits of HIV-1 clades B and E produced from transfected CHO cell lines was evaluated on its efficacy of reducing HIV-1 infection in 2546 injection drug users (89, 215). Pre-clinical studies showing that antibodies to gp120 could inhibit HIV-1 fusion events and that monovalent gp120 could elicit neutralizing antibodies in humans, combined with multiple phase I/II trials to arrive at an optimal dosing and administration schedule gave hope that this HIV-1 vaccine candidate, termed AIDSVAX, could be efficacious (156, 246). However, results of this clinical trial released after a 36-month follow-up of volunteers were very disappointing: there was no difference in HIV prevalence amongst the vaccine and placebo groups (215). Furthermore, no statistically significant effect of the vaccine was observed on reducing the HIV-1 viral load upon infection, a secondary end point measured in this trial (215). Altogether, these results had a devastating effect on the hopes of scientists to use an HIV protein subunit vaccine to elicit antibodies capable of preventing HIV-1 infection.
1.3.1.3 DNA-based HIV vaccine

Following this failure in eliciting protective anti-HIV-1 neutralizing antibodies, scientists sought to develop a vaccine candidate that could induce immunity that would protect from progression to AIDS, instead of preventing HIV acquisition. The aim of such a vaccine would be to stimulate a cellular response (mostly cytotoxic T lymphocytes) that would allow the reduction of viral load and provide considerable benefits to the health of the individual, as well as reducing rates of transmission and treatment demands. In early vaccine studies, although no sterile protection was observed, there was evidence of prolonged survival in rhesus monkeys who had been vaccinated with whole inactivated SIV (116, 251). It appeared that the mechanism of protection was lowering of the viral set point mediated by the T cells induced during vaccination (183). Subsequently, substantial efforts were invested in designing a DNA-based vaccine to elicit a cellular response capable of protecting against progression to AIDS upon HIV-1 infection. With this objective in mind, a phase IIb proof-of-concept clinical trial co-sponsored by the pharmaceutical company Merck and the US National Institute of Health (NIH) was started in 2004. The vaccine candidate consisted of a combination of three HIV antigens expressed in replication-defective recombinant adenoviral vectors (rAD) looking to generate a robust T cell response. In this trial, 3000 high-risk uninfected volunteers from around the world were immunized with a vaccine consisting of three rAD5 vectors: Ad5-gag, Ad5-pol and Ad5-nef (247). In 2007, this trial was abruptly stopped because not only did the vaccine fail to protect seronegative individuals against infection and lower their viral load by eliciting an effective cellular response, but it possibly even enhanced HIV-1 infection in individuals with prior immunity to adenoviruses (55, 158).

1.3.1.4 Prime-boost HIV vaccine

During the same period, another large test-of-concept vaccine trial started, administered by the Thai Ministry of Public Health and the US Military HIV Research Program, in collaboration with the NIH, the pharmaceutical company Sanofi Pasteur and Global Solutions for Infectious Diseases (224). This vaccine candidate consisted of a prime-boost regimen, where a DNA-based vaccine was first administered as a prime followed soon after by an injection with a protein subunit vaccine as a boost. Although the AIDSVAX bivalent gp120 vaccine had not demonstrated any significant capacity in reducing the risks of HIV infection in previous clinical trials, it was observed in a phase II trial that when this component was given after first priming
with a DNA-based canarypox-HIV-env-gag-protease vector (ALVAC-HIV), induction of pre-specified cellular and humoral immune responses resulted (197). Based on this partial success, 16,402 healthy men and women at low-risk of HIV infection were recruited for the RV144 advanced vaccine trial, where a placebo or prime-boost ALVAC/AIDSVAX regimen was administered using a specific schedule (224). After a three year follow-up of vaccinated individuals, results of the trial were made available in late 2009. In a subgroup of the study, a trend toward the prevention of HIV-1 infection among the vaccine recipients could be noticed: 51 HIV-1 infections were observed in the vaccine groups, whereas 72 infections where observed in the placebo group. These figures correspond to a 31.2% vaccine efficacy (p=0.04, 95% confidence interval=1.7 to 51.8) (224). On the down side of the study, it was found that vaccine recipients were not better adapted to reduce significantly the viral load upon HIV infection when compared to the placebo group. Although limited by many factors including poor statistical power, the efficacy of the ALVAC/AIDSVAX vaccine was encouraging and gave hope for the first time that the development of an HIV vaccine could be within reach. Follow-up research is currently underway to determine and characterize the immune correlates of protection induced by this vaccine.

1.3.2 Strategies guiding HIV vaccine design

1.3.2.1 SIV infection

As a strategy to help in the design of an HIV vaccine, researchers have been trying to better understand the specific immune correlates that make some hosts capable of either resisting infection, clearing the virus upon infection or simply living symbiotically with the virus. For this purpose, SIV infection represents a great case study. Indeed, African Green monkeys (AGMs) and Sooty Mangabeys (SMs) are natural hosts of SIV and infection in these monkeys never progresses to disease; on the other hand, SIV infection in Pig Tail monkeys (PTs) and Rhesus Macaques (RMs) results in a disease that closely resembles AIDS in humans (252). Interestingly, in both pathogenic and non-pathogenic SIV infections, the acute phase of infection is similar, characterized by a high level of viremia, substantial levels of T cell activation and depletion of the GALT CD4+ T cells (252). It is during chronic infection that a major difference between pathogenic and non-pathogenic SIV infection is observed. Although in both cases the SIV viral load remains high, in natural hosts of SIV, a down regulation of activated inflammatory genes is observed, whereas there is sustained activation of genes associated with inflammation in PTs and
RMas (209). Some of the prominent hypotheses postulated to explain these differences in sustained immune activation is a loss of Th17 cells / Treg cells balance in hosts of pathogenic infection and differential expression of receptors on the surface of the T central memory cells in the different monkey species, which allow the natural hosts of SIV infection to maintain homeostasis even under high viral loads (209, 252). The observation that SIV infection of the natural host leads to high viral load without evidence of immune suppression or disease progression has led to the idea that immune activation rather than viral replication is the primary mechanism by which HIV infection leads to AIDS. These findings give new insights to design vaccine and therapy strategies that seek to tune the immune system to prevent progression to AIDS, possibly by trying to down-regulate the sustained immune activation characteristically observed during HIV infection.

1.3.2.2 HIV-2

Another natural occurrence from which scientists can extract valuable information related to HIV vaccine design is the case of HIV type 2 (HIV-2). HIV-2 is thought to have originated from the transfer of the SIV from Western African Sooty Mangabeys (SMs) to humans approximately a decade after the introduction of HIV-1 into the human population from chimpanzees and this virus now infects approximately one million individuals in West Africa alone (159). At the amino acid level, there is a 60% similarity between the Gag and Pol proteins of the two subtypes, and a 30-40% similarity in the regions encoding Env, making the two HIV subtypes closely related retroviruses (108). In general, HIV-2 can be described as a natural human model of attenuated HIV infection. Indeed, upon HIV-2 infection, the CD4+ T cell decline is much slower and the number of AIDS-free survivors at five years post-infection is significantly higher (100% versus 67%) than with HIV-1 (126). Only 20-25% of HIV-2 infections lead to AIDS (syndromes indistinguishable from HIV-1 AIDS) and the remainder, termed long-term non-progressors (LTNP), show slow or no progression to AIDS and are characterized by an undetectable plasma viral load (65). Lower viral replication and enhanced host immune control have been put forward as explanations of the mild outcome of an HIV-2 infection (65). It is interesting to note that although the SMs are natural hosts to SIV and that HIV-2, which is closest to SIV from SMs, is also non-pathogenic in the majority of infections, the molecular mechanisms by which the host deals with the viral infection are quite different. As mentioned previously, SMs sustain a non-pathogenic high SIV viral load with a non-activated immune response. In the case of HIV-2
LTNPs, an initial immune activation is capable of reducing the virus to undetectable levels in the plasma through virus-specific CD8+ cytotoxic T cells, CD4+ T cells, Natural Killer (NK) cells and broadly neutralizing antibody responses (65). Unfortunately, HIV-2 infection provides no protection for subsequent HIV-1 infection, which precludes its use as a vaccine (65). Nonetheless, being able to answer the question as to why HIV-2 fails to cause disease in the majority of infected individuals, yet imitates HIV-1 in its ability to cause AIDS in others, might give invaluable clues for the crafting of an effective HIV-1 vaccine.

1.3.2.3 Elite controllers and long-term non-progressors

Lastly, numerous studies have shown that not all individuals are equally susceptible to HIV infection and that their progression to AIDS varies significantly. Of particular importance, 5 – 10% of HIV-1 infected individuals do not progress to AIDS for more than 7 years; they are also referred to as long-term non-progressors (LTNP) (69). Moreover, approximately 1% of total infected individuals, termed elite controllers, maintain a viral load at almost undetectable levels of <75 copies of RNA per millilitre (155). Amongst many factors, individuals vary in their susceptibility to the virus, viral set point upon infection, viral load sustained during chronic infection, rate of decline of CD4 T cells and the emergence of escape mutants. Understanding the molecular mechanisms by which these individuals mount such an impressive level of natural resistance against HIV could be the key to develop a vaccine capable of providing comparable resistance to the virus.

1.3.2.3.1 Genetic and immunological markers of HIV control

One of the earliest discovered genetic markers giving individuals resistance to HIV/AIDS was the naturally occurring deletion of 32 bases of the CCR5 gene (CCR5 Δ32), which introduces a premature stop codon resulting in a truncated protein product (166). CCR5 and CXCR4 are co-receptors for HIV-1 present on the surface of CD4+ T cells that are involved in viral entry, subsequent to the interaction of HIV-1 Env with the CD4 receptor. Homozygotes possessing the CCR5 Δ32 mutation are almost completely protected from HIV/AIDS whereas heterozygotes possessing this deletion delay the progression to disease by 2-4 years (66). The CCR5 Δ32 mutation is extremely rare in the African and Asian populations, but it is present at a frequency of approximately 15% in Northern Europe (94, 199).
It has been estimated in a recent study that HIV uses at least 250 host proteins to gain entry into target cells and complete its life cycle (31). It is therefore understandable that many genomic approaches have been designed to identify human genetic variants that regulate the course of HIV infection. To date, more than ten such sets of genes have been identified (138). Some of the host genetic factors associated with the differential regulation of HIV/AIDS that have conclusively been established include: 1) Human Leukocyte Antigen (HLA) polymorphic loci and their associated genes; 2) chemokine receptors (such as CCR5, as previously described) and chemokine ligands; 3) antiviral restriction factor tripartite interaction motif 5a (TRIM5a); 4) apolipoprotein B mRNA-editing enzyme, catalytic polypeptide-like 3G (APOBEC3); 5) dendritic cell-specific intracellular adhesion molecule-3-grabbing nonintegrin (DC-SIGN); 6) the killer cell IgG-like receptor (KIR) gene; 7) interferon regulatory factor 1; and 8) LY6 family of G [glycosylphosphatidylinositol (GPI)]-anchored proteins (138).

Moreover, large scale studies comparing the immune activation observed in elite controllers compared to that of progressors revealed 1) an over-representation of protective HLA-B*57 and B*27 alleles; 2) high ratios of functional CD4:CD8 T cells; 3) high generalized T cell activation levels; as well as 4) high proportions of polyfunctional T helper cells and cytotoxic T lymphocyte responses (17). Although these elements are characteristic of many elite controllers, they are not observed in all cases, suggesting that other yet unidentified factors are also involved in explaining the capacity of these rare individuals to naturally control HIV/AIDS. Ongoing studies of elite controllers, including those investigating host genetic polymorphisms, will surely help in defining the characteristics of an effective HIV-specific immune response.

### 1.4 Broadly neutralizing antibodies against HIV-1

In light of the predominant role played by the humoral branch of the immune system in protecting human hosts against various pathogens by vaccination, it is evident that one of the long sought-after goals in the quest for an HIV vaccine is the elicitation of neutralizing antibodies that would confer protection against HIV-1 infection (42, 143). These antibodies would need to interact with HIV-1 envelope (Env), target conserved regions and functional conformations of Env trimeric complexes and prevent HIV-1 fusion events with target cells (88, 237, 291, 294). As previously mentioned, past attempts at eliciting antibodies capable of neutralizing HIV-1 of diverse genetic backgrounds have met with only limited success, although
very recent studies have presented somewhat more encouraging results (14, 41, 42, 84, 110, 175, 224, 258).

On its surface, HIV-1 displays Env glycoproteins that mediate the membrane fusion events essential for the infection of CD4+ T-cells by the virus (75, 79, 93). Env is synthesized as a gp160 precursor, which subsequently is cleaved in the Golgi apparatus resulting in two subunits, gp120 and gp41. The two chains stay bound to each other and the heterodimers further assemble on the surface of the virion into non-covalently associated heavily glycosylated trimers (49). The crucial role of Env for viral infectivity and its presence on the surface of the virus should make it an ideal target for vaccine design. However, it has proven difficult to elicit a broad and potent neutralizing antibody response against HIV-1 Env both during the course of natural infection and by vaccination. This is thought to be predominantly caused by the elevated genetic diversity of the virus due to an error-prone reverse transcriptase. The extensive genetic diversity of HIV-1 and its ability to constantly adapt and escape immune responses represents a major hurdle in generating an antibody response capable of neutralizing all functional variants of HIV-1. In fact, it is estimated that there is up to 35% sequence variation in HIV-1 Env between subtypes and that the rate of evolution of the virus is up to a million times faster than that of animal DNA (96, 145). Other factors that preclude an effective humoral response to be mounted against HIV-1 Env include the masking of functional Env regions by conformational covering and by an extensive glycan shield, the ability of some conserved domains to partition to the viral membrane, as well as a high quantity of heterogeneous and non-functional Env displayed on the surface of HIV-1, which often mislead the humoral response (120, 131, 154, 162, 185, 186, 234, 285, 290).

In the mid 1990’s and 2000’s, the discovery of monoclonal antibodies capable of neutralizing a broad range of HIV-1 isolates on their own, termed broadly neutralizing monoclonal antibodies (bnmAbs) renewed the hope that a protective vaccine eliciting an effective neutralizing antibody response against HIV-1 Env was possible. Indeed, it was demonstrated that the passive administration of these bnmAbs prior to HIV or SIV exposure provided protection to the hosts upon challenge with the virus (58, 113-115, 174). These remarkable results suggested that if these bnmAbs could be elicited prior to HIV-1 exposure by vaccination, sterilizing protection could be achieved.
Subsequently, numerous studies were designed to further characterize neutralizing antibodies, and identify new anti-HIV-1 bnmAbs. From such studies, it was found that a humoral response generating neutralizing antibodies against HIV-1 can be detected early in HIV-1 positive individuals, but that the titers are often very low and viral control is seldom achieved by these neutralizing antibodies (91, 226, 228, 280, 282). More recent studies have shown that a considerable proportion (approximately 25%) of HIV-1 positive subjects who have been infected for at least one year show moderate to broadly neutralizing antibody responses (76, 236, 253, 259). Furthermore, approximately 1% of the HIV-1 subjects studied, termed elite neutralizers, had an unusually high neutralizing potency against many HIV-1 clades (253).

Although some HIV-1 infected individuals possess broadly neutralizing sera in which combinations of neutralizing monoclonal antibodies act together to neutralize a wide genetic range of viruses, monoclonal antibodies that are able to effectively neutralize various HIV-1 clades on their own were found to be very rare (241). Of the few well characterized ones, bnmAb b12 recognizes the site on gp120 responsible for binding to the CD4 receptor during the initial stages of fusion (Fig. 1.2a, site 1) (73, 301). The epitope of bnmAb 2G12 is a mannose cluster on the outer face of gp120 (Fig. 1.2a, site 2) (44, 239, 240, 271). Another site is that recognized by bnmAb X5: the CD4-induced co-receptor interacting site of Env (Fig. 1.2b, site 3) (188). Finally, three bnmAbs targeting gp41 have been identified: 2F5, 4E10 and Z13 (47, 191, 192, 194). These bnmAbs recognize conserved epitopes located in the membrane proximal external region (MPER) of gp41, between the second heptad repeat (HR2) and the TM domain (Fig. 1.2c, site 4) (235, 304, 306). Ongoing efforts to identify new bnmAbs are also underway, e.g. the International AIDS Vaccine Initiative (IAVI) Protocol G, which is a global collaboration that has collected blood samples from over 1,800 HIV positive elite neutralizers around the world with the objective to discover new bnmAbs. Early results of this study identified two new HIV-1 bnmAbs, PG9 and PG16, which recognize an epitope on the gp120 trimer and show particularly high potency and breadth (253, 278). With sustained efforts, it is highly probable that the IAVI Protocol G, as well as other such global collaborations will continue to identify novel anti-HIV-1 bnmAbs. Few would argue that understanding the neutralizing mechanism of these rare bnmAbs at the molecular level could yield invaluable clues for the rational design of a vaccine immunogen by identifying vulnerable sites on HIV-1 Env.
1.5 Broadly neutralizing monoclonal antibody 2F5

The neutralizing activity of sera from HIV-1 seropositive individuals has been demonstrated to be predominantly mediated by anti-gp120 antibodies, with anti-gp41 antibodies only playing a limited role (32). In fact, antibodies against gp41 are mainly directed against the immunodominant N-terminus region of gp41 and monoclonal antibodies recognizing epitopes outside this region have not been frequently described (198).

At the beginning of the 1990’s, Katinger and co-workers set out to identify human monoclonal antibody producing cell lines that could be important tools for diagnosis, research, and potential therapeutic agents (38). For this purpose, they screened asymptomatic HIV-1 positive volunteers of the Centres for Disease Control in Atlanta, GA. Donors in a late stage of HIV-1 infection were selected based on two criteria: 1) they had a lymphocyte count of at least 500 cells/mm$^3$ and 2) they had a CD4 T cells/CD8 T cells ratio greater than one (38). Subsequently, the sera of patients were screened for HIV-1 peptide reactivity and in the case that patients exhibited elevated serum titers for at least one of these peptides, their peripheral blood mononuclear cells (PBMCs) were isolated. This study resulted in the identification of 33 stable anti-HIV-1 human monoclonal antibody producing cell lines against different HIV-1 epitopes, including clones 4E10, 2G12 and 2F5 (38).

Human monoclonal antibody 2F5 of IgG3(κ) subtype, one of the broadly neutralizing monoclonal anti-HIV-1 antibodies isolated from this study, reacted with a core gp41 MPER epitope spanning residues 662 to 668 with the linear sequence ELDKWAS, a sequence found in approximately 70% of HIV-1 isolates currently described (26, 38, 192, 260, 306). Neutralization assays, fusion inhibition experiments and syncytium inhibition assays of different HIV subtypes revealed that 2F5 has the capacity to neutralize laboratory-adapted strains BIB, MN, RF, SF2, R640, and R548 with a 50% effective concentration (EC50) ranging from 0.7 to 9.7 μg/ml, as well as approximately 70% of a panel of primary isolates with an EC50 < 50 μg/ml (26, 220). Subsequent 2F5 IgG binding studies and screening of phage display libraries demonstrated that the DKW core epitope is essential for 2F5 recognition and binding (59, 182, 219). Binding of 2F5 is thought to occur during the fusion process, probably to the pre-hairpin intermediate form of gp41 (28, 64, 85, 103, 204, 208, 229, 243, 306). By binding to its target, this bnmAAb is
believed to prevent virus-host cell fusion by impairing the required conformational transitions of gp41 to a final six-helix post-fusion conformation (18, 28, 64, 85, 103, 204, 208, 211, 229, 243).

Previous crystal structures of 2F5 with peptides representing its core gp41 epitope revealed that 2F5 has evolved to recognize the gp41\(^{664}{\text{DKW}}^{666}\) residues in a type-I \(\beta\)-turn conformation (204, 208). In addition to binding to its primary epitope, evidence is accumulating that 2F5 also takes part in secondary interactions with gp41 and HIV-1: multiple reports have demonstrated the affinity of 2F5 for membrane components, possibly through its partly hydrophobic flexible elongated CDR H3 loop, and it has also been suggested that 2F5 might interact in a secondary manner with other regions of gp41, such as with residues located immediately N- and C-terminal to its core epitope or with more distantly located regions of gp41 like residues of the gp41 Fusion Peptide (FP) (7, 37, 109, 168, 176, 233).

1.6 Thesis rationale

In the quest for the design of an effective HIV-1 vaccine capable of eliciting a broadly neutralizing antibody response providing sterilizing immunity, existing anti-HIV-1 broadly neutralizing monoclonal antibodies are invaluable guides. By understanding the molecular details of broad neutralization of these rare antibodies, precious information regarding vulnerable sites of HIV-1 Env can be deduced and potentially mimicked in an immunogen. Anti-HIV-1 bnmAb 2F5 represents one of the very few guides available today for attempting to design an immunogen mimicking the HIV-1 gp41 MPER with three-dimensional characteristics that would allow elicitation of a broadly neutralizing antibody response upon vaccination. Before embarking on the tedious task of immunogen design, however, a detailed molecular understanding of how 2F5 recognizes its gp41 epitope and of how it inhibits HIV-1 fusion events is necessary. For this purpose, the present thesis seeks to attain the following goals:

1) Understand the molecular determinants of the breadth of 2F5 HIV-1 neutralization;

2) Refine our understanding of the 2F5-gp41 core epitope interaction and attempt to structurally characterize potential secondary interactions between 2F5 and gp41;

3) Functionally characterize 2F5 and investigate the requirement of particular regions of this rare antibody in mediating binding and neutralization.
To answer these questions, X-ray crystallography and structure determination methods will be used to provide an atomic level understanding of the 2F5-gp41 interaction. Furthermore, the functional characterization of 2F5 will be performed by using recombinant 2F5 Fab and strategically designed mutants to test their neutralization and binding capacities with various gp41-like epitopes. Altogether, the present work provides an incremental characterization of bnmAb 2F5 HIV-1 neutralization that will surely help to guide the design of immunogens seeking to elicit 2F5-like neutralizing antibodies by vaccination.
Chapter 2
Understanding the breadth of 2F5 HIV-1 neutralization


(S.B., R.C.H. and Annie Cunningham contributed to crystallization, data collection and initial structure determination of the following PDB entries reported in this chapter: 1U8H, 1U8I, 1U8J, 1U8L, 1U8M, 1U8N, 1U8O, 1U8P, 1U8Q, 1U91, 1U95 and 3IDG.)
2 Crystallographic definition of the epitope promiscuity recognized by 2F5

2.1 Introduction

Even if the characteristics of 2F5 interaction with its linear MPER consensus epitope have been previously described, a number of questions persist about the exact mechanism of 2F5 neutralization at a molecular level. One of them is to understand why 2F5 is a more potent neutralizing antibody, although its breadth across different HIV-1 isolates is more limited when compared to bnmAb 4E10 (26, 180). In an attempt to shed light on the exact molecular requirements for 2F5 recognition of its primary gp41 MPER epitope, we performed structural studies of 2F5 Fab’ with a variety of peptides. The remarkable breadth of possible 2F5 interactions reveals a somewhat surprising promiscuity of the 2F5 binding site. Furthermore, we link our structural observations with the natural variation observed within the gp41 MPER and discuss possible routes of 2F5 escape from a molecular standpoint. Finally, our discovery of 2F5’s ability to tolerate a rather broad spectrum of amino acids in its binding site, a spectrum that even includes non-natural amino acids, opens the door to new ways to design small molecule immunogens potentially capable of eliciting 2F5-like neutralizing antibodies.

2.2 Materials and Methods

2.2.1 gp41 MPER sequence alignment analysis

The Epilign sequence alignment and manipulation tool from the Los Alamos National Laboratory HIV database (4) was used to probe for gp41 MPER sequence variability. A search was performed for protein residues 659 to 683 of HIV-1 Env. It resulted in 1766 aligned sequences. A clustal W format file of all aligned sequences was then used with the program Weblogo (61, 244) to qualitatively represent variability at each amino acid position of the gp41 MPER. Subsequently, aligned sequences were manually divided by their clade information (Clade A: 94 sequences; clade B: 529 sequences; clade C: 469 sequences; clade D: 61 sequences; clade F: 19 sequences; clade G: 60 sequences; clade J: 2 sequences; clade AC: 41 sequences; clade AE: 71 sequences; clade AG: 75 sequences; clade BF: 68 sequences; clade BG: 4 sequences) and represented using Weblogo to show gp41 MPER sequence variability by clade.
2.2.2  2F5 Fab' production and crystal complex formation

2F5 IgG and 2F5 F(ab')₂ were a gift from Sanofi-Pasteur (Toronto, ON.) The bnmAb 2F5 amino acid sequence and the preparation and purification of its 2F5 Fab' fragment have been described previously (35). For the alkylation of cysteines partaking in inter-chain disulfide bonds, the 2F5 F(ab')₂ was first dialyzed in 20 mM Tris, pH 8.0 and subsequently incubated with a final concentration of 10 mM DTT for one hour at room temperature. Then, iodoacetamide was added to the solution to reach a final concentration of 25 mM and it was left to incubate for 30 minutes at room temperature. Finally, the solution was dialyzed extensively against four litres of a 20 mM Tris, pH 8.5 buffer before being concentrated for subsequent crystallization experiments.

In parallel, to develop an in-lab protocol for the digestion of 2F5 IgG into its antigen binding fragments, a series of enzymatic incubations with pepsin, papain and endopeptidase Lys-C were investigated. For these experiments, the 2F5 IgG was obtained from the NIH AIDS Research and Reference Reagent Program. For all experiments, 18 μl of 2F5 IgG at 2.5 mg/ml in the original buffer containing 2 mM acetate, pH 4.5, 10% maltose was used. For the pepsin digestion, the 2F5 IgG solution was initially supplemented with 2.5 μl of a 500 mM sodium acetate, pH 4.5 buffer, and subsequently with pepsin at a weight/weight (w/w) ratio of 1:100 or 1:1000 and incubated for 4 hours at 37°C or for 20 hours at 37°C. For the papain digestion, the 2F5 IgG solution was initially supplemented with 2.5 μl of a 500 mM phosphate, pH 6.0, 20 mM EDTA buffer and with 2.5 μl of a 100 mM cysteine solution. Subsequently, papain was added at 1:100 or 1:1000 (w/w) ratio and the resulting solution was incubated for 4 hours at 37°C or for 20 hours at 37°C. Finally, a previously published protocol for 2F5 IgG digestion was followed, which uses endoproteinase Lys-C enzyme (204). Briefly, following the reduction and alkylation protocol outlined above, the 2F5 IgG solution was dialyzed against a 50 mM Tris, pH 8.5, 1 mM EDTA buffer. Subsequently, endoproteinase Lys-C at a 1:1000 (w/w) ratio was added and the solution was incubated for 20 hours at 37°C. The results of these experiments are presented in Figure 2.1 and consist of the protein solution being analyzed by SDS-PAGE immediately following digestion, without the purification of resulting fragments. Clearly, pepsin digestion seems to yield the desired Fab’ fragment, although longer incubation time or a higher concentration of enzyme is probably required to obtain complete digestion of the 2F5 IgG. On the other hand, the papain and endoproteinase Lys-C digestions resulted mostly in the non-specific cleavage of the
IgG into many small polypeptide chains. These results support the use of pepsin for 2F5 IgG cleavage into its Fab’ fragment, which can then be used for crystallization experiments.

**Figure 2.1 Enzymatic digestion of 2F5 IgG into antigen binding fragments using various enzymes.** R+A (Reduced and alkylated).

Peptides were either gifts from A. Pedyczak and P. Chong, Sanofi-Pasteur, or obtained commercially. For crystallization, peptides dissolved in 20 mM Tris, pH 8.0 buffer at a concentration of 10 mM were added to an 8-10 mg/ml 2F5 Fab’ protein solution in the same buffer in a molar ratio ranging from 3:1 to 50:1. Crystal screening was performed using the commercially available Hampton Research Crystal Screen and Ammonium Sulfate grid screens. Initial crystal hits were identified from Crystal Screen I condition 41 (0.1 M HEPES, pH 7.5, 10% v/v 2-Propanol, 20% w/v Polyethylene glycol 4,000) and from the ammonium sulfate grid
screen condition 8 (0.1 M citric acid, pH 5.0, 1.6 M ammonium sulfate). Crystal complexes were refined using the hanging drop method at room temperature and most 2F5 Fab’-peptide crystals used for X-Ray diffraction experiments were obtained from a solution of 0.1 M citric acid, pH 5.6 and 1.6 M ammonium sulfate. As shown in Table 1, some 2F5 Fab’ – peptide mixtures did not form any crystals at all or only those of the free Fab’ under the conditions of the crystal screening procedure outlined above.

Table 2.1 List of peptide sequences (with associated MPER residue variation) that were used for co-crystallization experiments with 2F5 Fab’. Also listed are X-ray diffraction data collection and refinement statistics for each 2F5 Fab’-peptide complex that formed crystals (with PDB identifier).

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Amino acid sequence</th>
<th>Variation at position</th>
<th>Resolution range (Å)</th>
<th>Rmerge (%)</th>
<th>Complete (%)</th>
<th>Rfree (%)</th>
<th>PDB ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ELDKWAS</td>
<td>HXB2*</td>
<td>12.0–2.00</td>
<td>3.5 (31.3)</td>
<td>90.0 (93.3)</td>
<td>23.2</td>
<td>25.8</td>
</tr>
<tr>
<td>2</td>
<td>ALDKWAS</td>
<td>662</td>
<td>12.0–2.10</td>
<td>3.3 (36.8)</td>
<td>97.4 (96.9)</td>
<td>22.1</td>
<td>23.6</td>
</tr>
<tr>
<td>3</td>
<td>ELAKWAS</td>
<td>664</td>
<td>50.0–2.24</td>
<td>6.0 (38.6)</td>
<td>85.6 (83.2)</td>
<td>22.8</td>
<td>23.8</td>
</tr>
<tr>
<td>4</td>
<td>ELIKWAS</td>
<td>664</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>ELNKWAS</td>
<td>664</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>ELQKWAS</td>
<td>664</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>ELDWAS</td>
<td>665</td>
<td>17.0–2.60</td>
<td>8.8 (35.6)</td>
<td>88.3 (88.0)</td>
<td>22.3</td>
<td>24.9</td>
</tr>
<tr>
<td>8</td>
<td>ELDDWAS</td>
<td>662, 665</td>
<td>80.0–2.24</td>
<td>5.6 (24.6)</td>
<td>96.9 (100)</td>
<td>21.0</td>
<td>23.9</td>
</tr>
<tr>
<td>9</td>
<td>ELDWiHAS</td>
<td>665</td>
<td>80.0–2.24</td>
<td>4.6 (11.8)</td>
<td>91.3 (97.4)</td>
<td>22.9</td>
<td>22.7</td>
</tr>
<tr>
<td>10</td>
<td>ELD(Ome)WAS</td>
<td>665</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ELD(Nrg)WAS</td>
<td>665</td>
<td>80.0–2.24</td>
<td>5.7 (31.3)</td>
<td>93.2 (99.9)</td>
<td>22.9</td>
<td>23.6</td>
</tr>
<tr>
<td>12</td>
<td>ELD(Pal)WAS</td>
<td>665</td>
<td>50.0–2.25</td>
<td>6.3 (23.6)</td>
<td>97.4 (99.1)</td>
<td>22.0</td>
<td>23.8</td>
</tr>
<tr>
<td>13</td>
<td>ELDEWAS</td>
<td>665</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>ELDKAAS</td>
<td>666</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>ELDKYAS</td>
<td>666</td>
<td>50.0–2.40</td>
<td>5.4 (36.4)</td>
<td>99.2 (96.3)</td>
<td>21.8</td>
<td>22.6</td>
</tr>
<tr>
<td>16</td>
<td>ELDKFAS</td>
<td>666</td>
<td>50.0–2.56</td>
<td>9.6 (31.1)</td>
<td>92.2 (93.9)</td>
<td>20.5</td>
<td>23.1</td>
</tr>
<tr>
<td>17</td>
<td>ELDKHAS</td>
<td>666</td>
<td>20.0–3.02</td>
<td>9.5 (34.7)</td>
<td>94.8 (93.7)</td>
<td>21.1</td>
<td>21.4</td>
</tr>
<tr>
<td>18</td>
<td>ELDKIAS</td>
<td>666</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>ELDKLAS</td>
<td>666</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>ELDK(Nic)AS</td>
<td>666</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>ALDKWQN</td>
<td>662, 667, 668</td>
<td>30.0–2.10</td>
<td>8.3 (32.2)</td>
<td>99.9 (99.8)</td>
<td>21.2</td>
<td>23.0</td>
</tr>
<tr>
<td>22</td>
<td>ALDKWD</td>
<td>662, 667</td>
<td>20.0–3.86</td>
<td>4.2 (29.9)</td>
<td>98.5 (99.5)</td>
<td>22.4</td>
<td>22.7</td>
</tr>
<tr>
<td>23</td>
<td>ELDKWNSL</td>
<td>667</td>
<td>50.0–2.60</td>
<td>13.5 (35.2)</td>
<td>95.8 (90.3)</td>
<td>21.3</td>
<td>24.0</td>
</tr>
<tr>
<td>24</td>
<td>ELDKWKL</td>
<td>667</td>
<td>50.0–2.55</td>
<td>22.0 (30.0)</td>
<td>99.0 (98.5)</td>
<td>21.8</td>
<td>23.7</td>
</tr>
<tr>
<td>25</td>
<td>ELDKWAN</td>
<td>668</td>
<td>17.0–2.00</td>
<td>5.9 (39.9)</td>
<td>97.1 (89.0)</td>
<td>23.4</td>
<td>24.7</td>
</tr>
<tr>
<td>26</td>
<td>ELDKWAG</td>
<td>668</td>
<td>80.0–2.24</td>
<td>13.6 (27.3)</td>
<td>91.3 (96.5)</td>
<td>22.9</td>
<td>23.8</td>
</tr>
<tr>
<td>27</td>
<td>ECDKWC</td>
<td>663, 667</td>
<td>20.0–3.23</td>
<td>12.6 (20.2)</td>
<td>97.2 (98.7)</td>
<td>22.2</td>
<td>22.5</td>
</tr>
<tr>
<td>28</td>
<td>E(Dap)DKWES</td>
<td>663, 667</td>
<td>80.0–2.24</td>
<td>10.9 (24.8)</td>
<td>94.0 (87.9)</td>
<td>22.5</td>
<td>24.9</td>
</tr>
<tr>
<td>29</td>
<td>E(Dap)DKWS</td>
<td>663, 667</td>
<td>80.0–2.24</td>
<td>9.4 (20.0)</td>
<td>91.4 (88.1)</td>
<td>22.7</td>
<td>23.1</td>
</tr>
<tr>
<td>30</td>
<td>E(Dap)DKW(S)</td>
<td>663, 667</td>
<td>80.0–2.37</td>
<td>9.6 (34.5)</td>
<td>92.0 (95.3)</td>
<td>25.0</td>
<td>23.6</td>
</tr>
<tr>
<td>31</td>
<td>E(Dap)DKWDS</td>
<td>663, 667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>KLDKWAS</td>
<td>662, 665</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>ALDKWNNS</td>
<td>662, 665, 667, 668</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Collection and refinement statistics are given for each 2F5 Fab’-peptide complex that formed crystals along with the PDB identifier (ID). The absence of data indicates that no crystal complex was obtained. Values in parentheses are those for the highest-resolution bin.

* Reference sequence.
2.2.3 X-Ray diffraction analysis

All diffraction data were collected on crystals cooled to 100 K after soaking in a cryoprotectant of 25% glycerol mixed with mother liquor on a home source Rigaku FR-C rotating copper anode with a Mar345 detector and/or at NSLS synchrotron station X8C. Data were processed with the program packages DENZO and SCALEPACK (206). Most 2F5 Fab’ - peptide complexes adopted space group P2_12_12_1 (a~59.0 Å; b~65.0 Å; c~175.6 Å) and grew as rectangular prisms. Due to the long c-axis, high and low resolution data sets were collected and subsequently merged. The structure of the 2F5 Fab’ – peptide 1 crystal complex was determined by molecular replacement using the program AMoRe with the PDB entry 1CLZ as the search model (193). The constant and variable regions were used as independent models. The correct solution showed one Fab’ molecule per asymmetric unit and a correlation coefficient of 35.3 (R= 47.3%) using data to 3.3 Å. The CNS package was used for refinement (34). Real space refinement was done using the programs O and Coot (81, 132). Density for the peptides was clear and could be fitted unambiguously. After performing numerous cycles of real-space, positional and B-factor refinements, waters were included where peaks of > 3.5 σ were found in a difference map at an appropriate distance from a donor or acceptor atom. Most peptide complexes gave crystals isomorphous and isostructural to the 2F5 Fab’ – peptide 1 complex crystals. Figure images were generated using the program PyMol (70).

2.3 Results

2.3.1 gp41 MPER variation

Table 2.2 qualitatively depicts the sequence variation of the HIV-1 gp41 MPER. 15 amino acid positions out of the 25 contained in this region are conserved and show little to no variation: L_660, L_661, L_663, D_664, W_666, L_669, W_670, W_672, F_673, I_675, W_678, L_679, W_680, Y_681 and L_682. Except for aspartic acid at position 664 all conserved positions are occupied by non-polar or hydrophobic residues. Interestingly, one third of the conserved positions are tryptophan residues. Looking more closely, it is clear that sequence variation in the gp41 MPER is highly influenced by HIV-1 clade. Notably, position 662 is particularly variant in clades B and D, whereas position 665 varies significantly in clades C and D. Moreover, position 667 accommodates different residues in clade C and position 677 varies considerably in clade D.
Table 2.2 Natural sequence variation in HIV MPER and 2F5 neutralization (reported by Binley et al., J Virol 2004) across different clades.

<table>
<thead>
<tr>
<th></th>
<th>2F5 Neutralization</th>
</tr>
</thead>
<tbody>
<tr>
<td>All clades</td>
<td>60/90 viruses</td>
</tr>
<tr>
<td></td>
<td>67%</td>
</tr>
<tr>
<td>Clade A</td>
<td>11/12 viruses</td>
</tr>
<tr>
<td></td>
<td>92%</td>
</tr>
<tr>
<td>Clade B</td>
<td>23/29 viruses</td>
</tr>
<tr>
<td></td>
<td>79%</td>
</tr>
<tr>
<td>Clade C</td>
<td>0/12 viruses</td>
</tr>
<tr>
<td></td>
<td>0%</td>
</tr>
<tr>
<td>Clade D</td>
<td>5/11 viruses</td>
</tr>
<tr>
<td></td>
<td>45%</td>
</tr>
<tr>
<td>Clade F</td>
<td>3/3 viruses</td>
</tr>
<tr>
<td></td>
<td>100%</td>
</tr>
</tbody>
</table>
The primary epitopes of bnAbs 2F5, Z13 and 4E10 have been mapped to the gp41 MPER. A buried surface area analysis of gp41 residues from the crystal structures of 2F5 (3D0L), 4E10 (2FX7) and Z13 (3FN0) reveals that the five residues with the largest buried surfaces are L<sup>661</sup>, D<sup>664</sup>, K<sup>665</sup>, W<sup>666</sup> and W<sup>670</sup> for 2F5, W<sup>672</sup>, F<sup>673</sup>, T<sup>676</sup>, L<sup>679</sup> and W<sup>680</sup> for 4E10 and W<sup>670</sup>, N<sup>671</sup>, D<sup>674</sup>, I<sup>675</sup> and T<sup>676</sup> for Z13 (47, 134, 212). Clearly, the core residues recognized by 2F5 and 4E10 are...
more conserved than those recognized by Z13, which helps to explain the increased neutralization breadth of the former pair of bnmAbs (194, 306). In the case of 2F5, core residues are generally conserved, although significant variation is observed at position K_{665} in the case of HIV-1 clades B, C and D.

2.3.2 X-Ray diffraction experiments

Table 2.1 contains a list of 7-mer peptides that were tested for crystal complex formation with 2F5 Fab’ and statistics of data collection, processing and structure refinement for these peptides that formed complex crystals. Peptide 1 (ELDKWAS), represents the consensus HXB2 reference of the primary 2F5 epitope amino acid sequence from HIV protein gp41 (192). The structure of the complex of this peptide with 2F5 Fab’ was previously reported and described in detail (134, 204, 208). Briefly, the antibody was shown to recognize the gp41 MPER in the conformation of a type I β-turn centered on the DKW core residues. The present study has as its objective to investigate how tight the spatial restrictions are that govern the interaction between the gp41 core epitope and the corresponding 2F5 paratope. First, alanine substitutions were made at all positions of the ELDKWAS peptide. Subsequently, if such a substitution was tolerated and a crystal complex could be obtained, substitutions with amino acids found with a high frequency in the gp41 MPER were attempted. When alanine substitutions were not tolerated, only conservative substitutions were investigated. Figure 2.2 is an overlay of all peptides reported in this study and depicts the central β-turn conformation adopted by residues at position 664 through 666.
Figure 2.2 Overlay of all MPER representing peptides (except for constrained peptides) characterized in complex with 2F5 Fab’ reported in the present structural study. All peptides are seen to adopt a type-1 β-turn conformation. Peptides are shown in gray as a cartoon representation with side chain positions shown as stick models. An electrostatic potential of the 2F5 Fab’ paratope is shown in the background with red, blue and white regions representing negatively charged, positively charged and non-charged areas, respectively.

2.3.2.1 Asp664 substitutions

To study the extent of variability allowed in the DKW core itself, we started our analysis by investigating peptides carrying mutations in this core. Converting the Asp664 side chain to a methyl group (peptide 3, ELAKWAS) prevented complex formation with 2F5, consistent with the lack of observable binding found in alanine scanning mutagenesis studies on gp41 peptides (182). Similarly, substitutions D664N (peptide 5, ELNKWAS) and D664Q (peptide 6, ELQKWAS) did not form complexes with 2F5 Fab’. A crystal complex did form for peptide 4, ELEKWAS, which has the conservative D664E substitution (Figure 2.3). To achieve Fab’-peptide complex formation, the concentration of this peptide had to be 10-fold higher than the
standard used e.g. with the consensus ELDKWAS peptide. The structure of the ELEKWAS complex reveals that the side chain carboxylate makes the same polar contacts as aspartate with residues Arg(H)-95, His(L)-96 of the antibody, the amide bond of the peptide’s Trp666 and an adjacent water molecule. In addition, a slight shift in the glutamate side chain position compared to the aspartate side chain allows it to come into closer proximity with residues Tyr(H)-52 and Tyr(L)-94 of the antibody. Furthermore, the backbone of the peptide does not overlay exactly with reference peptide 1 and the B-factor of the α-carbon atom is 70.1 Å² as opposed to 23.7 Å² when aspartate is present. The entropic penalty paid to stabilize the longer glutamic acid side chain and especially limits on space leading to molecular crowding may be responsible for the reduced affinity of this peptide. These results indicate that the correct positioning of the negative charge at position 664 is essential for epitope recognition.

Figure 2.3 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions at position 664 of the MPER. The reference peptide ELDKWAS is shown in light gray whereas peptide ELEKWAS is shown in green. A white rendering of the 2F5 Fab’ paratope surface is shown in the background.
2.3.2.2 Lys665 substitutions

In a similar fashion to the Asp664 substitutions, we synthesized a peptide changing Lys665 to an alanine residue (peptide 7, ELDAWAS). Again, we were unable to generate a complex of this peptide with 2F5 Fab'. Replacing position 665 with a negatively charged glutamic acid (peptide 13, ELDEWAS) also failed to produce complex crystals.

The conservative substitution of Lys665 with arginine in peptide 8 (DLDRWAS) led to complex crystal formation with 2F5 Fab' (Figure 2.4B). As expected, the extended alkyl chain of arginine occupies the same space as that of lysine from the reference peptide, forming hydrophobic alkyl-π stacking interactions with the peptide tryptophan and heavy chain Tyr(H)-52. The positive charge at the end of the arginine side chain interacts in the same way with the Asp(H)-54 carboxylate as the corresponding positive charge of lysine but is no longer able to interact with Asp(H)-56. It should also be noted that two water molecules making hydrogen bonding interactions near Lys665 observed in the reference peptide structure are absent in the Fab' - DLDRWAS structure. Another allowed substitution is a histidine residue in place of Lys665 (peptide 9, ELDHWAS) (Figure 2.4D). Crystallization of this peptide with 2F5 Fab' was done using sodium citrate, pH 5.6 and at this pH, the histidine is protonated and positively charged making contact with the Asp(H)-56 side chain carboxylate. Alternate rotamer conformations of the imidazole ring could also lead to the positive charge making contacts with Asp(H)-54, which is also an interaction partner for Lys665 in the reference peptide. Another interesting feature in this structure is the tight semi-perpendicular stacking between the imidazole and indole ring of the peptide (the distance between the imidazole δ carbon and the indole carbons of the six-membered ring varies between 3.3Å and 3.8 Å) and partially parallel stacking arrangement with Tyr(H)-52 of the Fab’ (the distance between the histidine β,γ and ε carbons and the phenol ring carbons varies between 3.7 Å and 4.1 Å).
Figure 2.4 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions at position 665 of the MPER. The reference peptide ELDKWAS is shown in light gray whereas peptides A) ELD(Orn)WAS, B) ELDRWAS, C) ELD(Nrg)WAS, D) ELDHWAS and E) ELD(Paf)WAS are shown in green. A white rendering of the 2F5 Fab’ paratope surface is shown in the background.
Ornithine is an analogue of lysine but its alkyl chain is one methylene group shorter than that of lysine. When lysine was replaced with ornithine at position 665 (peptide10, ELD[Orn]WAS) a new water network formed around position 665 (Figure 2.4A). In the ELDKWAS structure, the positively charged amino group hydrogen bonds directly with the negatively charged side chains of 2F5 Fab’ residues Asp(H)-54 and Asp(H)-56. When ornithine is substituted for lysine, a water molecule now forms a bridge between the amino group at the tip of the shorter side chain of ornithine and Asp(H)-54 as well as Asp(H)-56. Two additional water molecules also create secondary bridging between the retracted ornithine amino group and side chains of the Fab’, including Ser(H)-53.

Nitroarginine corresponds to an extended arginine that carries an additional nitro group on its guanidinium group. In peptide 11 (ELD[Nrg]WAS), Lys665 was substituted with nitroarginine (Figure 2.4C). Mirroring the results of the closely related K665R substitution, the positively charged guanidinium group again forms a hydrogen bond with the negatively charged Asp(H)-54 side chain and stacking interactions with the surrounding ring systems. As in the case of the K665R substitution, however, no direct contact is observed to the side chain of Asp(H)-56. Furthermore, the additional negatively charged nitro group forms a hydrogen bond with Ser(H)-53 as well as the main chain peptide nitrogen of Gly(H)-33. The nitro group is now stationed in the same region previously occupied by two water molecules in the K665Ornithine substituted peptide.

p-Aminophenylalanine is a non-natural amino acid that was put in place of Lys665 in peptide 12 (ELDPafWAS), for which a crystal complex was obtained (Figure 2.4E). The structure revealed that the plane of the cyclic side chain is exactly superimposable with the imidazole ring of histidine in peptide 9 (ELDHWAS, K665H substitution). The amino group of this side chain now interacts with Asp(H)-54 and is bridged to Ser(H)-53 via an additional water molecule, in the same way as was seen in peptide 10 (ELD[Orn]WAS). Taken together, these results suggest that the 2F5 paratope can accommodate alternative interactions in the 665 binding region, as long as key hydrophobic stacking interactions and hydrogen bonding networks are conserved. However, no substitutes at position 665 reproduce exactly the interactions observed between 2F5 and Lys665.
2.3.2.3  Trp666 substitutions

Consistent with alanine scanning mutagenesis studies, peptide 14 (ELDKAAS) did not form a crystal complex (182). We synthesized a series of peptides with conservative amino acid substitutions at position 666: peptide 15 (ELDKYAS), peptide 16 (ELDKFAS), peptide 17 (ELDKHAS), peptide 18 (ELDKIAS), peptide 19 (ELDKLAS) and a non-natural amino acid substitution with norleucine (peptide 20) ELDK[Nle]AS. Of these, only peptides substituted with amino acids carrying aromatic side chains formed complexes with 2F5 Fab' (Figure 2.5A, 2.5B and 2.5C respectively) while substitutions with amino acids with hydrophobic alkyl side chains, i.e. isoleucine, leucine and norleucine, did not. In the three structures for which a crystal complex was obtained, the different rings all occupy the same space as the indole ring plane of Trp666. Hydrophobic alkyl-π stacking interactions are observed between the ring at position 666 and the Lys665 alkyl side chain, as well as van der Waals contacts with Pro(H)-98 and Val(H)-100K of 2F5. Furthermore, the position of a key water molecule located below the indole ring in the reference structure is conserved in all three structures. In the reference peptide structure, this water molecule bridges the ε nitrogen of the Trp666 indole ring with the backbone carbonyl Gly(H)-33.

In the case of the ELDKYAS peptide, the terminal tyrosine hydroxyl interacts directly with the backbone carbonyl of Gly(H)-33 as well as with the conserved water molecule below the ring. Interestingly, for the W666F substitution, the conserved water molecule is observed even though phenylalanine lacks a polar functional group to coordinate via hydrogen bonding. Finally, substitution of the Trp666 position with histidine required a buffer with a pH >7.0 to form a crystal complex with 2F5 Fab'. Obviously, a histidine residue is only tolerated in position 666 when uncharged. In this structure, the uncharged imidazole ring occupies a similar position as the indole, phenyl and phenol rings and makes a similar water-bridged contact to Gly(H)-33. Also of notice in this structure, the Asp664 side chain is slightly shifted towards the imidazole ring, making tighter contacts with the peptide backbone amide. Depending on the rotamer conformation of histidine, the Asp664 side chain could also directly interact via hydrogen bonding with the side chain of His666.
Figure 2.5 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions at position 666 of the MPER. The reference peptide ELDKWAS is shown in light gray whereas peptides A) ELDKYAS, B) ELDKFAS and C) ELDKHAS are shown in green. A white rendering of the 2F5 Fab’ paratope surface is shown in the background.

These results indicate that hydrophobic alkyl-π stacking interactions between a planar aromatic residue at position 666 and the extended alkyl chain of the previous residue at position 665 is an important structural feature of the DKW core β-turn and key to the stability of the epitope-2F5 Fab’ interaction. Linear alkyl side chains (norleucine) and branched alkyl side chains (isoleucine, leucine) are unable to form tight stable van der Waals interactions with the lysine side chain and residues of the 2F5 paratope, hence disrupting β-turn stability.
2.3.2.4 Substitutions outside of the DKW core

In the ELDKWAS reference structure, Glu662 interacts via hydrogen bonding with Arg(H)-58 and two water molecules. In the case of the D662A substitution (peptide 2, ALDKWAS), the alanine residue lacks any interaction with Arg(H)-58, as was expected. The loss of the glutamate side chain leads to a rearrangement in the water network around position 662. In particular, a water molecule appears in place of the glutamate functional group to interact with Arg(H)-58.

Leu663 is invariant across HIV-1 isolates and 2F5 interacts with conserved Leu663 through hydrophobic van der Waals contacts with residues His(L)-92 and Phe(L)-93 (204). Therefore, this position was only changed to residues that allowed the cyclization of the peptide in our efforts to favour β-turn formation, as described below.

Position 667 is particularly interesting to study at the molecular level because it shows significant variability across viruses of different clades. For peptide 21 (ALDKWQN), clear electron density is observed for residues Ala662 through Trp666 but quickly fades after that with no interpretable electron density for residues Gln667 and Asn668, which suggests that these amino acids do not adopt a unique stable conformation (Figure 2.6B). This could indicate an inability of these residues to make stabilizing interactions with the antibody or with other components of the linear sequence, such as Leu663. To further investigate the role of varying amino acids at position 667, peptides 22 (ALDKWD), 23 (ELDKWNSL) and 24 (ELDKWKSL) were synthesized. In the case of ALDKWD, the side chain of Asp667 interacts with His(L)-92 and Arg(H)-100H (Figure 2.6C). It replaces a water molecule interacting with His(L)-92 in the reference ELDKWAS structure. As previously documented, in the present P2₁2₁2₁ crystal system, the position of Arg(H)-100H is determined by crystal lattice interactions. Therefore, the interaction between Asp667 and Arg(H)-100H might not be representative of the interaction observed in a biological context (see reference (134) for complete discussion). Nevertheless, this crystal structure documents that 2F5 is able to accommodate an aspartic acid at position 667. When 2F5 Fab’ is crystallized in complex with the peptide ELDKWNSL, electron density for all residues is observed (Figure 2.6D). The asparagine side chain, however, does not engage in the same contacts with residues of the Fab’ that aspartate had formed. Instead, it packs tightly against Leu663, at a distance of ~3.5 Å and makes a hydrogen bond interaction with a conserved water, which in turn coordinates with Arg(H)-100H.
Figure 2.6 Crystal structures of 2F5 Fab’ in complex with peptides probing substitutions outside the DKW core, namely positions 662, 667 and 668 of the MPER. The reference peptide ELDKwas is shown in light gray whereas peptides A) ALDKWAS B) ALDKWQN, C) ALDKWD, D) ELDKWNLS, E) ELDKWKSL, F) ELDKWAN and G) ELDKWAG are shown in green. A white rendering of the 2F5 Fab’ paratope surface is shown in the background. For substitutions at position 667, a Fo-Fc electron density composite omit map at a 2.5 \( \sigma \) level are represented in a blue mesh.
This is a significant decrease in distance between the 663 and 667 positions, which is \(~4.0~\text{Å}\) in the reference ELDKWAS peptide. It is possible to see from the electron density that Leu663 moves back slightly as compared to the reference structure in order to make accommodations for a bulkier side chain at position 667. As well, this structure shows a significant difference in the direction of the C-terminal residues, which now fold back towards Trp666. This can be attributed to the addition of Leu669 to the peptide, which has been previously described to be part of a canonical \(\alpha\)-helical turn when the peptide epitope is extended \((134)\). Finally, when a lysine residue is introduced at position 667, no clear electron density is observed for residues Lys667, Ser668 and Leu669 (Figure 2.6E). This is similar to the structure of the Fab’ with peptide 21 (ALDKWQN), with the bulky glutamine amino acid at position 667. These data suggest that a residue with a bulky side chain at position 667 might prevent residues located at the C-terminus of the DKW core from adopting the proper conformation recognized by 2F5 because of clashing interactions, particularly with Leu663.

In the reference sequence, position 668 is occupied by serine. Previous reports demonstrated that Ser668 does not interact with 2F5 \((204)\). Substitutions at this position are therefore not expected to have major effects on the overall peptide structure as recognized by 2F5. The most commonly observed substitution at this position is S668N. Crystal structures of 2F5 Fab’ in complex with peptides 25 (ELDKWAN) and 26 (ELDKWAG) revealed no difference in core binding (Figures 2.6F and 2.6G). It is worth mentioning that the asparagine residue of peptide 25 forms hydrogen bonds to Arg(H)-100H, providing this peptide with additional interactions with the Fab’, as compared to the reference peptide. However, as was previously mentioned, this interaction with Arg(H)-100H is probably compromised by crystal lattice interactions.

### 2.3.2.5 Cyclic peptides

In an attempt to constrain the \(\beta\)-turn motif, peptides 27 (ECDKWCS), 28 (E(Dap)DKWES), 29 (EDDKW(Dap)S), 30 (EEDKW(Dap)S) and 31 (E(Dap)DKWDS) were synthesized and mixed with 2F5 Fab’. All but the E(Dap)DKWDS peptide formed crystal complexes. Peptide ECDKWCS has Leu663 and Ala667 substituted with cysteine residues \((i\text{ and } i+4\), respectively). Under oxidative conditions, a disulfide bond forms to constrain the \(\beta\)-turn. Figure 2.7A shows how the essential \(\beta\)-turn conformation and contacts seen in reference peptide 1 are conserved.
Figure 2.7 Crystal structures of 2F5 Fab’ in complex with peptides harboring β-turn stabilizing linkages between positions 663 and 667 of the MPER. The reference peptide ELDKWAS is shown in light gray whereas peptides A) ECDKWCS, B) ENDKW(Dap)S, C) EQDKW(Dap)S and D) E(Dap)DKWQS are shown in green. A white rendering of the 2F5 Fab’ paratope surface is shown in the background.

An alternative method to constrain the β-turn was to create a lactam bridge between residues at the same position. This was done by incorporating an aspartate (peptide 29) and glutamate (peptide 30) residue in the i position and an L-α-diaminopropionic acid (Dap) in the i+4 position, reacting to generate a lactam bridge. Figures 2.7B and 2.7C show the resulting complexes with 2F5 Fab’ using the aspartate and glutamate residue in the i position, respectively. The DKW core residues of both peptides superimpose in the same positions as reference peptide 1 without structural variation demonstrating that the lactam bridge is sufficient in constraining the β-turn conformation. The extra methylene group (γ-carbon) of the glutamate side chain does not adversely affect the conformation as it bulges out away from the Fab’ while the Dap β-carbon is
in the same position as the native alanine β-carbon. Furthermore, in these structures an additional water molecule is observed interacting with the carbonyl of the lactam bond.

When reversing the residues such that the Dap group is in the i position and a glutamate is at the i+4 position, a crystal complex forms with 2F5 Fab' (Peptide 28, Figure 2.7D). In this structure the α-carbon of the glutamate is shifted away from the optimal positions of the native alanine structure and unlike peptides 29 and 30, the methylene groups of the side chain point towards the Fab' body. The use of an aspartate residue at the i+4 position (E(Dap)DKWDS) did not allow for complex formation with 2F5 Fab'. The shorter side chain coupled with the backbone shift probably generated a highly strained peptide unable to maintain a stable β-turn and bind properly to 2F5 Fab'.

2.4 Discussion

Conformational flexibility of the gp41 MPER (α-helix, kinked α-helix, 3_10-helix, β-turn) has been suggested to be a characteristic feature of gp41 and to play an important role in HIV membrane fusion (28, 134, 204, 208, 212, 243, 264). This dynamic aspect of the MPER, in addition to its amphipathic properties, its ability to oligomerize and to interact with other proteins will put a significant constraint on its amino acid composition (9, 10, 187, 230, 261, 262, 297). Indeed, sequence variation within the MPER (especially for residues L_{660}, L_{661}, L_{663}, D_{664}, W_{666}, L_{669}, W_{670}, W_{672}, F_{673}, I_{675}, W_{678}, L_{679}, W_{680}, Y_{681} and I_{682}), is minimal, especially when compared to other regions of the Env protein. The relatively conserved nature of the MPER sequence helps to explain in part its recognition by three bnAbs with epitopes on the gp41 MPER (2F5, Z13 and 4E10). Furthermore, the neutralizing breadth of these three bnAbs (4E10 > 2F5 > Z13) seems to correlate directly with the conserved nature of the linear stretch of amino acids they recognize.

As shown in this study as well as in previous reports, 2F5 recognizes a core 664DKW666 β-turn motif (134, 204, 208). gp41 sequence analysis revealed that the aspartic acid and tryptophan residues of the 664DKW666 core are highly invariant and completely conserved, respectively, among all HIV isolates (Table 2.2). The crystal structures presented in this study expose a strict requirement for a negative charge at position 664. Indeed, no crystal complex was isolated when the peptide epitope had D664A, D664N or D664Q substitutions. Moreover, even though the D664E substitution yielded a complex crystal, the longer side chain of glutamate seems to
destabilize the peptide conformation (higher overall B factors for the peptide) with a loss in affinity as evidenced by the much higher concentration of peptide required to obtain complex crystal growth. Overall, it seems clear that 2F5 matured for an ability to recognize an invariant aspartic acid at position 664. Position 666 is able to accommodate residues with planar side chains (such as tyrosine, phenylalanine and histidine) but no crystal complex forms with residues possessing branched non-polar side chains at this position. Clearly, this indicates the requirement for specific hydrophobic alkyl-π stacking interactions between the planar side chain at position 666 with an extended alkyl chain at position 665, as well as van der Waals contacts with Pro(H)-98 and Val(H)-100K of 2F5. Although the conserved nature of positions 664 and 666 do not require 2F5 to be promiscuous in its recognition, the insights gained from the structures of 2F5 Fab’ with peptides having substitutions at these positions are useful for understanding key binding interactions and should help in immunogen design efforts.

On the other hand lysine 665, another residue participating in the β-turn motif recognized by 2F5, varies considerably in the known HIV-1 sequences. Overall, lysine 665 is seen to be most commonly substituted by serine, glutamine or threonine residues, with serine being the prevalent residue at position 665 in HIV-1 clade C isolates. The present study does not report on the structural determination of peptides with these substitutions at position 665. However, it is clear from the interaction between the positively charged side chain amine of lysine 665 and the two negatively charged aspartate residues (H)-54 and (H)-56 at the apex of the CDR-H2 that the recognition of 2F5 for its epitope relies heavily on these ionic interactions. In addition to the lack of a positive charge, the presence of a serine, glutamine or threonine residue at position 665 could lead to the disruption of proper hydrophobic stacking interactions with Trp666 and hence a destabilization of the β-turn. Therefore, we suggest that one of the major limitations of 2F5 neutralization breadth resides in its dependency for an extended positively charged residue at variable position 665.

All peptides with 665 substitutions that were able to form crystal complexes possessed both an extended chain capable of hydrophobic stacking interactions as well as a nitrogen containing functional group, positively charged in most structures, capable of interacting with either Asp(H)-54 and Asp(H)-56 of 2F5. It may be promising that compounds other than naturally occurring amino acids can be used to mimic the extended lysine and tryptophan conformations. Such peptide mimetics perhaps can be a foundation for the synthesis of a synthetic small
molecule vaccine. However, none of the substitutes used at position 665 in this study were able to recreate the exact interaction of a positive charge with both Asp(H)-54 and Asp(H)-56 and these peptides might therefore show limitations in their ability to elicit 2F5-like antibodies.

Position 662 is quite variable in gp41 MPER and is predominantly occupied in HIV-1 isolates by alanine and glutamic acid residues. The ability of 2F5 to make hydrogen bonding interactions with the carboxylate of Glu662 via Arg(H)-58 and Tyr(L)-94 suggests that this antibody has matured to interact with a negative charge at this position. However, the present crystal structure of 2F5 Fab’ in complex with a peptide harbouring the E662A substitution reveals that the antibody is able to adapt to the loss of the functional group at position 662. This analysis corresponds to neutralization assays which show that 2F5 is able to neutralize effectively isolates with mutations at position 662, although the efficacy at which it does so might be reduced (26).

Finally, position 667 is most commonly occupied by a small alanine residue, although it is particularly variable in clade C and F isolates where residues like lysine, asparagine, glutamine and aspartic acid are often observed. The crystal structures presented here reveal the ability of 2F5 to adapt its binding to accommodate for A667D and A667N substitutions, but that substitutions with residues possessing elongated side chains such as lysine and glutamine seem to disrupt proper binding at the C-terminus of the DKW core. As reported in Binley et al. (26) most viruses able to evade neutralization by 2F5 have mutations in the linear DKW core. However, it has also been previously observed that sequences with a conserved DKW motif but possessing 667 substitutions such as isolate SG364 (662ALDKWNQ668) are able to escape 2F5 neutralization (26). Combined with our structural studies these observations lead us to propose that a secondary 2F5 neutralization escape mechanism involves substitution from an alanine at position 667 to a residue with a bulky side chain, as is often observed in HIV-1 clade C isolates.

2F5 being such a potent HIV-1 neutralizing antibody for isolates possessing the LDKWA core, many efforts have been put forward to create immunogens possessing a constrained β-turn motif of this sequence in order to elicit 2F5-like antibodies (14, 54, 80, 83, 163). One of the approaches to constrain a β-turn conformation involves creating a connection between the i and i + 4 β-turn residues in a peptide, namely Leu-663 and Ala-667, either via disulfide linkage or lactam covalent bonds. The present crystal structures of such constrained peptides show the ability of 2F5 to interact with such peptides in a way that reproduces the interactions observed in the native peptide. In previous studies, these β-turn constrained gp41 peptides were tested to determine
whether they were sufficient to elicit antibodies capable of neutralizing HIV-1; however, despite generating very high titer antisera, HIV-1 neutralization was not observed (179). NMR analysis revealed that these constrained peptides formed a stable \( \beta \)-turn conformation in solution (179). However, the specific conformations of the DKW side chains that are adopted in solution are probably quite different than those found in the crystal structure. Indeed, a direct cation-\( \pi \) interaction between the terminal amino group of the Lys665 and Trp666 in solution differs from the stacking interaction observed in the crystal structure whereby the methylene groups of the extended Lys665 are stacked in an alkyl-\( \pi \) fashion against the Trp666 indole ring. This suggests that both proper DKW \( \beta \)-turn conformation and DKW side chain positions are crucial for designing immunogens capable of eliciting 2F5-like neutralizing antibodies.

Our structural analysis reveals multiple water molecules undergoing key interactions with the 2F5 paratope and the gp41 peptide epitope. Of particular interest are two conserved water molecules that mediate the binding of the gp41 epitope to the 2F5 paratope in all structures; one of them is located below the indole nitrogen of Trp666 and makes hydrogen bonding interactions with the backbone carbonyls of 2F5 Gly(H)-33 and Arg(H)-96 and the second one is placed next to Asp664 and hydrogen bonds to the 2F5 backbone carbonyl of Leu(L)-91 and the side chain NH group of Asn(H)-100L. These findings suggest that a synthetic immunogen looking to elicit 2F5-like neutralizing antibodies could take advantage of these water-binding sites by incorporating water displacement in its design, therefore decreasing the entropy of binding while maintaining the interaction enthalpy (Figure 2.8).
Figure 2.8 Representation of spatial and chemical requirements for MPER residues for 2F5 core recognition as deduced from the present complex crystal structures. The wall-eyed stereo representation is a translucent Connolly molecular surface coloured by group hydrophobicity (with a spectrum from red to blue showing hydrophobic to hydrophilic areas, respectively and white representing two conserved water molecules.) Also, a ball and stick representation shows the position of crucial nitrogen (blue), oxygen (red) and carbon (gray) atoms. The model design was based on 1) the conserved negative and positive charges present at positions 664 and 665, respectively, 2) the requirement for an aromatic ring at position 666, 3) the conserved alkyl functional group at position 663, 4) important amide and carbonyl atoms taking part in direct interactions with 2F5 and 5) two conserved water molecules mediating the epitope/paratope interaction. This visualization was generated using the program ChemBio 3D by CambridgeSoft.

While it remains uncertain why the MPER β-turn conformation is required for successful HIV fusion to CD4+ T-cells, this MPER motif has been the target of much scrutiny for vaccine design. The current study highlights the importance in immunogen design of the correct positioning of the negative charge of Asp664 and the hydrophobic interactions of Trp666, from which 2F5 gets most of its neutralization potency and breadth. Positions 665 and 667 were identified at the molecular level as a particular determinant of 2F5 neutralization potency but also of neutralization escape. We suggest that evolving 2F5 in vitro to rely less on variable residues and more on conserved MPER residues (such as Leu661 or Leu663) might be a way to increase both neutralization breath and potency, as was described for bnmAb Z13 and Z13e1 (194). Overall, recognition of the MPER by bnmAb 2F5 is characterized in this study to be partly promiscuous. Such an ability for antibodies to be slightly multi-specific in their recognition has been previously reported, and might confer antibodies a significant advantage for the recognition
of variable antigens (127, 128). Finally, we emphasize that other components of the 2F5 interaction with HIV-1 \textit{in vivo}, such as possible interactions of CDR H3 residues with membrane components, albeit very challenging, nevertheless, might have to be included in a successful immunogen. However, the correct presentation of the main determinant of 2F5 binding, the DKW core, should remain the priority in the design of immunogens intended to elicit 2F5-like neutralizing antibodies.
Chapter 3
Structurally refining the understanding of 2F5-gp41 core epitope interaction


(J.L.N. contributed significantly in the design of experiments described in this chapter. Rosy Hynes solved the crystal structure related to PDB ID: 2F5B presented here.)
3 Epitope conformation, antigen-recognition loop mobility and anion binding site.

3.1 Introduction

bnmAbs 2F5, 4E10 and Z13 bind to linear stretches of amino acids (304). The core sequence $^670\text{WNWFDITNW}^678^*$ is the overlapping epitope recognized by 4E10 and Z13 and adopts an $\alpha$-helical conformation when bound to the 4E10 antibody (46). For 2F5 bnmAb, peptide mapping and phage display experiments have identified the sequence $^662\text{ELDKWAS}^668$ as its central epitope (192, 220). Crystal structures illustrated that binding of 2F5 to this epitope is mediated by interactions with residues at the base of the antibody’s unusually long CDR H3 loop (204, 208). I now present eight additional crystal structures in an effort to characterize the binding properties of bnmAb 2F5 in more detail. First, I show that the extended CDR H3 loop remains mobile, even after binding to peptides representing the gp41 epitope and that previously observed stable conformations of this loop are the result of crystal packing artefacts. I also identify the extended conformation reported for residues C-terminal to the core epitope as caused by interactions of the peptide with a neighbouring molecule in the crystal lattice. Further, my results present evidence for the existence of an anion binding site at the base of the CDR H3 loop. This finding could be of particular interest in light of the recently reported ability of bnmAb 2F5 to interact with membrane components and the potential role of the antibody’s long CDR H3 in this process (109, 233, 234). Another goal of the current work is to investigate potential atomic level interactions of 2F5 outside of its core MPER epitope, e.g. with the N-terminal fusion peptide (FP) segment of gp41. Although no direct interaction between the FP residues and the antibody fragment could be detected in co-crystals of 2F5 Fab’ and a construct comprising FP linked to the MPER, my data still suggest a role for this N-terminal gp41 sequence in stabilizing an $\alpha$-helical conformation for residues located C-terminal to the gp41 ELDKWA core epitope (Figure 3.1). I believe that these novel features of 2F5 binding are important for a better understanding of the mechanism of 2F5 neutralization, especially when considering potential vaccine candidates.

* Peptide residues are numbered according to their position in the full amino acid sequence of HXBc2 Env gp160.
Figure 3.1 Overall arrangement of the 2F5 Fab’ fragment bound to its extended gp41 peptide epitope as seen in the crystal structure of the Hyb3K peptide (see text for details), PDB ID: 3D0L. The Fab’ light chain (yellow) and heavy chain (green) are represented according to their secondary structure. The gp41 epitope residues \(^{661}\)LELDKWASLW\(^{670}\) (salmon color) are represented both as a stick model and based on their secondary structure. The yellow tetrapod identifies a potential phosphate binding site.
3.2 Materials and Methods

The Fab’ fragment of the 2F5 antibody was prepared by enzymatic digestion of the full IgG as previously published (35). Synthetic peptides ELLELDKWASLN, LLELDKWASLW and LELELDKWASLW[amide] mimicking the 2F5 linear epitope of gp41 were purchased from Sigma-Genosys, Oakville, ON. The extended peptide HybK3 (sequence GIGALFLGFLGAAGS-KK-Ahx-KNEQELLELDKWASLWN) and the scrHybK3 peptide (sequence GIGAFGLLGFLAAGSKK-Ahx-KNEQELLELDKWASLWN) were synthesized by Fmoc chemistry as previously described (167). Lyophilized peptides were dissolved in either 20 mM Tris, pH 8.0 or DMSO and subsequently mixed with 2F5 Fab’ to a final molar peptide:protein ratio of 10:1. Crystals were grown using the hanging drop vapour diffusion method with the protein concentration at 9 mg/ml. The starting volume of the drops was 2 μl, 1 μl of protein solution mixed with 1 μl of reservoir solution.

Crystals of the free 2F5 Fab’ fragment were grown with a reservoir solution containing 0.5 M Li₂SO₄ and 15 % PEG 8000 (35). Crystals of the 2F5 Fab’ fragment in complex with peptides ELLELDKWASLN, LLELDKWASLW and LELELDKWASLW[amide] were grown with a reservoir solution composed of 0.1 M sodium citrate, pH 5.6, 16 % 2-propanol and 16-20 % PEG 4000. A reservoir solution of 0.1M sodium citrate, pH 5.6, 1.4-1.8 M ammonium sulfate was used to grow the crystals of the Fab’ fragment in complex with the HybK3 construct and the ELLELDKWASLWN peptide. When transferring crystals between the 0.1 M sodium citrate, pH 5.6, 1.4-1.8 M ammonium sulfate solution and the 0.1 M sodium citrate, pH 5.6, 16 % 2-propanol, 16-20 % PEG 4000 solution, a period of 18-36 hours was allowed for equilibration prior to collecting X-ray diffraction data. All crystallization experiments were carried out at room temperature.

Crystals of the 2F5 Fab’ in complex with the scrHybK3 construct were obtained above a reservoir solution containing 0.1 M sodium acetate, pH 5.6, 16% 2-propanol and 16% PEG 4000. This is in contrast to the 2F5 Fab’ – HybK3 complex crystals, which were obtained from a reservoir solution containing 0.1 M sodium acetate, pH 5.6, 1.6 M ammonium sulphate. The 2F5 Fab’ – scrHybK3 complex did not form crystals under these conditions. To verify the effect of the solution’s polarity on the conformation of the peptide epitope, 2F5 Fab’ – scrHybK3 crystals obtained in the low polarity 2-propanol, PEG 4000 solution were soaked for 36 hours in a high
polarity solution containing 0.1 M sodium acetate, pH 5.6, 1.6 M ammonium sulphate. All crystallization experiments were carried out at room temperature.

0.05-1 % Tween-20 was added to the 2F5 Fab’ solution (20 mM Tris/HCl, pH 8.0) to grow crystals of the Fab’-peptide complexes in crystal form II. The presence of the detergent induced crystal form II in crystals grown in both low and high ionic strength solutions.

Before flash-freezing in boiling nitrogen, crystals were soaked for a few seconds in reservoir solution supplemented with 15-20% glycerol. X-ray diffraction data were collected either on a Rigaku RUH3R rotating copper anode generator at 50 kV and 100 mA with Osmic multi-layer optics using an R-Axis IV++ detector or on a Rigaku FR-C copper anode generator run at 50 kV and 60 mA with Xenocs multi-layer optics and a Mar345 image plate detector or on a Rigaku MicroMax-007 HF at 40 kV and 30 mA with Osmic multi-layer optics and a Mar345 image plate detector. The 2F5 Fab’ – scrHybK3 crystals were thin needles, brittle in nature, difficult to handle and to reproduce, as well as very short-lived in a strong X-ray beam. For this reason, X-ray diffraction data were collected on a home source, a Rigaku MicroMax-007 HF at 40 kV and 30 mA with Osmic multi-layer optics and a Mar345 image plate detector.

Raw diffraction data were reduced using the programs HKL2000, SCALEPACK and DENZO (206), or XDS (136). Phases for all 2F5 Fab’ structures were calculated by molecular replacement using the programs AMoRe (193) or CNS (34) and PDB ID: 2F5B as the initial search model. Constant and variable regions were treated as independent units. Experimental electron density maps were calculated with CNS (34) and interpreted using the program Coot (81). The resulting model was refined with CNS (34). Solvent molecules were picked with the CNS algorithm using the following input parameters: 1) a peak over 3 $\sigma$ in the difference electron density map; and 2) a 2.0-3.2 Å distance between the water molecule and oxygen or nitrogen neighboring atoms. Statistical data for data collection and model refinement are given in Table 1. Figures were prepared using the program Pymol (70).
Table 3.1 Statistics for data collection and model refinement. Values in parentheses represent those in the highest resolution bin.

<table>
<thead>
<tr>
<th>Crystal data</th>
<th>2F5 F_{ab}^+</th>
<th>2F5 F_{ab}^-</th>
<th>2F5 F_{ab}^- \text{ELLEIKWASLWN}</th>
<th>2F5 F_{ab}^- \text{ELLEIKWASLWN}</th>
<th>2F5 F_{ab}^- \text{LLELDKWASLWN} \text{[amide]}</th>
<th>2F5 F_{ab}^- \text{LLELDKWASLWN}</th>
<th>2F5 F_{ab}^- \text{ELLEIKWASLWN}</th>
<th>2F5 F_{ab}^- \text{ELLEIKWASLWN}</th>
<th>2F5 F_{ab}^- \text{LLELDKWASLWN} \text{[amide]}</th>
<th>2F5 F_{ab}^- \text{LLELDKWASLWN}</th>
<th>2F5 F_{ab}^- \text{ELLEIKWASLWN}</th>
<th>2F5 F_{ab}^- \text{ELLEIKWASLWN}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crystal growth</td>
<td>0.5 M lithium sulfate, 15% PEG 8000</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
<td>16% PEG 4000, 16% \text{2-propanol}</td>
</tr>
<tr>
<td>PDB ID</td>
<td>2PR4</td>
<td>2PSL</td>
<td>2D0V</td>
<td>2PSM</td>
<td>2DR0</td>
<td>3DL</td>
<td>2P2</td>
<td>2P2</td>
<td>2P2</td>
<td>2P2</td>
<td>2P2</td>
<td>2P2</td>
</tr>
<tr>
<td>Space group</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
<td>P2_1 2_1 2_1</td>
</tr>
<tr>
<td>Unit cell axes a, b, c [Å]</td>
<td>53.6, 76.4, 94.7</td>
<td>58.0, 64.7, 177.7</td>
<td>58.2, 64.6, 177.7</td>
<td>58.5, 64.3, 178.2</td>
<td>63.8, 76.6, 93.9</td>
<td>63.8, 76.6, 93.9</td>
<td>63.9, 76.5, 95.9</td>
<td>63.9, 76.5, 95.9</td>
<td>63.3, 76.2, 94.0</td>
<td>63.3, 76.5, 94.0</td>
<td>63.3, 76.5, 94.0</td>
<td>63.3, 76.5, 94.0</td>
</tr>
<tr>
<td>Resolution range [Å]</td>
<td>(15.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
<td>(20.0–2.05)</td>
</tr>
<tr>
<td>Completeness [%]</td>
<td>92.9 (93)</td>
<td>98.9 (93)</td>
<td>99.9 (99)</td>
<td>99.9 (98)</td>
<td>99.9 (99)</td>
<td>99.9 (99)</td>
<td>99.9 (99)</td>
<td>99.9 (99)</td>
<td>99.9 (99)</td>
<td>99.9 (99)</td>
<td>99.9 (99)</td>
<td>99.9 (99)</td>
</tr>
<tr>
<td>Measured reflections</td>
<td>89,456</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
<td>74,598</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>26,045</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
<td>29,049</td>
</tr>
<tr>
<td>R_{	ext{int}} [%]</td>
<td>7.4 (31.3)</td>
<td>5.4 (31.4)</td>
<td>5.2 (30.0)</td>
<td>7.8 (33.3)</td>
<td>12.0 (31.7)</td>
<td>26.2 (31.4)</td>
<td>5.8 (30.2)</td>
<td>7.8 (33.3)</td>
<td>12.0 (31.7)</td>
<td>26.2 (31.4)</td>
<td>5.8 (30.2)</td>
<td>7.8 (33.3)</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution [Å]</td>
<td>20.0–2.05</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
<td>40.0–2.44</td>
</tr>
<tr>
<td>Number of protein atoms/solvent molecules</td>
<td>3271/266</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
<td>3410/267</td>
</tr>
<tr>
<td>RMSD bond length [Å]</td>
<td>0.009/1.4</td>
<td>0.007/1.5</td>
<td>0.008/1.8</td>
<td>0.013/2.0</td>
<td>0.008/1.8</td>
<td>0.013/2.0</td>
<td>0.008/1.8</td>
<td>0.013/2.0</td>
<td>0.008/1.8</td>
<td>0.013/2.0</td>
<td>0.008/1.8</td>
<td>0.013/2.0</td>
</tr>
<tr>
<td>Number of reflections used for R_{	ext{work}}/R_{	ext{free}}</td>
<td>25,445/1281</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
<td>25,822/1227</td>
</tr>
<tr>
<td>Ramachandran plot: preferred region/outliers (%)</td>
<td>92.7/1.9</td>
<td>92.7/1.6</td>
<td>94.7/0.9</td>
<td>92.4/2.1</td>
<td>93.5/1.9</td>
<td>84.7/4.9</td>
<td>94.0/0.7</td>
<td>94.9/0.7</td>
<td>94.0/0.7</td>
<td>94.9/0.7</td>
<td>94.0/0.7</td>
<td>94.9/0.7</td>
</tr>
<tr>
<td>R_{	ext{syn}}/R_{	ext{free}} [%]</td>
<td>24.0/27.0</td>
<td>19.7/23.4</td>
<td>21.9/23.3</td>
<td>24.2/26.7</td>
<td>22.3/23.8</td>
<td>22.2/26.3</td>
<td>22.2/26.3</td>
<td>22.2/26.3</td>
<td>22.2/26.3</td>
<td>22.2/26.3</td>
<td>22.2/26.3</td>
<td>22.2/26.3</td>
</tr>
</tbody>
</table>
| Average overall B-value [Å]
| 21.6 | 42.9 | 37.3 | 46.6 | 27.4 | 36.4 | 20.1 | 26.5 | 20.3 | 24.5 | 20.3 | 24.5 |
Real-space correlation analysis was performed by first calculating a simulated annealing omit map with the residues of interest omitted from the calculation, then calculating an F_o-F_c electron density map and finally by looking at the real-space correlation between the map and the model using the “Density fit” analysis function in Coot (81). Superposition of all epitope peptides to generate Figure 3.5 was performed using the SSM superpose function in Coot (81) based on the Cα atoms of the whole Fab’ – epitope peptide model. RMS deviations were calculated based on PDB ID: 3D0L using all assignable atoms.

3.3 Results

3.3.1 Structure of the ligand-free 2F5 Fab’

Crystals of the ligand-free 2F5 Fab’ belong to space group P2_12_1 with unit cell axes a = 63.6 Å, b = 76.4 Å, c = 94.7 Å; I will refer to this crystal form as form I. The antigen binding site of 2F5 Fab’ is in an environment where it is surrounded by bulk solvent and free of influences from crystal lattice neighbours. Lack of any interpretable electron density that would correspond to residues Pro98 through Pro100 of the heavy chain, i.e. the extended CDR H3 loop, is indicative of the intrinsically high degree of mobility displayed by the unusually long CDR H3 of bmAb 2F5.

3.3.2 Crystals of 2F5 Fab’-gp41 peptide complexes at low and high ionic strength

To test the impact of the ionic strength of the crystallization solutions on gp41 epitope conformations, I aimed to grow crystals of various 2F5 Fab’ – gp41 epitope peptide complexes in two solutions buffered at pH 5.6, one containing 16% 2-propanol and 16% PEG 4000 and the other containing 1.6 M ammonium sulfate as precipitants. In cases where no crystals grew in one solution, soaking experiments were performed, changing from one ionic strength solution to the other; the crystals proved surprisingly resilient and survived the large change in ionic strength quite well.
3.3.3 The extended CDR H3 loop of 2F5 in the presence of peptide antigen

Previously described crystals of 2F5 Fab’ in complex with various gp41 epitope peptides (PDB IDs: 2F5B, 2P8L, 1TJG, 1TJH, 1TJI) also adopted space group P2₁2₁2₁ but with unit cell axes: a~58.0 Å, b~65.0 Å, c~175.5 Å; I will refer to them as form II crystals (204, 208). Despite their larger unit cell, the distances between the epitope binding site and the CDR H3 loop of the 2F5 Fab’ on one side and neighbouring molecules in the crystal lattice on the other are much shorter than in form I. This led me to investigate whether the stable conformation of the CDR H3 loop observed in form II could be due to crystal packing interactions and not to the formation of the peptide complex.

To assure that as many parameters as possible were kept constant when comparing the two crystal forms, crystals of the 2F5 Fab’ in complex with the ⁶⁶⁹ELLELDKWASLWN⁷⁷¹ peptide epitope were grown in crystal forms I (PDB ID: 2P8M) as well as II (PDB ID: 2P8L). In both structures, the ELDKWA core epitope residues form a slightly distorted type 1 β-turn, as previously described (204, 208). In crystal form I, no electron density is observed for residues Leu100A to Arg100H of the CDR H3 (Figure 3.2A). In form II, on the other hand, I find a well-defined conformation of the extended CDR H3 loop because it is stabilized by interactions with neighbouring molecules (Figure 3.2B); the real-space correlation values calculated for these residues fully support the interpretation that this region of the 2F5 antibody is disordered/mobile (Figure 3.2C). In form II crystals, crystal packing artefacts near the antigen binding-site freeze out the movements of CDR H3. In contrast, the unrestricted interactions between 2F5 and its gp41 epitope observed in the form I crystals will be more akin to those occurring in a physiological environment.
Figure 3.2 Visualization of the effect of crystal packing on the conformation of the 2F5 Fab’ CDR H3 loop in two different crystal systems. The 2F5 Fab’ fragment is shown as a backbone trace colored corresponding to its B-factor values. The spectrum from blue to red and the thin to thick shape of the backbone trace represent lower to higher B-factor values. The 2F5 Fab’ in complex with peptide ELLELDKWASLWN was co-crystallized in both crystal forms I and II (PDB IDs: 2P8L and 2P8M). The gp41 epitope residues for which electron density is observed are shown as a salmon-colored stick model. Electron density is shown in blue as a 2Fo-Fc map with a 1.5 σ contour level. Water molecules were omitted for clarity. a. In crystal form I, the CDR H3 loop is free of restrictions. No interpretable electron density is observed for residues in the extended CDR H3 loop. b. In crystal form II, the CDR H3 loop interacts with two adjacent symmetry related molecules. These crystal contacts stabilize it in a specific conformation (as indicated by the relatively low B-factor). c. Real-space correlation graph of the 2F5 Fab’ CDR H3 (residues R95(H)-V102(H)). Correlation coefficients represent the fit between the model and the simulated annealing omit map, calculated omitting residues of the CDR H3. Crystals obtained in form I (unit cell dimensions a=64 Å, b=76 Å, c=94 Å) show a correlation of less than 1 for residues of the extended CDR H3 loop (T99(H)-I100F(H)) suggesting a high mobility for this region (red, yellow and light blue). Crystals obtained in form II (unit cell dimensions a=58 Å, b=64 Å, c=177 Å) show a positive correlation for all residues of the CDR H3 loop (dark blue). These data demonstrate how stability of the extended CDR H3 residues is impacted by different crystal packing forces, and how removing crystal packing artifacts results in a mobile CDR H3 extended loop.

3.3.4 Residues located at the C-terminus of the ELDKWA epitope core sequence

In both forms I and II of 2F5 Fab’ in complex with the peptide epitope ELLELDKWASLWN (PDB IDs: 2P8M and 2P8L, respectively), I did not find any
interpretable electron density for residues Ser668 to Asn671 of the gp41 peptide epitope. For the complex in crystal form II at least, this was in contrast to a previous report, in which well defined electron density was observed for all residues of the epitope up to Trp670 (PDB ID: 1TJH) (204). The only difference between the two studies is the amidation state of the C-terminus of the peptide used for complex formation. Capping of the peptide with an amide group at its C-terminus is a common strategy to remove the terminal charge to better mimic the biologically relevant neutral peptide bond. To probe the effect of C-terminal amidation, I performed a structural analysis of 2F5 Fab’ in complex with the linear peptide $^{661}\text{LELDKWASLW}^{670} \text{[amide]}$ (PDB ID: 2P8P), essentially a repetition of PDB ID: 1THJ and included for reasons of checking reproducibility, then compared it to that of the non-amidated form $^{660}\text{LELDKWASLW}^{670}$ (PDB ID: 3D0V); crystals of both complexes belonged to crystal form II.

In essence, my analysis of the 2F5 Fab’- $^{661}\text{LELDKWASLW}^{670} \text{[amide]}$ complex (PDB ID: 2P8P) fully reproduced the structure of PDB ID: 1TJH. In both cases, there is clear electron density for residues located C-terminal to the gp41 ELDKWA core (Figure 3.3A). Upon closer analysis of this region, however, one notices that the stable conformation of this part of the amidated gp41 epitope peptide is caused by the C-terminal amido group extending the peptide to just the right length so it can form a key hydrogen bond with the side chain of Asp72 of the heavy chain of the neighbouring Fab’ molecule. In addition, Trp670 now fits nicely into a groove on the surface of the same Fab’ molecule where it interacts with the side chain of Gln77 via its indole nitrogen. These contacts lock in the conformation of the C-terminal residues of the epitope peptide. Due to the strong interactions between epitope peptide and neighbouring molecules in the crystal lattice, however, we have to interpret this extended conformation of the peptide’s C-terminal residues as a crystallization artefact.

The crystal structure of 2F5 Fab’ with the non-amidated peptide $^{660}\text{LELDKWASLW}^{670}$ (PDB ID: 3D0V) reveals a flexible conformation for residues following the ELDKWA core; no interpretable electron density is observed for amino acids Ser668 to Trp670 of the peptide epitope (Figure 3.3B); again, the real-space correlation values present the same picture (Figure 3.3E). Although there are no detectable interactions between epitope peptide and crystal neighbours, Arg100H, located at the base of the elongated CDR H3 loop, is now in contact with the side chain of Asp72 of the heavy chain of the neighbouring Fab’ molecule. However, if the same complex is crystallized in form I (PDB ID: 2P8M) neither residues Ser668 to Trp670 nor
Arg100H of the CDR H3 loop are represented by interpretable electron density (Figures 3.3C and 3.3E). Again, form II crystals display their strong tendency to engage the epitope/paratope environment in crystal-crystal contacts, making it very difficult to obtain a valid and clear understanding of its appearance in a biologically relevant environment. For this reason, all subsequent structures were determined from crystals belonging to crystal form I.
No interpretable electron density for C-terminal peptide residues

No interpretable electron density for CDR H3 extended loop, including Arg100H

Sulfate ion

No interpretable electron density for CDR H3 extended loop
Figure 3.3 The 2F5 Fab’ paratope in complex with different gp41 epitope peptides. The gp41 epitope residues for which electron density is observed are shown as a salmon-coloured stick model. The 2F5 Fab’ is rendered as a cartoon model according to its secondary structure, with the heavy chain in green and the light chain in yellow. Surrounding symmetry related molecules are shown in purple. Electron density is shown in blue as a 2F_{o}-F_{c} map with a 1.5 σ contour level. Water molecules were omitted for clarity. a. Crystal structure of 2F5 Fab’ with peptide 661LELDKWASLW^{670} [amide] in crystal form II (PDB ID: 2P8P). Hydrogen bonds are observed between the peptide’s amidated C-terminus and Asp72 of the heavy chain of a symmetry related 2F5 Fab’ molecule as well as between the indole nitrogen of Trp670 of the peptide epitope and Gln77 of the heavy chain of the same symmetry related Fab’ molecule. These interactions stabilize an artefactual conformation of the C-terminus of the peptide epitope. b. Crystal structure of 2F5 Fab’ with peptide epitope 660LLELDKWASLW^{670} in crystal form II (PDB ID: 3D0V). Removing the amide group from the peptide’s C-terminus renders this part of the peptide mobile, as demonstrated by the lack of electron density for residues Leu669 and Trp670. Arg100H at the base of the long CDR H3 loop forms hydrogen bonds with Asp72 of the heavy chain of a symmetry related Fab’ molecule. Packing artefacts in crystal form II interfere with the proper binding of the C-terminal residues of the peptide epitope. c. Crystal structure of 2F5 Fab’ co-crystallized with epitope peptide 659ELLELDKWASLWN^{671} in crystal form I (PDB ID: 2P8M). No electron density is observed for residues Leu669, Trp670 and Asn671, suggesting that this area of the peptide epitope is flexible. Also, no electron density is observed for the 2F5 extended CDR H3 loop, including Arg100H. d. Crystal structure of 2F5 Fab’ co-crystallized with peptide Hyb3K, 514GIGALFLGFLGAAGS^{528}KK-Ahx-
655KNEQELLELDKWASLWN^{671} in crystal form I (PDB ID: 3D0L). Electron density for residues Leu669 and Trp670 at the peptide’s C-terminus is clearly visible. An intra-chain hydrogen bond between Trp666 (i) and Leu669 (i+4) shows the start of an α-helix for residues at the C-terminus of the ELDKWA core. A high electron density peak observed at the base of the CDR H3 loop is interpreted as a sulfate ion. Again, no electron density is observed for

(e) Real-space correlation of 2F5 Fab’ epitope/paratope residues and sulfate binding site

![Graph showing correlation of 2F5 Fab’ epitope/paratope residues and sulfate binding site]
residues of the 2F5 extended CDR H3 loop. e. Real-space correlation graph of the 2F5 epitope/paratope (gp41 residues $^{659}$ELLELDKWASLW$^{670}$, 2F5 CDR H3 residue Arg100H and sulfate ion at the base of the 2F5 CDR H3) derived from structures obtained in crystal form I. Correlation coefficients represent the fit between the model and the simulated annealing omit map calculated omitting residues of the 2F5 epitope/paratope (given above). Residues near the ELDKWA core show the highest correlation coefficients, while values for residues at its C-terminus fall sharply below 1 after residue Ser668, except in the case of the Hyb3K complex grown in ammonium sulfate where there is clear electron density observed for all residues of the epitope/paratope displayed in light red (PDB: ID 3D0L). Electron density for residue Arg100H at the base of the CDR H3 is only observed when the complete epitope ($\alpha$-helical turn for C-terminus residues) is formed. Also, the correlation coefficients seem to indicate that it is possible to titrate the sulfate ion in/out of the interaction site at the base of the 2F5 CDR H3 by growing/soaking the crystals in solutions of higher/lower ionic strength.

3.3.5 gp41 FP influences the conformation of an epitope peptide bound to the 2F5 paratope

Previous reports described improved binding of 2F5 to its epitope when the latter was presented linked to the FP section of gp41 (167, 168). To probe for these long-range influences, I co-crystallized the 2F5 Fab’ with the HybK3 peptide construct, $^{514}$GIGALFLGFLGAAGS$^{528}$KK-Ahx-K$^{656}$NEQELLELDKWASLWN$^{671}$, which connects two gp41 peptides, the FP and MPER regions, with a Lys-Lys-6-amino-hexanoic acid linker. Co-crystals, grown in form I (PDB ID: 3D0L), result in electron density maps that do not reveal the positions of any of the residues of the FP portion; only residues Leu661 to Trp670 show interpretable electron density. Interestingly, however, amino acids immediately following the type-1 $\beta$-turn, namely residues Trp666 to Leu669, assume dihedral angles defining the start of an $\alpha$-helix (Figure 3.3D). Residues Leu669 and Trp670 of the epitope now also fit nicely into a groove of the 2F5 paratope, making hydrophobic interactions with Pro98, Val100K and the side chain of Arg100H of the antibody heavy chain. When 2F5 Fab’ was co-crystallized with the peptide $^{659}$ELLELDKWASLWN$^{671}$, i.e. HybK3 without the linker and the FP region, interpretable electron density did not extend beyond Ser668 (PDB ID: 3DRO). The full interactions between the 2F5 paratope and this elongated epitope are represented in Table 3.2.
Table 3.2 Contact Surface Areas by Residue on Peptide Hyb3K and 2F5 Fab’.

<table>
<thead>
<tr>
<th>peptide residue</th>
<th>contact area (Å²)</th>
<th>Fab’ residue, light chain</th>
<th>contact area (Å²)</th>
<th>Fab’ residue, heavy chain</th>
<th>contact area (Å²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glu662</td>
<td>57</td>
<td>Tyr94</td>
<td>74.9</td>
<td>Arg100H</td>
<td>70.1</td>
</tr>
<tr>
<td>Leu663</td>
<td>41.2</td>
<td>Phe93</td>
<td>53.2</td>
<td>Pro98</td>
<td>34.3</td>
</tr>
<tr>
<td>Asp664</td>
<td>72.9</td>
<td>His92</td>
<td>45.4</td>
<td>Val100K</td>
<td>28</td>
</tr>
<tr>
<td>Lys665</td>
<td>75</td>
<td>Leu2</td>
<td>13.8</td>
<td>Arg95</td>
<td>25.7</td>
</tr>
<tr>
<td>Trp666</td>
<td>93.8</td>
<td>Glu27</td>
<td>13.7</td>
<td>Tyr52</td>
<td>25.4</td>
</tr>
<tr>
<td>Ala667</td>
<td>15.1</td>
<td>Leu81</td>
<td>11.3</td>
<td>Arg58</td>
<td>17.5</td>
</tr>
<tr>
<td>Ser668</td>
<td>0</td>
<td>Ala1</td>
<td>7.5</td>
<td>Asp56</td>
<td>15.2</td>
</tr>
<tr>
<td>Leu669</td>
<td>42</td>
<td>Pro95</td>
<td>5.7</td>
<td>Asp54</td>
<td>13.8</td>
</tr>
<tr>
<td>Trp670</td>
<td>81.2</td>
<td>Ser31</td>
<td>4.2</td>
<td>Asn100L</td>
<td>12.6</td>
</tr>
<tr>
<td>total:</td>
<td>554.1</td>
<td>Asp50</td>
<td>2.2</td>
<td>Phe32</td>
<td>9.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>total:</td>
<td>240.6</td>
<td>Gly100L</td>
<td>7.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Gly97</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>total:</td>
<td></td>
<td></td>
<td>277</td>
</tr>
</tbody>
</table>

These results suggest a role for the FP sequence of gp41 in inducing the α-helical turn observed in the HybK3 complex (Figure 3.3E). Real-space correlation analysis also shows that residues located C-terminal of the ELDKWA core tend to adopt a more stable conformation in a high ionic strength solution as opposed to one of lower ionic strength (Figure 3.3E).

3.3.6 The effect of the gp41 FP segment is specific

To test the specificity of the gp41 FP segment on the observed structure, unspecific hydrophobic interactions were mimicked within scrHybK3, a peptide with identical composition but the conserved FP heptad scrambled. In order to keep as many experimental parameters constant as possible, some of the 2F5 Fab’-scrHybK3 co-crystals grown in propanol and polyethylene glycol (PDB ID: 3DRT) were soaked in 1.6 M ammonium sulfate solution (PDB ID: 3EGS) to reproduce the high ionic environment in which co-crystals of 2F5 Fab’-HybK3 were obtained. However, the 2F5 Fab’-scrHybK3 co-crystals were of much lower quality i.e. maximum resolutions of 3.3 Å and 3.6 Å. Nonetheless, standard protocols produced models of the complexes that displayed reasonable statistics ($R_{crys} = 0.203$, $R_{free} = 0.229$ and $R_{crys} = 0.235$, $R_{free} = 0.249$, respectively, combined with good overall geometry and low B-factors of 19.0 and 23.5 Å²) (refer to Table 3.3 for complete crystallographic statistics). Finally, comparison of the two 2F5 Fab’ – scrHybK3 models with previously published 2F5 Fab’ crystal complexes showed r.m.s.d. values for all Cα atoms smaller than 0.2 Å, and r.m.s.d. values ranging from 0.2 to 0.6 Å for all atoms, confirming the close structural similarity between starting model and final model, i.e. the good quality of the phases underlying the electron density calculations.
Similarly to the co-crystal structures of 2F5 Fab’-HybK3, electron density maps did not reveal the positions of any of the hybrid Gly514-Ser528 FP residues in the 2F5 Fab’-scrHybK3 co-crystal structures (Figure 3.4). However, the interpretable electron density spanning the Leu661-Trp670 stretch confirmed the existence of striking differences between the HybK3 and scrHybK3 structures. In the HybK3 structure, amino acids Trp666-Ala667-Ser668-Leu669-Trp670, following the type-1 \( \beta \)-turn Asp664-Lys665-Trp666, assume dihedral angles compatible with a canonical \( \alpha \)-helix (Fig 3.4A). In contrast, 2F5 Fab’-scrHybK3 co-crystal diffraction data did not produce interpretable electron density for residues Ser668-Leu669-Trp670-Asn671 (low polarity solution, Figure 3.4B) and for residues Leu669-Trp670-Asn671 (high polarity solution, Figure 3.4C). Although a sulfate-containing solution seems to slightly increase the interpretable electron density located C-terminal to the type-1 \( \beta \)-turn, it is not sufficient to elicit a full canonical \( \alpha \)-helix in residues Trp666-Ala667-Ser668-Leu669-Trp670. From these results, we can conclude that the formation of the helical turn at the C-terminus of the DKW core residues seems to depend on the presence of both the correct FP sequence as well as sulfate ions.
Figure 3.4 Stereoimages of the crystal structures of 2F5 Fab’ in complex with different constructs. (A) Crystal structure of the 2F5 Fab’ cocrystallized with peptide Hyb3K, grown in a solution containing 1.6 M ammonium sulfate (PDB ID: 3D0L). (B) Crystal structure of 2F5 Fab’ cocrystallized with peptide scrHyb3K, grown without sulfate (16% PEG 4000 and 16% 2-propanol) (PDB ID: 3DRT). (C) Crystal structure of 2F5 Fab’
cocrystallized with peptide scrHyb3K, and subsequently transferred into a solution containing 1.6 M ammonium sulfate (PDB ID:3EGS). This procedure generated an environment comparable to the one for 2F5 Fab'-Hyb3K crystals, especially providing sulfate ions. The gp41 epitope residues for which electron density is observed are shown as a salmon-colored stick model. The 2F5 Fab' is rendered as a cartoon model according to its secondary structure, with the heavy chain in green and the light chain in yellow. Electron density is shown in blue as an Fo-Fc simulated annealing omit map with a 2.0σ contour level. To allow for easier comparison, all electron density maps were calculated at a resolution of 3.6 Å.

3.3.7 An anion-binding site at the base of CDR H3

Another hitherto not observed feature identified in the electron density map of the HybK3 complex is a high peak separate from the protein density and located at the base of the CDR H3 loop. It is surrounded by the peptide bond nitrogens of Ala100G and Gly100I as well as the positively charged guanidinium group of Arg100H at distances of 2.5 Å, 2.9 Å and 2.4 Å, respectively. Based on its intensity, its shape, its environment and the high concentration of sulfate in the mother liquor we interpret this peak as a sulfate ion (Figure 3.3D).

Our crystallographic data evaluated by real-space correlation analysis suggest that the anion binding site at the base of the 2F5 CDR H3 is only populated when the crystal is in a high ionic strength solution (Figure 3.3E). We also observed that it was possible to titrate the low-affinity sulfate ion out of the anion-binding site by transferring crystals from a high ionic strength solution to a low ionic strength solution (PDB ID: 3DRQ). Altogether, these data are a strong indication of the presence of a low-affinity anion binding site at the base of the 2F5 CDR H3 loop.

3.4 Discussion

Molecular interactions between the 2F5 paratope and its gp41 linear epitope have been described in detail (204, 208). This paper, presenting eight additional crystal structures, enlarges the area of antibody-epitope peptide interaction, identifies new structural features and also aims to illuminate the importance of taking into account the presence or absence of crystal contacts when interpreting electron density maps.
Figure 3.5 Superposition of the gp41 peptide epitope conformations found in eight different crystal structures. Superposition was performed using $C_\alpha$ atoms of the entire Fab’ fragment in addition to the epitope peptide. gp41 peptide residues are shown as stick models: PDB ID: 2P8L (green), PDB ID: 3D0V (cyan), PDB ID: 2P8P (magenta), PDB ID: 1TJH (yellow), PDB ID: 2P8M (salmon), PDB ID: 3DRO (light gray), PDB ID: 3D0L (blue), and PDB ID: 3DRQ (orange). The conformation of residues $651\text{LELDKWAS}^{668}$ when bound to 2F5 Fab’ is conserved in all structures, whereas the conformation of residues $669\text{LELDKWAS}^{670}$ varies significantly in structures 2P8P, 1TJH and 3D0L. RMS deviations for all atoms of these peptides are given in Table 3.4.

Obtaining both epitope complexed and uncomplexed 2F5 Fab’ crystals in the same crystal form, one in which the CDR H3 extended loop does not interact with neighbouring molecules in the crystal lattice, allowed us to demonstrate convincingly that in an aqueous environment this loop is mobile, both in its free form and when bound to its gp41 peptide epitope. The binding of epitope peptides to 2F5 alone will not cause the ordering of the loop. Possible explanations for this mobility include a requirement to access a recessed epitope as was found for gp120-specific bnMAbs (154), and/or a need to interact with a dynamic viral membrane (109), and/or to disrupt the structure of proteins coming into close proximity of gp41 in the process of membrane fusion.
Zwick et al. showed that non-conservative substitutions at certain positions of the CDR H3 loop led to a decrease in both binding affinity and 2F5 neutralization, suggesting a significant role for this region during these processes (305). Further studies, probably on a larger 2F5 epitope, are required to shed more light onto the exact role of the CDR H3 and the mechanisms underlying its actions.

Table 3.4 RMS deviations of gp41 epitope peptide residues for all structures presented in this work as well as for structure PDB ID 1TJH (Ofek et al. (2004)). The reference structure for RMS deviation comparison is PDB ID 3D0L. All structures adopt a very similar conformation for residues 661-LELDKWAS668. There is a significant difference in conformation for residues 669-LW670 in PDB ID 3D0L (α-helical turn) and PDB IDs 2P8P and 1TJH (extended conformation caused by crystal packing interactions).

<table>
<thead>
<tr>
<th>Structure PDB ID</th>
<th>RMSDs for residues 661-LELDKWAS668 (Å)</th>
<th>RMSDs for residues 669-LW670 (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D0L</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2P8P</td>
<td>1.36</td>
<td>5.64</td>
</tr>
<tr>
<td>3D0V</td>
<td>1.24</td>
<td>5.48</td>
</tr>
<tr>
<td>1TJH</td>
<td>5.19</td>
<td>5.48</td>
</tr>
<tr>
<td>2P8M</td>
<td>1.35</td>
<td>—</td>
</tr>
<tr>
<td>3DRO</td>
<td>0.48</td>
<td>—</td>
</tr>
<tr>
<td>3DRQ</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Another intriguing result is the conformation adopted by residues located at the C-terminus of the gp41 ELDKWA core. The electron density maps of 2F5 Fab’ in complex with the gp41 peptides 660-LELDKWASLW670 or 659-ELLELDKWASLWN671 show no interpretable density for residues Ser668 to Asn671, demonstrating the high mobility of the peptide antigen in this region. Ofek et al. reported in a structural study of 2F5 Fab in complex with the peptide 659-ELLELDKWASLW670[amide] a stable conformation for all residues located at the C-terminus of the ELDKWA core, a result that I was able to reproduce. In the complex with the amidated peptide, the terminal amide group anchors the peptide to the next molecule in the crystal lattice, stretching out the amino acid chain and placing Trp670 into a pocket of the neighbouring Fab’ molecule (204). Removal of the amide group reveals a much more flexible conformation for residues located at the C-terminus of the epitope core. In fact, the entire region consisting of the epitope peptide C-terminus and the base of the elongated CDR H3 loop possesses a high degree of flexibility. Real-space correlation analysis of the 2F5 paratope / epitope interaction in various
crystal complexes, which are free of lattice restrictions near the epitope binding site, confirms this statement.

Measurements of the relative binding strength of extended gp41 peptides to 2F5 as well as a subsequent alanine scan study have both demonstrated that residues located N-terminally of the 2F5 core epitope enhance affinity due to specific binding interactions (involving especially Glu659 and Leu661) whereas the nature of residues located C-terminally of the gp41 ELDKWA core has much less effect on the binding energy (179, 182). The results of our structural observations are in agreement with these biochemical findings. While residues located N-terminally of the 2F5 core epitope adopt a stable conformation and make specific contacts with the antibody, residues following Ala667 seem to be more mobile. Our results also agree with a recent model proposed by Sanchez-Martinez et al. (234), in which the authors suggest that the binding of 2F5 to the extended amino-terminal stretch of its gp41 epitope would contribute enough free energy to compete efficiently with the tendency of residues located at the C-terminus of the gp41 epitope to partition to the membrane. The correlation between antibody binding affinity and neutralization capabilities, however, still remains to be fully characterized. In alanine scanning studies, very few gp41 residues were essential for neutralization of virus by 2F5 (304). However, it was also seen that mutating 2F5 residues outside of the part of the paratope that interacted with the $^{664}$DKW$^{666}$ sequence led to both a decrease in binding affinity and in neutralizing capability, suggesting a link between these two processes (305). It also emphasizes the importance of considering residues of the gp41 epitope that lie outside of the ELDKWA core when trying to fully understand the mechanism of 2F5 neutralization (305).

One such longer-range interaction could be contributed by the N-terminal FP, which is suspected to come into close contact with the MPER (23, 167, 168). When analyzing the co-crystals of 2F5 Fab’ with the HybK3 peptide, which consists of the gp41 N-terminal FP linked with residues of the MPER ($^{514}$GIGALFLGFLAGAGS$^{528}$KK-Ahx-$^{655}$KNEQELLELDKWASLWN$^{671}$), no electron density is observed for any residues of the FP, which suggests that the 2F5 antibody does not interact tightly with this portion of gp41. Nevertheless in the same map, residues Ser668, Leu669 and Trp670 show clear electron density and arrange into the first turn of an α-helix. Previous NMR studies of the gp41 MPER conformation as well as the crystallographic analyses of the gp41 core six-helix bundle and of the 4E10 Fab-epitope complex all suggest that the gp41 MPER is prone to adopt a stable α-helical conformation, especially in a membrane
environment (18, 46, 50, 264). The apparent increased affinity towards HybK3 demonstrated by Lorizate et al. (167) compared to the linear epitope peptide alone would even argue for an FP-specific component. Our results show that the residues following Ser668 can assume such a helical conformation and therefore are consistent with the idea that 2F5 recognizes a transient gp41 fusion-intermediate β-turn conformation emerging from or on its way to a more energetically favoured α-helical gp41 MPER conformation. Although I do not see any firm location for the FP in the HybK3 complex structures and therefore no direct molecular interaction can be discussed, it seems that the interaction with the FP stretch that promotes the C-terminal helix in the HybK3 construct is specific, as evidenced from the 2F5 Fab’– scrHybK3 co-crystal structures. Molecular crowding is a possible mechanism explaining this effect of the FP on residues located at the C-terminus of the DKW core epitope.

The low quality of the 2F5 Fab’– scrHybK3 co-crystals was obvious not only from the low maximum resolutions of 3.3 Å and 3.6 Å obtained for the native and the soaked crystal, respectively, but also from the low I/σ(I), which in turn resulted in poor statistics for the diffraction data. Although the Protein Data Bank contains more than a dozen of entries with similar statistics, we felt that further data analysis was needed to establish whether the results could be interpreted in a meaningful manner. First, I confirmed that the crystal was not twinned by probing the data set with the Padilla-Yeates algorithm (http://nihserver.mbi.ucla.edu/pystats/). Subsequently, I refined the unit cell dimensions in space group P1 to find that both crystal complexes had unit cell dimensions less than 2/100th from the respective values of those determined in previous 2F5 Fab’ crystal complexes; the unit cell angles all refined to values deviating less than three/hundredths from the 90° of the assigned space group P2₁2₁2₁. Moreover, when density averaging was applied in space group P1 it closely reproduced the density calculated in space group P2₁2₁2₁. These findings established that the 2F5 Fab’– scrHybK3 crystals obtained were isomorphous with most of the 2F5 Fab’ complex crystals determined previously from much better data (a~64 Å, b~77 Å, c~94 Å, α=90°, β=90°, γ=90°). The 2F5 Fab’– scrHybK3 crystals could therefore be phased directly from a previous well-refined, high-resolution model (PDB ID: 3D0L), providing excellent starting phases. As the majority of crystal structure information is carried in the phases of the diffracted rays the resulting electron density map allowed for the clear determination of the presence or absence of ordered secondary structure, despite the relatively poor quality of the intensity measurements.
The crystal structure of the 2F5 Fab’ - HybK3 complex revealed a second feature of interest, a high electron density peak near residue Arg100H at the base of the CDR H3. We interpret this peak as a sulfate ion picked up from the crystallization solution. The three residues at the base of the CDR H3 loop that interact with the sulfate are all located in close proximity to the indole ring of the epitope residue Trp670 (distances < 4.5 Å). As the 2F5 epitope is known to be in close proximity to the phospholipid bilayer, we hypothesize that this sulfate ion might be mimicking the negatively charged phosphate of a lipid head-group, representing a possible site of interaction between the 2F5 antibody and the phospholipid bilayer (Figure 3.6). Recombinant 2F5 Fab with an R96A mutation in the paratope bound with decreased affinity to membrane components (Alam et al., Keystone HIV Vaccine meeting 2008 abstract book and personal communication). This effect on binding, together with our structural results raise the question whether positively charged residues of the paratope (such as Arg96 and Arg100H of the heavy chain) are important factors in mediating interactions with lipid head-group phosphates during 2F5 epitope binding, possibly in a multi-step mechanism (7).
Figure 3.6 Model representation of the interaction between bnmAb 2F5 and HIV-1 gp41. This figure was generated by performing a positional overlap of the 2F5 Fab'/gp41 epitope peptide $^{514}$GIGALFLGFLGAA$^{528}$GSKK-Ahx-$^{655}$KNEQELLELDKWASLWN$^{671}$ crystal structure (PDB ID: 3D0L) with the gp41 MPER $^{662}$ELDKWASLWNWFNITNWLWYIK$^{683}$ structure in a lipid environment obtained by NMR/EPR/SPR (represented in yellow, PDB ID: 2PV6) presented by Sun et al. (2008) (264). The 2F5 epitope is represented in green and the 2F5 Fab' fragment is depicted as a vacuum electrostatic model with blue indicating positively charged and red negatively charged regions; white represents non-polar regions of the molecule. The orientation of the bnmAb 2F5 relative to the viral membrane is chosen based on assigning the position of the sulfate ion (in green circle) to overlap with the head groups of the viral membrane, the electrostatic charges on the surface of the Fab' and the overlap of the α-helical MPER structures of the two models. Then, the mobile CDR H3 extended loop points towards the membrane, where it is hypothesized to interact with components of the membrane bilayer or with other parts of gp41 residing in/near the membrane. As there is no information available about the exact conformation of HR1, HR2 and FP of gp41 when binding to bnmAb 2F5 these parts have not been included in the model. The inset box shows a magnification of the 2F5 Fab’ interaction with its gp41 epitope. It displays the key residues of the 2F5 paratope (mostly CDR residues) involved in mediating the interaction with its antigen.

The crystal structures of 2F5 Fab’ in complex with various gp41 peptide epitopes reveal several key features of this antibody-antigen interaction: 1) whenever free of contacts caused by crystal
artefacts, the extended CDR H3 loop is mobile; this is true for its ligand-free as well as its
epitope-bound forms; 2) the interaction between the antibody and the gp41 ELDKWA epitope
core is absolutely critical and there are also close and specific contacts to residues located N-
terminally to the core, both aspects described previously; 3) residues located at the C-terminus of
the gp41 ELDKWA core do not interact as tightly with the antibody. However, in the presence of
a larger peptide containing the gp41 FP segment, these residues adopt a conformation consistent
with the start of an $\alpha$-helix; 4) at high sulfate concentrations, the electron density maps of 2F5
Fab'-peptide complexes contain a peak that may mark a binding site of phosphate groups of
negatively charged lipid headgroups.

Many questions persist about the detailed processes underlying HIV-1 neutralization by bnmAb
2F5. Past efforts at creating small molecule immunogens mimicking the 2F5 epitope or at
presenting the constrained $\beta$-turn in scaffold proteins have failed to generate neutralizing
activity, despite the ease with which antibodies with high affinities for the 2F5 epitope were
obtained (54, 117, 133, 171, 292, 300). Many reasons have been evoked to explain this failure,
such as the possibility that the immunogens do not retain the proper conformation of a transient
gp41 epitope recognized by 2F5 or that the immunogens do not represent the entire 2F5 epitope
(54, 133, 171, 292, 300). We believe that both aspects could well play a role in the failure of
generating bnmAbs. An effective immunogen capable of eliciting 2F5-like neutralizing
antibodies will therefore need to have a stable conformation, not only for the backbone
replicating the constrained $^{659}\text{ELLELDKWAS}^{668}$ type 1 $\beta$-turn, but also with the side chains
fixed in the positions observed in the Fab’ complex structures. In addition, residues at the C-
terminus of the ELDKWA core should probably adopt an $\alpha$-helical conformation and other, so
far unidentified parts of the full epitope might also have to be included. Immunogens might even
have to be presented in a phospholipid environment to elicit a fully competent neutralizing
antibody response.
Chapter 4
Determining the role of the 2F5 extended CDR H3


(N.H. and R.M. carried out the pseudovirus neutralization and cell-cell inhibition assays reported in this chapter. Isothermal calorimetry experiments were performed with support from S.G.T. and N.H. The recombinant 2F5 Fab expression protocol was developed with significant input from A.C.)
Ablation of the CDR H3 apex of the anti-HIV-1 broadly neutralizing monoclonal antibody 2F5 abrogates neutralizing capacity without affecting core epitope binding.

4.1 Introduction

The interaction between 2F5 and its minimal linear epitope \((662\text{ELDKWAS}^{668})\) has been extensively characterized in numerous structural studies (36, 134, 204). Briefly, the primary gp41 epitope as recognized by 2F5 assumes a \(\beta\)-turn conformation with core residues \(664\text{DKW}^{666}\) in the centre of the turn. These amino acids are flanked by residues adopting an extended conformation at the N-terminus and a canonical \(\alpha\)-helical turn at the C-terminus, respectively (134). One interesting characteristic of bnmAb 2F5 is its remarkably long CDR H3 loop, which contains 22 amino acids. On average, human CDR H3 loops encompass between 10 to 14 residues, although even larger CDR H3 loops with up to 32 residues have been reported ((289) and D. Burton, personal communications). Although most residues at the base of the 2F5 CDR H3 are involved in direct recognition of the core \(662\text{ELDKWAS}^{668}\) epitope along with the other five CDRs (Figure 4.3A) (63, 134), the role of residues located at the apex of the CDR H3 loop is still a matter of debate. In a previous Ala-scanning mutagenesis study, Zwick et al. demonstrated the involvement of this loop in neutralization and suggested that residues located at its apex play a role in core epitope recognition. This interpretation was based on direct ELISA binding studies to both core peptide epitope and gp41 ectodomain (305). However, such an involvement in core epitope binding is not apparent from structural studies (134, 204). Other studies documenting some affinity of 2F5 for membrane components have suggested that this loop interacts directly with viral membranes (7, 109, 168, 176, 234).

To investigate the role of the 2F5 CDR H3 apex residues in more detail, we first report on the production of sufficiently large amounts of recombinant 2F5 Fab of high purity. Expressing Fab in a bacterial system allowed us to easily create and produce mutants with changes in the 2F5 CDR H3 apex sequence. In an attempt to characterize the functional role of this element, two mutants were created. In one, the CDR H3 sequence of \(100\text{TLFGVP}^{100F}\) was replaced by a single Ser-Gly dipeptide, bridging the \(\sim5\ \text{Å}\) distance between the \(C_\alpha\)-atoms of T99 and A100G, the last two residues represented by electron density in our structural analysis of the native 2F5 Fab’
In the other mutant, the apical Phe100B(H) of the CDR H3 was changed to an alanine residue. Isothermal titration calorimetry (ITC) measurements, as well as competitive ELISA binding assays confirmed that the 2F5 CDR H3 apex does not directly affect core epitope recognition. In contrast, the correct amino acid composition of this Fab element was required for potent neutralization of pseudovirus infection and inhibition of Env-induced cell-cell fusion. With the aim of unraveling the molecular interactions responsible for this loss of neutralization, we performed ELISA binding assays of the 2F5 CDR H3 apex mutants with various extensions of the gp41 epitope, as well as with membrane components. We observed a clear difference in binding affinity between native and mutant versions of 2F5 Fab when residues comprising the 4E10 epitope and beyond were added to the C-terminus of the core 2F5 epitope and even more so if this extended 2F5 epitope was inserted into a phospholipid bilayer environment. In contrast, when probed against membrane components alone, no such differentiation was detected. The importance of these findings is discussed in detail.

4.2 Materials and Methods

4.2.1 Materials

Peptides representing different fragments of the gp41 MPER, 2F5ep (\(^{656}\text{NEQELLELDKWASLWN}\text{671}\)), 2F5preTM, (\(^{656}\text{NEQELLELDKWASLWNWFNITNWLWYIK}\text{683}\)), and 2F5preTM(9,10)Ala, (\(^{656}\text{NEQELLELAAAWASLWNWFNITNWLWYIK}\text{683}\)) were synthesized as C-terminal carboxamides by solid-phase synthesis using Fmoc chemistry and purified by HPLC at the Proteomics Unit of the University Pompeu-Fabra (Barcelona, Spain). Stock solutions for these peptides were prepared in DMSO (spectroscopy grade). Peptide \(^{662}\text{ELDKWAS}\text{668}\) was purchased from Genscript at a >75% purity level (Piscataway, NJ) and its stock solutions were prepared in 20 mM Tris, pH 8.0. 1-Palmitoyl-2-oleoylphosphatidylcholine (POPC), cholesterol (Chol) and 1-palmitoyl-2-hydroxy-\(sn\)-glycero-3-phosphate (sodium salt) (PA) were purchased in powder form from Avanti Polar Lipids (Birmingham, AL) and stock solutions were prepared in chloroform. Recombinant gp41 HxB2 (aa 541-682) (glycosylated), produced in \textit{Pichia pastoris}, was
purchased from Viral Therapeutics, Inc. (Ithaca, NY). In this construct, the 2F5 core epitope sequence does not contain any glycosylation sites.

### 4.2.2 Expression and purification of 2F5 Fab

The genes encoding the 2F5 Fab (light chain residues Ala$^{1}$ to Cys$^{214}$ and heavy chain residues Ile$^{2}$ to Cys$^{216}$) were synthesized (BioBasic Inc., Toronto, Ontario) and subsequently cloned into the pET-Duet 1 vector (Novagen, Gibbstown, NJ). Because the initial cloning of these genes made use of the constant region of another Fab (8F9) (267), the following deviations from the 2F5 consensus sequence are present in the 2F5 Fab constant region of all constructs: light chain 103 RVDVR$^{107}$ to 103 KLEIK$^{107}$ and E193A, as well as heavy chain 132 AGGA$^{135}$ to 132 SGGT$^{135}$, T114A, T116F, T195I and R210K. The thrombin site-containing tag DKTHLVPGRGSSHHHHHHH was added at the C-terminus of the 2F5 heavy chain for subsequent nickel affinity purification. Site-directed mutagenesis was used to create 2F5 Fab mutants T15(H)A and F100B(H)A (BioBasic Inc., Toronto, Ontario). For the 2F5 Fab delta CDR H3 mutant, residues 100 TLFGVPI$^{100}$ to 100 Ser-Gly forming the top part of the CDR H3 loop were replaced by a simple Ser-Gly dipeptide linker. Resulting plasmids were transfected into Rosetta-gami 2 (DE3) competent cells following a standard heat shock protocol and 2F5 Fab expression was subsequently induced with 0.1 mM IPTG. Cells were grown at 18°C for 36 hours in either Superbroth or Terrific Broth media containing 1mM of each chloramphenicol, ampicillin, streptomycin and tetracycline. After centrifugation, cells were re-suspended in lysis buffer (50 mM Hepes, pH 7.5, 500 mM NaCl, 5 mM imidazole, 5% glycerol and 100 mM iodoacetamide) containing DNase, lysozyme and EDTA-free Complete protease inhibitor (Roche, Mannheim, Germany). Cell lysis was performed by sonication. The cell lysate was loaded on a nickel affinity chromatography column (Qiagen, Mississauga, Ontario). After washing, the protein was eluted with 250 mM imidazole and further purified using a HiLoad Superdex 75 prep grade gel filtration column (GE Healthcare, Piscataway, NJ). Figure 4.1 shows a SDS-PAGE and Western Blot analysis documenting the 2F5 Fab content in different fractions isolated during the purification steps. From 12 liters of media culture, two to five milligrams of Fab were obtained reproducibly. Purity was estimated by semi-quantitative analysis of SDS-PAGE as approximately 80% using the ImageJ software (6). Proper folding of the Fab fragments was confirmed by circular dichroism in a Jasco J-810 spectropolarimeter.
Figure 4.1 Expression and purification of recombinant 2F5 Fab. A. SDS-PAGE electrophoresis. B. Western Blot using an anti-kappa light chain-HRP conjugated secondary antibody for specific detection. TB (Terrific Broth); IB (Inclusion Bodies); FT (Flow Through).

4.2.3 Expression and purification of non-glycosylated gp41 construct

The non-glycosylated gp41 ectodomain construct used in this study was designed based on the one previously described by Tan et al. (265). It spans residues 535-669 (which includes the 2F5 epitope) of gp41, with an SGGRGG-linker replacing native residues 580-627. An E. coli-codon-
optimized DNA sequence for this construct was synthesized (BioBasic Inc., Toronto, Ontario) and cloned into the pEt-26b vector (Novagen, Gibbstown, NJ) using the Ndel and XhoI restriction sites. As a result, the LE(H)6x sequence was added at the C-terminus of the gp41 construct. The plasmid was transfected into *E. coli* BL21(DE3) cells using a standard heat shock protocol. Cells containing the plasmid were grown in kanamycin-containing LB broth, induced with 0.1 mM IPTG and grown at 37°C overnight. After centrifugation, the cells were resuspended in a Tris-buffer saline (TBS) solution. Cell lysis was performed by sonication in the presence of lysozyme and DNase. After centrifugation, inclusion bodies were washed with TBS, centrifuged and subsequently dissolved in binding buffer (50 mM Hepes, pH 7.5, 500 mM NaCl, 5 mM imidazole, 5% glycerol) containing 8 M urea. After centrifugation, the supernatant was loaded directly on a nickel affinity chromatography resin (Qiagen, Mississauga, Ontario). After washing, the protein was eluted with the binding buffer solution containing 250 mM imidazole. Fractions containing the gp41 construct were collected, combined and dialyzed against 2 liters of 20 mM formate, pH 3.0 at 4°C for 16 hours. Figure 4.2 shows a SDS-PAGE analysis documenting the protein content in different fractions isolated during the purification steps. Proper folding of the gp41 construct into its post-fusion bundle of predominantly α-helical secondary structure was verified by circular dichroism using a Jasco J-810 Spectropolarimeter.

![Figure 4.2 SDS-PAGE electrophoresis analyzing the expression and purification of recombinant gp41 ectodomain. IB (Inclusion Bodies); Ni (Nickel); FT (Flow Through).](image-url)
4.2.4 Production of vesicles

Large unilamellar vesicles (LUVs) containing POPC, Chol and PA (2:1:0.6) were prepared by the extrusion method (119, 122). For extrusion, the mini-extruder set (Avanti-Polar Lipids, Birmingham, AL) with Whatman Nucleopore membranes with pores of 0.4, 0.2 and 0.1 μm was used according to the manufacturer’s protocol. For producing peptido-liposomes, 2F5preTM peptide stock solution in DMSO was added to the LUVs at a 1:100 peptide-to-phospholipid ratio. Partitioning of peptides on the surface of vesicles using this method has been previously characterized (122).

4.2.5 Isothermal titration calorimetry

Calorimetric titration experiments were performed with a high-sensitivity VP-ITC MicroCalorimeter (MicroCal, Northampton, MA) at 25°C. Proteins were dialyzed in buffer 5 mM HEPES and 100 mM NaCl, pH 7.4 overnight at 4°C. Protein and peptide solutions were degassed for 5 min before measurements. The antibody solution (3 – 5 μM) in the calorimetric cell was titrated with 2F5ep peptide dissolved in the same buffer (concentration 20 – 40 μM) in successive injections of 10 – 12 μL. The corresponding heat of peptide dilution into buffer was used to correct the data. The experimental data were analyzed according to a 1:1 binding model by means of Origin 7.0 (MicroCal Inc.). The fit of the binding curve yields the stoichiometry, the association constant, $K_a$, and the enthalpy, $\Delta H$, of the binding reaction. The free energy of binding $\Delta G$ and the entropy $\Delta S$ are determined by the basic thermodynamic expressions: $\Delta G = -RT\ln K_a = \Delta H - T\Delta S$, where $R$ and $T$ are the gas constant and the absolute temperature, respectively.

4.2.6 Cell-cell fusion and neutralization assays

Syncytium-formation assays were carried out using CHO-Env and TZM-bl as effector and target cells, respectively (obtained from the AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH, and contributed by C. Weiss and J. White and J. Kappes, respectively). Fusion was inhibited by incubating the 2F5 Fab constructs with CHO-Env cells for 90 minutes prior to co-culturing them with TZM-bl cells following the contributors' specifications.
Pseudoviruses were produced by transfection of human kidney HEK293T cells with the full-length env clone pHXB2-env (AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH, contributed by K. Page and D. Littman) using calcium phosphate, together with vectors pWXLP-GFP and pCMV8.91, encoding respectively a green fluorescent protein and an env-deficient HIV-1 genome (kindly provided by Dr. Patricia Villace, CSIC, Madrid). After 24 h, the medium was replaced with Optimem-Glutamax II (Invitrogen Ltd, Paisley, UK) without serum. Two days after transfection, the pseudovirus particles were harvested, passed through 0.45 μm pore sterile filters (Millex® HV, Millipore NV, Brussels, Belgium) and finally concentrated by ultracentrifugation in a sucrose gradient. Fab-induced neutralization was determined using TZM-bl target cells. Serial dilutions of Fabs were set up in duplicate in 96-well plates, and incubated for 1 h at 37°C with a 10-15 % tissue culture infectious dose of pseudovirus. After Fab-pseudovirus co-incubation, 10,000 target cells were added in the presence of 15 μg/ml DEAE-dextran (Sigma-Aldrich, St-Louis, MO). Neutralization levels after 72 hours were inferred from the reduction in the number of GFP-positive cells as determined by flow cytometry using a BD FACScalibur Flow Cytometer (Becton Dickinson Immunocytometry Systems, Mountain View, CA).

4.2.7 ELISA binding assays

4.2.7.1 Direct ELISA binding assays for peptides/proteins

Binding of the recombinant 2F5 Fab to its gp41 epitope was evaluated by direct ELISA with standard 96-well plates (Cayman Chemical Company, Ann Arbor, MI) coated with gp41 ectodomain constructs. Briefly, 50 μl of the recombinant gp41 at a concentration of 20 μg/ml in 0.1 M sodium bicarbonate, pH 9.6 was incubated at room temperature for 2 hours to coat the 96-well plates. Subsequently, the plates were blocked with 200 μl of 5% non-fat milk dissolved in TBS at 4°C for 16 hours. Serial dilutions of 2F5 Fab ranging from 100 μg/ml to 0.001 μg/ml were set up in triplicates and added to the gp41-coated wells. After 2 hours of incubation at room temperature and thrice washing with 200 μl of TBS, the secondary anti-human IgG (Fab specific)-alkaline phosphatase antibody produced in goat (Sigma-Aldrich, St-Louis, MO) was added to the plates at a 1:30,000 dilution in blocking buffer. Plates were subsequently incubated for 2 hours at room temperature. After washing, the reaction was developed by adding 100 μl per well of Alkaline Phosphatase Yellow (pNPP) Liquid Substrate System for ELISA (Sigma-
Aldrich, Oakville, Ontario) and analyzed after 30 minutes using a Spectramax 340PC instrument (Molecular Devices, Sunnyvale, CA) at a wavelength of 405 nm.

4.2.7.2 Direct ELISA binding assays for membrane components

Direct ELISA assays of binding to membrane components were performed following a slightly modified protocol from the one described previously (176). Briefly, chloroform stock solutions of POPC, cholesterol and PA were combined and diluted with ethanol to 10-100 nmol/ml; a range of 1-10 nmol/well was used as the antigen. Control wells contained 0.05 ml of the appropriate solvent to match the lipid plated, and readings from these control wells were subtracted from the readings of matching experimental wells for analysis. After evaporation overnight, the plates were blocked with TBS-0.3% gelatin for 2 h. Subsequent steps were performed as described above, with the exception that the detection could only be performed after an overnight incubation with the pNPP substrate (as opposed to a one hour incubation) due to a weak initial signal.

4.2.7.3 Competitive ELISA binding assays

Competitive ELISA assays were performed using the coating and blocking steps outlined above. Prior to adding to the gp41 coated plates, 1 µg/ml of 2F5 Fab was pre-incubated for a minimum of 30 minutes with serial dilutions of competitor ranging from 1 mM to 10 nM and set-up in triplicates. Competitors included: gp41 constructs in 0.1 M sodium bicarbonate, pH 9.6, ELDKWAS peptide in 20 mM Tris, pH 8.0, 2F5ep peptide in DMSO, 2F5preTM peptide in DMSO, 2F5preTM(9,10)Ala peptide in DMSO, POPC: Chol: PA (2:1:0.6) LUVs, and POPC:Chol:PA (2:1:0.6) LUVs with 2F5preTM peptide,. All LUVs were in a solution of 5 mM Hepes, pH 7.4, 100 mM NaCl. Incubation with secondary antibody and development with pNPP were performed as described for the direct ELISA protocol outlined above.

4.2.7.4 Analysis of ELISA data

ELISA data were analysed with the help of the GraphPad Prism 5 software. For competitive ELISA, absolute values were normalized and converted to percent of maximum response. Sigmoidal dose-response and log[inhibitor] vs. response non-linear regression analyses were used to derive EC<sub>50</sub> and IC<sub>50</sub> values for direct ELISA and competitive ELISA results, respectively. Finally, a one-way analysis of variance (ANOVA) followed by a Bonferroni post-
analysis test (significance level of 0.05) was used for the statistical analysis of IC$_{50}$ values obtained for the different Fab constructs.

4.3 Results

4.3.1 Supply of recombinant 2F5 Fab

2F5 Fab was expressed in *E. coli* Rosetta-gami 2 (DE3) cells. To characterize the role of the 2F5 CDR H3 apex residues, a 2F5 Fab mutant was designed (termed delta CDR H3), in which a Ser-Gly dipeptide linker replaces the seven CDR H3 apex residues 100-100F (Figures 4.3A and 4.3B). In a second mutant (termed F100B(H)A), the phenylalanine at position 100B of the heavy chain, located right at the tip of the CDR H3 loop, was replaced with an alanine, a minor change of overall hydrophobicity (Figure 4.3A). Finally, an Fab mutant with an unrelated threonine to alanine mutation at position 15 of the heavy chain (termed T15(H)A) was created to be used as a negative control (Figure 4.3A). Nickel-affinity and size-exclusion chromatography of native and mutant 2F5 Fabs resulted in the isolation of two to five milligrams per 12 liters of culture media of correctly folded Fabs of approximately 80% purity (Figures 4.3C and 4.3D).
Figure 4.3 Recombinant 2F5 Fab constructs. A. 2F5 light and heavy chain sequences of the variable region. FWR labels framework regions, whereas CDR marks complementary determining regions. 2F5 residues interacting with the core linear epitope, as inferred from the crystal structure (PDB ID: 3D0L), are marked by an *. The 2F5 Fab mutants used in this study (T15(H)A negative control, F100B(H)A and delta CDR H3) are indicated by arrows and boxes, respectively. B. Design of the 2F5 CDR H3 mutants. For the delta CDR H3 mutant, the Ser-Gly dipeptide linker bridges well the distance between the two structured endpoints of amino acids T99 and A100G. The figure was generated with the program Pymol (70) using the crystal structure 3D0L. C. SDS-PAGE gels and Western Blots showing the expression and purification of the 2F5 Fab constructs. D. Circular dichroism measurements of the 2F5 Fab constructs show the correct folding of all Fabs, with the expected signal for predominantly β-sheet secondary structure. [θ] is the molar ellipticity with units of deg*cm²/decimole.
4.3.2 Thermodynamics of 2F5ep binding to 2F5 Fab constructs

To test for the epitope-binding capacity of the recombinant 2F5 Fab and its mutants, isothermal titration calorimetry measurements were performed. For this purpose, the peptide 2F5ep (656NEQELLELDKWASLWN671) representing the nominal 2F5 epitope, was titrated into a solution containing the 2F5 Fab constructs. The heat signal produced following epitope-peptide injection allowed the determination of the thermodynamic parameters of binding. Titration profiles are shown for 2F5ep/wild-type 2F5 Fab and 2F5ep/delta CDR H3 2F5 Fab (Figures 4.4A and 4.4B, respectively) and the thermodynamic parameters of all 2F5 Fab constructs are presented in Table 1. The binding of this complex is energetically driven by exothermic enthalpy (\(\Delta H < 0\)) counterbalanced by unfavourable entropic (\(\Delta S < 0\)) changes. Both the enthalpic and entropic terms contribute almost in the same way to the free energy of 2F5ep binding to wild-type 2F5 Fab and its mutants, resulting in very similar and tight binding affinities (\(K_d \approx 20\) nM). These results constitute strong evidence that 2F5 CDR H3 apex residues are not involved in mediating core gp41 epitope interactions.

![Figure 4.4 Isothermal titration calorimetry.](image)

**Figure 4.4 Isothermal titration calorimetry.** Isotherms of 2F5ep (656NEQELLELDKWASLWN671) binding to (A) wild-type 2F5 Fab and (B) 2F5 Fab delta CDR H3. Upper panel: the heat released upon consecutive injections of 12 \(\mu\)L of the 2F5ep solution (26 \(\mu\)M) into native or delta CDR H3 2F5 Fab (3 \(\mu\)M) in the calorimetric cell. Lower panel: integrated heats (symbols) and non-linear least-squares fit (line) of the data with a 1:1 binding stoichiometry.
The derived thermodynamic parameters of binding for these complexes as well as the other 2F5 Fab mutants are listed in Table 4.1.

Table 4.1 Thermodynamic parameters of 2F5ep peptide (\textsuperscript{656}NEQELLELDKWASLWN\textsuperscript{671}) association to wild type 2F5 Fab and CDR H3 apex mutants. Values are means ± standard deviations.

<table>
<thead>
<tr>
<th>Complex</th>
<th>Stoichiometry</th>
<th>$K_d$ (nM)</th>
<th>$\Delta H$ kcal/mol</th>
<th>$\Delta S$ kcal/mol</th>
<th>$\Delta G$ kcal/mol</th>
</tr>
</thead>
<tbody>
<tr>
<td>2F5ep/wild type</td>
<td>1.03 ± 0.01</td>
<td>20 ± 1.5</td>
<td>-11.8 ± 0.01</td>
<td>1.4 ± 0.1</td>
<td>-10.4 ± 0.1</td>
</tr>
<tr>
<td>2F5ep/T15(H)A</td>
<td>0.99 ± 0.01</td>
<td>24 ± 4</td>
<td>-12.1 ± 0.2</td>
<td>1.8 ± 0.1</td>
<td>-10.3 ± 0.1</td>
</tr>
<tr>
<td>2F5ep/F100B(H)A</td>
<td>0.99 ± 0.01</td>
<td>21 ± 6</td>
<td>-11.6 ± 0.2</td>
<td>1.2 ± 0.2</td>
<td>-10.4 ± 0.2</td>
</tr>
<tr>
<td>2F5ep/delta CDR H3</td>
<td>1.1 ± 0.004</td>
<td>23 ± 2.4</td>
<td>-14.5 ± 0.1</td>
<td>4.1 ± 0.2</td>
<td>-10.5 ± 0.2</td>
</tr>
</tbody>
</table>

4.3.3 Neutralization assays

In an attempt to link linear epitope binding capacity with functionality of the 2F5 Fab constructs, two types of assays were carried out. Performing both cell-cell fusion inhibition assays and pseudovirus neutralization assays allowed us to ascertain the activity of Fabs independently of the system or the detection methods used. As seen in Figures 4.5A and 4.5B, the recombinant 2F5 Fab possesses the ability to abrogate gp41-mediated fusion with an approximate IC\textsubscript{50} of $10^{-7}$ M and $10^{-6}$ M in the pseudovirus infection and cell-cell mediated fusion assays, respectively. As expected, the negative control mutant possessing the T15(H)A mutation, a site unrelated to epitope recognition, is able to neutralize Env-mediated fusion at the same level as the wild-type 2F5 Fab. However, the Fab with the F100B(H)A mutation at the tip of the CDR H3 achieves half-neutralization only at a concentration nearly twentyfold that of wild-type 2F5 Fab. Moreover, very limited, if any, inhibition of gp41-mediated fusion is observed with the delta CDR H3 Fab mutant at the highest concentration tested (i.e. 4 μM). Altogether, these striking functional results clearly indicate the importance of the proper residues at the 2F5 CDR H3 apex for 2F5 inhibition of Env-mediated fusion.
Figure 4.5 Neutralization and cell-cell fusion inhibition assays with the different 2F5 Fab constructs. In these assays, pseudovirus or effector cells are pre-incubated with the recombinant 2F5 Fab constructs and fusion events are monitored after incubation with TZM-bl target cells. In both cases, wild-type 2F5 Fab and T15(H)A negative control are able to inhibit fusion events effectively, with an approximate IC₅₀ of 10⁻⁷ M and 10⁻⁶ M in the pseudovirus infection and cell-cell mediated fusion assays, respectively. The 2F5 Fab CDR H3 apex mutant F100B(H)A inhibits fusion 20-fold less than the native 2F5 Fab, whereas mutant delta CDR H3 is almost completely unable to inhibit Env-mediated fusion in the concentration range tested. A. Pseudovirus infection assays. The infection of TZM-bl target cells is monitored by flow cytometry as the GFP fluorescence signal of infected cells. B. Cell-cell fusion assays. CHO-Env effector cells are allowed to interact with TZM-bl target cells and fusion events are measured by the number of nuclei in syncytia after incubation. Graphs in both panels display means of 6 measurements in 3 independent experiments ±SD.

4.3.4 Binding affinities of recombinant 2F5 Fab and its mutants determined by ELISA

To support the ITC findings and to also allow for probing the binding of more complex systems, ELISA binding assays were performed. First, direct binding ELISA experiments were carried out
by coating two gp41 constructs on ELISA plates: 1) a glycosylated gp41 ectodomain comprising residues 541-682 and 2) a non-glycosylated gp41 construct produced in *E. coli* cells comprising residues 535-669, with an SGGRGG linker replacing the inter-helix residues 580-627. As seen in Figures 4.6A and 4.6B, the recombinant 2F5 Fab constructs bind to both gp41-coated plates. Interestingly in this system, modifications of the CDR H3 apex seem to reduce the ability of the 2F5 Fab to bind to the gp41 constructs coated on the ELISA plates. Furthermore, the extent of the binding difference seems to be influenced by which of the two gp41 constructs is coated on the plate. It is unclear if this observed difference in binding is due to the presence/absence of glycans on gp41, extra residues at the C-terminus, or less likely to the deletion of the inter-helix loop.

**Figure 4.6 Direct and competitive ELISA assays measuring the binding of recombinant 2F5 Fab constructs to two gp41 construct.**

A. Direct ELISA assays of the 2F5 Fab constructs binding to a glycosylated gp41 construct spanning residues 541-682. Clear differences in binding are observed for the 2F5 CDR H3 mutants compared to native 2F5 Fab, with the delta CDR H3 mutant almost unable to bind this construct in the concentration range tested. B. Direct ELISA assays of the 2F5 Fab constructs binding to a non-glycosylated gp41 construct spanning residues 535-669. Differences in binding to this gp41 construct are also observed for the 2F5 CDR H3 mutants, although to a smaller extent than for the glycosylated gp41 construct. C. and D. Competitive ELISA experiments
measuring the binding affinity of the different 2F5 Fab constructs to a glycosylated gp41 construct spanning residues 541-682 and a non-glycosylated gp41 construct spanning residues 535-669, respectively. In both cases, although saturation is not reached with the concentrations used and so an exact IC_{50} cannot be determined, it is possible to observe that all 2F5 Fab constructs have a similar apparent binding affinity. This discrepancy between direct and competitive ELISA results is thought to arise from the binding artifacts of the adsorbed epitope to the solid-phase ELISA plate (see text for a more complete discussion). From these results, we believe that competitive ELISA is a better experimental system for investigating relative binding affinities of the 2F5 Fab constructs to various epitopes.

Next, competitive ELISA experiments were performed to assess the binding affinity of the different 2F5 Fab CDR H3 constructs to gp41 in solution. In this protocol, the 2F5 Fab constructs were pre-incubated with a gp41 competitor before being added to the gp41-coated plates. The non-glycosylated gp41 (535-669) construct was selected for coating the ELISA plates in these competition ELISA assays because of the high signal observed in the previous direct ELISA experiments. As seen in Figures 4.6C and 4.6D, no significant difference in binding was observed between the various 2F5 Fabs in competition ELISAs with both gp41 constructs. Although the apparent binding affinity is low in solution (no saturation of the signal is observed even at high μM concentrations of the constructs), it is clear from the competitive ELISA curves that all four Fabs behave in a similar way. The weak binding affinity observed for this gp41 construct stems from its six-helix bundle post-fusion conformation, which is a low affinity conformation for recognition by 2F5, as previously described (90). These competitive ELISA results are in sharp contrast to what was observed in direct ELISA experiments. This difference between direct and competitive ELISA results for the behavior of the 2F5 Fab mutants strongly suggests that artifacts might be observed for the binding to the epitope adsorbed on the plate surface. Indeed, the height of the maximum signal observed for each Fab in competitive ELISAs was proportional to the apparent binding affinity observed in direct binding ELISAs (data not shown). This suggests that the 2F5 Fab CDR H3 mutants might have difficulty to access their adsorbed core epitope. For this reason, competitive ELISA was the method of choice to subsequently investigate the relative binding affinities of the 2F5 Fab constructs to various epitopes.
4.3.5 Relative binding affinity of recombinant 2F5 Fab constructs to different epitopes as determined by competitive ELISA

Competition ELISA experiments were conducted with a variety of epitopes in an attempt to characterize the role played by the 2F5 CDR H3 apex residues. First, competitive ELISA experiments with the ELDKWAS and 2F5ep peptides were performed. As expected, no significant differences in apparent binding affinities (IC50) were observed between the four Fabs with these peptides (Figures 4.7A and 4.7B). This finding supports the results obtained by ITC and suggests that residues at the apex of the 2F5 CDR H3 do not affect binding to the core linear epitope. Subsequently, C-terminal residues 672-683 were added to the competitor peptide making it a mimic of the full gp41 pre-transmembrane segment (2F5preTM). With this extended construct, statistically significant differences in IC50 of approximately one order of magnitude were observed between the wild-type 2F5 Fab and Fab mutant F100B(H)A, and of approximately two orders of magnitude between wild-type 2F5 Fab and Fab delta CDR H3 (Figure 4.7C). These results suggest that residues at the apex of the 2F5 CDR H3 loop play an important role in mediating interactions with the hydrophobic C-terminus of this extended gp41 epitope peptide. To verify the specificity of 2F5 recognition in this system, the negative control peptide 2F5preTM(9,10)Ala, with two of the core epitope residues Asp664 and Lys665 replaced by alanines, was used as a competitor and showed no apparent binding to the recombinant 2F5 Fab constructs up to the high μM range (data not shown). Thus, contribution of the CDR H3 apex to the overall 2F5preTM binding process appears to depend on the initial specific recognition of the core epitope sequence. However, some caution must be exercised when interpreting these results due to the unknown folding state of this aromatic-rich peptide in aqueous solution (10% DMSO), e.g. the breaking up of hydrophobic peptide aggregates by the hydrophobic CDR H3 apex residues would in turn favor core epitope binding. Other possible roles of CDR H3 residues in mediating interactions with this C-terminally extended MPER peptide include unmasking the core 2F5 epitope from possible hydrophobic interactions with C-terminal residues; mediating conformation changes of the peptide leading to 2F5 recognition of the core epitope in a β-turn conformation; and/or decreasing the off-rate of Fab-peptide binding.
Figure 4.7 Competitive ELISA binding assays of the 2F5 Fab constructs with various epitopes. Top panels show sample binding curves obtained from competitive ELISA assays for the different 2F5 Fab constructs from which IC_{50} values are calculated. Bottom panels summarize the IC_{50} values calculated from multiple experiments, which allow for a statistical analysis. All plates are coated with the recombinant gp41 construct. A. The linear peptide ELDKVAS was used as a competitor. As seen in both binding curves and derived IC_{50} values, wild-type 2F5 Fab and its mutants bind similarly to this core epitope. B. The linear peptide 2F5ep (656NEQELLELDKWASLWN671) was used as a competitor. Similarly, all 2F5 Fab constructs bind with comparable apparent affinity to this epitope, confirming the ITC data. C. In this experiment, the hydrophobic linear epitope 2F5preTM (656NEQELLELDKWASLWNWFNITNWLYIK683) is used as a competitor. In this case, significant differences in binding affinity are observed for the 2F5 Fab CDR mutants when compared to the wild-type 2F5 Fab. This suggests
a possible role for CDR H3 residues in interacting with residues located at the C-terminus of the gp41 MPER. Whether these CDR H3 apex-mediated interactions are specific or not cannot be inferred from these results (see text for a more complete discussion). D. An extended epitope consisting of the 2F5preTM peptide inserted in POPC:Chol:PA LUVs was used as a competitor. A significant difference in binding affinities is observed between all 2F5 Fab CDR H3 constructs when interacting with this more complex epitope. Altogether, these results imply a role for residues located at the apex of the 2F5 CDR H3 in mediating interactions with residues located at the C-terminus of the gp41 MPER, or/and with membrane components in the context of core epitope binding. A one-way analysis of variance (ANOVA) followed by a Bonferroni post-analysis test was used for the statistical analysis of IC50 values obtained for the different Fab constructs and the significance level where p < 0.05 is indicated on the graph.

Furthermore, when the 2F5preTM peptide was inserted into the POPC:Chol:PA LUVs, the recombinant 2F5 Fab constructs showed even more significant differential binding (Figure 4.7D). Core epitope recognition in the context of this liposome system yielded statistically significant differences in IC50 values of approximately one order of magnitude between the wild-type 2F5 Fab and Fab mutant F100B(H)A, and of almost three orders of magnitude between wild-type 2F5 Fab and Fab delta CDR H3. Also of note, overall binding of Fabs to the 2F5preTM peptide-liposome system was two orders of magnitude lower than Fabs binding to the 2F5preTM peptide alone. Whether this results from the concealed nature of the 2F5preTM peptide in a membrane environment decreasing its accessibility for 2F5 Fab recognition remains unclear.

4.3.6 2F5 Fab interaction with membrane components

The ability of the Fabs to interact with membrane components was assessed both by direct and competitive ELISA assays. For the direct assays, ELISA plates were coated with POPC, cholesterol and PA and both native 2F5 Fab and the CDR H3 mutants display affinity for the coated membrane components (Figure 4.8A). Modifying the hydrophobicity of the 2F5 CDR H3 by either point mutation or deletion of the extended loop therefore did not abrogate affinity for the coated membrane components. Interestingly, it seems like the deletion of bulky hydrophobic residues in the extended 2F5 CDR H3 loop allowed for a better interaction between Fab and coated membrane components. Also of notice in these experiments, the substrate development reaction needed to be carried out for an extended period of time before a signal could be
observed, even when high amounts of membrane components were coated on the ELISA plates, suggesting that these Fabs possess a rather low affinity for coated membrane components. Figure 4.8B depicts the results of competitive ELISA assays using POPC:Chol:PA LUVs as competitor. Up to the high μM range, these vesicles were not able to compete-off binding of the 2F5 Fab constructs to the gp41-coated plates. This finding further supports the previous observation that 2F5, when probed independently of core epitope binding, possesses a low affinity for membrane components (see for comparison Figure 4.7D).

**Figure 4.8 Direct and competitive ELISA assays measuring the binding of recombinant 2F5 Fab constructs to membrane components.** A. Direct ELISA binding assays of native 2F5 Fab and its mutants to lipid-coated plates (POPC, Chol and PA). All Fabs show interaction with the coated membrane components. Interestingly, deletion of the 2F5 Fab CDR H3 loop appears to slightly increase the ability of this Fab to interact with coated membrane components (see text for discussion). These results are representative of three measurements performed in triplicates on different ELISA plates and coated with different amounts of membrane components. B. Competitive ELISA experiments probing the binding of Fabs to gp41-coated plates after a pre-incubation with POPC:Chol:PA (2:1:0.6) LUVs competitor. No competition is observed for this membrane bilayer competitor up to the high μM range for any 2F5 Fab constructs, demonstrating a low affinity of 2F5 Fab for membrane components independently of core epitope binding.

### 4.4 Discussion

The production of high affinity antibodies to antigens is an essential step in protecting the host against undesired pathogens. In the course of HIV-1 infection, antibodies against HIV-1 epitopes are generated readily, although antibodies capable of neutralizing the virus generally only appear
after prolonged infection (76, 236, 253, 259). Moreover, only in exceptional cases have infected individuals generated monoclonal antibodies that are capable of neutralizing HIV-1 potently across a broad range of clades. One of these antibodies, bnmAb 2F5, has been extensively characterized in an effort to determine the molecular mechanism underlying its broad neutralization capability and with the objective to guide vaccine immunogen design efforts. Although the core recognition site of 2F5 has been mapped to the \textsuperscript{662}ELDKWAS\textsuperscript{669} MPER linear sequence of gp41, and the atomic interactions of the core antigenic complex have been described in detail (36, 134, 204), the exact mechanism of broad HIV-1 neutralization by 2F5 remains elusive. Interestingly, no contacts have been identified between the core gp41 epitope and the central seven amino acids of the 22 residue-long 2F5 CDR H3 loop. The importance of these CDR H3 residues for neutralization has been previously documented (305). Many publications have hypothesized that this region of 2F5 might be involved in a secondary role other than core epitope binding (7, 109, 134, 176, 204, 233, 234). Indeed, it has been reported that antibodies elicited during the course of HIV-1 infection might display elevated polyreactivity when compared to antibodies elicited during the course of other infections (200). It has also been suggested that the length of the CDR H3 loop, which plays a distinct role in determining antigen specificity, is related to the type of antigen recognized; antibodies raised against large antigens, such as viruses have a tendency to have longer CDR H3 loops than antibodies responsive to smaller antigens, such as peptides (57, 130, 245). It is hypothesized that antibodies with longer CDR H3 loops have an extended binding site that allows them to insert into cavities within an antigen (245). The 2F5 CDR H3 might indeed play a crucial role in recognizing a recessed conserved epitope, which stems from both its location at the membrane-partitioning interface of the virus and from the potential oligomerization of the gp41 MPER.

In the present study, a 2F5 Fab mutant with replacement of CDR H3 residues 100-100F by a Ser-Gly dipeptide linker, delta CDR H3, and the single-site 2F5 Fab mutant F100B(H)A were purified in bacterial cells. Both cell-cell fusion inhibition and pseudovirus neutralization assays showed almost complete loss of neutralization capacity by the delta CDR H3 2F5 Fab mutant and more than an order of magnitude reduction in the IC\textsubscript{50} of neutralization by the F100B(H)A mutant. These assays provided an additional confirmation of both the functionality of the produced recombinant 2F5 Fabs and the crucial role the residues at the apex of the 2F5 CDR H3 loop play in virus neutralization.
In an attempt to characterize the binding determinants of the 2F5 CDR H3 apex residues that confer neutralizing capacity on 2F5, ITC, direct ELISA and competitive ELISA measurements were performed. ITC measurements of the different 2F5 Fab constructs clearly indicated that binding to the nominal epitope peptide was not affected by mutations at the CDR H3 apex and that affinity for all 2F5 Fab constructs was in the low nM range. Competitive ELISA measurements using core peptides as competitors also showed no difference in relative binding affinity for the different 2F5 Fab constructs. These results are in agreement with the structural data that show that the molecular interactions between the 2F5 paratope and core gp41 epitope residues are mediated by 2F5 amino acids other than those at the apex of the CDR H3 loop (36, 134, 204). Altogether, we believe that the present neutralization and binding affinity results represent strong evidence indicating that the 2F5 CDR H3 apex residues are essential for neutralization but are not involved in core epitope binding. Therefore, this suggests one or more secondary sites of interaction between 2F5 and its HIV-1 target.

In this study, we report differences in binding profiles between direct ELISA and competitive ELISA experiments for various 2F5 Fab mutants assayed with the same gp41 constructs. Indeed, 2F5 CDR H3 mutants were observed to bind differently to gp41 constructs in direct ELISA experiments, whereas they showed a similar binding profile in competitive ELISA assays. Recognition of antibodies to the solid phase adsorbed antigen may be limited by multiple factors, including the conformation of the antigens on the plate surface and steric repulsion and/or attractive interactions between the antibody molecules themselves (195, 202, 274). It is therefore not unlikely that removing a flexible hydrophobic moiety at the apex of the 2F5 antigen binding site would cause artefacts in apparent binding affinity to the gp41 epitope bound to the solid phase. In contrast, in competitive ELISA, the free antibody available for binding to the adsorbed antigen is determined by the binding constant of the antibody to the competitor in solution (111). This allows for a more accurate comparison of relative binding affinity between antibodies. Furthermore, this process is free of the mass transport limitations present in direct ELISA assays (195, 201). Taken together, both direct and competitive ELISA results form supporting evidence that the 2F5 CDR H3 apex is important for mediating the access to a possibly recessed epitope, although any underlying specificity of the 2F5 CDR H3 apex residues is not located within the 2F5 core epitope itself.
In an attempt to determine if the 2F5 CDR H3 apex residues contributed to specific interactions with other components than its core epitope, additional ELISA assays were performed. The 2F5preTM(9,10)Ala peptide, which has both core epitope Asp and Lys residues mutated to alanines was used as a competitor in our ELISA assay; experiments in which plate-bound gp41 is the 2F5 capture entity revealed no such competition for 2F5 Fab binding up to high \( \mu \text{M} \) concentrations (data not shown). Similar results were obtained when POPC:Chol:PA liposomes alone were used as a competitor (Figure 4.8B). Together, these results indicate that 2F5 possesses low binding affinity, if any, for residues located at the C-terminus of the gp41 MPER alone or for membrane components alone. In addition, direct ELISA experiments suggest that the 2F5 hydrophobic CDR H3 apex contributes very little to interactions with lipid components on their own. In fact, it seems like removing the bulky 2F5 CDR H3 extended loop might even slightly improve the interaction between the Fab and lipid-coated plates, possibly by allowing easier access or by encouraging other types of interactions to take place. Such contacts might include ionic interactions between positively charged regions of the Fab and negatively charged phospholipid components of membranes, an interaction which has previously been proposed to be important in mediating possible viral membrane contacts by anti-MPER bnAbs (8, 134, 293). However, one should note that the binding observed in the present direct ELISA experiments probe the affinity for membrane components alone and outside of a bilayer environment and as such might be significantly limited in their relevance for the description of binding events \textit{in vivo}.

The competitive ELISA experiments show that the hydrophobic CDR H3 apex residues only make a definitive contribution to binding when the residues \(^{672}\text{WFNITNWLYIK}\(^{683}\) are added to the gp41 MPER C-terminus and – even more so – if the resulting extended peptide is placed in a membrane bilayer environment. Indeed, quite significant differences in binding affinities between the native 2F5 Fab and its CDR H3 mutants are observed for the extended epitope peptide 2F5preTM, both in solution and in a liposome environment. Taken together, our results and prior reports of weak interactions between 2F5 and both of these components individually (7, 109, 168, 176, 234, 279) allow us to hypothesize that the necessity of the apex of the elongated CDR H3 loop for 2F5 neutralization is caused by secondary interactions of much weaker affinity to either C-terminal MPER residues or components of a membrane bilayer, or both, in the context of the energetically dominant core epitope binding. Indeed, a dual interaction
of the extended CDR H3 of 2F5 with both membrane surfaces and C-terminally-located gp41 MPER residues is feasible since interfacial hydrophobicity is mainly based on aromatic and leucine residues, which can also contribute to establishing interactions among sequences embedded in a membrane milieu (231).

Recently, three other laboratories have reported on variations in HIV-1 neutralization capacity of both 2F5 and 4E10 CDR H3 mutants with different hydrophobicity characteristics (8, 203, 293). All studies, including the one presented here, agree that the hydrophobic CDR H3 extended loops of 2F5 and 4E10 are critical for HIV-1 neutralization, although the proposed mechanisms by which this effect is to be achieved are different. From SPR measurements of 2F5 and 4E10 mutants, Alam et al. deduce a two-step neutralization model, which postulates that the bnmAb first pre-concentrates on the membrane surface mediated by hydrophobic residues in its extended CDR H3 loop. In a second step, the bnmAb then interacts with its transiently exposed core epitope (8). A competing model presented by Xu et al., and also suggested previously by others (256, 293), argues that the CDR H3 extended loop of 4E10, in conjunction with positively charged pockets near the paratope, is responsible for the extraction of the gp41 core epitope partitioned at the membrane interface. The data gathered in the present study are consistent with the hydrophobic potential of the residues contained in the extended CDR H3 loop of 2F5 playing a measurable role in antigen binding only after core epitope recognition has occurred, contributing in a rather limited fashion prior to the latter high affinity interaction. Additional interactions of bnmAb 2F5 with hydrophobic amino acids located C-terminally of the gp41 MPER and/or with lipid components in the context of core epitope binding could allow for a metastable immunological complex to form and prevent the fusion cascade from proceeding. Certainly for 2F5, the answer to the question whether these additional hydrophobic interactions contribute to the extraction of its immersed MPER epitope, or whether they support the prevention of re-partitioning of the MPER to the membrane after being transiently exposed during the initiation of the fusion process, requires further investigation.

Although humans possess the ability of making long CDR H3 loops, the unusual origins of the D(H) segment of bnmAbs 2F5 (52 bp-long) and 2G12 (31 bp-long) suggest that making antibodies with specific CDR H3 loops effective in neutralizing HIV-1 is challenging for the immune system (148, 289). For 2F5, clues as to the requirement of such an unusual CDR H3 might be found in the present study, which clearly shows that the high affinity of 2F5 for its
primary epitope is not sufficient for neutralization but that one or more weaker secondary binding determinants might be required for this crucial activity, a kind of “intra-molecular avidity effect”. The present finding also helps to explain why past immunization protocols based on the presentation of the constrained core epitope of 2F5 have resulted in eliciting antibodies of high affinity but with strongly limited neutralization capacity. Our results support the hypothesis that immunogens seeking to elicit 2F5-like broadly neutralizing antibodies will require components in addition to the core epitope. The challenge of eliciting by immunization antibodies that are simultaneously highly specific, broad in recognition, slightly polyreactive and able to access recessed epitopes is significant and exemplified by the obvious difficulty of generating such a response during natural HIV-1 infection. Structural studies looking at the interactions of broadly neutralizing antibodies with more complex epitopes in more complete and biologically relevant environments might be able to provide further clues of how to generate such a challenging immunogen.
Towards an HIV vaccine capable of eliciting a broadly neutralizing antibody response

The path towards the design of an HIV-1 vaccine inevitably has to include the elicitation of an immune response capable of controlling this virus. Since the discovery of HIV-1 as the pathogen responsible for AIDS at the beginning of the 1980’s, scientists have pursued multiple routes in attempts to design an effective HIV-1 vaccine. Traditional vaccine approaches such as whole-killed virus or live-attenuated virus have proven ineffective in safely controlling HIV-1 infections. Furthermore, past large scale trials have failed to show complete sterilizing protection or the ability to delay the progression to AIDS by using protein subunit components or DNA-based vaccines. Surely, much creativity and ingenuity will be required to conceive an HIV-1 vaccine in the near future, an objective which is at the centre of the fight against the worldwide AIDS pandemic.

Arguably the most desired effect of an effective HIV-1 vaccine would be the elicitation of an immune response capable of preventing new HIV-1 infections from viruses of all clades and subtypes. For most other pathogens for which an effective vaccine was effectively engineered in the past, such a sterilizing immunity was provided by the neutralizing antibodies generated by vaccination. Because Env is the only viral protein present on the surface of HIV-1, it is obvious why efforts to design an HIV-1 vaccine looking to elicit a sterilizing broadly neutralizing antibody response will need to focus on this protein as the key immunogen. One area of research focus to design a vaccine immunogen consists of characterizing broadly neutralizing antibodies that have been isolated from rare HIV-1 elite neutralizer patients. Indeed, understanding the details of the interaction between these bnmAbs and their epitopes will allow employing a guided immunogen design strategy that attempts to present vulnerable and conserved sites of HIV-1 Env to the immune system. For this purpose, a study of the paratope-epitope interaction at the atomic level is required for these few anti-HIV-1 bnmAbs. This is why the current work characterizes the details of binding of bnmAb 2F5 using X-ray crystallography as the principal technique.
5.1 Requirements for 2F5 recognition and neutralization

Screening of antibodies present in the plasma of asymptomatic HIV-1 infected individuals allowed for the report of the isolation of bnmAb 2F5 in 1993. Early studies showed that this antibody was able to potently neutralize all tested lab-adapted HIV-1 strains, as well as 70% of primary isolates (26, 220). Subsequent peptide binding and mutational studies on this rare antibody revealed that it recognizes the linear sequence ELDKWAS as its primary epitope, which is located in the MPER of the gp41 subunit of HIV-1 Env. To gain an atomic level insight of the 2F5-gp41 interaction, X-ray crystallography studies of 2F5 complexes with the ELDKWAS peptide were performed and revealed that this peptide adopts a predominant type-1 \( \beta \)-turn fold upon binding to 2F5 (204, 208). With this structural information in hand, multiple immunogen design strategies looking to elicit 2F5-like antibodies by vaccination were attempted without success. Such immunogens included the ELDKWAS peptide alone, a type-1 \( \beta \)-turn constrained ELDKWAS peptide with disulfide or lactam linkages in surrounding residues, and graft of the ELDKWAS linear motif in scaffold proteins (18, 117). In these studies, it was found that the immunogens were antigenic and induced a strong antibody response upon vaccination, but that none of the antibodies were capable of neutralizing HIV-1 even though they showed high binding affinity to the 2F5 epitope. These results indicated that either 1) the immunogens failed to reproduce the exact vulnerable conformation of the gp41 MPER recognized by 2F5 and that the ELDKWAS immunogen design should be improved to elicit 2F5-like neutralizing antibodies or that 2) additional components to the 2F5-ELDKWAS high affinity interaction were required for 2F5 neutralization and that the failure of the immunogens to present such elements resulted in the failure to generate 2F5-like neutralizing antibodies. The work reported in this thesis attempted to investigate these two possibilities from a structural perspective.

5.1.1 The rational design of a small immunogen capable of eliciting 2F5-like antibodies

First, as reported in Chapter 2, it was attempted to get an atomic level understanding of the 2F5 breadth of neutralization both from a linear sequence perspective, as well as three-dimensional recognition of the core gp41 epitope by 2F5. An analysis of the variability of residues in the gp41 MPER reports on the invariant nature of 15 out of 25 amino acids comprising this region. Although gp41 residues D664 and W666 recognized by 2F5 are invariant across all HIV-1
sequences, residues K665 and A667 show significant variability and might therefore help to explain the inability of 2F5 to neutralize some 30% of primary isolates.

Subsequently, the ability of the bnmAb 2F5 to recognize 31 varying sequences of the gp41 MPER was evaluated at a molecular level. In 19 cases, resulting crystal structures show the various MPER peptides bound to the 2F5 Fab'. A variety of amino acid substitutions outside the $^{664}$DKW$^{666}$ core epitope are tolerated. However, changes at the $^{664}$DKW$^{666}$ motif itself are restricted to those residues that preserve the aspartate’s negative charge, the hydrophobic alkyl-$\pi$ stacking arrangement between the $\beta$-turn lysine and tryptophan, and the positive charge of the former. In addition, the suspected possible molecular mechanism of 2F5 escape by sequence variability at position 665 and 667 was characterized. Variability at these positions is most often observed in HIV-1 clade C isolates, which 2F5 is almost completely incapable of neutralizing. Sequence variability to a non-positively charged amino acid at position 665 led to the impossibility of obtaining a co-crystal complex with the 2F5 Fab’, while sequence variability at position 667 often resulted in a similar behavior. The structural determination presented in Chapter 2 suggest that a substitution at position 667 from an alanine residue to a residue with a bulkier side chain, such as lysine or glutamine, might prevent the formation of the exact type-1 $\beta$-turn conformation recognized by 2F5. Based on these results, a somewhat more flexible molecular model of epitope recognition by bnmAb 2F5 is proposed. This information will surely prove helpful in guiding future attempts to elicit 2F5-like bnmAbs by designing MPER-like immunogens either by small-molecule chemistry or protein scaffolding.

5.1.2 Understanding the 2F5 epitope type-I $\beta$-turn in a biological context

In an attempt to structurally understand the 2F5 neutralization mechanism outside of its ELDKWA core, Chapter 3 presents eight 2F5 Fab’ crystal structures in complex with various gp41 peptide epitopes. These structures reveal several key features of this antibody-antigen interaction. 1) Whenever free of contacts caused by crystal packing artefacts, the extended CDR H3 loop is mobile; this is true for ligand-free as well as epitope-bound forms. 2) Residues located at the C-terminus of the gp41 ELDKWA core do not seem to interact as tightly with the antibody. However, in the presence of a larger peptide containing the gp41 fusion-peptide segment, these residues adopt a conformation consistent with the start of an $\alpha$-helix. This $\alpha$-helical conformation for residues located at the C-terminus of the ELDKWA core is in contrast
to the β-turn conformation previously reported by Ofek et al (204). Structures reported in Chapter 3 reveal that this latter conformation was found to be a crystal packing artefact, which resulted from the amidation of the peptide C-terminus causing a favourable interaction between the C-terminal part of the peptide and a symmetry-related molecule in the crystal lattice. 3) At high sulfate concentrations, the electron density maps of 2F5 Fab'-peptide complexes contain a peak that may mark a binding site for phosphate groups of negatively charged lipid headgroups. These additional structural details are important and should be considered when designing potential vaccine candidates intended to elicit 2F5-like antibody production.

The structural increments of the flexibility of the 2F5 CDR H3, the presence of a possible anion binding-site at the base of the 2F5 CDR H3 loop and the start of an α-helix for residues C-terminal of the ELDKWA core allowed us to model the interaction of 2F5 with its gp41 epitope on the surface of HIV-1 virions in more detail. Indeed, this additional structural information regarding the gp41 MPER when bound to 2F5 fits nicely with biological data, as well as past structural characterizations of the gp41 MPER. Solution nuclear magnetic resonance (NMR) experiments investigating the structure of a 19-residue MPER peptide (residues 665 to 683) found it to be helical in dodecylphosphocholine micelles. Subsequently, a combination of NMR and spin-label electron paramagnetic resonance (EPR) studies of the gp41 MPER in a lipid environment revealed a partially inserted peptide with an amphipathic conformation consisting of a tilted N-terminal α-helix (residues 664-672) followed by a short hinge and a flat C-terminal helical segment (675-683) (264). Another NMR characterization of this region in water reported a 310-helix conformation for the gp41 MPER residues, whereas its incorporation in an elongated 36-residue peptide including part of the upstream gp41 C-heptad repeat showed the propensity of the 2F5 epitope to adopt an α-helical conformation (27, 28). All these conformations assumed by the gp41 MPER in different contexts exemplify the high degree of flexibility for this region, which is thought to play a crucial role during the membrane fusion cascade. In fact, in experiments in which the gp41 MPER region was deleted, a complete loss of Env-mediated cell fusion was observed, without however affecting the membrane-perturbation abilities of gp41 or the formation of its six-helix-bundle conformation (74). This evidence strongly suggests that the gp41 MPER is probably responsible for the bending of gp41, which is a movement thought to be required for the coming together of viral and target membranes. For better understanding this hypothetical bending event, the 2F5-HybK3 structure was particularly informative in that it
demonstrated that the helix' main axis forms an almost right angle with the amino-terminal extended chain. Thus, the type I $\beta$-turn element recognized by 2F5 further indicates a function of this region as a molecular hinge that could induce a $90^\circ$ directional change in the MPER chain path during the fusion process, and hence a major reorientation of other gp41 domains (Fig 5.1). The timeline of the different conformations adopted by the gp41 MPER ($\alpha$-helix, $3_{10}$-helix or type I $\beta$-turn) during the fusion process remains a matter of debate and requires further investigation.

Figure 5.1 Hypothetical model to exemplify the conformational changes of the MPER region during the fusion process. The turn-hinge structure (right) recognized by the antibody would disrupt the membrane-inserted amphipathic-at-interface $\alpha$-helix (red cylinder) and induce a change in gp41 chain path orientation. Energy coming from conformational changes in other parts of the system might induce the refolding of the turn and the preceding extended stretch into a more rigid $3_{10}$-helix, which might associate with the membrane interface (left). Hydrophobic-at-interface residues are designated by the green side-chains. The blue side-chain designates Lys665. (Adapted with permission from de la Arada et al., 2009 (63)).
5.1.3 An immunogen looking to elicit 2F5-like antibodies will need to incorporate other components in addition to its core epitope

The structures presented in the first chapters revealed that the \(^{100}\text{TLFGVP}^{100}\text{F}\) apex residues of the 2F5 CDR H3 loop do not directly interact with residues of its core gp41 epitope. However, residues at the apex of 2F5’s unusually long CDR H3 loop have been shown to be crucial for neutralization (305). In an attempt to gain better insight into the functional role of this element, a 2F5 Fab expression system was constructed in \textit{E. coli} so that native 2F5 Fab and two of its mutants, in which either the apical Phe100B(H) residue was changed to an alanine or the CDR H3 residues \(^{100}\text{TLFGVP}^{100}\text{F}\) were replaced by a Ser-Gly dipeptide linker could be recombinantly expressed (135). Isothermal titration calorimetry (ITC) and competitive ELISA binding assays rendered strikingly similar affinity constants (\(K_d \sim 20\text{ nM}\)) for linear peptide-epitope binding by 2F5 Fabs, independent of the presence or absence of the apex residues. Ablation of the CDR H3 apex residues, however, abolished the cell-cell fusion inhibition and pseudovirus neutralization capacities of 2F5 Fab. To further investigate the binding determinants that conferred neutralizing capacity on 2F5 through its CDR H3 extended loop, competitive ELISA were performed and the results suggest a role for 2F5 CDR H3 apex residues in mediating weak hydrophobic interactions with residues located at the C-terminus of the gp41 MPER and/or membrane components in the context of core epitope binding. The present data therefore imply an extended 2F5 paratope that includes weak secondary interactions that are crucial for neutralization of Env-mediated fusion. Such interactions, which remain to be better characterized and described structurally, will surely need to be eventually incorporated into an immunogen looking to elicit 2F5-like antibodies.

Furthermore, these secondary interactions of the elongated 2F5 CDR H3, in addition to those mediated by its primary paratope, might not be the only additional weak interactions that are required for the neutralization of HIV by bnAb 2F5. In fact, it is possible that 2F5 interacts weakly with other parts of gp41, such as regions located N-terminal of its core epitope, and that such interactions might be required for HIV neutralization. Supporting this idea is the comparison between the sequence of the variable region of 2F5 and its corresponding germline sequence, which reveals noticeable maturation for heavy chain residues Asn64, Thr65 and Arg82B of the framework 3 region, a site located immediately upstream of the 2F5 main paratope. Indeed, when compared to all human variable domain sequences in the Kabat database, the 2F5-precursor germline residues at these sites (Lys64, Ser65 and Asn82B) have a frequency
of 76%, 22% and 5%, respectively (37, 305). These residues mutate during the maturation process to much rarer residues, namely Asn64, Thr65 and Arg82B, which have frequencies of 0.5%, 1% and 3%, respectively, when compared to other human variable domain sequences at these positions (37). It has also been shown that the combination of \(^{64}\text{N-T-(X)}_{18}\text{R}^{82}\) seems to be unique for human antibodies (37). In fact, a previous study that attempted to generate an anti-idiotypic antibody to 2F5, an antibody that would mimic the gp41 epitope recognized by 2F5 by being the complement of the 2F5 paratope, yielded antibody 3H6, which recognizes these unique residues in this conserved helix-like motif (Fig 5.2). 3H6 was able to compete off the binding of 2F5 to its gp41 core epitope peptide by steric hindrance (37, 92, 150). Altogether, these observations make it tempting to speculate that 2F5 framework residues Asn64, Thr65 and Arg82B might play a so far unrecognized role in gp41 binding and therefore be part of an upstream extended 2F5 paratope possibly important for HIV-1 neutralization. It is obvious, however, that further studies will be required to determine the validity of such a hypothesis.

Expanding our structural and mechanistic knowledge of complexes between 2F5 and larger parts of gp41 will be necessary to better define the whole interaction interface between these two proteins. In order to perform such a structural characterization by X-ray crystallography, the relevant fusogenic state of gp41 as recognized by 2F5 (pre-fusion, fusion-intermediate or post-fusion) will first need to be determined. Recently, binding studies performed by surface plasmon resonance (SPR) of 2F5 with gp41 constructs representing attempts to mimic these different conformations suggest that 2F5 might have the highest affinity for a fusion-intermediate (\(K_d \sim 1.4\) nM), whereas it showed limited affinity for gp41 in its post-fusion six-helix bundle conformation (\(K_d \sim 1.4\) \(\mu\)M) and almost no detectable affinity for gp41 in its pre-fusion conformation (90). These 2F5 binding affinity properties help to explain why extensive efforts to co-crystallize 2F5 Fab’ with a bacterially expressed gp41 post-fusion six-helix bundle construct possessing the 2F5 epitope could not be achieved during this thesis (Appendix 1). These results also suggest that using a glycosylated fusion-intermediate construct of gp41 to obtain 2F5-gp41 co-crystals would surely have a higher chance of success. For this purpose, the only gp41 fusion-intermediate construct reported to date could be used, which consists of the following sequence: (HR2)-linker-[HR1-loop-HR2-MPER]-(trimerization tag), where HR1 and HR2 are the first and second heptad repeat domains of gp41, respectively (90). It is expected that in this construct, the N-terminal HR2 segments form a six-helix bundle with the HR1 segments upon trimerization,
whereas the C-terminal HR2 segments, which is constrained by the trimerization tag, remain extended and mimic the fusion-intermediate extended conformation. This construct has been shown to express well in both in *E. coli* or secreted by insect cells.

**Figure 5.2** Surface rendering of the Ab2/3H6 Fab–2F5 Fab′ complex with superimposed gp41 peptide. The crystal structure of 2F5 Fab′ variable domain with bound 13-mer peptide (ELLELDKWASLNW), representing gp41 residues 659–671, was superimposed onto the 2F5 Fab′ variable domain from the Ab2/3H6 Fab–2F5 Fab′ complex to show the proximity of Ab2/3H6 Fab to the 2F5 Fab′ paratope: Ab2/3H6 Fab (heavy chain, cyan; light chain, green), 2F5 Fab′ (heavy chain, purple; light chain, rose), gp41 peptide (yellow). The contact regions of Ab2/3H6 Fab and 2F5 Fab′ are coloured blue and red, respectively. Reproduced with permission from Bryson et al., 2008 (37).
5.2 The gp41 MPER as a target for HIV vaccine design

5.2.1 Roles of the gp41 MPER

One of the most crucial requirements that a successful HIV-1 vaccine immunogen will need to meet is to mimic a conserved region of Env that is essential for HIV-1 entry and hence cannot be readily mutated. As described in Chapter 2, the gp41 MPER is a particularly conserved region of HIV-1 Env, with 15 of 25 residues being completely invariable across all HIV-1 sequences reported to date (36). The gp41 MPER region spans residues 659-683 (184, 306). The conserved nature of this gp41 region is probably caused by the critical and diverse roles it plays during HIV-1 infection. The first involvement of the gp41 MPER during infection has been documented to be its binding to galactosyl ceramide receptors on epithelial cells that leads to mucosal infection mediated by transcytosis (9, 10, 297). Indeed, specific pre-incubation of the virus with IgA recognizing the ELDKWA epitope of the gp41 MPER resulted in the inability of HIV-1 to undergo transcytosis through epithelial cells (9, 29). This finding supports the potential of an MPER HIV-1 vaccine in that it could prevent initial infections by eliciting protective antibodies at mucosal sites. A detailed structural understanding of how the gp41 MPER interacts with the galactosyl ceramide receptor might give further indications as to which vulnerable conformations have to be included in an immunogen.

Additionally, the gp41 MPER has been implicated in gp41 oligomerization and membrane leakage through pore formation (187, 230, 261, 262). One of the characteristics of the gp41 MPER that could mediate such functions is its cholesterol-binding motif at residues 679-683 (276). Indeed, it has been shown that the HIV-1 membrane is highly ordered and is particularly rich in cholesterol (cholesterol:phospholipid ratio > 1) (222). It has been proposed that the MPER might be involved in the creation of membrane projections, where the bilayer is enriched in cholesterol/sphingomyelin (184). The MPER sequence could participate in the clustering of gp41 monomers (within the HIV-1 Env) through cholesterol interactions and hence destabilize the architecture of the bilayer at the fusion locus. This step has been hypothesized to be the main energetic barrier for initial bilayer merging, but a more in depth understanding of the fusion pore formation is required and could potentially lead to the elaboration of novel therapeutics or immunogen design efforts.
Finally, and as previously described, substitution and deletion studies have linked this unusually tryptophan-rich region to the fusion process of HIV-1 (23, 184, 190, 217, 232, 304). The role of the MPER in this process probably resides in a series of conformational changes and differential membrane partitioning that lead to required re-orientation of gp41 for host membrane interaction by the fusion peptide, and subsequent collapse into its six-helix bundle, which ultimately results in the fusion of host and viral membranes. Neutralizing antibodies or compounds (such as the commercially available drug Enfuvirtide, an entry inhibitor) capable of inhibiting the MPER-mediated fusion cascade are of particular interest.

Altogether, this wide array of roles for the gp41 MPER will put considerable pressure on sequence conservation and any change will certainly lead to a high cost in viral fitness. It is for this reason that the gp41 MPER represents a major object of interest in vaccine design strategies looking to elicit a broadly neutralizing antibody response.

5.2.2 Immunogenicity of the MPER during natural infection

Contrary to the gp41 N-terminal heptad repeat (residues 546-581) and to the hydrophilic loop that connects the two gp41 heptad repeats (residues 598-604), the gp41 MPER is not strongly immunogenic during the course of natural infection (184). In a recent study looking at the antibody reactivity of the serum of HIV-1 infected subjects to the ELDKWA peptide, recognition was present at a low frequency (15 to 35%) (257). Moreover, although the neutralizing power of these antibodies remains poorly characterized, a recent study showed that MPER-specific neutralizing antibodies are rarely found in HIV-1 infected sera (62). Whether this finding is due to the inability of the neutralization assays to detect low antibody titers, or to a real absence of neutralizing antibodies against the gp41 MPER in these studies remains to be determined. In fact, an important question that stems from these observations and that needs to be addressed is how much of an anti-MPER neutralizing antibody titer is required to confer protection.

Although rarely detected, anti-MPER neutralizing antibodies confer significant benefits in the few cases that they are observed. In a study comparing the sera of AIDS patients and asymptomatic individuals, it was observed that antibody titers to the ELDKWA epitope were significantly higher in the case of asymptomatic cases (257). Other studies have also reported that the antibody levels against the ELDKWA peptide over time were inversely associated with disease progression, and that almost all long-term non-progressors develop MPER binding
antibodies (30, 45, 98, 275). Passive immunization studies with the very few gp41 MPER-directed bnmAbs identified to date have showed that the presence of such antibodies prior to viral challenge can prevent its acquisition (115). In addition, passive transfer of a bnmAb cocktail, including anti-MPER antibodies, to patients undergoing cessation of antiretroviral therapy showed a significant delay in HIV-1 rebound in individuals with acute infection (270). Together, these studies highlight the importance of enhancing the immunogenicity of the gp41 MPER and give hope that generating sufficient anti-MPER antibody titers by vaccination could lead to protection against this virus.

5.2.3 bnmAbs targeting the gp41 MPER

The rare bnmAbs targeting the gp41 MPER give a unique opportunity to characterize this region in a vulnerable state that could be used as a vaccine immunogen. In addition to the bnmAb 2F5 characterized extensively in the present work, bnmAb Z13 and bnmAb 4E10 recognize overlapping linear segments of the MPER, namely residues 670WNWFDITNW678. This epitope is located immediately C-terminal to the linear epitope recognized by bnmAb 2F5, which consist of residues 662ELDKWAS668. The proximity of the epitopes of 2F5, Z13 and 4E10 results in the direct competition for binding to a synthetic peptide derived from gp41 (306). Interestingly, all three anti-MPER bnmAbs were initially isolated as IgG3, a selection that might be explained by the longer nature of this IgG subtype, which would allow the B-cell receptor to sufficiently extend to interact with an occluded epitope.

4E10 was initially isolated from an HIV-1 seropositive donor as a hybridoma and although the 4E10 epitope was originally mapped to residues 823AEGTDRV829, subsequent studies showed that the 4E10 epitope comprises in fact the linear sequence 670WNWFDITNW678 (38, 306). The crystal structure of the 4E10 Fab in complex with the MPER peptide 670WNWFDITNW678 revealed that residues of the epitope are recognized in a predominantly \(\alpha\)-helical conformation, where the 4E10 paratope interacts principally with core residues 672WFDIT676, with W672 being the one most buried in the paratope surface (47). Interestingly, the \(\alpha\)-helical conformation of the 4E10 epitope allows this bnmAb to interact with the face of the helix that has the most sequence conservation (no observed variation at positions W672, F673 and I675 across all HIV-1 sequences), which makes this bnmAb the one with the broadest activity of the extensively characterized neutralizing antibodies identified to date (26). On the opposite face of the helix are residues D674,
T/S$_{676}$ and N/K$_{677}$; these residues are much more variable and 4E10 only makes limited interactions with them. Mutagenesis studies revealed that the residues required for 4E10 neutralization are W$_{672}$, F$_{673}$ and W$_{680}$ located 7 amino acids away from the core epitope, which suggests that similarly to the 2F5 epitope, 4E10 could also interact with residues outside its core epitope (304). In fact, very similarly to the case of 2F5, the crystal structure of 4E10 with its MPER epitope revealed that the main interactions of the 4E10 paratope are mediated by CDRs other than CDR H3. The apex of the 18 amino acid long CDR H3 loop bends away from the peptide (47). More recently, mutagenesis studies of the 4E10 CDR H3 confirmed the necessity of this extended loop for mediating HIV-1 neutralization (8, 242, 293). Although the 4E10 and 2F5 linear epitopes are close in proximity and both bnmAbs have evolved elongated CDR H3 loops, which seem to be required for mediating secondary interactions to core epitope binding, these two bnmAbs might not neutralize HIV-1 in exactly the same way and hence should be further characterized separately. Indeed, 4E10 has much higher affinity towards membrane components than 2F5; it is also much better in extracting its epitope from a phospholipid bilayer (71, 109, 176, 256). These results indicate that much remains to be determined in understanding the exact neutralizing mechanism of anti-MPER bnmAbs, with the objective of mimicking their exact binding requirements in an effective immunogen.

More recently, using a different technology (bone marrow RNA phage display library), bnmAb Z13 was isolated from an HIV-1-infected individual with a neutralizing sera (306). The Fab library from which Z13 was isolated was initially screened against both a synthetic peptide composed of gp41 MPER residues 660-680, as well as the entire HIV-1$_{MN}$ sequence. The binding of Z13 to both sequences suggest that the epitope recognized by Z13 is exposed on the viral surface structure (184). Subsequently, random mutations of the CDR-L3 sequence led to the isolation of Z13e1, an Fab with both increased affinity for the MPER and increased neutralization potency (194). Scanning mutagenesis, peptide mapping, antibody competition studies, as well as the recently solved crystal structure of Z13e1 Fab with its core epitope peptide (W$_{670}$NWFDITN$^677$KKKK), all showed that Z13e1 interacts predominantly with residues Asn$_{671}$, Asp$_{674}$, and Asn$_{677}$, which are located on the hydrophilic face opposite to that recognized by 4E10 (194, 212, 306). Similarly to 2F5 and 4E10, Ze13e1 has an elongated CDR H3 loop (17 amino acids). It makes interactions with the core epitope only at its base; its apex residues are disordered in the crystal structure (212). The less conserved nature of the MPER residues
recognized by Ze13e1 (71% and 58% conservation for residues Asn<sub>671</sub> and Asp<sub>674</sub>, respectively) helps to explain why Z13e1 neutralizes HIV-1 with much lesser breadth than bnmAbs 2F5 and 4E10 (306). Nonetheless, the discovery of Ze13 has given a new target on which vaccine design can be based. Furthermore, it has strengthened the idea that anti-MPER antibodies are important in mediating HIV-1 neutralization in vivo and hence should be a major research focus for developing an effective immunogen.

### 5.2.4 Challenges associated with eliciting anti-MPER neutralizing antibodies

A substantial review by Montero et al. in 2008 summarized that a total of ten vaccines had been previously described that targeted the production of 2F5-like antibodies, as well as an additional six that attempted the development of a vaccine capable of producing antibodies against gp41 or the MPER, more specifically (184). However, all of these initial attempts to elicit 2F5-like and 4E10-like broadly neutralizing antibodies have failed. These results indicate that many challenges remain in the quest to produce such a vaccine.Outlined here are three main challenges associated with the development of a vaccine targeting the gp41 MPER: 1) a rather high flexibility for the MPER, 2) the proximity of the MPER to the membrane, and 3) the low immunogenicity of the MPER. Here, current and past attempts to address these challenges are briefly discussed.

The gp41 MPER is thought to play a major role in Env-mediated membrane fusion through its flexibility and the multiple conformations that it adopts. For vaccine design strategy, this causes a significant problem in choosing the appropriate MPER conformation to mimic. Past studies have reported on attempts to constrain the conformations recognized by bnmAb 2F5 and 4E10, either with β-turn-constrained peptides conjugated to outer membrane proteins or extended MPER epitope peptides constrained to stabilize a helical conformation (33, 178, 179). Although these immunogens were capable of eliciting antipeptide antibodies, these were not found to be neutralizing. Similar results were obtained when attempting to incorporate the 2F5 epitope in a carrier such as bovine serum albumin or tetanus toxoid, when attaching multivalent peptides on a carbohydrate scaffold or to a CD4 helper T-cell epitope (67, 68, 164, 196). Altogether, the results that antipeptide serum could be recovered without being neutralizing suggests that the immunogen most likely failed to present parts of the native antigen that is required to evolve a
neutralizing antibody, rather than not presenting the desired MPER conformation. Extensive studies, including the one presented in Chapter 2, have determined the exact high-affinity interactions necessary for inducing core epitope antibodies by immunization, and these will surely help in the design of an effective MPER immunogen. However, as documented in Chapter 4, it is now becoming evident that an effective immunogen looking to elicit 2F5-, 4E10- and Ze13e1-like antibodies will have to present, in addition to its high-affinity constrained MPER epitope, secondary components that are required to mediate neutralization. Indeed, it is extremely unlikely that the immunogens described above could have induced antibodies with a required extended CDR H3 loop. The questions that arise from these past studies are therefore not as much how to better constrain the MPER immunogen in a specific conformation, but rather how to better present the constrained gp41 MPER in a more biologically relevant environment that will lead to the elicitation of neutralizing antibodies.

Without doubt, the proximity of the gp41 MPER to the viral membrane is the biological context that is at the moment receiving the most attention when it comes to eliciting anti-MPER neutralizing antibodies. Indeed, many studies have made clear that the membrane plays an important role in shaping the structure of the MPER (230, 276, 295). Furthermore, the still unknown reason for the prevalence of the elongated hydrophobic CDR H3 loops of bnmAbs 2F5, Ze13e1 and 4E10 has tempted many to infer a role for this element in mediating interactions with the viral membrane, although the underlying evidence is equivocal (see Section 4.4 for a more complete discussion). Many immunogen design efforts have attempted to present the MPER in the context of a membrane with the objective to create an environment for the 2F5 and 4E10 epitopes reminiscent of that experienced by the viral Env. Some examples of such strategies include displaying the HIV-1 MPER on recombinant enveloped viruses; presenting the MPER fused to a TM domain, such as the influenza hemagglutinin (HA); and insertion of the MPER in liposomes (160, 171, 177, 296). Taken together, the results of these experiments showed that the presentation of the MPER in the context of the membrane could produce neutralizing antibodies, although at low titers. Furthermore, it seems from these studies that other factors such as the composition of the membranes and the types of adjuvant added to the immunogen for vaccination might have played a role in influencing the antibody response. Furthermore, it remains to be determined whether the low-titer neutralizing antibody response generated by inserting the MPER in a membrane environment resulted in a broadly neutralizing
antibody response. Although encouraging, these results showed only a small production of anti-MPER antibodies and further demonstrated that the MPER is only weakly immunogenic. Another significant challenge is therefore to identify methods of producing high titers of antibodies against this region.

One of the first strategies to generate a strong immunogenic response against the MPER was to introduce its sequence into a loop of another protein previously identified to be strongly immunogenic. Examples include the antigenic site B of HA, the V1/V2 loop of gp120 and three modified constructs of gp41 presented in the context of a virus-like-particle (VLP) (141, 157, 191). The results of these studies demonstrated that as long as the MPER was presented in the context of other immunodominant regions, its position in an immunodominant loop did not appear to enhance its targeted antibody response. It was proposed that if the other protein domains possessing immunodominant epitopes could be modified to decrease their immunogenicity, it could focus the antibody response on the MPER. For this purpose, strategies were designed based on successes in other research areas, e.g. that immunodominant regions be completely removed, hidden by glycosylation or that charged or bulky residues in these regions be substituted with less immunogenic ones (such as Gly, Ala, or Ser) (95, 129, 142, 169, 210, 272). Other strategies looking to improve the immune response to the targeted MPER region attempted to include other arms of the immune system. For example, Ho et al. inserted the ELDKWas MPER peptide in the framework region of an anti-HLA-DR-MAb, with the objective of aiding antigen presentation when the MAb would bind to the major histocompatibility complex (MHC) class II (60, 117). Another experiment fused a gp41 segment including the MPER to the human B-cell-activation factor (BAFF) with the objective for this region to act as a molecular adjuvant that would induce the development and maturation of follicular B cells recognizing this antigen (141). These creative experiments resulted in disappointing results; high titers of antibodies were generated by the immunogen-adjuvant but low anti-MPER titers could be detected.

Vaccinologists working on the HIV-1 MPER seem to agree that the generation of a successful vaccine targeted against this region will require a combination of approaches and components. Next generation immunogens will surely need to include a constrained MPER core, in addition to secondary elements, which will possibly lead to its required incorporation in a membrane environment. In addition, it is possible that an effective MPER immunogen will need to be
altered to remove immunodominant regions to focus the immune response on the weakly immunogenic MPER. Finally, the generation of an anti-MPER vaccine might require the use of novel strategies, including the discovery of optimal adjuvants. Such research will possibly be inspired by successes in other vaccine research areas, such as various prime-boost strategies that have resulted in enhanced immunogenicity (21, 72, 107).

5.3 Other anti-HIV-1 bnmAbs and Env vulnerability

Conserved regions on gp41 and gp120 Env as targets for vaccine design have been mostly characterized by the very few anti-HIV-1 bnmAbs that recognize them. In addition to the membrane proximal external region (MPER) of gp41 recognized by bnmAbs 2F5, Z13 and 4E10, other regions of HIV-1 Env have received significant attention for targeted vaccine design. These include the CD4 binding-site on gp120 (bnmAb b12), a CD4-induced gp120 co-receptor binding site (bnmAbs 17b and X5), a mannose cluster on the outer face of gp120 (bnmAb 2G12) and the variable loop 3 (V3) region on gp120 (bnmAbs 447-52D and F425-B4e8) (22, 40, 47, 102, 154, 238, 301). What becomes evident from looking at the properties of the isolated neutralizing monoclonal antibodies against these sites is that they have evolved special modes of humoral recognition suited to overcome viral evasion tactics. Understanding the unique characteristics and origins of these bnmAbs and the Env regions they recognize will surely help in guiding the design of immunogens capable of eliciting neutralizing antibodies against these vulnerable sites.

As previously described, HIV-1 initiates the fusion cascade by engaging the CD4 receptor on the host cell through its gp120 Env subunit. This suggests that the initial site of interaction for the CD4 receptor on gp120 should be exposed and available. However, structural studies of the gp120 core have revealed that the overall size of the gp120 CD4 site of contact is small and that this surface is potentially recessed on the Env viral spike (301). bnmAb b12, an antibody that has its epitope mapped to this CD4-binding site on gp120 was initially identified from the bone marrow of a long-term non-progressor HIV-1 positive individual by phage-display technology using a baculovirus-derived gp120 construct (227). Structural analysis revealed that b12 mediates its interaction with the gp120 CD4 binding site through its heavy chain complementary determining regions alone (301). It has been hypothesized that the small and recessed nature of the epitope might cause steric barriers for Fab recognition mediated by all six CDRs at the
juncture between heavy and light chains of the antibody (151). In the case of bnmAb b12, interactions with the CD4 binding site on gp120 is mostly restricted to a rigid elongated CDR H3 that protrudes 15 Å from the antibody paratope (235). It could possibly present a significant challenge for the immune system to elicit b12-like antibodies characterized by a “heavy-chain-only” recognition mechanism. This difficulty resides in the antibody variability being confined to only three CDRs (from the heavy chain), instead of the usual six (heavy chain + light chain) (151).

With approximately 25 N-linked glycans covering 500 residues, the HIV-1 gp120 is one of the most heavily glycosylated viral proteins identified to date. The gp120 high-mannose glycans are relatively conserved, and they are used to facilitate entry into CD4+ T cells via the interaction with other receptors, such as DC-SIGN (99). As these glycans are derived from the host’s endoplasmic reticulum and Golgi pathways during the HIV-1 life cycle, they appear as “self” to the immune system. They also sterically shield core protein epitopes from antibody recognition. Monoclonal antibody 2G12, initially isolated from the sera of an HIV-1 seropositive individual, is the only antibody with broad neutralization characteristics identified to date that has evolved the ability to see these gp120 glycans as “foreign” probably because of their unusual clustering on the surface of the glycoprotein, a motif not readily observed in host proteins (38). To achieve high-affinity binding, bnmAb 2G12 adopts a unique architecture in which the variable heavy chains are swapped on adjacent Fab arms. This feature expands the antibody paratope to 20 Å² x 60 Å², allowing it to interact with up to three N-linked glycans, with a binding affinity in the nanomolar range (44). The evolved domain-swapped structural design of bnmAb 2G12 is a unique feature that makes this antibody capable of neutralizing HIV-1 through high-affinity recognition of carbohydrate motifs.

The most highly conserved region on gp120 is the co-receptor binding site, which only forms after conformational changes induced by the encounter of the virus with the CD4 receptor (288). This conformational change requirement for surface exposure of a conserved region is a perfect example of yet another means employed by HIV-1 for immune evasion. Molecular modeling suggests that the access of antibodies to this surface is further complicated by steric hindrance (154). Nonetheless, bnmAbs whose epitope is the CD4-induced co-receptor binding site, such as 17b and X5, have been derived from the sera of various HIV-1 infected individuals (188, 266). Interestingly, these antibodies mostly arise from a single heavy-chain genomic precursor, V\textsubscript{H}1-69
Structural studies have revealed that antibody recognition of the CD4-induced coreceptor epitope requires a hydrophobic CDR H2 and an acidic CDR H3, two properties that are only found together in V$_{H}$1-69 (121). Elicitation of antibodies capable of recognizing the conserved HIV-1 CD4-induced coreceptor surface is therefore facilitated by the encoding of required recognition elements in germline genomic precursors. This feature greatly reduces the requirement for the generation of diversity in recognition.

Another region of interest is the highly immunogenic variable loop 3 (V3) on gp120, which is involved in determining co-receptor usage of the virus and is recognized by a considerable number of antibodies capable of some level of cross-clade neutralization, including bnmAbs 447-52D and F425-B4e8 (22, 40, 103). These bnmAbs recognize the V3 loop in a relatively sequence-independent manner, and most interactions with the hypervariable V3 loop are with the polypeptide backbone of the functionally conserved epitopes. Indeed, bnmAb 447-52D is adapted for recognition of the Gly-Pro-X-Arg motif in a β-sheet conformation, whereas bnmAb F425-B4e8 has evolved to specifically recognize only two residues (Ile and Arg) in an unusual five residue α-turn conformation (151). Such a sequence-independent recognition is necessary for neutralizing HIV-1 through interactions with an immunodominant region which is hypervariable with conservation of only a few residues from strain to strain.

Research analyzing the interactions of the immune system with the HIV-1 Env protein has succeeded in identifying novel modes of antibody recognition that had not yet been described before. A wide array of studies characterizing predominantly mouse antibodies led to the hypothesis that it is the persistent exposure of the immune system to an often high antigen load during HIV-1 infection that drives immune adaptation to create these unusual mechanisms of recognition. With this information in hand, the challenge then becomes to design strategies capable of re-eliciting such humoral immune responses by vaccination. However, in the hopes of providing sterilizing immunity, it might be important to focus future studies on HIV-1 viruses capable of mediating initial infection, as opposed to viruses that have resided in HIV-1 positive individuals for a long time, as these might have evolved to gain characteristics to evade immune recognition as opposed to promote initial infection. Indeed, recent studies assessing HIV-1 diversity in humans showed that infection can often be traced back to the transmission of a single virion (139).
5.4 Difficulties in eliciting bnmAbs against HIV-1

It is becoming increasingly clear that simple vaccination with Env protein subunits, such as the native Env spike, gp120 monomer or gp41 MPER, and including peptides have a very limited chance of eliciting an effective broadly neutralizing antibody response that would confer protection against HIV-1 infection. As described in Section 5.2.4 for the gp41 MPER, creative strategies will need to be employed to present the vulnerable sites of the HIV-1 Env molecule to the immune system in a way in which humans can mount a protective immunity against this deadly pathogen.

A significant challenge associated with this task, which was best exemplified by the different sites recognized by the few bnmAbs described above will be to overcome immune evasion by Env conformational masking. Comparing the time course of Env-mediated fusion with that of the low pH fusion induced by influenza HA suggests that HIV-1 entry is somewhat stochastic and lacks synchronicity (93). Incubation of HIV and SIV with cultured cells reached maximum fusion with a $t_{1/2}$ of 19 minutes; the rate limited step of Env-mediated fusion is related to the binding of gp120 to the CD4 receptor, an event initially observed in co-culture after a time lag of 10-15 minutes. The subsequent co-receptor engagement and the formation of the six-helix bundles have similar kinetics and are in the order of 10 minutes (189, 223, 281). Small molecule inhibitors capable of blocking Env-mediated fusion benefit from the relatively slow fusion process by binding fusion intermediates. However, the rarity of bnmAbs in HIV-1 infected individuals suggests that conserved Env regions presented to the immune system as a result of conformational changes are probably only exposed briefly during the lengthy course of the fusion process. In order to elicit broadly neutralizing antibodies against the gp120 CD4 binding site, the gp120 CD4-induced co-receptor binding site and the dynamic gp41 MPER by vaccination, it will be necessary to develop strategies to lock the Env spike in entry-relevant conformations for the effective presentation of conserved regions to the immune system.

In addition to the challenges associated with the design of a construct mimicking an entry-relevant conformation of the Env spike, a major objective will be to enhance the immunogenicity of the targeted conserved Env region. Indeed, as described for the case of the gp41 MPER, highly conserved regions of HIV-1 Env often have the unfortunate characteristic of being poorly immunogenic. To focus the immune response to the poorly immunogenic conserved regions, it
will be important to use construct design strategies such as the alteration of highly immunogenic variable Env domains, as well as the masking of these regions by PEGylation or glycosylation. Such attempts have been previously reported for a monomeric gp120 construct, where the immunodominant surfaces were masked by glycosylation to preferentially expose the b12 epitope (210). Although this construct was not able to generate broadly neutralizing antibodies in immunization experiments, it was found that glycan masking was successful in altering the specificity of elicited antibodies. Furthermore, it is highly probable that an effectively shielded construct will require the use of an adjuvant to boost the immune response targeted to the poorly immunogenic conserved Env region. Few adjuvant molecules have been extensively characterized to date but the development of such molecules capable of boosting the immune response in specific ways constitutes a very active field of research (137). In humans, the most commonly used adjuvant is aluminum salt particles on which antigen can be adsorbed (Alum) (140). Its use in vaccine formulation has been shown to result in the enhancement of antibody responses to the antigen by predominantly eliciting a T helper 2-biased immune response producing humoral immunity. The careful use of this adjuvant, or others, in vaccine regimens that present focused cross-reactive determinants has the potential to drive the production of effective antibody responses against HIV-1.

Over the years, starting with sequence and then structural characterization of bnmAbs b12, 2G12, 4E10 and 2F5, it quickly became evident that a property that all these rare antibodies had in common is an unusually elongated CDR H3 hypervariable loop. Recent studies have demonstrated that these specific CDR H3 loops are required to confer neutralizing capacities on these antibodies (8, 52, 203, 242, 293). The origin of these long CDR H3 loops has been long debated, and it is unclear whether they originate from unusual recombination events of variable (V), diversity (D) and joining (J) segments during the initial maturation of immunoglobulins or through subsequent extensive somatic hypermutation during chronic HIV-1 infection (148, 289). The analysis of several hundred human monoclonal antibodies has revealed that antibodies with an elongated CDR H3 are in general associated with chronic viral infections (184). Evidence of the anti-HIV-1 bnmAbs identified to date suggests that it is highly probable that immunogens seeking to elicit an effective humoral response will need to generate antibodies with such elongated CDR H3 loops. To be able to do so might require novel vaccination strategies that mimic the course of chronic viral infections, such as the administration of the vaccine regimen
on multiple occasions or a vaccine delivered by time-release preparations (184). Furthermore, the possible requirement to elicit antibodies possessing elongated CDR H3 loops will require the careful design of pre-clinical animal model immunization experiments. Indeed, it has been shown that some animals, such as mice, are unable to produce long CDR H3 regions, whereas other animals like rabbits are capable of doing so (157). It would be extremely unfortunate if the design of an immunogen capable of eliciting broadly neutralizing antibodies in humans did not progress to clinical trials because of its inability to produce such a humoral response in animal models.

Finally, another important aspect that needs to be considered for HIV-1 vaccine design has to do with the effector functions of elicited antibodies. Indeed, human antibodies can occur as either unswitched antibodies, IgM and IgD, or class-switched antibody isotypes: IgG (of subclass 1, 2, 3 or 4), IgA (of subclass 1 or 2) and IgE (221). The different antibody isotypes determine their localization and functions in humoral immunity. Whereas the Fab portion of the antibody determines the antigen-binding specificity, the Fc portion mediates interactions with complement components and Fc receptors, which in turn determine effector function capabilities of the antibody such as complement activation, inactivation through opsonisation and antibody-dependant cellular cytotoxicity (ADCC). Numerous in vitro studies have shown that the inhibitory activity of anti-HIV-1 Env neutralizing antibodies can be altered by their effector functions (5, 118, 123, 170, 263). Indeed, it is possible that in addition to direct neutralization of the virus by binding to the functional Env spike through its Fab fragment, part of the bnmAb activity comes from mechanisms mediated by its Fc region (213). As a noticeable example, it was reported that mutations eliminating the Fc receptor-binding site of bnmAb b12 significantly reduced the ability of this antibody to protect macaques by prophylaxis (112). In another study, analysis of passive immunization trials with bnmAbs 2G12, 2F5 and 4E10 revealed that complement activation by these monoclonal antibodies did not significantly contribute to treatment success, suggesting that the in vivo activity of these antibodies was due to either direct neutralization or Fc receptor-mediated mechanisms (124). Other investigations reported that a 2F5 IgG3 to IgG1 subclass switch did not lead to changes in affinity or specificity but that 2F5 IgG1 had slightly higher Fcγ receptor effects, mostly mediated by Fcγ receptors I and IIb, which are present on the surface of macrophages, dendritic cells and polymorphonuclear leukocytes (such as basophils and neutrophils) (149, 213). Furthermore, class switching experiments of 2F5
and 2G12 IgGs to polymeric IgM and IgA led to the discovery that the newly produced antibodies could effectively interfere with HIV-1 entry across a mucosal epithelial layer \textit{in vitro}, whereas their IgG counterpart could not (286). Altogether, these results emphasize the need to better understand the antibody-mediated effector functions for the few bmmAbs identified to date to fully uncover their mechanism of neutralization (269). Advancements in this field could lead to the introduction of new vaccination strategies targeting the elicitation of specific antibody effector functions capable of mediating HIV-1 neutralization.

5.5 Future Directions

5.5.1 Structural characterization of HIV-1 Env

Three-dimensional structures have become increasingly prominent in HIV vaccine research. Without a doubt, the landmark work reported by Kwong and colleagues in 1998 describing the first high-resolution snapshot of the exterior of gp120 was a pivotal point in structure-based HIV-1 vaccine design and is still a paradigm for research on the Env structure and function to this day (153). More specifically, this structure characterized the interactions between a deglycosylated monomeric gp120 core, the Fab 17b and a two domain fragment from CD4 at a maximum resolution of 2.5 Å. Subsequently this group and others also described crystal structures of 1) gp120 core structures from different HIV-1 clades (IIIB, YU2 and JR-FL), 2) a complex of this gp120 construct possessing a variable loop (V3), and 3) core gp120 in complex with a different Fab (X5) (120, 152). Finally, the structure of a fully glycosylated unliganded gp120 from an SIV isolate at a maximum resolution of 4 Å was solved (51). These four published structures, in addition to several other studies reporting co-crystal structures between Fab molecules and domains of gp120 have provided the bulk of structural information related to HIV-1 Env interactions with receptor, co-receptors, and antibodies (reviewed in (218)).

On the other hand, the gp41 subunit has been much less well-characterized structurally. In the late 1990’s, the crystal structures of the HIV-1 and SIV gp41 heptad repeats were solved and revealed a self-forming six-helix coiled-coil bundle (50, 172, 283). This conformation was presumed to be that of post-receptor binding in the HIV-1 Env-mediated fusion process based on resemblance to other fusion processes. Indeed, such a coiled-coil structural feature had been previously described in the hemagglutinin membrane spanning subunit (HA2), in the transmembrane subunit of Moloney murine leukemia virus, as well as in the snarepin fusion
machinery involved in intracellular fusion events (39, 48, 82, 255). It is from these much better structurally characterized systems that the current understanding of the HIV-1 gp41-mediated membrane fusion is derived.

These atomic resolution breakthroughs have allowed establishing working models of HIV-1 Env-mediated fusion, as well as guide past and current drug and vaccine design strategies. In addition, various electron microscopy experiments were developed to provide a better global structural perspective of the native HIV-1 Env. Initially, a cryoelectron microscopy study of SIV viruses at approximately 3 nm resolution reported that Env was organized in a three-lobed gp120 trimer supported by a stalk made of three separated gp41 legs (302). This tripod-like arrangement for the transmembrane protein on the surface of the virus had previously been described in the case of the Moloney murine leukemia virus and was hence appealing (87). However, a subsequent cryoelectron microscopy study with a 2.8 nm resolution also focused on the tomography of Env on the surface of SIV performed by another group revealed that although the gp120 membrane-distal configuration was trimeric, it was supported by tightly interacting monomers forming the gp41 trimer, hence having an overall mushroom-like structure (298). These two contradicting structures emphasized the clear limitations of this lower-resolution technique and further stressed the development of better methodologies and unbiased averaging algorithms to bring an additional level of confidence to this promising technology. Since, much progress has been made in cryoelectron tomography techniques which led to the determination of a mushroom-shaped native HIV-1 Env structure, a b12-bound HIV-1 Env structure, as well as a CD4-17b HIV-1 Env bound structure (165). When interpreted with caution, these structures have yielded invaluable information into better understanding the trimeric nature of Env, the conformational changes of the gp120 trimer induced by CD4 binding, as well as the molecular-level recognition of HIV-1 Env by neutralizing and non-neutralizing antibodies.

What is now required is to continue the structural characterization of missing components of the fusion cascade and hence expand our knowledge of the intricacies and vulnerabilities of HIV-1 Env. Namely, the next challenges in HIV-1 Env related structural biology will be to characterize at the atomic level constructs that are increasingly representative of the conformations adopted by the native Env during the fusion process. Long sought-after and high-priority crystal structures which have yet to be solved include the native and hypothetical fusion-intermediate conformations of gp41, trimeric gp120, a complex between gp120 and co-receptor molecules
(CCR5 and CXCR4), a complex between gp120 and alternative receptors such as galactosyl ceramide, as well as the native gp120-gp41 trimer of dimers. In addition, to better understand the neutralization mechanisms of the rare bnmAbs that recognize HIV-1 Env, structural determinations would be invaluable in the case of 2F5 and 4E10 binding to their core epitope in the presence of membrane components and of elongated gp41 constructs. As well, it would be interesting to determine the crystal structure of b12 in complex with a gp120 trimer to identify the exact binding determinants of this specific interaction, including the identification of any unforeseen roles of the light chain of b12 in mediating binding. Furthermore, significant information could be gained by characterizing the interaction of 2G12 with carbohydrates on the surface of gp120 at the atomic level. Finally, it is evident that the structural characterization of the two newly identified trimer-specific bnmAb PG9 and PG16 in complex with native HIV-1 Env now constitutes a major research effort. Altogether, much remains to be done in the quest to better understand the HIV-1 fusion process at an atomic-level resolution with the objective to better guide immunogen vaccine-design and drug-design efforts targeted at the few vulnerable sites of HIV-1 Env.

In the meantime until a native Env trimer crystal structure becomes available, it could also prove informative to carry out further experiments on the molecular characterization of HIV-1 Env by other techniques, such as cryoelectron tomography studies and small-angle X-ray scattering.

5.5.2 Conclusions

In a few instances so far, structure-assisted therapeutic discovery has been proven successful. This includes the discovery of a potent HIV-1 protease inhibitor, Darunavir, which interacts with the backbone of the protease active site to combat drug resistance (101). It is however still not decided if such a structure-based approach can be successful in designing an HIV-1 vaccine capable of providing sterilizing immunity. It is a daunting task to predict which one of the few HIV-1 Env conserved and vulnerable sites can lead to the design of an effective immunogen that will elicit a potent and broad humoral response. What is becoming increasingly clear, however, is that such a vaccine will probably have to include many components and possibly an elaborate vaccination schedule. Advances in our understanding of the HIV-1 Env structure at the atomic level promises to provide additional tools and strategies to pave the way forward for the design of a much-needed HIV-1 vaccine.
References


2. 01/06/2010, posting date. Antiretroviral drugs used in the treatment of HIV infection. U.S. Food and Drug Administration. [Online.]


immunodeficiency virus type 1 gp120 glycoprotein critical for interaction with the CD4 receptor. Cell 50:975-85.


defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. Cell 86:367-77.


**Appendix 1: Attempt at crystallizing the 2F5 Fab’ in complex with the gp41 ectodomain in a post-fusion six-helix bundle conformation**

In order to determine the full scope of interactions between the bnmAb 2F5 and its epitope on HIV-1 gp41, a co-crystal structural analysis of this complex would be very valuable. Indeed, an increasing number of reports are now documenting secondary interactions between 2F5 and gp41 that are located outside of the primary epitope-paratope, and some of these weak interactions have been shown to be essential for HIV-1 neutralization (7, 109, 134, 176, 204, 233, 234). Although the interactions between 2F5 and its linear core epitope has been described in detail, a better structural understanding of the whole interaction interface between these two proteins would be extremely useful in guiding the design of an immunogen seeking to effectively elicit 2F5-like neutralizing antibodies. It is for this purpose that a significant portion of the present thesis was devoted to attempting the co-crystallization of the 2F5 Fab’ with the gp41 ectodomain in its lowest energy form, the six-helix bundle representing the post-fusion conformation of gp41.

**Materials and Methods**

The production of 2F5 Fab’ for crystallization was described in Chapter 2 of this thesis.

Two gp41 ectodomain constructs were designed for co-crystallization experiments with 2F5 Fab’. The first construct was obtained through collaboration with Aventis Pasteur and its design has been previously reported (146). Briefly, a stable gp41 ectodomain construct named Q4 (sequence

\[
\begin{align*}
535\text{MLTVQARQL LSGIVQQNQ \text{ LLRAIEAQQH LLQLTVWGIK}}
\end{align*}
\]

QLQARILAVE RYLKDQQLDG IWGCSGKDDC TTAVPWNASR SNKSLEQIWN
NMRTWMEWDR EINNYTSIHS LIEESQNQKE KNEQELLELD KWA^{667}SGGGGS
HHHHH)

was produced in *E. coli* and purified from inclusion bodies. It had been previously described how the bacterial recombinant gp41 ectodomain forms insoluble aggregates at neutral pH, an effect that is predominantly mediated by residues located in the loop region between the N and C heptad repeats (146). For this reason, in this region several mutations were introduced to
further increase the solubility of the construct at neutral pH. These mutations were L593D, L602D, I603D, W614R, H624N and T625M.

The second gp41 ectodomain construct used for co-crystallization with 2F5 Fab’ was described in Chapter 4 of this thesis and spans residues 535-669 (which includes the 2F5 epitope) of gp41, with an SGGRGGG-linker replacing native residues 580-627. This construct, termed gp41-linker also contains a (His)$_6$ tag at its C-terminus, immediately after the 2F5 epitope.

Both constructs have an approximate molecular weight (MW) of 10-14 kDa, but self-assemble into trimers of MW ~35 kDa. Both gp41 ectodomain constructs were expressed in E. coli, purified from inclusion bodies using the same procedure and yielded approximately 50-100 mg of purified protein per liter of media culture (refer to Chapter 4 of this thesis for a complete description of the preparation protocol.)

Results

Attempts at the formation of a homogeneous 2F5 Fab’-gp41 Q4 heterocomplex

Although mutations had been added in the gp41 Q4 construct, it was apparent that its solubility at neutral pH was limited. Krell et al. documented this threshold to be 80 μg/ml (146). The major difficulty in attempting to isolate a complex between 2F5 Fab’ and gp41 Q4 for crystallization was therefore to find a buffer in which the two proteins were stable. gp41 Q4 is highly soluble at a pH of 3, but becomes increasingly insoluble as the pH rises. The paratope/epitope interaction between 2F5 and gp41 is mediated, amongst others, by an arginine/glutamic acid ionic interaction, which could be disrupted at such a low pH where the glutamic acid residue probably becomes protonated. Indeed, competitive ELISA binding assays suggested that a minimum pH of 4.2 was needed for the binding of 2F5 to its gp41 core epitope (data not shown).

The optimal protocol found for mixing the 2F5 Fab’ with the gp41 Q4 construct consisted of first diluting the gp41 Q4 to 20 μg/ml in 50 mM formate, pH 3.0. Then, this solution was dialysed extensively against 2 L of 10 mM sodium acetate, pH 4.4, 200 mM guanidine-HCl, 5% glycerol. Finally, a 1:1 stoichiometric amount of 2F5 Fab’ was added at a concentration of 40 μg/ml in the same dialysis buffer. The 2F5 Fab’ + gp41 Q4 solution could then be concentrated to a maximum concentration of 2.8 mg/ml (Fig. A1.1A).
In order to isolate a homogeneous 2F5 Fab'-gp41 Q4 heterocomplex, size exclusion chromatography was used. It was found that the easily-aggregating properties of gp41 Q4 made it unstable on a Superose 12 HR 10/30 and that it had a tendency to bind to this matrix, even in the presence of 2F5 Fab’. For this reason, it was necessary to use a buffer of 10 mM sodium acetate, pH 4.4, 800 mM guanidine-HCl, 10% dioxane and 5% ethanol to obtain the proper elution of the gp41 Q4 construct on this size exclusion matrix. In this running buffer, a peak containing both 2F5 Fab’ and gp41 Q4 could be isolated, which eluted at an earlier volume than both individual proteins, suggesting the presence of a higher molecular weight species (Fig A1.1B and C). Re-running this isolated peak on the same size exclusion column resulted in a single peak at the same elution volume (data not shown).

To confirm complex formation in solution, a Native-PAGE electrophoresis experiment was performed. Because of the sensitivity of the gp41 Q4 complex to pH changes, the electrode-buffer consisted of a mixture of acetic acid/β-alanine, pH 4.5, whereas the separation gel was made with an acetic acid/KOH, pH 4.3 solution. For gel polymerization, because TEMED is not active at low pH, 0.01% riboflavin was used followed by exposure of the casted solution to light for several minutes. The presence of a (faint) third band with slower migration properties in the Native-PAGE electrophoresis experiment suggests the formation of a 2F5 Fab’-gp41 Q4 complex in solution. However, the assembly and stability of this complex seems to be weak, as after electrophoresis migration, the resulting intensity for the bands corresponding to the individual proteins is much higher than that of the 2F5 Fab’-gp41 Q4 complex (Figure A1.1D).
Figure A1.1 Attempt at the formation and isolation of a 2F5 Fab'-gp41 Q4 complex. (A) SDS-PAGE electrophoresis of complex formation in buffer of 10 mM sodium acetate, pH 4.4, 200 mM guanidine-HCl, 5% glycerol. (B) Chromatograph for running the 2F5 Fab'-gp41 Q4 mixture on a Superose 12 HR 10/30 size exclusion column in a running buffer containing 10 mM sodium acetate, pH 4.4, 800 mM guanidine-HCl, 10% dioxane and 5% ethanol. The major elution peaks in this running buffer for 2F5 Fab' alone and gp41 Q4 alone are 14.1 ml and 20.0 ml, respectively (data not shown). (C) SDS-PAGE electrophoresis analysis of the isolated size exclusion chromatography peak eluted at 13.71 ml. Both 2F5 Fab’ and gp41 Q4 proteins appear to be present in this peak. (D) Native-PAGE electrophoresis analysis at pH 4.5 of the isolated size exclusion chromatography peak eluted at 13.71 ml. In addition to the two individual bands of the 2F5 Fab’ and gp41 Q4 proteins, a faint third band seem to be present, indicating the presence of a 2F5 Fab'-gp41 Q4 complex in solution at pH 4.5.

A major limitation of the present protocol was a low recovery of protein from the size exclusion experiment. Indeed, only 30% of the protein put on the column eventually eluted, a phenomenon
probably explained by the easily-aggregating properties of the gp41 Q4 construct. In addition, the complex nature of the running buffer could hinder the chances of crystallizing this protein complex. Nonetheless, it was possible to concentrate the 2F5 Fab’-gp41 Q4 solution recovered from size exclusion chromatography up to 3 mg/ml. This complex mixture, in addition to 2F5 Fab’-gp41 Q4 complexes mixed at different molar ratios and not further purified by size exclusion chromatography were used to perform vapor diffusion crystallization experiments using standard Hampton Research and Qiagen commercial screening solutions with solutions at a pH around 4.5. The only crystals obtained from these screens came from a well solution of 0.1 M sodium cacodylate, pH 5.6, 0.2 M ammonium sulfate and 30% PEG 4000. Unfortunately, X-ray diffraction experiments and structural analysis of these crystals revealed that only the 2F5 Fab’ was present in the crystals. These results further confirmed the labile nature of the 2F5 Fab’-gp41 Q4 complex and emphasized the requirement of an improved gp41 ectodomain construct.

*Attempts at the formation of a homogeneous 2F5 Fab’-gp41-linker heterocomplex*

In an attempt to resolve the aggregating properties of the gp41 Q4 construct, the gp41-linker was designed. Briefly, in this construct, residues 580-627 of the inter-helices loop, which had previously been implicated in the aggregation properties of the recombinant bacterially expressed gp41 ectodomain were replaced by a an SGGRGG-linker. After purification from inclusion bodies, this construct also showed maximum solubility at pH 3.0 and limited solubility at neutral pH. However, a new range of pH ranging from 8.5 and higher was found for keeping this new gp41-linker construct soluble. Binding of 2F5 to gp41-linker at different pH values was verified by competitive ELISA (Fig A1.2A). The approximate IC$_{50}$ values derived from this experiment suggest that affinity of binding of 2F5 to gp41-linker at a pH around 4-4.5 is around one to two order of magnitudes better than at pH 9-10.

Size exclusion chromatography experiments were performed to attempt to isolate a homogeneous 2F5 Fab’-gp41-linker complex. At pH 9.0, the gp41-linker construct eluted on a Superdex 75 HR 10/30 column in running buffer of 0.1 M Tris, pH 9.0, 0.1 M NaCl, 5% glycerol at an apparent molecular weight corresponding to its expected trimer configuration (33 kDa). When gp41-linker was pre-mixed with 2F5 Fab’ in the running buffer, size exclusion chromatography revealed no apparent complex formation and the two observed peaks
corresponded to the elution profile of the individual proteins (Fig A1.2B). The gp41-linker was found to be much less stable at pH 4-4.5 and to behave similarly to the gp41-Q4 construct on a size exclusion column (data not shown). The deletion of the aggregation-prone gp41 inter-helices loop did not therefore result in an easier isolation of the 2F5 Fab’-gp41 ectodomain complex by size exclusion chromatography.

Nonetheless, a number of crystallization screening experiments were set-up with different mixtures of 2F5 Fab’-gp41-linker. Because it is uncertain how many 2F5 Fab’ molecule bind to the gp41 ectodomain trimer (possibility of three binding sites), 1:1, 2:1 and 3:1 molar mixtures of 2F5 Fab’-gp41-linker were made in pH 4-4.5 and pH 8.5-9. These mixtures were subsequently concentrated to concentrations of approximately 5 mg/ml and hanging drop vapor diffusion crystal screening experiments were set-up with commercially available solutions. Unfortunately, no protein crystals were obtained from these crystal screening experiments.

Figure A1.2 Attempt at the formation and isolation of a 2F5 Fab’-gp41-linker complex. (A) Competitive ELISA binding assays of the 2F5 IgG pre-incubated with the gp41-linker as a competitor at different pHs. The
Competitive ELISA experiments were performed as described in chapter 4. (B) Chromatograph for running the 2F5 Fab’-gp41-linker mixture at a concentration of 1 mg/ml on a Superdex 75 HR 10/30 size exclusion column in a running buffer containing 0.1 M Tris, pH 9.0, 0.1 M NaCl, 5% glycerol. (C) SDS-PAGE electrophoresis analysis of the 2F5 Fab’-gp41-linker mixture used for crystallization screening experiments.

Conclusion

To improve the chances of obtaining a crystal for a complex between two proteins, the prior isolation of a homogeneous species is preferable. Because of the low affinity between 2F5 and gp41 in its lower energy post-fusion six-helix bundle conformation (K_d in the low μM range (90)), such a homogeneous complex was found to be difficult to isolate by size exclusion chromatography. Furthermore, the absence of glycans on the gp41 ectodomain constructs due to their bacterial origin led to significant solubility problems, except in small ranges of extreme pH values. From these results, it becomes evident that a complex between 2F5 and its full gp41 ectodomain would be most relevant and achievable if the gp41 construct was glycosylated and adopting its fusion-intermediate conformation, for which 2F5 has increased affinity (K_d in the low nM range (90)).
Copyright Acknowledgements


