THE INFLUENCES OF CONFORMATIONAL DYNAMICS ON T CELL RECEPTOR SPECIFICITY AND CROSS-REACTIVITY

A Dissertation

Submitted to the Graduate School of the University of Notre Dame in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

by

Daniel R. Scott

Brian M. Baker, Director

Graduate Program in Chemistry and Biochemistry

Notre Dame, Indiana

April 2012
Abstract

by

Daniel R. Scott

Multi-specificity is a hallmark of T cell receptor (TCR) recognition, as a high volume of antigen must be identified by the T cell arm of the immune system. The mechanism by which a TCR engages a potential ligand and distinguishes its composition as either foreign or self – and doing this many times over with a diverse array of ligands – is a remarkable operation that remains mechanistically ambiguous. Protein-protein interactions have been characterized by a wide variety of biophysical and structural methods, which have resolved many key attributes of the interaction process. This is certainly the case of the TCRs and their antigen, peptides bound to cell-associated major histocompatibility complex (pMHC). However, holes in the arguments for how the recognition process is carried out are extremely confounding, calling for new approaches to more fully characterize the mechanism of TCR-pMHC recognition.
It has been well-established that protein function is heavily influenced by the
dynamic properties of a protein. Conformational diversity of TCRs has been revealed
through crystallographic and thermodynamic data, yet the pre-existing flexibility of the
receptor at its binding surface has gone relatively unidentified. Here we address the
fundamental dynamics of the A6 TCR protein, and consider the potential effects its
conformational variability may impose on its recognition strategy. By the means of
time-resolved fluorescence anisotropy and molecular dynamics simulations, distinct and
measureable flexibilities of the antigen-binding loops of A6 have been determined.
Towards the ultimate goal of elucidating the receptor’s recognition mechanism, a
comprehensive analysis of the data from both structural and energetic perspectives
have been critically performed.
I would like to dedicate this thesis to my parents, Alan M. and Alice W. Scott, whose countless sacrifices, unending patience, and unconditional love have given me every chance to succeed. In return, I hope to make them proud.
CONTENTS

Figures ............................................................................................................................. vi
Tables .............................................................................................................................. xiii
Preface ............................................................................................................................. xiv
Acknowledgments .......................................................................................................... xv

Chapter 1: Introduction ................................................................................................. 1
  1.1 T Cell Immunity ........................................................................................................ 1
  1.2 T Cell Diversity and Regulation ............................................................................ 5
  1.3 Structural Basis for T-Cell Receptor Recognition .................................................. 7
  1.4 Conformational Transitions of Proteins Upon Binding – Roles in Molecular
      Recognition ................................................................................................................ 9
  1.5 Folding Funnels Suggest Pre-existing Equilibria of Protein Conformations .......12

Chapter 2: Materials and Methods .............................................................................. 18
  2.1 Time-Resolved Fluorescence Anisotropy (TRFA) .................................................. 18
      2.1.1 Theory .............................................................................................................. 18
      2.1.2 Experimental procedures of TRFA .................................................................. 21
          2.1.2.1 DNA mutagenesis, protein expression and purification of
                 TCRs .............................................................................................................. 21
          2.1.2.2 Protein labeling ....................................................................................... 25
          2.1.2.3 Steady-state fluorescence measurements ................................................. 28
          2.1.2.4 Time-resolved fluorescence measurements ............................................. 29
  2.2 Molecular Dynamics (MD) Simulations .................................................................. 32
      2.2.1 Background .................................................................................................... 32
      2.2.2 Production of MD trajectories ....................................................................... 35
      2.2.3 Analysis of MD trajectories .......................................................................... 40

Chapter 3: CDR3 Hypervariable Loops Possess Independent Pre-Binding Energy
   Landscapes That Promote T Cell Receptor Binding and MHC Restriction ............... 41
  3.1 Summary .................................................................................................................. 41
  3.2 Results ..................................................................................................................... 45
3.2.1 The free A6 structure depicts rearrangements of the CDR3 loops in binding ................................................................. 45
3.2.2 Time-resolved fluorescence anisotropy reveals disparate CDR3 loop dynamics ......................................................... 48
3.2.3 MD simulations of the free A6 TCR .............................................. 51
3.2.4 Superimposition of the dynamic CDR3α loop onto HLA-A2 ........... 58
3.3 Discussion ......................................................................... 58
3.3.1 Dynamic rates of CDR sampling imply a pre-existing equilibrium of binding ............................................................... 58
3.3.2 CDR3 alpha loop maintains MHC specificity via conformational change .................................................................. 59
3.3.3 Future work ...................................................................... 61

Chapter 4: Mimicking Protein/Probe Dynamics Using Molecular Dynamics Simulations 63
4.1 Summary ........................................................................... 63
4.2 Results ................................................................................ 67
4.2.1 Fluorescence anisotropy reveals differential dynamics between the α and β chains of the TCR binding domain .................. 67
4.2.2 Molecular dynamics verify native-like flexibility for the A6-F5M constructs ................................................................ 75
4.2.3 Correlation functions of the simulated F5M probe agree well with TRFA flexibility values of the A6-F5M mutants ............ 76
4.2.4 Backbone vector correlation functions illustrate probe-independent CDR loop dynamics for four of the five mutants .... 81
4.3 Discussion ......................................................................... 86
4.3.1 Cross-validation of TRFA and computational methods for resolving site-specific protein dynamics ............................... 86
4.3.2 Re-examining the performance of fluorescence anisotropy ...... 87
4.3.3 Implications towards an A6 TCR binding mechanism ............. 89

Chapter 5: Correlated Motions Between A6 and Tax/HLA-A2 via MD Simulations – The Influence of Dynamics on the Entropy of Binding .................................................. 91
5.1 Summary ........................................................................... 91
5.2 Results ................................................................................ 96
5.2.1 Correlation values produced from the free A6 MD simulations are consistent with the previously-described dynamics of the CDR loops .................................................................... 96
5.2.2 The bound A6 MD simulation reveals an astonishingly low amount of entropically-costly interactions with Tax/HLA-A2 ligand .... 100
5.2.3 Distances and correlation of motion between contact atoms confirms CDR3α’s mode of MHC restriction .................... 106
5.3 Discussion ........................................................................ 112
5.3.1 Implications towards TCR binding mechanism .................... 112
Chapter 6: Two TCRs Similar to the A6 TCR – Examining the Interdependency of Loop Dynamics

6.1 Introduction ..................................................................................................................... 114
6.2 Results ............................................................................................................................ 117
   6.2.1 B7 and A6 have differential dynamics of their shared Vβ germline loops .......................................................... 117
   6.2.2 The high-affinity A6 c134 displays different dynamics that the wild type A6 ........................................................... 120
6.3 Discussion ....................................................................................................................... 124

Chapter 7: Conclusions ......................................................................................................... 127

7.1 Rationale Behind the Investigations ............................................................................... 127
7.2 The Role of CDR Loop Dynamics in pMHC Recognition ........................................... 128
7.3 Applications of Time-Resolved Fluorescence Anisotropy and Molecular Dynamics Simulations in Determining TCR Dynamics ......................................................... 130

Appendix A: Fluorescence Anisotropy Data for the Single-Chain 2C TCR ...................... 133
   A.1 Steady-State Fluorescence Anisotropy ........................................................................ 133
   A.2 Time-resolved Fluorescence Anisotropy ................................................................. 134

Bibliography .......................................................................................................................... 137
Figure 1.1 Antigen presentation and recognition of the class I MHC complex. Peptide fragments from endogenously-degraded proteins (Ag) are loaded onto the major histocompatibility complex (MHC). Class I MHCs are recognized specifically by cytotoxic T cells, which possess CD8 co-receptor proteins, via T cell receptor (TCR) engagement. ...........................................................2

Figure 1.2 The TCR-pMHC complex. (A) A relatively conserved orientation is observed in the TCR-pMHC interaction, positioning the Vα domain over the N-terminal portion of the peptide, while the Vβ domain typically interacts with the peptide’s C-terminal half. (B) The conformations of the specific binding loops of the TCR are more unique from one TCR-pMHC complex to the next. Shown here is the complex of the A6 TCR with an altered peptide ligand Tax P6A, associated with the HLA-A2 (a human MHC allele). Adapted from the PDB 1QRN structure (Ding et al., Immunity; 1999); visualized with PyMOL (DeLano Scientific LLC). .............4

Figure 1.3 A schematic of TCR gene rearrangement. T cell diversity is derived from random gene rearrangement of TCR-encoded DNA. Different variable and constant segments are annealed together, with additional nucleotide insertions and deletions at the joints of the DNA fragments. This image depicts the construction of the mRNA-encoded β polypeptide chain of an arbitrary TCR. .........6

Figure 1.4 Representations of pre-existing dynamics and its impact on ligand association. The folding of a protein can be depicted by a fluid progression of continuously sampling alternative energy states until the most stable conformation is reached (the “native” state; top image). Energy landscapes for folding, as shown here, illustrate this concept of proteins having multiple pathways to reach their native structure. A multi-specific protein can form complexes with various ligands by accessing relatively degenerate substates, each of which are energetically preferential to a specific ligand or ligands (cartoon depictions of pMHC are represented here, with two distinctly different peptides corresponding to the preferential substate of the TCR). ......................14

Figure 1.5 A schematic of two opposing protein binding mechanisms. Conformational selection and induced fit models differ in the order of events leading to final complex formation between protein (P) and ligand (L). For conformational
selection, structural rearrangement occurs before the protein’s association with ligand, while the induced fit model depicts rearrangement only after initial contacts have been formed.

Figure 2.1 Vector map of pET-28a. The gene construct for the single-chain 2C TCR was inserted into this vector for bacterial expression. pGMT7, the vector used for expressing all the human TCRs of this study, is not shown. Image taken from Novagen.

Figure 2.2 Labeling sites of the TCRs. Residues mutated to cysteine are underlined and in red.

Figure 2.3 Series of chromatograms for the A6 TCR. (A) S200 size exclusion column of the recently refolded TCR. (B) MonoQ column (anion-exchange) for separating dimmers of different charges. (C) S75 size exclusion column after labeling.

Figure 2.4 SDS PAGE gels of labeled A6 TCR. (A) Visible gel stained with Coomassie and SeeBlue Plus2 (Invitrogen) standard molecular weights (in kDa). Left of the standard “ladder” are various A6 samples with reducing agent (DTT) added; to the right are corresponding non-reduced samples. (B) The same gel under UV light, indicating the presence of fluorescence within specific bands (those polypeptide chains containing the cysteine mutant). UV image taken with KODAK DS 1D LE camera.

Figure 2.5 Components of a time-resolved anisotropy decay. Data collection was continued until either the parallel or perpendicular decay exceeded the other by 10,000 counts.

Figure 2.6 Examples of anisotropy decays with overlaying fitted models. The IRF function denotes time zero of the decay curves, and colored lines for each curve are the least-squares double exponential fits. These six samples are taken from the A6 TCR data.

Figure 2.7 Ranges of lifetime values measured for the A6 TCR. Notably, free fluorescein behaves similarly to bound fluorescein-5-maleimide, which free in solution has a much faster lifetime value. Error bars represent the standard deviation of averaged lifetime values for multiple measurements.

Figure 2.8 Atomic positions of the FSM molecule. Numbering of atoms corresponds with the attached FSM molecule in Table 2.1.

Figure 3.1 Free and bound positions of the antigen binding loops of A6. View of the backbone conformations as perceived from below the antigen (the Tax peptide is shown in pink stick for orientation). All bound TCRs were superimposed with the free A6 TCR by their Va-Vß backbone atoms before the non-binding portions of
the TCRs were removed for imaging. Visualized with PyMOL (DeLano Scientific LLC).

Figure 3.2 Free and bound positions of the CDR3 loops of A6. View of the backbone conformations as perceived from the side of the TCR-pMHC complex. See Fig. 3.1 caption for details on superimposition/imaging.

Figure 3.3 Labeling sites of the A6 CDR3 loops for TRFA studies. Side chains are of the native residues; each residue is singly mutated to cysteine for linkage to the fluorescent probe.

Figure 3.4 Anisotropy curves of the A6 CDR3 loops. Data was normalized to each residue’s own maximum value, which is set as $t = 0$ in the plot. Only one of at least seven data sets for each residue is shown.

Figure 3.5 Atomic fluctuations in the Vα-Vβ domains of the 200-ns MD simulations of A6. Values are calculated for the α carbon atoms as B-factors, and are in units of Å$^2$. Germline loops (CDR1, CDR2, and HV4) are plotted in red, CDR3α in orange, and CDR3β in blue; all other residues are plotted in black. The inset shows 4-ns PDB snapshots of the two loops throughout the 200-ns simulations, depicting their conformational diversity as they can be related to the B-factor values.

Figure 3.6 Psi/phi plots of key A6 CDR3α residues depicting conformational changes during simulations. Scattered plots represent the psi/phi values of the residue during each time step of the 200-ns simulations in black (recorded every 200 fs). Values of the psi/phi angles found in the A6 crystal structures are shown with the corresponding symbols in the legend.

Figure 3.7 Psi/phi plots of key A6 CDR3β residues depicting conformational changes during simulations. See caption for Fig. 3.6 for details.

Figure 3.8 RMSD calculations of the CDR3 loops in the 200-ns MD simulations of free A6. Backbone atoms of the CDR3α (top) and CDR3β (bottom) were superimposed with either the free, Tax-bound, or Tel1p-bound structure. Values were calculated every 200 fs.

Figure 3.9 RMSD calculations of the A6 CDR3 loops in the 200-ns “reverse” MD simulations of bound A6. See caption for Fig. 3.8 for details.

Figure 3.10 Energy landscapes of the CDR3 loops. A more discriminate two-state equilibrium exists for the CDR3α backbone (A), with a larger barrier separating the states than the more conformationally promiscuous CDR3β (B). Image taken from Scott et al., JMB; 2011.
Figure 4.1 **Labeling sites of the A6 TCR.** This ribbon structure of A6 illustrates the 19 backbone sites where cysteine mutations were made (yellow segments) for fluorescence labeling, both in the CDR/HV loops, and the S19α and N120α reference sites away from the binding loops (adapted from PDB 3QH3). ..............69

Figure 4.2 **Flexibility ranges of A6 in terms of \( f_i/\theta_i \) values.** Each position within the measured CDR/HV4 loops is denoted by a diamond with the residue number embedded. The values for the negative and positive reference sites of S19α and N120α are depicted as black and red vertical lines, respectively. ......................71

Figure 4.3 **Chemical structures of the fluorescein-5-maleimide (F5M) and BODIPY-FL (BDY) fluorophores.** Each probe has a maleimide functional group which allows thiol linkage to the cysteine side chain. For the F5M molecule, the two carbons encircled green were used to define the vector for the time correlation analysis of the MD simulations. Figures were taken from Invitrogen (© 2012 Life Technologies Corporation). .................................................................73

Figure 4.4 **Steady-state anisotropy measurements of F5M- and BDY-labeled A6 CDR loops.** A total of 11 sites compared on five of the A6 binding loops, using two chemically and structurally diverse probes......................................................74

Figure 4.5 **Correlation analysis of the F5M and BDY measurements..........................74**

Figure 4.6 **Atomic fluctuations of A6-WT and A6-F5M MD simulations.** (A) B-factors for the \( \alpha \) carbons of the A6 V\( \alpha \) and V\( \beta \) domains. The five A6-F5M MD simulations are shown in color, whereas A6-WT simulations starting with unbound (black) and bound (grey) coordinates are shown for reference. (B-D) Close-up views of the CDR loops which were perturbed by cysteine-probe incorporation.............77

Figure 4.7 **Time correlation functions of the F5M probe from MD simulations.** The average of ten 20 ns segments is shown for each mutant in color, having been calculated from the decorrelation of a defined vector within the F5M molecule. An unbound F5M molecule was simulated for 20 ns for reference (black). The upper-right inset shows the 1 ns region that was used for fitting. ......................78

Figure 4.8 **Derived order parameters of the backbone vectors in the A6-F5M and A6-WT MD simulations.** For the five independent simulations of A6-F5M, the backbone flexibilities of the mutated cysteine are quantified by time correlation functions. N-H, C\( \alpha \)-C\( \beta \), and C-O vectors are all included for the most comprehensive characterization of the backbone. The \( C_{\infty} \) values are cross-referenced with values from the wild-type residues of the A6-WT simulation (darker bars). ..........................................................82

Figure 4.9 **Correlation analysis of the A6-WT and A6-F5M order parameters.............83**
Figure 4.10 **A6 A99Cβ-F5M simulation depicting probe-protein interaction.** Each image represents a 2 ns snapshot of the average coordinates during a 40 ns segment of the simulation. The dynamics of the F5M molecule (stick) attached to the CDR3β loop (grey) at position 99 is inhibited by interactions with the CDR1α (green) and CDR3α (orange) loops.

Figure 4.11 **Correlated motion of the F5M molecule with the backbone of the attached loop.** Vectors within F5M were cross-examined with Cα-Cβ atoms for the five A6-F5M simulations; all 200 ns were included. Cross-correlation values range from -1 (anti-correlated motion; blue) to 1 (correlated motion; red). The residues of the loop which were included in the analysis are above its respective value, including the CYS residue to which the probe is directly attached.

Figure 5.1 **The A6-Tax/HLA-A2 complex.** The five polypeptide chains of the complex are labeled, with the A6 αβTCR heterodimer oriented at the top. Shown here is the complex of A6 with the altered peptide ligand Tax P6A, associated with the HLA-A2 molecule and accessory β2M (PDB 1QRN; Ding et al., Immunity; 1999). The simulation was performed with the coordinates of this structure, only the Tax residue of A6 was mutated back to the native P6. Visualized with PyMOL (DeLano Scientific LLC).

Figure 5.2 **Order parameters of the 200-ns free A6 MD simulations.** Values of each backbone vector for selected residues in the α (A) and β (B) chains of A6.

Figure 5.3 **Order parameters of the 50-ns bound A6 MD simulations.** Values of each backbone vector for selected residues in the α (A) and β (B) chains of A6.

Figure 5.4 **Comparing the free and bound RMSD values of the A6 binding loops.** Values are the average deviation of the loop’s backbone atoms in reference to the starting coordinates of the free A6 simulations (blue) and the bound A6-Tax simulations (red). Error bars represent the standard deviation of the RMSD values intermittently calculated from the simulations (every 200 fs for the free A6 simulations; every 1 ps for the bound A6-Tax simulations).

Figure 5.5 **Psi/phi plots of key CDR3 residues in A6-Tax simulations.** Conformational diversity in the CDR loops are depicted by backbone dihedral angles of CDR3α (left-hand side) and CDR3β (right-hand side) residues. For reference, dihedral angles are depicted for the starting bound structure (◊), the native structure (∆), and the free A6 structure (*).

Figure 5.6 **RMSD values for the CDR3β loop in A6-Tax simulations.** Backbone atoms of the CDR3β loop are superimposed to the free A6 structure (black) and the starting bound structure (red). Averages and standard deviations of both plots are given.
Figure 5.7 Contact distances of the A6-Tax/HLA-A2 MD simulations. Structural distances taken from the PDB 1QRN structure are shown for reference (blue). Distances between atoms of A6 residues (highlighted in lime) and atoms of the Tax peptide (cyan) or HLA-A2 (yellow) are average distances (in Å) intermittently calculated from the A6-Tax simulations (every 1 ps); error bars represent the standard deviation of these values. Alpha (A) and beta (B) chain residues of A6 are selected based on close contact in either the 1QRN or 1AO7 PDBs. ..........107

Figure 5.8 Correlated motions of the A6-Tax/HLA-A2 intermolecular contacts. Cross-correlation values represent the level of concerted motion between two defined atoms. A value of +1 is the highest level of correlation, while -1 is the highest anti-correlation two atoms can possess. Residues are highlighted as in Fig. 5.7; alpha (A) and beta (B) chain residues of A6 are depicted..........................109

Figure 6.1 Loop positioning of the A6 and B7 TCRs bound to Tax/HLA-A2. Complexes are superimposed by the backbone atoms of their HLA-A2 heavy chains. (A) The Vα loops of A6 and B7 have very similar loop positioning and conformations. (B) The Vβ loops, despite nearly 95% sequence similarity, have large displacements in their binding solutions to Tax. Side chain residues in each of the CDR loops are displayed to more clearly depict the translational deviations of the entire germline loops (E30 – red; I54 – purple) and the C-terminal half of the CDR3 loop (E105 – green). Darker colored loops belong to A6.............................................116

Figure 6.2 Differential dynamics of the Vβ germline loops of the A6 and B7 TCRs. (A) Representative anisotropy curves for positions within the CDR1β and CDR2β loops of the two TCRs. All measurements are with the F5M probe. (B) Average fitted values of the anisotropy curves, comparing the A6 and B7 TCRs’ identical positions in the Vβ germline loops. Error bars represent the standard deviations of at least four independent measurements of F5M-labeled TCRs. .................119

Figure 6.3 Lifetime measurements of the F5M probe attached to the A6 and B7 TCRs. Error bars, when present, represent the standard deviation of at least two independent measurements. .................................................................120

Figure 6.4 Distinguishing the dynamics of several CDR loops in the A6 and A6 c134 TCRs. All measurements were collected using the F5M probe. (A) Representative anisotropy curves of the wild type and high-affinity variant of A6. (B) Steady-state anisotropies of samples corresponding to (A). ......................121

Figure 6.5 Lifetime measurements of the F5M probe attached to the A6 and A6 c134 TCRs. Error bars, when present, represent the standard deviation of at least two independent measurements. .................................................................123
Figure 6.6 **Hypervariable CDR3β sequences for the A6 and B7 TCRs.** Identical residues of the TCRs are highlighted in red. ................................................................. 124

Figure 6.7 **Identical backbone conformations of the A6 and A6 c134 TCRs bound to Tax/HLA-A2.** The CDR loops of A6 investigated in this study are shown, completely superimposable to the high-affinity variant c134, and are therefore not distinguishable. Side chains of the mutated residues of the c134 CDR3β loop are labeled. ........................................................................................................ 125

Figure A.1 **SSFA measurements at 25°C.** ................................................................. 133

Figure A.2 **Anisotropy curves at 25±1°C.** ................................................................. 134

Figure A.3 **Diagram for the TRFA 2C data.** ................................................................. 136
TABLES

Table 2.1 Electrostatic potential values of the simulated F5M probe ......................37
Table 3.1 RMSD values of the various bound A6 CDR loops from their free positions ....46
Table 3.2 Correlation times and amplitudes from the TRFA measurements .................50
Table 4.1 Fitted values from the A6 TRFA measurements .....................................70
Table 4.2 Correlation times and amplitudes of the F5M molecule in MD simulations ...79
Table 5.1 Average order parameters of the A6 MD simulations ................................97
Table 6.1 The correlation times and amplitudes of the A6 and A6 c134 TCRs ..........122
Table A.1 TRFA data for the 2C TCR ..................................................................135
The interaction between the T cell receptor (TCR) protein and its cell-associated ligand is a critical element in T cell activation. Proper surveillance of the organism by the T cell is dependent upon effective discrimination between foreign and self-antigens — yet naïve T cells selected during lymphocyte development have only been exposed to and tested against self-antigen. The remarkable discrepancy of a TCR in its interaction with antigen is easy to appreciate but difficult to comprehend. To attain an accurate depiction of the TCR recognition mechanism, investigations must include a wide range of experimental approaches to adequately assess the chemical, structural, and dynamic processes of TCR function. Successfully elucidating the strategy by which a TCR “sees” the ligand surface is a critical step towards discovering new applications of TCRs for vaccine design and cancer immunotherapies.
ACKNOWLEDGMENTS

I would like to formally give credit to the people who have made my work possible. First, the co-workers within the Baker lab: I could not have survived the long hours in the lab without you. To my collaborators: Dr. P. Kamat, Dr. S. Corcelli, Dr. K. Tvrdy, Dr. D. Miller, and the Center for Research Computing: all of whose technical assistance was instrumental in my research, as well as in my development as a biophysicist. Finally, to Dr. Brian M. Baker: a true teacher if I have ever seen one.
CHAPTER 1: 
INTRODUCTION

1.1 T Cell Immunity

The survival of an organism depends heavily on its ability to combat disease and infection. For vertebrates, the adaptive immune response has evolved as a defense mechanism for fighting off foreign, as well as aberrantly-derived, invaders that pose a threat to an organism’s health and sustainability. Adaptive immunity relies on memory – previous encounters with foreign antigen equip the system with strategies to respond to later infections more effectively. This, of course, is the premise for vaccination. Hence, our understanding of the immune system’s basic operations can cultivate life-changing breakthroughs in medicine.

The two major cellular entities of the adaptive immune system are B cells and T cells. Their names are derived from the location of their maturation stage: B cells differentiate from precursor lymphocytes within bone marrow, while T cells mature within the thymus. Despite their common lineage from immature lymphocytes, B cells and T cells operate distinctly from one another. B cells release their receptor proteins (called immunoglobulins, or antibodies) into their surroundings, which then adhere to soluble antigen. In contrast, T cell receptor (TCR) proteins remain intact with the
cellular membrane, and recognize complementary cell-associated antigen (Fig. 1.1). The TCR’s antigen is composed of the host cell’s major histocompatibility complex (MHC) and a bound peptide, which has been degraded within the cell. Thus, the T cell and its receptors have the remarkable task of determining a threatening component (peptide) in the presence of a self-entity (the MHC scaffold protein). The fact that the fate of the cell, and subsequently the organism, depends on such a subtle differentiation is nothing short of spectacular.

Figure 1.1 Antigen presentation and recognition of the class I MHC complex. Peptide fragments from endogenously-degraded proteins (Ag) are loaded onto the major histocompatibility complex (MHC). Class I MHCs are recognized specifically by cytotoxic T cells, which possess CD8 co-receptor proteins, via T cell receptor (TCR) engagement.
The deft precision by which the TCR/pMHC recognition process is carried out has consumed several decades of research time and resources, yet a clear, unanimous depiction of this mechanism has not transpired. A significant source of complication in understanding the T cell activation process comes from the unknown roles played by other T cell co-receptor proteins. A T cell also expresses either a CD4 or CD8 co-receptor molecule, depending on the type of T cell, in addition to three accessory complexes: CD3εδ, CD3εγ, and CD3ξξ. While evidence suggests these extracellular co-receptors stabilize the TCR/pMHC complex and induce downstream signaling, much more information must be revealed about the timing and synergy of their interactions.

At the forefront of T cell activation, however, is the initial TCR/pMHC interaction. The nature of the peptide, as determined by its direct contact to the TCR, dictates how the T cell responds to the host cell. The binding surface of the TCR, therefore, must be optimized for specificity to its pMHC ligands. The TCR is comprised of two polypeptide chains, forming either an αβ or γδ heterodimer, which fold elegantly into a highly conserved tertiary structure of predominantly beta sheet motifs (Fig 1.2A). A composite surface of eight β-turn loops, which are referred to as complementarity determining regions (CDRs) and hypervariable (HV) regions, is thus created by this fold to produce the pMHC binding site. Since the number of antigenic peptides is extraordinarily large, the adaptive immune system has evolved an exquisite process to introduce randomization into the binding loops of the TCR.
Figure 1.2 The TCR-pMHC complex. (A) A relatively conserved orientation is observed in the TCR-pMHC interaction, positioning the Vα domain over the N-terminal portion of the peptide, while the Vβ domain typically interacts with the peptide’s C-terminal half. (B) The conformations of the specific binding loops of the TCR are more unique from one TCR-pMHC complex to the next. Shown here is the complex of the A6 TCR with an altered peptide ligand Tax P6A, associated with the HLA-A2 (a human MHC allele). Adapted from the PDB 1QRN structure (Ding et al., Immunity; 1999); visualized with PyMOL (DeLano Scientific LLC).
1.2 T Cell Diversity and Regulation

The defining moment in the lifetime of a T cell takes place during the recombination of its TCR-encoded genome. By the random selection and annealing of many available variable (V), diversity (D), and joining (J) genes, a developing T cell will almost certainly express a unique TCR from its brethr (Fig. 1.3). In humans, for example, T cell receptors can be derived from a combination of approximately 70/61 V/J gene segments for its α chain, and 52/2/13 V/D/J gene segments for its β chain.\(^6\) Additional nucleotide deletion and insertion processes at the junctions of the annealed genes further diversify the receptor repertoire, allowing a single T cell to express a clonotypic receptor protein from a possible \(10^{15}-10^{18}\) distinct primary sequences.\(^6-8\) Thus, an expansive TCR repertoire is produced to defend against an almost limitless number of chemically and structurally distinct antigenic peptides.

The process of V(D)J gene recombination only affects the domain which directly interacts with the pMHC complex, and is thus called the variable domain. The largest amount of sequence variation is concentrated at the third CDR loop (CDR3) of each chain as a result of the N- and P- nucleotide deletion and insertions, whereas the rest of the variable domain is derived from the non-fragmented sequence of the V gene region. Consequently, the T cell repertoire within any given species will have common “germline” CDR1, CDR2, and HV loops amongst various TCRs, which differ only in their CDR3 sequence. After the V(D)J region of each chain, additional segments of non-fragmented splices are added to the 3’ end which encode the constant and
Figure 1.3 A schematic of TCR gene rearrangement. T cell diversity is derived from random gene rearrangement of TCR-encoded DNA. Different variable and constant segments are annealed together, with additional nucleotide insertions and deletions at the joints of the DNA fragments. This image depicts the construction of the mRNA-encoded β polypeptide chain of an arbitrary TCR.

transmembrane domains of the TCR heterodimer.

The creation of such a vast and random repertoire of TCRs requires the immune system to weed out those T cells which are dysfunctional and would disrupt the homeostasis of the organism. During their development within the thymus, T cells undergo selection processes by interacting with other immune system cells to determine their self-tolerance and efficacy. Simultaneous to this selection stage, T cells are differentiated based on which co-receptor molecule they express. T cells which maintain the CD4 co-receptor function as helper T cells, and recognize only class II MHC molecules. In contrast, those expressing CD8 are cytotoxic T cells, which possess specificity towards class I MHCs.
This trial-by-fire examination eliminates almost 99.9% of T cells before they are released into the periphery, a vivid indication of just how sensitive the TCR recognition process must be in order to differentiate between healthy and diseased cells. Investigations of the physical processes of T cell recognition, both in the thymic development stages and during bodily surveillance, seem to have produced more questions than answers. Yet advancements in our understanding of this immunological mystery have produced exciting improvements in such areas as vaccine design, organ transplants, and cancer therapies that have certainly benefited our health and quality of life. Much more can be accomplished, however, if a lucid depiction of the mechanisms of TCR/pMHC binding can be more clearly attained.

1.3 Structural Basis for T-Cell Receptor Recognition

Particular focus has been paid to the tertiary structures of the TCR and the pMHC, both in their free states as well as in complex with each other. The first crystallographic structure solved for a TCR/pMHC complex was of the human TCR A6 recognizing the virally-derived Tax peptide some 15 years ago, and roughly 50 subsequent crystal structures have since been obtained. In all cases, both proteins have their transmembrane domains removed to reduce aggregation, while leaving essential components of their binding domains intact. The current library of structures contains both human and murine TCRs in complex with both self and foreign pMHCs. Several of these TCRs have been artificially engineered to increase their affinity to the ligand.
Conversely, some pMHC molecules include altered peptide ligands as a means to gain a perspective into the sensitivity of a TCR’s recognition ability.\textsuperscript{16-19}

Given the large range of available TCR/pMHC complexes, attempts have been made to find common binding strategies that may underlie the mechanisms of antigen recognition. For example, TCRs have been shown to bind their pMHC antigens with a relatively conserved diagonal docking orientation, positioning their germline loops (CDRs 1 and 2) typically adjacent to the MHC helices that form the peptide-binding groove (Fig. 1.2B).\textsuperscript{20} In addition, TCR residues across several complexes have been identified as making conserved contacts to the MHC molecule.\textsuperscript{21} This has led to the hypothesis that germline loops of the TCR and the binding domain of the MHC molecule have evolutionarily co-developed restricted “rules of engagement.”\textsuperscript{22} Reinforcing this concept is that the germline CDR loops of a TCR are often observed with identical conformational states upon binding to multiple ligands.\textsuperscript{16} This is in stark contrast to the CDR3 loops, which commonly adopt alternative structures from one complex to the next.\textsuperscript{23,24} Exceptions to these “rules” have been observed,\textsuperscript{24,25} yet they have not discouraged the search for more evidence of this proposed binding theory.

One must keep in mind that a library of structures will inevitably have a bias towards those which can crystallize more readily – the microscopic subset of TCR/pMHC complex structures do not necessarily reflect the strategies used throughout the entire TCR repertoire. Therefore, the appearance of any universal binding mechanism may be a result of this bias (and could possibly be the cause of the bias!). As literally billions of
remaining TCR structures are still unsolved, it would be prudent to continue our investigations open-mindedly with regard to how the recognition process is carried out. While the impact of having crystal structures to assist our understanding of TCR/pMHC interactions cannot be overstated, alternative methods of investigation have begun to shed light on essential components of the recognition strategy, unable to be derived from structure alone.

One of the major shortcomings of using crystal structures to depict the binding mechanism is that they are static representations of the TCR and pMHC proteins. Thus they do not reveal the “process” – only the “before” and “after” ground-state conformations – of the TCR and/or pMHC interaction. Dynamic states of the proteins throughout the binding process can influence critical components of the interaction not observable by structural elucidation. This information has been extracted for a great number of other protein-protein complexes by many methods, and would be invaluable when applied to the TCR/pMHC binding mechanism.

1.4 Conformational Transitions of Proteins Upon Binding – Roles in Molecular Recognition

The notion of protein dynamics influencing function goes back many decades. One of the earliest theories which suggested its relevance, the “induced fit” model, was an extension of the lock and key model of binding proposed by Fisher in the late 19th century. It postulated that while molecular recognition was indeed dependent on the protein and ligand having complementary topographies, proper binding only comes
after a conformational change by the protein. Two steps are required for successful recognition: (1) an initial engagement of the protein with its ligand, and (2) a conformational melding of the protein-ligand interface that stabilizes the complex to a final ground state. The induced fit model, initially relying on interpretations of kinetic studies, has suited the description of many protein-protein interactions being investigated throughout the years.\textsuperscript{27-29} It has therefore deeply influenced the perception of the molecular recognition process, including that of TCR/pMHC binding.\textsuperscript{30-33}

The theory’s impact has been especially amplified with the aid of X-ray crystallography, which has made it possible to elucidate structural information for binding proteins. Structural rearrangements due to ligand binding have been observed for a large number of protein-ligand systems, further verifying the induced fit model as a reliable explanation for molecular recognition processes.\textsuperscript{34-36} The protein’s ability to conform to the ligand via side-chain rotamerizations and/or larger-scale backbone alterations could have two functions. First, conformational melding may occur to form energetically favorable contacts that are critical in the recognition process. Second, and not mutually exclusive from the first, rearrangement may steer the interface from penalizing steric clashes that could ultimately disrupt the interaction. In either case, the lifetime of the protein’s association with its ligand would be extended, increasing the efficacy of the recognition process, and thereby initiating the protein’s molecular function.
A very appealing notion resonates from the induced fit prescription for binding, with regard to its “conform or bust” strategy. Unwanted activation of a biomolecule can have detrimental effects on intra- and intercellular processes. Thus the second step of the induced fit mechanism (protein “refolding”) may have evolved to serve as an extra security measure to minimize inadvertent recognition that may be caused by purely lock and key mechanisms.\textsuperscript{37} Essentially, the recognizing protein has an opportunity to disengage from the ligand before it becomes fully “activated.” One could speculate that TCR/pMHC interactions are a particularly ideal example of the induced fit strategy differentiating between activating and non-activating ligands. The respective ligands being referred to here, of course, are the foreign peptides displayed from diseased cells, and the peptides of normally-expressed proteins within properly-functioning cells. Structural evidence of disparate TCR adaptability between foreign and self-pMHC has not yet transpired, but the expansion of the TCR/pMHC structural library will undoubtedly provide more opportunities to examine this prospective strategy of flexibility-induced protein activation.

Intriguingly, the structural rearrangements of a protein upon binding are not necessarily isolated at the reaction interface. Mechanical forces, for instance, stimulated by the binding of a protein to ligand could alter the overall structure of the molecule in regions distal to the binding site.\textsuperscript{38} For multi-binding proteins, the notion of this conformational “side-effect” has evolved into the allostERIC model of binding, where association to one ligand is necessary for optimizing the protein to bind a second
ligand.\textsuperscript{39} It has been suggested that the activity of certain proteins (e.g., enzymes) are regulated by these subsequent conformational changes. Evidence of an allosteric mechanism has been observed for TCR/pMHC complexes as well, both by structural analysis as well as fluorescence studies.\textsuperscript{40,41} This brings into question the role of the T cell co-receptor proteins of CD4 (helper T cells) or CD8 (cytotoxic T cells), as well as the role of the accessory proteins CD3εδ, CD3εγ, and CD3ξξ. Could the suspected conformational changes caused by TCR/pMHC engagement recruit any of these proteins, ultimately submitting an activation signal to the T cell? Speculation about this question has led to several proposed strategies,\textsuperscript{42-44} although this remains a relatively new topic of examination.

1.5 Folding Funnels Suggest Pre-existing Equilibria of Protein Conformations

The composition of proteins can range from very simple constructs of one polypeptide chain to an intricate assembly of multiple chains. Nonetheless, all proteins, starting out as amorphous strings of amino acids, typically achieve their final tertiary structure via a complex mechanism of folding processes. The protein folding pathway has been elegantly depicted by energy well and landscape models,\textsuperscript{45-47} which take into account the myriad of arrangements that protein segments can sample on the polypeptide’s route to reaching a final ground state. This process is not necessarily straight-forward, especially for larger proteins with more complex folds. Hence, the folding pathway as modeled by common energy landscapes account for reversibility in its mechanism, allowing the protein to backtrack in search of a more energetically and
kinetically favorable pathway. One could imagine that if folding were not reversible, misfolding would occur too frequently for cellular processes to cope.

This depiction of protein folding – intermediates being populated along the way towards an energy-minimized state – has impacted the mechanisms of protein binding as well (Fig. 1.4). Although it was often assumed that every protein had one definitive native fold, it was possible for the last of the intermediate states adjacent to the lowest energy well of the native structure to be occasionally populated (due to $k_B T$ thermal energy fluctuations). As measurements became available for addressing the dynamics of proteins, it became convincingly evident that proteins adopt multiple conformations. A protein’s native state, therefore, had to be redefined as the most-populated, lowest-energy state from a potentially vast number of nearly isoenergetic substates. Now it was postulated that as long as the kinetic exchange between these substates was fast enough (within the timescales of effective ligand binding), any of these conformations could be utilized for recognition.

This revelation in the dynamic state of the protein stood in opposition to the induced fit model – no longer should one exclusive structure of the protein be assumed as it engages its ligand. However, it was not necessarily assumed that a protein which undergoes structural changes upon binding (as seen from crystal structures) could access its exact binding-competent state before ligand association. In other words, one or more of the alternative substates that were accessed independently in the protein’s folding energy landscape could be an active form of the protein (Fig. 1.4). This
Figure 1.4 **Representations of pre-existing dynamics and its impact on ligand association.** The folding of a protein can be depicted by a fluid progression of continuously sampling alternative energy states until the most stable conformation is reached (the “native” state; top image). Energy landscapes for folding, as shown here, illustrate this concept of proteins having multiple pathways to reach their native structure. A multi-specific protein can form complexes with various ligands by accessing relatively degenerate substates, each of which are energetically preferential to a specific ligand or ligands (cartoon depictions of pMHC are represented here, with two distinctly different peptides corresponding to the preferential substate of the TCR).
argument for multi-state specificity was originally made within the Modon-Wyman-Changeux model (named after their founders),\textsuperscript{54} and later evolved into the conformational selection, or pre-existing equilibrium, models of molecular recognition (Fig. 1.5 shows conceptual difference between conformational selection and induced fit models). These equivalent theories essentially postulate that the conformational fluctuations of a molecule are not just random transitions in the folding pathway, but purposeful in the grand scheme of molecular recognition. Evidence of inherent protein dynamics all but ensured that the induced fit model would have an alternate strategy to compete with when elucidating a specific recognition scheme.

---

**Figure 1.5** A schematic of two opposing protein binding mechanisms. Conformational selection and induced fit models differ in the order of events leading to final complex formation between protein (P) and ligand (L). For conformational selection, structural rearrangement occurs before the protein’s association with ligand, while the induced fit model depicts rearrangement only after initial contacts have been formed.
While the dynamic nature of proteins makes conformational selection not only possible, but logical as well, it would be just as illogical to assume that upon binding, the dynamics of the two proteins in complex cease entirely. The energy landscape for each protein would most likely be altered, where new states may be accessible and exchange between conformations more concerted and interdependent. This stands as a plausible theory for how dissociation occurs: while in complex, one, if not both, of the molecules shift into a binding-incompetent, energetically-taxing state that causes instability and thus dissociation. The kinetics rates of protein-protein interactions are thus intrinsically dependent on the kinetic exchange of dynamical states of the proteins themselves. In this regard, a compromise between the induced fit and conformational selection models, although seemingly an admittance of each theory’s limitations, emphasizes the fine line between kinetically and energetically favorable events between protein and ligand.

Hybrid mechanisms have evolved due to the complex behavior of cross-reactive proteins with paradoxically high specificity. TCR/pMHC interactions seemingly fall within this category – multi-specificity is a hallmark of T cell activation, given the range of antigens a restricted repertoire of T cells must recognize. Focusing on just the TCR side of the interaction, determining the pre-binding flexibilities of the receptor protein could reveal which of the two mechanisms exist, and if both, how each contributes to recognition. Towards this goal, robust dynamic information has been obtained regarding the CDR loops of the A6 TCR – and to a lesser extent, the A6 c134, B7, and 2C
TCRs – using complementary techniques of molecular dynamics (MD) simulations and time-resolved fluorescence anisotropy (TRFA). The detection of possible binding mechanisms is thoroughly addressed, and future work towards this goal is considered.
2.1 Time-Resolved Fluorescence Anisotropy (TRFA)

2.1.1 Theory

By attaching a fluorescent probe to a specific site of a protein, the dynamic behavior of the protein can be inferred from the observation of the probe’s fluorescent properties. Specifically, the technique of TRFA involves the excitation of the attached probe with pulses of polarized light, and the recording of the fluorescence emission at both parallel ($I_p$) and perpendicular ($I_x$) angles to that of the excitation source. By exciting the transition dipoles of fluorophores with a distinct orientation (called photoselection), measurements of the solution’s fluorescence of two distinct polarities will quantify the probe’s angular deviation from that original orientation. The equation to describe the time-dependent anisotropy of the probe:

$$r(t) = \frac{I_p(t) - I_x(t)}{I_p(t) + 2 \cdot I_x(t)}$$

normalizes the difference of the orthogonal fluorescence intensities by their sum, eliminating any intensity-dependency. According to this equation, the theoretical maximum anisotropy is defined as one for all probes, irrespective of concentration,
quantum yield, lifetime, and so on. However, probes with transition dipoles parallel to the incident light are not exclusively excited – they just stand as the most probable to be excited. Therefore, the population of excited probes will already include a small but significant amount of non-uniformity even before time-dependent dynamics of the probe can decorrelate orientation. Statistically, the maximum anisotropy able to be measured for a random solution is calculated as 0.4 (although this value can be increased for oriented systems, and as one approaches single-molecule techniques, a value of one can be acquired). The time-dependent decay of anisotropy from its initial value \( r_0 \) can now be evaluated strictly as the dynamics of the probe during its fluorescence lifetime (from excitation to emission).

The timescale on which anisotropy measurements report dynamics is highly dependent on the average lifetime of the fluorophore. For example, a longer lifetime allows a probe to potentially lose a greater degree of correlation with its original orientation before it fluoresces. Thus, probes with various lifetimes can report on different dynamics of the protein to which it is attached, at timescales both earlier than and concurrent with that of the fluorescence. The phenomenon of fluorescence generally occurs within the nanosecond time regime, making TRFA a useful technique for determining relatively fast protein dynamics. The kinetics of many protein-protein interactions, including TCR-pMHC engagement, indicate that nanosecond dynamics of the binding domains of proteins may be relevant in recognition events. The exchange rates between pre-existing conformations of the TCR binding loops can
potentially influence how the receptor “sees” its pMHC ligands. Determining relative flexibilities of the CDR loops, therefore, can be instrumental in elucidating the binding mechanisms of the TCR.

A protein in solution has a propensity to rotate at a rate proportional to its molecular weight. This is described by a rotational diffusion time constant:

$$\theta = \frac{\eta V}{RT}$$

where the time constant ($\theta$) is dependent on the temperature ($T$) and viscosity ($\eta$) of the solution. This time constant can be theoretically derived, and for proteins the size of the TCRs worked with in these studies, fall in the low-to-mid nanoseconds when near room temperature.\textsuperscript{63} Fluorescence anisotropy measurements, therefore, can be sensitive to rotational motion. In order to resolve the dynamics of the loops occurring at a distinct timescale from the tumbling of the protein, a multi-exponential decay model is applied:

$$r(t) = r_{\infty} + \beta f \cdot e^{-t/\theta_f} + \beta s \cdot e^{-t/\theta_s} + \ldots + \beta n \cdot e^{-t/\theta_n}$$

where $r_{\infty}$ is the residual anisotropy (which was constrained to zero for all reported measurements), and $\beta x$ terms represent the amplitudes which correspond to the relative magnitudes of the separated correlation times. The number of exponentials used is limited by the resolution of the instrumentation, since ultrafast femtosecond lasers are capable of detecting a wider timescale of motions. In these studies, however, an LED source with a relatively long 1.3 ns pulse duration is used, thus requiring only a
two-exponential model – to detect the faster anisotropy decay attributable to the local flexibility of the protein ($\theta_f$), versus the decay produced mainly by the overall tumbling of the molecule ($\theta_s$). It must be noted that any slower conformational dynamics of the local protein segment may bleed into the rotational tumbling time constant, thus diluting its value by an indeterminate amount. Thus, variation of the slow time constant from theoretical approximations ($\theta_s$) could be observed.

2.1.2 Experimental procedures of TRFA

2.1.2.1 DNA mutagenesis, protein expression and purification of TCRs

The site-specific labeling of the receptor protein with a fluorophore requires covalent attachment via side-chain chemistry. The most common linkage sites are to amino groups of lysine side chains or the thiol groups of cysteines. Cysteine is a much less naturally-occurring amino acid in protein sequences, and for the TCRs being studied here, exist solely as disulfide-bonded cystines. There is minimal risk, therefore, for undesired covalent labeling. The gene constructs for the polypeptide chains of the human (A6, A6 c134, and B7) and mouse (2C) TCRs were independently inserted into plasmid vectors (pGMT7 and pET-28a, respectively; Fig. 2.1), which include antibiotic-resistant and IPTG-inducible lac operon genes for specific protein expression.

For the A6 TCR, the Cole-Sewell variant was the construct of choice for these experiments. It has an engineered disulfide bond in the constant domain, near where the transmembrane domain of the TCR has been genetically removed. It is significantly
Figure 2.1 Vector map of pET-28a. The gene construct for the single-chain 2C TCR was inserted into this vector for bacterial expression. pGMT7, the vector used for expressing all the human TCRs of this study, is not shown. Image taken from Novagen®.

more durable than other A6 constructs previously used in our laboratory, enhancing the refolding quality of the single cysteine mutants that are necessary for the fluorescence studies. The B7 and A6 c134 constructs also have equivalent disulfide bonds engineered in their constant domains. For the 2C TCR, a single-chain construct was utilized. This variant has the constant domains of each polypeptide chain removed, and possesses a 25-residue linker connecting the variable domains of the N-terminus of the alpha chain with the beta chain's C-terminus. In addition, a 6-histidine tag was engineered at the N-
terminus of the beta chain sequence for nickel column chromatography. The single-chain 2C construct has been shown to produce equivalent binding affinities to its pMHC ligands as that of the original 2C heterodimer. 66

The individual chains of single-cysteine TCR mutants were made using a site-directed mutagenesis kit (Quikchange®, Stratagene) and oligonucleotide primers with lengths of 25-35 base pairs (Integrated DNA Technologies). Residues for cysteine substitutions were chosen based on visual inspection of residues’ solvent accessibilities in the free TCR crystal structures, to promote optimal conditions for refolding and fluorescent probe labeling. The mutation sites selected are identified in Fig. 2.2, in addition to the reference sites S19Cα (a reference measurement for rigidity, located in a β-sheet of the TCR scaffold) and N120Cα (a reference measurement for flexibility, residing in an unstructured loop coil between the variable and constant domains) of A6. The individual α and β chains of A6, B7, and A6 c134 were expressed in BL21 (DE3) pLysS E. coli, and the single-chain 2C was expressed in BL21 (DE3) Codon Plus E. coli. Cells were harvested and lysed after inducing with 0.5mM IPTG, and the inclusion bodies (IBs) of the proteins were washed with a detergent-based buffer (Triton X-100 surfactant). The A6, B7, and A6 c134 IBs were solubilized in 8 M urea, and concentrations were determined by Bradford Assays using Coomassie staining.

For αβ TCR assemblies (A6, B7, and A6 c134), both chains were refolded together at approximately 1.5:1 ratios in 1L of refold buffer (2.5 M urea, 50mM Tris·HCl, 2mM
Figure 2.2 **Labeling sites of the TCRs.** Residues mutated to cysteine are underlined and in red.
EDTA, pH 8.0) and dialyzed against 10 mM Tris-HCl, pH 8.3 for 32-48 hours at 4°C. The refolded αβTCR dimer was then purified by alternating rounds of ion exchange (DEAE, MonoQ) and size exclusion chromatography (S200, S75), and buffer-exchanged into an optimal labeling buffer (20 mM phosphate, 150 mM NaCl, pH 7.0).

The 2C IBs were solubilized with 1-2 g guanidine HCl (s) and 2.5-5.0 μL β-mercaptoethanol (a reducing agent) and heating at 37°C. The denatured protein was then added to 400mL of refold buffer (3 M urea, 50 mM Tris-HCl, 2 mM reduced/0.2 mM oxidized glutathione, pH 8.0), then allowed to refold by drop-wise dilution for 24-30 hours using 2L of dilution buffer (50 mM Tris-HCl, 200 mM NaCl, pH 8.0) at 4°C. Ni-charged ProBond™ resin (Invitrogen) was then added to the protein and equilibrated for 18-24 hours. The protein-resin suspension was collected using vacuum filtration, then poured into a column for elution of the protein via affinity chromatography (500 mM imidazole, pH 7.4). The eluted protein was purified using size exclusion chromatography (S200, S75) and buffer-exchanged into labeling buffer (representative chromatograms are shown in Fig. 2.3).

2.1.2.2 Protein labeling

TCR mutants were reacted with a ten-fold excess concentration of fluorophore and 10-50 μM TCEP-HCl (reducing agent), alongside a parallel reaction with wild-type TCR for determining levels of non-specific labeling. Stock solutions of fluorophore (fluorescein-5-maleimide, F5M, and BODIPY-FL, BDY; Invitrogen) were prepared at 10
Figure 2.3 **Series of chromatograms for the A6 TCR.** (A) S200 size exclusion column of the recently refolded TCR. (B) MonoQ column (anion-exchange) for separating dimmers of different charges. (C) S75 size exclusion column after labeling.
mM in labeling buffer. In the case of BODIPY-FL, solubility was enhanced with up to 2% dimethyl sulfoxide (DMSO). After mixing for approximately 45 minutes in the dark at room temperature, excess label was separated out by dialysis and size exclusion chromatography. Final concentrations of A6 TCR were 7-21 µM, and labeling efficiencies varied amongst TCR mutants between 17 and 93%. \( \frac{A_{280}}{A_{495}} \) absorbance measurements confirmed that non-specific labeling of wild-type TCR never exceeded 2.1%, ensuring that the following lifetime and anisotropy measurements would report on predominantly site-specific fluorescence. Other TCRs had similar labeling efficiencies to A6. UV images of reduced/non-reduced SDS-PAGE gels verified that fluorescence emanated only from the chain with the cysteine mutant (representative gels are shown in Fig. 2.4). Intensity readings from steady-state fluorescence measurements (Beacon 2000; PanVera) also verified that the wild-type had relatively insignificant amounts of non-covalently associated label (see Section 2.1.2.3 for details).

Issues with the 2C TCR arose, where a significant amount of non-specific labeling was observed (~20%). Cleaving the 6-His tag with thrombin at a cutting site previously inserted into the sequence greatly diminished mislabeling, especially for the neutrally-charged BDY probe. Centrifugation (using a 10 kDa cutoff filter; Millipore\textsuperscript{8}) and an additional size exclusion column run were performed to effectively separate the ~1.9 kDa cleaved product from the still-active 2C protein.
2.1.2.3 Steady-state fluorescence measurements

Samples were constantly illuminated with a lamp source filtered at a wavelength near the absorption maximum of the probes ($\lambda_{\text{max}} = 492$ and 504 nm for F5M and BDY, respectively). Simultaneous fluorescence intensity and anisotropy measurements were collected at 25°C. Blank measurements were taken with aqueous phosphate buffer (20 mM phosphate, 75 mM NaCl, pH 7.4). Concentrations of samples were modified to fluoresce at an optimal intensity for anisotropy measurement (typically between 2000 and 7000 intensity units). This optimal concentration usually fell in the 100-400 nM range.
range. Samples of the wild-type (WT) TCRs needed a higher concentration, typically by at least an order of magnitude. At least ten data sets were collected per sample, spaced by approximately 30 seconds, to verify that the samples had properly equilibrated in the chamber. Intensity ratios of TCR mutants to the WT were calculated to verify minimal non-specific labeling. Average values of anisotropy measurements are reported in milli-anisotropy units (mA).

2.1.2.4 Time-resolved fluorescence measurements

Time-correlated single-photon counting (TCSPC) fluorescence lifetimes and anisotropies of labeled TCR mutants were measured at 25±1°C using a 5000U FluoroCube spectrofluorometer (HORIBA Jobin Yvon IBH Ltd.; Glasgow, UK). The samples were excited with a 457-nm NanoLED pulsed laser diode at a repetition rate of 1 MHz. Lifetime measurements were taken with the excitation and emission polarizers set to vertical and 54.7° from vertical (referred to as the “magic” angle, to eliminate anisotropic effects), respectively. The anisotropy measurements were collected as the emission polarizer was toggled between parallel ("p") and perpendicular ("x") positions relative to the vertically oriented excitation polarizer until a minimum peak difference of 10,000 counts was reached (representative curves shown in Fig. 2.5). A G-factor was experimentally determined to adjust for the polarization bias of the detection instrumentation. The resulting sum \(I_S = I_p + 2G \cdot I_x\) and difference \(I_D = I_p - G \cdot I_x\) decays were analyzed using the impulse reconvolution method provided by the instrument’s DAS6 software analysis package (IBH). This method corrects for pulse delays and non-
Data collection was continued until either the parallel or perpendicular decay exceeded the other by 10,000 counts.

zero pulse widths using a scattering sample (aluminum silica solution) to produce an instrument response function (IRF). The final reconvoluted function produces two correlation times, $\theta_f$ and $\theta_s$, from a double exponential decay fit (representative curves shown in Fig. 2.6). The fitted data for the sum decays determined in the TCSPC Anisotropy program, which are simply the total emission decays, closely agreed to those values produced using the TCSPC Lifetime program (lifetimes for the A6 TCR are shown in Fig. 2.7).
Figure 2.6 Examples of anisotropy decays with overlaying fitted models. The IRF function denotes time zero of the decay curves, and colored lines for each curve are the least-squares double exponential fits. These six samples are taken from the A6 TCR data.
Figure 2.7 *Ranges of lifetime values measured for the A6 TCR.* Notably, free fluorescein behaves similarly to bound fluorescein-5-maleimide, which free in solution has a much faster lifetime value. Error bars represent the standard deviation of averaged lifetime values for multiple measurements.

2.2 Molecular Dynamics (MD) Simulations

2.2.1 Background

Explicit-solvent MD simulations stand as a powerful tool to complement the fluorescence anisotropy studies of TCR flexibility. The two methods are both sensitive to sub-nanosecond dynamics, but the computational studies resolve additional features
that, for TRFA, are unattainable. While TRFA reports on the timescale, and potentially the magnitude, of local protein rearrangements, it cannot produce information regarding the actual conformation(s) being sampled by the protein. When deducing binding mechanisms based off of the pre-existing motions of the protein, this is a crucial element that MD simulations can determine with high resolution. MD conformational analysis can be performed by calculating torsion angle values for the backbones (or side chains) of individual residues, root-mean-square deviation (RMSD) fitting of selected atoms relative to known conformations in crystal structures, and producing intermittent PDB files for visual inspection. Combining these methods of analysis adds dimension to the perspective of the protein’s dynamic properties, while also providing an opportunity for testing the accuracy of the simulations (i.e., by showing that the protein is capable of native-like sampling of various conformations seen in crystal structures). Other advantages to performing MD simulations are that no perturbations are necessary to the protein, direct dynamics are observable, and as a single-molecule depiction of the protein, rare events can be resolved.

Conversely, one of the main limitations to explicit-solvent MD simulations is the time capacity of the trajectory length. Due to computational costs, oftentimes only relatively fast timescales are able to be observed (microseconds or less is typical, although millisecond timescales are beginning to be reached). For determining any reliable kinetic information, such as rates of exchange, conformational sampling needs to occur at a statistically-relevant frequency, meaning simulations must be run orders of
magnitude longer than the actual process. In the case of calculating time correlation functions, as is done in these studies, it is recommended to extend simulations to at least one order of magnitude longer than the desired time constant. While this may not seem difficult to achieve, again, statistically-reliable values are heavily dependent on averaging many observables. Thus, accumulating great lengths of MD trajectories is crucial for producing justifiable time constant values.

It also must be acknowledged here that MD simulations are only a representation of the actual behavior of the protein, as well as the solvent. In this regard, the force fields which are applied to the system inarguably limit the accuracy of the simulations. While force field parameters are continuously updated and modified based on high level quantum mechanics calculations as well as experimental agreement, they will always possess some inaccuracies when manipulating the atoms within a simulation. Thus, the selection of the force field (along with other initial parameters) becomes a critical aspect of the simulation production, and should be done so with consideration of what aspects of the system are being characterized.

In the case of studying the pre-existing dynamics of a TCR protein, the goal here is to address the conformational diversity of the CDR loops’ backbones, looking for any evidence of potentially binding-competent populations being sampled. Accurate treatment of torsion angles is thus a requirement, and the parm99 force field (denoted as “ff99SB”) has been recently modified to comply with such elements. Specifically, the improvement of glycine and alanine side chains has produced more accurate secondary
structure behavior with respect to PDB survey analysis and even NMR relaxation experiments.\textsuperscript{69} Even this force field has shown limitations in side chain rotamerization, recently prompting further refinement of its parameters.\textsuperscript{70} With the focus here on mainly backbone dynamics, however, these limitations may not have too large an effect on the outcome of this analysis.

Another aspect to consider that will have a potential effect on the dynamic behavior of the protein is the ensemble constraints. In this study, the microcanonical NVE ensemble is applied, meaning the Number of particles, Volume, and Energy are kept constant. The rationale behind selecting NVE comes from the fact that thermostats applied in the comparable NVT ensemble could induce non-physical perturbations to the protein, detrimentally influencing its dynamics.

Additional parameters applied to the system in creating the MD simulations are more briefly mentioned in the next section.

2.2.2 Production of MD trajectories

For all simulations in this study, atomic coordinates were directly taken from PDB structures (“free A6” – 3HQ3; “bound A6” and “A6-Tax complex” – 1QRN), with any crystal waters removed. For the A6-Tax complex, the pAla6 residue was mutated to Tax’s native residue, pPro6 (PDB 1QRN is used because of its more complete Cα domain of A6). Simulations were prepared and performed using the AMBER10 package. All structures had disulfide bonds added, were neutralized using sodium or chloride counterions, and solvated with SPC/E waters within a 10 Å cubic box, all by using the
teLEaP program. “Prmtop” and “inpcrd” files for these final systems were also created in teLEaP. Energy minimization and pre-equilibration steps were performed as follows (Kristina Furse as author of protocol; see website:

http://corcellicodelibrary.wikispaces.com/Amgber+MD):

1. Energy minimization with solute restrained at 500.0 kcal/mol/Å² force constant (500 cycles)
2. Energy minimization with no restraints (1000 cycles)
3. Defrost with temp ramp: 0.0 K → 300.0 K in NVT ensemble; solute restrained at 50.0 kcal/mol/Å² (20 ps)
4. Five subsequent segments where restraints on solute were incrementally relaxed in NPT (20 ps each; restraints reduced from 25.0 → 15.0 → 5.0 → 0.5 → 0.05 kcal/mol/Å²)
5. NPT segment with no restraints (50 ps)
6. Final volume of box reset to average during previous step, followed by NVT segment (50 ps)
7. Final temperature scaled to 300.0 K, followed by NVE segment (100 ps)

A 2-fs time step was collected using the SHAKE algorithm during all stages.

Systems were equilibrated for 2 ns in the NVE ensemble before production stages.

For the simulations of the five F5M-labeled A6 TCRs, the electrostatic potential (ESP) values for the fluorescein-5-maleimide atoms were derived using the GAUSSIAN/09 package (Table 2.1), as described by the Amber tutorial (Ross Walker as corresponding author of tutorial; see website:

http://ambermd.org/tutorials/advanced/tutorial1). Bond types, atom types, and partial charges of the F5M molecule were assigned in teLEaP. The cysteine mutation and
TABLE 2.1

ELECTROSTATIC POTENTIAL VALUES OF THE SIMULATED F5M PROBE

<table>
<thead>
<tr>
<th>Atom Name</th>
<th>F5M (Attached to A6)</th>
<th>F5M (Free)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Atom Type</td>
<td>ESP/Charge</td>
</tr>
<tr>
<td>O1</td>
<td>O2</td>
<td>-0.794790</td>
</tr>
<tr>
<td>O2</td>
<td>OA</td>
<td>-0.416518</td>
</tr>
<tr>
<td>O3</td>
<td>O2</td>
<td>-0.883984</td>
</tr>
<tr>
<td>O4</td>
<td>O</td>
<td>-0.864524</td>
</tr>
<tr>
<td>O5</td>
<td>O</td>
<td>-0.858423</td>
</tr>
<tr>
<td>O6</td>
<td>O</td>
<td>-0.629499</td>
</tr>
<tr>
<td>O7</td>
<td>O</td>
<td>-0.619033</td>
</tr>
<tr>
<td>N8</td>
<td>N</td>
<td>-0.413919</td>
</tr>
<tr>
<td>C9</td>
<td>CA</td>
<td>0.434055</td>
</tr>
<tr>
<td>C10</td>
<td>CA</td>
<td>0.112405</td>
</tr>
<tr>
<td>C11</td>
<td>CA</td>
<td>-0.372651</td>
</tr>
<tr>
<td>C12</td>
<td>CA</td>
<td>-0.316351</td>
</tr>
<tr>
<td>C13</td>
<td>CA</td>
<td>0.002894</td>
</tr>
<tr>
<td>C14</td>
<td>CA</td>
<td>0.561249</td>
</tr>
<tr>
<td>C15</td>
<td>CA</td>
<td>0.510881</td>
</tr>
<tr>
<td>C16</td>
<td>C</td>
<td>0.902128</td>
</tr>
<tr>
<td>C17</td>
<td>CA</td>
<td>-0.156519</td>
</tr>
<tr>
<td>C18</td>
<td>CA</td>
<td>0.163017</td>
</tr>
<tr>
<td>C19</td>
<td>CA</td>
<td>0.080847</td>
</tr>
<tr>
<td>C20</td>
<td>CA</td>
<td>-0.197572</td>
</tr>
<tr>
<td>C21</td>
<td>CA</td>
<td>0.143699</td>
</tr>
<tr>
<td>C22</td>
<td>CA</td>
<td>-0.115229</td>
</tr>
<tr>
<td>C23</td>
<td>CA</td>
<td>-0.739521</td>
</tr>
<tr>
<td>C24</td>
<td>CA</td>
<td>-0.696506</td>
</tr>
<tr>
<td>C25</td>
<td>CA</td>
<td>-0.602449</td>
</tr>
<tr>
<td>C26</td>
<td>CA</td>
<td>-0.537707</td>
</tr>
<tr>
<td>C27</td>
<td>CA</td>
<td>0.962772</td>
</tr>
<tr>
<td>C28</td>
<td>CA</td>
<td>0.920787</td>
</tr>
<tr>
<td>C29</td>
<td>C</td>
<td>0.875763</td>
</tr>
<tr>
<td>C30</td>
<td>C</td>
<td>0.641063</td>
</tr>
<tr>
<td>C31</td>
<td>CT</td>
<td>-0.554817</td>
</tr>
<tr>
<td>C32</td>
<td>CT</td>
<td>0.235200</td>
</tr>
<tr>
<td>Atom Name</td>
<td>Atom Type</td>
<td>ESP/Charge</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>H33</td>
<td>HC</td>
<td>0.006127</td>
</tr>
<tr>
<td>H34</td>
<td>HC</td>
<td>-0.021781</td>
</tr>
<tr>
<td>H35</td>
<td>HA</td>
<td>-0.001407</td>
</tr>
<tr>
<td>H36</td>
<td>HA</td>
<td>0.111620</td>
</tr>
<tr>
<td>H37</td>
<td>HA</td>
<td>0.182845</td>
</tr>
<tr>
<td>H38</td>
<td>HA</td>
<td>0.186000</td>
</tr>
<tr>
<td>H39</td>
<td>HA</td>
<td>0.144964</td>
</tr>
<tr>
<td>H40</td>
<td>HA</td>
<td>0.046773</td>
</tr>
<tr>
<td>H41</td>
<td>HA</td>
<td>0.012168</td>
</tr>
<tr>
<td>H42</td>
<td>HA</td>
<td>0.193139</td>
</tr>
<tr>
<td>H43</td>
<td>HA</td>
<td>0.161787</td>
</tr>
<tr>
<td>H44†</td>
<td>HC</td>
<td>0.201020</td>
</tr>
<tr>
<td>N45*</td>
<td>N</td>
<td>(-0.423900)</td>
</tr>
<tr>
<td>C46*</td>
<td>CT</td>
<td>(-0.054300)</td>
</tr>
<tr>
<td>H47*</td>
<td>H</td>
<td>(0.290100)</td>
</tr>
<tr>
<td>H48*</td>
<td>H1</td>
<td>(0.062700)</td>
</tr>
<tr>
<td>H49*</td>
<td>H1</td>
<td>(0.062700)</td>
</tr>
<tr>
<td>H50*</td>
<td>H1</td>
<td>(0.062700)</td>
</tr>
</tbody>
</table>

Values determined by the GAUSSIAN/09 package for later implementation into the MD simulations. The ESP charges of the F5M molecule set for attachment to A6 was calculated using a 6-atom capping group (-NHCH3), designated with (*) and assigned constant charge values (in parentheses). This cap is removed before the probe is incorporated with the coordinates of the A6 molecules. Covalent attachment of the F5M molecule to the sulfur atom of the cysteine residue reduces the double bond in the maleimide group, thus requiring an additional hydrogen atom, designated with (†).
Figure 2.8 Atomic positions of the F5M molecule. Numbering of atoms corresponds with the attached F5M molecule in Table 2.1.
covalently linked probe were incorporated into the A6 PDB 3HQ3 coordinates at five positions of significant interest: S100C (CDR3α), M27C (CDR1β), N28C (CDR1β), A99C (CDR3β) and R102C (CDR3β).

2.2.3 Analysis of MD trajectories

The trajectory files were analyzed for $\psi/\phi$ backbone dihedrals, atomic fluctuations, correlation functions, PDB average structures, interatomic distances, and DCCM values using the ptraj program in the AMBER10 suite. All trajectories were stripped of solvent before analysis for faster calculations, and for most analyses, required rms fitting to backbone atoms of a reference molecule (typically the pre-equilibrated structure, unless mentioned otherwise). Further details of the setup and rationale behind the ptraj analyses will be specified in the following chapters where appropriate.
CHAPTER 3:
CDR3 HYPERVARIABLE LOOPS POSSESS INDEPENDENT PRE-BINDING ENERGY LANDSCAPES THAT PROMOTE T CELL RECEPTOR BINDING AND MHC RESTRICTION

3.1 Summary

This chapter discusses findings from a recent published article\textsuperscript{18} which supports the idea of the A6 TCR using pre-existing conformations of its hypervariable CDR3 loops to establish cross-reactivity and MHC restriction with the HLA-A2 allele.

T cell activation against virally infected or aberrant cells requires the recognition of an antigenic peptide-MHC complex by the T cell receptor (TCR) protein. This protein-protein interaction has been exhaustively studied in both structural and thermodynamic capacities, with the data for each new TCR-pMHC system producing important, yet sometimes contradicting, evidence towards already-established TCR recognition mechanisms. If a universal binding strategy does exist, however, the evidence for this mechanism has yet to be unveiled, or at least recognized. One likely reason for this oversight is that insufficient attention has been paid to the site-specific dynamics of the proteins’ surfaces. Until this relevant aspect of the interaction is resolved, our understanding of TCR recognition will remain incomplete.
A stout library of TCR and pMHC structures, both bound and unbound, has been constructed in order to characterize the contact surfaces and structural adaptability of the proteins at the interface of the complex. Crystal structures, however, fall short in describing the level of conformational dynamics that each protein inherently possesses at their respective binding surfaces, so any conclusions drawn thus far regarding the contributions of pre-existing flexibility towards TCR recognition are largely speculative.

Crystal structures depict the “before and after” snapshots of the protein surfaces – critical pieces to the binding mechanism puzzle. Yet we assume that crystal structures encapsulate the only potential state of a protein, and subsequently fail to observe key dynamic properties of the system. Until recently, it had been almost a foregone conclusion that any structural rearrangements of the TCR or pMHC are the result of induced fit.\textsuperscript{23,30,41,71} This is almost assuredly an effect of static structures dominating our perspective of binding.

The conformational isomer ("conformer") model proposed by Holler and Kranz\textsuperscript{37} adds a different perspective to TCR binding, accounting for the receptor’s inherent dynamic properties. This model suggests that a TCR’s pre-existing equilibrium of states generates an exponentially greater antigen repertoire, as every conformer binds to unrelated ligands, while each one of those conformers can additionally bind to similar pMHCs via molecular mimicry. This expansion of antigenic specificity, i.e., cross-reactivity, by TCRs has been predicted by Mason\textsuperscript{72} and other colleagues as an essential means to adequately counter a large repertoire of antigen. By elucidating the pre-
existing dynamics of the receptor’s binding surface, we will reveal a more definitive TCR-pMHC binding mechanism applicable to multiple antigens.

Information regarding CDR loop dynamics has been elucidated for antibodies, which are structurally and functionally analogous to TCRs. Both molecular dynamic (MD) computations and nuclear magnetic resonance (NMR) experiments are common methodologies in antibody (Ab) studies, with attention mostly paid towards loop rigidification upon binding and maturation, loop stabilization via hydrogen bonding and hydrophobic side-chain burial, and antigen/Ab binding energetics. Much less focus has been on TCR loop dynamics, yet promising reports by Hare et al. and Michieclin et al. are paving the way for a greater understanding of TCR loop behavior.\textsuperscript{73,74} In addition, we have recently postulated that differential dynamics of the α2–helix of the MHC molecule govern the cross-reactivity of the αβ TCR A6 in its recognition of two structurally similar pMHC antigens.\textsuperscript{75} Both fluorescence anisotropy and molecular dynamic simulations were applied to perceive these differences in conformational mobility of the MHC molecule, and stand as ideal techniques to determine pre-existing TCR dynamics as well.

Structures for the A6 TCR in complex with the HTLV-1 viral Tax peptide, several variants of the Tax peptide, and the Tel1p and HuD peptides, all presented by the class I MHC molecule HLA-A*0201 (HLA-A2), have been solved.\textsuperscript{13,14,16,17,75,76} Both equilibrium and kinetic binding studies have also been performed on many of these ternary complexes.\textsuperscript{16,77,78} For a structural perspective of binding to its various ligands, the free A6 structure was solved.\textsuperscript{18} Superimposing the unligated A6 TCR with the
The aforementioned A6-pMHC complexes reveals translational shifts of the receptor’s complementarity determining regions (CDRs) across the interface, yet only the CDR3 loops undergo significant conformational rearrangements. Specifically, the CDR3α loop appears to take on one common conformation upon recognition to various pMHC ligands, while the binding solutions of the CDR3β loop are more randomly rearranged.

In this study, measured CDR3 loop dynamics via TRFA and MD simulations have helped us elucidate the means by which the A6 receptor may recognize more than one antigenic peptide. Interestingly, distinct flexibilities exist between these loops, where the dynamics of the CDR3α loop is much more restricted than that of CDR3β. The greater flexibility of the CDR3β loop allows the TCR to exist in a collection of specific, binding-competent states, as the binding conformation of CDR3β in the A6-Tax/HLA-A2 complex is clearly sampled in the MD simulations of free A6. Other Tax variants are also recognized using one of the many CDR3β’s pre-existing conformations, including the Tax variant Y5F peptide (LLFGEPYV). We surprisingly found that Tax Y5F binds to A6 with an almost identical binding free energy as wild-type Tax/HLA-A2, whereas other Tax variants have drastically lower free energies upon binding. Altogether, these diverse and complementary approaches of TCR-pMHC interactions and dynamics present a clearer picture of the A6 TCR’s binding strategy, and may lead to a better understanding of the TCR’s inherent traits of cross-reactivity and MHC restriction.
3.2 Results

3.2.1 The free A6 structure depicts rearrangements of the CDR3 loops in binding

When comparing the free A6 structure (PDB 3QH3) to the bound A6-pMHC complexes already solved, it is apparent that the CDR1 and CDR2 germline loops of A6 undergo little to no conformational rearrangement upon binding (Table 3.1 and Fig. 3.1). It should be pointed out, however, that modest whole-body loop shifts occur for several of these germline loops, especially CDR2α and CDR1β. The overall lack of conformational diversity by the backbone residues of the germline loops is typically seen in TCR-pMHC tertiary structures, and is consistent with an established two-step TCR binding mechanism that portrays the peripheral loops as being involved in a “lock and key type of binding” to the MHC molecule. In contrast, both CDR3 loops are observed to rearrange significantly after antigen recognition, also common to previously-studied bound and unbound TCR structures. The CDR3α loop regularly forms a conserved binding solution to various peptides, and in comparing to the free A6 structure, is the result of a large (~90°) hinge twist a few residues on either side of the loop apex, T93 and G102 (Fig. 3.2). The CDR3β loop, in contrast, has varying conformations among its binding states, which are themselves unique to the loop’s free conformation (Fig. 3.2). The CDR3 loops of the A6 TCR, therefore, are shown to possess binding-relevant structural adaptability, as opposed to the conformationally static germline loops.
TABLE 3.1

RMSD VALUES OF THE VARIOUS BOUND A6 CDR LOOPS FROM THEIR FREE POSITIONS

<table>
<thead>
<tr>
<th>Peptide Recognized by A6</th>
<th>Vα-Vβ Region</th>
<th>CDR1α</th>
<th>CDR2α</th>
<th>CDR3α</th>
<th>CDR1β</th>
<th>CDR2β</th>
<th>CDR3β</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tax</td>
<td>1.42</td>
<td>0.78</td>
<td>1.40</td>
<td>2.46</td>
<td>1.74</td>
<td>1.19</td>
<td>1.59</td>
</tr>
<tr>
<td>Tax P6A</td>
<td>1.45</td>
<td>0.91</td>
<td>1.53</td>
<td>2.49</td>
<td>1.90</td>
<td>1.14</td>
<td>1.63</td>
</tr>
<tr>
<td>Tax V7R</td>
<td>1.35</td>
<td>0.71</td>
<td>1.28</td>
<td>2.42</td>
<td>1.87</td>
<td>0.94</td>
<td>1.81</td>
</tr>
<tr>
<td>Tax Y8A</td>
<td>1.36</td>
<td>0.77</td>
<td>1.38</td>
<td>2.45</td>
<td>1.93</td>
<td>0.98</td>
<td>1.66</td>
</tr>
<tr>
<td>Tax Y5K-IBA</td>
<td>1.36</td>
<td>0.73</td>
<td>0.95</td>
<td>2.39</td>
<td>2.01</td>
<td>1.06</td>
<td>2.26</td>
</tr>
<tr>
<td>Tax Y5F</td>
<td>1.42</td>
<td>0.72</td>
<td>1.40</td>
<td>2.46</td>
<td>2.11</td>
<td>0.97</td>
<td>1.60</td>
</tr>
<tr>
<td>Tax Y5F&lt;sup&gt;F&lt;/sup&gt;</td>
<td>1.37</td>
<td>0.75</td>
<td>1.31</td>
<td>2.41</td>
<td>1.92</td>
<td>0.95</td>
<td>1.78</td>
</tr>
<tr>
<td>Tel1p</td>
<td>1.34</td>
<td>0.70</td>
<td>1.24</td>
<td>2.43</td>
<td>1.93</td>
<td>1.06</td>
<td>1.85</td>
</tr>
<tr>
<td>HuD</td>
<td>1.41</td>
<td>0.82</td>
<td>1.33</td>
<td>2.43</td>
<td>1.96</td>
<td>1.00</td>
<td>1.75</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td><strong>1.38</strong></td>
<td><strong>0.76</strong></td>
<td><strong>1.31</strong></td>
<td><strong>2.44</strong></td>
<td><strong>1.93</strong></td>
<td><strong>1.03</strong></td>
<td><strong>1.77</strong></td>
</tr>
<tr>
<td><strong>St. Dev.</strong></td>
<td><strong>0.04</strong></td>
<td><strong>0.07</strong></td>
<td><strong>0.16</strong></td>
<td><strong>0.03</strong></td>
<td><strong>0.10</strong></td>
<td><strong>0.09</strong></td>
<td><strong>0.21</strong></td>
</tr>
</tbody>
</table>

Values are reported in Å; averages and standard deviations across the bound structures are given at the bottom of the table. The TCRs were first fit to the backbones of the free A6 Vα-Vβ domains before calculations were performed.
Figure 3.1 **Free and bound positions of the antigen binding loops of A6.** View of the backbone conformations as perceived from below the antigen (the Tax peptide is shown in pink stick for orientation). All bound TCRs were superimposed with the free A6 TCR by their Vα-Vβ backbone atoms before the non-binding portions of the TCRs were removed for imaging. Visualized with PyMOL (DeLano Scientific LLC).
3.2.2 Time-resolved fluorescence anisotropy reveals disparate CDR3 loop dynamics

To assess the inherent backbone flexibility of the A6 CDR3 loops, we have performed time-resolved fluorescence anisotropy (TRFA) experiments using fluorescein-5-maleimide (F5M) as a probe. Site-directed single cysteine mutations were made within the sequences of the CDR3 loops to allow for the thiol-reactive probe to be linked at positions which were determined to have high solvent accessibility in the free A6 crystal structure. Two residues of the CDR3α loop (S100C and W101C) and four residues within the CDR3β loop (A99C, G100C, G101C, and R102C) were labeled to
comprehensively portray the overall dynamics of each loop at the pico-to-nanosecond timescale (Fig. 3.3).

The anisotropy results show that the CDR3β loop of A6 possesses a heightened level of conformational diversity relative to CDR3α. The consistently faster decay of the anisotropy curves of the four CDR3β positions depict the distinct dynamics between the two loops (Fig. 3.4), while fitting parameters of the curves quantitatively detail their inherent flexibilities (Table 3.2). The fast correlation times ($\theta_f$) are longer than 0.89 ns for the CDR3α positions, while the CDR3β correlation time values are as fast as 0.55 ns, and do not exceed 0.80 ns. This signifies a significantly faster rate of depolarization by the probes when attached to CDR3β residues – most attributable to greater local backbone flexibility. Additionally, the relative amplitudes of the fast decay process for the CDR3α positions range between 19-26%, compared to 38-52% for CDR3β, which is

![Diagram showing labeling sites of the A6 CDR3 loops for TRFA studies. Side chains are of the native residues; each residue is singly mutated to cysteine for linkage to the fluorescent probe.](image-url)

**Figure 3.3** *Labeling sites of the A6 CDR3 loops for TRFA studies.* Side chains are of the native residues; each residue is singly mutated to cysteine for linkage to the fluorescent probe.
Figure 3.4 Anisotropy curves of the A6 CDR3 loops. Data was normalized to each residue’s own maximum value, which is set as $t = 0$ in the plot. Only one of at least seven data sets for each residue is shown.

<table>
<thead>
<tr>
<th>Position</th>
<th>$\theta_f$ (ns)</th>
<th>$\theta_s$ (ns)</th>
<th>$f_f \times 100%$</th>
<th>$f_f/\theta_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDR3α</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S100</td>
<td>1.13±0.51</td>
<td>18.3±1.0</td>
<td>19.2±4.5</td>
<td>0.23±0.15</td>
</tr>
<tr>
<td>W101</td>
<td>0.89±0.34</td>
<td>17.5±1.5</td>
<td>26.3±3.2</td>
<td>0.35±0.17</td>
</tr>
<tr>
<td>W101 (BDY)</td>
<td>1.40±0.08</td>
<td>13.6±0.3</td>
<td>37.3±1.2</td>
<td>0.27±0.01</td>
</tr>
<tr>
<td>CDR3β</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A99</td>
<td>0.55±0.12</td>
<td>16.9±1.0</td>
<td>41.7±2.4</td>
<td>0.80±0.21</td>
</tr>
<tr>
<td>G100</td>
<td>0.79±0.26</td>
<td>15.0±1.2</td>
<td>39.4±2.4</td>
<td>0.56±0.20</td>
</tr>
<tr>
<td>G101</td>
<td>0.80±0.19</td>
<td>11.0±1.1</td>
<td>52.3±1.9</td>
<td>0.68±0.15</td>
</tr>
<tr>
<td>G101 (BDY)</td>
<td>1.12±0.02</td>
<td>14.6±0.3</td>
<td>54.5±0.1</td>
<td>0.49±0.01</td>
</tr>
<tr>
<td>R102</td>
<td>0.67±0.12</td>
<td>11.8±0.4</td>
<td>38.6±2.1</td>
<td>0.60±0.13</td>
</tr>
</tbody>
</table>
further evidence that these loops have contrasting levels of conformational freedom. These results ultimately indicate that the CDR3β loop possesses backbone flexibility observed at the faster time regimes of protein dynamics, potentially allowing the loop to sample conformational substates necessary for recognition of distinct antigenic peptides. In contrast, the CDR3α loop is more likely to achieve its routine structural rearrangement seen in A6-pMHC tertiary structures via slower dynamics (or strictly an induced fit rearrangement).

3.2.3 MD simulations of the free A6 TCR

Additional dynamic information was obtained from a total of 200 ns of MD simulations of the free A6 structure, making it possible to assess the backbone flexibility and conformational sampling of the CDR3 loops. First, the atomic fluctuations of backbone atoms within each residue were resolved as “B-factor” values. Fig. 3.5 shows values for the α carbons atoms of the Vα-Vβ domains of A6. Peaks represent segments of the A6 TCR which possess significantly higher levels of backbone motion. Ignoring the N- and C-terminus residues, the CDR3 backbone residues easily dominate the rest of the protein in terms of atomic disorder. Strikingly, the apex of the CDR3β b-factor peak is at least 30% greater than the peak of the CDR3α residues. These results are analogous to the time-resolved fluorescence data, which depicts the CDR3β loop as having a heightened level of backbone flexibility to the CDR3α loop. Because the MD trajectory length extends beyond the timescale for that of the fluorescence anisotropy measurements, the computational results are potentially even more sensitive for
elucidating conformational flexibility of the CDR3 loops.

We also considered the actual backbone conformations that the CDR3 residues sample during the simulations. Dihedral angles of selected residues were recorded over the course of the simulation, while the resulting φ/ψ plots depict the sampling of distinct conformations by the residue of interest. Our particular focus is on how well these conformations of the free A6 TCR match with dihedral angles of the same residues in A6-pMHC crystal structures. The results indicate that key residues of both CDR3 loops can access their bound conformations within the timescales of the MD simulations (Fig. 3.6 and 3.7). CDR3α’s T93 and G102, most notably, switch conformations that are
occupied within the bound crystal structures, controlling the hinge motion from free A6 to its conserved binding conformation. More significant sampling is performed by the CDR3β residues, almost all of which match remarkably well with their corresponding crystal structure positions. The ability to interconvert between free and bound states is a significant finding in terms of the receptor’s recognition mechanism, closely following the perspective detailed by the conformer model proposed by Holler and Kranz.37 Thus conformational diversity seems to be the key feature in A6’s cross-reactive nature.

Analysis of the MD simulations in the form of time-resolved RMSD calculations further depicts the bound-like sampling of the CDR3 loops, with CDR3α interconverting between free- and bound-like conformations at a much slower rate than CDR3β (Fig. 3.8). Of the five 40-ns MD segments, the occupancy of the bound-like state of CDR3α occurs just once (segment 3), whereas significant fluctuations between various conformations is evident in nearly all the segments for the CDR3β loop.

An additional 200-ns MD simulation was produced with the starting coordinates of A6 in its bound state (PDB 1QRN). This “reverse” simulation was analyzed with similar scrutiny as the “forward,” free MD simulation, and revealed that the CDR3β loop was able to revert back to free-like conformations, while the CDR3α loop remained locked in its bound state (Fig. 3.9). By starting A6 in both its free and bound structure, it clearly depicts the two hypervariable loops as possessing differential kinetic exchange
Figure 3.6 Psi/phi plots of key A6 CDR3α residues depicting conformational changes during simulations. Scattered plots represent the psi/phi values of the residue during each time step of the 200-ns simulations in black (recorded every 200 fs). Values of the psi/phi angles found in the A6 crystal structures are shown with the corresponding symbols in the legend.
Figure 3.7 **Psi/phi plots of key A6 CDR3β residues depicting conformational changes during simulations.** See caption for Fig. 3.6 for details.
Figure 3.8 RMSD calculations of the CDR3 loops in the 200-ns MD simulations of free A6. Backbone atoms of the CDR3α (top) and CDR3β (bottom) were superimposed with either the free, Tax-bound, or Tel1p-bound structure. Values were calculated every 200 fs.
Figure 3.9 RMSD calculations of the A6 CDR3 loops in the 200-ns “reverse” MD simulations of bound A6. See caption for Fig. 3.8 for details.
rates, and may indicate independent roles of the CDR3 loops in recognizing a diverse array of pMHC ligands.

3.2.4 Superimposition of the dynamic CDR3α loop onto HLA-A2

One of the most intriguing observations from the MD trajectory snapshots is the conformational space occupied by the CDR3α loop in relation to the superimposed position of the HLA-A2 α1 helix. Specific residues within this region of the HLA binding groove (R65 and K66) have been identified as an energetic “hot spot” in previous alanine mutation studies. The simulations predict, in comparison with the superimposed Tax-HLA-A2 structure, a major clash between CDR3α residues and the guanidinium group of the HLA R65 residue. The 100-fold decrease in $k_{on}$ observed for the mutation of R65 to alanine is a testament to how significant this residue is in catalyzing the A6-Tax/HLA complex formation. CDR3α’s rearrangement into its common binding state could be driven by its interaction with the side chain of R65 and other HLA-A2 residues, whose suboptimal shape complementarity with CDR3α appears to be bypassed by the loop’s inherent conformational diversity.

3.3 Discussion

3.3.1 Dynamic rates of CDR sampling imply a pre-existing equilibrium of binding

As evidenced in our studies, the more structurally promiscuous CDR3β loop is readily equipped with its pre-existing equilibrium of conformational substates to recognize multiple antigens. From an energy landscape perspective, the free energy
barriers separating these substates are relatively small (Fig. 3.10B). A much different level of conformational flexibility is observed for CDR3α, which demonstrates a slower and very structurally discriminate exchange between two substates. The relative populations of the two conformations of CDR3α cannot be characterized sufficiently by the 400ns sampling times of MD simulation, as there are not enough exchange events to statistically evaluate their relative stabilities. Because of the limited interconversion, however, it can be deduced that the energy barrier separating these CDR3α substates is high, preventing frequent exchange (Fig. 3.10A). With the CDR3α loop much more kinetically restricted, there appears to be a bimodal recognition mechanism at play, governed by loop flexibility. This binding strategy can be related to a two-step mechanism previously proposed, yet postulates for the first time that pre-existing flexibility may be a determinant factor in how these steps are regulated.

This investigation also reveals that structures of an unligated protein must not be viewed as its one true conformation. In this instance, such a misinterpretation could have lead to pre-emptive, incomplete conclusions regarding the protein’s structural rearrangements upon binding, ultimately confusing the binding mechanism that is our goal in elucidating.

3.3.2 CDR3 alpha loop maintains MHC specificity via conformational change

What controls the rate-limiting CDR3a conformational switch? In order to realize just how this conformation may be selected upon binding, we superimposed the MD snapshots with the bound A6-Tax ternary structure. Interestingly, we have found a
Figure 3.10 **Energy landscapes of the CDR3 loops.** A more discriminate two-state equilibrium exists for the CDR3α backbone (A), with a larger barrier separating the states than the more conformationally promiscuous CDR3β (B). Image taken from Scott *et al.*, *JMB*; 2011.
probable cause of the binding-competent state of CDR3α – the free-like backbone conformational space of the loop invades the structural space of the Tax/HLA ligand, most notably the Arg65 and Lys66 side chains of the MHC α1 helix. These clashes are even more striking than those seen in the superimpositions of the free/bound A6 structures. The Arg65/Lys66 residues have previously been identified as energetic hot spots for several TCR/pMHC ligand systems, and may have a large effect on the association of the TCR to the pMHC binding surface.

Altogether these analyses of CDR3 loop dynamics have laid groundwork on mapping a TCR-pMHC binding mechanism for the A6 TCR with its large array of antigens. Interpretation of the structural rearrangements of the CDR3 loops would have predicted the random (conserved) changes of the CDR3β (CDR3α) loops as exclusively peptide-scanning events. Integrating loop dynamics with the structural analysis has shown MHC-dependent interactions which may govern the CDR3 loop binding solutions, an argument that has previously been made strictly for the peripheral, germline-encoded CDR1 and CDR2 loops.

3.3.3 Future work

Notably, our focus has predominantly been on dynamic backbone conformations, so a better understanding of how side-chain dynamics affects overall recognition is not yet considered. Subsequent backbone adaptability and side-chain rearrangements to the pMHC surface are likely components of the multi-exponential kinetics often witnessed in protein-protein interactions. A closer analysis of side chain
dynamics and their correlation with the motions of the backbone loops would provide more precise evidence of the binding mechanism at play.

Additionally, the measurements of CDR3 loop dynamics presented in this study only extend outwards into the longer nanosecond timescale range, and are only the “tip of the iceberg” when elucidating binding-relevant timescales. Further mining of the MD simulations and additional TRFA measurements are necessary to determine the conformational mobility of the entire TCR binding surface, as well as site-specific dynamics away from the binding site that may communicate downstream signaling to induce the immune response (co-receptors, TCR crowding, etc).
CHAPTER 4:

MIMICKING PROTEIN/PROBE DYNAMICS USING MOLECULAR DYNAMICS SIMULATIONS

4.1 Summary

Protein-protein interactions are relied upon to manage a myriad of processes within and on the surfaces of cells, thus sustaining life for the organism. In particular, the success of our immune system is dependent on the recognition of invading pathogens and aberrant cells via receptor-ligand interactions. T cell receptors (TCRs) are membrane proteins which initiate an immune response against diseased cells upon the encounter of cell-associated antigenic peptide/major histocompatibility complex (pMHC) proteins. The TCR-pMHC interaction is inherently complex due to the recognition of threatening peptide components in the presence of a self-entity, the MHC molecule. Attempts to derive a binding model that portrays the deft precision of this recognition mechanism have sparked great debate over the years. One constant, however, in the various TCR-pMHC perspectives is that conformational flexibility of the TCR binding site plays a major role in recognition, and gives the TCR its hallmark properties of cross-reactivity towards many different antigenic peptides.

This chapter details the combination of experimental and computational studies of A6 TCR dynamics, which entail cysteine mutations and probe attachments into
various sites along the TCR binding interface for fluorescence depolarization measurements. The technique of time-resolved fluorescence anisotropy (TRFA) is an established biophysical method for depicting site-specific flexibility in macromolecules. With the aid of TRFA and supporting molecular dynamics (MD) simulations, we have recently shown that the pre-existing flexibility of two complementarity determining region (CDR) loops of the A6 TCR each independently contribute to cross-reactivity and specificity. This is accomplished via conformational diversity and rapid sampling by the CDR3β loop, while CDR3α adapts to the ligand with slower, more methodical loop dynamics.

Importantly, however, protein dynamics measured by time-resolved fluorescence anisotropy (TRFA) carries several assumptions: the local protein and the probe to which it is attached are concerted in motion, the probe does not significantly interact with neighboring protein segments, and the probe and/or cysteine mutation does not perturb the native dynamics of the protein segment. In an attempt to characterize the pre-existing flexibility of a T cell receptor protein's antigen-binding loops, TRFA measurements of 19 independently labeled sites of the A6 TCR were collected. To confirm the aforementioned assumptions regarding TRFA's viability, five of these labeled A6 molecules were computationally generated and described via molecular dynamic (MD) simulations, with three aims in mind: (1) to reproduce the experimental anisotropy curves, (2) to search for any structural evidence for possible protein-probe interactions that may affect the experimental data, and (3) to assess the
level of correlation between the motion of the probe with that of the labeling site’s backbone dynamics.

As expected, there was an extensive range of flexibilities across the A6 binding surface as detected by both TRFA and MD simulations. The time correlations of the probe that were calculated from the five simulated A6-F5M systems closely compared to that of the TRFA data, showing the simulations’ ability to duplicate the experimental system. In addition, the CDR loops which incurred the cysteine/probe perturbations predominantly reflected native-like conformational flexibility, with respect to the previously reported A6 simulations. Hence, the agreement between experiment and simulation seemed to cross-validate their success in depicting the A6 TCR’s site-specific flexibility.

A major discrepancy, however, was observed for several of the CDR 1 and 2 loops, which were reported by the TRFA measurements as possessing significantly high degrees of backbone dynamics. This had not been depicted by the B-factor calculations of the A6 backbone from any of the MD simulations. The contradiction between simulation and experiment warranted closer examination of the protein’s motion with that of the probe, to test the ability of the probe’s decorrelation to report actual backbone flexibility. Subsequently, dynamic cross-correlation matrixes (DCCM) values were calculated for the five simulated A6-F5M systems, which revealed that the probe was capable of losing much of its correlation despite relative rigidity of its attached protein segment. This was especially the case for one site labeled at the CDR1β loop –
from where most of the disagreement between TRFA and MD data stemmed. Thus, it has been shown that TRFA may produce “false positive” results in cases where the protein is conformationally static.

Conversely, evidence suggested that large conformational changes in the protein backbone can create opportunity for the attached probe to interact with neighboring protein segments, thus downplaying loop flexibility reported by the fluorescence depolarization. For one of the A6-F5M systems labeled at CDR3β, considerable probe effects were observed for extensive periods, thereby reducing the probe’s ability to lose correlation and properly reflect the larger CDR3β dynamics observed in the B-factor values.

Altogether this analysis of TRFA’s ability to resolve site-specific dynamics of the attached protein’s backbone has yielded questionable results. However, showing the ability to account for the divergence between TRFA results and actual backbone motion has actually re-confirmed that MD simulations are a powerful tool for determining protein flexibility with a remarkable level of accuracy. Thus, the capacity of the simulations to imitate real events creates great opportunity to translate knowledge regarding protein dynamics to understanding biological function.
4.2 Results

4.2.1 Fluorescence anisotropy reveals differential dynamics between the α and β chains of the TCR binding domain

To reveal the inherent flexibility of the CDR loop binding surface of the A6 TCR, the fluorescence depolarization of fluorescein-5-maleimide (F5M), covalently linked to a total of 19 single-cysteine mutants engineered throughout the A6 sequence, was measured. This data set includes the six mutants within the 3α and 3β loops measured in earlier work.\(^\text{18}\) The attached F5M reporter molecule has an average lifetime of approximately four nanoseconds, making it sensitive to the early nanosecond dynamics that are evidenced to contribute to TCR/pMHC recognition.

The anisotropy decays for each F5M-labeled A6 mutant were fit to a two-exponential decay model (see Methods), separating the slower time constant mainly associated with the global tumbling of the molecule (\(\theta_s\)) from the faster correlation time attributed to the local flexibility of the labeled protein segment (\(\theta_f\)). It is important here to acknowledge that slower-timescale flexibility of the local segment will be separated out with the global tumbling, making \(\theta_s\) a diluted representation of the actual rotational correlation time of the protein. Corresponding relative amplitudes \((f_f\) and \(f_s\)) indicate the fractional loss of anisotropy from the two respective decay processes. With the focus being on the local dynamics of the CDR loops, the ratio of \(f_f\) to \(\theta_f\) was adopted as a representation of loop flexibility, where a larger ratio depicts a more dynamic region (as a result of a larger amplitude and faster decay of the fluorescence).
The TRFA data spans all CDR loops, as well as the HV4α (Fig. 4.1), which in the unique case of the A6 TCR makes contact to its pMHC ligands. At least two cysteine mutants for each loop were engineered, and up to four mutants in some, having great success overall. Unfortunately, two mutants (Q30Cα and N52Cα) were unable to express despite multiple attempts, the latter of which leaves CDR2α as the only loop for which just one position has been measured. In addition, reference measurements were taken for residues distal to the CDR loops (at positions S19α and N120α) to put our TRFA measurements into context. S19α is a solvent-exposed residue within the TCR’s β-sheet framework, and was expected to produce flexibility values associated with a conformationally-restricted backbone region. In contrast, N120α is located at an unstructured segment of the polypeptide chain (within the linker region between the variable and constant domains), and was anticipated to report a standard value for greater backbone flexibility.

The results illustrate a distinct variability in the dynamics of the α-chain loops in comparison to those of the β-chain (Table 4.1). The 𝜃ᵣ values indicate that the CDR2α and CDR3α loops have the slowest decay processes (0.89-1.25 ns), as well as the smallest 𝑓ᵣ amplitudes (19-28%), which are significantly lower than the S19α “rigid” reference measurement (33%). Collectively, the 𝑓ᵣ/𝜃ᵣ ratios describe the CDR2α and CDR3α loops as more rigid than the remaining A6 loops. In contrast, the CDRβ loops have consistently faster 𝜃ᵣ values (0.52-0.80 ns), and in the case of the 1β and 2β loops, possess fractional amplitudes that either approach or exceed the “flexible” N120Cα
Figure 4.1 **Labeling sites of the A6 TCR.** This ribbon structure of A6 illustrates the 19 backbone sites where cysteine mutations were made (yellow segments) for fluorescence labeling, both in the CDR/HV loops, and the S19α and N120α reference sites away from the binding loops (adapted from PDB 3QH3).
<table>
<thead>
<tr>
<th>Mutant</th>
<th>$\theta_f$ (ns)</th>
<th>$\theta_s$ (ns)</th>
<th>$f_f$ (100%)</th>
<th>$f_f/\theta_f$</th>
<th>$r_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alpha Chain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative Reference</td>
<td>S19</td>
<td>0.50 ± 0.13</td>
<td>11.8 ± 0.9</td>
<td>32.8% ± 2.8%</td>
<td>0.70 ± 0.23</td>
</tr>
<tr>
<td>Positive Reference</td>
<td>N120</td>
<td>0.57 ± 0.06</td>
<td>9.6 ± 2.4</td>
<td>67.0% ± 1.4%</td>
<td>1.19 ± 0.12</td>
</tr>
<tr>
<td>CDR1α</td>
<td>D26</td>
<td>0.79 ± 0.19</td>
<td>17.5 ± 1.1</td>
<td>33.8% ± 1.8%</td>
<td>0.46 ± 0.14</td>
</tr>
<tr>
<td></td>
<td>R27</td>
<td>0.81 ± 0.10</td>
<td>11.6 ± 0.8</td>
<td>49.7% ± 4.8%</td>
<td>0.62 ± 0.07</td>
</tr>
<tr>
<td>CDR2α</td>
<td>S51</td>
<td>1.25 ± 0.13</td>
<td>17.5 ± 1.5</td>
<td>27.9% ± 1.5%</td>
<td>0.23 ± 0.03</td>
</tr>
<tr>
<td>HV4α</td>
<td>K68</td>
<td>0.88 ± 0.04</td>
<td>16.0 ± 0.7</td>
<td>53.6% ± 1.1%</td>
<td>0.61 ± 0.03</td>
</tr>
<tr>
<td></td>
<td>A69</td>
<td>0.74 ± 0.09</td>
<td>16.6 ± 1.3</td>
<td>41.6% ± 3.3%</td>
<td>0.57 ± 0.10</td>
</tr>
<tr>
<td></td>
<td>S70</td>
<td>0.83 ± 0.08</td>
<td>18.1 ± 5.8</td>
<td>71.4% ± 4.1%</td>
<td>0.87 ± 0.09</td>
</tr>
<tr>
<td>CDR3α</td>
<td>S100</td>
<td>1.13 ± 0.51</td>
<td>18.3 ± 1.0</td>
<td>19.2% ± 4.5%</td>
<td>0.23 ± 0.15</td>
</tr>
<tr>
<td></td>
<td>W101</td>
<td>0.89 ± 0.34</td>
<td>17.5 ± 1.5</td>
<td>26.3% ± 3.2%</td>
<td>0.35 ± 0.17</td>
</tr>
<tr>
<td><strong>Beta Chain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDR1β</td>
<td>D26</td>
<td>0.71 ± 0.08</td>
<td>9.3 ± 0.7</td>
<td>56.5% ± 1.5%</td>
<td>0.80 ± 0.10</td>
</tr>
<tr>
<td></td>
<td>M27</td>
<td>0.76 ± 0.12</td>
<td>10.4 ± 0.4</td>
<td>42.1% ± 2.0%</td>
<td>0.57 ± 0.11</td>
</tr>
<tr>
<td></td>
<td>N28</td>
<td>0.52 ± 0.05</td>
<td>6.7 ± 1.4</td>
<td>68.3% ± 1.6%</td>
<td>1.31 ± 0.11</td>
</tr>
<tr>
<td>CDR2β</td>
<td>A52</td>
<td>0.68 ± 0.10</td>
<td>14.2 ± 5.1</td>
<td>78.9% ± 0.8%</td>
<td>1.18 ± 0.16</td>
</tr>
<tr>
<td></td>
<td>I54</td>
<td>0.56 ± 0.03</td>
<td>13.3 ± 0.7</td>
<td>62.7% ± 0.8%</td>
<td>1.12 ± 0.07</td>
</tr>
<tr>
<td>CDR3β</td>
<td>A99</td>
<td>0.55 ± 0.12</td>
<td>16.9 ± 1.0</td>
<td>41.7% ± 2.4%</td>
<td>0.80 ± 0.21</td>
</tr>
<tr>
<td></td>
<td>G100</td>
<td>0.79 ± 0.26</td>
<td>15.0 ± 1.2</td>
<td>39.4% ± 2.4%</td>
<td>0.56 ± 0.20</td>
</tr>
<tr>
<td></td>
<td>G101</td>
<td>0.80 ± 0.19</td>
<td>11.0 ± 1.1</td>
<td>52.3% ± 1.9%</td>
<td>0.68 ± 0.15</td>
</tr>
<tr>
<td></td>
<td>R102</td>
<td>0.67 ± 0.12</td>
<td>11.8 ± 0.4</td>
<td>38.6% ± 2.1%</td>
<td>0.60 ± 0.13</td>
</tr>
</tbody>
</table>

A two-exponential decay fit was universally applied to acquire all data shown. Values represent the averages and standard deviations of at least seven independent measurements using the F5M probe. The faster of the two correlation times ($\theta_f$), along with its normalized amplitude ($f_f$) are used to derive an overall flexibility ratio value. The initial anisotropy ($r_0$) represents the amplitude of the overall fitted curve.
Finally, the 1α and HV4α loops possess moderate flexibility values, including several of their positions overlapping with the 1β and 3β ranges (Fig. 4.2).

Figure 4.2 **Flexibility ranges of A6 in terms of $f/\theta_f$ values.** Each position within the measured CDR/HV4 loops is denoted by a diamond with the residue number embedded. The values for the negative and positive reference sites of S19α and N120α are depicted as black and red vertical lines, respectively.

The broad distribution of flexibility observed throughout the pMHC-binding surface of A6 is equivalent to, or even exceed, the $f/\theta_f$ ranges observed in analogous studies of protein dynamics using TRFA. This is not surprising, given the diverse topology associated with the CDR loop composite surface. More peculiar is the wide
range of flexibility observed within an individual CDR loop. In the most extreme case of
CDR1β, the two adjacent positions of M27 and N28 have over a two-fold difference in
their fθ/fθ ratios. This was initially interpreted as a fundamental lack of synchronization
of the loops: one segment of the loop can move independently, and at different rates,
than others.

To decipher if probe effects play a role in the large variability observed within
measurements of the same CDR loop, measurements were repeated on 11 of the 19
positions using BODIPY-FL maleimide (BDY) as an alternate probe. BPY differs from F5M
in charge (neutral versus negatively-charged, respectively) and structure (more
rotational freedom for BDY in its linker arm), allowing for different physical effects to be
reported by each probe (Fig. 4.3). The 11 BDY-labeled mutants span across five
different CDR loops, and in using steady-state anisotropy, a very similar flexibility profile
was conserved for the loops when compared to the F5M-labeled mutants, with just one
exception. While four of the five CDR loop flexibility profiles modestly shift in the
direction of higher flexibility when BDY is used, the S70Cα position depicts the HV4α
loop flexibility as shifting in the opposite direction (Fig. 4.4). This probe-dependent
effect revealed for the S70Cα position, while isolated, warrants further investigation of
probe-protein interactions and the possible adverse effects they may have on the TRFA measurements. Despite this one anomaly, however, the overall high level of correlation between the F5M and BDY measurements ($R^2 = 82\%$ correlation; lowered to 69\% when the S70 measurement is included; Fig. 4.5) gives confidence in the fluorescence anisotropy measurements reporting relative flexibility of A6’s CDR binding surface.
Figure 4.4 **Steady-state anisotropy measurements of F5M- and BDY-labeled A6 CDR loops.** A total of 11 sites compared on five of the A6 binding loops, using two chemically and structurally diverse probes.

Figure 4.5 **Correlation analysis of the F5M and BDY measurements.**
4.2.2 Molecular dynamics verify native-like flexibility for the A6-F5M constructs

Previously, the backbone flexibility of the A6 TCR was characterized using molecular dynamics (MD) simulations, which from preliminary analysis, portrayed a much different dynamic landscape than the anisotropy measurements. Specifically, the 460 ns of total simulation time (260 ns of A6 starting with its unligated structure, and 200 ns starting in a bound conformation) of the A6 molecule illustrated both CDR3 loops, particularly the 3β loop, as by far the most flexible segments of the pMHC-binding surface of A6. The TRFA data, however, portrays CDR3β as equivalent in flexibility to the 1α and HV4α loops, and even surpassed by the 1β and 2β loops, which in the MD simulations demonstrate very limited backbone flexibility. The flexibility of CDR3α is also downplayed by the TRFA data, with only the one measurement of the 2α loop being more rigid. This undoubtedly complicates how we assess dynamics of the protein in terms of a recognition process.

In an effort to reconcile the discrepancies of the previous simulations from the TRFA measurements, additional MD simulations were performed, this time incorporating the cysteine mutant and F5M molecule into five various sites of the A6 TCR coordinates: S100α, M27β, N28β, A99β, and R102β. This allows for a direct comparison and in-depth scrutiny of the anisotropy measurements by (1) duplicating the fluorescence depolarization curves in the analogous form of time correlation functions, (2) investigating the probe-protein interactions that may influence the anisotropy results, and (3) correlating the motion of the probe with that of the
backbone to which it is attached. The atomic coordinates of the free A6 TCR (PDB 3QH3) were subsequently modified with coordinates of a cysteine residue and thiol-linked F5M molecule, and simulated for five independent 40-ns trajectories (200 ns of total simulation time per A6-F5M variant).

To verify that the presence of the cysteine mutation and probe do not cause non-native dynamics in any part of the protein, the atomic fluctuations (in the form of B-factors) of the backbone α carbons of the A6-F5M TCRs were calculated. The overall B-factor profiles for the α and β binding domains of all five A6-F5M mutants generally follow the trends set by the previous wild-type A6 simulations (Fig. 4.6). In addition, sampling of the backbone dihedral angles shows no clear evidence that the cysteine mutations and attached probes affect the structural integrity of the A6-F5M constructs (data not shown).

4.2.3 Correlation functions of the simulated F5M probe agree well with TRFA flexibility values of the A6-F5M mutants

Next the time-dependent decorrelation of the probe were examined, the results of which were compared to the fluorescence anisotropy curves reported by the TRFA experiments. In each of the A6-F5M constructs, a vector was defined in the F5M molecule as two carbon atoms within the 3-membered ring, mimicking the transition dipole of the fluorescence emission (Fig. 4.3; circled in green). A time correlation function was subsequently calculated for ten 20-ns segments of simulations for each A6-F5M mutant, with the average of the ten independently-obtained functions accepted as
Figure 4.6 Atomic fluctuations of A6-WT and A6-F5M MD simulations. (A) B-factors for the α carbons of the A6 Vα and Vβ domains. The five A6-F5M MD simulations are shown in color, whereas A6-WT simulations starting with unbound (black) and bound (grey) coordinates are shown for reference. (B-D) Close-up views of the CDR loops which were perturbed by cysteine-probe incorporation.
the best representation of the decorrelation of the F5M probe (Fig. 4.7). In addition, one 20-ns simulation was performed for an unbound F5M molecule to establish a decorrelation “baseline” for the probe’s native dynamics.

Figure 4.7 Time correlation functions of the F5M probe from MD simulations. The average of ten 20 ns segments is shown for each mutant in color, having been calculated from the decorrelation of a defined vector within the F5M molecule. An unbound F5M molecule was simulated for 20 ns for reference (black). The upper-right inset shows the 1 ns region that was used for fitting.

A two-exponential decay model provided the best χ² values and residuals around the ultrafast timescales of the probe’s decorrelation. The faster of the two time constants recovered by the fits for each A6-F5M mutant were extremely consistent (26-
37 ps), and fall near the same time regime as the 57 ps-time decay determined for the unbound F5M molecule allowed to rotate freely (Table 4.2). The faster time constant was therefore attributed to the protein-independent flexibility of the probe, and refer to this constant as \( \theta_p \) for the MD simulations. Notably, the ultrafast decorrelation of the probe is not obtained by the TRFA experiments due to resolution limits of the excitation/detection system.

### TABLE 4.2

CORRELATION TIMES AND AMPILITUDES OF THE F5M MOLECULE IN MD SIMULATIONS

<table>
<thead>
<tr>
<th>Probe Type</th>
<th>( \theta_p ) (ps)</th>
<th>( f_p )</th>
<th>( \theta_{loop} ) (ps)</th>
<th>( f_{loop} )</th>
<th>( C^\infty )</th>
<th>( f_{loop}/\theta_{loop} )</th>
<th>( f_p/\theta_p^\Delta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free F5M ( ^* )</td>
<td>56.6</td>
<td>0.995</td>
<td>--</td>
<td>--</td>
<td>0.005</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Free F5M ( ^\dagger )</td>
<td>10.6</td>
<td>0.678</td>
<td>--</td>
<td>--</td>
<td>0.322</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>A6 S100Ca-F5M</td>
<td>37.3</td>
<td>0.256</td>
<td>520</td>
<td>0.269</td>
<td>0.475</td>
<td>0.52</td>
<td>0.23</td>
</tr>
<tr>
<td>A6 M27Cβ-F5M</td>
<td>30.1</td>
<td>0.187</td>
<td>428</td>
<td>0.264</td>
<td>0.549</td>
<td>0.62</td>
<td>0.57</td>
</tr>
<tr>
<td>A6 N28Cβ-F5M</td>
<td>34.4</td>
<td>0.421</td>
<td>224</td>
<td>0.306</td>
<td>0.273</td>
<td>1.37</td>
<td>1.31</td>
</tr>
<tr>
<td>A6 A99Cβ-F5M</td>
<td>26.4</td>
<td>0.210</td>
<td>309</td>
<td>0.342</td>
<td>0.448</td>
<td>1.11</td>
<td>0.80</td>
</tr>
<tr>
<td>A6 R102Cβ-F5M</td>
<td>34.8</td>
<td>0.203</td>
<td>425</td>
<td>0.228</td>
<td>0.569</td>
<td>0.54</td>
<td>0.60</td>
</tr>
</tbody>
</table>

One- or two-exponential decay fits were the best representation of the averaged vector correlation functions produced by the free and attached probes, respectively. The correlation function of free probe was calculated both by allowing the probe to rotate freely (denoted by \( ^\dagger \)) or superimposed to the double-bonded carbon atoms of the maleimide group (denoted by \( ^* \)). The \( C^\infty \) value is an order parameter-equivalent of the correlation function. Finally, the ratio of the slower correlation time constant associated with protein dynamics (\( \theta_{loop} \)) and its relative amplitude (\( f_{loop} \)) are directly compared to the experimentally derived values (denoted by \( ^\Delta \)).

The slower time constant (\( \theta_{loop} \)) from the two-exponential fit is consequently a reflection of the dynamics of the protein (and certainly includes an unknown component of slower-timescale motions of the probe as well). As the A6 protein was disallowed from global, rotational tumbling (using the “rms” fitting command during
ptraj analysis), the protein’s local flexibility (e.g., the motion of the specifically labeled CDR loop) is suspected to be the most dominant contribution to the probe’s decorrelation. The derived $f_{\text{loop}}/\theta_{\text{loop}}$ values of the various A6-F5M mutants have nearly a 3-fold spread, demonstrating the sensitivity of the correlation functions in resolving dynamical differences of the labeled CDR loops. Strikingly, the time constants associated with the protein flexibility are unanimously consistent with experimental data across all five mutants (Table 4.2). The N28Cβ-F5M mutant possesses the fastest timescales of dynamics, with a $\theta_{\text{loop}} = 224$ ps, as well as the second largest associated amplitude (30.6%). The other labeled site of the CDR1β loop, M27Cβ-F5M, has a significantly slower $\theta_{\text{loop}}$ value (428 ps) and a smaller amplitude (26.4%), illustrating remarkable similarity to the wide range of flexibility reported by the TRFA data for this loop. The CDR3β loop mutants also adhere to the trend of the experimental data: A99Cβ is the more flexible of the 3β mutants ($f_{\text{loop}}/\theta_{\text{loop}} = 1.11$, compared to 0.54 for position R102Cβ), yet does not approach the flexibility of the N28Cβ mutant. Finally, the S100Cα-F5M simulation produced the slowest dynamics of probe motion as a function of protein flexibility ($\theta_{\text{loop}} = 520$ ps), although it was not as distinct from the other four positions as the TRFA data demonstrated. This is mostly a result of a significantly higher fractional amplitude (27%, opposed to the TRFA amplitude of 19%), whereas the other four A6-F5M constructs have a reduced amplitude value in the MD results. Here one must consider that in each technique, a different time constant is being separated from the segmental flexibility: for the TRFA measurements, rotational
tumbling is the other resolvable attribute of the probe’s decorrelation, while the MD simulations have the ultrafast probe component that must be separated. For these reasons, caution must be taken when judging the accuracy of the amplitude values.

Overall, the probes within the A6-F5M simulations appear to behave equivalently with those in the fluorescence anisotropy experiments (experimental versus theoretical flexibility values of the $f/\theta$ ratio have a $R^2 = 84\%$ correlation).

4.2.4 Backbone vector correlation functions illustrate probe-independent CDR loop dynamics for four of the five mutants

To more closely investigate the actual backbone dynamics of each CDR loop, a dynamic comparison of the N-H, Cα-Cβ, and C-O vectors for the five A6-F5M mutants was performed, side-by-side with those of the unlabeled A6-WT simulation. Again, time correlation functions of these vectors were fitted with an exponential decay model. However, as this analysis takes advantage of dynamic properties analogous to that of NMR spin relaxation methods, the $C_\infty$ values of the correlation function (the minimum “plateau”/asymptote of the curve) will be addressed here, equivalent to how order parameters ($S^2$ values) are calculated in the NMR relaxation experiments.

In Figure 4.8, similar dynamic profiles emanate from four of the five labeling sites in comparison to their respective A6-WT residue, with the A99Cβ-F5M mutant being the lone exception. This demonstrates that the dynamics of the backbone for four of the five sites do not appear to be affected by the presence of the probe or the cysteine mutation ($R^2 = 90\%$ correlation for these four sets of values; lowers to 53% with the
Figure 4.8 Derived order parameters of the backbone vectors in the A6-F5M and A6-WT MD simulations. For the five independent simulations of A6-F5M, the backbone flexibilities of the mutated cysteine are quantified by time correlation functions. N-H, Cα-Cβ, and C-O vectors are all included for the most comprehensive characterization of the backbone. The C∞ values are cross-referenced with values from the wild-type residues of the A6-WT simulation (darker bars).

addition of A99Cβ-F5M; Fig. 4.9), just as was concluded from the earlier B-factor calculations. A99Cβ-F5M, on the other hand, has significantly higher C∞ values for all three of its backbone vectors with the presence of the probe/cysteine perturbations. This reduction in disorder (C∞ is inversely proportional to flexibility, just as for order parameters) may be due to the added bulk of the cysteine’s sulfhydryl group, and of course, the F5M molecule, restricting the native motion of the A99 backbone. In contrast, the other four labeled residues possess larger native side chains that more resemble the added bulkiness of the cysteine/probe perturbation. Additionally, the A6-
Figure 4.9 Correlation analysis of the A6-WT and A6-F5M order parameters.

WT simulation reveals that the C-O vector of A99 is the most dynamic backbone vector elucidated in this study, alluding to the likelihood that greater backbone motion creates more opportunity for probe-protein interactions. This is observed within the 41-80 ns-segment of the A6 A99Cβ-F5M trajectory, where the F5M probe is sandwiched between two adjacent loops for an extended period of time (Fig. 4.10). This will obviously have an effect on the native flexibility of the A99 residue, and its associated CDR3β loop, as the probe acts like a “ball and chain” to the covalently-linked A99/CDR3β. If this is occurring in the experimental realm as well, then the flexibility of the site is probably being under-estimated by the TRFA data.

In contrast, the probe attached to the N28Cβ position is almost completely void of protein clashes, consequently giving the probe a larger amplitude of decorrelation ($f_\rho$,
Figure 4.10 **A6 A99Cβ-F5M simulation depicting probe-protein interaction.** Each image represents a 2 ns snapshot of the average coordinates during a 40 ns segment of the simulation. The dynamics of the F5M molecule (stick) attached to the CDR3β loop (grey) at position 99 is inhibited by interactions with the CDR1α (green) and CDR3α (orange) loops.

= 42%) than the probe of the other A6-F5M mutants (19-26%; Table 4.2). This very likely explains the high degree of dynamics indirectly attributed to the backbone of the CDR1β loop. Hence, the probe’s decorrelation as measured via fluorescence anisotropy (and also by the simulated F5M vector) may be characterizing solvent accessibility and/or regional protein surface dynamics more significantly than discussed or presumed in the TRFA literature. Going back to the free F5M simulation, the rotation of the probe was restrained by superimposing the sulfhydryl-proximal atoms of the maleimide functional group, after which the time correlation functions were re-calculated. As suspected, the
function looked very similar to that of the A6N28Cβ-F5M probe, suggesting that the majority of the probe’s decorrelation was due to protein-independent motion.

To quantify actual correlation of motion between the probe and the covalently-linked residue’s backbone, dynamic cross-correlation matrices (DCCM) were calculated for the α carbons of the CDR loop with the carbons of the 3-membered aromatic ring of the F5M molecule (as this is the location of the transition dipole of the fluorescence). From Fig. 4.11, maximum correlation of the probe’s ring motion with the mutated cysteine residue was observed for most of the A6-F5M mutants, with the only exception being R102Cβ (yet its highest correlation is still with residues of the 3β loop). This is a good indication that the fluorescence decay of the probe is most directly reporting the backbone mobility of the loop to which it is attached. However, the level of correlation is seemingly weak (14.6-48.0% range for the five probe sites relative to the cysteine’s Cα-Cβ vector). Due to the elongated separation of the probe and backbone vectors, and the subsequent degrees of freedom throughout the probe’s linker arm (in addition to the thiol and Cβ-S bonds), a great deal of correlation is lost between the motion of the probe and the attached backbone. The detection of site-specific protein flexibility, therefore, does not appear to be the lone cause of probe decorrelation.
Figure 4.11 Correlated motion of the F5M molecule with the backbone of the attached loop. Vectors within F5M were cross-examined with Ca-Cβ atoms for the five A6-F5M simulations; all 200 ns were included. Cross-correlation values range from -1 (anti-correlated motion; blue) to 1 (correlated motion; red). The residues of the loop which were included in the analysis are above its respective value, including the CYS residue to which the probe is directly attached.

4.3 Discussion

4.3.1 Cross-validation of TRFA and computational methods for resolving site-specific protein dynamics

Critical information for protein dynamics can be elucidated by fluorescence spectroscopy and computational methods, but a combination of both produces a much clearer depiction of the role conformational flexibility has on the protein’s binding mechanism. Here, it was determined that the A6 TCR protein possesses varying pre-existing flexibility across its numerous antigen-binding loops. In the process, the
viability of MD simulations in duplicating experimentally-derived values related to probe-protein dynamics was highlighted. Relative to the TRFA data, simulations have depicted site-specific mobility of the A6’s TCR loops with seemingly more accuracy. In addition, conformational aspects of protein mobility that can be elucidated from the MD simulations (ψ/φ torsion angles, RMS deviations, etc.) significantly enhance the interpretations of binding-relevant dynamics. Implications towards a preliminary TCR binding mechanism involving the hypervariable loops have been revealed in the analysis of MD simulations. Here, the roles of backbone conformational diversity and surface-sensitive dynamics throughout the entire A6 binding domain were further revealed.

4.3.2 Re-examining the performance of fluorescence anisotropy

Given the significant disconnect between the motion of the probe and the backbone to which it is attached, where does this leave TRFA as a reliable technique for depicting protein dynamics? In the case of the TCR binding surface, which is a composite of many CDR loops of various lengths and side chain chemistries, it seems probable that a probe will occasionally interact with surrounding protein components. The non-native size and degrees of freedom that a probe possesses is likely to increase this probability. Thus, the influence of the probe’s surrounding environment needs to be taken into consideration, both when choosing a labeling site, as well as when analyzing TRFA results.

A similar investigation has been reported in which simulations were used to test the accuracy of fluorescence anisotropy, although it was not near to the depth of this
study. Briefly, it was found that for a labeled AB loop of bacteriorhodopsin, the attached probe had reflected dynamics of two side chains proximal to the attached site. One of these residues was adjacent to the labeled cysteine and possessed the highest RMS fluctuations of the AB loop, illustrating that a probe can report on dynamics of the most flexible protein segment, not necessarily that of just the attached residue. Remarkably, the other residue was even further removed sequentially from the labeled site, but the dynamics of the lysine’s side chain were clearly correlated due to its proximity to the probe’s range of motion. Anisotropy measurements encompassing such a large spread of residues within a protein segment were occasionally observed for the A6-F5M systems of this study (see cross-correlation values in Fig. 4.11). Therefore, the dynamics of the A6 TCR by fluorescence anisotropy must be closely scrutinized before claiming site-specific flexibility. It may be reasonable, therefore, to assign reported flexibility by TRFA data to the overall loop dynamics, and possibly even to the surface area to which the probe can reach.

In the specific case of the germline Vβ loops (CDR1 and CDR2), whose various labeling sites collectively reported greater flexibility than all other CDR loops, the F5M probe could be picking up on greater dynamics of its surroundings (such as the nearby CDR3β loop, shown to possess the greatest backbone dynamics of all loops). Alternatively, the probe may simply possess a greater solvent accessibility (SA), as was depicted in the A6 N28Cβ-F5M simulations. If the former is the case, then the β chain would appear to be more dynamic overall than the recognition surface of the α chain. If
the latter is the case, it could be argued that increased solvent accessibility observed for the probe would equate to a higher SA for the side chain being represented by the probe. This could lead to more promiscuous conformational sampling by the side chain in the native TCR, thus influencing its capabilities of binding. Here lies another added dimension that MD simulations can detail that most experimental techniques cannot – native side chain mobility at nanosecond timescales (although this more specific aspect of TCR dynamics is not addressed in this study).

4.3.3 Implications towards an A6 TCR binding mechanism

Structural analysis of A6/HLA-A2 with various peptides illustrates that the TCR has more evenly distributed contacts by its alpha chain (each CDRα loop has at least 3 contact points to the pMHC) than its beta chain, which is clearly dominated by CDR3β contacts to the ligand. The TRFA measurements of the CDR loops reveal that the alpha loops, overall, are more conformationally constrained and/or less solvent-exposed (depending on the interpretation of the data), than the beta loops. These data could allude to an unorthodox molecular scanning mechanism, divergent from the conventional scheme which depicts the evolutionarily-developed CDR1 and CDR2 germline loops of both chains making rigid-body association to the pMHC platform. The CDR3 loops would then be optimally oriented in order to contact the peptide within the MHC binding groove, enabling the loops to “scan” the peptide, possibly involving loop conformational adaptability to accommodate distinct ligands.
This investigation may reveal an alternative strategy specific for the A6 TCR, where the binding surface of the α chain performs the docking event, whereas the β chain scans its complementary half of the pMHC binding surface. The more stabilized CDR loops of the alpha chain would significantly contribute to a more energetically favorable binding solution, as the entropic loss upon binding would be minimized since the loops are already conformationally fixed. The more dynamic beta chain surface may contribute to the cross-reactivity of the TCR as it can conform to multiple antigenic pMHC. To illustrate this point, it is helpful to refer to the cross-recognition of Tax and Tel1p antigens by the A6 TCR, where the C-terminal region of the pMHC diverges in structure between Tax and Tel1p. This region of each pMHC is complementary to the beta chain CDR loops of A6. It could be speculated that the enhanced flexibility of the A6 beta loops is essential for the TCR to cross-react with the dynamically diverse Tax and Tel1p antigens.

Unfortunately, the uncertainty in interpreting the TRFA measurements leaves the discussion of A6’s binding mechanism to remain open-ended. While still providing insight into the nanosecond dynamics of the TCR binding surface as a whole, it is difficult to specify the locational cause of the differential dynamics observed by TRFA alone. The accuracy of the MD simulations presented here, however, has created momentum towards determining a more complete recognition scheme for A6. Future work will require cross-analysis with other experimental data that places the dynamic properties of the protein in a more energetically-relevant context to binding.
CHAPTER 5:
CORRELATED MOTIONS BETWEEN A6 AND TAX/HLA-A2 VIA MD SIMULATIONS – THE
INFLUENCE OF DYNAMICS ON THE ENTROPY OF BINDING

5.1 Summary

Selective recognition of foreign antigen by T-cell receptor (TCR) proteins is crucial for an appropriate and timely immune response. T cell-mediated immunity is dependent upon the interaction of the TCR protein with a targeted cell’s major histocompatibility complex (MHC) and its bound peptide. While a TCR must be specific in its recognition of antigenic peptide-MHC (pMHC), it must also be able to engage and respond to multiple ligands. This is necessary for the T cell repertoire to adequately patrol the organism’s periphery against an almost limitless number of potential antigens. The paradoxical traits of TCR specificity and cross-reactivity allude to a complex binding mechanism that has yet to be resolved.

Currently, a wide range of molecular recognition models have been proposed for TCR cross-reactivity, predominantly supported by structural, thermodynamic, and biophysical data. Unfortunately, a key component is missing in these emerging models: the dynamics of the binding surfaces of the TCR and pMHC. Conformational plasticity of the complementarity determining region (CDR) loops of the TCR has often been
suggested as a mean for cross-reactivity, but interpreting the dynamical behavior of the TCR from static crystal structures or global thermodynamic data is presumptuous at best.

The αβTCR A6 recognizes the HTLV-1 Tax viral peptide bound to the class I MHC molecule HLA-A*0201 (Tax/HLA-A2), along with two structural mimics of Tax: Tel1p/HLA-A2 (peptide from Saccharomyces cerevisiae) and HuD/HLA-A2 (a self-pMHC). The receptor’s proven ability to cross-react with multiple ligands, ten of which are depicted in high-resolution crystal structures and quantified by numerous binding studies, justifies the analysis of this receptor’s dynamic behavior. One study in particular – double mutant cycles conducted throughout the A6-Tax/HLA-A2 interface – has resolved the energetic contributions of individual side-chain/side-chain interactions to the overall binding affinity of this complex (Piepenbrink, K. H. et al., manuscript in progress). These studies have added an important dimension in defining contacts that can normally only be inferred when analyzing the complex’s crystal structure.

Complementary to the double mutant cycles are computational studies of the A6-Tax/HLA-A2 complex, which through steered molecular dynamics, report binding contributions of CDR residues at atomistic levels. Detailing the level of flexibility for each individual CDR loop, in the context of their “energetic favorability” to the recognition of Tax/HLA-A2, can elucidate important aspects of the overall binding strategy of the A6 TCR.
We have recently shown that the varying flexibility of the two CDR3 loops of A6 each independently contribute to cross-reactivity and specificity.\textsuperscript{18} This was established by means of both computational and fluorescence anisotropy techniques. To expand our investigation throughout the entire antigen-binding surface, we calculated order parameters from backbone vector correlation functions processed from MD simulations of the A6 TCR, both in its free state (using the coordinates from PDB 3QH3) and in complex with its well-characterized Tax/HLA-A2 ligand (PDB 1QRN; Fig. 5.1).

MD simulations of the A6 system have already shown their accuracy in depicting conformational dynamics at multiple levels. First, a simulation of the free A6 TCR previously predicted bound-like conformations of the CDR loops that are observed in crystal structure complexes of the A6 TCR, and conversely, one of the CDR loops in its bound structure partially reverted back to its free conformation.\textsuperscript{18} The ability to sample real states of the TCR, in both “forward” and “reverse” pathways, without any biasing towards these conformations is a testament to the robustness of these simulation methods. In addition, the same simulations of free A6 produced remarkable consistency with the results of the fluorescence anisotropy measurements, depicting the backbone of the CDR3α loop as much more dynamically restricted than CDR3β (Chapter 4). Finally, our confidence in the vector correlation function analysis we apply here has been solidified by the close agreement between probe vectors simulated and the corresponding experimental correlation times obtained by fluorescence anisotropy.
Figure 5.1 The A6-Tax/HLA-A2 complex. The five polypeptide chains of the complex are labeled, with the A6 αβTCR heterodimer oriented at the top. Shown here is the complex of A6 with the altered peptide ligand Tax P6A, associated with the HLA-A2 molecule and accessory β2M (PDB 1QRN; Ding et al., Immunity; 1999). The simulation was performed with the coordinates of this structure, only the Tax residue of A6 was mutated back to the native P6. Visualized with PyMOL (DeLano Scientific LLC).
By assessing the pre-existing flexibilities of the CDR loops, and then depicting their organization upon binding, we will have a novel opportunity to address the contributions of dynamic effects (conformational entropy, shape complementarity, peptide scanning, etc) to the recognition mechanism of A6. Our results show that the intrinsically disordered CDR3β loop maintains a level of flexibility in the complex, despite it being buried at the center of the interface. It deviates from its original conformation more than all other CDR loops, with evidence from ψ/φ plots that some of its key residues revert back to its free state. This would seem to minimize the entropic costs upon binding, while providing evidence that the loop does not contribute towards specificity via static interactions. On the other hand, the CDR3α stays locked into a very restricted conformation (one that is conserved throughout all ten of A6’s bound structures), maintaining contacts to the HLA-A2 surface that have been shown to provide energetically favorable interactions (Piepenbrink, K. H., et al., manuscript in progress).

The remaining CDR loops show more subtle contributions to the recognition mechanism, but could steer the interaction to a final complex. For example, several residues within the CDR1 and CDR2 loops are shown to maintain their moderate backbone flexibility with respect to their order parameters in the free and bound A6 simulations. This indicates that the germline loops of the A6 TCR contribute to pMHC binding by conserving conformational entropy. Thus the “glue” we speak of for an interaction does not necessarily result from a static interaction, but a give-and-take of
enthalpic and entropic contributions, as well as possibly an avoidance of “bad” interactions. Overall, these results call into question the conventional perceptions of the TCR-pMHC interaction. Future work will reveal if other TCRs use a similar strategy to A6 with respect to their loops’ changes in flexibility upon binding.

5.2 Results

5.2.1 Correlation values produced from the free A6 MD simulations are consistent with the previously-described dynamics of the CDR loops

To portray the backbone flexibility of the CDR loops in an alternative capacity, we calculated the vector correlation functions of the amide, Cα-Cβ, and carbonyl bonds of prominent residues within the CDR loops of the 200-ns, unbound A6 MD simulation. The amplitude of these functions (denoted as “C∞”) when fit to a multi-exponential decay model are equivalent to NMR-derived order parameters, and will hereafter be referred to as such. After averaging the sets of C∞ values for each residue (Table 5.1; individual values for each backbone vector can be found in Fig. 5.2), our findings reiterated that CDR3β was by far the most flexible segment of the pMHC-binding surface of A6. Particularly the glycines of the CDR3β loop portray remarkably low order parameter values (0.328 ± 0.008 for G100β; 0.384 ± 0.109 for G101β) relative to the remaining CDR loop residues (mean/stdev of 0.825 ± 0.049; next-lowest value of 0.678 ± 0.074 for G53 of the CDR2β loop). This differentiation in conformational diversity, whereby just one of the seven pMHC-binding loops possesses a collectively high level of
TABLE 5.1
AVERAGE ORDER PARAMETERS OF THE A6 MD SIMULATIONS

<table>
<thead>
<tr>
<th>Residue</th>
<th>Alpha Chain</th>
<th>200-ns Free A6 MD Simulations</th>
<th>50-ns Bound A6-Tax MD Simulations</th>
<th>( \Delta C_\infty ) (Free - Bound)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Avg. ( C_\infty )</td>
<td>St. Dev.</td>
<td>Avg. ( C_\infty )</td>
</tr>
<tr>
<td>CDR1( \alpha )</td>
<td></td>
<td>0.825</td>
<td>0.049</td>
<td>0.831</td>
</tr>
<tr>
<td></td>
<td>D26</td>
<td>0.841</td>
<td>0.023</td>
<td>0.865</td>
</tr>
<tr>
<td></td>
<td>R27</td>
<td>0.823</td>
<td>0.017</td>
<td>0.866</td>
</tr>
<tr>
<td></td>
<td>G28</td>
<td>0.847</td>
<td>0.004</td>
<td>0.889</td>
</tr>
<tr>
<td></td>
<td>Q30</td>
<td>0.895</td>
<td>0.052</td>
<td>0.891</td>
</tr>
<tr>
<td>CDR2( \alpha )</td>
<td></td>
<td>0.873</td>
<td>0.057</td>
<td>0.892</td>
</tr>
<tr>
<td></td>
<td>Y50</td>
<td>0.841</td>
<td>0.049</td>
<td>0.854</td>
</tr>
<tr>
<td></td>
<td>S51</td>
<td>0.809</td>
<td>0.121</td>
<td>0.799</td>
</tr>
<tr>
<td>HV4( \alpha )</td>
<td></td>
<td>0.864</td>
<td>0.045</td>
<td>0.808</td>
</tr>
<tr>
<td></td>
<td>K68</td>
<td>0.839</td>
<td>0.039</td>
<td>0.804</td>
</tr>
<tr>
<td></td>
<td>A69</td>
<td>0.866</td>
<td>0.048</td>
<td>0.860</td>
</tr>
<tr>
<td>CDR3( \alpha )</td>
<td></td>
<td>0.907</td>
<td>0.023</td>
<td>0.909</td>
</tr>
<tr>
<td></td>
<td>T93</td>
<td>0.866</td>
<td>0.039</td>
<td>0.881</td>
</tr>
<tr>
<td></td>
<td>T98</td>
<td>0.837</td>
<td>0.036</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td>D99</td>
<td>0.795</td>
<td>0.012</td>
<td>0.921</td>
</tr>
<tr>
<td></td>
<td>S100</td>
<td>0.786</td>
<td>0.058</td>
<td>0.923</td>
</tr>
<tr>
<td></td>
<td>W101</td>
<td>0.751</td>
<td>0.041</td>
<td>0.878</td>
</tr>
</tbody>
</table>
The average $C_{\infty}$ values for the N-H, Cα-Cβ, and C-O backbone vectors were calculated for key residues within the A6 TCR. Values for both the free A6 and bound A6-Tax simulations are given, along with a difference value ($\Delta C_{\infty}$), which will be positive for a more disordered backbone upon binding, whereas backbone “stiffening” will produce a negative value.

disorder, initially implies a recognition mechanism of molecular scanning by the hypervariable CDR3β loop of the A6 TCR.

However, there are moderate flexibilities scattered amongst the remaining CDR and HV4α loops, with “moderate” being defined as the average order parameters between 0.6 and 0.8. Values higher than 0.8 will be considered negligible in terms of possessing conformational diversity, a common interpretation of order parameters.
Figure 5.2 Order parameters of the 200-ns free A6 MD simulations. Values of each backbone vector for selected residues in the α (A) and β (B) chains of A6.
based on similar studies (Tjandra, Copie, Case). In particular, N52α, K68α, W101α, G102α, and E30β each contain at least one backbone vector which possess C∞ values that are less than 0.75. Notably, these residues are depicted as engaging the peptide and/or the MHC molecule in the vast majority of A6’s co-crystal structures, as well as shown to contribute highly favorable interaction energies with the A6-Tax/HLA-A2 ligand (Piepenbrink, K. H, et al., manuscript in progress). Thus, the modest mobilities of individual residues within the germline and CDR3α loops, while not as extreme as that possessed by CDR3β, may provide evidence for the receptor’s ability to cross-react with its pMHC ligands.

5.2.2 The bound A6 MD simulation reveals an astonishingly low amount of entropically-costly interactions with Tax/HLA-A2 ligand

As we wish to more closely investigate the backbone dynamics of each CDR loop after binding to its native Tax/HLA-A2 ligand, we next calculated the order parameters from a 50-ns MD simulation of the A6 TCR and pMHC in complex (Table 5.1 and Fig. 5.3). This will allow for an evaluation of the degree of organization upon binding of the A6 TCR loops.

From previous examination of the free A6 molecule, we witnessed the rare sampling of the CDR3α loop to a bound-like conformation. This pre-existing, two-state equilibrium is picked up by the order parameters in the loops’ C-terminal residues of S100-G102. It is apparent these residues are significantly restricted in comparison to their moderate flexibility in the unbound state (ΔC∞ values of -0.13, -0.14, -0.13 for the
Figure 5.3 Order parameters of the 50-ns bound A6 MD simulations. Values of each backbone vector for selected residues in the α (A) and β (B) chains of A6.
three respective residues). A large reduction of the CDR3α loop’s average RMS deviation from its starting coordinates (0.44 ± 0.09 Å, a -52% relative change from the free state simulations; Fig. 5.4), as well as obvious restrictions in the dihedral angle sampling of the individual residues’ backbones (Fig. 5.5; left), support this claim. This depicts the backbone of CDR3α as becoming extremely organized upon binding, potentially causing entropic penalties with compensating enthalpic binding energies.

Figure 5.4 **Comparing the free and bound RMSD values of the A6 binding loops.** Values are the average deviation of the loop’s backbone atoms in reference to the starting coordinates of the free A6 simulations (blue) and the bound A6-Tax simulations (red). Error bars represent the standard deviation of the RMSD values intermittently calculated from the simulations (every 200 fs for the free A6 simulations; every 1 ps for the bound A6-Tax simulations).
In comparison, the CDR3β loop, while losing much of its disorder after binding ($\Delta C_\infty \geq -0.18$ for each of the A99-R102 apex residues), still maintains a high enough level of backbone flexibility for several of its pMHC-interacting residues (G100, G101, and R102) to revert back to their unbound dihedral angles (Fig. 5.5; right). The loop’s divergence from the bound state is also clearly evident from its RMS deviation ($1.14 \pm 0.11$ Å – only a -30% relative change from the free A6 simulation). Remarkably, this value is even larger than its RMSD when superimposed to its free structure ($0.84 \pm 0.08$ Å; Fig. 5.6), revealing that the conformational sampling of the loop, in the presence of ligand, is capable of resembling the unbound state. Hence, the backbone of this A99-R102 segment of the CDR3β, in contrast to its CDR3α counterpart, is less restricted by the Tax/HLA-A2 ligand surface.

As mentioned earlier, CDR2α’s N52 and HV4α’s K68 residues are pMHC-binding residues according to the A6-Tax/HLA-A2 structure, and are shown to have moderate backbone flexibility in A6’s free state. Upon binding, their carbonyl backbones slightly increase in disorder ($\Delta C_\infty = +0.04$ and $\Delta C_\infty = +0.13$ for the respective residues). These residues appear to contribute to the conformational variability for their respective loops, as observed in the calculated RMS deviations (each loop having a +32 and +34% relative change in disorder, respectively). The CDR1α backbone also shows an average increase in RMS deviation (+27% relative change) when comparing the bound to the free simulations. Surprisingly, none of the individual vectors of the residues within this
Figure 5.5 *Psi/phi plots of key CDR3 residues in A6-Tax simulations.* Conformational diversity in the CDR loops are depicted by backbone dihedral angles of CDR3α (left-hand side) and CDR3β (right-hand side) residues. For reference, dihedral angles are depicted for the starting bound structure (◊), the native structure (Δ), and the free A6 structure (*).
loop increase in disorder upon binding according to the order parameters, so the source of this RMSD increase is admittedly puzzling. This could be an extreme instance where the minimization/equilibration steps force a change in conformation before MD production starts, which probably occurs in all the CDR loops to some degree.

Now focusing on peptide-contacting residues of A6, we observed E30β also retains its moderate backbone flexibility, with its amide vector showing undiminished...
disorder upon binding ($\Delta C_\infty = +0.07$). This CDR1β residue is the only β-chain germline residue of A6 that makes contact to the Tax/HLA-A2 ligand (to the pY8 hydroxyl group), yet like several of the pMHC-contacting residues of A6, appears to lose minimal conformational entropy upon binding. The crystal structure of this complex serves as evidence of a mobile E30β residue, as the side chain’s electron density is absent (PDB A1O7). S31α and R95β, on the other hand, are residues which possess negligible inherent backbone flexibility, and are both observed to make intimate contact with the peptide’s Y5 residue. These instances of apparent local rigid-body association by a specific residue of the CDR1α and CDR3β loops could potentially serve as an initial docking event where no entropic costs have to be paid.

Overall, these data indicate an entropically-favorable binding mechanism by the A6 TCR, where the germline and CDR3β loops can be hardly depicted as rigidifying when engaged to the pMHC ligand.

5.2.3 Distances and correlation of motion between contact atoms confirms CDR3α’s mode of MHC restriction

To deduce the level of dynamic communication between “contact” atoms in the TCR/pMHC complex, here we derived the time-dependent distances of intermolecular atoms (Fig. 5.7), defined both structurally (from the Tax_{WT} and Tax_{P6A} complexes, the latter of which was used for starting coordinates for the bound simulation) and energetically ($\Delta\Delta G_{int}$ by Piepenbrink, K. H., et al., manuscript in progress). In addition,
Figure 5.7 **Contact distances of the A6-Tax/HLA-A2 MD simulations.** Structural distances taken from the PDB 1QRN structure are shown for reference (blue). Distances between atoms of A6 residues (highlighted in lime) and atoms of the Tax peptide (cyan) or HLA-A2 (yellow) are average distances (in Å) intermittently calculated from the A6-Tax simulations (every 1 ps); error bars represent the standard deviation of these values. Alpha (A) and beta (B) chain residues of A6 are selected based on close contact in either the 1QRN or 1AO7 PDBs.
we calculated the correlation of motion between these contact atoms to assess their concerted dynamics (Fig. 5.8).

Unquestionably, the most restricted interactions occur with the residues of the CDR3α loop. Across this section of the TCR/pMHC interface, nearly all the CDR3α residues maintain a close proximity to their pMHC binding counterparts. For instance, the atoms involved in the T98α-R65, D99α-R65, and S100α-pG4 interactions each have an average distance of less than 3.5 Å. In addition, these atomic pairs interact with highly-correlated dynamics (0.42-0.58 cross-correlation values). Even larger correlation values are observed for the W101α and G102α residues with R65 of HLA-A2, which were previously determined to have the largest reduction in disorder upon binding for residues of the CDR3α loop. These results indicate that the restriction of the CDR3α loop, in terms of backbone mobility upon binding, translates into enthalpically-driven interactions more so with the MHC molecule, and not the peptide.

In contrast, the other CDRα loop residues less frequently maintain contact with the pMHC ligand, and have inconsistent correlations with their respective interatomic contacts. For example, Y50 of CDR2α and Q155 of HLA-A2 do not maintain close enough distance to seemingly interact (6.59 ± 0.81 Å), yet calculations show a similar extent of positive correlation as that for the CDR3α interactions. Intriguingly, position 155 is a conserved contact residue for nearly all TCR-pMHC-I complexes, and has been shown to contribute significant binding energy by its interaction with A6. The entire CDR1α-
Figure 5.8 Correlated motions of the A6-Tax/HLA-A2 intermolecular contacts. Cross-correlation values represent the level of concerted motion between two defined atoms. A value of +1 is the highest level of correlation, while -1 is the highest anti-correlation two atoms can possess. Residues are highlighted as in Fig. 5.7; alpha (A) and beta (B) chain residues of A6 are depicted.
pMHC network of atomic contacts, however, have no substantial correlated motion that also possess favorable interaction energies. Paradoxically, this cannot be solely attributed to the large distances between supposed “contact” atoms: S31α’s hydroxyl group is intermittently within H-bond distance to pY5’s, and the side chain of Q30α is proximal to the MHC residues of Y159 and T163 for potential VDW contacts. Altogether, the inability of these calculations to decipher a tight interactive process for α-chain loops other than CDR3α signifies the remarkable and exclusive specificity that the CDR3α loop contributes to the A6-Tax/HLA-A2 recognition process.

The CDR3β loop, which we earlier pointed out partially re-samples its free state, displaces its G100 and G101 residues from the peptide and MHC surface upon binding. This apparently causes an anti-correlated dynamic between the CDR loop and peptide. For instance, the A99β-pY8, G100β-pV7, G101β-pV7, and P103β-Y5 contacts all display either negligible or moderately negative correlation. This is significant, as the structural diversity of this loop across the various A6-peptide/HLA-A2 complexes is exclusive for these C-terminal residues (A99 through P103). The inability of these residues to be dynamically influenced by the peptide is revealing of its cross-reactive character. A6’s recognition of certain pMHC molecules may be determined by how well the CDR3β can maintain its independent conformational diversity – those which restrict its motion may not be recognized due to entropic costs of organization.
5.3 Discussion

5.3.1 Implications towards TCR binding mechanism

To gain a more comprehensive perception of the role of TCR loop dynamics in the recognition of pMHC, we examined the degree of conformational diversity across A6’s pMHC-binding interface, both in its free and bound state. Our thorough analysis revealed that CDR3α is its most conformationally restricted TCR loop upon binding. In close agreement with recently observed energetic and dynamic data from our group, CDR3α’s ligand compatibility involves HLA-A2 interactions more than with the Tax peptide. This suggests a divergence from a more conventional TCR recognition scheme, which depicts the MHC platform making conserved docking interactions with the evolutionarily-developed CDR1 and CDR2 germline loops, not the hypervariable CDR3 loops. The precise conformity of CDR3α with the Tax/HLA-A2 ligand is consistent with all of A6’s other ternary structures, indicating that the reduced dynamics of this loop is crucial to producing energy-stabilizing contacts above the threshold for recognition.

This leads to a close examination of the dynamic behavior of the CDR3β loop, which seems to play ying to the 3α’s yang. This loop tolerates a more diverse pMHC surface as it maintains an ability to sample non-Tax-specific conformations. The relatively poor shape complementarity at this region of the interface depicts a recognition strategy that does not require a conformationally restricted CDR3β. It could be speculated, therefore, that antagonistic ligands inhibit this loop’s dynamic freedom to a point of driving entropic penalties of binding too high. The various conformations
CDR3β takes with the ten known pMHC-bound structures of A6, consolidates this notion of a more promiscuous loop even after binding. Admittedly, comparing the dynamic data for A6 bound to Tax and the remaining static structures is a significant stretch; additional MD simulations performed with one or more of these A6-peptide/HLA-A2 complexes is needed to confirm these speculations.

Focusing now on the germline CDR loops (including the HV4α loop) which bind to Tax/HLA-A2, the results indicate that several of their backbones scattered throughout these loops maintain their pre-binding level of disorder in the presence of pMHC. This creates non-specific interactions not only with the α-helices of the HLA-A2 molecule, but the Tax peptide as well. Just as for the CDR3β loop, this lack of discrimination between antigenic peptide and the MHC molecule evokes an entropically-driven mechanism of binding, as these germline loops are capable of forming favorable interactions without increased organization. Overall, this seems to be a consistent trend, where the peripheral loops contribute favorably to the total free energy of binding via flexibility, not docking rigidly as predicted by the two-step model.
CHAPTER 6:
TWO TCRS SIMILAR TO THE A6 TCR – EXAMINING THE INTERDEPENDENCY OF LOOP DYNAMICS

6.1 Introduction

While rearrangement of TCR-encoded genes during T cell development generates a high level of diversity, it also allows for TCRs with shared germline-encoded sequences to co-exist within the lymphocyte repertoire. Thus, it often occurs where TCRs with shared Vα and/or Vβ sequences simultaneously appear in T cell lines activated against a particular antigen.\textsuperscript{88,89} This implies a level of degeneracy amongst TCRs of common germline lineage. Recognition by two similar TCRs could be driven by similar interactions with the same antigen in terms of conserved contacts being formed, as well as imposing equivalent dynamic influences on the binding mechanism.

In a vacuum, identical germline loops of different TCRs should display correspondingly identical dynamics. However, the presence of varying sequences in the hypervariable CDR3 loop, and its potentially unrelated monomeric α or β chain-pair, could result in divergent dynamics for these germline loops. As we have seen in Chapter 4, TRFA is a plausible method for examining the secondary effects on loop dynamics
from perturbations in neighboring loops, due to the probe’s sensitivity to local environments.

The A6 TCR studied exclusively in previous chapters stands as an ideal system to study with regard to interdependent loop dynamics. One reason for this is the identification of a TCR which shares the same Vβ germline sequence as A6: the B7 TCR. B7 also recognizes Tax/HLA-A2, with a very similar binding affinity compared to A6 (\(\Delta \Delta G^0 = 0.3\) kcal/mol at 25°C), and cross-reacts with related Tel1p and HuD antigenic peptides. While there is no free B7 crystal structure, its ternary complex with Tax/HLA-A2 has been solved and related to A6’s binding mode to the antigenic ligand (Fig. 6.1).

One might initially predict that the equivalent recognition of the same ligand by the A6 and B7 TCRs would, at least partially, result from the contributions by the identical sequences shared by the TCRs’ CDR1/2β germline loops. In the structural analysis of the two bound TCRs, however, the most significant divergence in ligand recognition is between these two sequentially conserved CDRβ loops. Whereas the genetically unrelated α chain of the two TCRs align closely (the largest deviation being only 1.0 Å within the CDR2 loop), the β chain of B7 is tipped in a different orientation relative to the pMHC binding groove. The positions of the B7 CDR1β and CDR2β loops are subsequently shifted from where the corresponding A6 loops reside (a deviation of approximately 5 Å for both loops). Subsequently, structural contacts identified for B7 versus A6 vary significantly: the CDR2β loop of B7 makes two contacts with the Tax
Figure 6.1 Loop positioning of the A6 and B7 TCRs bound to Tax/HLA-A2. Complexes are superimposed by the backbone atoms of their HLA-A2 heavy chains. (A) The Vα loops of A6 and B7 have very similar loop positioning and conformations. (B) The Vβ loops, despite nearly 95% sequence similarity, have large displacements in their binding solutions to Tax. Side chain residues in each of the CDR loops are displayed to more clearly depict the translational deviations of the entire germline loops (E30 – red; I54 – purple) and the C-terminal half of the CDR3 loop (E105 – green). Darker colored loops belong to A6.
ligand (Y48β to HLA R65, and I54β to HLA Q72) whereas this loop for A6 makes no contacts.

In light of these differential positions of identical sequences binding to the same ligand, the dynamics of these germline loops are of particular interest. If these TCRs possess different dynamics due solely to variation in the hypervariable CDR3β loop, it may reveal the underpinnings of the evolution of T cell immunity.

In addition, variants of A6 have been engineered which significantly increase the binding affinity with the Tax/HLA-A2 ligand. The mutation of \(^99\)AGGR\(^{102}\) in the CDR3β loop to MSAE (named the A6 c134 variant),\(^92\) for instance, improves Tax recognition by nearly 100-fold (unpublished data). Potentially, these mutations create more favorable contacts with the ligand, or possibly have an influence on the dynamics of its neighboring loops. Measuring the dynamics of the CDR3α and CDR1β of A6 c134 loops may reveal pertinent details regarding the alterations on the recognition mechanism between native and high-affinity A6.

6.2 Results

6.2.1 B7 and A6 have differential dynamics of their shared Vβ germline loops

Five sites within the CDR1β and CDR2β loops of B7 were mutated to cysteine and labeled with the F5M probe, identical positions as labeled within A6 in Chapter 4. Under equivalent conditions as the A6 measurements, the anisotropy data for the B7 TCR
should indicate if the chemical and structural variations of the CDR3β loop relative to A6 affect the dynamics of the shared germline loops of the two TCRs.

The results show a deviation in dynamics for the CDR1β loop, which is positioned much closer to the hypervariable CDR3β loop and the unrelated α-chain loops (Fig. 6.2). Position N28β, in particular, has a significant decrease in its flexibility value with respect to the A6 measurements ($\Delta f_i/\sigma_i = -0.37$). Since the only perturbation of the β chain is the sequence of the hypervariable loop, it could be speculated that the CDR3β of the B7 TCR interacts more prevalently with its CDR1β germline loop than for the A6 TCR.

Recalling from Chapter 4, it was determined that the N28Cβ-F5M sample was reporting high values due to the probe’s complete avoidance of contacts with neighboring loops, as opposed to all other positions investigated by MD simulations. Interestingly, the lifetime of the probe is not significantly altered when comparing A6 and B7, an indication that the dynamic influence is directly between the loops, and not a result of probe-protein interaction (Fig. 6.3).

In contrast, the further-removed CDR2β loop does not reveal any extreme differences of dynamics between the A6 and B7 TCRs. Thus the contacts observed for
Figure 6.2 Differential dynamics of the Vβ germline loops of the A6 and B7 TCRs. (A) Representative anisotropy curves for positions within the CDR1β and CDR2β loops of the two TCRs. All measurements are with the F5M probe. (B) Average fitted values of the anisotropy curves, comparing the A6 and B7 TCRs’ identical positions in the Vβ germline loops. Error bars represent the standard deviations of at least four independent measurements of F5M-labeled TCRs.
the two B7 residues in the ternary structure that are absent in A6-Tax/HLA-A2 do not appear to be a consequence of altered dynamics of the CDR2β loop.

6.2.2 The high-affinity A6 c134 displays different dynamics that the wild type A6

Similar to the approach and rationale of the B7-A6 dynamic comparison, A6 c134 was labeled in loops surrounding the mutated CDR3β loop, and within the altered loop itself. Direct comparison of the anisotropy data measured for the same positions in the wild-type A6 TCR will reveal if loop interdependency affects loop dynamics, and therefore plays a potential role in antigen recognition.

It is evident from the TRFA data that CDR3α of A6 c134 has an increased dynamics relative to the wild-type measurements (Fig. 6.4). Both positions within the
Figure 6.4 Distinguishing the dynamics of several CDR loops in the A6 and A6 c134 TCRs. All measurements were collected using the F5M probe. (A) Representative anisotropy curves of the wild type and high-affinity variant of A6. (B) Steady-state anisotropies of samples corresponding to (A).
CDR3α loop, S100 and W101, possess a higher level of flexibility according to the $f_i/\theta_f$ values (Table 6.1). The $^{99}$MSAE$^{102}$ sequence present in the opposing CDR3β of the high-affinity variant appears to have a significant effect on the conformational flexibility of CDR3α. Importantly, lifetime measurements for the positions S100 and W101 agree well between the A6 and A6 c134 TCRs, suggesting that probe-protein interactions do not significantly skew the observed differential dynamics (Fig. 6.5).

### TABLE 6.1

THE CORRELATION TIMES AND AMPLITUDES OF THE A6 AND A6 C134 TCRS

<table>
<thead>
<tr>
<th>Mutant</th>
<th>$\theta_f$ (ns)</th>
<th>$\theta_s$ (ns)</th>
<th>$f_i$ (100%)</th>
<th>$f_i/\theta_f$</th>
<th>$r_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A6 c134</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDR3α</td>
<td>S100 0.69 ± 0.14</td>
<td>18.0 ± 0.9</td>
<td>37.8% ± 4.3%</td>
<td>0.58 ± 0.15</td>
<td>0.264 ± 0.002</td>
</tr>
<tr>
<td></td>
<td>W101 0.78 ± 0.09</td>
<td>14.9 ± 1.8</td>
<td>39.7% ± 2.2%</td>
<td>0.52 ± 0.08</td>
<td>0.265 ± 0.002</td>
</tr>
<tr>
<td>CDR1β</td>
<td>D26 0.67 ± 0.05</td>
<td>9.9 ± 0.4</td>
<td>63.1% ± 0.8%</td>
<td>0.95 ± 0.08</td>
<td>0.218 ± 0.006</td>
</tr>
<tr>
<td>CDR3β</td>
<td>E102 0.85 ± 0.02</td>
<td>16.1 ± 0.5</td>
<td>40.1% ± 1.1%</td>
<td>0.47 ± 0.01</td>
<td>0.311 ± 0.001</td>
</tr>
<tr>
<td><strong>WT A6</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDR3α</td>
<td>S100 1.13 ± 0.51</td>
<td>18.3 ± 1.0</td>
<td>19.2% ± 4.5%</td>
<td>0.23 ± 0.15</td>
<td>0.277 ± 0.020</td>
</tr>
<tr>
<td></td>
<td>W101 0.89 ± 0.34</td>
<td>17.5 ± 1.5</td>
<td>26.3% ± 3.2%</td>
<td>0.35 ± 0.17</td>
<td>0.290 ± 0.021</td>
</tr>
<tr>
<td>CDR1β</td>
<td>D26 0.71 ± 0.08</td>
<td>9.3 ± 0.7</td>
<td>56.5% ± 1.5%</td>
<td>0.80 ± 0.10</td>
<td>0.246 ± 0.010</td>
</tr>
<tr>
<td>CDR3β</td>
<td>R102 0.67 ± 0.12</td>
<td>11.8 ± 0.4</td>
<td>38.6% ± 2.1%</td>
<td>0.60 ± 0.13</td>
<td>0.293 ± 0.015</td>
</tr>
</tbody>
</table>
A more modest increase in dynamics is observed for the CDR1β position of D26, while the labeling site within the CDR3β at position 102 indicates a slight decrease in flexibility due to the high-affinity mutations in the adjacent residues. This position was selected as the best labeling site in the CDR3β as the residue appears to contribute the least to the acquisition of high-affinity binding – a mutation of the native R102 residue to glutamine (99MSAQ102) instead of glutamate (MSAE, as used in this study) actually increases binding affinity by another order of magnitude.92
6.3 Discussion

The findings here indicate that variations in CDR loop sequence, imposed either naturally by gene rearrangement, or artificially engineered, can have propagating effects on the dynamics of the TCR binding surface. The former case is exemplified by the comparison of the B7 and A6 TCRs, which share the same germline Vβ loops, yet were determined to bind to the same antigen (Tax/HLA-A2) with disparate binding positions by the CDRβ loops. As the TRFA studies indicate here, the variation in loop positioning could very well be a result of the dynamic interplay between the germline loops and the hypervariable loops of the two TCRs. Fig. 6.6 compares the residues of their corresponding CDR3β loops, revealing only seemingly minor deviations in their sequence and chemical attributes. Other than an additional residue for the A6 hypervariable loop, and two arginines (A6) in the place of tyrosines (B7), the sequences are nearly identical. However, the TRFA data implies that these differences may be significant in the governance of the dynamics of the β-germline loops. Mindfully, the distinct α-chain loops of the two TCRs may also have an indirect influence on the β-chain dynamics.

<table>
<thead>
<tr>
<th></th>
<th>A6</th>
<th>B7</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDR3β</td>
<td>CASRPGLAGGRP EQ</td>
<td>CASSYPGGGY - EQ</td>
</tr>
</tbody>
</table>

Figure 6.6 Hypervariable CDR3β sequences for the A6 and B7 TCRs. Identical residues of the TCRs are highlighted in red.
In the case of the high-affinity variant of A6, we see an even more drastic perturbation on neighboring loop dynamics, despite only a four-residue alteration in the CDR3β loop. This observation is even more striking when considering that the structures of the two TCRs bound to Tax/HLA-A2 are identical (except for the side chains of the mutated residues; Fig. 6.7). For the CDR3α loop, possessing heightened dynamics relative to the wild-type would have a major impact on the energy landscape discussed in Chapter 3. If the altered CDR3β loop creates an opportunity for the CDR3α loop to access its binding-competent state more frequently (i.e., the energy barrier separating its two distinct states is lowered), it could be the rationale for how the higher binding affinity is achieved.

![Figure 6.7 Identical backbone conformations of the A6 and A6 c134 TCRs bound to Tax/HLA-A2.](image)

The CDR loops of A6 investigated in this study are shown, completely superimposable to the high-affinity variant c134, and are therefore not distinguishable. Side chains of the mutated residues of the c134 CDR3β loop are labeled.

125
The concept of dynamic interdependency of CDR loops raises fundamental questions of how proteins with complex tertiary structures are used in recognition and overall biological function. These preliminary results are a starting point for examining the effect of mutations on TCR binding, a method for developing cancer immunotherapies growing in popularity. By understanding how adjacent CDR loops interact from a dynamic standpoint, the creation of high-affinity drugs may more predictable and therefore more cost-effective.
CHAPTER 7:  
CONCLUSIONS

7.1 Rationale Behind the Investigations

The necessity for T cell engagement with antigen-presenting cells to solicit the correct outcome (either evasion or cell death) is a particularly intriguing feature of the adaptive immune system. Multi-specificity towards a large range of pMHC antigens would appear to put even more pressure on the T cell’s responsibility to decide the fate of the large population of cells with which it encounters. As complex as this recognition process may seem, its obvious success is a testament to the precision by which T cell surveillance is governed.

The focus of these investigations have revolved around the understanding of T cell receptor (TCR) dynamics, and addressing the binding mechanisms that can be elucidated from such findings. Current models describing the TCR-pMHC interaction underestimate the role played by the pre-existing flexibility of the TCR’s complementarity determining region (CDR) loops. This is evidenced by the abundance of structural studies performed on TCR and pMHC systems, towards the goal of expanding the library of TCR structures to a point of discovering trends for TCR binding. Far less effort has been placed into investigations that address the dynamic aspects of
the receptor protein, although a few NMR and molecular dynamics (MD) studies should be acknowledged.\textsuperscript{73,86} The prevalent role of dynamics in protein-protein interactions has been documented;\textsuperscript{57} here protein dynamics have been assessed towards the understanding of the TCR-pMHC binding mechanism.

7.2 The Role of CDR Loop Dynamics in pMHC Recognition

The results of the time-resolved fluorescence anisotropy (TRFA) and MD studies have shown that the individual CDR loops of a TCR have differential dynamics relative to one another, both in terms of conformational diversity (e.g., the number of states it can access) as well as kinetics (e.g., how frequently distinct conformations are exchanged between). The independent dynamics possessed by each CDR loop appear to induce distinct functions in the pMHC recognition process, as attested here for the A6 TCR.

Particularly, the CDR3 hypervariable loops, possessing a heightened level of inherent flexibility relative to the germline loops, seem to regulate each other’s dynamic contributions to binding. The CDR3\(\alpha\) loop is depicted as having a slower exchange between two distinct conformations, while its CDR3\(\beta\) counterpart more frequently exchanges between a greater number of isoenergetic states. Thus, the CDR3\(\alpha\) loop is more conducive to establishing an energetically-stabilizing interaction (specificity), while the promiscuous nature of CDR3\(\beta\)’s dynamics is capable of broadening the potential number of antigen with which the TCR can react (cross-reactivity). The pairing of these two hypervariable loops during T cell development may be one of many possible strategies of the adaptive immune system to regulate the affinities of the TCRs towards
foreign antigen. This energetic threshold governing T cell activation, however, cannot be overcome during the receptor’s engagement with self-antigen during negative selection and in the periphery.

Relating CDR loop dynamics to the A6-Tax/HLA-A2 recognition mechanism requires a closer examination of the enthalpic and entropic effects to the overall energy of binding. The conformational selection of the binding-competent state of the CDR3α loop is seemingly an enthalpically favorable event. The loop is reported to lock into a very restricted conformation after binding to the Tax/HLA-A2, reducing its level of conformational entropy. By doing so, however, the CDR3α loop is allowed the opportunity to create long-lasting bonds with HLA-A2 residues, overcompensating for the entropic penalties with stabilizing enthalpic interactions. The overall interaction energy of the CDR3α loop ($\Delta G^0$), identified by both experimental (Piepenbrink et al., unpublished work) and computational\(^86\) approaches, is in complete agreement with this dynamic assessment.

In contrast, the CDR3β loop is much more indiscriminate in the selection of its binding conformation, permitting its residues to sample its free state even after the A6’s interaction with the Tax/HLA-A2 surface. The contributions of this loop, with respect to the binding energy, can therefore be perceived as imposing minimal entropic penalty, but a lack of enthalpic gains. Notably, this qualitative assignment of energetics should be limited to the segment of the loop in which conformational diversity was observed (the $^{100}\text{GGR}^{102}$ residues), as the favorable contributions by the flanking residues of L98
and P103 appear to be enthalpically-driven (relative to results from Piepenbrink et al., unpublished data). The dynamic profile of CDR3β, therefore, suggests a lack of specificity towards the antigenic peptide by this particular loop.

Surprisingly, the general correlation of motion and contact between the CDR loops of A6 and Tax/HLA-A2 is quite low. This implies that, apart from the CDR3α loop, this TCR-pMHC complex lacks dynamic conformity between the recognizing molecule and its ligand and is therefore an entropically-driven event. Interestingly, this was an observation made previously in earlier work done within our lab, but the specific location of the enthalpic and entropic contributions could not have been elucidated without the investigations presented here. The conclusions to which we have come regarding the dynamic effects on the A6-Tax/HLA-A2 recognition process thus stand in solid agreement with relevant thermodynamic studies.

7.3 Applications of Time-Resolved Fluorescence Anisotropy and Molecular Dynamics Simulations in Determining TCR Dynamics

A significant portion of this investigation was devoted to verifying the accuracies of the techniques implemented for studying the dynamics of TCRs. This was deemed a necessity due to the assumptions or limitations that each technique inherently carries within its interpretations of protein dynamics.

With respect to TRFA measurements, potential inaccuracies stem from the fluorescent probe’s bulkiness. In this study, a substantial disconnect has been observed between the motion of the probe’s 3-membered ring (the location of the transition
dipole, from which the fluorescence emanates) and the motion of the backbone to which the probe is attached. The level of disconnect is directly proportional to the degrees of physical separation – the further removed the source of the signal is from the segment of the protein to which dynamics are being assigned, the less concerted the dynamics of these two segments can be. From this relationship it can be deduced that the most ideal probes for TRFA studies would be those with the fewest degrees of freedom separating the transition dipole from the covalently-attached backbone of the protein. This makes a strong argument for the application of intrinsic tryptophan (Trp) fluorescence anisotropy, although this method has obvious limitations as well: low quantum yield, indistinguishable fluorescence between various tryptophans, and only sites which possess a naturally-occurring Trp residue can be assessed for dynamics.

In addition, the propagation of dynamics imposed from the backbone to the probe is relative to the extent of backbone motion itself. Native probe dynamics must be acknowledged and accounted for in the interpretations of TRFA data. This is why several studies, when the appropriate instrumentation allows, resolve a separate time constant that is assigned to the probe’s loss of correlation (REF X). For a protein segment that undergoes large conformational transitions, subsequent displacement of the probe will obscure the contributions of the probe’s independent dynamics to the observed anisotropy. Conversely, a relatively rigid protein segment will impose little probe displacement, allowing the probe’s protein-independent motions to dominate the anisotropy measurement. Thus, two protein segments of equivalently low protein
flexibility could be observed as having differential dynamics, if the probe of one segment is allowed to more freely decorrelate within its environment than the other. In this scenario, solvent accessibility is more resolved in these measurements than is actual protein dynamics.

This unavoidable aspect of TRFA measurements – through the application of bulky probe attachment – is an unsettling realization. Conclusions drawn regarding protein dynamics from TRFA data alone, therefore, would hold little weight. The fact that these alarming details regarding TRFA were elucidated by the MD simulations performed on the A6-F5M systems, however, gives that much more credence to the level of accuracy by which computational approaches can predict real protein (and probe) dynamics. In these investigations we have observed remarkable agreement between MD sampling and crystallographically-solved conformations (Chapter 3). The additional level of resolution that can be obtained for the conformational flexibility of the CDR loops presents a strong case for using MD simulations to complement experimental studies such as TRFA. Finally, it can be said with great confidence that MD simulations should be more frequently used in the pursuit of TCR-pMHC binding mechanisms revolving more around dynamics.
APPENDIX A:

FLUORESCENCE ANISOTROPY DATA FOR THE SINGLE-CHAIN 2C TCR

A.1 Steady-State Fluorescence Anisotropy

Figure A.1 SSFA measurements at 25°C.
A.2 Time-resolved Fluorescence Anisotropy

Figure A.2 Anisotropy curves at 25±1°C.
<table>
<thead>
<tr>
<th>Position-Label</th>
<th>$\theta_r$ (ns)</th>
<th>$\theta_s$ (ns)</th>
<th>$f_f$ ($\times100%$)</th>
<th>$f_f/\theta_f$</th>
<th>$r_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Alpha Chain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S17Cα-F5M</td>
<td>0.815 ±0.026</td>
<td>10.0 ±0.2</td>
<td>41.2 ±0.6</td>
<td>0.506 ±0.016</td>
<td>0.259 ±0.002</td>
</tr>
<tr>
<td>S17Cα-BDY</td>
<td>1.537 ±0.155</td>
<td>13.9 ±0.8</td>
<td>30.1 ±2.0</td>
<td>0.197 ±0.015</td>
<td>0.257 ±0.006</td>
</tr>
<tr>
<td>S27Cα-F5M</td>
<td>0.850 ±0.047</td>
<td>10.0 ±0.1</td>
<td>47.0 ±1.0</td>
<td>0.554 ±0.036</td>
<td>0.269 ±0.003</td>
</tr>
<tr>
<td>S27Cα-BDY</td>
<td>0.978 ±0.083</td>
<td>9.9 ±0.5</td>
<td>51.6 ±1.2</td>
<td>0.530 ±0.032</td>
<td>0.265 ±0.002</td>
</tr>
<tr>
<td>S69Cα-F5M</td>
<td>0.687 ±0.090</td>
<td>10.7 ±1.0</td>
<td>64.3 ±2.8</td>
<td>0.951 ±0.139</td>
<td>0.233 ±0.010</td>
</tr>
<tr>
<td>S69Cα-BDY</td>
<td>0.801 ±0.008</td>
<td>7.4 ±0.1</td>
<td>66.3 ±0.3</td>
<td>0.827 ±0.005</td>
<td>0.188 ±0.001</td>
</tr>
<tr>
<td>F100Cα-F5M</td>
<td>0.651 ±0.100</td>
<td>9.3 ±0.7</td>
<td>43.0 ±1.0</td>
<td>0.676 ±0.117</td>
<td>0.315 ±0.010</td>
</tr>
<tr>
<td>F100Cα-BDY</td>
<td>0.927 ±0.038</td>
<td>11.8 ±0.2</td>
<td>50.5 ±0.2</td>
<td>0.546 ±0.021</td>
<td>0.287 ±0.003</td>
</tr>
<tr>
<td><strong>Beta Chain</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T20Cβ-F5M</td>
<td>1.290 ±0.201</td>
<td>10.4 ±0.4</td>
<td>23.0 ±1.4</td>
<td>0.180 ±0.022</td>
<td>0.298 ±0.003</td>
</tr>
<tr>
<td>N30Cβ-F5M</td>
<td>0.690 ±0.051</td>
<td>9.0 ±1.4</td>
<td>47.1 ±2.0</td>
<td>0.687 ±0.068</td>
<td>0.289 ±0.017</td>
</tr>
<tr>
<td>N30Cβ-BDY</td>
<td>1.080 ±0.036</td>
<td>12.7 ±0.1</td>
<td>30.1 ±0.5</td>
<td>0.279 ±0.012</td>
<td>0.292 ±0.002</td>
</tr>
<tr>
<td>S68Cβ-F5M</td>
<td>0.871 ±0.034</td>
<td>9.7 ±0.1</td>
<td>43.7 ±0.7</td>
<td>0.502 ±0.025</td>
<td>0.262 ±0.003</td>
</tr>
<tr>
<td>S68Cβ-BDY</td>
<td>1.21</td>
<td>7.9</td>
<td>61.6</td>
<td>0.509</td>
<td>0.224</td>
</tr>
<tr>
<td>G97Cβ-F5M</td>
<td>0.721 ±0.058</td>
<td>9.3 ±0.2</td>
<td>43.5 ±0.2</td>
<td>0.606 ±0.049</td>
<td>0.318 ±0.004</td>
</tr>
<tr>
<td>G97Cβ-BDY</td>
<td>1.053 ±0.025</td>
<td>9.9 ±0.0</td>
<td>46.4 ±0.3</td>
<td>0.441 ±0.013</td>
<td>0.286 ±0.002</td>
</tr>
</tbody>
</table>
Figure A.3 Diagram for the TRFA 2C data.


