Spring 5-17-2010

High-Throughput Engineering and Analysis of Class II Mhc/Peptide Binding by Yeast Co-Display

Wei Jiang
Unversity of Pennsylvania, jiangwe@seas.upenn.edu

High-Throughput Engineering and Analysis of Class II Mhc/Peptide Binding by Yeast Co-Display

Abstract
Polymorphisms of major histocompatibility complex (MHC) and molecular mechanisms of their antigen-presenting specificity and promiscuity have great impact on T cell-mediated immune responses and related diseases. Challenges in elucidating the characteristics of antigenic peptide binding by MHC motivate the development of high throughput experimental tools to quantitatively analyze interactions between hundreds of MHC allelic proteins and various peptide sequences. We demonstrated such a method by co-displaying target peptides and class II MHC (MHC-II) on the yeast surface in an intracellular association-dependent manner. The optimized yeast co-display system enabled quantitative mapping of side-chain preferences and general motifs for peptides binding to MHC-II by site-directed mutagenesis or peptide library screening, and also allowed rapid tailoring of MHC-II peptide binding specificity by directed evolution approaches, which derived MHC-II allelic mutants with altered peptide binding specificity or hyper-promiscuity. Comparison of these experimentally engineered mutants with naturally discovered MHC-II proteins recovered valuable information about structure-function relationship in the evolutionary mechanisms for polymorphic MHC-II molecules, which could direct future immunotherapeutic innovation.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Chemical and Biomolecular Engineering

First Advisor
Wei Jiang

Keywords
yeast surface display, MHC-II, directed evolution, peptide binding, yeast co-display

Subject Categories
Biochemical and Biomolecular Engineering

HIGH-THROUGHPUT ENGINEERING AND ANALYSIS OF CLASS II MHC/PEPTIDE BINDING BY YEAST CO-DISPLAY

Wei Jiang

A DISSERTATION

in

Chemical and Biomolecular Engineering

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2010

Supervisor of Dissertation

Signature_____

Eric T. Boder, Associate Professor

Chemical and Biomolecular Engineering, University of Tennessee

Graduate Group Chairperson

Signature_____

Matthew J. Lazzara, Assistant Professor of Chemical and Biomolecular Engineering

Dissertation Committee

David J. Graves, Associate Professor of Chemical and Biomolecular Engineering

Dimitri S. Monos, Professor of Pathology and Laboratory Medicine, Children's Hospital of Philadelphia

Scott L. Diamond, Professor of Chemical and Biomolecular Engineering

ACKNOWLEDGEMENTS

First of all, I would like to thank my advisor Dr. Eric Boder for his help and guidance on my Ph.D. project for the past five and a half years. Eric is such a nice person that I really felt comfortable and confident to do research abroad with him. During countless friendly communications with him, both of my academic background and English presentation skills have been improved dramatically. Most importantly, the freedom Eric gave me upon designing projects and performing experiments helped me developing the potential to think and work independently. I would also like to thank Dr. Dimitrios Monos, Dr. Scott Diamond, and Dr. David Graves, for being in my thesis committee and proposing helpful suggestions to my project. As an immunologist, Dr. Monos offered a lot of great opinions and valuable immunological background support on my research. I would like to give special thanks to Dr. Lawrence Stern (University of Massachusetts Medical School) for sharing the plasmid pDLM1-drb1s, which provides the necessary MHC-II gene to start my project.

In the Boder Lab, many people deserve my acknowledgements. Andrew Nields was the first one tutoring me through most of experimental techniques for months and assisting me on the MHC-II engineering project, which he had been working on for many years. Even though we did not have too long overlap periods in the lab, Andrew did help me a lot to cross both research and language barriers. I would like to thank Ranganath Parthasarathy, an experienced chemist and bioengineer with whom I spent the most years in lab. Ranganath provided both technical expertise and experimental assistance on my thesis work and actually made my first two years of life in the lab

easier and more bearable after lab relocation to Tennessee. Jeong Lee, Lauren Pepper, and Kyudam Oh not only shared their interesting ideas with me in research but also befriend to me in my life. Sheldon Park, Shyam Subramanian, Choi Fong Lu and the undergraduate student Nick Dang also assisted me a lot during my first years of Ph.D. study at Penn.

There were many people at University of Tennessee who have provided assistance to my research. Nancy Neilsen (College of Veterinary Medicine) spent a lot of time and patience on showing me technical details on how to set up a cell sorter for better performing one of my key experiments, which indeed greatly helped me to understand the theoretical background beyond the experiment itself and to perform troubleshooting thereafter. Dianne Trent (Department of Comparative Medicine) and Shige Eda (Department of Forestry, Wildlife, and Fisheries) also kindly operated their equipments for me to complete many flow cytometric analyses. Joe May (Division of Biology) gave me a lot of suggestions on how to generate better DNA sequencing results and offered couples of reruns for my samples with unsuccessful results. Other members joining the Boder lab after relocation: Vince Price, Ellen Messenger, Maryam Raeeszadeh Sarmazdeh, and Sathish Dharmalingam, together created an enjoyable lab atmosphere for my research in Tennessee.

Finally, I would like to thank my parents – Zuozhao Jiang and Deling Zheng in China for their continuous understanding, encouragement and support both on my life and oversea study, especially during the time when I was writing this dissertation.

This thesis work was funded by the U.S. National Science Foundation and the U.S. National Institutes of Health.

ABSTRACT

HIGH-THROUGHPUT ENGINEERING AND ANALYSIS OF CLASS II MHC/PEPTIDE BINDING BY YEAST CO-DISPLAY

Wei Jiang

Eric T. Boder

Polymorphisms of major histocompatibility complex (MHC) and molecular mechanisms of their antigen-presenting specificity and promiscuity have great impact on T cell-mediated immune responses and related diseases. Challenges in elucidating the characteristics of antigenic peptide binding by MHC motivate the development of high throughput experimental tools to quantitatively analyze interactions between hundreds of MHC allelic proteins and various peptide sequences. We demonstrated such a method by co-displaying target peptides and class II MHC (MHC-II) on the yeast surface in an intracellular association-dependent manner. The optimized yeast co-display system enabled quantitative mapping of side-chain preferences and general motifs for peptides binding to MHC-II by site-directed mutagenesis or peptide library screening, and also allowed rapid tailoring of MHC-II peptide binding specificity by directed evolution approaches, which derived MHC-II allelic mutants with altered peptide binding specificity or hyper-promiscuity. Comparison of these experimentally engineered mutants with naturally discovered MHC-II proteins recovered valuable information about structure-function relationship in the evolutionary mechanisms for polymorphic MHC-II molecules, which could direct future immunotherapeutic innovation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1. Background and significance

1.1. Overview

Biological evolution makes creatures on the earth more and more sophisticated as well as better and better developed. For example, to protect against diseases caused by various agents in the surrounding, organisms have developed a defending bio-system called immune system[1-3] by gradually accumulating and coordinating subtle helpful immunological features supplied by specifically differentiated cells and naturally evolved molecules over thousands or millions of generations. Natural evolution is slow and results are determined by complex fitness constraints existing in nature. Nonetheless, evolution remains a powerful paradigm for optimizing biological systems. Thus, *in vitro* methods for directed evolution[4, 5], wherein researchers can control fitness drivers guiding change, have garnered substantial interest in the last two decades. These methods are commonly and most simply applied to proteins[6, 7], the basic unit of biological function.

Most highly developed vertebrates, such as mammals and humans, have a relatively complicated immune system consisting of many types of organs, tissues, cells, and proteins, which connect with each other in an elaborate and dynamic process. Within this network, CD4+ T cell, a sub-group of lymphocytes, serves as a director to control and optimize the function of other immune effector cells[8] and a booster to maximize the efficacy of immunity. The role of these lymphocytes is so important that dysfunction or deficiency of functional CD4+ T cells in human individuals will result in fatal immune disorders, such as severe combined immunodeficiency (SCID)[9] and acquired immune

deficiency syndrome (AIDS) caused by a notorious virus, HIV[10]. It is of great importance to understand the mechanism for stimulation of CD4+ T cell responses and find a possible way to modulate their activation and function. Studies of CD4+ T cells at the molecular level indicate that two signals are involved in the process of their epitope specific activation[11, 12]. The primary pathway for transmitting signal 1 is predominantly determined by the molecular interaction between T cell receptor (TCR) and class II major histocompatibility complex (MHC-II) associated with antigenic peptide processed and presented by professional antigen-presenting cells (APCs)[11, 13]. Therefore, peptide-binding specificity of MHC-II proteins plays an important role in defining the epitope specificity of T cell activation and evolution of MHC-II for altered specificity implicates potentials in regulation of helper T cell-mediated immune responses.

As one of the most polymorphic membrane proteins[14, 15], MHC has already been selected for millennia by nature[16, 17] so that each MHC-II allele is poly-specific for a large set of antigenic peptides derived from both endocytosed pathogens and cytosolic proteins[18-22] and only a few alleles are required and expressed by each individual for presentation of peptides and initiation of T cell responses to overwhelm a verity of foreign invaders. However, environmental changes enforce improvement of MHC's peptide selectivity such that CD4+ T cell mediated immune responses could function against mutagenic pathogens. Laboratorial protein engineering techniques enable an initial evaluation of these evolutionary processes experimentally, which not only help to understand structure-function relationship for naturally evolved MHC-II molecules but also shed light on artificially modulating immune responses for therapeutic purpose.

This thesis work is focused on 1) development of a quantitative, high throughput *in vitro* protein engineering methodology – yeast co-display, for studying the interaction between various peptides and diverse MHC-II molecules (Chapter 2 and 3); 2) characterization at the molecular level of peptide binding specificity for selected MHC-II allele (HLA-DR1), determination of side chain preference for some pocket-like region in the binding site of DR1 (or pocket profile[23]), and prediction of DR1 specific ligands using the novel technology (Chapter 4); 3) modification of pockets within the peptide binding groove of DR1 for improved and/or altered peptide binding specificity by directed evolution, which has a potential to regulate T cell activation (Chapter 5).

1.2. Central role of CD4+ T cells in adaptive immune responses

1.2.1. Brief introduction for Immune system and lymphocytes

The immune system that most vertebrates have not only inherits the oldest innate immunity, but also develops a more evolved adaptive immunity, first appeared in jawed vertebrates[2] and became increasingly specialized with further evolution. Adaptive immune system consists of macrophages and other phagocytic cells, lymphocytes and their derivatives such as antibodies[1, 3, 24]. Lymphocytes are the only cell types in the body capable of specifically recognizing and distinguishing different antigenic determinants and are responsible for the adaptive immune response, specificity and memory[25]. A lot of congenital and acquired immunodeficiencies are related to molecular defects in the development and function of lymphocytes or their products[9]. Therefore, many immunotherapeutic approaches and vaccinations aim at modifying and optimizing the performance of lymphocytes or its mediators[26-28].

There are three major subsets of lymphocytes[25]: 1) Natural killer (NK) cells, a part of innate immune system, can recognize and kill infected cells directly without a need for additional activation; 2) B cells, main components of adaptive immune system, are responsible for secreting antibodies, which stimulate humoral immune responses by directly capturing antigens; 3) T cells, another group of main components in adaptive immune system, only respond to antigens processed and presented by APCs via a MHC-invoked secretory pathway and stimulate cell-mediated immune responses. Two important effector cells differentiated from T lymphocytes involved in MHC-restricted activations are T helper cells (T_h cells, expressing CD4 coreceptors for MHC-II binding) and CD8+ cytotoxic T lymphocytes (CTLs, expressing CD8 coreceptors for MHC-I binding).

1.2.2. Differentiation and Function of CD4+ T cells

Not like NK cells, antibodies or CTLs, T_h cells do not have the ability to eliminate infected pathogens by themselves, however, taking up almost half of the lymphocyte population in human body, they do play an important role in adaptive immunity or even innate immunity. Because of their regulatory function, these effector cells of CD4+ T cells are usually designated as helper cells. Of particular interest is that CD4+ T cells may differentiate into two different subsets of effector cells, T_h1 and T_h2, which produce distinct cytokines and perform distinct effector functions. Once being activated, T_h1 cells are essential in stimulating antibody secretion and determining B cell antibody class switch, in promoting the activation and proliferation of CTLs[8], and in regulating activity of phagocytes such as macrophages, neutrophils, and other leukocytes responsible for

cell-mediated immunity (Figure 1-1). T_h2 cells on the hand mainly function as suppressers for these effector cell-mediate responses. Both cell subsets are involved in formation of tissue injury and inflammation as well as generation of memory T cells, which may persist for years after acute immune responses are down-regulated and disappeear[25].

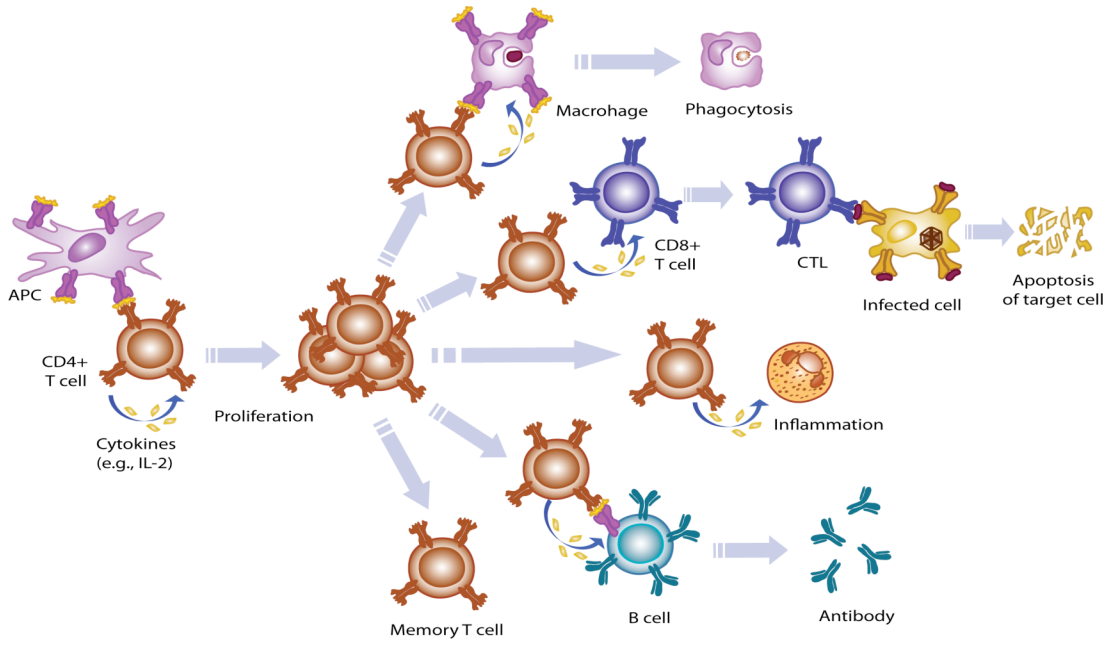

Figure 1-1 Proliferation and function of CD4+ T cells. CD4+ T cells are activated by antigenic peptide presented by professional APCs and driven to the clonal expansion by cytokines (e.g. interleukin-2). The activated and differentiated T_h cells can further activate and differentiate macrophages, CD8+ T cells, and B cells, or introduce inflammation. The main function of activated macrophages, CD8+ T cells, and B cells are to lyse phagocytosed antigens, to kill antigen-infected cells or tumor cells, and to secrete antigen-specific antibodies, respectively. Once antigens are eliminated, T_h cells will be down regulated and entering a memory phase. (Adapted from Abbas, A.K., Lichtman, A.H. & Pober, J.S. Cellular and molecular Immunology, 4/e (Saunders, philadelphia, 2000)).

Both B cell and macrophages are professional APCs for T_h cells[29], so immature B cells or macrophages can process and present antigen fragments to specifically activate CD4+ cells, which in turn will fully activate B cells and macrophages by cytokines (e.g., interferon-γ (IFN-γ)[30, 31]) or CD40-mediated signaling[32-34] (Figure 1-2) for antibody secretion and phagocytosis respectively (Figure 1-1).

In contrast, CD8+ T cells do not serve as professional APCs to T_h cells and their development and function do not necessarily require presence of T_h cells. In a strong innate immune response, the primary activation of CD8+ T cells can be elicited by dendritic cells (DCs), if microbe directly infects DCs[35], or if cross-presentation[36, 37] of microbial antigens is sufficient. However, the participation of T_h cells is usually required for CD8+ T cell responses to tumor cells[38], viral infections[39, 40] and for expansion of memory CD8+ T cells[41, 42]. Communication between these two types of T cells is largely mediated by professional APCs, such as DCs[43], via interleukin-2 (IL-2) signaling[41] (Figure 1-1) or costimulatory signaling mediated by CD40:CD40 ligand[39, 44, 45] or CD28:B7-1/B7-2 pathway[42] (Figure 1-2).

Thus, CD4+ helper T cell actually plays a central role in adaptive immunity. The possibility of controlling or adjusting CD4+ T cell development and function at molecular level would present intriguing potential for immunotherapy of numerous diseases related to T_h cell-mediated responses[46-48].

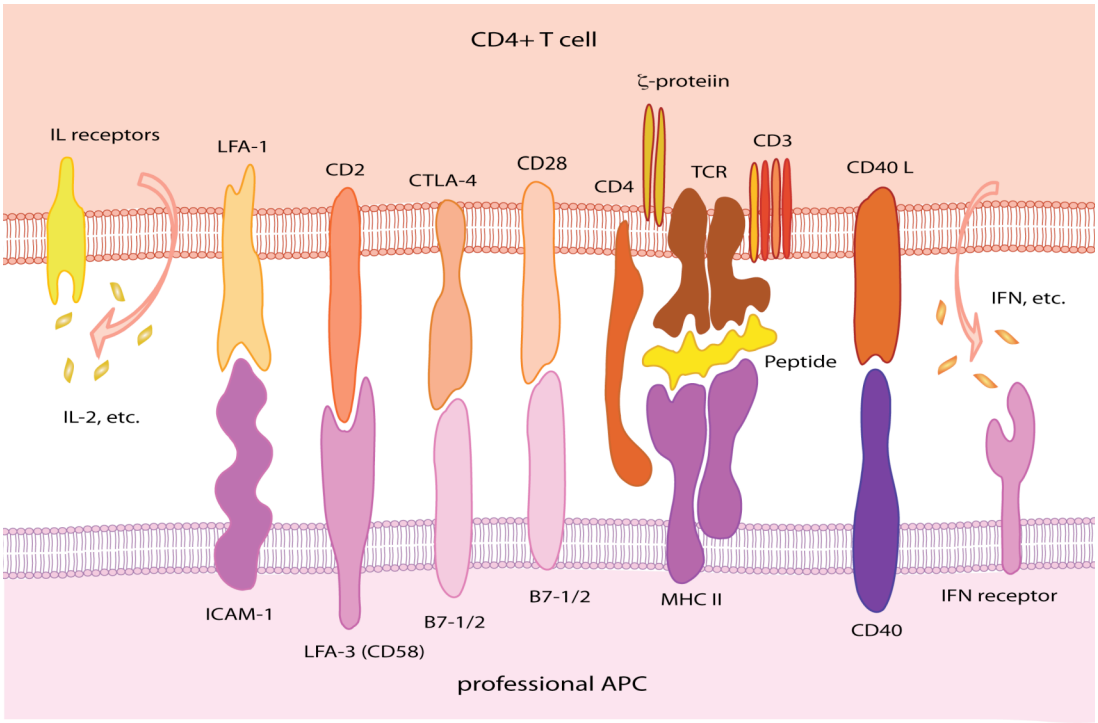

Figure 1-2 Signaling pathways between CD4+ T cell and APCs.

1.2.3. Activation of CD4+ T cells by APCs

In order to enter the functional stage, naïve T cells need to be activated and differentiated into effector T cells by specialized APCs, called "professional APCs" (Figure 1-1). As mentioned earlier, DCs, B cells and macrophages can all serve as professional APCs to present antigenic peptides for CD4+ T cell recognition and activation[29, 35]. However, the amount of MHC-II and cotimulators they can express, the phase of adaptive immune responses at which they can activate CD4+ T cells, the locations where they meet with T cells, and the functional feedback received from T_h cells are all different[25].

B cells and macrophages with lower levels of MHC-II mainly present antigen to differentiated effector T_h cells, and the process of antigen presentation mostly takes place in lymphoid organs. Other than those places, macrophages can also mediate T cell

responses in tissues, which might cause delayed type hypersensitivity. The main purpose of this antigen presenting process is to further activate B cells and macrophages themselves and fully accomplish their function. DCs on the contrary, resident in epithelia and tissues, can capture antigens and transport them to secondary lymphoid organs, including lymph nodes and spleen and present them to naïve CD4+ T cells. They can express large amount of MHC-II and costimulators on the surface and are the most effective APCs for initiating primary T cell responses.

Adoptive transplantation of autologous APCs such as DCs has been used in immunotherapy to generate more effective T cell-mediated immune responses[26, 49]. However, the labor intensity and high expense on producing APCs and the unreliable quality and quantity of the available APCs that can be used in adoptive immunotherapy all limit their usage in clinical applications. Therefore, investigators have started to develop various artificial APCs (aAPCs) to enhance the generation of antigen specific T cells[27]. No matter which method is going to be used, first of all, it is critical to clarify the molecular mechanism that leads to the activation of CD4+ T cells by APCs.

1.2.4. Signaling pathways for CD4+ T cell activation

T cell activation needs both antigen recognition and costimulation. After antigens are specifically processed and presented by professional APCs to CD4+ T cells, T cells will receive both "signal 1" through TCRs and "signal 2" via costimulator receptors from professional APCs[11, 12].

The primary signaling pathway, which determines the specificity of CD4+ T cell-mediated adaptive immune responses, is formed by the interaction between CD3 and

ζ protein-associated TCRs (TCR complexes) and APC-processed antigenic peptide-coupled MHC-II molecules (peptide/MHC-II complexes) along with the CD4 coreceptor[12, 13] (Figure 1-2). The antigen specificity of T cell response is greatly dependent on antigenic peptides processing and presentation by APCs and peptide-MHC complex recognition of TCRs on the T cell surface.

Other than the first signaling pathway, there are several intermolecular ligand-receptor interactions that function to deliver the second signal which is necessary for optimization of T cell activation and differentiation: CD28:B7-1/B7-2, the principal pathway for delivering second signals for T cell activation[12, 13]; CD2: CD58, signal transducer as well as intercellular adhesion molecules. Some other ligand-receptor pairs might only serve as adhesion molecules, such as integrin LFA-1[50]: ICAM-1/ICAM-2, however, they are necessary for keeping T cells and APCs in close contacting with each other for signaling transduction[25] (Figure 1-2).

Although the appearance of costimulatory molecules is substantial, specialized MHC-II proteins on the surface of APCs perform the main task of displaying cell-associated antigens for recognition by CD4+ T cells. Therefore, intensive studies using *in vivo* and *in vitro* methods have been carried out and will keep focusing on elucidating the cellular and molecular basis of antigen processing inside APCs as well as interaction between antigenic peptide and MHC molecules, which is also a key to the construction of aAPCs and to the design of vaccines for modulating T cell responses. One purpose of this thesis work is trying to develop a quantitative, high throughput method for better understanding the characteristics of MHC-II molecules that determine their selectivity for peptides.

1.3. MHC-II and antigenic peptide binding

1.3.1. Classification and evolution of polymorphic MHC

The MHC locus is a large gene cluster found in the genome of most vertebrates. The main products encoded by these diverse genes are expressed as trans-membrane glycoproteins on the surface of APCs, which have the ability to degrade and process antigens into short fragments (peptides) and transport them using MHC proteins to the surface for recognition by antigen-specific T cells. Two well-known subgroups of MHC proteins are 1) class I (MHC-I), expressed on the surface of all nucleated cells and responsible for presenting intracellular peptides to CD8+ T cells, and 2) class II (MHC-II), expressed on the surface of professional APCs, such as DCs, macrophages and B cells and responsible for presenting extracellular peptides to CD4+ T cells[51]. In humans, MHC proteins are referred to as human leukocyte antigen (HLA), and most intensely studied HLA alleles are HLA-A, HLA-B, HLA-C (for class I) and HLA-DR, HLA-DP, HLA-DQ (fcor class II).

MHC is the most polymorphic gene in the human genome. Up till today, multiple alleles have been found for all the nine classical genes (the heavy chain for each of class I alleles and two chains for each of class II alleles), among which the most conspicuously diverse loci, HLA-A, HLA-B, HLA-C and HLA-DRB, have 893, 1431, 569 and 814 known gene alleles that can encoding 681, 1165, 431 and 637 proteins respectively (Table 1-1).

Table 1-1 HLA alleles numbers (adapted From IMGT/HLA Database, assigned October 2009)

Gene	*HLA class I*			*HLA class II*					
	A	**B**	**C**	**DRA**	**DRB**	**DQA1**	**DQB1**	**DPA1**	**DPB1**
Alleles	893	1,431	569	3	814	35	106	28	136
Proteins	681	1,165	431	2*	637	26	77	16	118

Gene	*HLA-DRB Alleles*								
	DRB1	**DRB2**	**DRB3**	**DRB4**	**DRB5**	**DRB6**	**DRB7**	**DRB8**	**DRB9**
Alleles	722	1	52	13	19	3	2	1	1
Proteins	572	0	42	7	16	0	0	0	0

* The two DRA proteins have only one difference at the cytoplasmic tail.

The evolutionary force that has created and maintained such astounding allelic diversity is postulated to be balancing selection[52, 53], a broad term that identifies any kind of natural selection in which no single allele is absolutely most fit. This striking feature of MHC genes as well as proteins not only creates fertile grounds for evolutionary biologists but also provides resourceful platforms for protein engineers. Although a high degree of MHC polymorphism is found in human population, an individual can only express approximately 18 class I or class II proteins, which are usually individually specific, so transplantation rejection and autoimmune diseases might occur when trying to transplant organs or tissues between individuals. Because of these limitations, immunologists and bioengineers would always like to investigate the common and distinct molecular properties of MHC proteins and develop methods for monitoring the evolution of MHC molecules for improved immune tolerance as well as diversity. The

in vitro methodology we will talk about in this thesis allows discovery of the molecular relationship among MHC-II alleles in a functionality-dependent manner.

1.3.2. Structure of MHC-II proteins

An MHC-II protein is a non-covalently associated heterodimer consist of an α chain (32-34kDa) and a β chain (29-32kDa). Each subunit contains two extracellular domains[54-59] (Figure 1-3A), a trans-membrane spanning domain and a cytoplasmic tail. The N-terminal α1 and β1 domains interact to form the peptide-binding groove, composed of two α-helical walls and a floor formed by eight strands of anti-parallel β-sheets (Figure 1-3B). Polymorphisms of MHC-II are mainly due to the existence of various possible residues at the same position within the peptide binding groove[15]. In contrast with class I, the class II molecules leave both ends of the peptide-binding groove open for accommodating longer peptides up to 30 residues[59, 60]. This peptide binding domain is responsible for loading peptides and presenting them to T cells[61]. The α2 and β2 domains fold into immunoglobulin (Ig) domains, responsible for CD4 coreceptor binding and are relatively conserved among various alleles of a particular gene[62].

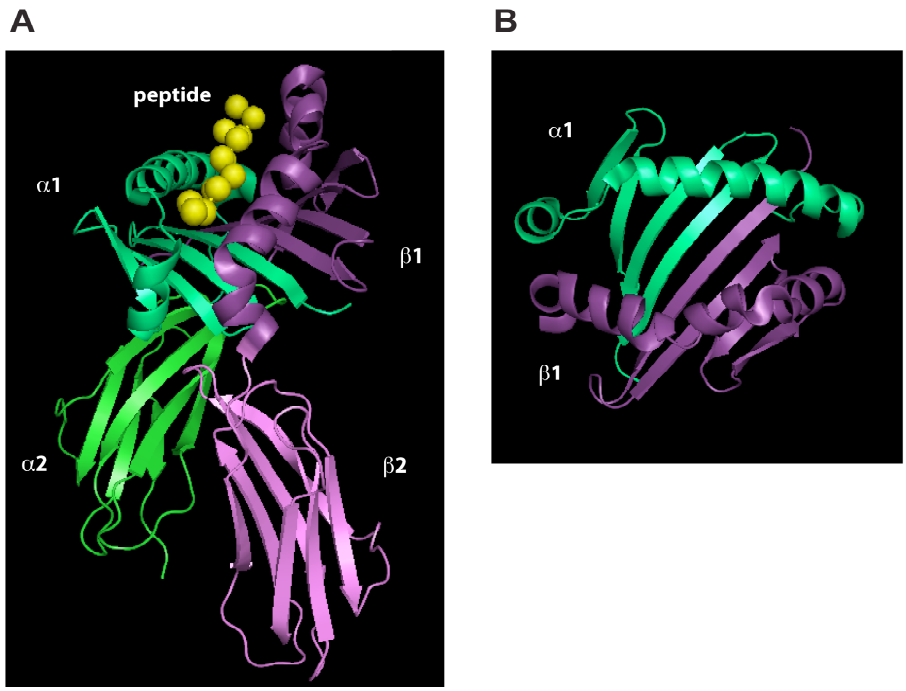

Figure 1-3 Three-dimensional structure of human leukocyte antigen-DR1 in complex with FLU peptide[54]. **A.** Extra-cellular domains of DR1 associated with FLU peptide (derived from influenza hemagglutinin residues 306-318: PKYVKQNTLKLAT, also known as HA306-318 elsewhere[54]; and we use "FLU" henceforward in order to distinguish it from another HA epitope tag derived from the same protein). **B.** Top view of the peptide-binding site formed by α1 and β1 domains. Images were made using PyMOL software (Delano, W.L., **The PyMOL Molecular Graphics System** (2002) on world wide web http://www.pymol.org) by rendering molecular coordinates in the code file 1DHL.

1.3.3. MHC-II pathway of antigen processing in APC

The maturation of MHC-II and the peptide binding and transport by MHC-II take place inside professional APCs upon the exposure to foreign antigens[29, 63] (Figure 1-4). After capturing antigens, these specialized APCs will internalize the extracellular proteins into vesicular compartments, such as endosomes and lysosomes, where proteolytic enzymes degrade antigens into short fragments (peptides) under the acidic condition. Newly assembled MHC-II in endoplasmic reticulum (ER) are transported through Golgi complex to late endosomes and lysosomes with the help of invariant chain (Ii), which weakly occupies the peptide-binding groove of immature MHC-II. After

membrane fusion of different endosomes and lysosomes, MHC-II and antigenic peptides enter the same vesicle, called the MHC class II compartment, or MIIC[64], which contains all components required for peptide/MHC-II association such as enzymes, MHC-II, the Ii (or invariant chain-derived peptides), and HLA-DM (product of a non-polymorphic gene located at the locus of MHC-II gene cluster) with the ability to catalyze peptide binding[65]. To remove Ii from the peptide-binding site, proteolytic enzymes first cleave Ii into class II-asociated invariant chain peptide (CLIP), and then DM helps to catalyze the conformational change[66] of MHC-II for peptide exchange between CLIP and high affinity antigenic peptides. Once mature MHC-II molecules are associated and stabilized by antigenic peptides[66-69], the peptide/MHC-II complex can be delivered by exocytic vesicles to the surface of APC for recognition by CD4+ T cells.

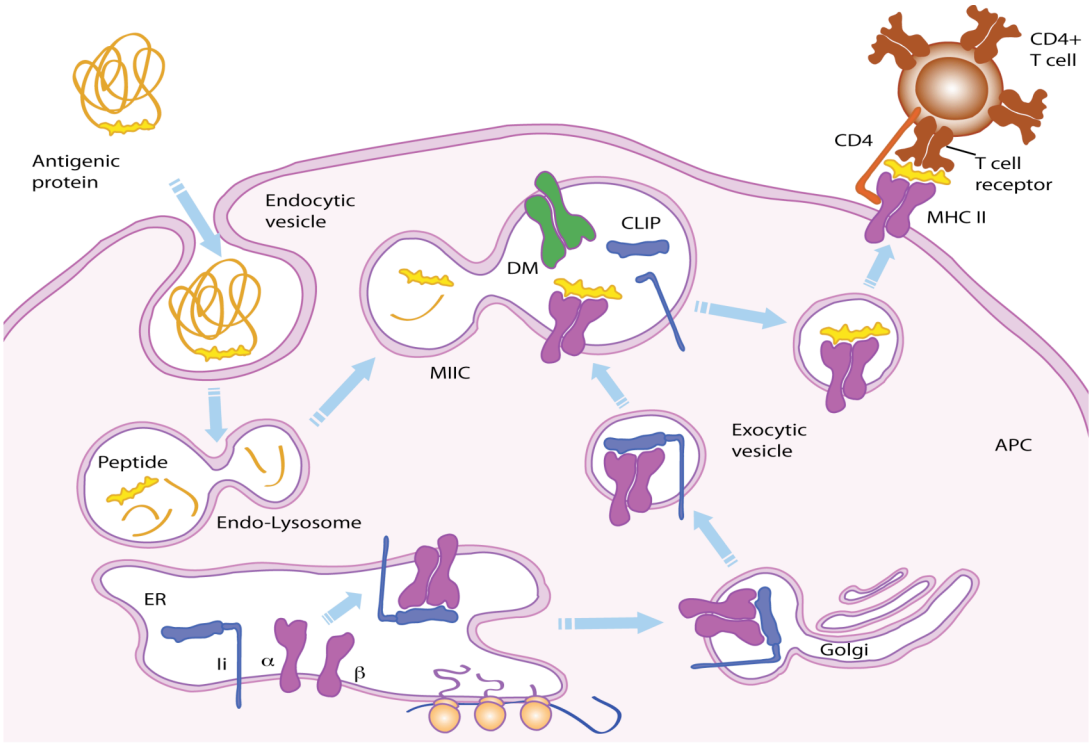

Figure 1-4 Peptide processing in APC via MHC-II maturation pathway.

1.3.4. Peptide binding characteristics of MHC-II

There are several important features of the interaction between MHC-II proteins and specific peptides[21, 70]. First of all, even though each MHC-II molecule contains a single peptide binding groove capable of accommodating one peptide at a time, it has the ability to specifically recognize and present a large set of peptides derived from enormous numbers of antigens[18-20]. This explains why each individual expressing only a few different MHC-II alleles (normally less than 12) could defend against infections of various foreign agents.

Secondly, peptides bound to MHC-II share common structural features that favor their interaction. The optimal length of specific peptides is between 12 to 16 residues, with 9 consecutive residues directly contacting with the binding groove[59]. Some have their side chains pointing down to pocket-like region within the groove serving as anchors for MHC-II binding, whereas the rest face up with their side chains away from the groove for recognition by TCRs (Figure 4-1). This feature reflects the antigen specificity of both MHC-II and TCRs.

Thirdly, the association is relatively stable with a very low off-rate and the peptide binding affinity of MHC-II is higher than the peptide/MHC-II complex binding affinity of TCR[25]. This allows peptide/MHC complexes to persist long enough on the surface of APCs for T cell recognition.

Fourthly, MHC-II shows a broader specificity (promiscuity) for peptide binding than the specificity of antigen recognition by TCRs[25]. They can bind different sources of peptides derived from both self and foreign antigens, which TCRs will distinguish for T

cell stimulation. This actually increases the chance of autoimmunity and the tolerance of self-like foreign antigens.

All the information about peptide/MHC-II interactions may be used in vaccine design for enhancing CD4+ T cell-mediated adaptive immunity or in immunotherapy for treating T cell-related diseases[26-28, 38, 46-49, 71, 72]. However, most of these strategies based on applying known T cell specific epitopes presented by self MHC-II molecules for stimulating CD4+ T cell activation, which are limited by the restricted set of epitopes recognized by the few types of self expressed MHC-II alleles found in each individual. Therefore, there will always be some nonspecific epitope or mutated epitope bearing antigens with the capability of escaping the recognition of MHC-II as well as CD4+ T cells and causing sever immune diseases to certain MHC-II alleles expressing individuals. In these cases, immunotherapeutic methods relying on MHC-II-specific antigens will become less effective, and alternative strategies such as improving the range of peptides presentable by MHC-II or changing the specificity of certain alleles could be useful.

Even though each individual could only bear a few MHC alleles and will inevitably come across with the problem of transplantation rejection when a foreign allele showing up in their bodies, hundreds or even thousands of alleles do exist in nature among different individuals and their protein products share similar structures and functions. This large freedom of the balancing selection for MHC-II in nature suggests that any self-MHC allele has the potential of being mutated or modified subtly without affecting most of its characteristics, possibly including self-tolerance. Furthermore, if a selection mechanism is properly set up, the peptide binding properties such as specificity and

affinity of an MHC-II molecule should be able to be "directedly" edited or evolved by introducing subtle changes such as single mutations in the protein sequence leaving most of the other properties and functions unchanged. Creating and applying this kind of self-derived or self-evolved tolerant MHC-II molecules in the development of vaccines and aAPCs would become a breakthrough in vaccination and immunotherapy.

In this thesis work, we will take the first step of this big project on evaluating the possibility of developing novel MHC-II molecules with minimum residue substitutions, which can specifically recognize and present targeted peptides for T cell activation.

1.4. Protein engineering technologies in this thesis work

1.4.1. Cell-surface display for target proteins

Display of recombinant proteins on the surface of cells or viral particles not only serves as an expression system for correct assembly of proteins but also provides a platform for directly exploring and improving protein functions such as immunogenicity, ligand binding specificity and enzymatic activities, etc. by directed evolution[73-89]. Up till today, various cellular systems such as bacteriophages, bacteria, yeast, baculovirus-insect cells and baculovirus/retrovirus-mammalian cells have been developed and widely used for displaying and engineering proteins of interest. There is no expression system perfectly suitable for all kinds of *in vitro* engineering applications, so advantages and disadvantages usually need to be considered and compromised for acquiring a specific accomplishment.

Phage display[90, 91] and bacterial display[79, 80, 92] are largely applied for expressing short peptides and simple polypeptides, which require more relaxed folding

environments and less modifications, because these prokaryotic organisms are normally lack of the post-translational machinery that helps properly assembling eukaryotic especially human proteins. However, the simplicity of culturing and cloning these lower hierarchy single-cell systems make them priorities for engineering a great number of well studied eukaryotic polypeptides by construction and screening of libraries of variants with a size of 10^8~10^{10}. Investigation of large pool of variants and screening for better fitness has become more and more useful via directed evolution for understanding function of various proteins nowadays, thus, other than cellular platform based displaying system, ribosome display[93] or even mRNA display[94-96] have been derived for expression of even larger libraries for protein engineering in a cell-free environment without limitation of transfection and/or transformation efficiency. The latter two systems are very useful for screening a large variety of short peptides in some cases, but less preferred for polypeptides, whose function relies on tertiary structure.

On the other hand, insect cell display[76, 97] and mammalian cell display[75, 82, 98, 99] inherit the advantage of eukaryotic expression systems, which provide more accurate protein folding regulatory and auxiliary mechanism and better similarity to *in vivo* conditions. As a result, large membrane proteins with sophisticated structures or membrane protein complexes consist of multiple non-covalently associated monomers are better to be studied within these higher-level organisms. For the cost, one must sacrifice diversity and the number of protein variants that can be tested, or, labor intensity as well as time and expense cost has to be increased dramatically to obtain similar library of protein mutants for screening novel functionality, which sometimes is difficult and impossible for these multicellular organisms even with modern

biotechnology. Additionally, even with the help of baculoviral[100] or retroviral[89] technology, the difficulty of applying mammalian cells to engineer heterologous proteins make them less commonly used than other displaying systems. The ability of producing viral particles (e.g. baculovirus or retrovirus) for displaying interest polypeptides though provides more flexibility to extend the application of these two kinds of virus infected eukaryotic displaying systems[84, 89].

In a lot of cases, diversity and complexity of certain eukaryotic protein are both desired when performing specific structure-function analysis, such as the one in this thesis work; hence, yeast display[73, 74, 101] represents an ideal *in vitro* research platform, which has the ability to translate and modify proteins under control of eukaryotic expression though a little different from mammalian cells, while lose little molecular cloning advantage as a single cellular expression system. Theoretically, one can generate a combinatorial library with up to 10^9 individual clones using yeast, which is able to cover combinations of 20 standard amino acids in 7 different sites within a relatively complex protein. It is for sure that this single-cell eukaryotic displaying system has its potential on occupying more and more research fields in protein engineering.

In this thesis, we are going to display proteins of interest on the surface of yeast and extend the display technology to fully utilize yeast to characterize interaction between two or more proteins, e.g. peptides and DR1 heterodimer.

To display a protein of interest on the surface of yeast, the gene coding for protein of interest can be cloned to upstream[102] or downstream[74] of *AGA2* gene in a yeast shuttle vector. After being transformed into an engineered yeast *Saccharomyces cerevisiae*

strain EBY100[74] (***a*** *GAL1-AGA1::URA3 ura3-52 trp1 leu2Δ1 his3Δ200 pep4::HIS2 prb1Δ1.6R can1 GAL*), heterologous genes in the vector will be translated to protein of interest fused by Aga2 protein (Aga2p), which is the subunit of a native yeast surface protein, **a**-agglutinin[103, 104]. Directed by a signal peptide, the protein fusion will enter protein secretory pathway, where Aga2p will associate with the other subunit of **a**-agglutinin (Aga1p) by forming two disulfide bonds. Once transported outside yeast, the assembled **a**-agglutinin will anchor protein of interest on the surface by the designed covalent linkage between Aga2p and protein of interest (Figure 1-5).

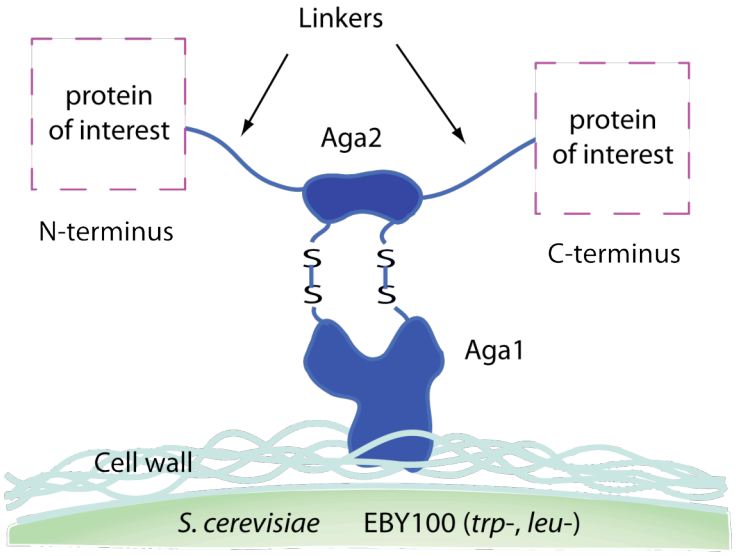

Figure 1-5 Classical yeast surface display of target protein. Protein of interest can be covalently linked to either N-terminus or C-terminus of Aga2p, which serves as the anchor for surface-displaying protein of interest.

1.4.2. Analyzing immunofluorescently labeled cells by flow cytometry

Detection of protein on yeast surface is mainly carried out by applying immunofluorescenty labeled yeast to a fluorescence detector – flow cytometer. To label

proteins on the surface of yeast, cells are usually enriched in growth media and then switched to induction media for a certain time for optimal protein expression. Consequently, cells are collected by centrifugation and labeled with a primary labeling reagent specific for protein of interest displayed on yeast surface. In some cases, the primary reagent is conjugated with a fluorophore for direct fluorescent labeling. In other cases, a secondary reagent conjugated with a fluorescent dye is used for indirect fluorescent labeling. These conjugated fluorescent dyes can be detected via flow cytometry. All labeling work performed in this thesis work used indirect labeling, which normally gives the stronger signal.

Other than proteins of interest themselves, small epitope tags such as HA, V5, 6His, c-Myc, etc., are wildly used and fused to nonfunctional ends of target protein enabling indirect checking of expression and presentation of protein of interest on the surface of yeast. Due to the small size (5-20 residues), epitope tags will not affect the tagged protein's biochemical properties, but can be recognized by epitope-specific antibodies, which also allows immunofluorescent labeling. Fluorescent labeling can be performed for one epitope or more at the same time as long as primary and secondary reagents have lowest cross-reactivity against each other. Double labeling by using an antibody specific for protein of interest and another one specific for a fused epitope tag was used frequently in this thesis work.

Flow cytometer allows suspension of labeled yeast cells flowing dropwise through a detecting area at a high speed, where fluorophores coupled on yeast surface can be excited by a laser beam, and emit fluorescent light with certain wavelength, intensity of which can be recorded in fluorescent detector with a specific wavelength cut-off half

mirror. At the same time, forward light scattering (FSC) and 90° side light scattering (SSC) can also be observed, which correlate with volume of yeast cells and their intracellular complexity, respectively. The dot-plot of these two parameters always shows a certain pattern (e.g. Figure 3-1 (i)), which is significantly different from that of bacteria or mammalian cells, allowing a first validation for properly cultivation of yeast and an easy checking for contamination.

Sometimes, yeast displaying target protein is desired to be recovered from a large pool of yeast cells, a more developed flow cytometer, called cell sorter, will make this possible. Cell sorter can isolate individual cells via an electronic field and direct them into the corresponding collection tubes. This process is usually called fluorescent-activated cell sorting (FACS).

After flow cytometry or cell sorting, flow cytometric data can be analyzed in different plots. The simplest is using a one-dimensional histogram by cell events (or cell numbers) vs. fluorescence intensity of the fluorophore used for labeling protein of interest (e.g. Figure 2-4 (i) and (ii)). This histogram will normally exhibit a distribution of cell numbers at different intensities, the mean of which statistically represents how well the labeled protein is displayed on the surface of yeast. Usually, yeast cells themselves will give an auto-fluorescent background in most flow data, which can be determined by preparing a few unlabeled or irrelevant yeast controls in parallel. If double labeling of two proteins on yeast surface is performed, other than histograms, a two-dimensional dot-plot can be generated by plotting the fluorescence intensity of one fluorophore against the other (e.g. Figure 2-4 (iii)). The dot-pot will normally display cell populations with different fluorescence intensities, where information of two

proteins or epitopes displayed on yeast surface can be examined simultaneously. It is critical to set up appropriate compensation when running a double-labeled sample, because a lot of times, the emission from one fluorophore can overflow into channels used to detect other fluorophores, especially for some strong signals. In those cases, single labeled controls need to be used to adjust the compensation so that no signal interfering occurs.

Most recent flow cytometers are equipped with at least two lasers, which can excite fluorophores with excitation wavelengths in a fairly large range, thus, if properly selected, two fluorophores with excitation wavelengths far from each other can be used for double labeling and excited by two lasers respectively on flow cytometer without any compensation. This will greatly increase the efficacy and accuracy of flow detection. In this thesis work, we are going to take advantage of this feature so that no compensation is needed most of times.

Chapter 2. Development of yeast co-display – a novel methodology for characterizing MHC-II/peptide binding

2.1. Introduction

Antigen specificity of CD4+ T cell-mediated immune response is primarily determined by peptide presented by MHC-II molecules on the surface of APCs. In depth characterization of peptide binding by MHC-II is critical to understanding issues in vaccine design, autoimmune disease, infectious disease progression, and transplantation rejection[28, 46, 72].

MHC-II is a transmembrane protein, whose extracellular portion consists of two heterologous chains, α and β, each of which containing two domains[54]. Mature MHC-II proteins function to capture antigenic peptides processed inside professional antigen-presenting cells (APCs) and present them on the surface of APCs for recognition by CD4+ T cells to mediate adaptive immunity. The peptide-binding site of MHC-II formed by α1 and β1 domains contains several pocket-like regions, which prefer to accommodate specific side chains of anchor residues on peptides[58, 59]. The engaged peptides are always heterogeneous in size with both termini extended beyond the binding groove[60]. Even though a lot of common features for peptide binding have already been discovered, the majority of peptide binding motif and anchor preference information for distinct alleles of MHC-II still needs to be further characterized so that specific target peptide/MHC-II pairs can be applied for immunotherapeutic application via activating specific T cell clone.

MHCs are the most polymorphic glycoproteins known in nature, out of which each individual only expresses a few alleles, so it is challenging to investigate peptide-binding properties of all these trans-membrane proteins *in vivo* or *ex vivo*. Therefore a variety of *in vitro* methods have been developed for studying the interaction between peptide and MHC-II. A routinely used *in vitro* approach entails purifying soluble recombinant MHC-II molecules from different expression systems such as B cell lines[105], insect cells[106], yeast[107], or *Escherichia coli*[108-112] and then characterizing binding of these molecules to different peptides generated either chemically by solid-phase synthesis[113, 114] or genetically by cell or non-cell display technologies. The former one using synthesized peptides enables coupling of all kinds of quantitative assays for acquiring binding affinity and kinetics[18, 112, 115-123], whereas the latter allows construction of big libraries of peptide sequences for a high throughput screening and investigation[90]. However, the labor-intensive preparation of soluble MHC-II proteins and lengthy binding assays limit the efficacy and throughput of these methods for mapping MHC-II binding specificities across the large number of existing alleles. Other attempts focused on assaying binding of soluble peptide by MHC-II expressed natively on mammalian cell surface[124-128], which still did not completely get away from the difficulty of manipulating MHC-II proteins.

Alternatively, engineered cell-surface display systems such as phage[129-131], yeast[102, 132-134], baculovirus-insect cells[76], and mammalian cells[98, 135] have been tried for displaying recombinant extracellular domain of MHC molecules, suggesting a potential to engineer MHC-II proteins by directed evolution such that any possible allele can be examined for peptide binding. However, to stabilize the recombinant MHC proteins, an

antigenic peptide was always genetically introduced by covalent linkage, which not only brings in unknown artificial effects but also restrain the flexibility of using these methods to quantify peptide binding. Sometimes, acid/base half of leucine zipper was attached to C-terminus of either chain of MHC-II to facilitate protein association and surface anchorage[76], which was not necessary in other display system but could alter MHC-II native folding. These limitations make current cell-surface display system less proper for characterizing the interaction between peptides and MHCs. The lack of a rapid, efficient, robust, and quantitative methodology for characterizing the peptide binding specificity and promiscuity of MHC-II alleles remains a bottleneck. Several two-hybrid expression systems[76, 81, 88, 102, 136-141] suggest a way to manipulate multi-proteins in their native forms within the same cell and a possibility to improve current cell-surface display technology to study interaction between peptide and MHC-II.

Herein, we report a novel, quantitative, high throughput methodology, yeast co-display (Figure 2-1), for characterizing and engineering peptide-binding specificity of MHC-II. As a model system, the extracellular domains of the MHC-II human leukocyte antigen HLA-DR1 were expressed as a secreted heterodimer in *Saccharomyces cerevisiae.* Simultaneously, an antigenic peptide known to bind to HLA-DR1 was surface-displayed. Intracellular binding between the MHC-II and the peptide antigen thus anchored the soluble MHC-II to the cell surface upon secretion, allowing detection by immunofluorescence. The relative abundance of MHC-II compared to peptide on the cell surface depended on the strength of binding between these species. Accordingly, the relative binding of different peptides and/or MHC-II variants can be assayed by

genetically manipulating either partner, enabling the application of directed evolution approaches for high throughput characterization or engineering.

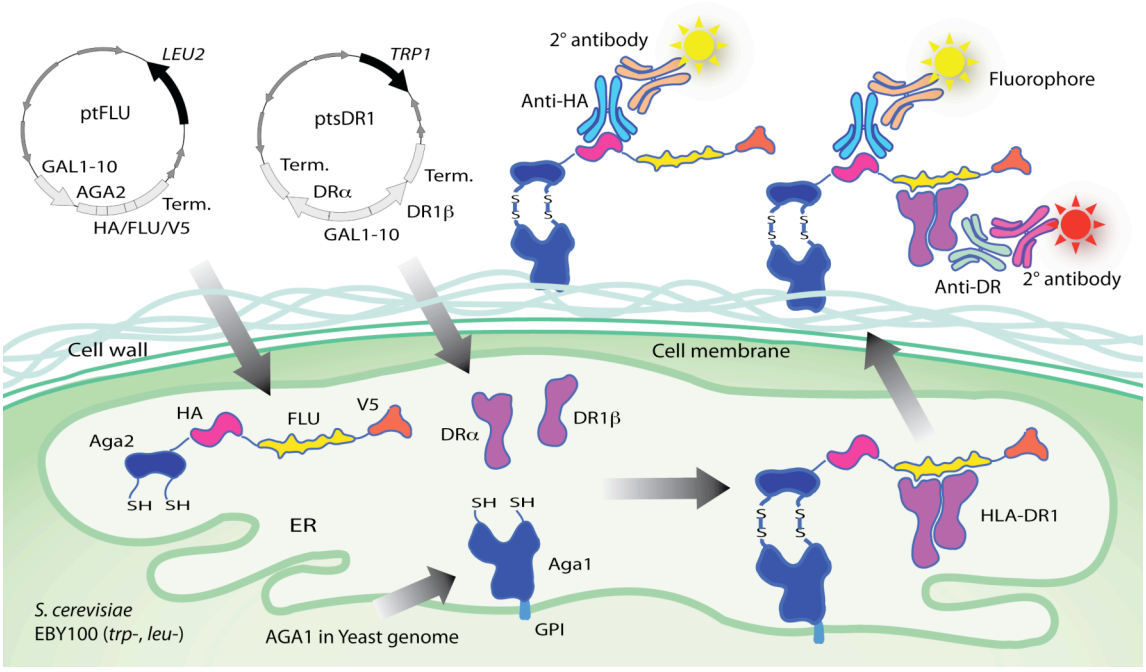

Figure 2-1 Design of the yeast co-display system for HLA-DR1. Two plasmids direct the expression of MHC-II ectodomain heterodimer and a peptide respectively in *Saccharomyces cerevisiae* yeast stain EBY100 (*URA+, trp-, leu-*). Aga2p-fused FLU peptide (derived from influenza hemagglutinin residues 306-318: PKYVKQNTLKLAT) is expressed flanked by HA tag (influenza hemagglutinin 98-106: YPYDVPDYA, different from FLU) and V5 tag (short peptide: GKPIPNPLLGLDST), enabling detection of peptide levels by immunofluorescent staining with antibodies specific for either tag. The α and β chain of HLA-DR1 ectodomain are expressed from separate cassettes and assembled in endoplasmic reticulum (ER) as a soluble non-covalently associated heterodimer. FLU peptide is anchored to the cell surface via native processing and secretion of the assembled **a**-agglutinin protein (composed of the Aga1p and Aga2p subunits) as described[74], and the HLA-DR1 heterodimer is anchored by noncovalent binding to FLU. Relative fluorescence levels of different fluorophores coupled to anti-tag and anti-DR reagents indicate the level of saturation of available peptides by bound HLA-DR1. GPI: glycosyl phosphatidylinositol, an anchor helping **a**-agglutinin to transport through the secretory pathway[104].

2.2. Materials and Methods

2.2.1. Construction of plasmids

Primers or oligonucleotides were synthesized commercially either by Invitrogen (Carlsbad, CA) or by Integrated DNA Technologies, Inc. (Coralville, Iowa). Other than

stated, most enzymes were purchased from New England Biolabs (NEB, Ipswich, MA). Taq (Qiagen, Germantown, MD or NEB) or Vent DNA polymerase catalyzed amplification of polymerase chain reaction (PCR) products. Enzymatic digested PCR products and plasmids were all confirmed by gel electrophoresis and purified before ligation, which was catalyzed by T4 DNA Ligase (NEB or Invitrogen). Either QIAquick Gel Extraction Kit (Qiagen) or Wizard SV Gel and PCR Cleanup System (Promega, Madison, WI) was used for purification of DNA or PCR products without significant differences. Either chemical- or electro- transformation in *Escherichia coli* strain DH5α (Invitrogen) was performed for cloning and amplifying plasmids, which can be isolated by using QIAprep Spin Miniprep Kit (Qiagen) and confirmed by DNA sequencing (DNA Sequencing Facility, Department of Genetics, University of Pennsylvania, PA or Molecular Biology Resource Facility, Division of Biology, University of Tennessee, TN)

- Plasmids for classical surface display of HLA-DR1

The gene expressing the extracellular domain of HLA-DR1 β chain allele *DRB1*010101* (1-190) was cloned from plasmid pDLM1-drb1s (kindly provided by Dr. L. Sten, Umass) by PCR using N-terminal primer W3 (5'–TAATCCCGGGGACCTAA GTATGTGAAGCAGAATACACTGAAGCTGGCAACCGGAGGTGGTTCACTAGT GCCACGGGGCTCTGGAGGAAAGCTTGGAGACACCCGACCACGTTTCTT) and C-terminal primer W4 (5'–TACTGGATCCAGAACCACCACCACCAGAACCACCAC CACCAGAACCACCACCACCGCTAGCTGCTCTCCATTCCACTGTGAG), such that sequence encoding the 13-residue FLU peptide (derived from influenza hemagglutinin residues 306-318: PKYVKQNTLKLAT) and a thrombin recognition site were added at

the N-terminus. The PCR product was ligated into plasmid Z47[102] in place of the original expression cassette coding for allele *DRB1*0401* via annealing of sticky ends obtained by XmaI/BamHI double digest, to create pfluDR1 for surface-displaying FLU-fused HLA-DR1 (Figure 2-2). Another vector pDR1 expressing "empty" HLA-DR1 was constructed likewise but without adding FLU peptide by using another N-terminal primer W1 (5'–AGCTCCCGGGGAGACACCCGACCACGTTTCTT) instead of W3.

- Plasmids for classical surface display of FLU peptide

The entire expression cassette (GAL1-10//AGA2//HA//scFv 4-4-20//MFα Term.) was excised from pCT302[74] by double digest using KpnI and SacI and cloned into KpnI/SacI partially digested yeast shuttle vector pRS315[142] to create a plasmid with a *LEU2* selectable marker. Expression cassette (scFv 4-4-20 //MFα Term.) within NheI and SacI sites in this plasmid was substituted by a PCR product encoding (FLU//SphI//MFα Term.), obtained by using N-terminal primer W20 (5'–TCTAGCTAGCCCTAAGTATGTGAAGCAGAATACACTGAAGCTGGCAACCTAAGCATGCCAACAGTGTAGATGTAACAA) and C-terminal primer W21 (5'–CTGTGAGCTCAATTCTCTTAGGATTCG), to construct pFLU. Similarly, W34 (5'–TCTAGCTAGCCCTAAGGCGGTGAAGCAGAATACACTGAAGCTGGCAACCTAAGCATGCCAACAGTGTAGATGTAACAA) was used instead of W20 for constructing pFLUA, which directs expression of a FLU analogue with Tyr308 substituted by Ala (encoded by underlined GCG) at the putative anchor position P1 (refer to Figure 4-1).

To add V5 epitope tag at C-terminus of FLU or its analogues, an oligonucleotide encoding FLU followed by a short spacer (GGGP) and V5-tag was generated by PCR

using N-terminal primer 0118 (5'–AAGGACAATAGCTCGACGATT) and C-terminal primer W48 (5'–TCTAGCATGCTTAAGTAGAATCGAGACCGAGGAGTGGATTTG GTATAGGCTTCCCGGGTCCACCTCCGGTTGCCAGCTTCAGTGTATTC), and inserted into vectors pFLU and pFLUA via annealing of sticky ends generated by PstI/SphI double digest, for constructing ptFLU (Figure 2-6A, construct a) and ptFLUA respectively.

In order to clone FLU to the N-terminus of Aga2p, a restriction site EcoRI in ptFLUD was changed to AatII by site directed mutagenesis for future cloning. Briefly, A pair of reverse complementary oligos W95 (5'–ACCCCGGATCGACGTCCCTACTT CAT) / W96 (5'–ATGAAGTAGGGACGTCGATCCGGGGT) carrying the mutagenic site was designed to prime the synthesis of mutant strand by PCR following the manufacturer's recommended procedures in QuikChange II site directed mutagenesis kit (Stratagene, La Jolla, CA). After being digested by DpnI for 1 h, mutant plasmids were transformed into electro-competent DH5α using the MicroPulser Electroporation Apparatus (BIO-RAD, Hercules, CA) and recovered by Miniprep kit. Following that, an AatII/SphI fragment was amplified from a modified surface display plasmid (will be referred to again in next paragraph) by W97 (5'–CGTCGACGTCATGAAGGTTTTGA TTGTC) and W98 (5'–ATGAGCATGCTTAAGCGTAGTCTGGAACG) and inserted into AatII/SphI sites of the mutant plasmid for introducing an expression cassette (Syn-Pre-Pro leader//XmaI//an irrelevant protein//speI//6His//AGA2//HA), which allows target protein followed by a polyhistidine tag to attach to Aga2p N-terminus. Finally, gene encoding the irrelevant protein was replaced by a PCR product encoding FLU

peptide, generated using primers 0115 (5'–CAACAAAAAATTGTTAATATACCT) and W106 (5'–TAATGACTAGTGGTTGCCAGCTTCAGTGTATTCTGCTTCACATACT TAGGACCCCGGGCTTCTCTCTTGTCCAAAG), to construct ptNFLU (Figure 2-6A, costruct b).

- Plasmids for expressing soluble HLA-DR1

Z47 was modified first for displaying an irrelevant heterodimeric protein. In the resulting plasmid, XhoI recognition site at the downstream of gene cassette promoted by GAL1 promoter was removed and replaced by a Myc-tag (EQKLISEEDL) flanked by two unique restriction sites NcoI and SphI, thus, the other XhoI left in the plasmid also became unique for later cloning work. Two expression cassettes (*DRB1*010101* (1-190)//Flag-tag (DYKDDDDK)) and (*DRA*0101* (1-192)//6His) were synthesized by PCR using corresponding primer pairs W38 (5'–CACGCGGCCGATCAAAGAAGAACATGTGATCATCC) / W43 (5'–TACTCCATG GGTGGTGGTGGTGGTGGTGACCACCGTCGACGTTCTCTGTAGTCTCTGGGAG) and W1 / W44 (5'–TAGACTCGAGTTACTTATCGTCATCATCCTTGTAATCTCCA CCACTAGTTGCTCTCCATTCCACTGTGAG), and cloned into the modified plasmid sequentially via annealing XmaI/XhoI and EagI/NcoI digested cohesive ends respectively, to create plasmid ptsDR1 (Figure 2-6B).

2.2.2. Buffers and media for yeast cultivation and treatment

Other than stated, most chemicals and reagents mentioned below were purchased from Fisher Scientific (Fair Lawn, New Jersey). Buffers used in this chapter include: phosphate buffer (5.4 g/l Na_2HPO_4, 9.70 g/l $NaH_2PO_4 \cdot 2H_2O$, pH 6.0); citrate buffer

(14.7 g/l sodium citrate (Sigma-Aldrich Inc., St. Louis, MO), 4.29 g/l citric acid monohydrate, pH 5.0); phosphate buffered saline (PBS; 137mM NaCl, 2.7mM KCl, 10.1mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4); Tris buffered saline (TBS; 137 mM NaCl, 20mM Tric-Cl, pH 7.6); Factor Xa buffer (100mM NaCl, 2mM CaCl, 20mM Tris-Cl, pH 8.0).

Glucose (Dextrose) is used as carbon source in all growth media: rich medium such as YPD (10 g/l Yeast extract (BD Diagnostics Systems, Sparks, MD), 20 g/l Peptone (BD Diagnostic Systems), 20 g/l Glucose) can be used for enriching all strains without selectivity, which is suitable for preparing competent cells; minimal media allowed yeast growing under different nutrient dropout conditions. For example, SD-CAA (20 g/l glucose, 6.7 g/l Yeast N_2 Base w/o Amino Acids (BD Diagnostic Systems), 5.4 g/l Na_2HPO_4, 9.70 g/l $NaH_2PO_4 \cdot 2H_2O$, 5g/l Bacto Casaminoacids (CAA, Difco)) was used for cells requiring *ura-* and *trp-* condition and SD-SCAA (Bacto Casaminoacids in SD-CAA was replaced by 0.62 g/l *Leu/Trp/Ura* dropout supplement mixtures of amino acids (Clontech, Mountain View, CA)) provided *ura-*, *trp-* and *leu-* condition. D-Tryptophan (Sigma) was sometimes added in above minimal media to make those conditions suitable for *trp-* yeast strains. For induction of protein expression promoted by GAL1-10 promoter, glucose was replaced by galactose as the carbon source, while all other nutrients stay the same. Induction media used in this chapter include rich medium YPG and minimal media SG-CAA and SG-SCAA for corresponding nutrient dropout condition as stated above.

2.2.3. Antibodies and labeling reagents

Mouse anti-DR monoclonal antibody (mAb) L243 (BD Biosciences, San Jose, CA), rabbit anti-HA polyclonal antibody H6908 (Sigma) and mouse anti-V5 Mab (Invitrogen, Carlsbad, CA) were used for specifically labeling HLA-DR1, HA-tag and V5-tag on the surface of yeast. Alexa Fluor 488 or 647 fluorophore conjugated goat anti-rabbit, Alexa Fluor 647 or tandem dye PE-Alexa Fluor 647 conjugated goat anti-mouse (Invitrogen), and R-phycoerythrin (PE) conjugated goat anti-mouse (Sigma) antibodies were used as secondary labeling reagents for primary antibodies from corresponding species. Streptavidin-PE (Sigma-Aldrich) was used to label biotinylated peptides.

2.2.4. Generation and cultivation of yeast strains

Saccharomyces cerevisiae yeast stain EBY100 (*URA+, trp-*, *leu-*) was used as the parent strain to accommodate plasmids with either *TRP1* or *LEU2* nutrient marker. In this thesis work, all plasmids directing expression of HLA-DR1 have *TRP1* marker, and all plasmids expressing peptide fusions contain *LEU2* marker. Plasmids were transformed into electro-competent EBY100 (lab made) by electroporation following 'MicroPulser Electroporation Guide' (BIO-RAD) and transformants were selectively grown up under corresponding nutrient dropout condition on glucose containing minimal medium agar plates (containing 1 M sorbitol) by incubating at 30 °C for 3 – 5 days. For example, EBY100 transformed with pfluDR1, pDR1 or ptsDR1 were selected under *ura-* , *trp-* and *LEU+* condition on SD-CAA medium agar plates. Transformants with ptFLU were able to grown up under *ura-*, *TRP+* and *leu-* condition on SD-SCAA agar plates

with D-tryptophan (TRP; 20μg/ml) added. Co-displaying strains can grow under *ura*-, *trp*- and *leu*- condition on SD-SCAA agar plates.

A single colony was picked from agar plates and inoculated into 2–3 ml corresponding glucose-containing liquid minimal medium and enriched to a cell density around OD_{600} of 3.0–7.0 in a 30 °C shaker for 16–24 h. After that, enough cells were pelleted by centrifugation and transferred to 2–3 ml corresponding galactose-containing liquid medium to an initial OD_{600} around 1.0 to induce the expression of target proteins in yeast. Induction using minimal medium was usually performed at 30 °C for 16–24 h, while that using rich medium was usually carried out at 20 °C for 48 h.

2.2.5. Peptide binding assay for classical surface displayed HLA-DR1

Yeast strain EBY100 bearing pDR1 was cultured under *ura*-, *trp*- and *LEU*+ condition as described previously. 4×10^6 galactose-induced yeast were collected by centrifugation and resuspended in 40 μl of one of the following solutions: YPG (pH 6.5), PBS (pH 7.4), citrate buffer (pH 5.0), citrate buffered SG-CAA (pH 5.0, using citrate buffer instead of phosphate buffer). Biotinylated FLU peptide and β_2m peptide (derived from human β_2 microglobulin 52-64: SDLSFSKDWSFYL, an irrelavent peptide specifically bound by class I MHC proteins) (Abgent, Inc., San Diego, CA) were added to a final concentration of 100 μM. Cells were incubated with peptides at 30 °C for 1 day and then washed twice with 500 μl ice cold PBS + 1% (wt/vol) bovine serum albumin (BSA) before staining with streptavidin-PE diluted 1:20 in 100 μl PBS + 1%

BSA on ice for 30-45 min. Cells were then washed with 500 μl ice cold PBS + 1% BSA twice and finally resuspended in 700 μl PBS + 1% BSA for analysis on a BD FACSCalibur flow cytometer (BD Biosciences). EBY100 transformed with pCT302, which will display an irrelavent single chain antibody on the surface, were used as a negative control.

2.2.6. Surface-stripping assay by DTT or Factor Xa treatment

Approximately 10^6 yeast cells per sample were harvested by centrifugation after induction and washed with 400 μl TBS 1–2x. For DTT treatment, pellets were resuspended in 20 μl of reducing buffer [50mM Tris-Cl pH 8.0 containing 0.04mM, 0.2mM, or 1mM DTT (added before use, (Sigma)] and incubated at 4°C for 24 h with gentle shaking. For Factor Xa treatment, pellets were resuspended in 20 μl of Factor Xa buffer with 20 μg/ml Factor Xa protease (NEB) added right before use. One set of samples were incubated at 4 °C for 24 h, another set was incubated at 23 °C for 6 h followed by 4 °C for another 18 h, and a third set was incubated at 23 °C for 48 h. After treatment, cells were collected by centrifugation and washed at 1–2x with 500 μl ice cold PBS + 1% BSA before immunofluorescent labeling.

2.2.7. Immunofluorescent labeling of yeast

Labeling was performed in a 2-step manner. After induction or other treatments, approximately 10^6 cells were collected by centrifugation and washed with 400 μl ice cold PBS + 1% BSA 1–2x. Cell pellet was resuspended in 20–50 μl PBS + 1% BSA with primary antibodies. After staining for 40–60 min on ice or at room temperature,

cells were spun and washed with 400 μl ice cold PBS + 1% BSA again. For the secondary labeling, cells were stained with specific reagents in 40–50 μl PBS + 1% BSA on ice for 30–60 min. A final wash was applied before resuspending cells in 500–700 μl PBS + 1% BSA for flow cytometry.

2.2.8. Flow cytometric analysis of labeled cells

Fluorescently labeled yeast cells were analyzed by flow cytometry, where forward scatter, side scatter and fluorescent signals in corresponding channels (e.g. Alexa Fluor 488 in FL1 or PE in FL2) were recorded by acquiring at least 10,000 cell events. Flow cytometers used in this chapter as well as this thesis work are as follows: Accuri C6 (Accuri Cytometers Inc., Ann Arbor, MI), Beckman Coulter XL (Beckman Coulter Inc., Brea, CA), FACSCalibur, FACSVantage (also a cell sorter), and LSR II (BD Biosciences).

2.3. Results

2.3.1. Classical yeast display of functional HLA-DR1 heterodimer

MHC-II is originally expressed on the surface of APCs for antigenic peptide presentation. To verify that wild-type MHC-II heterodimers can be expressed in a functional form competent for binding to soluble peptides in yeast, the extracellular domain of HLA-DR1 both with and without a covalently fused antigen peptide was surface displayed by fusion to the N-terminus of the endogenous yeast adhesion receptor subunit Aga2p, as described previously for HLA-DR4 (with peptide linked to the β chain N-terminus)[102]. The antigen used in these studies is a 13 amino acid peptide

(designated FLU) from a naturally HLA-DR1-presented fragment of influenza hemagglutinin (HA306-318), a well-studied peptide known to bind tightly to HLA-DR1[59]. Similar to Z47[102], the expression cassettes of plasmids pfluDR1 or pDR1 directs the expression of HLA-DR1 ectodomain on the surface of yeast, with or without FLU fused at N-terminus of β chain (Figure 2-2).

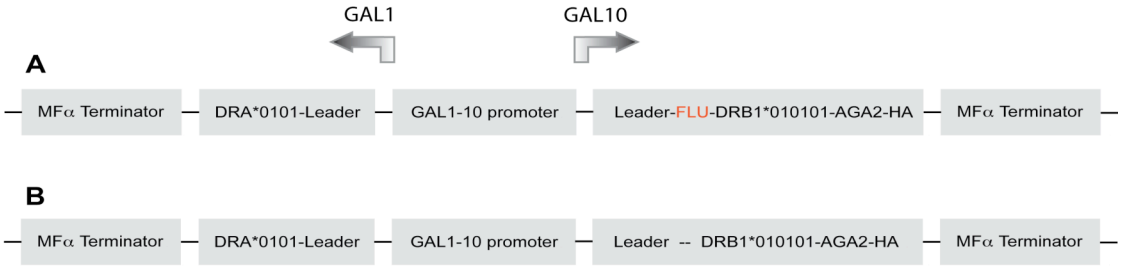

Figure 2-2 Expression cassettes for classical yeast display of HLA-DR1 heterodimer. Galactose inducible bi-directional yeast promoter GAL1-10 initiates the expression of extracellular domains of both chains for HLA-DR1: *DRA*0101* (1-–192) and *DRB1*010101* (1–190) in plasmid pfluDR1 (construct **A**) or pDR1 (construct **B**). The α chain (DRα) is expressed solubly and the β chain (DR1β) is fused to the N-terminus of Aga2p so that once correctly assembled, the non-covalent heterodimer can be led to the secretory pathway and anchored to surface for immunofluorescent detection. Construct **A** with FLU peptide covalently linked to DR1β by a flexible linker is used for displaying FLU/HLA-DR1 complexes, while construct **B** is for displaying "empty" HLA-DR1 on yeast surface.

Saccharomyces cerevisiae yeast stain EBY100 (*URA+, trp-*) was used as parent strain to accommodate either pfluDR1 or pDR1, both containing a *TRP1* nutrient marker, and selectively cultured under *ura-* and *trp-* condition for protein expression. The surface display of peptide-fused or "empty" HLA-DR1 was assessed by simultaneous immunofluorescent labeling with anti-HLA-DRα antibody and antibodies specific for the HA epitope tag included at the Aga2p C-terminus (Figure 2-3, Figure 2-4).

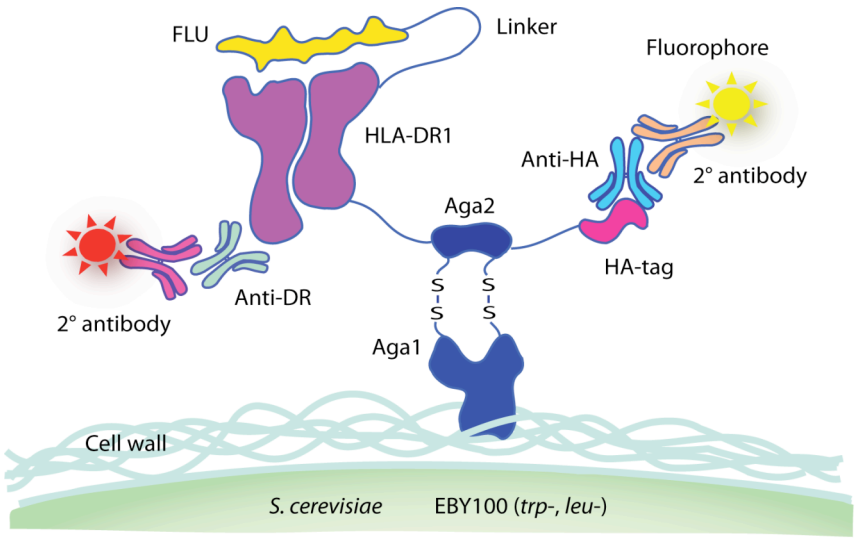

Figure 2-3 Classical yeast display and immunofluorescent detection of HLA-DR1 heterodimer. This is a schematic illustration of yeast surface displayed FLU/HLA-DR1 complexes to Aga2p N-terminus (corresponding to the expression construct **A** in Figure 2-2). The Aga2p C-terminal HA-tag enables immunofluorescent labeling using specific antibodies, ensuring determination and verification of Aga2p fusion expression on yeast surface. Simultaneously, the expression and non-covalent association of soluble DRα and Aga2p fused FLU-DR1β can be validated by functional DR-specific antibodies.

Yeast displaying either form of HLA-DR1 are exclusively HA positive (Figure 2-4(iii)), confirming that detected anti-DR signals correlate to the corrected association of DRα and DR1β or simply the expression of HLA-DR1 heterodimer. The expression level of FLU-fused HLA-DR1 (Figure 2-4B(ii)) is significantly higher than that of the "empty" form (Figure 2-4C(ii)), while the expression level of Aga2p in these two scaffolds are similar, as represented by anti-HA signals (Figure 2-4(i)). Enhanced expression of peptide-fused HLA-DR1 suggests that specific peptides might stabilize MHC-II folding in yeast similarly to in natural APCs[63], while the majority of the empty HLA-DR1 fails to express (~5-fold reduction in cellular fluorescence in Fig. 2-4(ii) and (iii)), consistent with a prior study[134] and putatively due to degradation by the secretory quality control machinery[143].

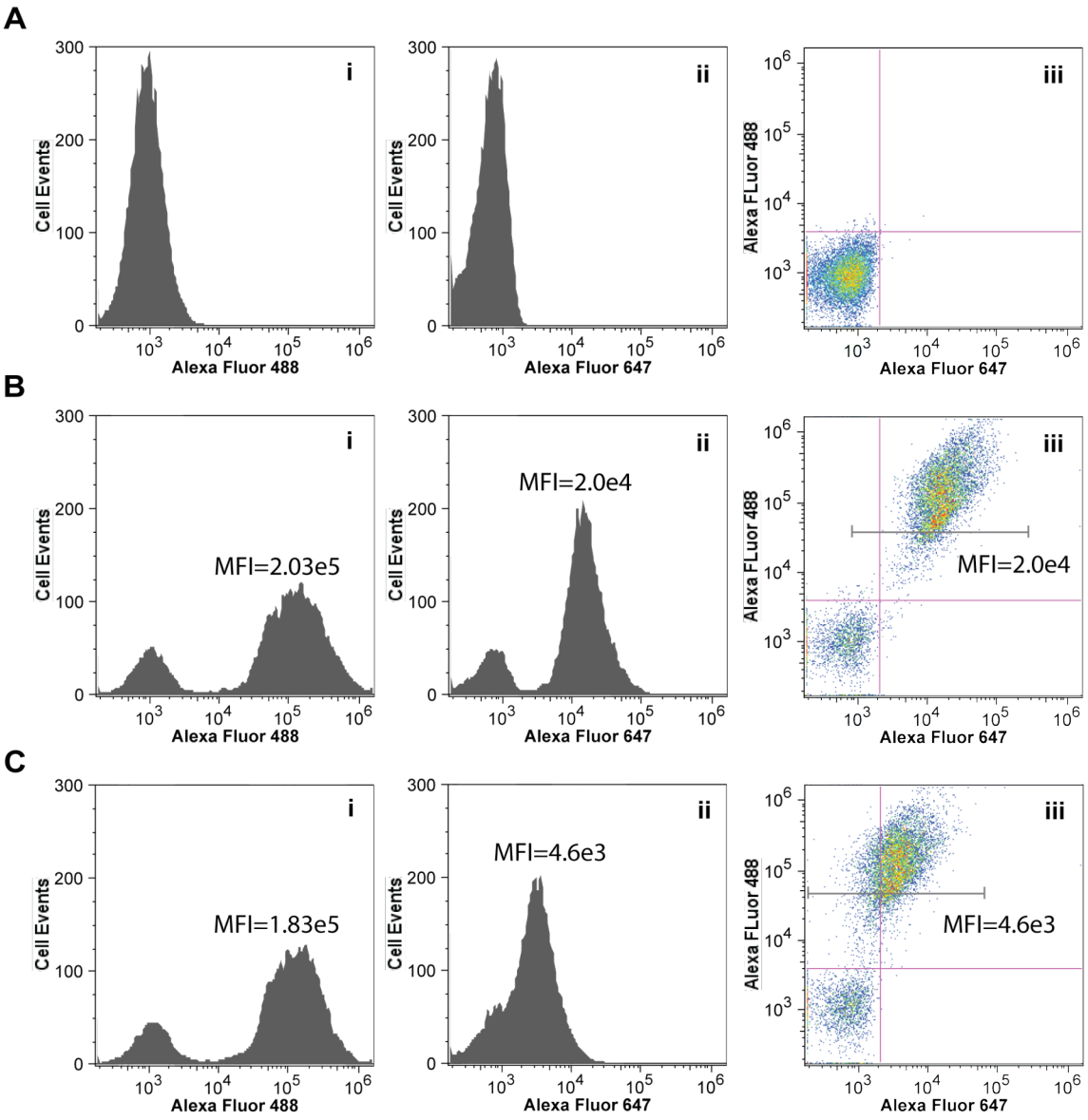

Figure 2-4 Fluorescence intensity from labeled yeast displaying FLU-linked or "empty" HLA-DR1. Aga2p and HLA-DR1 levels of yeast expressing **A.** an irrelevant soluble protein, **B.** Aga2p-fused HLA-DR1 with FLU-covalently linked to DR1β N-terminus, or **C.** "empty" Aga2p-fused HLA-DR1. Cells were double-labeled with H6908 and L243 followed by Alexa Fluor-488 goat anti-rabbit and Alexa Fluor 647-goat anti-mouse highly cross-adsorbed secondary antibodies. Mean fluorescence intensity (MFI) of HA signal and DR signal for HA positive cell population was indicated. Flow cytometric data collected on Accuri C6 were plotted in (i), (ii) univariate histograms, or (iii) two-dimensional dot plots.

Detection of FLU-linked HLA-DR1 on the surface suggests the peptide binding capability of HLA-DR1 expressed in yeast, however, the flexible linker connecting FLU and DR1β (Figure 2-3) might provide an unknown positive effect on HLA-DR1

refolding and FLU association. To validate this functionality, post-induced yeast cells displaying "empty" HLA-DR1 were incubated with synthetic biotinylated peptides and stained with streptavidin-PE before detection by flow cytometry (Figure 2-5). Synthetic FLU peptide bound to HLA-DR1-displaying yeast, but not to a control strain displaying an irrelevant protein, while a control peptide derived from β2 microglobulin demonstrated no binding to either yeast. Similar results were obtained whether binding proceeded in citrate buffer at pH 5.0 or several other buffers or growth media at pH 5.0–7.4 (Figure 2-5), in agreement with the demonstrated pH-independence of peptide recognition by HLA-DR1[144].

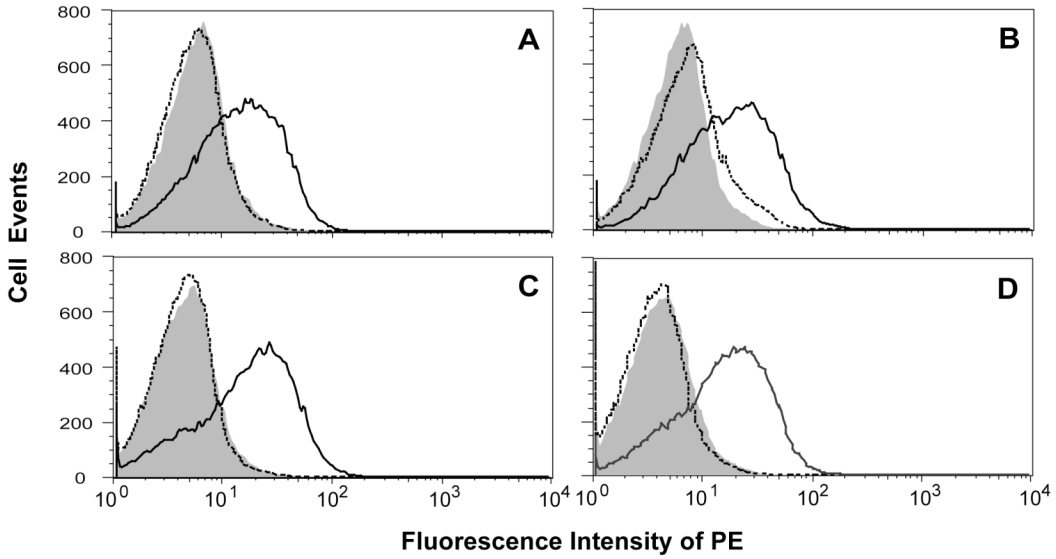

Figure 2-5 Peptide binding capability of classical yeast displayed HLA-DR1 heterodimer. Yeast displaying "empty" HLA-DR1 (solid curve) or an irrelevant protein (shaded) were incubated with 100 μM biotinylated FLU; HLA-DR1-displaying yeast were also incubated with 100 μM biotinylated control peptide (dashed curve). Peptide incubation was carried out in different media or buffers at various pH: **A.** YPG (pH 6.5), **B.** PBS (pH 7.4), **C.** citrate buffer (pH 5.0), **D.** citrite buffered SG-CAA (pH 5.0). All samples were stained with streptavidin-PE and analyzed by flow cytometry on BD FACSCalibur.

Successful expression of functional MHC-II heterodimer as a fusion to yeast surface proteins enables study and modification of this highly polymorphic molecule using

directed evolution or other biomolecular techniques, however, difficulties still exist for characterizing and engineering their functionality, such as peptide binding capability, which require manipulations of two or more proteins inside or outside yeast. For instance, the time and labor intensity and high expense on chemically synthesizing peptides for binding assay significantly limit the variety of peptide sequences tested; the variation on mimicking peptide binding conditions outside yeast (on the surface of yeast) could affect the accuracy of determining peptide binding affinity of MHC-II; the artificial linker in between DR1β and Aga2 might introduce unknown steric impact on the folding of HLA-DR1 molecules; etc. Thus, a way to co-express Aga2p-peptide fusion and soluble heterodimeric MHC-II in yeast, which allows interaction of peptide and MHC-II taking place inside ER and display of the peptide/MHC-II complex on the surface, would be an alternative strategy to overcome most of these problems.

2.3.2. Surface display of FLU peptide and their capability for anchoring soluble HLA-DR1

As mentioned in Chapter I, proteins of interest can be attached to either terminus of Aga2p; hence, two plasmids ptFLU and ptNFLU with a *LEU2* nutrient marker (Figure 2-6A) were constructed: the former one directs expression of Aga2p C-terminal FLU peptide flanked by two epitope tags (HA-tag and V5-tag), whereas the latter one enables FLU peptide attached at N-terminus of Aga2p. In both expression constructs, fused epitope tags allow immunofluorescent labeling by specific antibodies, which indirectly represent the display level of Aga2p-linked FLU peptide.

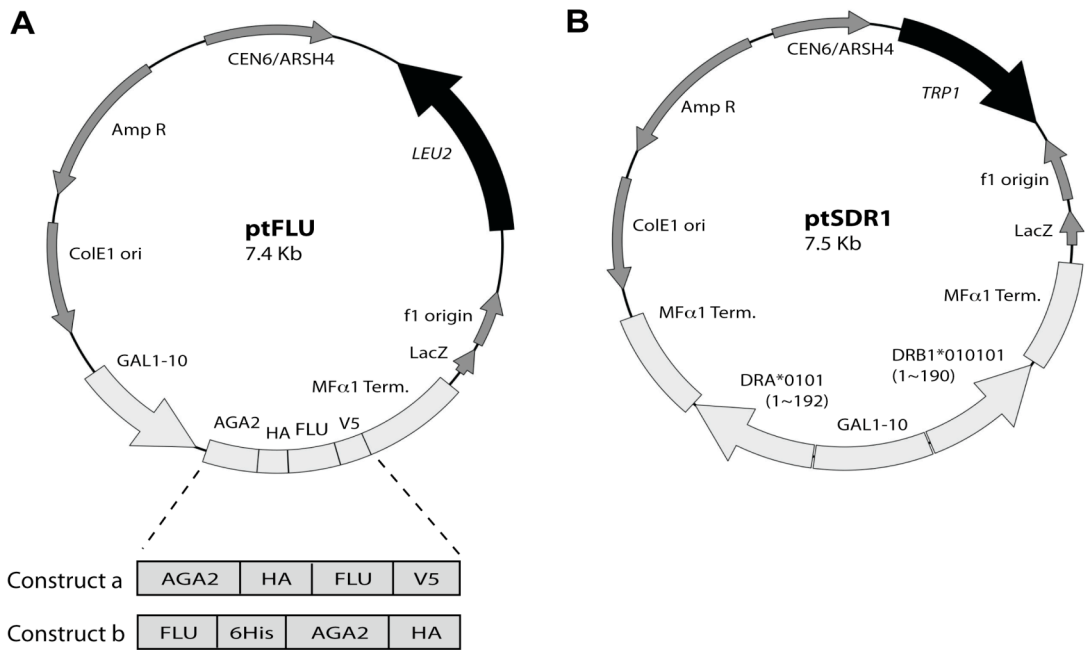

Figure 2-6 Map of yeast shuttle vectors for yeast co-display. **A.** A *LEU2* nutrient marker bearing plasmid is constructed for displaying peptide fusion covalently linked to Aga2p either at its C-terminus (Construct a in plasmid ptFLU) or at the N-terminus (Construct b in plasmid ptNFLU). In construct a, FLU peptide is flanked by HA and V5 epitope tags, whereas in construct b, FLU is overhung at the N-terminus of Aga2p with a 6His epitope tag fused in between. **B.** The plasmid ptsDR1 with a *TRP1* nutrient marker allows simultaneous expression of DRα [DRA*0101(1-192)] and DR1β [DRB1*010101(1-190)] extracellular domains in secreted forms from separate cassettes under the control of a bi-directional promoter GAL1-10.

EBY100 (*URA+, trp-, leu-*) transformed with ptFLU or ptNFLU (both *LEU+*) was selectively grown up and induced under *ura-, trp+, leu-* condition. Immunofluorescent labeling against HA-tag and flow cytometric analysis of labeled yeast demonstrates the successful surface display of FLU peptide using either construct, but a better surface expression level by construct b (Figure 2-7A). To evaluate HLA-DR1-binding probability of surface-displayed FLU, another plasmid ptsDR1 with a *TRP1* nutrient marker (Figure 2-6B) for secreting soluble HLA-DR1 heterodimer was constructed and transformed into yeast displaying either form of FLU peptide fusion. The resulting yeast strains were cultured under *ura-, trp-, leu-* condition and labeled using DR-specific

antibodies. The overlay of flow cytometric data indicates that HLA-DR1 can be anchored on the surface of yeast no matter which terminus of Aga2p was used for FLU displaying, but an obvious difference of detectable HLA-DR1 amounts between these two strains exists (Figure 2-7B). Construct a (Figure 2-6A) with lower FLU expression level actually leading to the higher DR1-anchoring capability was chosen for performing yeast co-display in the rest of this thesis work.

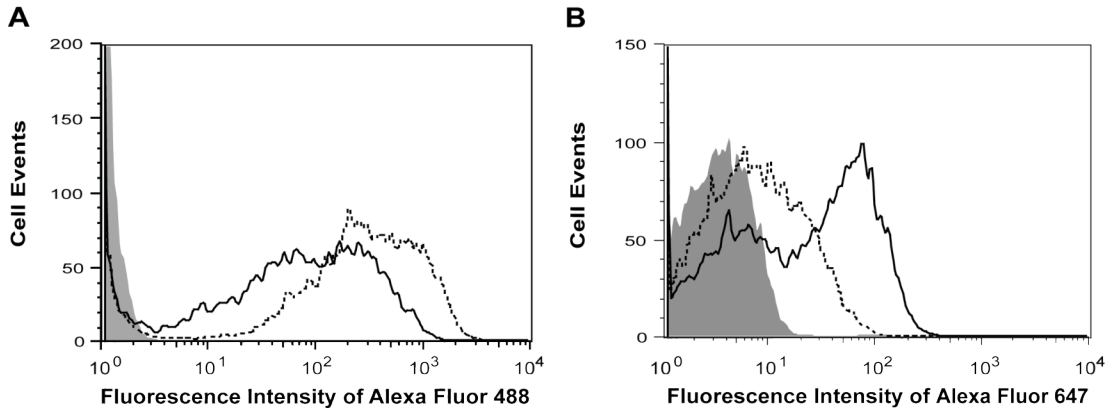

Figure 2-7 Surface levels of two FLU fusion constructs and anchored soluble HLA-DR1. A. Yeast displaying FLU peptide as a Aga2p C-terminal fusion (solid curve) or N-terminal fusion (dashed curve) were induced in SG-SCAA + tryptophan at 30 °C and labeled sequentially with HA-tag-specific H6908 and Alexa Fluor 488 conjugated secondary antibody and analyzed on BD FACSVantage. **B.** The same yeast strains as in **A** were transformed with plasmid expressing soluble HLA-DR1, induced in YPG at 20 °C and labeled sequentially with DR-specific L243 and a tandem dye PE-Alexa Fluor 647 conjugated secondary antibody and analyzed on Beckman Coulter XL. In both cases, yeast labeled with corresponding secondary antibody only (shaded) represents an auto-fluorescent background cell population in the overlayed histograms.

To further confirm the display of FLU to Aga2p C-terminus on the surface of yeast using construct a, HA-tag and V5-tag flanking FLU peptide were simultaneously stained with corresponding antibodies. In comparison to non-peptide displaying strain (Figure 2-8A), yeast expressing FLU fusion exhibit comparable HA-tag and V5-tag display levels, which is independent to the expression of soluble HLA-DR1 (Figure 2-8B and C). The FLU display level on the surface of yeast can be represented by the

background-corrected mean fluorescence intensity value normalized to the background intensity:

$$cMFI = \frac{MFI_{(HA+)} - MFI_{(HA-)}}{MFI_{(HA-)}} \qquad \text{Equation 2-1}$$

where $MFI_{(HA-)}$ and $MFI_{(HA+)}$ represent mean fluorescence intensity of the HA-tag coupled Alexa Fluor 647 emission for negative and positive cell populations, respectively (Figure 2-8 (i)). Analysis using anti-V5-tag staining resulted in equivalent *cMFI* values (Figure 2-8 (ii)), confirming the surface display of FLU peptide.

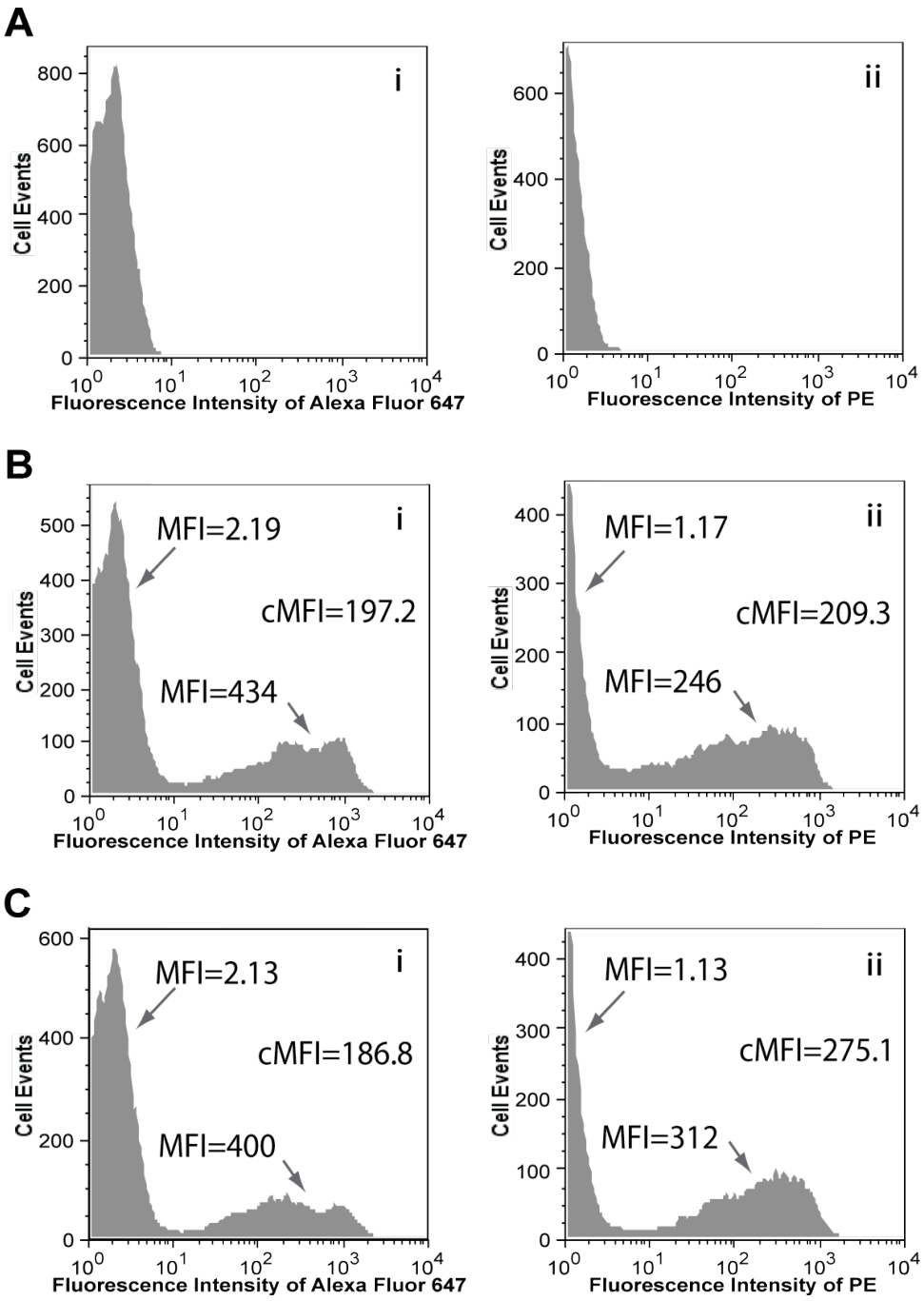

Figure 2-8 Simultaneous detection of both HA and V5 epitope tags flanking FLU peptide. Yeast expressing **A.** soluble HLA-DR1, **B.** Aga2p-fused FLU, or **C.** both were induced in YPG at 20 °C and labeled with antibodies specific for (i) HA-tag or (ii) V5-tag and analyzed on BD FACSCalibur. *cMFI* was calculated using MFI measured by Flowjo software (Tree Star Inc.) for positive and negative cell populations.

2.3.3. Yeast co-display of FLU and FLU-anchored soluble HLA-DR1

It was suggested previously that a two plasmids based co-display methodology (Figure 2-1) can be derived from the classical yeast display approach such that the MHC-II and target peptide are expressed from separate yeast shuttle vectors (Figure 2-6). MHC-II heterodimers that bind the Aga2p-fused target peptide are thus anchored to the cell surface via Aga2p (Figure 2-7), whereas MHC-II that do not bind the target peptide or bind instead to endogenous peptides present in the secretory pathway will be either secreted as soluble species or potentially degraded by the secretory quality control machinery, as suggested by our classical surface display results. To verify the peptide-binding-dependent cell surface display of HLA-DR1, plasmids directing expression of HLA-DR1 or Aga2p-FLU were transformed into EBY100. Yeast expressing combinations of these proteins were immunofluorescently co-stained to detect the level of FLU peptide and HLA-DR1 on the surface and analyzed by flow cytometry (Fig. 2-9). Yeast expressing only the soluble HLA-DR1 ectodomain show no detectable staining, while those co-expressing soluble HLA-DR1 and surface-displayed FLU demonstrate correlated signals for both (Fig. 2-9B and D). Based on the known peptide-binding motif of HLA-DR1[90, 145], Ala was substituted for Tyr at the P1 major anchor position of the FLU peptide (peptide P1-Ala; PK**A**VKQNTLKLAT; see Figure 4-1). As anticipated, this substitution abolished detectable HLA-DR1 on the cell surface (Figure 2F) without affecting peptide display level (Figure 2C and E), confirming the specific anchoring of HLA-DR1 to the surface via interaction with FLU.

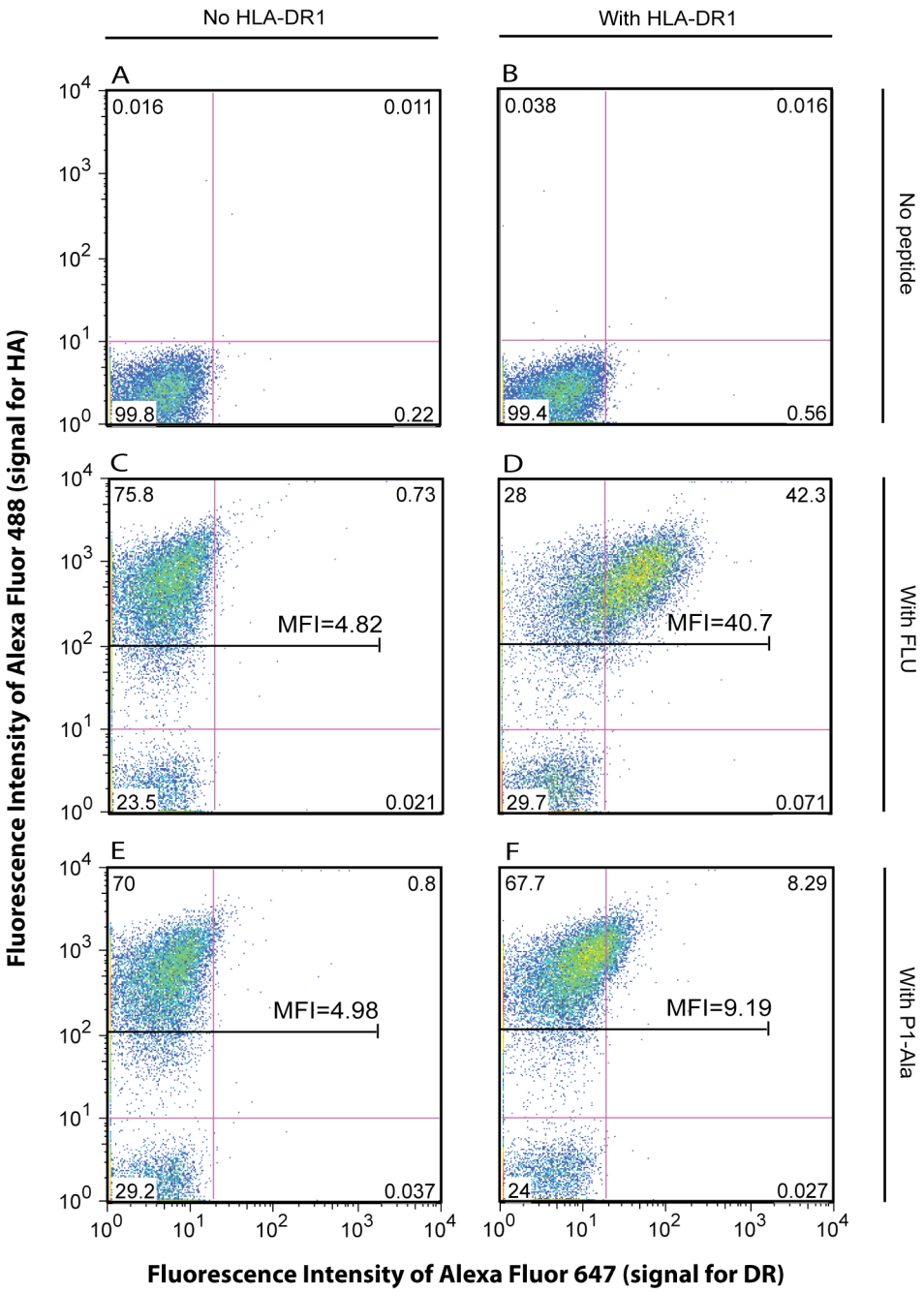

Figure 2-9 Yeast co-display and surface detection of FLU peptide and FLU-bound soluble HLA-DR1. A. Yeast parent strains or yeast expressing **B.** HLA-DR1, **C.** Aga2p-fused FLU alone, **D.** Aga2p-fused FLU with HLA-DR1, **E.** Aga2p-fused P1-Ala alone, or **F.** Aga2p-fused P1-Ala with HLA-DR1 were induced in corresponding galactose-containing minimal media at 30 °C, co-stained by anti-HA and anti-DR antibodies, and analyzed on BD LSR II. MFI of DR signal was generated using Flowjo for HA positive cell population. Proportions of the whole cell events were indicated at the corner of each quad. P1-Ala is a FLU analogue with Tyr308 replaced by Ala.

2.3.4. Stripping of FLU/HLA-DR1 complex off the surface of yeast

Anchorage of all Aga2p fusions to Aga1p is through two disulfide bonds, which is reducible by reducing reagents (Figure 2-10), such as Dithiothreitol (DTT) and β-mercaptoethanol (βME). Other than that, reducing reagents can also degrade HLA-DR1, which contains three disulfide bonds (one on α chain and two on β chain).

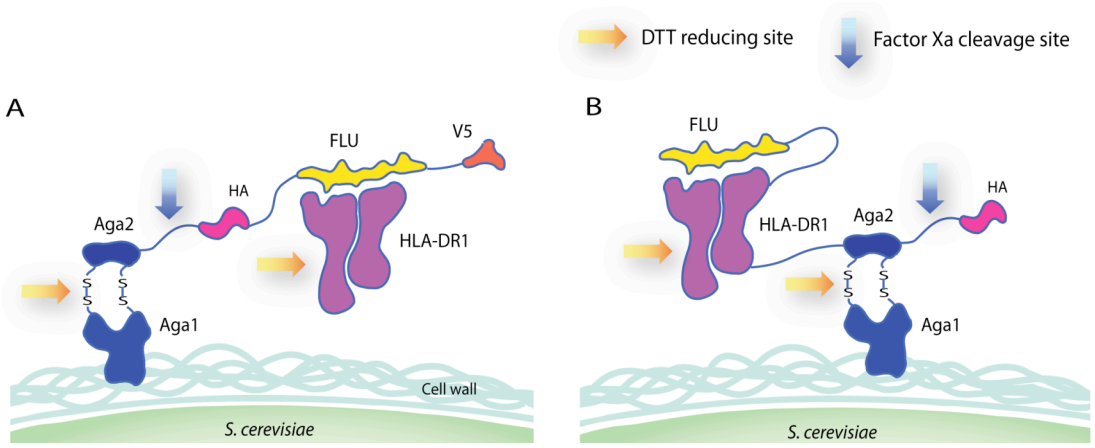

Figure 2-10 DTT and Factor Xa reacting sites in two FLU/HLA-DR1 displaying schemes: A. co-display, and **B.** classical surface-display.

To evaluate the effect of reducing reagents, different concentrations of DTT were applied to yeast strains displaying FLU/HLA-DR1 complexes by co-display or classical surface display. Cells before or after DTT treatment were co-stained to detect HA-tag and HLA-DR1 by flow cytometry (Figure 2-11). As expected, mean fluorescence intensity of either anti-HA-tag or anti-DR signal decreased as increasing DTT concentration (Figure 2-12 and Table 2-12), indicating the surface anchorage of FLU/HLA-DR1 complexes by Aga2p for both displaying schemes. It is worthy of noticing that almost all FLU/HLA-DR1 complexes were stripped off the surface of co-displaying yeast at the concentration of 1 mM, while the negative showed little influence

by DTT except for a small population moving towards higher intensity possibly due to contamination from DTT preparation.

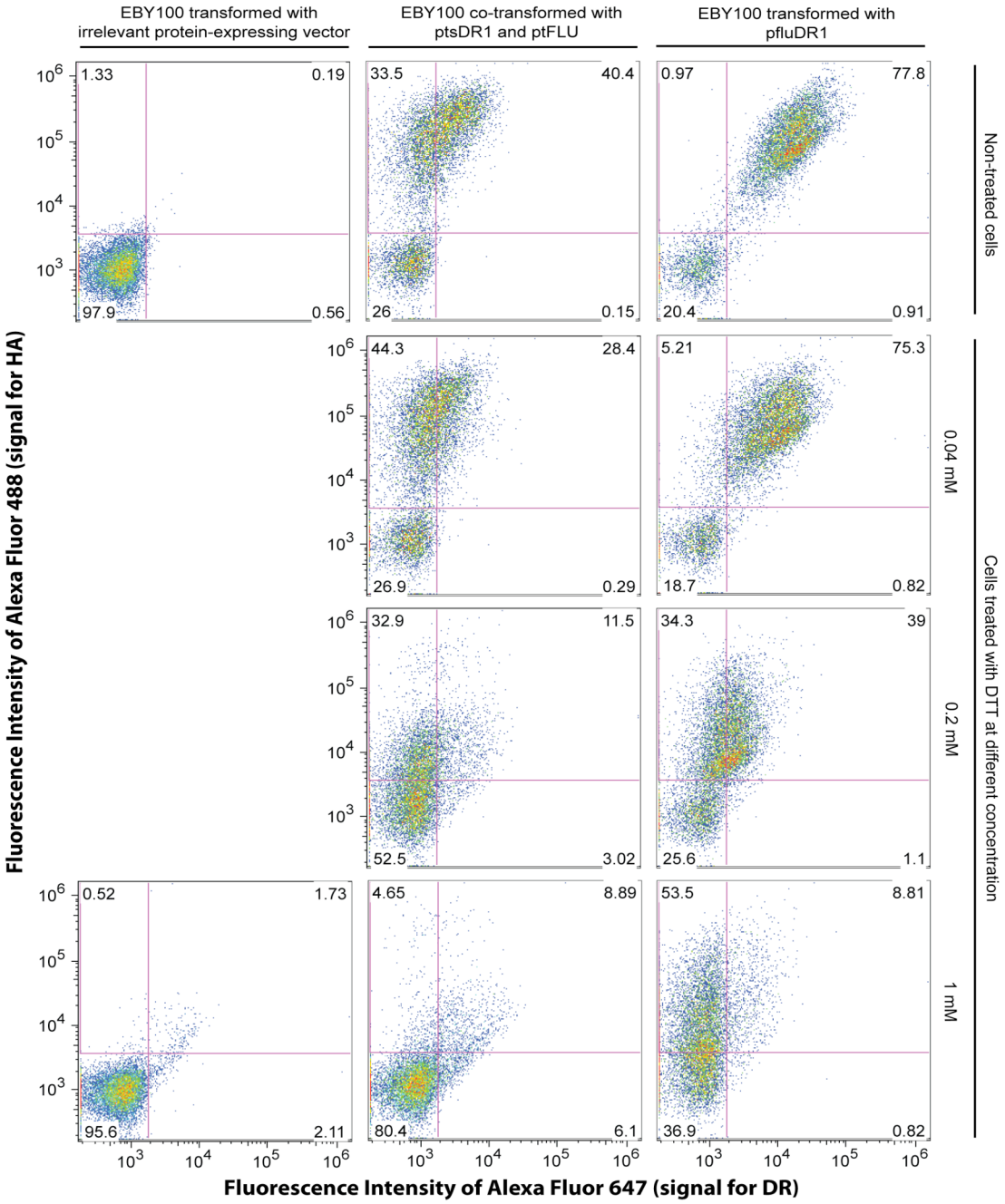

Figure 2-11 Effect of DTT on surface display of Aga2p or FLU/HLA-DR1 complexes. Yeast expressing an irrelevant soluble protein (left column), co-displaying Aga2p-FLU and FLU-anchored soluble HLA-DR1 (middle column) or displaying FLU-linked HLA-DR1 by classical approach (right columnw), collected before (top row) or after (other rows) DTT treatment, were double labeled using antibodies specific for HA-tag and HLA-DR1 and applied to Accuri C6 for dot-plot analysis.

Table 2-1 MFI for HA or DR signal on the surface of yeast before or after DTT treatment.

		EBY100 co-transformed with ptsDR1 and ptFLU		EBY100 transformed with pfluDR1	
		$MFI_{(HA)}$	$MFI_{(DR)}$	$MFI_{(HA)}$	$MFI_{(DR)}$
Non-treated		192000	2721	149000	19439
Treated with DTT	**0.04mM**	138000	1902	110000	10613
	0.2mM	12290	1618	37300	2231
	1mM	6471	1336	22857	1232

NOTE: MFI represents mean fluorescence intensity for the major cell population shown in Figure 2-11.

A short peptide (Ile-Glu-Gly-Arg) constructed in between Aga2p and HA-tag enables removal of Aga2p C-terminal protein fusion by Factor Xa cleavage (Figure 2-10). To examine this effect, yeast strains displaying FLU/HLA-DR1 by either displaying scheme were collected and incubated with Factor Xa protease at two different temperatures (4 °C and 23 °C). Flow cytometric data of yeast co-displaying FLU and HLA-DR1 (Figure 2-12 (middle column)) exhibits a noticeable decrease for both HA and DR signals, as suggested by calculated *MFI* values (Table 2-2). Longer treatment (~2 days) by Factor Xa at 23 °C resulted a substantial decline of the amount of FLU/HLA-DR1 complexes (Figure 2-13 and Table 2-3), further confirming the fact that FLU peptide is fused to the C-terminus of Aga2p and HLA-DR1 is bound by FLU peptide (or at least peptides to the C-terminal downstream of Aga2p//HA-tag cassette, which do not confirm previous results). On the other hand, the data of yeast displaying FLU-linked HLA-DR1 by classical approach (Figure 2-13 (right column)) shows decrease only for HA signal but not DR signal (see Table 2-2), demonstrating that FLU/HLA-DR1 complexes were linked to the other terminus of Aga2p, where Factor Xa cleavage had no impact on. Effect of both DTT and Factor Xa protease confirmed

previously conclusion that soluble HLA-DR1 was specific anchored on the surface of yeast by interaction with FLU peptide.

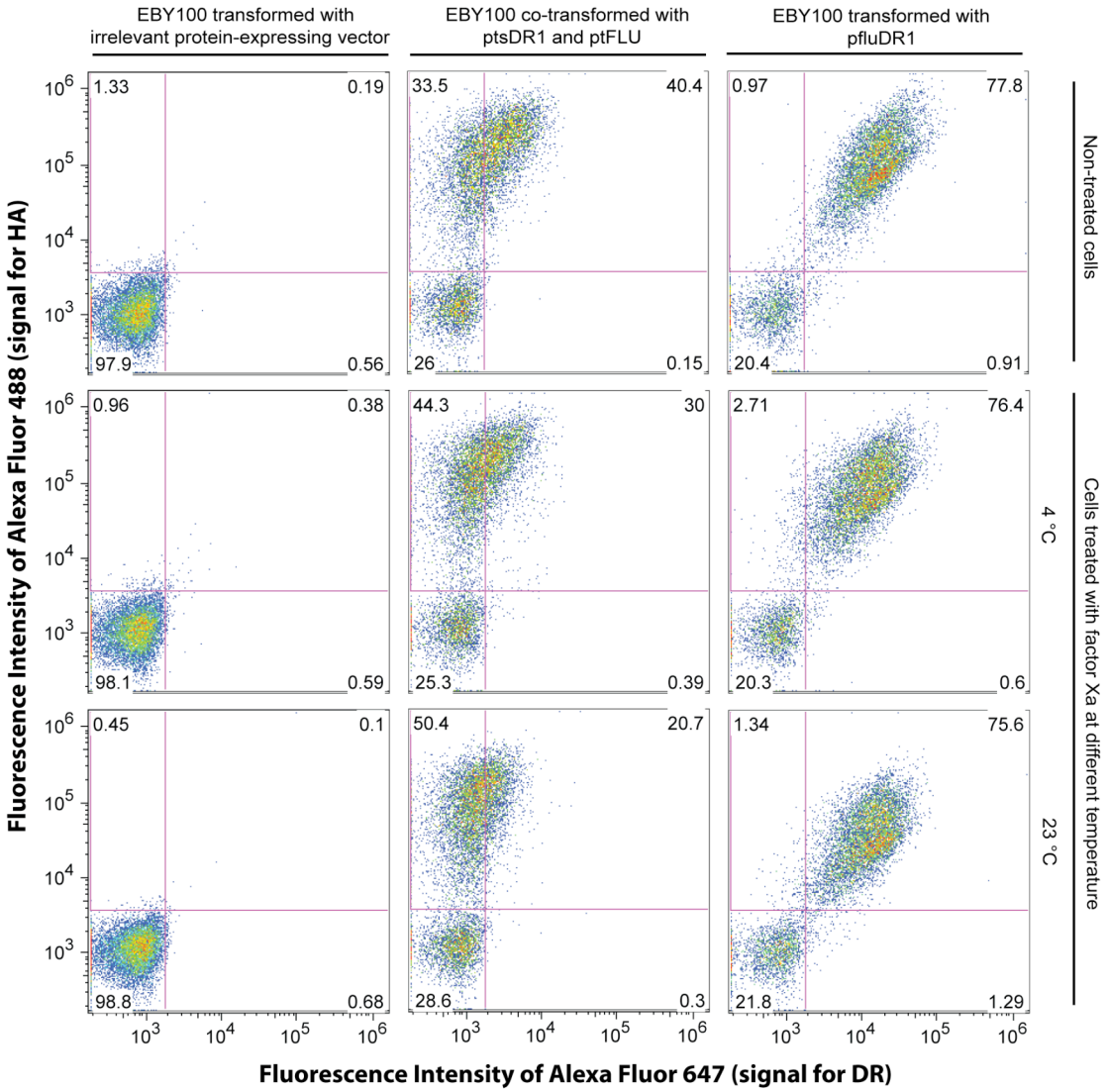

Figure 2-12 Effect of Factor Xa protease on surface display of FLU/HLA-DR1. Yeast expressing an irrelevant soluble protein (left column), co-displaying Aga2p-FLU and FLU-anchored soluble HLA-DR1 (middle column) or displaying FLU-linked HLA-DR1 by classical approach (right columnw), collected before (top row) or after (other rows) Factor Xa treatment were double labeled for HA-tag and HLA-DR1 and applied to Accuri C6 for flow analysis.

Table 2-2 MFI for HA or DR signal on the surface of yeast before or after Factor Xa treatment.

		EBY100 co-transformed with ptsDR1 and ptFLU		EBY100 transformed with pfluDR1	
		$MFI_{(HA+)}$	$MFI_{(DR)}$	$MFI_{(HA+)}$	$MFI_{(DR)}$
Non-treated		192000	2721	149000	19439
Treated with factor Xa	**4 °C**	201000	1983	123000	13712
	23 °C	125000	1498	50682	14489

NOTE: MFI represents mean fluorescence intensity for HA-positive cell population shown in Figure 2-12.

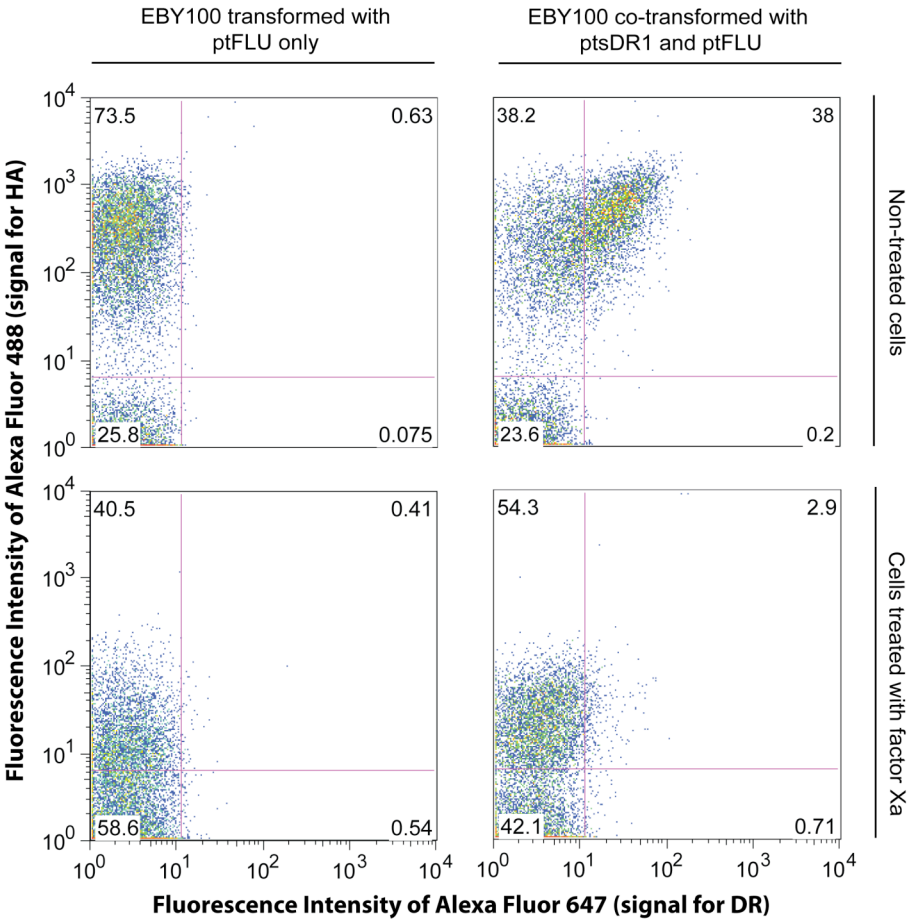

Figure 2-13 Effect of Factor Xa on FLU/HLA-DR1 display level after longer incubation at 23 °C. Yeast displaying Aga2p-FLU (left column) or co-displaying Aga2p-FLU and FLU-anchored soluble HLA-DR1 (right column), collected before (top row) or after (bottom row) Factor Xa treatment, were double labeled using antibodies specific for HA-tag and HLA-DR1 and applied to BD FACSVantage (cell sorter) flow cytometric analysis.

Table 2-3 MFI for HA or DR signal on the surface of yeast before or after longer Factor Xa treatment.

	EBY100 co-transformed with ptsDR1 and ptFLU		EBY100 transformed with pfluDR1	
	$MFI_{(HA+)}$	$MFI_{(DR)}$	$MFI_{(HA+)}$	$MFI_{(DR)}$
Non-treated	394	3.27	442	18.1
Treated with factor Xa	21.4	3.34	29	4.71

NOTE: MFI represents mean fluorescence intensity for HA-positive cell population shown in Figure 2-13.

2.3.5. Verification of genotype-phenotype linkage between plasmids and co-displayed proteins

In order to extend the co-display approach to library screening, the genotype-phenotype linkage between plasmids and displayed proteins must be maintained; therefore, experiments were performed to determine whether HLA-DR1: (1) bound to FLU intracellularly and was subsequently exported to the surface, (2) was secreted "empty" into the culture medium (or with weakly bound endogenous peptides) and then bound to FLU, or (3) a combination of these mechanisms.

Yeast cells displaying FLU peptide (Figure 2-9C) or P1-Ala analogue (Figure 2-9E) were co-cultured with yeast expressing soluble HLA-DR1 ectodomain (Figure 2-9B) and double labeled using anti-HA and anti-DR antibodies before flow analysis (Figure 2-14A). Approximately 50% HA-positive cell population shown in the dot-plot indicates an unaffected peptide expression level on the surface of either peptide-displaying strain, whereas no obvious DR-positive cell population were observed, suggesting failure of secretion of soluble HLA-DR1 by corresponding yeast strain or failure of interaction between surface-displayed FLU and soluble HLA-DR1 outside yeast under the culturing

condition. Furthermore, yeast cells displaying FLU were incubated in conditioned medium from a culture of soluble HLA-DR1-expressing yeast (Fig. 2-14 B and C), including the co-displaying strain that ensured secretion of soluble HLA-DR1 (Fig. 2-14B(i)). In all cases, no significant binding of HLA-DR1 to the surface of any yeast was observed regardless of media used, indicating that resorting of secreted HLA-DR1 between different cells in the same culture fails to occur and suggesting that HLA-DR1/FLU binding occurs intracellularly within the secretory pathway.

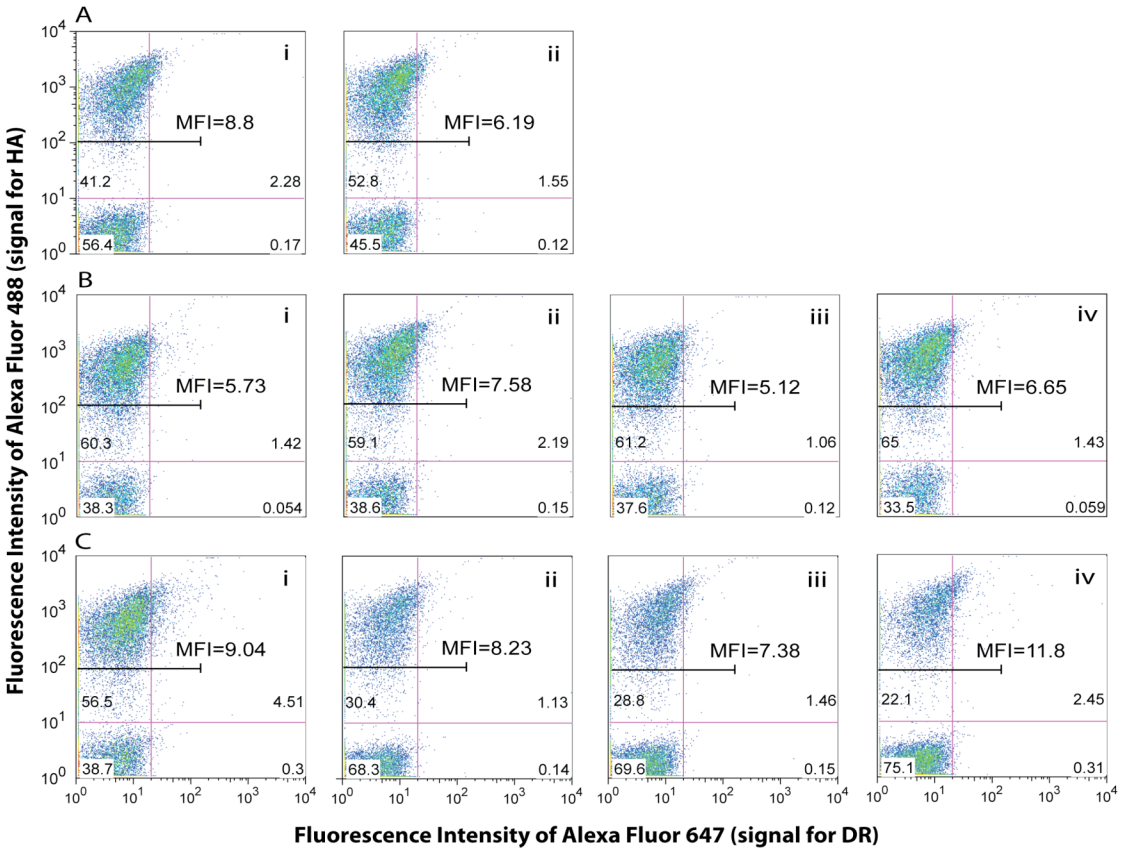

Figure 2-14 Control experiments for examining possible switching of soluble HLA-DR1 molecules among different yeast cells in the same culture. **A**. Co-culturing. 10^7 yeast cells expressing (i) Aga2p-FLU or (ii) P1-Ala were co-induced with 10^7 yeast cells expressing soluble HLA-DR1 in 2 ml SG-CAA + TRP at 30 °C. **B**. Medium switching. After induction in minimal media, 2.5 × 10^6 yeast cells displaying FLU peptide were incubated in 500 µl medium preconditioned by growing yeast (i) co-expressing Aga2p-FLU and HLA-DR1, (ii) expressing soluble HLA-DR1, (iii) co-expressing Aga2p-P1-Ala and HLA-DR1, or (iv) expressing Aga2p-FLU peptide at 30 °C. **C**. Minimal media used in **(B)** for induction was replaced by rich medium YPG. Cells were all double labeled for HA-tag and HLA-DR1 molecules by H6908 and L243 before applied on BD LSRII for flow cytometric analysis.

2.4. Discussion

2.4.1. Yeast is suitable for expressing functional heterologous multimeric protein complexes

The eukaryotic protein displaying and engineering system, *Saccharomyces cerevisiae* has been demonstrated to have the ability to assemble heterologous multimeric proteins either as recombinant single chain derivatives[74, 132, 133, 146], or more natively as non-covalently associated heterodimers, such as MHC-II alleles DR4[102], Fab antibody fragments[88]. In this chapter, we also showed that the heterodimeric HLA-DR1 can be displayed on the surface of yeast with or without covalently linked FLU peptide (Figure 2-4). The noticeablly better expression level of FLU-linked form on yeast surface agrees with the stabilization effect of FLU peptide on assembly of DR1 molecules[67, 68], suggesting a similarity of the *in vivo* environments between yeast and human cells. Furthermore, we confirmed the peptide binding ability of "empty" form of HLA-DR1 on yeast surface (Figure 2-5), which indicates the potential of yeast surface display as a strategy for characterizing peptide/MHC interaction. However, the time and labor consumed on chemically synthesizing peptides individually would eventually restrain this method to generate high throughput results. Another disadvantage of most available strategies used for characterizing peptide/MHC-II interactions (including the one just mentioned and those mentioned in introduction of this chapter) is the *in vitro* laboratory environments where interaction between peptides and MHCs takes place. As we all know, in their native biological environments, proteins are not independently existing but mostly within a complex network, where molecular interactions among

proteins, nucleic acids, lipids, chemicals, etc. are occurring all the time (known as interactome[147, 148]), so, the simple *in vitro* peptide binding condition might not accurately reflect the true scenario for peptide/MHC-II association happening inside ER of professional APCs.

2.4.2. Yeast co-display contains a variety of advantages for studying peptide/MHC-II interaction

Most of above problems can be solved by the novel method – yeast co-display – we developed in this chapter. Intensive characterization and validation were performed for an interacting pair: HLA-DR1 and DR-specific FLU peptide. These two molecules are translated from gene cassettes carried by two separated yeast vectors, both of which can be accommodated by the same yeast cell. Under regulation of secretory signal peptides, both molecules will enter endoplasmic reticulum, where the association between FLU and HLA-DR1 starts occurring at a more *in vivo* like environment. Along with exocytic process, the Aga2p fused FLU/HLA-DR1 complex will be delivered outside yeast and anchored on the surface of yeast as design (Figure 2-1). In this genetic design, we took advantage of the fact that the binding groove of MHC-II is more open and allows both ends of peptide extended outside[60], and covalently attached FLU peptide to the yeast surface protein, **a**-agglutinin. This small trick not only provides a linkage for co-displaying FLU and soluble HLA-DR1, but also eliminates possible artifacts caused by covalently linking HLA-DR1 to anchoring proteins in other display strategies, such as Aga2 protein for yeast display[74], gIII or gVIII coat proteins for phage display[139], hybrid

produces of lipoproteins and membrane proteins for *E.coli* display[80], envelope glycoprotein gp64 for baculovirus-insect cell display[76].

By site directed mutagenesis and enzymatic cleavage (Figure 2-9E and F, Figure 2-10 ~ 2-13, and Table 2-1 ~ 2-3), the surface anchorage of soluble HLA-DR1 by FLU peptide but not other yeast surface structural molecules has been evaluated and confirmed. The fact that a single substitution in the FLU peptide sequence was able to disrupt the binding of HLA-DR1 obviously implies the peptide specificity of HLA-DR1 produced in our co-display system, which we will further examine in the following chapter more quantitatively.

It is worth noticing that the display level of FLU/HLA-DR1 complexes by co-displaying yeast is significantly lower than using classical surface display (Figure 2-11 top row). This could be a direct reflection for the non-covalent association of FLU to HLA-DR1 in yeast, because there is no artificial linker introduced in between these molecules any longer. Another interesting result is that no intercellular switching of soluble HLA-DR1 molecules via FLU anchoring on the surface of yeast is detectable (Figure 2-14), validating that the interaction of these two molecules mainly takes place inside yeast cells and both dissociation and re-association kinetics happened outside yeast is probably slow enough such that phenotype on the surface of a yeast cell is directly related to its genotype but not others. This observation not only clarifies the binding location but also suggests FACS as an appropriate screening and sorting strategy for discovering HLA-DR1 mutants specific for target peptide by directed evolution.

As a new *in vitro* technology for studying *in vivo* interaction between peptide and MHC-II, yeast co-display bears most advantages described earlier for other methods,

such as the ability to perform quantitative analysis of peptide binding (the analysis in Figure 2-8 has already set up an basis), the potential to generate peptide libraries and/or MHC mutant libraries for high throughput assay, etc. In chapter III, we will discuss how to quantify peptide binding and how to maintain optimal experimental conditions so as to ensure the reliability of the quantitation. In the following two chapters, we will show some applications of this quantitative, high throughput methodology for identification of HLA-DR1 binding motif and prediction of HLA-DR1-specific ligands and for characterization and engineering of peptide binding specificity of MHC-II allelic variants.

Chapter 3. Optimization of yeast co-display for quantitative analysis

3.1. Introduction

Yeast co-display provides us a powerful tool to study the interaction between at least two molecules. It has been shown in chapter II that, by doing some simple cloning work we could culture yeast to produce both FLU peptides and HLA-DR1 proteins, in return, yeast will give a feedback about the interaction between these two molecules via presenting them on the surface for detection by flow cytometry. Only knowing whether a given MHC-II molecule can bind a target peptide or not is not enough sometimes, we would also like to know how well they bind with each other so that we can understand which residue at which position will favor or inhibit the interaction and what kind of side chain of a anchor residue is preferable to the peptide binding, etc. Indeed, yeast co-display system does have the potential to quantitatively determine the relative peptide binding affinity for MHC-II molecules. As demonstrated by experiments in Chapter II, the expression level of peptide on yeast surface is independent and supposed to be constant if yeast is always induced for a certain time, so theoretically, the better an MHC-II can bind a peptide, the more MHC-II should be anchored and displayed on the surface of yeast after being induced for a certain time.

Although the direct counting of how many peptides each yeast can display and how many MHC-II molecules are anchored by peptides on yeast surface is impossible, we can always get some related information about them by immunofluorescently labeling

MHC-II or peptide and measuring the fluorescence intensity using some detection techniques, such as flow cytometry as described in Chapter I. Due to the morphology and diverse size of budding yeast, the integrated fluorescence intensity of each yeast will be different, so measuring thousands or millions of yeast by flow cytometer and generating a distribution of global fluorescence intensity of different yeast cells is relatively more meaningful. The mean value of the distribution (*MFI*) can properly represent the relative number of MHC-II or peptides on the surface of the average yeast cell.

It is clear that the biological expression system will never tell us the exact amount of proteins produced and associated inside the cell before the result is turned out on its surface, so in order to quantitatively characterize peptide binding specificity of MHC-II by yeast co-display, optimal yeast culturing and labeling conditions have been determined so that the generated *MFI* from flow data with apparently small standard deviation could be used to accurately reflect how well an MHC-II binds to a peptide. This optimization actually contains two meanings. First one is to investigate appropriate experimental conditions for better proteins expression and interaction as well as more accurate labeling and surface detection. The expression of FLU peptide essentially is an indicator of protein expression level in yeast cells and an optimal expression level obviously provides enough binding blocks for efficient protein-protein interaction. A better interaction between FLU peptide and HLA-DR1 molecule in yeast will result in more detectable HLA-DR1 anchored for easier surface detection. Consequently, proper labeling for proteins on yeast surface will largely decrease the inaccuracy and variance of fluorescent detection by flow cytometry. Second meaning is that the optimized

working condition has no need to be the best, but at least a local optimum with relatively broad resistance to variations caused by different reagents and instrument parameters so that the quantitative assay generated later on can be more reliable. This will also ensure that mutagenesis introduced to FLU and HLA-DR1 molecules in the following chapters have no significant impact on quantitation by yeast co-display. Normally, an optimum should provide some resistance to small errors, so we will mainly focus on seeking for optimal experimental conditions in this chapter without worrying about the second meaning too much as it could be obtained automatically once working condition is optimized.

Variables for culturing yeast such as medium, temperature, time length, and cell density, etc, are direct effectors controlling the expression and association of proteins in yeast. Variables for fluorescent labeling including choices of primary antibodies and secondary fluorescently labeling reagents, concentration or dilution of reagents used, etc, are factors probably introducing errors for flow detection but not affecting the real amount of proteins on yeast surface. Actually, all these variables are more or less connected with each other, which makes the evaluation really difficult. As mentioned previously, we do not have to find out the best condition, so to make things easier and still reasonable, the number of variables were grouped and varied a few at a time to maximize the efficacy of the whole optimization process.

The most direct and reasonable phenotypic signal for evaluating the optimum of experimental conditions is the amount of HLA-DR1 on the yeast surface. In this chapter, we proposed a quantitative way to measure the relative HLA-DR1 amount in the materials and methods section and then evaluated several sets of variables based on

the HLA-DR1 measurement for searching an optimal experimental condition in result section. First of all, issues about medium, time and temperature selected for inducing yeast to produce proteins for interaction were analyzed. Following that, cell densities at different stages were evaluated for eliciting better detectable DR-signals on yeast surface. Finally, proper antibodies and labeling reagents were selected and the concentrations of them were titrated for an accurate surface detection.

3.2. Materials and Methods

3.2.1. Media for yeast growth and induction

SD-SCAA was used for cell enrichment. Either a minimal medium SG-SCAA or a rich medium YPG was picked for induction of GAL1-10 promoted protein expression. Both SD-SCAA and SG-SCAA were filter sterilized. For YPG, five different batches were used in experiments done in this chapter (Table 3-1).

Table 3-1 Different batches of YPG media used in this chapter.

Batch Number	Component	Vendor	Sterilization method
YPG 1	Yeast extract	BD Biosciences	Autoclaved
	Peptone	Fisher Sientific	
	Galactose	Acros Organics	
YPG 2	Yeast extract	BD Biosciences	Filtered
	Peptone	Fisher Sientific	
	Galactose	Acros Organics	
YPG 3	Yeast extract	Invitrogen	Autoclaved
	Peptone		
	Galactose	Acros Organics	Filtered
YPG 4	Yeast extract	BD Biosciences	Autoclaved
	Peptone		
	Galactose	Acros Organics	
YPG 5	Yeast extract	BD Biosciences	Autoclaved
	Peptone		
	Galactose	Acros Organics	Filtered

3.2.2. Primary and secondary labeling reagents

Immunofluorescent labeling for induced yeast cells was carried out mostly in a 2-step manner, and primary antibodies and secondary reagents with fluorophores conjugated are listed in the following.

One of the following DR-specific mouse Mab: L243 or biotinylated LB3.1 (purified from hybridoma culture supernatant at National Jewish Center, Denver, CO) was used to specifically label functional HLA-DR1 in the primary labeling step. Different secondary reagents used for labeling corresponding primary antibodies are Alexa Fluor 488 or 647 conjugated goat anti-mouse antibody (all Alexa Fluor dyes were purchased from Invitrogen), tandem dye PE-Cy5 conjugated streptavidin (BD Biosciences Pharmingen).

Polyclonal antibody H6908 and Alexa Fluor 488 - goat anti-rabbit were used as primary and secondary labeling reagents, respectively, for detecting FLU-linked HA-tag.

3.2.3. General culturing procedure for yeast co-display

Colonies for EBY100 transformants with two plasmids containing both *TRP1* and *LEU2* nutrient markers are first selectively grown up on SD-SCAA agar plates (containing 1 M sorbitol) under *trp-, leu-,* and *ura-* condition and then inoculated into 2 ml SD-SCAA liquid medium in an autoclaved glass test tube. The yeast culture was then enriched to a proper OD_{600} (an optimal range of the density will be determined in this chapter) in a 30 °C shaker at 225 rpm. To induce the expression of both FLU fusion and sDR1, enough yeast cells (an optimal amount will be determined in this chapter) are collected by centrifugation for ~1 min at maximum speed in a microcentrifuge. After withdrawing the supernatant, cell pellet was resuspended by 2 ml of either SG-SCAA or YPG medium and transferred to a new glass tube. Induction is carried out at some temperature for a certain time length before harvest for surface detection (optimal induction condition will be determined in this chapter).

Above is just a general yeast cultivation procedure, specific OD_{600}, medium, time length, temperature are indicated in each experiment stated later on.

3.2.4. General fluorescent labeling procedure

After induction, a pellet of 10^6 yeast cells were collected from each sample by centrifugation into a 1.5 ml Eppendorf tube and washed once with at least 300 μl ice cold PBS + 1% BSA. Cells were then stained with a proper dilution of a primary antibody specific for either DR or HA epitope tag in PBS + 1% BSA at room

temperature (RT) or on ice for 30-60 min. Following that, cells were spun and washed once again with ice cold PBS + 1% BSA and then incubated with corresponding secondary labeling reagents diluted in PBS + 1% BSA at RT or on ice for 30-60 min. Finally, cells were washed with PBS + 1% BSA one more time and resuspended in 500~750 μl ice cold PBS + 1% BSA for analysis by flow cytometer.

Above is just a general labeling procedure, specific antibody chosen and exact dilution used will be indicated in each experiment stated later on.

3.2.5. Quantitative analysis of flow cytometric data

A direct way of evaluating one working condition versus another is to compare the relative amount of HLA-DR1 on the surface of yeast prepared using different conditions, because the only reason one can detect HLA-DR1 molecules by yeast co-display is due to the association and anchorage of HLA-DR1 by FLU peptide expressed on yeast surface. Herein, we proposed one quantitative approach for determining relative amount of HLA-DR1 on the surface of yeast by flow cytometric analysis. In most experiments carried out in this chapter, both a DR-positive and a DR-negative co-displaying yeast strain were used under each working condition. After flow cytometry, data were first plotted by SSC vs. FSC for checking and gating the right yeast cell population (Figure 3-1 (i)). Then the fluorescent signal for labeled HLA-DR1 molecules on the surface of gated cells was displayed as a cell event distribution in a histogram (Figure 3-1 (ii)). By using Flowjo, one can measure the mean fluorescence intensity for the distribution. The HLA-DR1 display level on the surface of yeast can be estimated by the background-corrected mean fluorescence intensity value normalized to the background intensity:

$$cMFI_{(DR)} = \frac{MFI_{(DR)}(+) - MFI_{(DR)}(-)}{MFI_{(DR)}(-)} \qquad \text{Equation 3-1}$$

where $MFI_{(DR)}(+)$ and $MFI_{(DR)}(-)$ represent mean fluorescence intensity of DR signals on the surface of DR-positive and negative yeast strains, respectively.

The $cMFI_{(DR)}$ value can be used to reflect the relative HLA-DR1 amount on the surface of yeast cultured and labeled under a certain condition. The negative subtraction and normalization are both necessary to set up a base line for comparing data collected at different experimental condition and minimize unnecessary errors introduced by different parameters set up on flow cytometers. It is worthy of noticing that due to the difficulty to discriminate negative and positive, the measurement of MFI (Figure 3-1 (ii)) includes the whole cell population, which actually brings down all absolute values of *cMFI* after the background subtraction in Equation 3-1, but has no impact on relative value when we compare them with each other. Additionally, due to the truth that negative strain could change its surface turning-out at different culturing condition as well, there is actually no exact base line to compare data collected on different flow cytometers or even on different dates. Therefore, most comparison was done for flow data collected on the same flow cytometer in the same day, and results from different dates were then put together to draw a final conclusion.

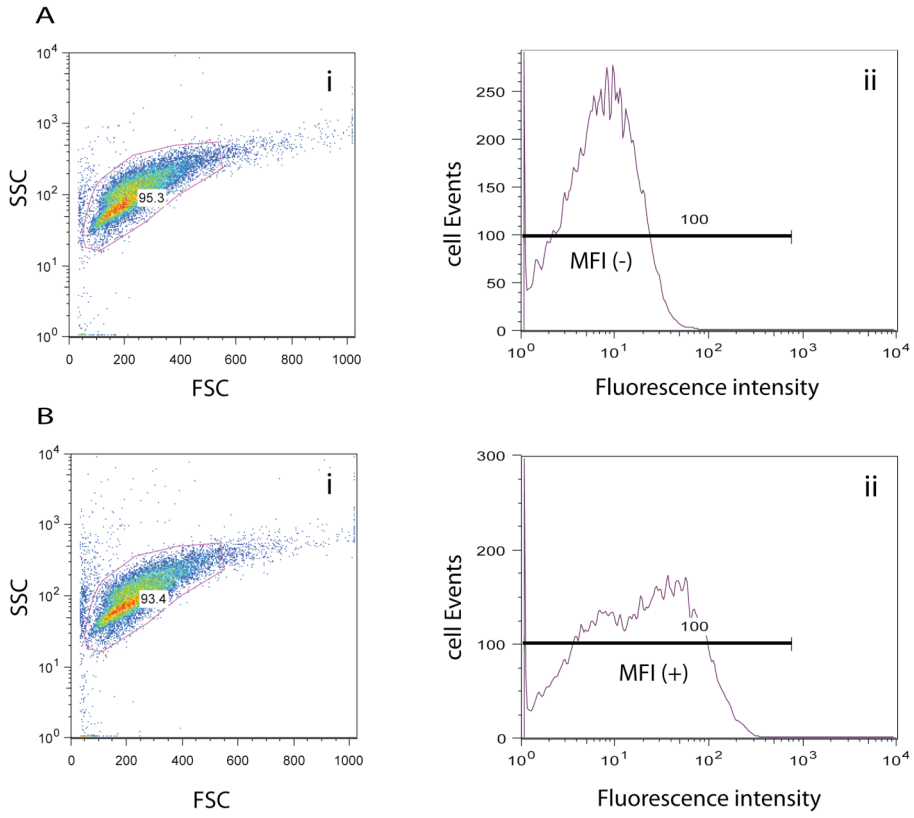

Figure 3-1 Quantitative analysis of relative HLA-DR1 amount on yeast surface using flow cytometric data generated from anti-DR-labeled yeast. Data of both **A.** DR- and **B.** DR+ cell samples were first plotted in (i) a dot plot by side scatter (SSC) vs. forward scatter (FSC) representing information of cell morphology and intracellular organelle complexity, and then gated for the majority of cell population (red line, ~90% of total) in order to separate real yeast from contaminants. The gated cell subset was then plotted in (ii) a histogram for fluorescence intensity of fluorophore labeling HLA-DR1 molecules. Mean fluorescence intensity (MFI) of the whole subset was measured for both **A** and **B.**

3.3. Results

3.3.1. Evaluation of medium, temperature and time for induction

The growth and induction of yeast cells are usually sensitive to a couple of things which people might not notice most of the times, such as air flow inside the shaker, shaking speed, size and shape of the flasks or tubes, which will affect the aeration for the cell culture and thus the cell growth. To minimize the influence of these factors, all yeast strains used for quantitative analysis were inoculated and induced in a constant

volume - 2 ml of medium in an autoclaved glass culture tube, and the shaking speed was set to 225rpm for all experiments. By setting other effectors unchanged, we will evaluate medium and temperature chosen for induction at different time length first.

Co-displaying yeast strains were first selected out on SD-SCAA agar plates, which will inhibit the growth of background or parent strains. After being enriched in the minimal liquid medium SD-SCAA, cells were switched to different induction media for expressing proteins. Herein, we tested two kinds of media: a rich medium YPG and a minimal medium SG-SCAA both containing galactose with the ability to turn on the GAL1-10 promoter for expression of both Aga2-FLU fusion and soluble HLA-DR1. The cell growth rate in YPG is dramatically faster than that in SG-SCAA, so the induction in YPG was tried mainly at 20°C. Occasionally, huge difference between the final densities of induced cells was observed when different batches of YPG media were used (OD_{600} not shown). To determine whether YPG, made of components bought from different vendors and prepared differently by autoclaving or filtering, would affect the expression and interaction of FLU and HLA-DR1, different batches of YPG were examined in some experiments (Table 3-1, Figure 3-2A~C and Figure 3-3D). Results suggest that not only the final cell density is different, but the HLA-DR1 amount on cell surface is also different. However, there is no obvious relationship between different batches and the detectable HLA-DR1 amounts. For example, YPG 3 can provide a much higher *cMFI* than YPG 2 when induced for 90 h (Figure 3-2A), but a lower *cMFI* when induction time became shorter (Figure 3-3D). In contrast with using other bathes, cells in YPG 3 didn't show obvious loss of DR signal after 90 h induction comparing with 20 h (Figure 3-2A, B and Figure 3-3D), where the yeast extract and peptone

purchased from Invitrogen for making YPG 3 could have an impact. Hence, it is difficult to predict which batch of YPG will give the best expression or the highest association of proteins. Theoretically, YPG provides the most nutrients and materials for cells to make proteins, however, the uncertainty and richness of components in YPG will lead to an unpredictable cell growth behavior, which we would like to avoid when performing quantitation. On the contrary, although not rich, SG-SCAA indeed provides enough nutrients to keep cells growing slowly while making proteins as well as to inhibit background cells growing. Henceforward, we will characterize the induction in SG-SCAA more intensively.

It was suggested that induction in YPG with shorter time tends to favor the detection of HLA-DR1 molecules on yeast surface (Figure 3-2B and C), so we tried to induce yeast in SG-SCAA at two different temperatures using shorter time frames. Within 2 Days, cells induced in SG-SCAA at 30 °C can display an equivalent amount of HLA-DR1 in comparison with those induced in YPG at 20 °C and both exhibited a similar feature that the longer they were induced, the fewer HLA-DR1 we can detect on cell surface (Figure 3-2B and C, also see Figure 3-3A~C), whereas cells induced in SG-SCAA at 20 °C do need longer time to present more HLA-DR1 (Figure 3-2C), suggesting that protein expression in minimal media is slower than in rich media. Considering higher temperature could favor the protein binding kinetics, we stuck to 30 °C induction in SG-SCAA and explored behaviors at even shorter time frames, where DR signal no longer increased with increasing time (Figure 3-2D). Together with previous results (Figure 3-2C), a local optimum for induction time can be searched out between 14 h and 20 h when inducing co-displaying yeast in SG-SCAA at 30 °C.

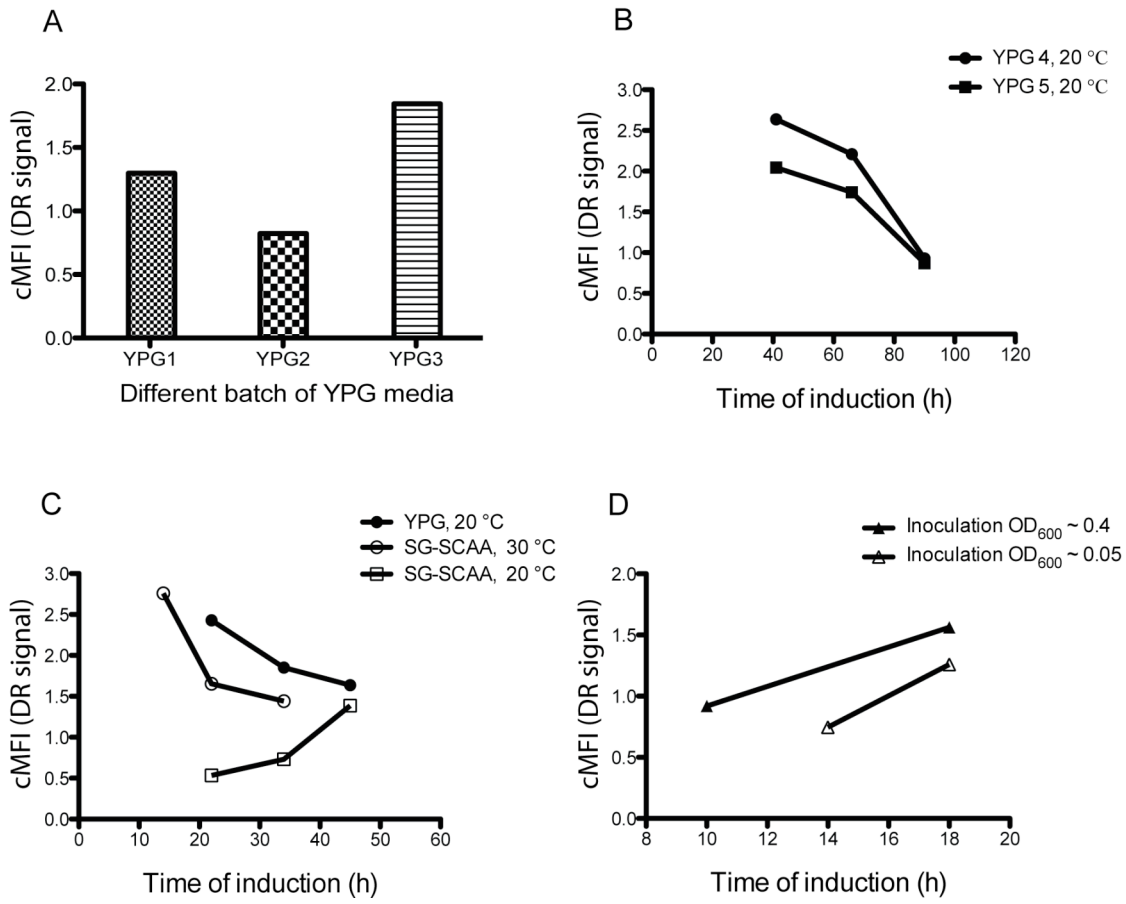

Figure 3-2 Effect of medium, temperature and time length used for yeast induction on HLA-DR1 surface display level. Shown are four separate experiments. $cMFI_{(DR)}$ were calculated as described in Equation 3-1 for the co-displaying yeast strain prepared at different conditions. **A.** Cells were induced in different batches of YPG at 20 °C for 90 h. **B.** Time lengths for induction were varied when using two different YPG media. **C.** Induction was carried out in different media at different temperature for various time lengths. Biotinylated LB3.1 and PE-Cy5 conjugated streptavidin were used for DR-labeling. **D.** The initial cell density (OD_{600}) for inoculation and the time length for induction were both varied. Another change in this set of experiment is that cells collected for induction were not spun down but with some volume of SD-SCAA carried over to the medium for induction. (Other than noted, L243 and Alexa Fluor 488 conjugated goat anti-mouse antibody were used for the two-step yeast labeling. The staining was carried out on ice and flow cytometry was done on Beckman Coulter XL.)

3.3.2. Optimization of cell density for induction

We introduced another factor – cell density when seeking for a local optimum of induction time when performing induction in SG-SCAA at 30 °C to broaden the applicability of the optimized condition. Since most of the time yeast is inoculated from plates directly into culture tubes, it is impossible to check the starting yeast age. Indeed,

a little more HLA-DR1 molecules were detected when inoculating at a higher density (OD_{600}~0.4) than lower one (OD_{600}~0.05) (Figure 3-2D), partially reflecting that starting with fresh cells tends to result better than older ones for a certain growth length, because old ones normally contain more dead cells which won't divide. This postulation was also confirmed by the experimental result being talked about in Chapter IV (Figure 4-3). Bearing this in mind, we tried to inoculate yeast as fresh as possible henceforward. However, if cells are too old, several passages using fresh medium will always dilute out older ones, so it is not necessary to control the starting cell density for inoculation.

Densities of cells collected and initiated for induction are parameters that could affect the final HLA-DR1 surface display level. Yeast collected from SD-SCAA culture at both an OD_{600} of 3.0 and 6.0 exhibit optimal DR signals after being induced for 16 h regardless of what OD_{600} used for initiating induction (Figure 3-3A and B), suggesting that an optimal induction time could be somewhere closed to 16 h. A direct comparison of these two local optimums (Figure 3-3C) further confirms that cells collected at a smaller OD_{600}~3.0 and induced in SG-SCAA at a smaller OD_{600}~0.5 will display more detectable HLA-DR1 than those with bigger OD_{600}. Additionally, serving as a parallel control, results of yeast induced in YPG at different initial densities (Figure 3-3D) confirmed that collecting cells at a relatively earlier growth stage helped to display more HLA-DR1. This makes us to speculate that cells should be collected and induced within the exponential phase of growth or before stationary phase, after which the changes in cell wall structure could possibly influence surface display and detection of target proteins.

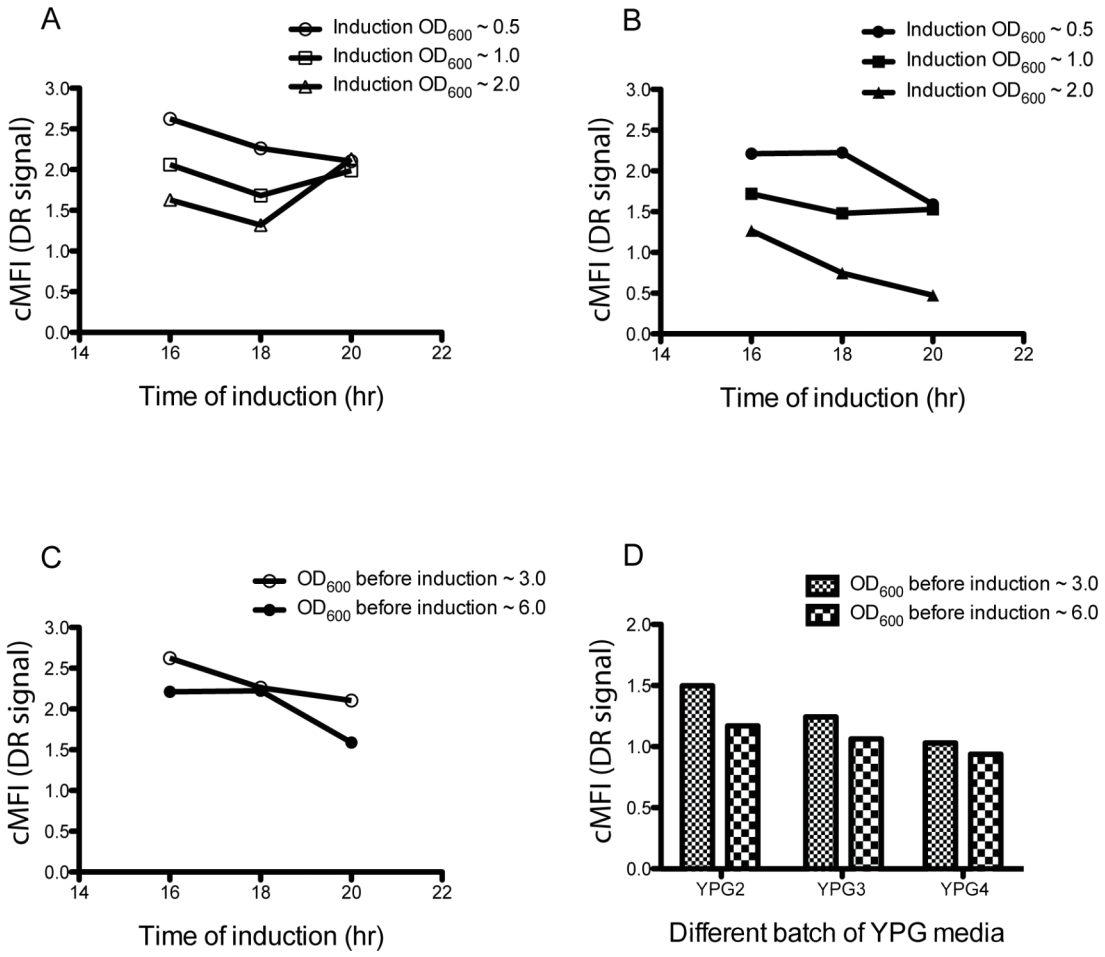

Figure 3-3 Effect of densities of cells collected and initiated for induction on HLA-DR1 surface display level. Co-displaying yeast grown up in SD-SCAA were collected at either **A.** OD_{600} ~ 3.0 (open symbols) or **B.** OD_{600} ~ 6.0 (filled symbols), spun down and resuspended in SG-SCAA to various densities at an OD_{600} of 0.5, 1.0 or 2.0 as indicated, and then induced at 30°C for various time lengths. **C.** Further comparison of the DR signal from yeast collected at the two different OD_{600} and induced with a starting OD_{600} of 0.5. **D.** Cells collected at different OD_{600} after enrichment were also resuspended in different batches of YPG media to an OD_{600} of 1.0 and induced at 20 °C for 20 h. After induction, all samples in were collected and labeled at the same time for flow cytomeric analysis. $cMFI_{(DR)}$ were calculated as described previously for yeast prepared at different conditions.

Before performing more experiments searching for optimal densities of cells collected and initiated for induction, we defined the growth pattern for co-displaying yeast strain in SD-SCAA medium. 2 ml of SD-SCAA fresh media in glass culture tubes were inoculated by two-day-old cell cultures (used to be kept at 4 °C) to an OD_{600} of 0.02 and incubated in a 30 °C shaker at 225rpm. OD_{600} measured at different time

points plotted in log scale against growth time (Figure 3-4) reveals a curve for yeast growing in SD-SCAA medium. The curve is actually linear before 20 h (OD_{600}~4.0), reflecting an exponential growth phase, and then starts to tilt down a little bit, indicating the beginning of stationary phase or a deceleration stage for entering stationary, where apoptosis of older cells in the culture tube emerges. Furthermore, the doubling time for yeast cells growing in SD-SCAA can be determined using this curve, which is around 3.5 hours. It is clear that OD_{600}~3.0 (~19 h post-inoculation) is within the exponential phase whereas OD_{600}~6.0 (~24 h post-inoculation) is towards the end of this partial growth curve closed to the stationary phase, confirming our assumption that cells collected in exponential growth stage tended to display more HLA-DR1 than those collected outside.

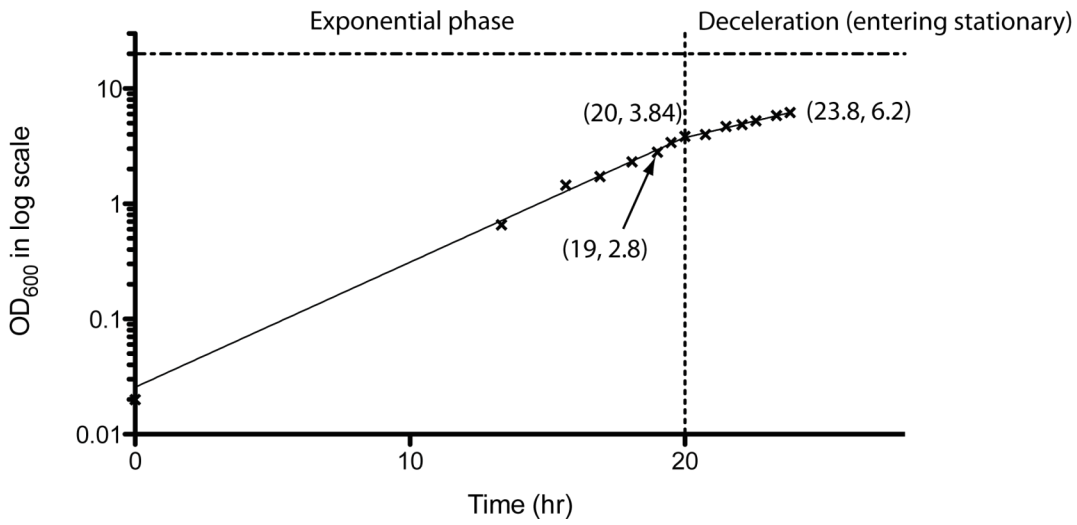

Figure 3-4 Cell growth curve for co-displaying yeast cultured in SD-SCAA.

To further evaluate the effect of growth stage and starting cell density on induction of proteins expression and association, more OD_{600} were examined (Figure 3-5). It is suggested that as long as not collected too early, the yeast will always display a similar

amount of detectable HLA-DR1 after proper induction, while a relative higher *cMFI* is observed at OD_{600}~5.2 (Figure 3-5A). The suggested range of OD_{600} is about 2.0 to 6.0, which is a little broader than that obtained from previous results. Meanwhile, the checking for different initial cell densities of induction culture shows an optimum around OD_{600}~0.5 (Figure 3-5B), which is consistent with previous observation (Figure 3-3A).

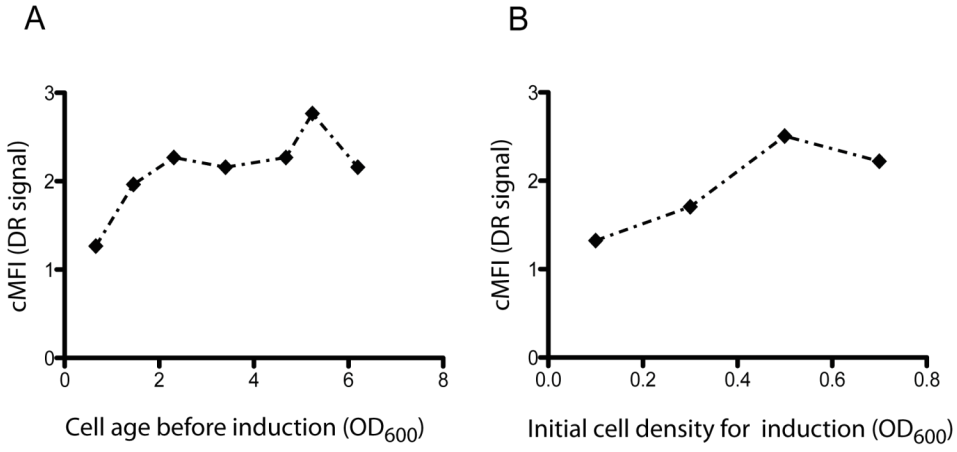

Figure 3-5 Further evaluation for the effect of cell age and initial cell density. A. Cells were grown up to various ages (OD_{600}), and then switched to SG-SCAA to an OD_{600} of 0.75. **B.** Cells were collected at an OD_{600} of 3.7 and resuspended in SG-SCAA to various densities (OD_{600}). In both experiments, immunofluorescent labeling was performed using biotinylated LB3.1 and streptavidin-PE-Cy5. $cMFI_{(DR)}$ were calculated and plotted against OD_{600} measurements representing densities of cells collected or initiated for induction.

In sum, one local optimum condition for culturing co-displaying yeast would be: inoculate a 2 ml SD-SCAA medium by a single colony (or an aliquot from a few days old culture) in a culture tube and grow cells in 225 rpm shaker at 30 °C for 16~20 h until the density reach an OD_{600} of 2.5 to 5.0, then switch cells into 2 ml SG-SCAA to an OD_{600}~0.5 and incubate in 225 rpm shaker at 30 °C for 16~18 h before harvest.

3.3.3. Optimization of immunofluorescent labeling

Using the culturing method optimized in previous section, we can have yeast to express and display an optimal amount of HLA-DR1 for surface detection. In this section, different immunofluorescent labeling conditions will be evaluated for maximizing the detectable signal for flow cytometric analysis. Two-step labeling procedures are wildly used for indirect immnofluorescence methods, because of their higher specificity and better fluorescence intensity. Thereby, all labeling procedures stated in this thesis work are in a two-step manner. In previous experiments, small changes have already been tried for evaluating different labeling conditions, such as using different amount of wash buffers, using different time length for incubations, and using different labeling reagents, etc. It turned out that washing 10^6 cells with 200–400 μl PBS + % BSA once or twice, incubating cells with different DR-specific antibodies for 30–60 min and using different fluorophores, all showed no obvious impact on yeast labeling and HLA-DR1 detection. However, Alexa Fluor dyes[149, 150] are said to be more stable, brighter, and less pH-sensitive than common dyes such as fluorescein, rhodamine and the newer cyanine series, so we will use them for the rest of this thesis work.

First, antibodies used for labeling HLA-DR1 were titrated for a better working dilution. L243 and Alexa Fluor 488 – goat anti-mouse antibodies were selected because both antibodies are commercially available with suggested concentration and working dilutions (Table 3-2). A series of dilutions for both were prepared around a reasonable range and tested for labeling and surface detection of sDR1 (Figure 3-6A and B). The roughly horizontal curve indicates that within the range tested, both L243 and Alexa Fluor 488 – goat anti-mouse are sufficient to couple most detectable DR molecules on

yeast surface without losing essential information. For L243 (Figure 3-6A), the result basically agree with the suggested dilution by supplier (Table 3-2), and more case-specifically, dilutions around 1:4–1:2 seems to be an optimum for yeast co-displaying FLU and HLA-DR1. For Alexa Fluor 488 – goat anti-mouse, dilution as low as 1:200 is still enough to cover all DR proteins labeled by primary antibodies L243 on the whole yeast surface, due to the high concentration of the antibody itself (Table 3-2).

Table 3-2 Antibodies with working dilution recommended by corresponding suppliers.

Antibody name	Concentration	Working dilution
L243	0.025 mg/ml	1:1
H6908	0.5 mg/ml	1:25
Alexa Fluor dye-secondary labeling reagents	2 mg/ml	NR

NR, no recommendation.

Up till now, we have already constructed a working condition for producing and detecting an optimal amount of HLA-DR1 on co-displaying yeast surface. To quantitatively determine the relative binding of FLU by HLA-DR1, the relative amount of FLU peptide on the surface needs to be calculated using Equation 2-1 and used as a comparison base line (will explain this in the discussion section in detail). Here, we evaluated the titration of antibodies used for labeling FLU-fused HA-tag so that more accurate amount of FLU peptide on yeast surface can be determined later for quantitation. Again, commercially available antibodies, H6908 and Alexa Fluor 488 – goat anti-rabbit were examined at a series of dilutions for HA signal by flow cytometry (Figure 3-6C and D). The Alexa Fluor dye conjugated secondary shows similar horizontal curve to that used to label HLA-DR1 (Figure 3-6D), whereas H6908 has to be

used in a relatively smaller dilution to stay concentrated enough for coupling all HA epitopes on yeast surface (Figure 3-6C). It also suggests that an optimum might be achieved while using a dilution around 1: 25, consisting with the vendor's suggestion (Table 3-2).

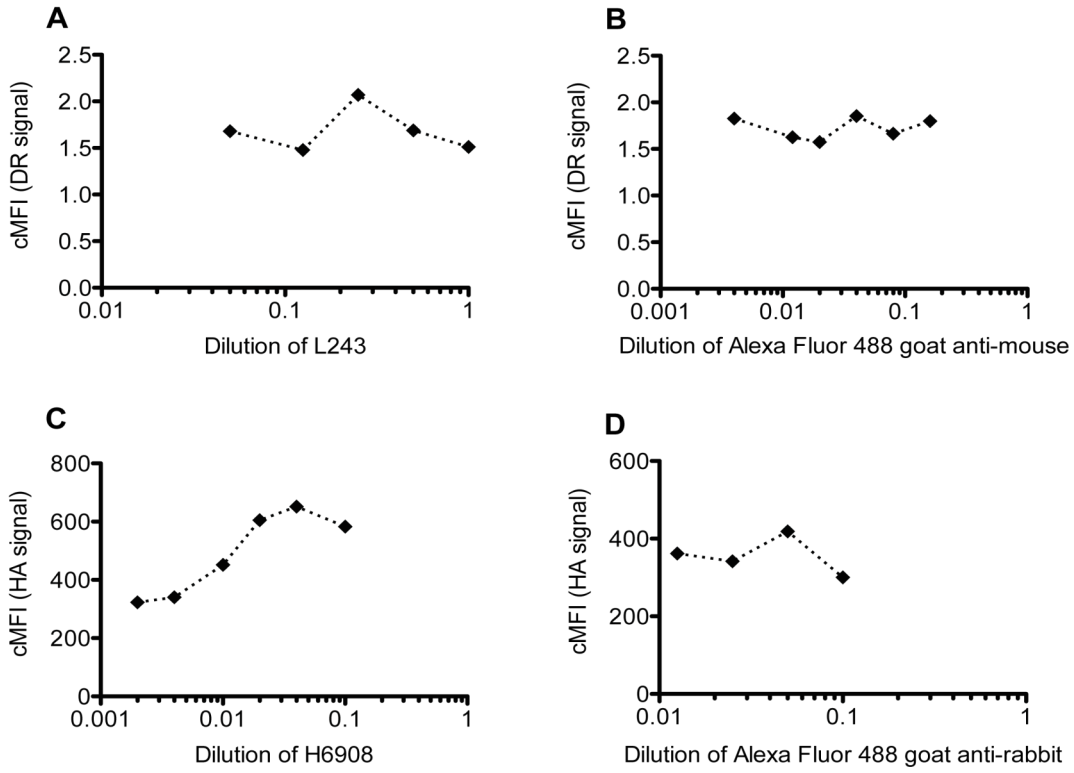

Figure 3-6 Titration of antibodies used for immunofluorescent labeling of HLA-DR1 and HA-tag on the surface of yeast. Yeast co-displaying FLU and HLA-DR1were induced using the culturing condition optimized in previous sections and labeled for either **A** and **B.** HLA-DR1, or **C** and **D.** HA epitope tag. **A** and **B.** L243 and Alexa Fluor 488 goat anti-mouse antibodies were used for labeling HLA-DR1. Staining was performed at room temperature (RT) for 40–50 min. **C** and **D.** H6908 and Alexa Fluor 488 goat anti-rabbit antibodies (highly cross-adsorbed with lowest cross-reactivity to mouse antibodies) were used for labeling HA-tag. Staining was performed at room temperature (RT) for 30 min followed by on ice for 10 min. Flow cytometric data was collected on BD LSR II flow cytometer.

3.4. Discussion

3.4.1. Optimal working condition for yeast co-display

Optimization for cultivation and post-induction manipulation of yeast co-displaying FLU peptide and HLA-DR1 protein has been carried out. Variables possibly determining or affecting the detectable DR-signal have been tested in groups in a reasonable manner searching for a local optimum that can be used reliably for quantitation. Quantitative analysis of flow cytometric data confirms that differences do exist when disturbing certain factors, so it is essential to adjust variables and optimize experimental conditions so that consistent and reliable flow cytometric results can be generated for quatitation.

Results suggest that the immunofluorescent labeling of HLA-DR1 or FLU on the surface of yeast cells is not a big factor for introducing errors for flow cytometry, as long as the concentration of labeling reagents used every time is closed to the optimal or exceed the saturation amount. The labeling procedure is relatively robust without causing too much variance when performing at different temperatures (RT or on ice) for a certain range of time lengths. Actually, the vessels used for labeling also have no impact on the data at all as suggested in the following chapters, where labeling were largely performed in 96-well plates instead of 1.5 ml eppendorf tubes. In those cases, more wash steps were included to remove as many non-specific coupling reagents as possible.

Indeed, cultivation of yeast is the dominant part, where significant differences between MFI values calculated from flow data could emerge. Therefore cares need to be

taken when growing and inducing yeast cells. First of all, minimal media is better than rich media on keeping the constancy of nutrient combination and providing the time efficiency at higher temperature (for example, 30 °C instead of 20 °C). Next, cells are better to be collected at the exponential growth stage before switching for induction (for example, OD_{600}~3.0), where too early or too late is neither recommended. Then, initial cell density of the induction culture should not be too high so that plenty of nutrients could be used for protein assembly (for example, OD_{600}~0.5). Finally, temperature and time length selected for induction should enable both optimal protein expression and protein-protein interaction (for example, 16h). An optimized protocol for yeast co-displaying FLU peptide and sDR1 is suggested as follow:

- *Culturing:*
    1) Inoculate a 2 ml SD-SCAA medium by a single colony (or an aliquot from an old culture, but not too old) in a culture tube
    2) Put the tube into a 225 rpm shaker with an angle and grow cells at 30 °C for 16–20 h until the density reaches an OD_{600} of 2.5 to 5.0
    3) Collect 10^7 cells by centrifugation and withdraw most medium
    4) Resuspend cell pellet in 2 ml SG-SCAA medium to an OD_{600}~0.5
    5) Put the tube back to the same shaker and induce cells at 30 °C for 16–18 h
    6) Harvest enough cells by centrifugation and withdraw as much medium as possible for fluorescently labeling. (Alternatively, cell culture can be stored at 4 °C for a couple of days without significantly affecting the flow cytometric results, but this is not recomended)
- *Labeling (for 10^6 cells):*

1) Resuspend 10^6 cells by 300–400 μl ice cold PBS + 1% BSA in a 1.5 ml eppendorf tube and wash cells by thoroughly vortexing
2) Spin cells down and withdraw as much supernatant as possible
3) label cells by primary antibodies appropriately diluted in PBS + 1% BSA (specifically, 1:2.5 for L243 and 1:25 for H6908) at room temperature for 30 min and then on ice for 10 min
4) Spin cells down and repeat step 1) and 2)
5) label cells by secondary reagents appropriately diluted in PBS+1%BSA (specifically, 1:80 for all kinds of Alexa Fluor dye conjugated secondary reagents) at room temperature for 25–30 min followed by on ice for 10 min
6) Repeat step 4)
7) resuspend cell pellet by 500–700 μl ice cold PBS + 1% BSA for flow cytometry

NOTE: these procedures are actually case sensitive, so for cultivation and labeling of yeast cells under other nutrient dropout conditions, they might not be suitable and another optimum might need to be characterized.

3.4.2. Quantitative analysis of co-displayed FLU and HLA-DR1

We now proposed a quantitative analysis for defining the relative binding of peptide to MHC-II molecules using optimized yeast co-display. First of all, both a co-displaying yeast strain and a peptide-only-displaying strain have to be simultaneously stained for peptide-fused HA epitope tag and MHC-II for flow detection. As described in Chapter 1 (1.4.2), cross-reactivity of labeling reagents and fluorescent signal overflowing are both issues we should pay attention to when performing double labeling. Henceforward, L243/Alexa Fluor 647 – goat anti-mouse and H6908/Alexa Fluor 488 – goat anti-mouse

will be used for labeling HLA-DR1 and FLU respectively when performing quantitative analysis. Both secondary reagents purchased form Invitrogen are 'highly cross-adsorbed' with lowest cross-reactivity for other species, ensuring no cross-reaction taking place during labeling (this has been checked and proved in some experiments with data not shown). The far separation of excitation and emission wavelengths between these two fluorophores also minimizes signal overflowing, where no compensation required at all.

Next, flow cytometric data (take the two samples in Figure 2-9C and D for example) were analyzed using Flowjo (Figure 3-7). The display level of FLU peptide on yeast surface can be determined by calculating *cMFI* for HA signal using Equation 2-1. Different from Equation 3-1, the background correction here is based on the same sample. One reason is that there are two distinct populations in the same culture, indicating an obvious population of negative cells missing FLU-displaying plasmids and having no contribution to the surface display of FLU at all. Another reason is that in contrast with the dependency of detectable amount of HLA-DR1 on FLU expression, displaying and detecting FLU peptide on yeast surface is apparently independent. Therefore, eliminating the background will not affect analysis of DR signal at all, but increase the accuracy as shown in the following. To reflect how well HLA-DR1 bind to FLU peptide on the surface of yeast, the ratio of $MFI_{(DR)}$ to $cMFI_{(HA)}$ instead of $MFI_{(DR)}$ (in Equation 3-1) was used for negative subtraction and normalization in the following equation:

$$DR\text{ - }ratio = \frac{\left[\frac{MFI_{(DR)}}{cMFI_{(HA)}}\right](+) - \left[\frac{MFI_{(DR)}}{cMFI_{(HA)}}\right](-)}{\left[\frac{MFI_{(DR)}}{cMFI_{(HA)}}\right](-)} \qquad \text{Equation 3-2}$$

where $\left[\frac{MFI_{(DR)}}{cMFI_{(HA)}}\right](+)$ and $\left[\frac{MFI_{(DR)}}{cMFI_{(HA)}}\right](-)$ represent the ratio on the surface of the co-displaying yeast and peptide-only-displaying yeast, respectively.

This equation can generate relative binding values for different MHC-II proteins to different peptides by comparing *DR-ratio* of corresponding co-expressing yeast. In the following two chapters, we are going to introduce mutations in FLU peptide and HLA-DR1 protein to evaluate relative binding of different DR mutants to different FLU variants, where *DR-ratio* normalized by that of the wild type co-displaying strain will be compared frequently.

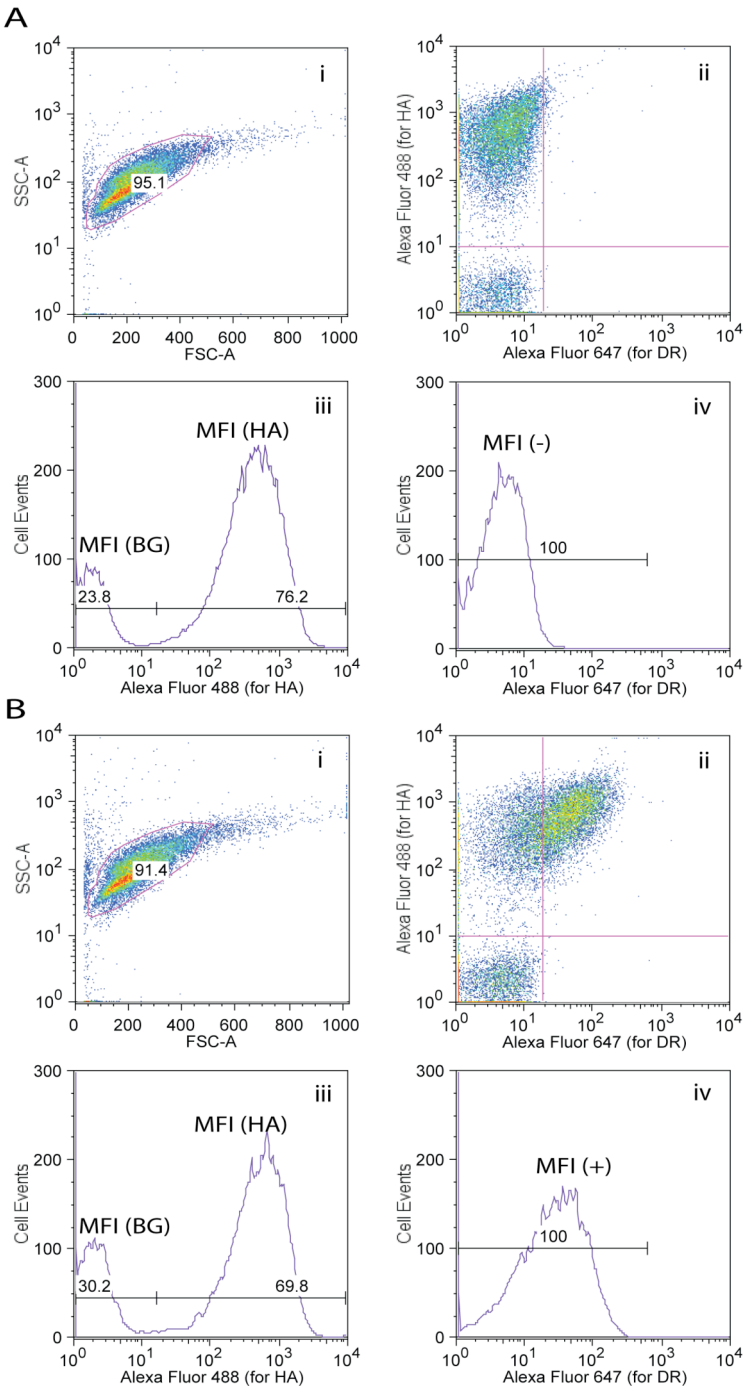

Figure 3-7 Quantitation for relative binding of MHC-II to peptides on the surface of yeast. Flow cytometric data of **A.** peptide-only-displaying yeast or **B.** co-displaying yeast were first assessed (i) to gate the gate the correct cell population (red line). Gated cells was then analyzed in (ii) a dot plot showing FLU and HLA-DR1 surface display level respectively, where HA+ and HA- populations can be further gated. An easier way to gate and generate mean values ($MFI_{(HA+)}$ and $MFI_{(HA-)}$) for the two populations is using (iii) a univariate histogram for HA signal only. Detection of HLA-DR1 on yeast surface is largely due to their binding to the displayed FLU peptide, so HA+ subgroup was further plotted in (iv) the univariate histogram for generation of $MFI_{(DR)}(+)$ in **B** and $MFI_{(DR)}(-)$ in **A**.

Chapter 4. Identification of P1 pocket profile and prediction of HLA-DR1-specific ligands

4.1. Introduction

Naturally evolved immune systems can function properly for detecting, preventing and eliminating foreign pathogens without influencing most of self bio-systematic balances, predominantly due to their capability to specifically recognize and associate with epitopes (usually short peptides) in a target pathogen. Hence, immunotherapy and vaccination used to rely on the original form of pathogens for initiating immune responses now have an alternative strategy via directly applying epitope-based reagents to stimulate specific immune responses[28]. Identification and prediction of pathogenic epitope-like peptides have become more and more essential.

CD4+ T cell-mediated immune responses are restrictedly initiated by TCR-specific epitopes associated and presented by MHC-II molecules. Thus, epitope-specificity of CD4+ T cells is largely determined by antigenic peptide binding specificity of MHC-II proteins. A few MHC-II alleles expressed by an individual able to present peptides derived from many kinds of self or foreign antigens indicates the fact that one MHC-II allele is supposed to recognize and bind a group of structurally similar peptides for selection by T cell receptor, as a result of which CD4+ T cell-mediated immune responses will be initiated. Investigating the promiscuous peptide sequences recognized by MHC-II molecules will greatly help to define peptide-binding specificity of CD4+ T cells and predict T cell specific epitopes, which can be used for vaccine design and immunotherapeutic application.

The earliest and most directly method of defining MHC-specific epitope peptides is to analyze and pool-sequence peptides eluted off peptide/MHC-II complexes purified from specialized APC cell lines[22, 60, 151-156]. This method provides enormous information for defining binding motif of MHC-II molecules by aligning naturally processed peptides, but lack of freedom on manipulation of residues at positions of interest for determining side chain preference and detecting unknown MHC-II ligands. Thereafter, people started to synthesize soluble peptides and assay for their capability to bind "empty" MHC-II molecules on the surface of APCs[124] or solubly produced from an EBV-transformed homozygous B cell line[18, 118-123]. Theoretically, one can examine the binding of any peptide to a given MHC-II, but the disadvantage of labor intensity on synthesizing peptides actually makes it a really high cost and low throughput method even with a modern peptide synthesizer. In addition, the difficulty on acquiring and culturing mammalian cell lines could be another problem, which was partially solved by expressing and purifying soluble "empty" MHC-II proteins in other easily manipulated expression systems[106, 108].

Investigation of various peptides with different sequences and lengths for MHC-II binding requires high throughput experimental approaches and probably computational prediction. For example, libraries of peptides can be constructed by using phage display, and screened by mixing phage particles with biotinylated HLA-DR1[20, 90]. This engineering system allows generating information of hundreds of DR-specific peptide sequences, which is suitable for identifying peptide binding motif and anchor preferences of MHC-II proteins[145] or even predicting promiscuous MHC-II ligands and T cell specific epitopes with the help of computational programs TEPITOPE[23]. As one

of the most successful experimental approaches, phage display and peptide screening have little to no ability to tell the difference of affinity between different peptides bound by HLA-DR1, which, complimentarily, has to be further determined by synthesizing peptides screened out of phage library and mixing them with soluble produced HLA-DR1 molecules for binding assay. This greatly limits the applicability of phage display on quantifying peptide-binding specificity of given MHC-II molecules.

Other than above experimental approaches, people have also developed several computational methods for predicting immunogenic epitopes[157]. Some are algorithms or programs such as TEPITOPE, depending on binding motifs and ligand matrices of MHC-II determined by experimental data. Others are models based on 3D structure and molecular dynamic simulation, for example, a predictive computational model by global minimization of potential energy between interacting atoms was proposed for examining interactions between the entire array of naturally occurring amino acids and the P1 pocket within the binding groove of HLA-DR1[158]. However, these estimations based on simplified models are less meaningful without support of experimental data.

Yeast co-display we have developed and optimized in previous chapters on the contrary, not only has the ability to create libraries of various peptides for screening MHC-binders experimentally, but also contains a potential to quantitatively determining relative binding of each peptide to MHC-II. In this chapter, we will examine these features of the novel engineering system and show its applicability in characterization of side chain preference of HLA-DR1 and in prediction of MHC-II ligands.

Experimental results from other work[20, 90, 145] as well as predictive molecular dynamic calculations[158] confirm the observation obtained from X-ray crystallographic

studies of HLA-DR1[59] that there are five pocket-like regions within the peptide-binding site of HLA-DR1 molecule (Figure 4-1), which are critical for holding specific side chains of anchor residues in peptide and determining peptide binding specificity. Among these pockets, the largest and deepest one closed to N-terminus of the bound peptide (named as P1), is demonstrated to be the dominant position for peptide-HLA-DR1 interaction. The hydrophobic feature of P1 pocket demands most accommodated peptides containing an anchor residue with a large hydrophobic side chain towards their N-terminus. Specifically, aromatic residues such as Trp, Tyr, Phe show the most favorability to P1 pocket, followed by other hydrophobic residues such as Met, Ile, Leu, and Val. The known anchor preference of P1 pocket can be used to validate the applicability of yeast co-display for quantitatively determining peptide-binding motif. Pockets other than P1 are consecutively accommodating the fourth, sixth, seventh and ninth residues from P1 anchor residue of the associated peptide and designated as P4, P6, P7, P9 respectively, which exhibit less importance than P1 pocket but noticeable additive effects on peptide binding[145].

In this chapter, site directed mutagenesis was first coupled for replacing and saturating P1 anchor of FLU peptide with all natural amino acids and yeast co-display was used for quantifying the relative binding of these P1 variants to HLA-DR1. Subsequently, a small library of peptides was constructed and screened by directed evolution approaches allowing isolation and identification of HLA-DR1-specific peptides in a high throughput manner.

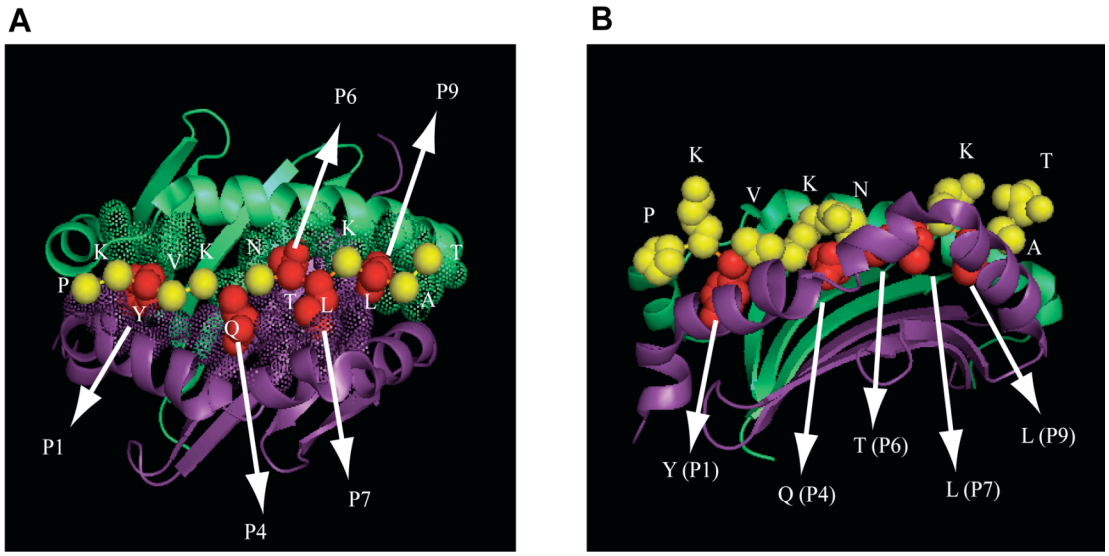

Figure 4-1 Anchor-Pocket regions in the structure of HLA-DR1 associated with FLU peptide. A. Top view of the peptide-binding groove formed by α1 and β1 domains of HLA-DR1, represented by green and purple cartoon structures respectively. Residues in FLU peptide are shown as spheres, where red ones depict anchors with side chain buried inside pockets formed by residues in the binding groove, represented by green or purple dots**. B.** Side view perpendicular to the FLU peptide extending direction with side chain of all peptide residues shown. Red ones again suggest the five main anchors and yellow ones either point up to bind T cell receptor (especially the four in the middle) or extend out of the binding groove. The deepest anchor (P1) near the amino-terminus of FLU peptide dominates the peptide association.

4.2. Materials and Methods

4.2.1. P1 variant-expressing plasmid construction and yeast transformation

Site directed mutagenesis following "Quikchange II" protocol (as described in Chapter 2 section 2.2.1) was used to introduce residue substitutions at P1 anchor of FLU peptide for construction of ptFLUX (X denotes any single letter designation for amino acids). For example, ptFLUF was generated via PCR using a pair of reverse complementary primers W53 (5'–GCTAGCCCTAAGTTCGTGAAGCAGAATAC) /

W54 (5'–GTATTCTGCTTCACGAACTTAGGGCTAGC) carrying mutated nucleotides (underlined) encoding Phe at P1 anchor.

Plasmid ptFLUX was transformed into either EBY100 or EBY100 with ptsDR1 by electroporation for expressing P1-variant of FLU peptide or co-expressing P1 variant and soluble HLA-DR1 in yeast.

4.2.2. P1 anchor preference assay via yeast co-display

Yeast were prepared using previously optimized culturing and labeling conditions with a minor modification. 96-well plates instead of 1.5 ml Eppendorf tubes were used sometimes for a high throughput labeling. For washing steps, cells in each well of the plate were spun down by centrifugation of the plate at 4 °C, 3000 × g for 5 min and washed with more than 200 μl ice cold PBS + 1% BSA at least twice. In the final step, cells were resuspended in 150–300 μl PBS + 1% BSA for flow cytometric analysis. *DR-ratio* (using Equation 3-2) was then calculated for each co-expressing yeast strain and normalized by wild type (FLU peptide) for quantitatively determining the relative binding of MHC-II to different peptide variants on the surface of yeast.

4.2.3. Construction of peptide library

An oligonucleotide mixture W124 (5'–TGGTGGTGGTTCTGCTAGCNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSNNSGGAGGTGGACCCGGGAAGCCTATAC, PAGE purified; IDT) with 15 degenerate NNS codons was designed for generating a library of 15 amino-acid-long random peptide sequences. Another two primers W123 (5'–CGCTCTGCAGGCTAGTGGTGGTGGTGGTTCTGGTGGTGGTG

GTTCTGGTGGTGGTGGTTCTGCTAGC) and W125 (5'–TTGGCATGCTTAAGTAG AATCGAGACCGAGGAGTGGATTTGGTATAGGCTTCCCGGGTCC) allowed amplification and extension of the degenerate oligonucleotides by PCR such that both ends had at least 50 base pairs (bps) sequence homologous to the linear backbone vector, separately prepared by digesting ptFLUD with NheI and XmaI. Both the PCR generated insert and the open backbone vector were prepared at a large scale and concentrated by Microcon YM-10 cartridges (Millipore, Billerica, MA) before the final purification by agarose gel electrophoresis. Purified DNAs were precipitated using Pellet Paint Co-Precipitant kit (Novagen) following the manufacturer's instructions after estimating the concentration by absorption at 260 nm. Approximately 2.25 μg degenerated gene inserts and 1.5 μg linear vectors were mixed in 2 μl of sterile nuclease-free water and transformed into 100 μl freshly made electro-competent cells (EBY100 previously transformed with ptsDR1) via two separate eletroporation reactions. Transformants were recovered from the two cuvettes (Fisher) and transferred into 80 ml SD-SCAA liquid medium for construction of a yeast library co-expressing randomized peptides and HLA-DR1. Samples of the yeast culture were plated at different dilutions on SD-SCAA agar plates for estimation of peptide library size (number of distinct transformants), after which the culture was incubated at 30 °C to enrich real transformants. Once the library size was determined, the grown up culture (OD_{600}~7.0) was passaged at least twice in 50–150 ml fresh SD-SCAA medium to eliminate non-transformants. During each passage, number of cells transferred to fresh medium was at least 10 fold larger than the library size to maintain the diversity. A separate electroporation reaction was performed by transforming only 0.75 μg vector but no insert into 50 μl competent cells, which were

plated at different dilutions as well for estimation of background caused by non-linearized vector transformants.

4.2.4. Library screening and FACS sorting

After non-transformants being diluted away, the combinatorial library of yeast transformants bearing both random peptide sequence-expressing plasmid and ptsDR1 were prepared under the optimal condition as described in a larger scale. Briefly, library culture was enriched in SD-SCAA to an OD_{600} around 4.0 and transferred to 50 ml SG-SCAA in a 250 ml baffled flask to initiate induction at an OD_{600} of 0.5 for 17–18h. A scaled up double staining was performed for 1.5×10^7 cells from the yeast library using antibody pairs H6908/Alexa Fluor 488 – goat anti-rabbit and L243/Alexa Fluor 647 – goat anti-mouse at similar dilutions as suggested in the discussion of Chapter 3. Labeled yeast was then sorted on a FACSVantage flow cytometer (Becton-Dickinson) with sorting gate set to collect double-positive cells and sorting speed set around 3,000 events/s. Sorted cells in citrate buffered SD-SCAA (pH 5.0, to minimize bacterial contamination) were enriched and prepared in an enough scale to cover the diversity for next round of sorting until yeast with desired phenotype (double-positive for Alexa Fluor 488 and 647 signals) occupies a constant proportion of the total cell population.

4.2.5. Positive clone isolation and plasmid recovery

A dilution of cell samples were spread on SD-SCAA agar plates after the fourth round of sorting for isolation of individual clones. 10 colonies were picked randomly from the plate and inoculated for growth and induction followed by labeling for flow cytometric analysis as described. Plasmids in clones with the desired phenotype were

rescued using Zymoprep Yeast Plasmid Miniprep Kit (Zymo research, Orange, CA) and transformed into electro-competent DH5α for discrimination of peptide-expressing plasmids from ptsDR1 by *E. coli* colony PCR using vector-specific primer pairs 0118 / W126 (5'–GTACGAGCTAAAAGTACAGT). Peptide-expressing plasmids were finally isolated from the right *E.coli* cells by normal miniprep and sequenced as described in section 2.2.1. All recovered plasmids were retransformed into fresh yeast to verify that plasmids conferred the observed binding phenotype.

4.3. Results

4.3.1. Quantitative mapping of P1 pocket profile by saturation mutagenesis

Side chain preference for anchor residues accommodated by pockets in the peptide-binding site of MHC-II (defined as "anchor preference" or "pocket profile") can be determined by amino acid saturation of selected anchor position for a target MHC-II-specific peptide. For instance, P1 anchor-pocket position was chosen to be further characterized for its anchor preference in this section. Tyr308 at P1 anchor of FLU peptide was replaced by all other natural amino acids (except for Cys, which could form disulfide bond) using site-directed mutagenesis. The sequence of resulting P1 variant-expressinging plasmids ptFLUX (X refers to any single letter designation for amino acids) were confirmed by sequencing gene of interest portion of these plasmids.

- **Verification for expression of P1 variants on yeast surface**

Yeast expressing P1-variants only (Figure 4-2) or co-expressing P1-variants and HLA-DR1 (Figure 4-3) co-stained for HA and V5 epitope tags indicates that HA-tag and

V5-tag are always present on the surface of yeast simultaneously and the relative ratio of these two tags is constantly closed to 1 for all P1 variants, which is independent of HLA-DR1-secretion. Double-positive cell population also represented a relatively constant percentage of total cells for both types of stains: 60–80% for the one without expression of HLA-DR1 and 40–60% for the one with. The HLA-DR1-expressing yeast cells used in this experiment were inoculated from relatively old plates, so the expression level of P1 variants was a little lower than that of strains without HLA-DR1 expressing plasmid, confirming previous conclusion in Chapter 3 that inoculation of older cells tended to result in lower detectable signals of displayed proteins. Other independent experiments suggested that expression level of P1 variants has no difference for yeast expressing HLA-DR1 or not (data not shown).

These results and observations convinced that mutation at P1 anchor of FLU peptide and introduction of HLA-DR1-expressing plasmid both have no impact on proper expression and display of various Aga2-P1-variant fusions by yeast.

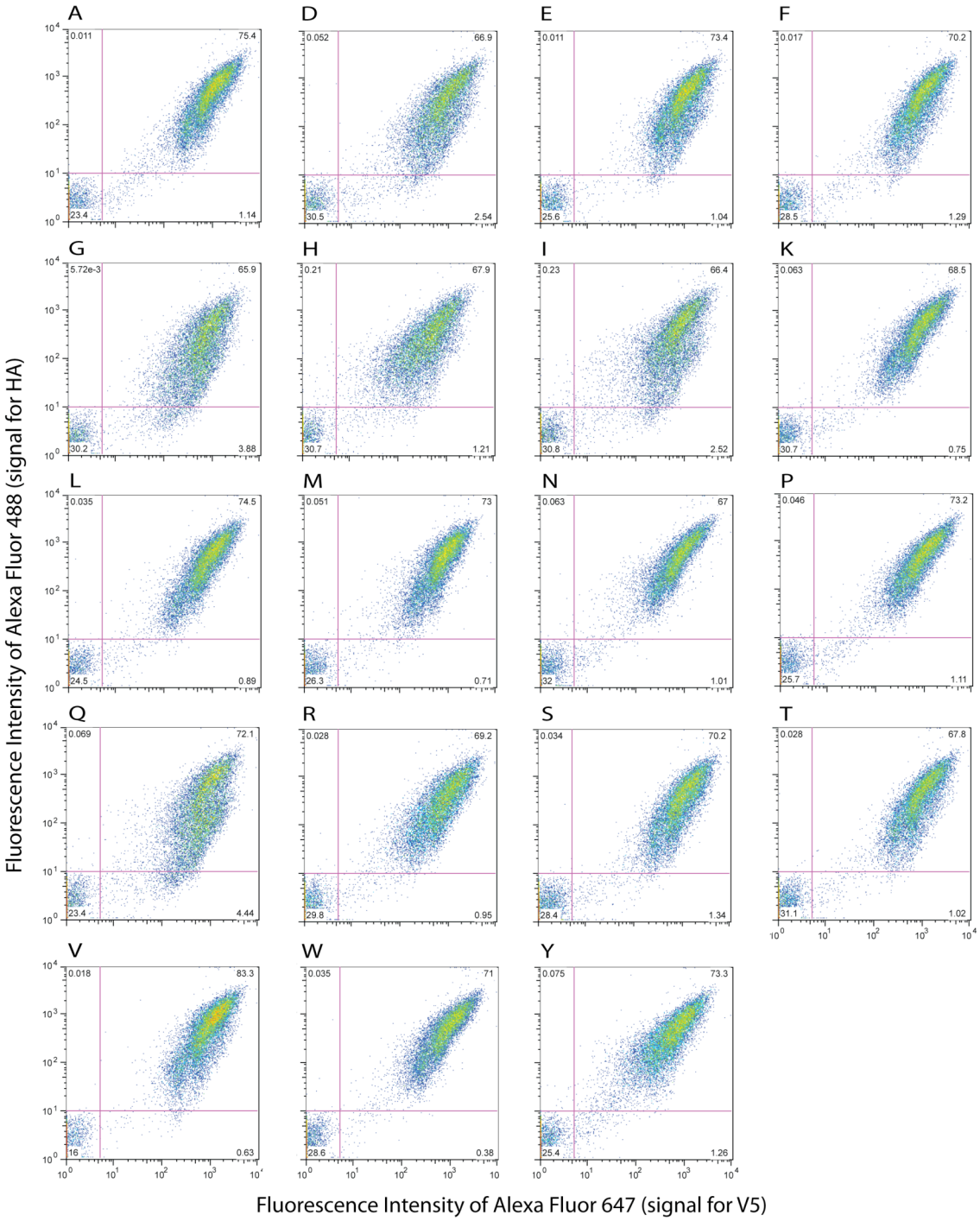

Figure 4-2 Fluorescent detection of HA-tag and V5-tag simultaneously on the surface of yeast displaying P1 variant of FLU peptide. The specific residue substitution at P1 anchor position of surface-displayed FLU peptide is indicated to the top left of each plot using the single letter designation of amino acids. Alexa Fluor 488 and 647 were used for labeling HA-tag and V5-tag respectively. Data were collected on BD LSR II flow cytometer.

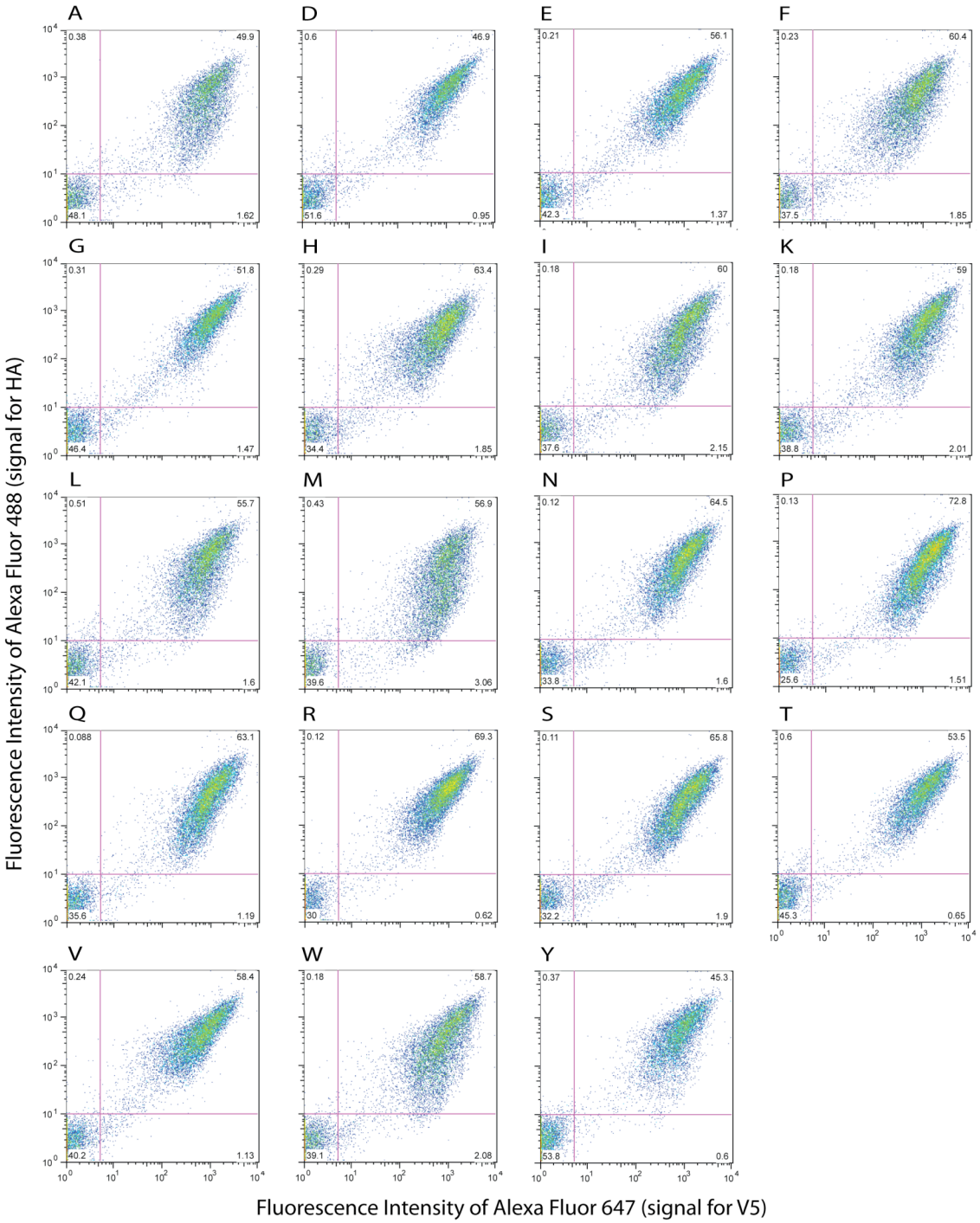

Figure 4-3 Fluorescent detection of HA-tag and V5-tag simultaneously on the surface of yeast co-expressing Aga2p-P1-variant and HLA-DR1. The specific residue substitution at P1 anchor position of surface-displayed FLU peptide is indicated to the top left of each plot using the single letter designation of amino acids. Alexa Fluor 488 and 647 were used for labeling HA-tag and V5-tag respectively. Data were collected on BD LSR II flow cytometer.

- **Relative binding of HLA-DR1 to P1 variants determined by yeast co-display**

Knowing the proper expression of P1 variants, yeast cells co-expressing one P1 variant (except for Cys) and HLA-DR1 molecules were double labeled using anti-HA-tag and anti-DR antibodies in order to detect variant-binding dependent DR signal. A two-dimensional dot-pot (Figure 4-4) exhibits two separated cell populations for all samples, where the double-negative auto-fluorescent background occupied approximately 20–30% of total cell population, indicating that all P1 variants were displayed as expected on yeast surface. However, some strain shows only HA-positive while others contain a double-positive cell population. The double-positive population bearing yeast strains were displaying following P1 variants: Phe, Tyr, Trp, Met, Leu, Ile, Val respectively, which all contain hydrophobic side chain, consistent with the side chain preference of P1 pocket found elsewhere[145].

The relative peptide binding of HLA-DR1 to all these P1 variants can be further estimated by using the quantitative analysis proposed in the discussion of Chapter 3. An average generated from four independent flow experiments (Figure 4-5) suggests that the peptide binding affinities follow the order Phe ≥ Tyr ≥ Trp > Met ≥ Leu ≥ Ile ≥ Val, with other amino acids tested yielding much weaker or no binding to DR1, in agreement with the peptide binding motif of DR1 determined using phage display and/or other methods[145]. However, compared to other methods, our system is even more sensitive for detecting the subtle differences between all those weaker or non-binders, which would be especially useful for characterizing naturally existing weak interactions between epitopes and epitope-specific MHCs.

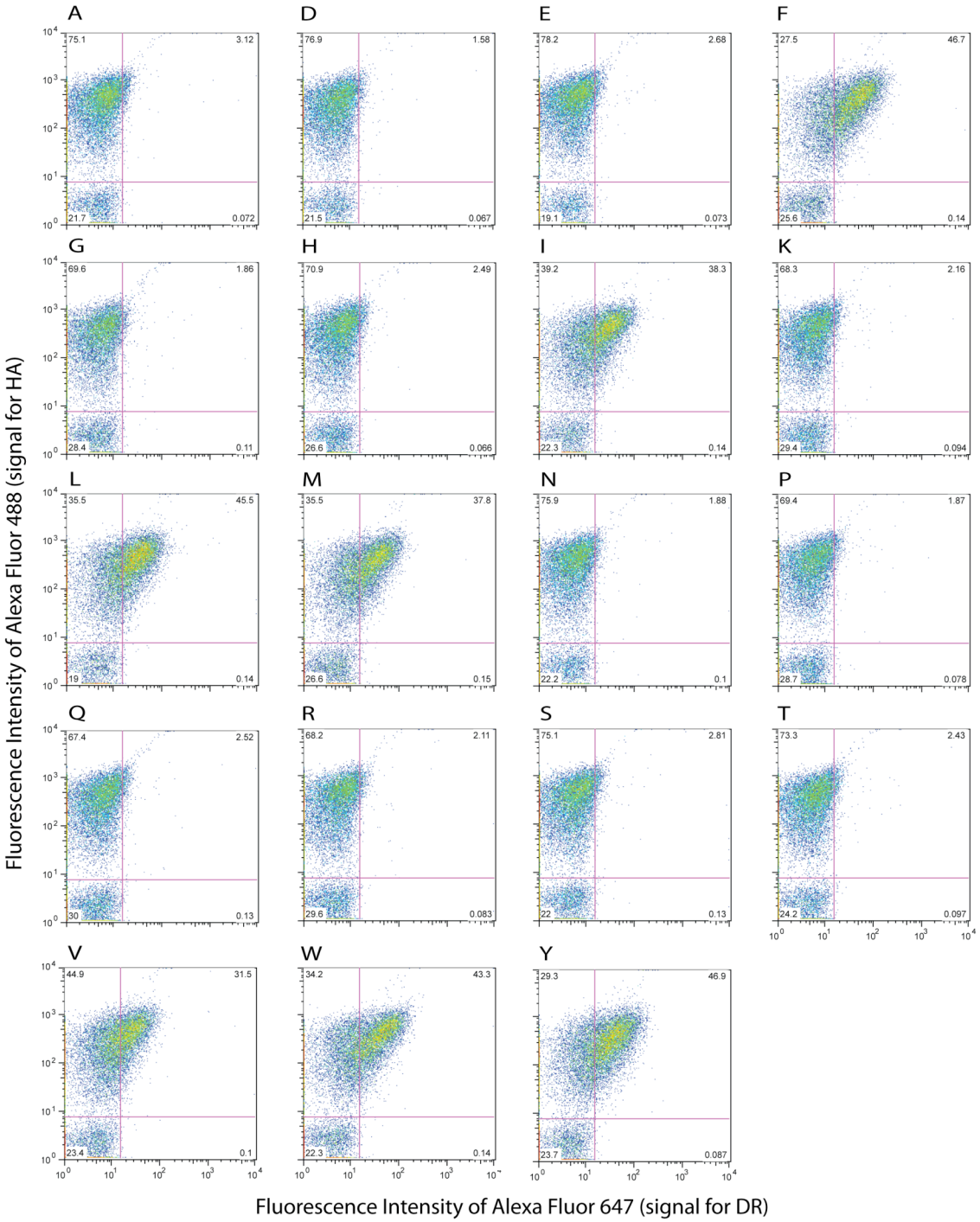

Figure 4-4 Simultaneous detection of P1-variant and HLA-DR1 on the surface of co-expressing yeast. Yeast Co-expressing HLA-DR1 and a peptide with the indicated P1 residue were co-stained for HA-tag and HLA-DR1 to determine the surface level of both peptide and HLA-DR1 and analyzed by flow cytometry.

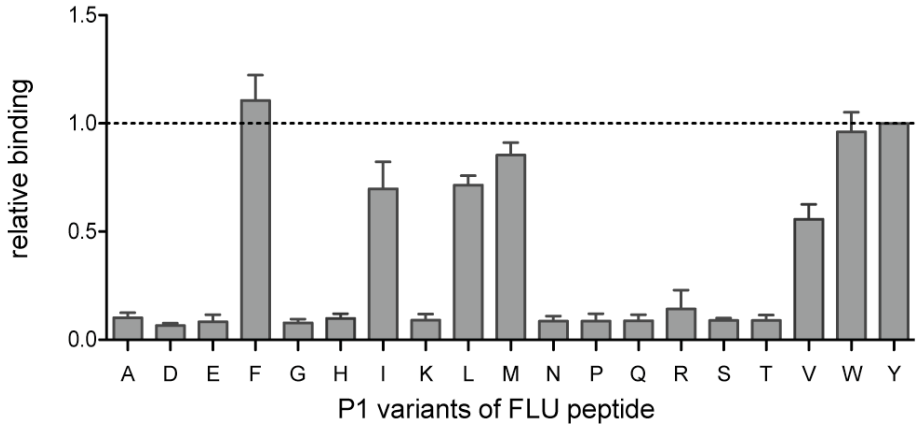

Figure 4-5 P1 pocket profile of HLA-DR1 analyzed by yeast co-display. *DR-ratio* was calculated using Equation 3-2 based on flow cytometric data (Figure 4.4) of yeast co-expressing HLA-DR1 and FLU analogues with the indicated P1 anchor residue. Relative binding level of HLA-DR1 to different peptide variants were generated by normalizing *DR-ratio* by the value for yeast co-displaying FLU with Tyr at P1 and HLA-DR1, indicated as the dashed line. Error bars represent standard error of the mean (SEM) determined from four independent experiments.

The ability to discriminate the difference of peptide binding among P1 variants also confirms that yeast co-display is an appropriate approach for quantitatively characterizing peptide/MHC-II interaction. Using similar method, determining anchor preference of any pockets of a selected MHC-II molecule will be straightforward. Theoretically, a matrix could be generated for all possible combinations of anchor residues allowing peptides to bind any given MHC-II protein, which will be greatly helpful for predicting MHC-II specific ligands.

4.3.2. Determination of HLA-DR1 specific ligands by screening peptide library

By saturation mutagenesis at a specific anchor of peptides, information such as anchor reference of a specific pocket can be generated, however, the labor intensity will limit the output once more anchors are involved. To generate more information about

the whole amino acid sequence of peptides bound by MHC-II molecules or to further determine peptide binding specificity, alternative approaches can be screening of yeast co-expressing HLA-DR1 and library of randomized peptides for DR-positive clones.

- **Construction and screening of peptide library**

Degenerate oligonucleotides encoding randomized peptides with 15 amino acids in length were used for amplification of mutated gene of interest cassettes [$(SGGGG)_3$ linker//randomized peptide//V5-tag] by PCR. These mutated gene inserts along with enzymatic-linearized vector (from ptFLUD instead of ptFLU in order to minimize false positive caused by wild type FLU peptide) were then co-transformed into yeast strain bearing HLA-DR1-expressing plasmid by electroporation. The insert and vector were ligated to form randomized peptide-displaying plasmids by homologous recombination inside yeast cells. After several passages, yeast cells displaying millions of distinct peptides were selectively enriched with background or parent strains all diluted out. Size of the yeast library was estimated by plating a series of diluted samples of the library culture on agar plates. For the first trial, the library size was designed to be $1–2 \times 10^6$, which is amenable by FACS screening and later characterization. Once the procedure was approved to be applicable, the size of yeast library can be scaled up easily to 10^8 or even 10^9 by increasing the concentration of insert and vector DNA and the number of electroporation reactions to obtain more information.

The combinatorial peptide library was cultured and double labeled for HA-tag and HLA-DR1 signal using the same procedure as that for wild type co-displaying strain at a bigger scale, and sorted by fluorescent-activated cell sorting (FACS) on a FACSVantage cell sorter for four rounds. The sorting gate was set for HA and HLA-DR1 double-

positive cell population and the sorting stringency was increased gradually (not too much) every round by moving the gate toward higher fluorescence intensity. Significant amount of double positive was already observed after the second around (Figure 4-6). Ratios of sorted cells to total cells in the four rounds are 0.6%, 5.8%, 14.8%, and 7.9% respectively. The percentage of cell population within the sorting gate increased each round for the first three rounds and stayed constant (dropped a little due to the restrict stringency) after the third one, suggesting no more sorting rounds needed. According to the ratio of sorted cells in each round, we can also estimate that the real number of distinct positive clones sorted out of peptide library with a size of $1–2 \times 10^6$ is roughly $10–10^2$.

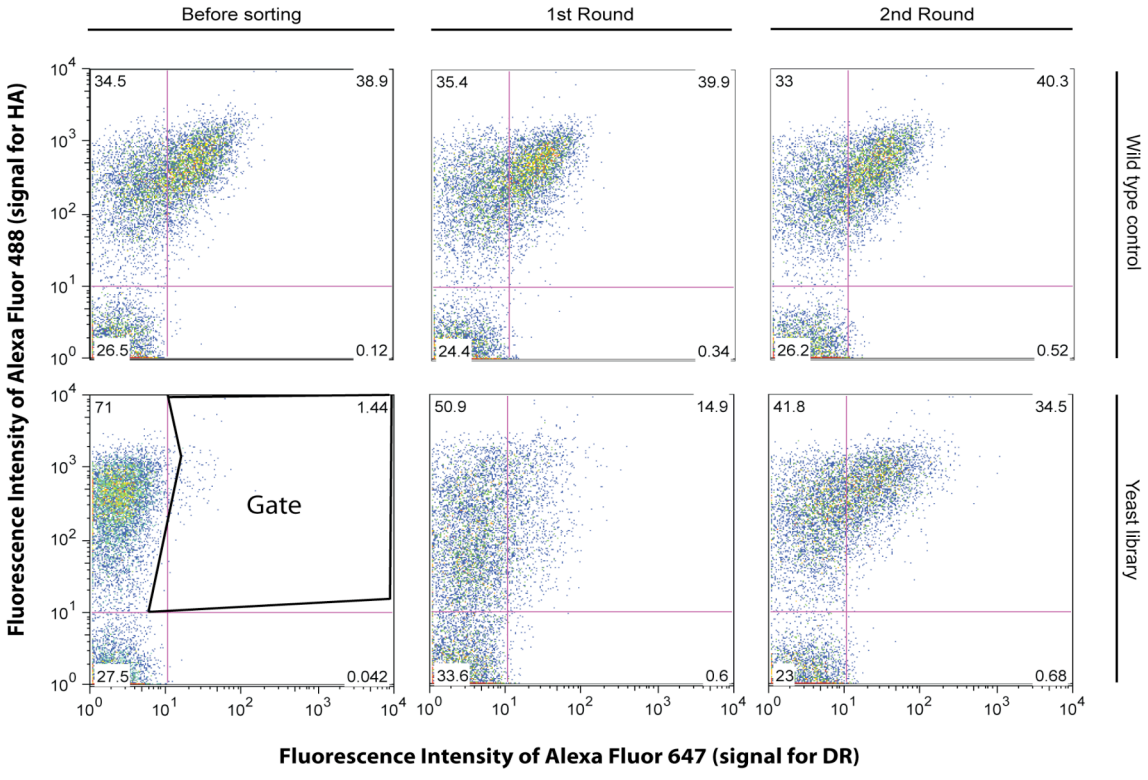

Figure 4-6 Fluorescent-activated cell sorting of a library of randomized peptides for HLA-DR1 specific ligands by yeast co-display. Wild type (yeast co-displaying FLU and HLA-DR1) was used as a double positive control for each round. The original sorting gate was draw with a little relaxed stringency to ensure the recovery of all possible genuine HLA-DR1 binders in the library.

- **Characterization of HLA-DR1-specific ligands extracted from yeast libray**

Samples of the sorted cells were plated on SD-SCAA agar plates to isolate individual clones. Ten colonies, a proper sample number representing the sorted positive strains, were inoculated and induced for further characterization of the interaction between HLA-DR1 and the displayed peptide. The ten selected yeast clones were double labeled for HA-tag and HLA-DR1 and applied for flow cytometric analysis (data not shown). $cMFI_{(DR)}$ values corresponding to the amount of HLA-DR1 anchored by peptides on the surface of these ten clones were calculated based on the flow data (using Equation 3-1), and $cMFI_{(DR)}$ of yeast co-displaying FLU/HLA-DR1 was used for normalization (Figure 4-7). Seven of these ten clones display peptides with improved HLA-DR1-anchoring ability than FLU peptide, and only one clone shows no detectable HLA-DR1 on its surface.

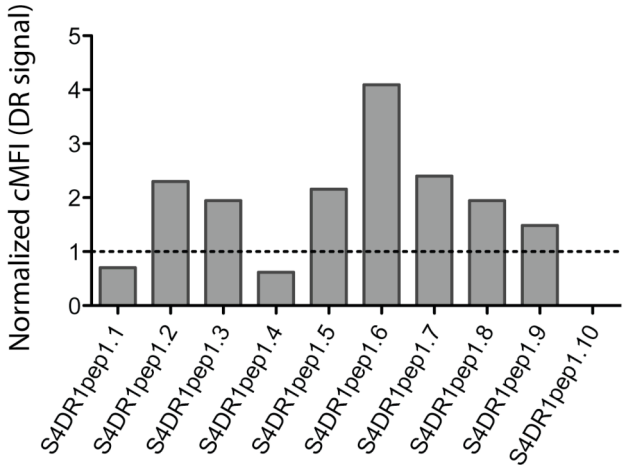

Figure 4-7 Relative HLA-DR1 amount on the surface of selected clones from yeast library co-expressing randomized peptide sequences and HLA-DR1. By using flow data, *cMFI* for DR coupled fluorophores was generated using Equation 3-1 for all selected clones (designated as S4DR1pep1.n, where n refers to clone number) and normalized by that of yeast co-displaying FLU/HLA-DR1 (dashed line).

Plasmid rescuing and DNA sequencing enables further analysis of these peptides. Plasmids were recovered from the seven positive yeast clones by yeast mini-prep, which normally contains two kinds of plasmids: one for displaying peptide (ptFLU based plasmid), the other for secreting soluble HLA-DR1 (ptsDR1 based plasmid). To screen the plasmid prep and select out the one for displaying peptide, the prep was transformed into *E. coli* strain DH5α by electroporation and a portion of the transformants was spread on LB agar plates with ampicillin added. Colony PCR allows discrimination of these two kinds of plasmids because each *E. coli* cell can only accommodate one plasmid. It is worthy of noticing that DH5α containing peptide-expressing plasmid tends to grow faster than the one containing HLA-DR1-expressing plasmid, so bigger colonies on the plate would always be the right ones. Plasmids encoding peptide sequences were then isolated from the right *E. coli* colonies by normal mini-prep and sequenced to acquire sequences of sort-out peptides (Fig 4-8). Six of these seven clones gave distinct sequences while one sequence showed up twice, which suggested that the original ten clones represented a good sample of the whole sort-out strains.

```
S4DR1pep1.2   IEGRYPYDVPDYALQASGGGGSGGGGSGGGGS-----ASMLRAHLRGRLCAQWFGGGPGKPIPNPLLGLDST
S4DR1pep1.3   IEGRYPYDVPDYALQASGGGGSGGGGSGGGGS-----ASLKRLGPAGVWHLPPQGGGPGKPIPNPLLGLDST
S4DR1pep1.5   IEGRYPYDVPDYALQASGGGGSGGGGSGGGGSGGGGSASYSYLQYLLGPKLYYRGGGPGKPIPNPLLGLDST
S4DR1pep1.6   IEGRYPYDVPDYALQASGGGGSGGGGSGGGGS-----ASLPLNRLKYYAWLLRSGGGPGKPIPNPLLGLDST
S4DR1pep1.7   IEGRYPYDVPDYALQASGGGGSGGGGSGGGGS-----ASMLHRLVGLLSVECKSGGGPGKPIPNPLLGLDST
S4DR1pep1.8   IEGRYPYDVPDYALQASGGGGSGGGGSGGGGS-----ASMVKWMLMGRLAARKLGGGPGKPIPNPLLGLDST
S4DR1pep1.9   IEGRYPYDVPDYALQASGGGGSGGGGSGGGGS-----ASMLRAHLRGRLCAQWFGGGPGKPIPNPLLGLDST
```

Figure 4-8 Sequences of peptide fusions displayed by sorted DR-positive clones. The sequence starts from Factor Xa cleaving site (IEGR) followed by HA-tag (pink), $(SGGGG)_3$ linker, randomized peptide (underlined), and V5-tag (orange).

To verify that higher HLA-DR1 amount on yeast surface is due to peptide binding but not other positive effect introduced by cell sorting and to compare relative binding of

HLA-DR1 to these sorted peptide and FLU peptide, plasmids encoding the six distinct peptides were re-transformed back to yeast strain containing HLA-DR1-expressing plasmid and tested for the phenotype by flow cytometry again. As a control, these plasmids were also transformed into parent strain EBY100 for checking false positives obtained by sorting yeast co-displaying library. Strains without HLA-DR1-secretion will give a value close to zero for genuine HLA-DR1 binders because there is no HLA-DR1 produced and anchored by these yeast at all, while those with HLA-DR1 expression should generate higher DR-ratio due to their apparently stronger HLA-DR1-binding ability than FLU as suggested from previous results (Figure 4-7). Five out of six re-transformed clones show true phenotypes in agreement with expectation (Figure 4-9). On the contrary, one clone shows no difference between strains with and without HLA-DR1 expressing plasmid, indicating that the DR-signal detected by flow cytometry is actually a false positive probably due to the interaction between displayed peptide and DR-labeling antibodies or fluorophores.

The whole trial experiment suggests that, a combinatorial yeast library of randomized peptides with 1–2 × 10^6 distinct clones will yield 10–20 distinct peptide sequences with the capability to bind HLA-DR1, excluding 10% false positives.

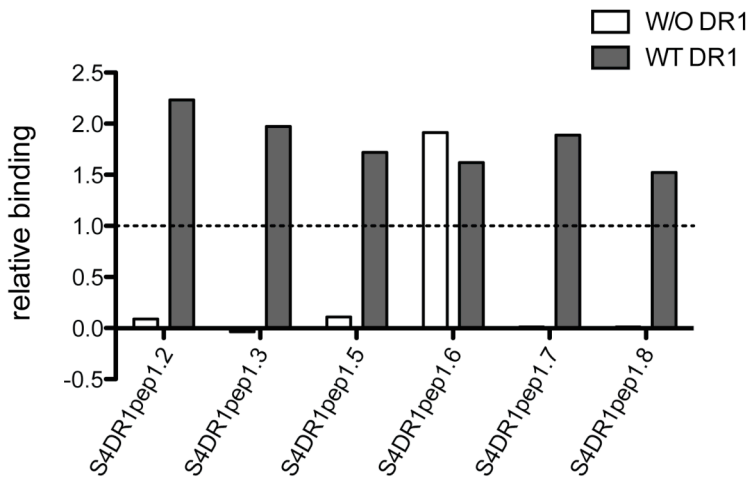

Figure 4-9 Relative binding of HLA-DR1 to peptides sorted out of yeast library. Plasmids encoding randomized peptides were recovered from selected positive clones and re-transformed into parent strain with (filled bar) or without (open bar) HLA-DR1-expressing plasmid. Double labeling and flow detection of these transformants (data not shown) were performed for generation of *DR-ratio* values. *DR-ratio* for yeast co-displaying HLA-DR1 and FLU (dashed line) was used for normalization.

4.4. Discussion

In last two chapters, we have developed yeast co-display for characterizing and engineering peptide binding specificity of MHC-II molecules and optimized its working condition for quantitatively determining the interaction between peptides and MHC-II. In this chapter, one application of this quantitative, high throughput method has been explored for characterization of peptide binding specificity.

4.4.1. Yeast co-display can quantitatively determine pocket profiles

The ability of defining anchor preference or pocket profile by yeast co-display has been proved by substituting Tyr at P1 anchor of FLU peptide with all natural amino acids and calculating relative binding level of P1 variants to HLA-DR1 displayed on corresponding yeast surface (Figure 4-5). Specifically, expression of P1 variants on the surface of yeast has been demonstrated to be irrelevant to the reside substitution within

FLU peptide and also independent of possible binding by HLA-DR1 as well (Figure 4-2 and 4-3). The equivalent amount of HA-tag and V5-tag detected by flow cytometry confirms the relatively constant display level of peptide variants and suggests a standard for quantitative analysis (refer to Figure 2-8, Figure 4-2 and 4-3). Relative binding of HLA-DR1 to all 19 P1 variants (except for Cys) generated using flow cytometric data of corresponding co-displaying yeast shows perfectly agreement with the P1 pocket profile obtained by soluble HLA-DR1/peptide binding assay[145]. Our quantitative analysis also shows an preferable binding order for favorable anchor residues: Phe ≥ Tyr ≥ Trp > Met ≥ Leu ≥ Ile ≥ Val, confirming that anchor preference of P1 pocket is restricted, where aromatic side chains are most favorable to be accommodated followed by other larger hydrophobic side chains.

Due to the relatively independent and additive nature of different pockets of HLA-DR proteins contributing for peptide binding[145], anchor preferences of other pockets of HLA-DR1 can be similarly defined using yeast co-display simply by saturating other anchor residues with all natural amino acids. To extend this application, pocket profiles of other DR alleles can be determined likewise by yeast co-display, which would largely broaden the predictability of software such as TEPITOPE for evaluation and identification of DR-specific ligands.

4.4.2. Potential of yeast co-display for determining binding motif of DR-specific ligands in a high throughput manner

Other than performing anchor preference analysis as discussed above, another feature of yeast co-display for generation of high throughput informative results has

been tested by constructing a library of peptides in an easily manipulative scale. This can also be used as an alternative method for determining pocket profile or generally identifying binding motif of MHC-II proteins. From this initial attempt, we can estimate that a combinatorial library of randomized peptides with 10^6 distinct yeast clones will yield 10–100 HLA-DR1 specific peptide sequences with 10–20% experimental errors introduced by false positives. Theoretically, information of 10^3–10^4 DR1-specific ligands can be generated if a yeast library (with a size up to 10^9 as stated in Chapter 1) is properly designed and sorted. Comparison of these peptide sequences will help to discover most correlation between peptide binding groove of HLA-DR1 molecules and specific peptides.

Although only a few peptides were sorted out from the small yeast library, the alignment of their sequence (Figure 4-8) did exhibit some common features which are consistent with the binding motif determined elsewhere. For example, first, the N-terminal residue of all five genuine DR1 binders has a large hydrophobic side chain, which has the potential to serve as the putative P1 anchor for their binding to HLA-DR1. The whole sequence of each peptide is relatively hydrophobic due to the existence of more hydrophobic residues, which agrees with the truth that most DR1 pockets favor for accommodating hydrophobic residues[59] though FLU has a Asn at P4 anchor. However, the false positive also show all these features, indicating that more intensive characterization is needed to clarify the exact anchors for these peptides, so that binding motif based on these anchor preference information could be statistically determined if more genuine binders were sorted and characterized.

The work in this chapter has shown evidences of the quantitative and high through put features of yeast co-display on characterizing peptide-MHC-II binding. It is suggested that discovery of the database for promiscuous HLA-specific ligands would become possible within decades.

Chapter 5. Engineering HLA-DR1 by directed evolution for altered peptide binding specificity

5.1. Introduction

Intensive study of anchor preferences or binding motifs of certain MHC-II alleles (e.g. results in Chapter 4) discovers their peptide-binding specificity, but the molecular basis and evolutionary motivation eliciting these kinds of properties remains unclear. Therefore approaches such as crystallography, NMR, and molecular modeling, etc., have been widely applied attempting to learn the rules and limitations governing the architecture of peptide/MHC-II complexes by isolating them from all the other biological constraints imposed on the naturally occurring proteins[55-59, 159-165]. Introducing mutations at interested sites of these molecules and analyzing the structure-function relationship with the help of X-ray crystal structures or stimulated models are of great value for understanding the significance of specific residues to peptide binding and presentation[69, 119, 165-167]. However, information in protein crystal structures is not sufficient to guide decisions for mutagenesis studies, so a high throughput screening of MHC-II allelic variants for altered peptide binding specificity is required to help narrowing down criteria important for peptide association. A few attempts performed in other work suggested a potential of using classical yeast surface display to screen libraries of single-chain MHC-II mutants, but none of them successfully expressed wild type proteins or suggested useful information about mutation sites in regard of peptide

binding[133, 168]. Therefore, a robust high throughput *in vitro* screening method is of great demand.

In addition, screening DR-variants for altered or desired peptide binding properties could find application in immunotherapy and vaccine design. For example, applying soluble MHC-II or its multimeric products to associate epitope peptides and present them to CD4+ T cells[169-174] has been demonstrated to be an efficient strategy in therapy of autoimmune diseases[46-48, 71, 175, 176] and in initiating antitumor immune responses[38], but the restricted set of peptides that can be presented by available solubly produced MHC-II allelic proteins hamper widely clinical application of peptide/MHC-II complexes. The possibility of genetically engineering peptide-binding site of some MHC-II protein so that they can specifically bind and present designed target peptide could help expand recently limited MHC-II presenting capability.

Alignment of known natural selected MHC-II proteins shows that most polymorphic residues locate at or have their side chains involved in forming pockets and stabilizing peptide binding or contacting T cell receptors[15] (e.g. Figure 5-8), suggesting an evolvability of MHC-II proteins for altered peptide specificity. Protein Engineering via experimental evolution and selection imposes a possible path to select better fitness of MHC-II functionalities. Yeast co-display actually represents such an high throughput engineering and screening method.

Since P1 anchor-pocket region is known to be the determinant for peptide binding[59, 90, 145], the initial trial will be focused on this region. According to the pocket profile generated in Chapter 4, some P1 variants are not bound by wild type HLA-DR1. These variants are proper representatives that emerge outside the restricted group of HLA-

DR1-specific peptides. To possibly bind these targets, specific modifications or mutations of HLA-DR1 need to be introduced (or in nature, we would say evolved). Therefore, a combinatorial library of HLA-DR1 mutants will be constructed using yeast co-display system and screened for positive alones regaining P1 variant binding ability. Based on the structural and polymorphic information of HLA-DR1 in complex with FLU peptide, mutation sites will be intensively evaluated within the P1 pocket of the peptide-binding site, where residues are mainly involved in peptide association.

In this chapter, we will first show a laboratory strategy for *in vitro* engineering HLA-DR1 protein for altered peptide association properties by construction and screening two generations of co-displaying yeast libraries, and then discuss the possible evolutionary significance of polymorphic peptide binding regions where essential substitutions could take place in the lab or in the nature.

5.2. Materials and Methods

5.2.1. Construction of HLA-DR1 mutants library in yeast

The GAL1-10 promoter flanked by fragments of *DRA*0101* and *DRB1*010101* genes encoding the α1 and β1 domains, respectively, was amplified from plasmid ptsDR1 via 20 cycles of PCR reactions using two synthetic oligonucleotides: W90 (5'–CGCCAGACCGTCTCCTTCTTTGCCATATCCACATGSNNSNNCTCATCACCATC AAAGTCSNNCATAAACTCGCCTGATTGGT, PAGE purified; IDT) carrying three NNS degenerate codons enabling saturated mutations at positions α 24, α 31, and α 32 on α1 domain, and W93 (5TGGGTCTTTGAAGGATACACAGTCACCTTAGGCTCA

ACTCGCCGCTGCACSNNSNNGCTCTCSNNAACCCCGTAGTTGTGTCTGC, PAGE purified; IDT) carrying three NNS codons at positions β 86, β 89, and β 90. The mutagenized PCR products were agarose gel purified and extended twice each by PCR with another two pairs of primers. Primers for the first extension were W89 (5'–TTTGTCCACAGCTATGTTGGCCAATGCACCTTGAGCCTCAAAGCTGGCAAAT CGTCCAAATTCTTCAAGCCGCCAGACCGTCTCCTTCTT, PAGE purified; IDT) and an ultramer W92 (5'–TCCAATCTCCATTCTGGATCAGGCCTGTGGACACCAC CCCAGCCTTCTCTTCCTGGCCGTTCCGGAACCACCTGACTTCAATGCTGCCTG GATAGAAACCACTCACAGAGCAGACCAGGAGGTTGTGGTGCTGCAGGGGCT GGGTCTTTGAAGGATACAC, PAGE purified; IDT). The second pair, W88 (5'–GTTGGAGCGCTTTGTCATGATTTCCAGGTTGGCTTTGTCCACAGCTATGTTGG) and W91 (5'–CTGTTTCCAGCATCACCAGGGTCTGGAAGGTCCAATCTCCATTC TGGATC) added 50 bps of sequence homologous to the linearized yeast shuttle vector prepared by MscI/StuI enzymatic digest of ptsDR1. Both the PCR generated insert and the open backbone vector were prepared at a large scale to ensure enough DNA for highly efficient homologous recombination and concentrated by Microcon YM-10 cartridges (Millipore) before the final purification by agarose gel electrophoresis. Purified DNAs were precipitated using Pellet Paint Co-Precipitant kit (Novagen) following the manufacturer's instructions after estimating the concentration by absorption at 260 nm. Approximately 24–30 μg insert and 6 μg vector were mixed in 6 μl sterile nuclease-free water and transformed into EBY100 bearing a plasmid directing surface display of P1-mutated FLU via six separate electroporation reactions. Transformants were pooled into 150 ml SD-SCAA liquid medium and aliquots were

plated at different dilutions on SD-SCAA agar plates for determination of library size. Yeast cultures were enriched at 30 °C and passaged twice by transferring at least 10^9 cells into 250 ml fresh SD-SCAA medium to dilute background cells and maintain the HLA-DR1 mutant library diversity. A separate reaction of electroporation by transforming only 1μg vector but no insert into competent cells served as the background control.

5.2.2. FACS Sorting for Library of HLA-DR1 mutants

Culturing, labeling and screening of HLA-DR1 mutants' library were similar to methods described in Chapter 4 section 4.2.4, except: 1) The primary/secondary labeling reagent pair: Biotinylated L243 (BD Biosciences) or Biotinylated LB3.1 / Alexa Fluor 647 – streptavidin (Invitrogen), instead of L243 / Alexa Fluor 647 – goat anti-mouse antibody, were used for labeling HLA-DR1 when performing cell sorting for Val- and P1-Ala-binding HLA-DR1 mutant libraries in some rounds, which provided no impact on phenotype detection and positive recovery at all; 2) In the first round, approximately 10^8 cells were examined using enrich mode with an event rate of 5,000–10,000 s^{-1} to recover as many positive cells as possible.

5.2.3. DNA shuffling and backcrossed library screening

Yeast cells recovered after the fourth round of sorting were enriched in SD-SCAA and applied directly for yeast miniprep using Zymoprep Yeast Plasmid Miniprep Kit. Plasmids isolated from the culture were used as template for PCR amplification of mutagenic gene inserts encoding cassettes ((DRα(1–61)//GAL1-10//DR1β(1–146)) by primers W88 and W91. At the same time wild type gene inserts were cloned from

ptsDR1 by similar PCR reactions using primer pairs: W88/W91, W120 (5'–GGAGGTACATTGGTGATCG) / W121 (5'–CAGGTGTAAACCTCTCCAC) or just 0117 (5'–GTTACATCTACACTGTTGTTAT), products of which had a little different lengths but all cover the whole mutated cassette sequence. After concentration and gel purification, 0.5 mg mutant and 2.5 mg wild type DNA were mixed in 100 μl of Buffer A (1 mM $MgCl_2$, 50 mM Tris-HCl pH7.4) in an Eppendorf tube and degraded by quickly adding 1.5 U DNase I (10 U/μl, Thermo Scientific, Pierce, Rockford, IL) and gently shaking the tube for 14 – 18 min at room temperature, after which the DNase I was inactivated immediately by incubating the tube at 90 °C for at least 10 min. The degraded DNA fragments were mostly shorter than 250 bps as confirmed by 2% Agarose (Invitrogen) gel electrophoresis. 0.05 U/μl PfuUltra high fidelity DNA polymerase (2.5 U/μl, Stratagene) in a 25 μl primerless PCR reaction [programmed as (96 °C, 3 min) for 1 cycle, (94 °C, 1 min, 55 °C, 1 min, 72 °C, 1 min 40 s + 5 s/ cycle gradient) for 40 cycles, (72 °C, 10 min) for 1 cycle, 4 °C to infinity] allowed 5 ng/μl gel purified DNA fragments to reassemble into new gene cassettes with the original length but different combination of mutations. Another gene piece encoding cassette GAL1-10 amplified from the wild type ptsDR1 was also added in the primerless PCR mixture to decrease the mutation carry-over within the promoter region. 4 μl raw primerless PCR products were diluted 50 fold in a modified PCR mixture containing primer pairs W88/W91 or W120/W121, 1× ThermolPol Reaction buffer (NEB) and two kinds of polymerases: 0.05 U/μl PfuUltra and 0.05–0.1 U/μl Taq (5 U/μl, NEB) for amplification of the right shuffled gene cassettes by 20 cycles of reactions. The shuffled DNA was

confirmed by 1% gel electrophoresis and purified using the Wizard Cleanup Kit (Promega).

4–5 μg shuffled insert DNA's along with 0.8–0.9 μg vector (MscI/StuI double cut ptsDR1, same as the one used in section 5.2.1) were transformed into 100 μl freshly made competent cells (EBY100 containing a P1-variant-expressing plasmid) via 2 electroporation reactions. Transformants in both cuvettes were transferred into 50 ml SD-SCAA to construct the backcrossed library of HLA-DR1 mutants. Culturing and labeling of this library was the same as described in Chapter 4 section 4.2.3 and 4.2.4. Backcrossed libraries were only sorted for one round at normal mode on FACSVantage.

5.2.4. Positive clone isolation and characterization

The isolation of positive clones and recovery of corresponding plasmids were carried out in a similar manner as described in Chapter 4 section 4.2.5, with clone numbers varied. *E. coli* transformation was used again as mentioned previously but followed by Colony PCR with [DRα(1–61)//GAL1-10//DR1β(1–146)] expression cassette-specific primers W7 (5'–AAGGATACACAGTCACCTTAGGCTCA) and W10 (5'–AACTCGGCCTGGATGATCACAT) for discriminating HLA-DR1 mutant-expressing plasmids from P1-variant-expressing plasmids in the yeast miniprep mixture. Recovered plasmids were retransformed into EBY100 by electroporation for false positive evaluation and resulting transformants were subsequently transformed with ptFLUX for P1 anchor preference analysis as described in Chapter 4 section 4.2.2.

5.3. Results

Different from the peptide library constructed in Chapter 4, HLA-DR1 mutant library results in non-directly tethered protein variants outside yeast cells, where possible self-neighbor interchange raises certain issues. It has already been demonstrated that HLA-DR1 molecules secreted and anchored by FLU peptide on the surface of yeast are rather self-specific than intercellular switchable, indicating that in the library each yeast will mainly present HLA-DR1 mutants secreted by its own via P1-variant-anchorage. This genotype-phenotype linkage will greatly favor for yeast library screening and DR positive clone recovery by flow cytometry or FACS sorting.

5.3.1. Construction and screening of HLA-DR1 mutant library

In order to edit HLA-DR1 molecules for altered peptide-binding specificity, nonspecific peptides will be set up as targets and a combinatorial library of HLA-DR1 mutants need to be screened against these target peptides for sorting of target-binding mutants. From results described in Chapter 4, yeast co-display system has already discovered some P1 variants of FLU peptide showing lower or no binding ability to wild type HLA-DR1 proteins (Figure 4-5), which can be used as proper target peptides for screening HLA-DR1 mutant-containing yeast library. Herein, initial trial was taken placed by selecting three representative peptides: 1) P1-Val (PK**V**VKQNTLKLAT) – weak binder with hydrophobic side chain at P1 anchor; 2) P1-Ala (PK**A**VKQNTLKLAT) – non-binder with hydrophobic side chain at P1; and 3) P1-Glu (PK**E**VKQNTLKLAT) – non-binder with hydrophilic side chain at P1. It is obvious that screening difficulty increases one after another for these three variants, because P1

pocket of wild type HLA-DR1 protein favors for larger hydrophobic side chain as shown in Figure 4-5.

Different purposes of protein engineering by directed evolution require different criteria for designing and screening library of protein mutants. In this trial, we are testing the possibility of altering minimum residues of HLA-DR1 molecules so that only the peptide binding specificity will be modified but not other properties or function. Accordingly, criteria for construction HLA-DR1 mutants-secreting library include: 1) mutations should be focused within regions related to peptide-binding site, e.g., pockets, which specially serve as functional group for peptide recognition and association; 2) mutations should not greatly affect structural stability of HLA-DR1 protein, e.g., disulfide bonds, non-covalent interactions; 3) mutations should not inhibit peptide presenting ability of HLA-DR1 mutants to T cell receptor, so that the HLA-DR1 mutant remains a potential for immunotherapeutic application; 4) mutations should not influence other functional regions of HLA-DR1 molecules, e.g., CD4 reactive region, DM binding region. Based on the crystal structure of HLA-DR1 in complex with FLU peptide[59], residues chosen to be saturated to all natural amino acids are Phe α24, Ile α31, Phe α32, Gly β86, Phe β89, Thr β90, all located around P1 pocket of the binding site containing at least one atom within 5 Å of at least on atom of P1 anchor residue (Tyr308) of FLU peptide (Figure 5-1). All these residues were closed to the floor of peptide binding groove, where is far from the TCR accessible peptide-presenting surface as well as HLA-DR1 heterodimer-associating interface.

Three pairs of nested primers were used to amplify mutagenic PCR product encoding expression cassette [DRα(1–61)//GAL1-10//DR1β(1–146)] bearing the six designed

mutation sites via 60 reaction cycles totally. More PCR cycles actually provide undersigned random mutations within the insert gene cassette, which might help stabilize folding of HLA-DR1 mutants interfered by the six modified residues and provide more structure-function linkage related to peptide binding.

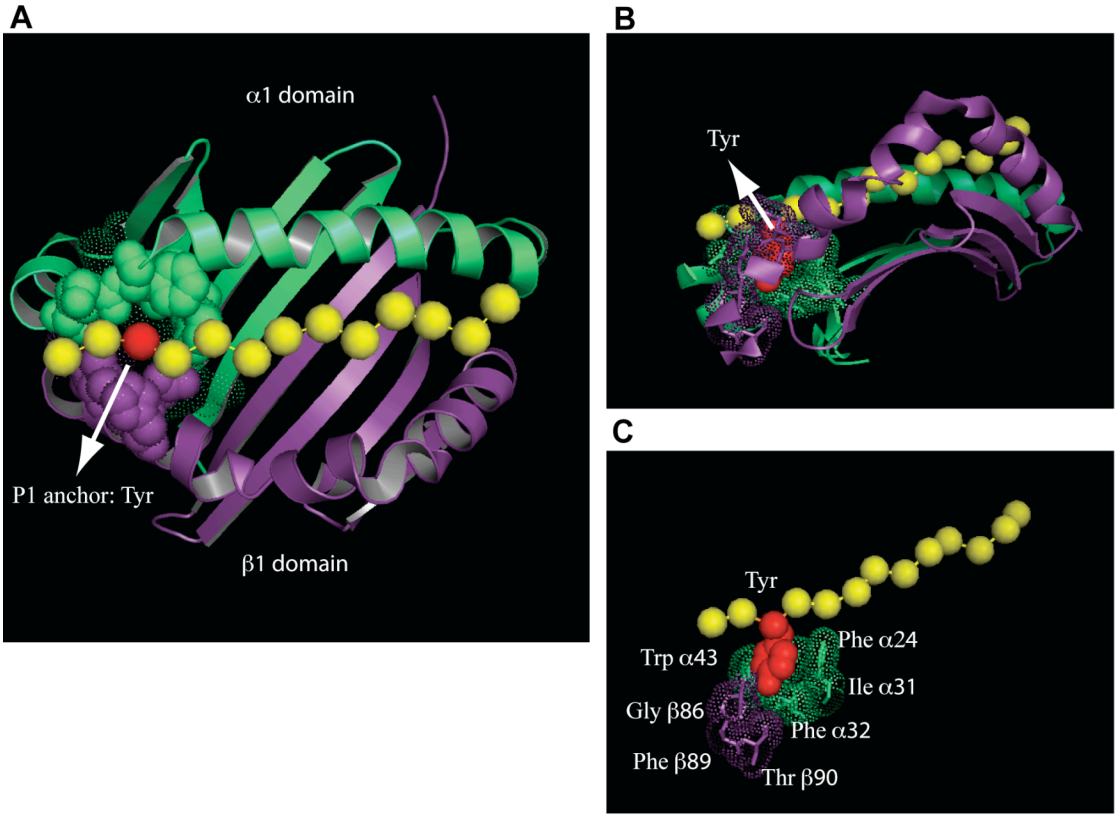

Figure 5-1 P1 pocket in peptide binding site of HLA-DR1. A. Top view of P1 pocket formed by HLA-DR1 residues surrounding P1 anchor (red sphere) of FLU peptide. **B.** side view of how P1 pocket holds the side chain of P1 anchor residue, Tyr308. **C.** Residues in P1 pocket have at least one atom within 5 Å of at least on atom of Tyr.

Mutated gene cassette mixture and cut vector were concentrated and co-transformed into yeast strain expressing P1-Val, P1-Ala, or P1-Glu peptides to construct three HLA-DR1 mutant-secreting yeast libraries for sorting P1-variant-binding HLA-DR1 mutants: P1-Val-bing mutants, P1-Ala-bing mutants, or P1-Glu-bing mutants. Homologous

recombination of insert and vector DNA took place inside yeast cell to form a HLA-DR1 mutant-expressing plasmid, millions of which were then selectively enriched after several passages in minimal growth medium. Plating of diluted samples from each library suggested an estimation of their sizes: 3×10^7 (P1-Val), 6×10^7 (P1-Ala), and 2×10^7 (P1-Glu). Size of background determined by transformants with vector only is three orders of magnitude less. Consequently, each library constructed includes almost all combinations of 20 natural amino acids in the six selected positions, which is supposed to be 20^6.

The three combinatorial yeast libraries were cultured and double labeled for HA and DR signal using the same procedure as that for wild type co-displaying strain at a bigger scale, and sorted by FACS on a FACSVantage cell sorter for four rounds. Sorting gates were all set for HA and DR double positive population with variations in each round. Enrichment mode and relaxed stringency were set up for the first round to collect as many positive clones as possible. Normal mode and a little more stringent sorting were used for the rest rounds. Ratios of sorted cells to total cells in each round (Table 5-1) suggest that roughly 10^3, 10^2, and $10–10^2$ distinct positive clones were sorted out of P1-Val, P1-Ala, and P1-Glu -displaying yeast libraries, respectively. It is also worthwhile noticing that the sorting of P1-Val-displaying library is relatively easier than the other two as suggested by the percentage in the third round of sorting, wherein significant numbers of positive clones had already been screened out for P1-Val-displaying.

Table 5-1 Ratios of sorted cells to total cells after each round of sorting.

Sorting rounds	Libraries		
	P1-Val-displaying	P1-Ala-displaying	P1-Glu-displaying
1	2.2%	5.1%	0.5%
2	0.6%	0.5%	0.3%
3	9.1%	0.3%	1.1%
4	21.6%	22.2%	ND

ND, not determined.

5.3.2. Characterization of positive clones containing target HLA-DR1 mutants

After the fourth round of sorting, samples of sorted clones from each of the three libraries were plated on agar plates. Individual colonies were randomly selected and inoculated for induction and further characterization. Number of colonies picked for P1-Val, P1-Ala, and P1-Glu -displaying libraries are ten, ten, and eighteen respectively. Flow analysis of these yeast clones (data not shown) indicated most of them carrying the phenotype as sorted, except for some from P1-Glu-displaying library. Additionally, the detectable DR-positive signal indicates that mutations within these sorted HLA-DR1 mutants have no impact on the binding and coupling of DR-specific mAb L243. $cMFI_{(DR)}$ representing the amount of HLA-DR1 mutants anchored by P1 variants on the surface of these clones were calculated (using Equation 3-1), and normalized by wild type (Figure 5-2). Most HLA-DR1 mutants show improved corresponding P1-variant-binding ability than wild type FLU peptide, and it is easier to find out mutants with higher binding affinity to weak binder (P1-Val, Figure 5-2A) than non-binder (P1-Ala or P1-Glu, Figure 5-2B or C). In addition, not all clones derived from the P1-Glu-diplaying

library exhibit detectable HLA-DR1 molecules (e.g., clone S4E1.10 in Figure 5-2C) or display HLA-DR1 with improved binding affinity (e.g. clone S4E1.1, S4E1.16 and S4E1.17 in Figure 5-2C), again indicating that the hydrophilic side chain is not as favorable as hydrophobic ones in P1 pocket even with designed residue substitutions.

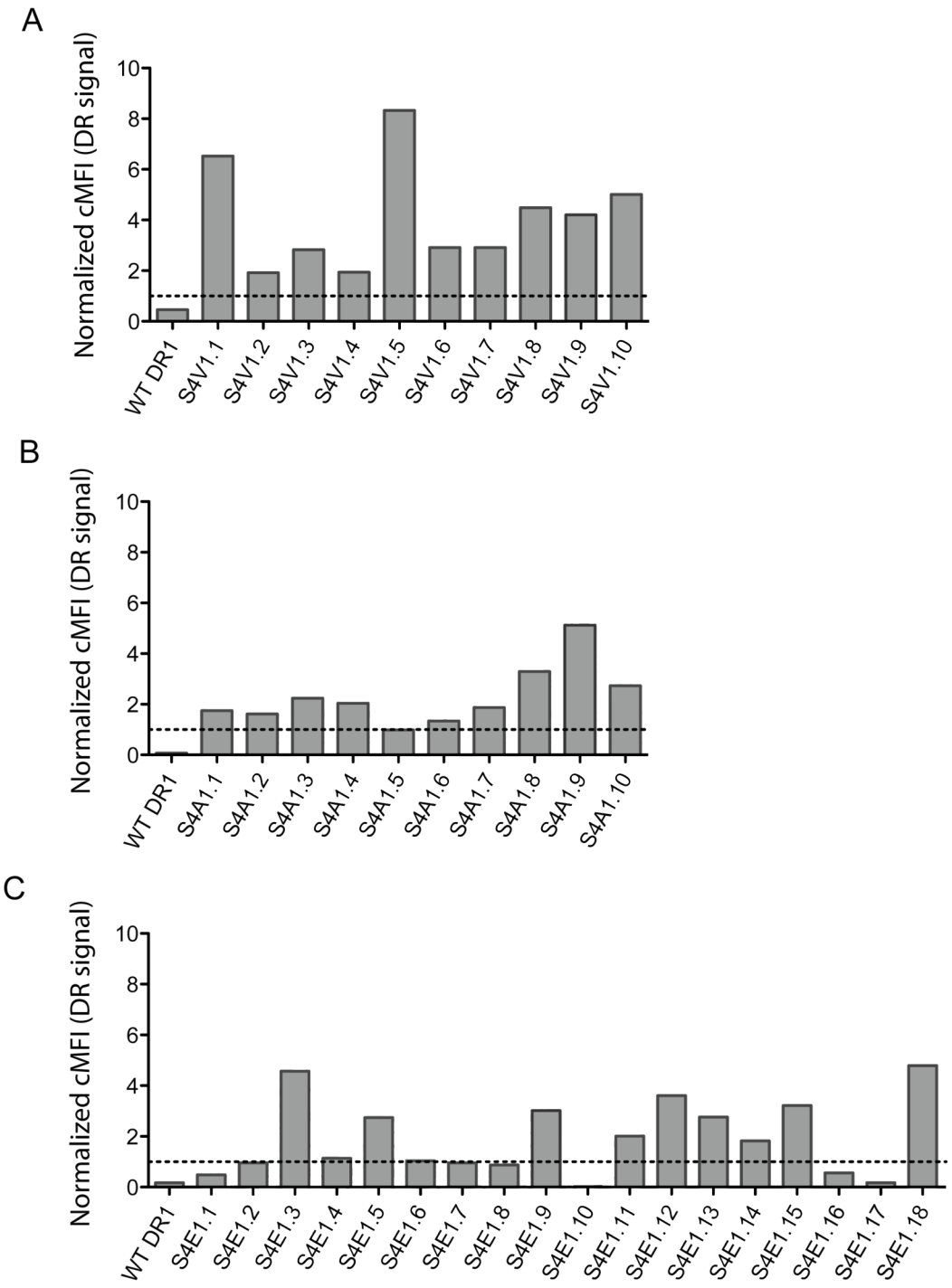

Figure 5-2 Relative HLA-DR1 amount on the surface of selected positive clones co-displaying HLA-DR1 mutants and target P1 variant: A. P1-Val, **B.** P1-Ala, or **C.** P1-Glu. The wild-type-normalized cMFI value representing relative HLA-DR1 amount was generated using flow data for selected clones (designated as S4X1.n, where X refers to **A.** V, **B.** A, or **C.** E, and n refers to clone numbers) from corresponding HLA-DR1 mutant libraries. The dashed line indicated cMFI value of yeast co-displaying wild type HLA-DR1 and FLU peptide.

In order to find out more sequence-structure-function related information, plasmids were rescued from selected positive yeast clones by yeast mini-prep. As mentioned in Chapter 4, the prep is a mixture of two kinds of plasmids, which can be distinguished by screening of *E. coli* transformants on LB agar plates. This time, the plate was incubated at 37 °C a little longer allowing the smaller colonies to grow big enough for colony PCR and inoculation. *E. coli* mini-prep of grown-up culture was performed to isolate the ones containing HLA-DR1 mutant-expressing plasmid, which can be used as template for DNA sequencing. However, the yield of these plasmids was sometimes not high enough for sequencing; alternatively, gene cassettes encoding mutated regions of HLA-DR1 mutants were amplified from yeast mini-prep or *E. coli* mini-prep by PCR, and then cleaned up by electrophoresis gel extraction and purification before DNA sequencing.

Sequences of selected HLA-DR1 mutant clones obtained by DNA sequencing (Table 5-2) suggest that other than the six selected mutation sites, random mutations do exist within the PCR amplified gene cassette. The observation that all P1-Val and P1-Ala-binding mutants are different from each other confirms that the number of distinct clones sorted out of these two libraries is larger than the sample number. In contrast, the same clone was sequenced multiple times for P1-Glu-binding mutants, indicating a relatively small number of distinct clones sorted out of this library. This is consistent with the estimation made by sorted cell percentage as described previously, confirming the difficulty for sorting P1-Glu-binding mutants. Another interesting result is that most P1-Val-binding and P1-Ala-binding mutants bearing an Aα61S mutation located at the restriction enzyme recognition site used for linearizing vector ptsDR1. This mutation may be result from homologous recombination, but provide no positive effect on peptide

binding (actually, this assumption was confirmed by sequence analysis after backcrossing shown in table 5-3). Except for this mutation, consensus mutations or single mutation bearing mutants are seldom observed by alignment of these sequences, but besides the six heavily mutated sites, α54, β11, β30, β37, β57, β81 show certain importance on peptide binding. Backcrossing of these multiple-mutation sequences with wild type HLA-DR1 can be used to pinpoint the most essential mutations that contribute to the altered peptide binding capability and to derive more conclusive results.

Mutations in selected HLA-DR1 mutants.

		α chian											β chain																																
		16	22	24	31	32	45	50	51	54	60	61	11	17	18	25	26	30	32	37	52	53	55	57	71	81	86	89	90	91	92	100	105	108	109	112	113	114
WT HLA-DR1		P	F	F	I	F	L	R	F	F	L	A	L	F	F	R	L	C	Y	S	E	L	R	D	R	H	G	F	T	V	Q	T	K	P	L	H	N	L
P1-Val-binding	s4V1.1			Y	L							S								F							T	A	L										
	s4V1.2											S															I	R	M										
	s4V1.3				A	L				L																	L	V	F		K								
	s4V1.4			C	L							S															L		S										
	s4V1.5			L	L											Q	F										F	N	P										
	s4V1.6											S															M	L	A										
	s4V1.7											S															L	M	S	E						X			
	s4V1.8			L	S								F													Q	L	G	I										
	s4V1.9																										F	L	A					X					
	s4V1.10			I	T	L														T							F	N	L										
P1-Ala-binding	s4A1.1				S	L																					L	P	H										
	s4A1.2												H														F	R	H										
	s4A1.3	H			M	L														F							F	S	K										
	s4A1.4																										F	S	L										
	s4A1.5	S			L							S	I		L											Y	M	W									K		
	s4A1.6											S			S												F	T	L										
	s4A1.7																	Y									F	H	P		R	S	N						
	s4A1.8																										F	A	M										
	s4A1.9																	Y				M					F	L	L										V
	s4A1.10				L	I				L			R														L	L	L										
P1-Glu-binding	s4E1.1																				K		Q		M		L	Y	P										
	s4E1.2																										L	H	H										
	s4E1.3		V			S					M													Y			L	I	Y										
	s4E1.6						H																				I												
	s4E1.8																										I	L	L										
	s4E1.11				L	L			Y				F						N								L	L	A										
	s4E1.12							G						L										Y			L	I	M						Q				
	s4E1.15		V			S					M													Y			L	I	Y										
	s4E1.18		V			S					M													Y			L	I	Y										

Shaded: designed mutation sites. X: deletion.

Table 5-2

Three representative HLA-DR1 mutants, S4V1.5, S4A1.9 and S4E1.3, with the highest relevant P1-variant-binding ability were chosen from these three characterized mutant pools for P1 anchor preference analysis. EBY100 was first transformed with the plasmid secreting one of these three HLA-DR1 mutants, and then transformed with one of 19 P1-variant-displaying plasmids ptFLUX to generate corresponding co-expressing strain. All resulting strains were cultured and double labeled with H6908 and L243 for flow analysis as described before. Relative binding of each mutant to any P1 variant (Figure 5-3) determined using Equation 3-2 followed by normalization shows significantly higher overall peptide binding ability of all HLA-DR1 mutants than wild type, which is supposed to be the result of sorting for better binding signal. Other than that, P1 anchor preferences of these three HLA-DR1 mutants (Figure 5-3B~D) are all different from that of wild type (Figure 5-3A or Figure 4-5), indicating a possibility of engineering peptide binding specificity of MHC-II protein using yeast co-display by directed evolution. Among these mutants, S4V1.5 and S4A1.9 still prefer to accommodate hydrophobic side chains in their P1 pockets; however, residues with smaller hydrophobic side chains tend to be more favorable because sorting was set for against peptide variants with smaller hydrophobic side chain (Val or Ala) at P1 anchor. S4E1.3 on the other hand, shows a hyper bidning promiscuity, suggesting the P1 pocket is no longer dominant for determining peptide-binding specificity.

The ability of screening co-expressing yeast library by FACS for acquiring HLA-DR1 mutants with improved peptide binding ability and altered specificity validates the fact that the interaction between peptide and MHC-II molecules does occur inside yeast

cells and the switching of MHC-II proteins among different yeast cell surface is a rare event or flow non-detectable in our experimental circumstance as evaluated in Chapter 2.

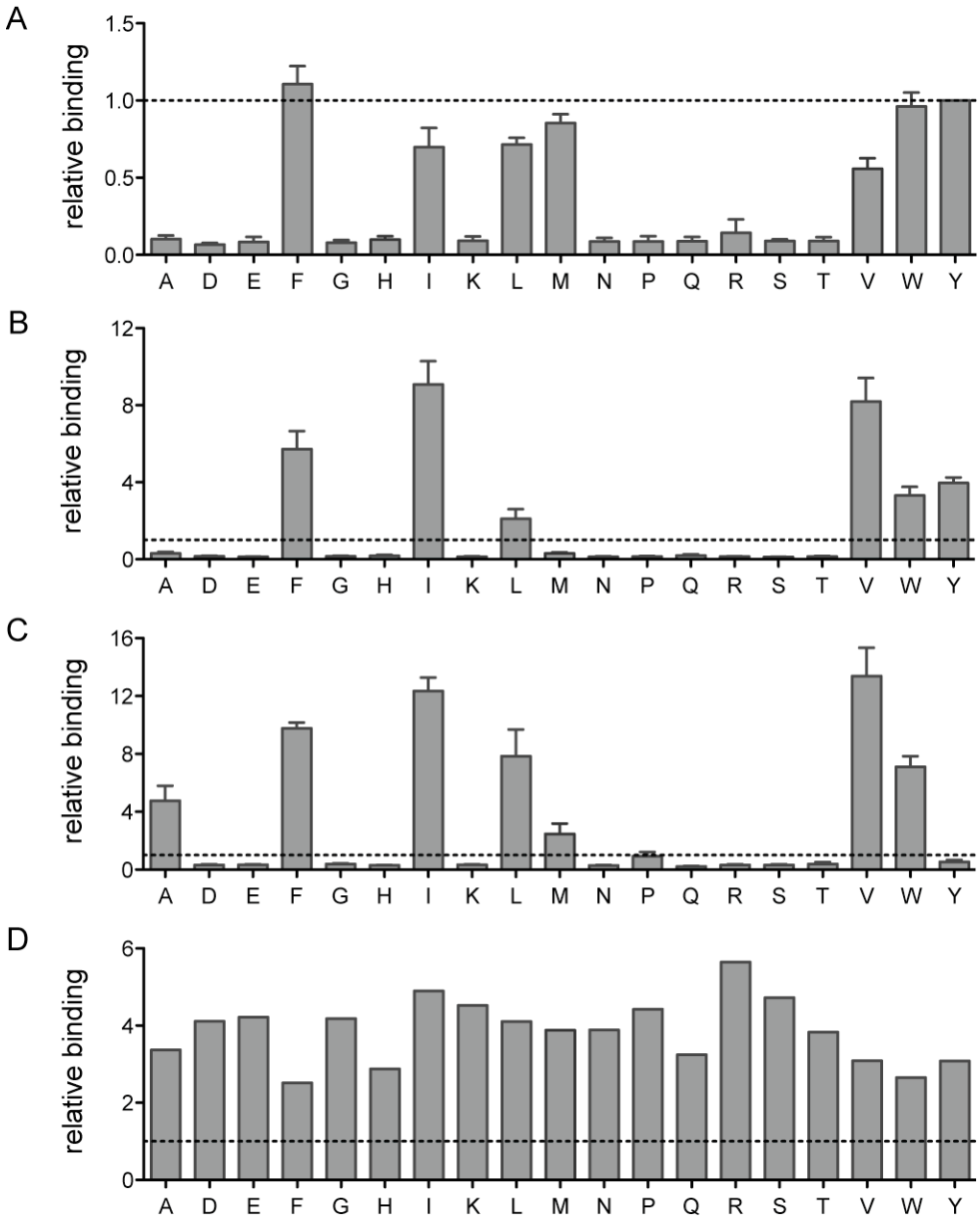

Figure 5-3 P1 anchor preference for HLA-DR1 and its variant-binding mutants. *DR-ratio* was calculated using Equation 3-2 based on flow cytometric data of yeast co-expressing HLA-DR1 mutant and one of the indicated P1-variants. Relative binding level of **A.** wild type HLA-DR1, or **B.** P1-Val-binding mutant S4V1.5, or **C.** P1-Ala-binding mutant S4A1.9 **D.** P1-Glu-binding mutant S4E1.3 to indicated P1-variant were generated by normalizing *DR-ratio* by the value for yeast co-displaying FLU with Tyr at P1 and HLA-DR1 (dashed line). Error bars represent standard error of the mean (SEM) determined from four independent experiments.

5.3.3. Backcrossing and new library construction and screening

Although P1-variant-binding mutants with changed peptide binding properties were obtained from previously constructed yeast libraries, mutations in these mutants are too diverse, which brings too much difficulty to structure-function analysis. Shuffling mutagenized gene cassettes with extra amount of wild type provides an effective approach to separate these mutations and enables construction and screening of yeast library containing single mutation bearing mutants[177, 178].

Three plasmid mixtures were recovered from three sorted mutant pools, respectively, by yeast mini-prep and used as templates for amplifying mutated gene portions [DRα(1–61)//GAL1-10//DR1β(1–146)], which carrying all mutations to be shuffled. After gel purification, each of the three PCR products was mixed with five folds more concentrated wild type DNA and was chopped up into small fragments by DNase I. Primerless PCR drives these small fragments to reassemble into longer pieces DNA, which were then selectively amplified using specific primers to recover shuffled gene cassettes. Enough shuffled insert and cut vector (same as the one used to construct previous three HLA-DR1 mutant libraries) were co-transformed into corresponding P1-variant-displaying yeast strain, respectively, to construct backcrossed HLA-DR1 library by homologous recombination. Combinations of all mutations after shuffling is not too many due to the small amount of essential mutations found in previous libraries, so the size of new libraries was designed to be a little smaller: 8×10^6 (P1-Val), 7×10^6 (P1-Ala), and 1.5×10^7 (P1-Glu). The three backcrossed libraries were sorted against the three P1-variants individually by FACS. Different from previous ones, only one round of sorting yield significant amount of double positive clones for all three libraries.

5.3.4. Characterization of positive clones of backcrossed libraries

Spreading sorted clones on agar plates enabled isolation and characterization of individual colonies. More than twenty clones were randomly picked from each of these three libraries for flow cytometric characterization. Similar results were observed by comparing the normalized $cMFI_{(DR)}$ histograms generated before (Figure 5-2) and after DNA shuffling (Figure 5-4), suggesting a correlation between these two sets of mutants. Because of fewer rounds of sorting were carried out, more negative clones without phenotype of interest (normalized $cMFI_{(DR)} \approx 0$) were included.

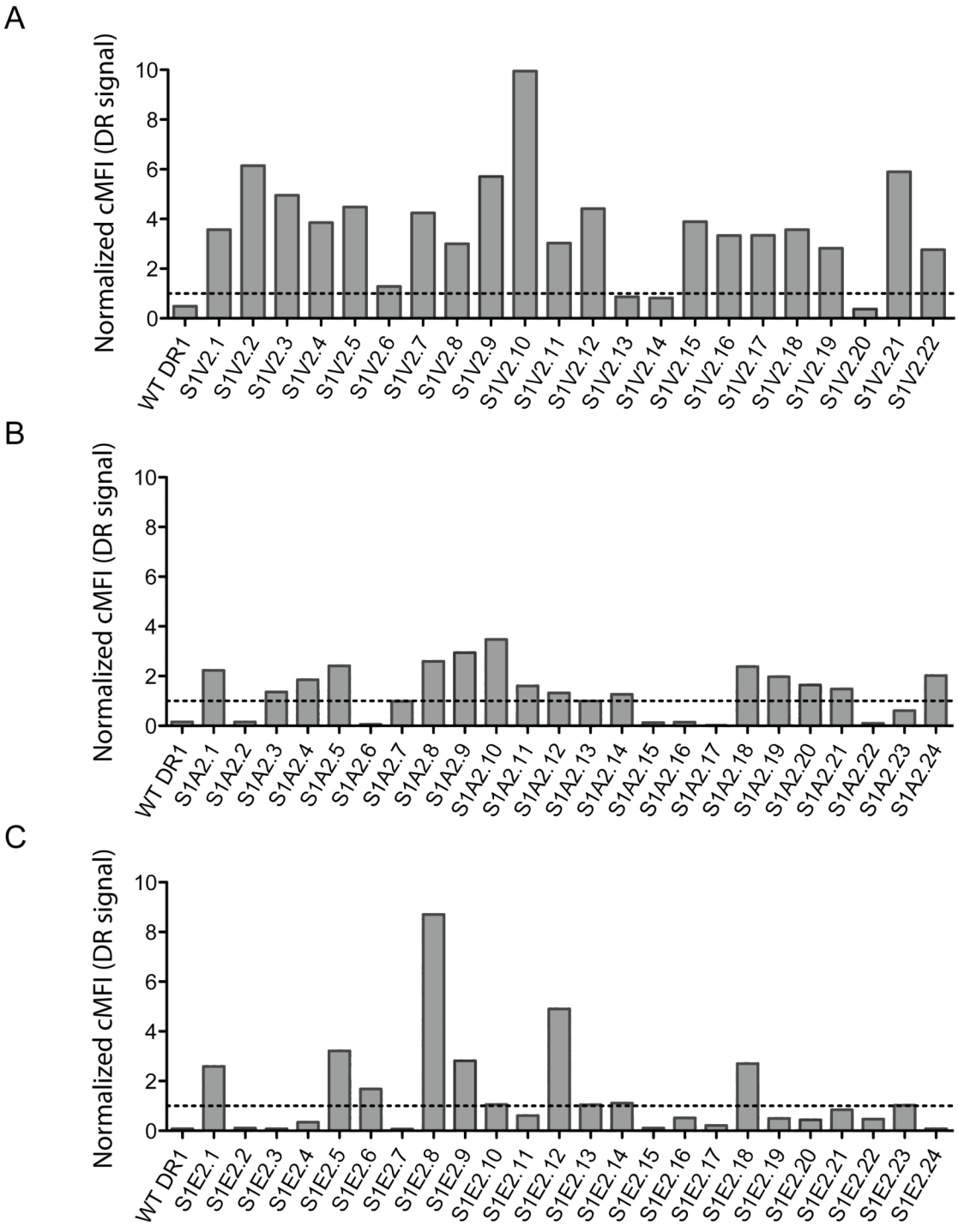

Figure 5-4 Relative HLA-DR1 amount on the surface of clones sorted out of wild type-backcrossed libraries for co-displaying of HLA-DR1 mutants and target P1 variant: **A.** P1-Val, **B.** P1-Ala, or **C.** P1-Glu. The wild-type-normalized *cMFI* value representing relative HLA-DR1 amount was generated using flow data for selected clones (designated as S1X2.n, where X refers to **A.** V, **B.** A or **C.** E, and n refers to clone numbers) from corresponding backcrossed libraries. The dashed line indicated cMFI value of yeast co-displaying wild type HLA-DR1 and FLU peptide.

Similarly, to obtain sequence information, plasmids were rescued from selected positive yeast clones by yeast mini-prep, and either used for *E. coli* transformation and plasmid isolation or used as template for amplifying PCR product for DNA sequencing. Comparing with previous one (Table 5-2), alignment of sequences of more HLA-DR1 mutants from backcrossed libraries (Table 5-3) discovers less mutations or single mutation bearing clones with improved P1-variant-binding ability, which helps to determine more essential mutations, such as α22, β11, β26, β30, β37, β57, β86. Of particular interest is the phenomenon that much less mutations in α chain, including the three designed mutation sites, were carried over to backcrossed mutants than β chain from previous mutant pools, indicating neutral or negative contributions of most mutations in α chain, which agrees with the fact that α chain is more conserved under natural evolution whereas most polymorphic regions are located at β chain. Stop codon observed at the end of β1 domain of some mutant (mutant S1V2.10) implies that β2 domain may be unnecessary for peptide binding. Same sequences can be found in all three mutant pools, especially for those bound by P1-Ala and P1-Glu, suggesting that proper sample number was chosen to acquire enough information from each pool.

Table 5-3 Mutations in HLA-DR1 mutants after being backcrossed with wild type.

		α chain								β chain																	
		6	10	22	31	32	44	60	61	11	26	30	37	48	57	63	86	88	89	90	91	92	112	117	120	126
	WT	V	A	F	I	F	R	L	A	L	L	C	S	R	D	S	G	S	F	T	V	Q	H	C	S	S
P1-Val-binding	S1V2.1												F													
	S1V2.3																F									
	S1V2.4											Y													G	
	S1V2.5	A															F		L	F						
	S1V2.6									I																
	S1V2.8									F									L	I						
	S1V2.9																L		L	Q						
	S1V2.10						G							Q			L		L	V		*				
	S1V2.11																M		Y	L						N
	S1V2.12							M	S			Y														
	S1V2.15																L		M	L				X		
	S1V2.16																F		L	Q						
	S1V2.17																W		L	W						
	S1V2.18																L		M	S	E					
	S1V2.19																L		L	L						
	S1V2.22																L		L	L						
P1-Ala-binding	S1A2.1									R																
	S1A2.3									H																
	S1A2.4																F		M	Q						
	S1A2.5									R																
	S1A2.7									H																
	S1A2.8																F		L	L						
	S1A2.9																F		M	Q						
	S1A2.10																F		A	M						
	S1A2.11									V																
	S1A2.12									V																
	S1A2.13																F		L	Q						
	S1A2.14		T														F		T	N						
	S1A2.18																F		A	M						
	S1A2.19																F		I	A						
	S1A2.20																F									
	S1A2.21																F									
	S1A2.23																L		L	H						
	S1A2.24																F		A	M						
P1-Glubinding	S1E2.1			V		S									Y		T			I						
	S1E2.5																I		Q	I			Y			
	S1E2.6														Y											
	S1E2.8														Y		I		Q	M						
	S1E2.9							M			F															
	S1E2.10				L	L									Y	G										
	S1E2.11																I		M	L						
	S1E2.12				L			M							Y											
	S1E2.13														Y											
	S1E2.14														Y											
	S1E2.16																I		Q	M						
	S1E2.18										F						I		Q	M						
	S1E2.19																L		L							
	S1E2.20																L		I	M						
	S1E2.21			V													I		Q	M						
	S1E2.22																F	N	V	A						
	S1E2.23														Y											

Shaded: designed mutation sites. *: stop codon. X: deletion.

Like previous characterization, seven mutants: S1V2.1, S1V2.3, S1A2.1, S1A2.10, S1E2.6, S1E2.16 and S1E1.18, most of which show multiple times in table 5-3, were chosen for P1 anchor preference analysis (Figure 5-5). Only S1V2.3 and S1A2.10 exhibit different P1 anchor preferences from wild type HLA-DR1, and the patterns of their relative binding histograms look very similar, indicating some correlation between these two clones. S1V2.3 is a single mutation-bearing mutant, the mutation Gα86F, which is also include in S1A2.10, is likely to be responsible for the changed P1 anchor preference. Therefore, other mutations within S1A2.10 might just provide some effect on stabilizing protein folding or just neutral mutations. In contrast with the two with altered peptide binding specificity, the other five selected mutants just show an increasing overall binding ability to all variants without changing the preference, which suggest that DNA shuffling and backcrossing against wild type may get rid of some positive contributions from rare mutations which cell sorting can not tell and recover. These results also tell us that fewer mutations outside a specific pocket region have the potential to alter anchor preference of that pocket.

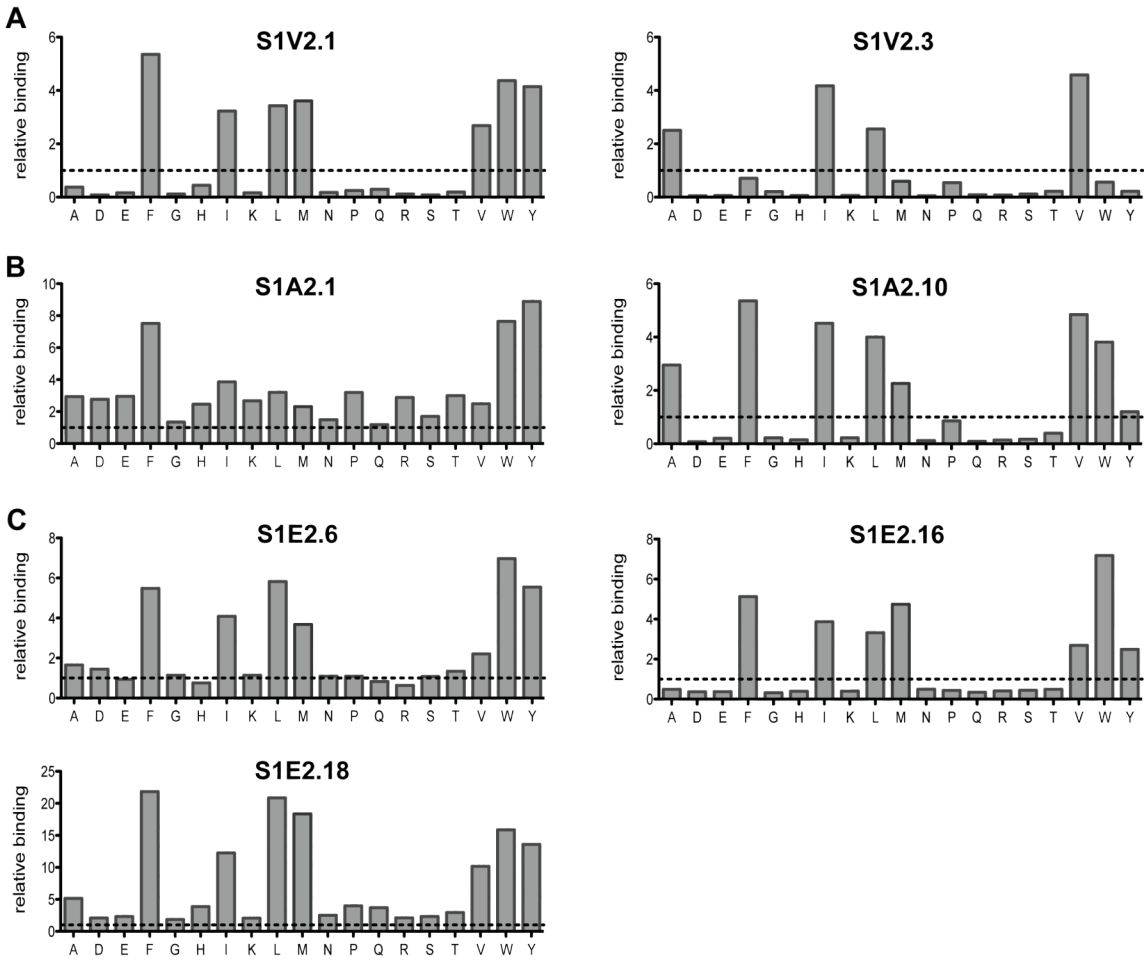

Figure 5-5 P1 anchor preference for selected HLA-DR1 mutants from backcrossed libraries: A. P1-Val-binding **B.** P1-Ala-binding and **C.** P1-Glu-binding. Relative binding level of different mutants to FLU analogues with indicated residues at P1 anchor were determined as decribed previously. The dashed line again represents the value for yeast co-displaying wild type FLU peptide and HLA-DR1.

5.3.5. Evaluation for false positive sorted from yeast library

S4V1.5, S4A1.9 and S1V2.3, S1A2.10 obtained before and after DNA shuffling, respectively, show obvious altered P1 anchor preferences, which demonstrate the applicability of using yeast co-display to engineer peptide binding specificity of HLA-DR1. However, the observation that S4E1.3 has no binding preference to all P1 variants and five out of seven mutants sorted out of backcrossed libraries exhibit a similar increase for binding to each P1 variant, suggesting that the phenotype might due to false

positives, for instance, possible binding of mutagenized HLA-DR1 molecules to yeast surface proteins. To evaluate this possibility, parent strain EBY100 was transformed with plasmids encoding these mutants and tested for HA and DR double positive signals without including any peptide-expressing plasmid. Because S4V1.5, S4A1.9, S1V2.3 and S1A2.10 have already shown no binding to several P1 variants (Figure 5-3 B and C, Figure 5-5A and B), there is no need to test these true positives; on the other hand, they can be used as controls (Figure 5-6A and B) to evaluate other ones. Five of the other six mutants gave similar flow cytrometric results (Figure 5-6C~H), whereas only S1E2.18-expressing plasmid transformed yeast shows a small DR-positive population. It is interesting to notice that S1E2.16 and S1E2.18 share most mutations but Lβ26F only found in S1E2.18 (Table 5-3), which could be the one bringing in the false positive phenotype. However, S4V1.5 containing the same mutation Lβ26F still represents a true positive clone, suggesting that the false positive observed for S1E2.18 is possibly due to an overall false positive effect added up by non-obvious ones contributed by single mutations. Nonetheless, the percentage of false positive is only 10% or less, which will not significantly affect the accuracy of applying yeast co-display for manipulation of HLA-DR1 molecules.

Along with the checking for false positives performed in chapter 4 (Figure 4-9), this control experiment again points out that directed evolution using the high throughput yeast co-display system normally provides conclusive information with 90% accuracy, which is statistically high enough for quantitatively evaluating characteristics of interaction between MHC-II proteins and various peptides.

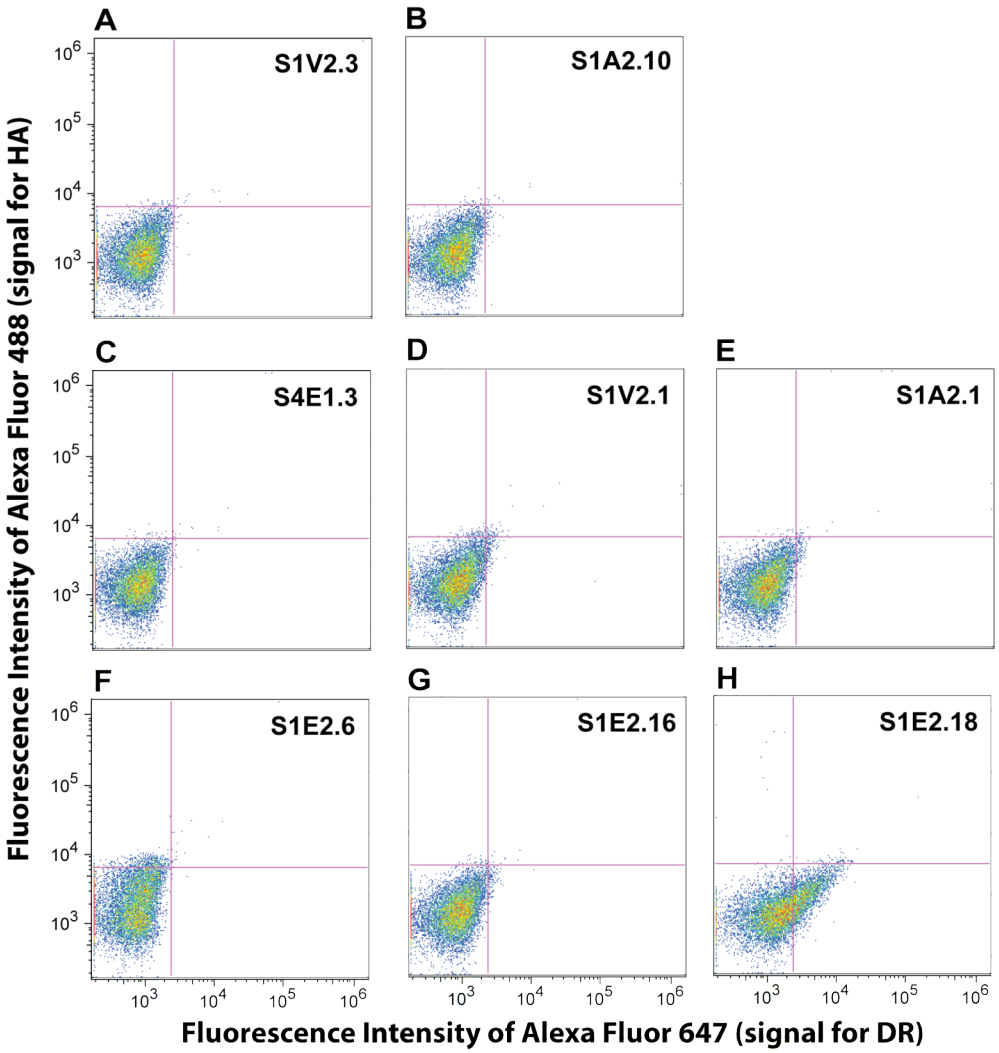

Figure 5-6 Flow cytometric analysis for examination of false positive. EBY100 transformed with one of these HLA-DR1 mutant-expressing plasmids was cultured under corresponding nutrient dropout condition and double labeled to detect both HA and DR signal. Data was collected on Accuri C6 flow cytometer.

5.4. Discussion

5.4.1. Appropriate strategy for engineering peptide binding specificity of MHC-II

Following previous chapter, we have examined another more profound application of yeast co-display on engineering peptide binding specificity of HLA-DR1, product of a well-characterized MHC-II allele, in order to explore the possibility of tailoring the antigen presenting spectrum for MHC-II molecules evolved under laboratorial situation. Directed evolution strategy such as constructing and screening a combinatorial yeast library of tens of millions of HLA-DR1 mutants has been applied throughout this chapter for selecting novel binding phenotypes in a high throughput manner.

Although HLA-DR1 mutants are secreted solubly and bound by yeast-surface-tethered P1 variants non-covalently, no intercellular switching of these soluble molecules was observed, which enables recovery of the mutagenic genotype corresponding to screened phenotype easily. This genotype-phenotype linkage has been further confirmed for clones sorted out of both the original libraries and the backcrossed libraries by recovering the HLA-DR1-expressing plasmid from positive clones and retransforming them into corresponding parent strains (for example, after transforming the plasmid encoding S4V1.5 back into parent strain displaying P1-Val peptide, the same phenotype (Figure 5-3B) was observed just as the positive clone sorted out of P1-Val-displaying library (Figure 5-2A), so it is reliable to use FACS for screening positive phenotype in co-displaying yeast libraries and obtaining corresponding genotype information.

The laboratorial evolution of HLA-DR1 discussed in this chapter includes two generations of yeast libraries. The original yeast libraries were constructed for screening HLA-DR1 mutants bearing mutation sites concentrated near P1 pockets against a DR non-specific P1 variant. Sequence analysis of selected P1-variant-associated HLA-DR1 mutants indicates that although a little more leucine's (which should be more because it can be encoded by six codons) appearing among mutated residues, the whole mutagenesis strategy did not cause too much bias and if combining different sites, mutations almost include all possible amino acids (Table 5-2). Indeed, this generation has already yield significant amount of mutants capable of associating with FLU variants used to escape the recognition of HLA-DR1 molecules and these mutants have started to exhibit altered peptide specificity (Figure 5-3), but most mutants contain multiple substitutions. If time permitted, one could test these substitutions one by one to define their effects on defining peptide binding motifs and pocket profiles by making single mutation bearing mutants. To accelerate the process for selection and determination of useful mutations, backcrossed libraries aiming for pinpointing the most essential substitutions within the mutagenized gene cassette were constructed and screened for similar phenotypes against corresponding P1-variants. Alignment of genotypic sequences suggests an obvious correlation between the two generations of libraries (Table 5-2 and 5-3), where the backcrossed ones significantly shrink the range of mutation sites to regions possibly dominating the altered phenotypes. Characterization of HLA-DR1 mutants extracted from selected clones also implies the additive contribution of each single-substitution bearing mutant (Figure 5-5) to the phenotype of a multi-substitution bearing mutant (Figure 5-3). This overview clearly shows that the

searching strategy herein is appropriate for engineering HLA-DR1 molecules and discovering important sites within peptide binding groove, which will imply more structure-function relationship with the help of crystallographic analysis.

5.4.2. Common features of essential mutations in HLA-DR1

By simply comparing the sequence alignments in Table 5-2 and Table 5-3, we found that mutations in β chain are more diverse while those in α chain are more conserved, and the number of mutations in α chain dramatically declines after backcrossing such that even the three designed mutation positions have no more mutations left, especially for P1-Val and P1-Ala-binding mutants (the case of P1-Glu-binding mutants is a little special and we will further discuss this in detail later). The more conserved feature of DRα on defining peptide anchor preference nicely agrees with the discovery in nature that there is only one single extracellular domain of HLA-DRA protein known till today comparing with more than six hundred HLA-DRB proteins (Table 1-1). This again indicates the outstanding capability of yeast co-display to reflect and mimic the natural evolution of MHC-II proteins.

In contrast to the total decreased number, mutations are significantly enriched after backcrossing in a few sites: α22, α60, β11, β26, β30, β37, β57, β86, most of which have been observed in mutants binding to different P1 variants. Other than these, α54 and β81 also imply certain importance as both of them show up for both P1-Val and P1-Ala-binding mutants in original libraries (Table 5-2). Together with α31, α32, β89, β90, which are designed to be intensive mutated, almost all these essential mutation positions suggested by yeast co-display are either located in P1 pocket or on the floor of

peptide binding groove (Figure 5-7A, α31, α32, β81, β89, β90, responsible to form P1 pocket, are not shown in this figure). Space spanning of these residues in wild type HLA-DR1 suggest a direct contact or correlation between side chains of peptide anchors and these residues, except for α60, which is further away from the peptide-binding site (Figure 5-7B). Mutation Lα60M observed in S1V2.12, S1E2.9 and S1E2.12 is actually coexisting with other essential mutations at sites mentioned earlier, so this site might not critical on its own or just a neutral mutation carried over from original libraries, so we will not consider it when further discussing the structure-function correlation.

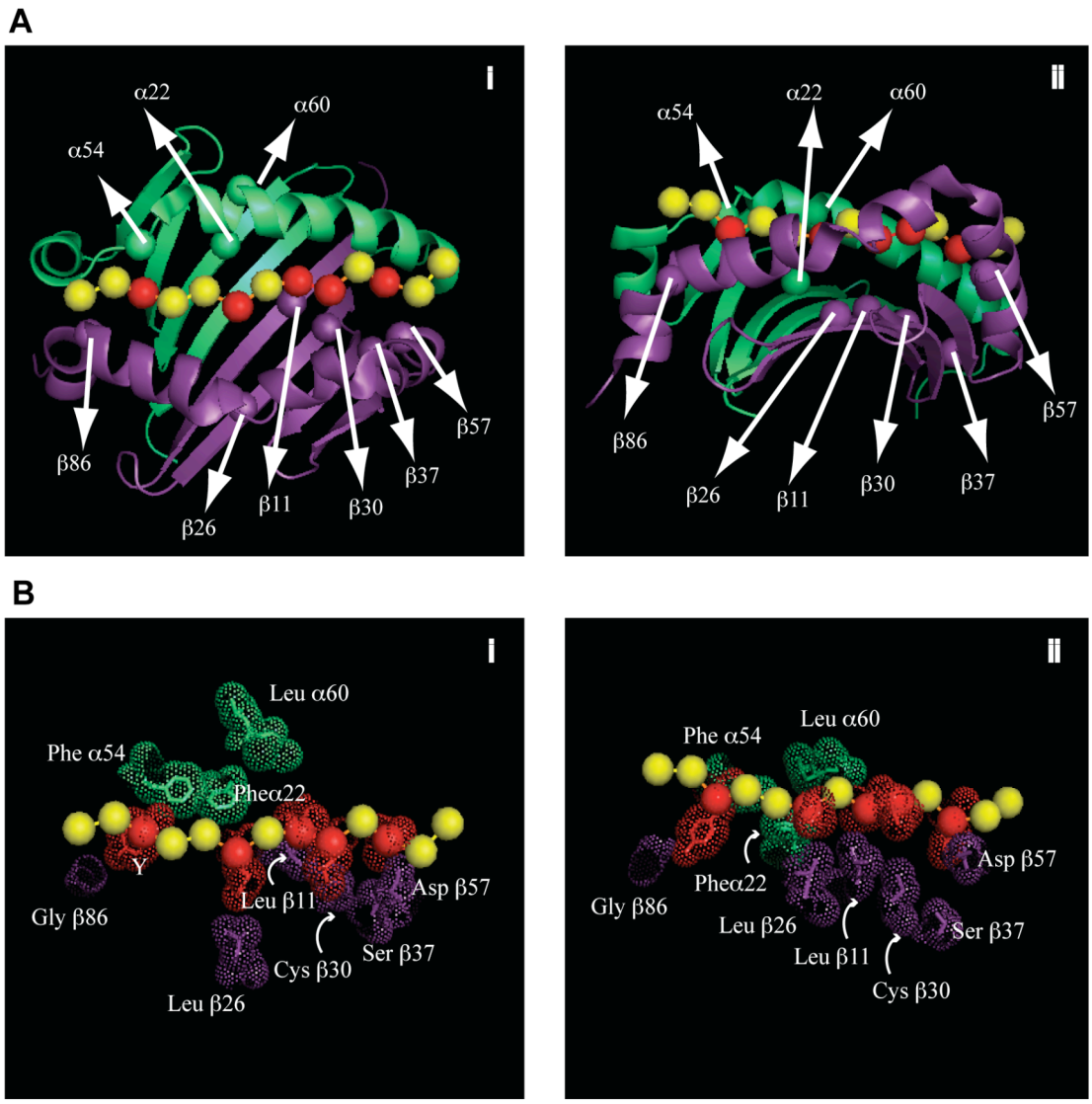

Figure 5-7 Essential mutation sites implied by directed evolution via yeast co-display for altering peptide binding properties. **A.** Location of these positions corresponding to HLA-DR1 structure. **B.** Space spanning for side chains of wild type residues at these positions. In all images, greens and purples represent α1 and β1 domains of HLA-DR1 or corresponding residues in either domain. FLU peptide is shown as stick-connected red or yellow spheres, where reds represent anchor residues with side chain dots also dipicted in **B**. In **A** and **B**, top view (i) and side view (ii) are both presented.

An alignment of dozens of representative protein sequences for extracellular domain encoded by naturally discovered DRB alleles indicates that most polymorphic sites are concentrated in β1 domain, and responsible for peptide binding or TCR binding (Figure 5-8). All essential mutation sites (β11, β26, β30, β37, β57, β86) found in β chain

suggested by yeast co-display are right among these polymorphic regions, and several substitutions (Lβ11R, Lβ26F, Cβ30Y, Sβ37F, or even Hβ81Y) even exist in some genuine alleles. Except for the relatively conserved bimorphic position β86 in P1 pocket, all other essential mutation sites allow various alterations (up to 8) in nature (red rectangles in Figure 5-8). In addition, it seems none of these sites affect the interaction between MHC-II and TCR, DM or CD4 molecules. Because mutants isolated by yeast co-display appear consistent with the natural evolution of MHC-II molecules, yeast co-display seems like a potent tool for probing the molecular basis for MHC-II evolution in defining peptide specificity.

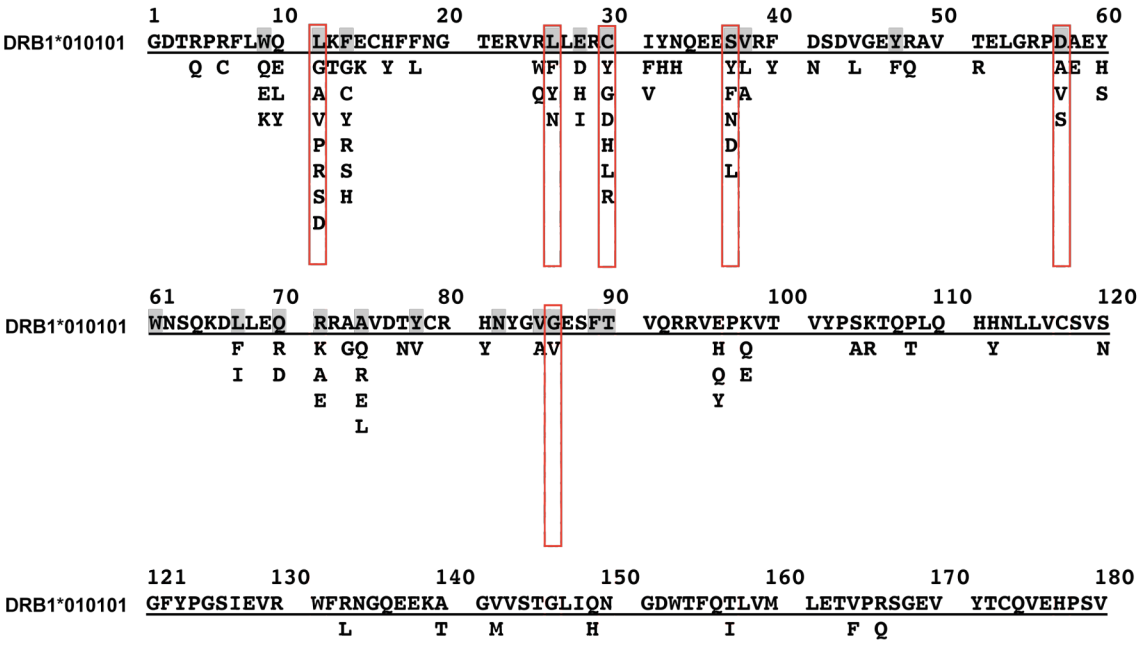

Figure 5-8 Polymorphic sites of protein encoded by DRB alleles. Alternative residues at variable positions of DRB extracellular domain (β1 and β2) are listed under corresponding position of protein sequence for DR1β chain (DRB1*010101) by alignment of selective mature protein sequences in each subgroup of DRB alleles (DRB1*010101, DRB1*030101, DRB1*040101, DRB1*0411, DRB1*070101, DRB1*080101, DRB1*090102, DRB1*100101, DRB1*110101, DRB1*120101, DRB1*130101, DRB1*140101, DRB1*150101, DRB1*160101, DRB3*01010201, DRB3*0201, DRB3*030101, DRB4*01010101, DRB5*010101, DRB5*0202). Putative peptide contacting residues are shaded. Essential mutation sites implied by directed evolution via yeast co-display are illustrated in red rectangles.

5.4.3. Residues in P1 pocket defining peptide binding specificity

The conserved feature of α chain in nature, where little or no neutral drift is suggested, implies that this chain is structurally or functionally intolerant to mutations; here, α chain mutations were isolated from the primary libraries, suggesting that mutations at these sites were not highly detrimental to folding, assembly or function of the heterodimer. However, most mutations found at α24, α31, α32, and α54 disappear after backcrossing, implying that these mutations were not important for binding to the target peptide. In contrast, the three intensively mutated sites on β chain almost carry alterations all the time, where β89, β90 permit various substitutions though a little hydrophobic prone after backcrossing, while β86 mainly allows hydrophobic residues even for accommodating P1-Glu with a negative charged side chain at P1 anchor. After backcrossing, phenylalanine was greatly enriched at β86 in mutants specific for hydrophobic P1 variants, especially for P1-Ala (Table 5-3), which agrees perfectly with the fact discovered in other work that the β86 Gly-Val dimorphism plays an substantial role in affecting peptide binding[69, 119, 152, 179-181]. This enrichment results in a single mutation-bearing mutant S1V2.3 (Gβ86F), which exhibits a significant different side chain preference for its P1 pocket (Figure 5-5A) and tends to lose binding ability to aromatic side chain and bigger side chain bearing residues including the wild type Tyr in comparison with the P1 pocket profile of wild type HLA-DR1 (Figure 4-5 or Figure 5-3A), which is also confirmed by previous characterization of a similar mutation at β86[69]. Furthermore, S4V1.5 (Figure 5-3B), S4A1.9 (Figure 5-3C) and S1A2.10 (Figure 5-5B) bearing the mutation Gβ86F are all exhibiting a similar P1 pocket profile to S1V2.3

though minor differences exist among them, probably resulting from the non-shared mutations in these clones.

As shown previously, in naturally-occurring DRB proteins, the bimorphic site β86 has either Gly or Val (Figure 5-8), whose small hydrophobic side chains allow P1 pocket to have enough space for aromatic or bigger hydrophobic side chain to fit in (Figure 5-3A, 5-7B). This also explains why Val and Ala at P1 anchor of FLU analogues are less favored, and mutants specific for them need longer side chain substitutions in order to occupy the space and provide enough attractive force to hold small anchors (e.g. Figure 5-9A). These characterizations all support the fact that β86 is one of the dominant residues within P1 pocket of HLA-DR1 for defining the pocket profile, or even suggest a possible role of β86 in defining the significantly different peptide binding specificity of another MHC-II allele, HLA-DQ, whose product has a hydrophilic β86 residue[22, 58, 122, 154, 155]. Substitutions for other residues forming P1 pocket such as Hβ81Y, Fb89L, Fbb89A, or even Fa32L, Fα54L found in conserved chain in Table 5-2, seem not necessary but could presumably donate positive contribution to maintain the hydophilicity and to stabilize P1 pocket by saving more space for the Phe-substitution at β86 (e.g. comparing Figure 5-9B to Figure 5-7B).

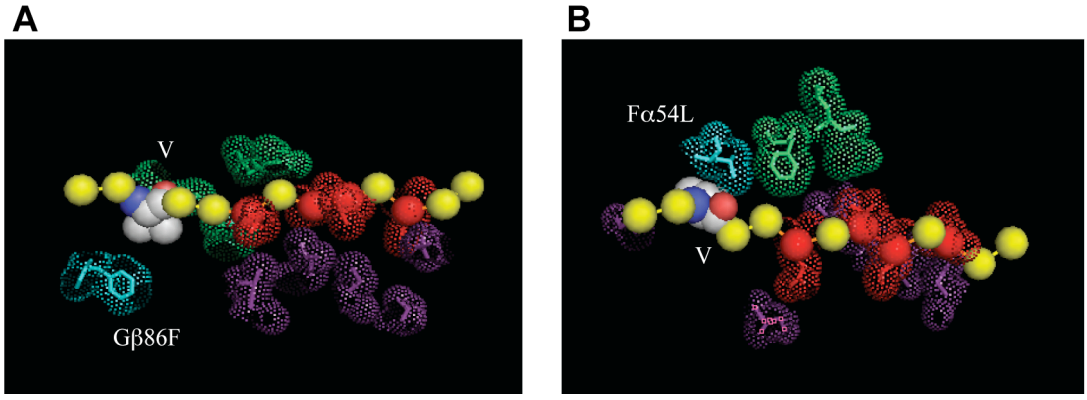

Figure 5-9 Selected examples of substitutions at essential mutation sites within P1 anchor-pocket region dominating or affecting P1 Pocket profile. A. Gβ86F/P1-Val. **B.** Fα54L/P1-Val. Same as shown previously, greens and purples represent essential residues in α1 and β1 domains of HLA-DR1 respectively, presumably responsible for peptide binding. P1 variants of FLU peptide is illustrated as spheres either in red for anchors or yellow otherwise, connected by yellowish sticks with P1 anchor residue replaced by corresponding variant (carbon/hydrogen, nitrogen and oxygen atoms are shown in white, blue and red respectively). Space spanning for side chains of all residues at essential mutation sites and all peptide anchors are indicated using dots. Cyan suggests specific mutations at different sites in HLA-DR1 mutants. Mutagenesis was performed by using the software PyMOL. Briefly, wild type amino acids was replaced by mutated residues arbitrarily first, and then one of the two conformations predicted by PyMOL was selected to simply illustrate the putative side chain space spanning.

However, due to hydrophobic nature of P1 pocket[59, 145], hydrophilic or charged side chains such as Glu are much less favored such that replacing Gly itself at β86 by Ile, Leu, Phe, or Thr would not be enough any more, alterations around P1 pocket especially in the conserved chain become inevitable. Few phenotypic positive clones sorted out of original P1-Glu-displaying library also confirm the difficulty for HLA-DR1 mutants to accommodate P1-Glu peptide. Although most mutations in mutants bound by P1-Glu are only shown once in table 5-2 and possibly neutral, Fα31L, Fα32S, Fα32L Lα45H, Rα50G, Fα51Y are all located at P1 area. Along with Gβ86T, Fβ89Q, Tβ90M, all these putative alterations in P1 pocket tend to convert the hydrophilicity of P1 pocket, which potentially favor anchors such as Glu. Therefore, DNA shuffling and backcrossing actually limit the carry-over of contributions from all these mutations and lead mutants

even harder to prefer binding to P1-Glu (Figure 5-5C). The only example in our characterized clones that suggests a possibility to make Glu a preferred P1 anchor is S4E1.3 (Fα22V, Fα32S, Lα60M, Dβ57Y, Gβ86I, Fβ89Q, Tβ90M), which unselectively bind to all P1 variants with similar binding affinity (Figure 3D). This ultra promiscuity implies an intermediate state from where if more alterations being introduced, resulting mutants could specifically prefer Glu instead of hydrophobic residues. To compromise, dominating nature of P1 pocket on peptide specificity may be dramatically disturbed and importance of other pockets possibly arise, because several changes outside P1 pockets are observed in S4E1.3 and these changes remain after backcrossing.

5.4.4. Significance of essential mutations occurred outside P1 pocket

Clones selected from the original three libraries all contain multiple mutations (Table 5-2), which makes structure-function relationship harder to interpret, but of particular interest is that random mutations outside the six designed mutation sites have been extensively observed throughout the whole mutagenized gene cassette [DRα(1–61)//GAL1-10//DR1β(1–146)], especially along the β chain even considering the more residues included in the cassette. A couple of these mutations were remaining in the cassettes as the only mutation after DNA shuffling and backcrossing against wild type sequence (Table 5-3), a process that theoretically would eradicate most nonessential substitutions. Therefore, these sites found by sorting original libraries appear to convey significant influence on peptide binding even they are not directly engaged in association with P1 anchor. Based on the screening and characterization of phenotypic positive

clones selected from two generations of yeast libraries, residues outside P1 pocket that might introduce important contribution to P1 variant associations include α22, β11, β26, β30, β37, β57 (Figure 5-7). Mutations occurring in these sites are more or less favorable for accommodation of other anchor residues on peptide regardless of alteration in P1 anchor-pocket region. For example, Lβ11R, Cβ30Y (not single mutation, but clearly important), Sβ37F and Dβ57Y found in S1A2.1, S1V2.4, S1V2.1 and S1E2.6 actually exhibit larger space occupancy (Figure 5-10A~D) than wild type (Figure 5-7B), which could provide stronger molecular interaction between the mutated residues and P6, P7, P9 anchors respectively. Other than that, even though Fα22V and Lβ26F are not observed as single mutations, their contributions for peptide binding can be judged by comparing S1E 2.21 or S1E2.18 versus S1E2.16 (Figure 5-5C). The effect of Fα22V is possibly connected to stabilization of P4 anchor Gln by replacing the aromatic side chain to smaller hydrophobic one (Figure 5-10F). The effect of Lβ26F is not obvious due to the simple mutagenesis analysis by PyMOL (Figure 5-10E), but could also be due to some correlation with P4 anchor considering its importance in some other DR proteins with Phe at β26 (e.g. DR2[182]), though a little bit false positive was observed for S1E2.18 (Figure 5-6H).

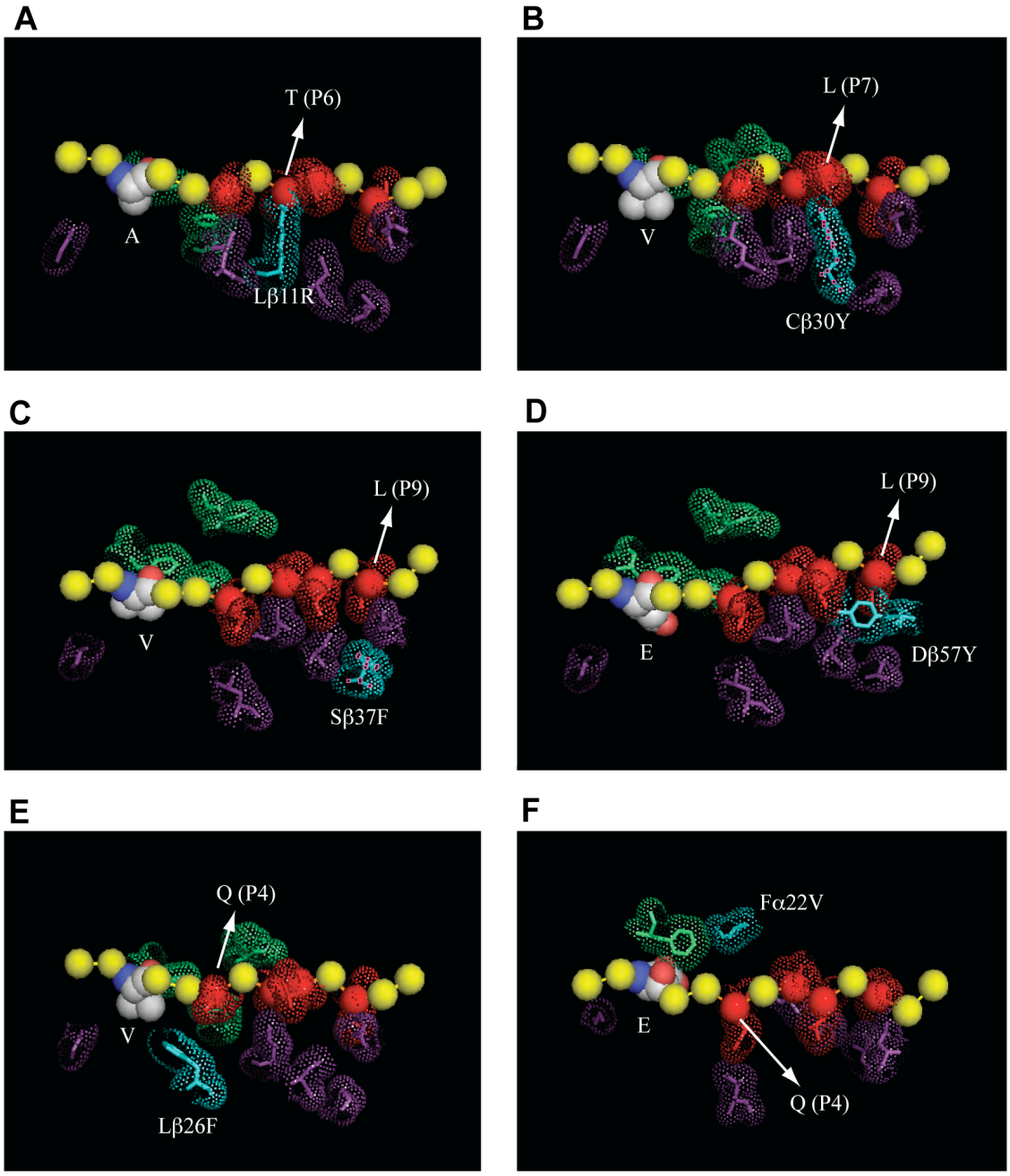

Figure 5-10 Substitution at essential mutation sites outside P1 anchor-pocket region favoring for accommodation of P1 variants. **A.** Lβ11R/P1-Ala. **B.** Cβ30Y/P1-Val. **C.** Sβ37F/P1-Val. **D.** Dβ57Y/P1-Glu. **E.** Lβ26F/P1-Val. **F.** Fα22V/Glu-variant. Color illustrations are the same as previous figure. Mutagenesis was performed by PyMOL in a similar manner to previous figure.

Relative binding assay shows HLA-DR1 mutants containing only these mutations rather obtain stronger overall binding ability than altering P1 anchor preference (Figure 5-5). Alternative explanation of this observation is that higher expression level (probably related to stability) instead of better peptide binding occurred due to these mutations, which agrees with the hypothesis proposed in other work that mutations Lβ11H, Lβ26F and Dβ57A are responsible for increasing the stability of a single-chain derivative of HLA-DR1 displayed by yeast[133]. The fact that the anchor preference governed by P1 pocket can hardly be changed by introducing substitution outside this pocket confirms the dominance of P1 pocket on determining peptide binding specificity and the relatively independent peptide-binding nature of P1 pocket region to elsewhere.

However, different from S1V2.1 (Sβ37F), the ability of mutants S1A2.1 and S1E2.6 to associate with variants that are not bound by wild type HLA-DR1 suggests that substitutions Lβ11R and Dβ57Y occurring in P6 and P9 pockets could compensate a little affinity loss caused by P1 anchor alteration, such that weak- or non-binder could associate. Although due to a lack of single mutation in P4 and P7 pockets we did not characterize whether such substitution could provide similar compensation, essential mutations around all these four anchors indeed have been identified for altered binding affinity, suggesting similar effects of all these pockets on acquiring binding ability to P1 variants that wild type does not bind. In addition, elsewhere studies suggest that site-directed mutations applied at sites β11, β13, β30, β67, β70, β71, β74, β78 forming P4, P6, and possibly P7 pockets have impact on peptide binding by other DR alleles such as DR3 or DR7[119, 166]. The importance of β26, β37 and β57 for P4, P9 pocket of DQ proteins has also been postulated for determining their specific peptide association[56, 58,

[167, 183, 184]. Therefore we postulate that the effect of any of the other four pockets is roughly equal on stabilization of peptide binding for DR proteins without influencing the binding specificity determined by P1 pocket, and the contribution for peptide binding from one of these four pockets can be obtained by modifying residues in others.

5.4.5. Evolutionary hint for MHC-II in peptide binding specificity

Comparing with P1-Val or P1-Ala, the huge hydrophilicity change in P1-Glu peptide requires more residue alterations in essential positions, which are closely related to natural evolution of MHC-II molecules. In humans, HLA-DQ, another subclass of MHC-II alleles, encode some proteins specific for binding and presenting antigenic peptides with Glu or Gln at P1 anchor, and these proteins have significantly different residues in their P1 pockets as well as other pockets[22, 58, 122, 154, 155]. For example, the structure of HLA-DQ8 in complex with an immunodominant peptide insulin B (derived from insulin B chain 9-23: SHLVEALYLVCGERG, with the N-terminal Glu as P1 anchor and C-terminal Glu as P9 anchor)[58] suggests that Gluβ86, Hisα24 and Gluα31 instead of Glyβ86, Pheα24 and Ileα31 in HLA-DR1 are responsible for forming P1 pocket for DQ8 to hold Glu (P1) in the insulin B peptide. Obviously, the hydrophilicity of P1 pocket in DQ8 is completely reversed, and P1 is demonstrated to be less important than some other pockets such as P4 and P9 in determining peptide binding specificity of DQ molecules, in agreement with our findings for mutants (e.g. S4E1.3) with relatively better binding to P1-Glu mentioned previously. It is also known that protein sequences of DQA are not conserved at all (Table 1-1), consistent with the observation in our results that more alterations are involved in α chain for mutants obtaining P1-Glu

binding ability (Table 5-2 and 5-3). Particularly, DQ8 has Leuβ26 and Alaβ57 at P4 and P9 pockets respectively responsible for accommodation of corresponding anchors Tyr (P4) and Glu (P9) in insulin B peptide. Even though mutations Fα22V, Lβ26F and Dβ57Y occurred around P4 and P9 pockets in HLA-DR1 mutants S4E1.3, S1E2.1, S1E2.6, S1E2.9, S1E2.18, S1E2.21 may not contribute to P1-Glu binding the same way as Leuβ26 and Alaβ57 towards insulin B peptide, β26 and β57 are selected out of yeast library even though we did not design any of these sites for mutagenesis. All these similarity between the mutations we found in mutants bound by P1-Glu peptide and the essential residues in DQ8 specific for Glu (P1) bearing insulin B peptide, gives us an obvious hint that there is a molecular motivation for evolving MHC-II by switching peptide binding dominancy in between different pockets to broaden peptide binding spectrum. This might be the reason why MHC-II molecules have been differentiated into three major subtypes for recognition and presentation of various antigenic peptide sequences derived from different sources.

Additionally, the result that single-mutation-bearing HLA-DR1 mutants favor for binding to P1-Glu is hardly found to exhibit a similar anchor preference to S4E1.3 indicates that single mutation might only lead to a relatively moderate alteration and the genuine change in peptide specificity can only be obtained by combining multiple such moderate alterations. This phenomenon can be observed by comparing the relatively lesser polymorphism among different proteins within a subtype of MHC-II to the more significant sequence differences between DR and DQ proteins.

5.4.6. Summary of engineering HLA-DR1 for altering peptide binding specificity

The peptide binding specificity of HLA-DR1 molecule is predominantly governed by residues forming P1 pocket, essentially the naturally bimorphic site β86, which is also responsible for controlling the hydrophilicity of this pocket with significant contributions dedicated by other residues in P1 pocket. Residues outside P1 pocket region provide relatively independent and equal importance on stabilizing peptide association.

To specifically associate with hydrophobic P1 variants, one hydrophobic substitution at β86 is presumably sufficient and those at other potentially polymorphic sites in or outside P1 pocket could strength the specificity of this alteration though not necessary. This indicates the relationship between different alleles found in DR subtype (e.g. DR1, DR2, DR3, DR7 as mentioned previously). On the other hand, to specifically bind hydrophilic P1 variants, appropriate alterations applied to β86, conserved sites in α chain within P1 pocket, and essential sites around other pockets are all required for accumulating all necessary contributions to this specificity conversion. To maintain structural stability, more residue substitutions might be needed in other non-polymorphic sites, as a result of which the dominant feature of P1 pocket could significant altered, implying a structural rationale for differentiation of different MHC-II subtypes (DR, DQ, etc.). Another hypothesis is that hydrophilicity might be one critical motivation for the gradual switching of dominant pocket during natural selection and could possibly guide the experimental evolution strategies.

Chapter 6. Conclusions of this thesis work and the future

6.1. Development of a quantitative, high throughput engineering platform for studying protein-protein interactions

In this thesis work, we developed a powerful *in vitro* methodology – yeast co-display for characterizing and engineering *in vivo* protein-protein interaction by using a well-characterized immune response related protein pair FLU/HLA-DR1. It has been clarified that the assembly of HLA-DR1 and their baiting by target peptide mainly take place inside yeast via secretory pathway and dissociation on the surface is a slow process which will not affect discovering the direct linkage between phenotype on the surface to genotype inside the same yeast. Furthermore, the surface display of HLA-DR1 is actually peptide-binding-dependent. Therefore, interest phenotype of corresponding peptide-variants/MHC-II-mutants binding on the surface of target yeast clones recovered after several rounds of FACS cell sorting in a high throughput manner can be further analyzed quantitatively via yeast co-display.

On one hand, peptide molecules can be manipulated either by saturating an interested anchor with all natural amino acids for determination of anchor preference or by randomizing the entire sequences for characterization of HLA-DR1 binding specificity. Therefore, both specific pocket profiles and general binding motif can be generated by yeast co-display. On the other hand, MHC-II molecules can be modified by directed evolution method for altering their peptide binding specificity, which can be applied for designing novel proteins for immunotherapeutic purpose or for understanding molecular

basis of natural evolved polymorphic MHC-II proteins with the help of crystallographic analysis or molecular modeling.

6.2. Generation of peptide binding spectrum of MHC-II for various applications

Importance on modification of MHC-II molecules and on identification of MHC-II ligands has been emphasized several times throughout this thesis. Herein we are showing a map of peptide binding spectrum for HLA-DR1 and its mutants generated by collecting and quantitatively analyzing several yeast co-display data, which would provide a guidance for future mapping the database of promiscuous HLA-specific ligands experimentally or computationally (Figure 6-1).

In this three-dimensional map, X-axis starts with the wild type FLU peptide followed by alternative P1 variants, and to a extend, more randomized peptides. Y-axis lists HLA-DR1 as well as its mutants, which normally contain more substitutions following the axis direction. Z-axis represents the relative binding of each peptide to each protein. Because libraries constructed by yeast co-display would yield false positive, a set of background controls for each peptide are also included at the origin of Y-axis.

Looking at this map, first, it is easy to tell the false positive such as peptide sequence S4DR1pep1.6 generated by sorting randomized peptide library against wild type HLA-DR1. Second, anchor presences for each protein can be compared for further sequence-structure-function relationship analysis. For example, S1V2.3 (Gb86F) prefers smaller hydrophobic side chains to bigger ones in comparison with wild type HLA-DR1, and S1A2.1 (Lβ26F) obtains the ability to bind most non-binders of wild type HLA-DR1.

These findings are directly related to the single substitutions within these two mutants as we discussed in Chapter 5. Third, relative binding of different DR proteins to a specific variant can be compared, which could affect the determination of using specific protein to present target peptide for possible immunotherapeutic applications. Fourth, interaction between DR proteins and randomized peptide sequences can be used to identify and compare peptide motifs of each protein. If more alterations are introduced into wild type FLU peptide, data can be added in between P1 variants and randomized peptides and more sequence-function relationship could be acquired. If more DR proteins are tested for their association with different peptides, data can also be inserted in corresponding Y-axis. The potential of constructing and combining peptide libraries with MHC-II mutant libraries using yeast co-display would accelerate the generation of peptide/MHC-II interaction data to fill this indefinite binding spectrum, which would be extremely useful for guiding us to understand peptide binding properties of MHC-II molecules and applying created MHC-II with defined peptide specificity for vaccine design or other purposes.

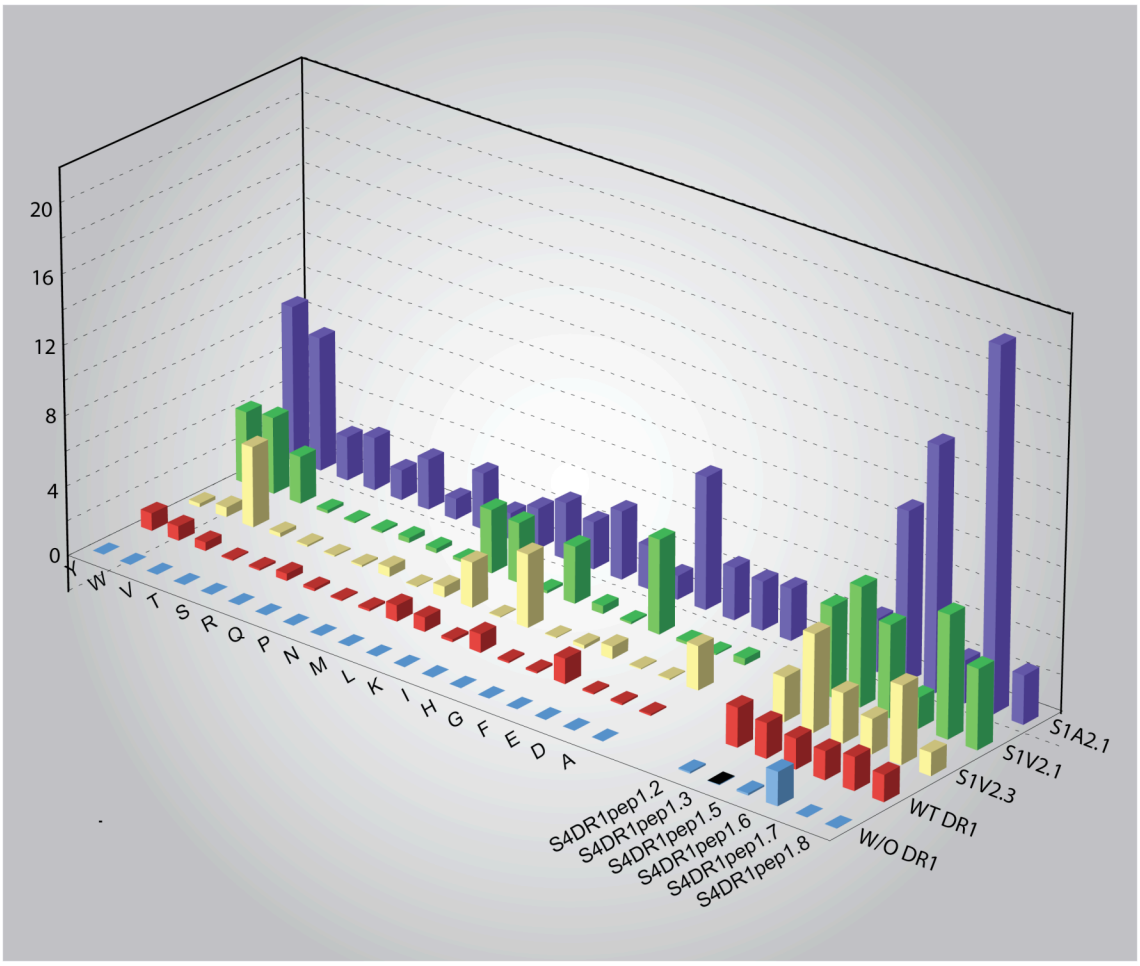

Figure 6-1 An example of peptide binding spectrum of DR proteins determined by yeast co-display.

6.3. Potential applications of yeast co-display

Interaction of FLU and HLA-DR1 molecules occurring inside yeast cells better mimics the is *in vivo* scenario compared to other binding assays, but it belongs to *in vitro* engineering methods, which decrease biological complexity and concentrate on the interested functionality of targets. Therefore, phenomena observed by yeast co-display could be subtly different from those occurring in original biological environment, where accessory molecules are extensively involved[65]. It is suggested that the presence of HLA-DM is important for acceleration of weaker peptide (e.g. Ii) dissociation[185] (Figure

1-4) and DM reportedly functions by interacting with DR molecules from the N-terminal side of accommodated peptides[128, 186], which would provide more restraints on screening HLA-DR1 mutants specific for target P1 variant. The ability to assemble HLA-DR in a secreted form therefore entering the secretory pathway and to co-express multiple plasmids simultaneously suggests a potential of this yeast system wherein DM could be co-expressed (e.g., with an ER retention signal) in a similar way and the peptide binding of DR molecules would be examined in the presence of this catalyst.

The capability to characterize and modulate protein-protein interactions in yeast using co-display will also show its promising application in various ligand-receptor interaction studies. For instance, in the immune system, other than peptide/MHC-II protein pairs, a lot of other protein pairs like antigen-antibody, peptide/MHC complex-TCR, IFA-ICAM, and CD40-CD40L are being intensive studied these days due to their significant role in CD4+ T cell development and function (Figure 1-2). It is possible to study the interaction between these protein pairs using yeast co-display system and to determine the function-structure relationship of them. However, covalent linker in between anchor protein Aga2 and at least one protein of these ligand-receptor pairs has to be introduced most of the times, which could be a drawback in the future.

The success of expressing two heterologous proteins encoded by three genes in yeast actually extends yeast expression system for engineering more structurally complicated molecules. Instead of making single chain derivatives, essential immunity related membrane proteins such as MHCs, antibodies, and T cell receptors (TCRs), are all able to be displayed more natively as heterodimer on yeast surface for function studies.

Additionally, yeast display of a homotrimer like hemagglutinin would also become possible.

Yeast co-display remains some other advantages of surface-display technology. One example is the possibility of direct purification of soluble HLA-DR1. The solubility of HLA-DR1 produced in this system and the freedom of its dissociation from the surface though a little slow make routinely purification of recombinant MHC-II from this system become possible. It is demonstrated that the FLU/HLA-DR1 complex on the surface of yeast can be stripped off by treatment of Factor Xa (Chapter 2), so some easier approaches can be developed for extracting soluble HLA-DR1 or FLU/HLA-DR1 from yeast cells. However, although the affect is not yet shown in our system, the high-mannose yeast glycosylation and the low protein expression level are still major limitations for assembling human glycoproteins. Engineered yeast strains with fully humanized glycosylation pathway[187, 188] would probably help improved the performance of our system for producing recombinant proteins in the future. Another advantage of the surface-display system is that yeast co-displaying FLU/HLA-DR1 complexes has the potential to stimulate T cell response by direct contact of FLU/HLA-DR1 complex and TCR. This is partially proved by experiments elsewhere[134]. Finally, applying yeast co-display technology in the area of whole yeast vaccine design[189, 190] would shed light on development of novel immunotherapeutic strategies.

Chapter 7. References

1. Beck, G. & Habicht, G.S. Immunity and the invertebrates. *Sci Am* **275**, 60-63, 66 (1996).

2. Litman, G.W. Sharks and the origins of vertebrate immunity. *Sci Am* **275**, 67-71 (1996).

3. Litman, G.W., Cannon, J.P. & Dishaw, L.J. Reconstructing immune phylogeny: new perspectives. *Nat Rev Immunol* **5**, 866-879 (2005).

4. Chen, K. & Arnold, F.H. Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. *Proc Natl Acad Sci U S A* **90**, 5618-5622 (1993).

5. Romero, P.A. & Arnold, F.H. Exploring protein fitness landscapes by directed evolution. *Nat Rev Mol Cell Biol* **10**, 866-876 (2009).

6. Deem, M.W. Evolution and evolvability of proteins in the laboratory. *Proc Natl Acad Sci U S A* **101**, 3997-3998 (2004).

7. Earl, D.J. & Deem, M.W. Evolvability is a selectable trait. *Proc Natl Acad Sci U S A* **101**, 11531-11536 (2004).

8. Ahmed, R., Butler, L.D. & Bhatti, L. T4+ T helper cell function in vivo: differential requirement for induction of antiviral cytotoxic T-cell and antibody responses. *J Virol* **62**, 2102-2106 (1988).

9. Cunningham-Rundles, C. & Ponda, P.P. Molecular defects in T- and B-cell primary immunodeficiency diseases. *Nat Rev Immunol* **5**, 880-892 (2005).

10. Douek, D.C., Picker, L.J. & Koup, R.A. T cell dynamics in HIV-1 infection. *Annu Rev Immunol* **21**, 265-304 (2003).

11. Schwartz, R.H. A cell culture model for T lymphocyte clonal anergy. *Science* **248**, 1349-1356 (1990).

12. Bromley, S.K. et al. The immunological synapse. *Annu Rev Immunol* **19**, 375-396 (2001).

13. Smith-Garvin, J.E., Koretzky, G.A. & Jordan, M.S. T cell activation. *Annu Rev Immunol* **27**, 591-619 (2009).

14. Complete sequence and gene map of a human major histocompatibility complex. The MHC sequencing consortium. *Nature* **401**, 921-923 (1999).

15. Reche, P.A. & Reinherz, E.L. Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. *J Mol Biol* **331**, 623-641 (2003).

16. Hughes, A.L. & Yeager, M. Natural selection at major histocompatibility complex loci of vertebrates. *Annu Rev Genet* **32**, 415-435 (1998).

17. Martinsohn, J.T., Sousa, A.B., Guethlein, L.A. & Howard, J.C. The gene conversion hypothesis of MHC evolution: a review. *Immunogenetics* **50**, 168-200 (1999).

18. O'Sullivan, D. et al. On the interaction of promiscuous antigenic peptides with different DR alleles. Identification of common structural motifs. *J Immunol* **147**, 2663-2669 (1991).

19. Chicz, R.M. et al. Specificity and promiscuity among naturally processed peptides bound to HLA-DR alleles. *J Exp Med* **178**, 27-47 (1993).

20. Hammer, J. et al. Promiscuous and allele-specific anchors in HLA-DR-binding peptides. *Cell* **74**, 197-203 (1993).

21. Suri, A., Lovitch, S.B. & Unanue, E.R. The wide diversity and complexity of peptides bound to class II MHC molecules. *Curr Opin Immunol* **18**, 70-77 (2006).

22. Falk, K., Rotzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.G. Pool sequencing of natural HLA-DR, DQ, and DP ligands reveals detailed peptide motifs, constraints of processing, and general rules. *Immunogenetics* **39**, 230-242 (1994).

23. Sturniolo, T. et al. Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol* **17**, 555-561 (1999).

24. Flajnik, M.F. & Du Pasquier, L. Evolution of innate and adaptive immunity: can we draw a line? *Trends Immunol* **25**, 640-644 (2004).

25. Abbas, A.K., Lichtman, A.H. & Pillai, S., Edn. 6. (Saunders, philadelphia; 2007).

26. Nestle, F.O. Dendritic cell vaccination for cancer therapy. *Oncogene* **19**, 6673-6679 (2000).

27. Oelke, M., Krueger, C., Giuntoli, R.L., 2nd & Schneck, J.P. Artificial antigen-presenting cells: artificial solutions for real diseases. *Trends Mol Med* **11**, 412-420 (2005).

28. Sette, A. & Fikes, J. Epitope-based vaccines: an update on epitope identification, vaccine design and delivery. *Curr Opin Immunol* **15**, 461-470 (2003).

29. Vyas, J.M., Van der Veen, A.G. & Ploegh, H.L. The known unknowns of antigen processing and presentation. *Nat Rev Immunol* **8**, 607-618 (2008).

30. Schroder, K., Hertzog, P.J., Ravasi, T. & Hume, D.A. Interferon-gamma: an overview of signals, mechanisms and functions. *J Leukoc Biol* **75**, 163-189 (2004).

31. Zhao, Y., Wilson, D., Matthews, S. & Yap, G.S. Rapid elimination of Toxoplasma gondii by gamma interferon-primed mouse macrophages is independent of CD40 signaling. *Infect Immun* **75**, 4799-4803 (2007).

32. Klaus, S.J., Berberich, I., Shu, G. & Clark, E.A. CD40 and its ligand in the regulation of humoral immunity. *Semin Immunol* **6**, 279-286 (1994).

33. van Kooten, C. & Banchereau, J. CD40-CD40 ligand. *J Leukoc Biol* **67**, 2-17 (2000).

34. Stout, R.D., Suttles, J., Xu, J., Grewal, I.S. & Flavell, R.A. Impaired T cell-mediated macrophage activation in CD40 ligand-deficient mice. *J Immunol* **156**, 8-11 (1996).

35. Palucka, K. & Banchereau, J. How dendritic cells and microbes interact to elicit or subvert protective immune responses. *Curr Opin Immunol* **14**, 420-431 (2002).

36. Bevan, M.J. Cross-priming. *Nat Immunol* **7**, 363-365 (2006).

37. Fonteneau, J.F., Larsson, M. & Bhardwaj, N. Interactions between dead cells and dendritic cells in the induction of antiviral CTL responses. *Curr Opin Immunol* **14**, 471-477 (2002).

38. Wang, R.F. Enhancing antitumor immune responses: intracellular peptide delivery and identification of MHC class II-restricted tumor antigens. *Immunol Rev* **188**, 65-80 (2002).

39. Sarawar, S.R., Lee, B.J., Reiter, S.K. & Schoenberger, S.P. Stimulation via CD40 can substitute for CD4 T cell function in preventing reactivation of a latent herpesvirus. *Proc Natl Acad Sci U S A* **98**, 6325-6329 (2001).

40. Ndhlovu, L.C., Loo, C.P., Spotts, G., Nixon, D.F. & Hecht, F.M. FOXP3 expressing CD127lo CD4+ T cells inversely correlate with CD38+ CD8+ T cell activation levels in primary HIV-1 infection. *J Leukoc Biol* **83**, 254-262 (2008).

41. Williams, M.A., Tyznik, A.J. & Bevan, M.J. Interleukin-2 signals during priming are required for secondary expansion of CD8+ memory T cells. *Nature* **441**, 890-893 (2006).

42. Obar, J.J. et al. CD4+ T cell regulation of CD25 expression controls development of short-lived effector CD8+ T cells in primary and secondary responses. *Proc Natl Acad Sci U S A* (2009).

43. Ridge, J.P., Di Rosa, F. & Matzinger, P. A conditioned dendritic cell can be a temporal bridge between a CD4+ T-helper and a T-killer cell. *Nature* **393**, 474-478 (1998).

44. Bennett, S.R. et al. Help for cytotoxic-T-cell responses is mediated by CD40 signalling. *Nature* **393**, 478-480 (1998).

45. Schoenberger, S.P., Toes, R.E., van der Voort, E.I., Offringa, R. & Melief, C.J. T-cell help for cytotoxic T lymphocytes is mediated by CD40-CD40L interactions. *Nature* **393**, 480-483 (1998).

46. Casares, S. et al. Down-regulation of diabetogenic CD4+ T cells by a soluble dimeric peptide-MHC class II chimera. *Nat Immunol* **3**, 383-391 (2002).

47. Masteller, E.L. et al. Peptide-MHC class II dimers as therapeutics to modulate antigen-specific T cell responses in autoimmune diabetes. *J Immunol* **171**, 5587-5595 (2003).

48. Sharma, S.D. et al. Antigen-specific therapy of experimental allergic encephalomyelitis by soluble class II major histocompatibility complex-peptide complexes. *Proc Natl Acad Sci U S A* **88**, 11465-11469 (1991).

49. Mayordomo, J.I. et al. Bone marrow-derived dendritic cells pulsed with synthetic tumour peptides elicit protective and therapeutic antitumour immunity. *Nat Med* **1**, 1297-1302 (1995).

50. Pepper, L.R., Hammer, D.A. & Boder, E.T. Rolling adhesion of alphaL I domain mutants decorrelated from binding affinity. *J Mol Biol* **360**, 37-44 (2006).

51. Germain, R.N. Antigen presentation. The second class story. *Nature* **353**, 605-607 (1991).

52. Hedrick, P.W. Balancing selection and MHC. *Genetica* **104**, 207-214 (1998).

53. Hedrick, P. Evolutionary genomics: foxy MHC selection story. *Heredity* **93**, 237-238 (2004).

54. Brown, J.H. et al. Three-dimensional structure of the human class II histocompatibility antigen HLA-DR1. *Nature* **364**, 33-39 (1993).

55. Dessen, A., Lawrence, C.M., Cupo, S., Zaller, D.M. & Wiley, D.C. X-ray crystal structure of HLA-DR4 (DRA*0101, DRB1*0401) complexed with a peptide from human collagen II. *Immunity* **7**, 473-481 (1997).

56. Henderson, K.N. et al. A structural and immunological basis for the role of human leukocyte antigen DQ8 in celiac disease. *Immunity* **27**, 23-34 (2007).

57. Jardetzky, T.S. et al. Crystallographic analysis of endogenous peptides associated with HLA-DR1 suggests a common, polyproline II-like conformation for bound peptides. *Proc Natl Acad Sci U S A* **93**, 734-738 (1996).

58. Lee, K.H., Wucherpfennig, K.W. & Wiley, D.C. Structure of a human insulin peptide-HLA-DQ8 complex and susceptibility to type 1 diabetes. *Nat Immunol* **2**, 501-507 (2001).

59. Stern, L.J. et al. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* **368**, 215-221 (1994).

60. Chicz, R.M. et al. Predominant naturally processed peptides bound to HLA-DR1 are derived from MHC-related molecules and are heterogeneous in size. *Nature* **358**, 764-768 (1992).

61. Hennecke, J., Carfi, A. & Wiley, D.C. Structure of a covalently stabilized complex of a human alphabeta T-cell receptor, influenza HA peptide and MHC class II molecule, HLA-DR1. *EMBO J* **19**, 5611-5624 (2000).

62. Wang, J.H. et al. Crystal structure of the human CD4 N-terminal two-domain fragment complexed to a class II MHC molecule. *Proc Natl Acad Sci U S A* **98**, 10799-10804 (2001).

63. Watts, C. The exogenous pathway for antigen presentation on major histocompatibility complex class II and CD1 molecules. *Nat Immunol* **5**, 685-692 (2004).

64. Stern, L.J., Potolicchio, I. & Santambrogio, L. MHC class II compartment subtypes: structure and function. *Curr Opin Immunol* **18**, 64-69 (2006).

65. Busch, R., Doebele, R.C., Patil, N.S., Pashine, A. & Mellins, E.D. Accessory molecules for MHC class II peptide loading. *Curr Opin Immunol* **12**, 99-106 (2000).

66. Natarajan, S.K., Assadi, M. & Sadegh-Nasseri, S. Stable peptide binding to MHC class II molecule is rapid and is determined by a receptive conformation shaped by prior association with low affinity peptides. *J Immunol* **162**, 4030-4036 (1999).

67. Germain, R.N. & Hendrix, L.R. MHC class II structure, occupancy and surface expression determined by post-endoplasmic reticulum antigen binding. *Nature* **353**, 134-139 (1991).

68. Sadegh-Nasseri, S. & Germain, R.N. A role for peptide in determining MHC class II structure. *Nature* **353**, 167-170 (1991).

69. Sato, A.K. et al. Determinants of the peptide-induced conformational change in the human class II major histocompatibility complex protein HLA-DR1. *J Biol Chem* **275**, 2165-2173 (2000).

70. Engelhard, V.H. Structure of peptides associated with class I and class II MHC molecules. *Annu Rev Immunol* **12**, 181-207 (1994).

71. Huan, J. et al. Monomeric recombinant TCR ligand reduces relapse rate and severity of experimental autoimmune encephalomyelitis in SJL/J mice through cytokine switch. *J Immunol* **172**, 4556-4566 (2004).

72. Lu, X. et al. Suppression of major histocompatibility complex class II-associated invariant chain enhances the potency of an HIV gp120 DNA vaccine. *Immunology* **120**, 207-216 (2007).

73. Boder, E.T., Midelfort, K.S. & Wittrup, K.D. Directed evolution of antibody fragments with monovalent femtomolar antigen-binding affinity. *Proc Natl Acad Sci U S A* **97**, 10701-10705 (2000).

74. Boder, E.T. & Wittrup, K.D. Yeast surface display for screening combinatorial polypeptide libraries. *Nat Biotechnol* **15**, 553-557 (1997).

75. Chervin, A.S., Aggen, D.H., Raseman, J.M. & Kranz, D.M. Engineering higher affinity T cell receptors using a T cell display system. *J Immunol Methods* **339**, 175-184 (2008).

76. Crawford, F. et al. Use of baculovirus MHC/peptide display libraries to characterize T-cell receptor ligands. *Immunol Rev* **210**, 156-170 (2006).

77. Ernst, W. et al. Baculovirus surface display: construction and screening of a eukaryotic epitope library. *Nucleic Acids Res* **26**, 1718-1723 (1998).

78. Fernandez-Gacio, A., Uguen, M. & Fastrez, J. Phage display as a tool for the directed evolution of enzymes. *Trends Biotechnol* **21**, 408-414 (2003).

79. Francisco, J.A., Campbell, R., Iverson, B.L. & Georgiou, G. Production and fluorescence-activated cell sorting of Escherichia coli expressing a functional antibody fragment on the external surface. *Proc Natl Acad Sci U S A* **90**, 10444-10448 (1993).

80. Georgiou, G. et al. Display of heterologous proteins on the surface of microorganisms: from the screening of combinatorial libraries to live recombinant vaccines. *Nat Biotechnol* **15**, 29-34 (1997).

81. Ito, J. et al. Regulation of the display ratio of enzymes on the Saccharomyces cerevisiae cell surface by the immunoglobulin G and cellulosomal enzyme binding domains. *Appl Environ Microbiol* **75**, 4149-4154 (2009).

82. Lee, J.H., Goulian, M. & Boder, E.T. Autocatalytic activation of influenza hemagglutinin. *J Mol Biol* **364**, 275-282 (2006).

83. Li, Y. et al. Directed evolution of human T-cell receptors with picomolar affinities by phage display. *Nat Biotechnol* **23**, 349-354 (2005).

84. Makela, A.R. & Oker-Blom, C. Baculovirus display: a multifunctional technology for gene delivery and eukaryotic library development. *Adv Virus Res* **68**, 91-112 (2006).

85. Narita, J. et al. Display of active enzymes on the cell surface of Escherichia coli using PgsA anchor protein and their application to bioconversion. *Appl Microbiol Biotechnol* **70**, 564-572 (2006).

86. Scott, J.K. & Smith, G.P. Searching for peptide ligands with an epitope library. *Science* **249**, 386-390 (1990).

87. Shusta, E.V., Holler, P.D., Kieke, M.C., Kranz, D.M. & Wittrup, K.D. Directed evolution of a stable scaffold for T-cell receptor engineering. *Nat Biotechnol* **18**, 754-759 (2000).

88. van den Beucken, T. et al. Affinity maturation of Fab antibody fragments by fluorescent-activated cell sorting of yeast-displayed libraries. *FEBS Lett* **546**, 288-294 (2003).

89. Wolkowicz, R. & Nolan, G.P. Retroviral technology--applications for expressed peptide libraries. *Front Biosci* **8**, d603-619 (2003).

90. Hammer, J., Takacs, B. & Sinigaglia, F. Identification of a motif for HLA-DR1 binding peptides using M13 display libraries. *J Exp Med* **176**, 1007-1013 (1992).

91. Smith, G.P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317 (1985).

92. Charbit, A., Boulain, J.C., Ryter, A. & Hofnung, M. Probing the topology of a bacterial membrane protein by genetic insertion of a foreign epitope; expression at the cell surface. *EMBO J* **5**, 3029-3037 (1986).

93. Hanes, J. & Pluckthun, A. In vitro selection and evolution of functional proteins by using ribosome display. *Proc Natl Acad Sci U S A* **94**, 4937-4942 (1997).

94. Roberts, R.W. & Szostak, J.W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proc Natl Acad Sci U S A* **94**, 12297-12302 (1997).

95. Gold, L. mRNA display: diversity matters during in vitro selection. *Proc Natl Acad Sci U S A* **98**, 4825-4826 (2001).

96. Fukuda, I. et al. In vitro evolution of single-chain antibodies using mRNA display. *Nucleic Acids Res* **34**, e127 (2006).

97. Domingo, D.L. & Trowbridge, I.S. Characterization of the human transferrin receptor produced in a baculovirus expression system. *J Biol Chem* **263**, 13386-13392 (1988).

98. Boen, E., Crownover, A.R., McIlhaney, M., Korman, A.J. & Bill, J. Identification of T cell ligands in a library of peptides covalently attached to HLA-DR4. *J Immunol* **165**, 2040-2047 (2000).

99. Kessels, H.W., van Den Boom, M.D., Spits, H., Hooijberg, E. & Schumacher, T.N. Changing T cell specificity by retroviral T cell receptor display. *Proc Natl Acad Sci U S A* **97**, 14578-14583 (2000).

100. Makela, A.R. & Oker-Blom, C. The baculovirus display technology--an evolving instrument for molecular screening and drug delivery. *Comb Chem High Throughput Screen* **11**, 86-98 (2008).

101. Boder, E.T. & Wittrup, K.D. Optimal screening of surface-displayed polypeptide libraries. *Biotechnol Prog* **14**, 55-62 (1998).

102. Boder, E.T., Bill, J.R., Nields, A.W., Marrack, P.C. & Kappler, J.W. Yeast surface display of a noncovalent MHC class II heterodimer complexed with antigenic peptide. *Biotechnol Bioeng* **92**, 485-491 (2005).

103. Shen, Z.M. et al. Delineation of functional regions within the subunits of the Saccharomyces cerevisiae cell adhesion molecule a-agglutinin. *J Biol Chem* **276**, 15768-15775 (2001).

104. Lipke, P.N. & Kurjan, J. Sexual agglutination in budding yeasts: structure, function, and regulation of adhesion glycoproteins. *Microbiol Rev* **56**, 180-194 (1992).

105. Gorga, J.C., Horejsi, V., Johnson, D.R., Raghupathy, R. & Strominger, J.L. Purification and characterization of class II histocompatibility antigens from a homozygous human B cell line. *J Biol Chem* **262**, 16087-16094 (1987).

106. Stern, L.J. & Wiley, D.C. The human class II MHC protein HLA-DR1 assembles as empty alpha beta heterodimers in the absence of antigenic peptide. *Cell* **68**, 465-477 (1992).

107. Kalandadze, A., Galleno, M., Foncerrada, L., Strominger, J.L. & Wucherpfennig, K.W. Expression of recombinant HLA-DR2 molecules. Replacement of the hydrophobic transmembrane region by a leucine zipper dimerization motif allows the assembly and secretion of soluble DR alpha beta heterodimers. *J Biol Chem* **271**, 20156-20162 (1996).

108. Altman, J.D., Reay, P.A. & Davis, M.M. Formation of functional peptide complexes of class II major histocompatibility complex proteins from subunits produced in Escherichia coli. *Proc Natl Acad Sci U S A* **90**, 10330-10334 (1993).

109. Stockel, J. et al. Refolding of human class II major histocompatibility complex molecules isolated from Escherichia coli. Assembly of peptide-free heterodimers and increased refolding-yield in the presence of antigenic peptide. *J Biol Chem* **269**, 29571-29578 (1994).

110. Gauthier, L. et al. Expression and crystallization of the complex of HLA-DR2 (DRA, DRB1*1501) and an immunodominant peptide of human myelin basic protein. *Proc Natl Acad Sci U S A* **95**, 11828-11833 (1998).

111. Frayser, M., Sato, A.K., Xu, L. & Stern, L.J. Empty and peptide-loaded class II major histocompatibility complex proteins produced by expression in Escherichia coli and folding in vitro. *Protein Expr Purif* **15**, 105-114 (1999).

112. Justesen, S., Harndahl, M., Lamberth, K., Nielsen, L.L. & Buus, S. Functional recombinant MHC class II molecules and high-throughput peptide-binding assays. *Immunome Res* **5**, 2 (2009).

113. Merrifield, B. Concept and early development of solid-phase peptide synthesis. *Methods Enzymol* **289**, 3-13 (1997).

114. Merrifield, R.B. Solid-phase peptide synthesis. *Adv Enzymol Relat Areas Mol Biol* **32**, 221-296 (1969).

115. Joshi, R.V., Zarutskie, J.A. & Stern, L.J. A three-step kinetic mechanism for peptide binding to MHC class II proteins. *Biochemistry* **39**, 3751-3762 (2000).

116. Kasson, P.M., Rabinowitz, J.D., Schmitt, L., Davis, M.M. & McConnell, H.M. Kinetics of peptide binding to the class II MHC protein I-Ek. *Biochemistry* **39**, 1048-1058 (2000).

117. Roche, P.A. & Cresswell, P. High-affinity binding of an influenza hemagglutinin-derived peptide to purified HLA-DR. *J Immunol* **144**, 1849-1856 (1990).

118. Jardetzky, T.S. et al. Peptide binding to HLA-DR1: a peptide with most residues substituted to alanine retains MHC binding. *EMBO J* **9**, 1797-1803 (1990).

119. Krieger, J.I. et al. Single amino acid changes in DR and antigen define residues critical for peptide-MHC binding and T cell recognition. *J Immunol* **146**, 2331-2340 (1991).

120. O'Sullivan, D., Sidney, J., Del Guercio, M.F., Colon, S.M. & Sette, A. Truncation analysis of several DR binding epitopes. *J Immunol* **146**, 1240-1246 (1991).

121. Johansen, B.H., Vartdal, F., Eriksen, J.A., Thorsby, E. & Sollid, L.M. Identification of a putative motif for binding of peptides to HLA-DQ2. *Int Immunol* **8**, 177-182 (1996).

122. Sidney, J., del Guercio, M.F., Southwood, S. & Sette, A. The HLA molecules DQA1*0501/B1*0201 and DQA1*0301/B1*0302 share an extensive overlap in peptide binding specificity. *J Immunol* **169**, 5098-5108 (2002).

123. Southwood, S. et al. Several common HLA-DR types share largely overlapping peptide binding repertoires. *J Immunol* **160**, 3363-3373 (1998).

124. Ettinger, R.A., Papadopoulos, G.K., Moustakas, A.K., Nepom, G.T. & Kwok, W.W. Allelic variation in key peptide-binding pockets discriminates between closely related diabetes-protective and diabetes-susceptible HLA-DQB1*06 alleles. *J Immunol* **176**, 1988-1998 (2006).

125. Santambrogio, L., Sato, A.K., Fischer, F.R., Dorf, M.E. & Stern, L.J. Abundant empty class II MHC molecules on the surface of immature dendritic cells. *Proc Natl Acad Sci U S A* **96**, 15050-15055 (1999).

126. Vacchino, J.F. & McConnell, H.M. Peptide binding to active class II MHC protein on the cell surface. *J Immunol* **166**, 6680-6685 (2001).

127. Wettstein, D.A., Boniface, J.J., Reay, P.A., Schild, H. & Davis, M.M. Expression of a class II major histocompatibility complex (MHC) heterodimer in a lipid-linked form with enhanced peptide/soluble MHC complex formation at low pH. *J Exp Med* **174**, 219-228 (1991).

128. Zhou, Z., Callaway, K.A., Weber, D.A. & Jensen, P.E. Cutting edge: HLA-DM functions through a mechanism that does not require specific conserved hydrogen bonds in class II MHC-peptide complexes. *J Immunol* **183**, 4187-4191 (2009).

129. Kurokawa, M.S. et al. Expression of MHC class I molecules together with antigenic peptides on filamentous phages. *Immunol Lett* **80**, 163-168 (2002).

130. Le Doussal, J., Piqueras, B., Dogan, I., Debre, P. & Gorochov, G. Phage display of peptide/major histocompatibility complex. *J Immunol Methods* **241**, 147-158 (2000).

131. Vest Hansen, N., Ostergaard Pedersen, L., Stryhn, A. & Buus, S. Phage display of peptide / major histocompatibility class I complexes. *Eur J Immunol* **31**, 32-38 (2001).

132. Brophy, S.E., Holler, P.D. & Kranz, D.M. A yeast display system for engineering functional peptide-MHC complexes. *J Immunol Methods* **272**, 235-246 (2003).

133. Esteban, O. & Zhao, H. Directed evolution of soluble single-chain human class II MHC molecules. *J Mol Biol* **340**, 81-95 (2004).

134. Wen, F., Esteban, O. & Zhao, H. Rapid identification of CD4+ T-cell epitopes using yeast displaying pathogen-derived peptide library. *J Immunol Methods* **336**, 37-44 (2008).

135. Zhu, X. et al. A recombinant single-chain human class II MHC molecule (HLA-DR1) as a covalently linked heterotrimer of alpha chain, beta chain, and antigenic peptide, with immunogenicity in vitro and reduced affinity for bacterial superantigens. *Eur J Immunol* **27**, 1933-1941 (1997).

136. Fields, S. & Song, O. A novel genetic system to detect protein-protein interactions. *Nature* **340**, 245-246 (1989).

137. Hu, X., Kang, S., Chen, X., Shoemaker, C.B. & Jin, M.M. Yeast surface two-hybrid for quantitative in vivo detection of protein-protein interactions via the secretory pathway. *J Biol Chem* **284**, 16369-16376 (2009).

138. Jeong, K.J., Seo, M.J., Iverson, B.L. & Georgiou, G. APEx 2-hybrid, a quantitative protein-protein interaction assay for antibody discovery and engineering. *Proc Natl Acad Sci U S A* **104**, 8247-8252 (2007).

139. Krebber, C. et al. Selectively-infective phage (SIP): a mechanistic dissection of a novel in vivo selection for protein-ligand interactions. *J Mol Biol* **268**, 607-618 (1997).

140. Nyfeler, B., Michnick, S.W. & Hauri, H.P. Capturing protein interactions in the secretory pathway of living cells. *Proc Natl Acad Sci U S A* **102**, 6350-6355 (2005).

141. Stagljar, I. & Fields, S. Analysis of membrane protein interactions using yeast-based technologies. *Trends Biochem Sci* **27**, 559-563 (2002).

142. Sikorski, R.S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae. *Genetics* **122**, 19-27 (1989).

143. Vembar, S.S. & Brodsky, J.L. One step at a time: endoplasmic reticulum-associated degradation. *Nat Rev Mol Cell Biol* **9**, 944-957 (2008).

144. Sette, A. et al. Effect of pH on MHC class II-peptide interactions. *J Immunol* **148**, 844-851 (1992).

145. Hammer, J. et al. High-affinity binding of short peptides to major histocompatibility complex class II molecules by anchor combinations. *Proc Natl Acad Sci U S A* **91**, 4456-4460 (1994).

146. Holler, P.D., Chlewicki, L.K. & Kranz, D.M. TCRs with high affinity for foreign pMHC show self-reactivity. *Nat Immunol* **4**, 55-62 (2003).

147. Tarassov, K. et al. An in vivo map of the yeast protein interactome. *Science* **320**, 1465-1470 (2008).

148. Yu, H. et al. High-quality binary protein interaction map of the yeast interactome network. *Science* **322**, 104-110 (2008).

149. Berlier, J.E. et al. Quantitative comparison of long-wavelength Alexa Fluor dyes to Cy dyes: fluorescence of the dyes and their bioconjugates. *J Histochem Cytochem* **51**, 1699-1712 (2003).

150. Panchuk-Voloshina, N. et al. Alexa dyes, a series of new fluorescent dyes that yield exceptionally bright, photostable conjugates. *J Histochem Cytochem* **47**, 1179-1188 (1999).

151. Davenport, M.P., Ho Shon, I.A. & Hill, A.V. An empirical method for the prediction of T-cell epitopes. *Immunogenetics* **42**, 392-397 (1995).

152. Davenport, M.P. et al. Naturally processed peptides from two disease-resistance-associated HLA-DR13 alleles show related sequence motifs and the effects of the dimorphism at position 86 of the HLA-DR beta chain. *Proc Natl Acad Sci U S A* **92**, 6567-6571 (1995).

153. Falk, K., Rotzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.G. Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature* **351**, 290-296 (1991).

154. Godkin, A. et al. Use of eluted peptide sequence data to identify the binding characteristics of peptides to the insulin-dependent diabetes susceptibility allele HLA-DQ8 (DQ 3.2). *Int Immunol* **9**, 905-911 (1997).

155. Suri, A., Walters, J.J., Gross, M.L. & Unanue, E.R. Natural peptides selected by diabetogenic DQ8 and murine I-A(g7) molecules show common sequence specificity. *J Clin Invest* **115**, 2268-2276 (2005).

156. Wicker, L.S. et al. Naturally processed T cell epitopes from human glutamic acid decarboxylase identified using mice transgenic for the type 1 diabetes-associated human MHC class II allele, DRB1*0401. *J Clin Invest* **98**, 2597-2603 (1996).

157. Tong, J.C., Tan, T.W. & Ranganathan, S. Methods and protocols for prediction of immunogenic epitopes. *Brief Bioinform* **8**, 96-108 (2007).

158. Androulakis, I.P., Nayak, N.N., Ierapetritou, M.G., Monos, D.S. & Floudas, C.A. A predictive method for the evaluation of peptide binding in pocket 1 of HLA-DRB1 via global minimization of energy interactions. *Proteins* **29**, 87-102 (1997).

159. Brunger, A.T. et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921 (1998).

160. Driscoll, P.C. et al. Two-dimensional nuclear magnetic resonance analysis of a labeled peptide bound to a class II major histocompatibility complex molecule. *J Mol Biol* **232**, 342-350 (1993).

161. Ghosh, P., Amaya, M., Mellins, E. & Wiley, D.C. The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. *Nature* **378**, 457-462 (1995).

162. Lee, C. & McConnell, H.M. A general model of invariant chain association with class II major histocompatibility complex proteins. *Proc Natl Acad Sci U S A* **92**, 8269-8273 (1995).

163. Painter, C.A., Cruz, A., Lopez, G.E., Stern, L.J. & Zavala-Ruiz, Z. Model for the peptide-free conformation of class II MHC proteins. *PLoS One* **3**, e2403 (2008).

164. Schmitt, L., Boniface, J.J., Davis, M.M. & McConnell, H.M. Conformational isomers of a class II MHC-peptide complex in solution. *J Mol Biol* **286**, 207-218 (1999).

165. Wilson, N., Fremont, D., Marrack, P. & Kappler, J. Mutations changing the kinetics of class II MHC peptide exchange. *Immunity* **14**, 513-522 (2001).

166. Doebele, R.C. et al. Point mutations in or near the antigen-binding groove of HLA-DR3 implicate class II-associated invariant chain peptide affinity as a constraint on MHC class II polymorphism. *J Immunol* **170**, 4683-4692 (2003).

167. Sato, A.K., Sturniolo, T., Sinigaglia, F. & Stern, L.J. Substitution of aspartic acid at beta57 with alanine alters MHC class II peptide binding activity but not protein stability: HLA-DQ (alpha1*0201, beta1*0302) and (alpha1*0201, beta1*0303). *Hum Immunol* **60**, 1227-1236 (1999).

168. Starwalt, S.E., Masteller, E.L., Bluestone, J.A. & Kranz, D.M. Directed evolution of a single-chain class II MHC product by yeast display. *Protein Eng* **16**, 147-156 (2003).

169. Appel, H., Gauthier, L., Pyrdol, J. & Wucherpfennig, K.W. Kinetics of T-cell receptor binding by bivalent HLA-DR. Peptide complexes that activate antigen-specific human T-cells. *J Biol Chem* **275**, 312-321 (2000).

170. Cochran, J.R., Cameron, T.O. & Stern, L.J. The relationship of MHC-peptide binding and T cell activation probed using chemically defined MHC class II oligomers. *Immunity* **12**, 241-250 (2000).

171. Fox, B.S., Quill, H., Carlson, L. & Schwartz, R.H. Quantitative analysis of the T cell response to antigen and planar membranes containing purified Ia molecules. *J Immunol* **138**, 3367-3374 (1987).

172. Krogsgaard, M. et al. Agonist/endogenous peptide-MHC heterodimers drive T cell activation and sensitivity. *Nature* **434**, 238-243 (2005).

173. Nag, B. et al. Stimulation of T cells by antigenic peptide complexed with isolated chains of major histocompatibility complex class II molecules. *Proc Natl Acad Sci U S A* **90**, 1604-1608 (1993).

174. Quill, H. & Schwartz, R.H. Stimulation of normal inducer T cell clones with antigen presented by purified Ia molecules in planar lipid membranes: specific induction of a long-lived state of proliferative nonresponsiveness. *J Immunol* **138**, 3704-3712 (1987).

175. Hackett, C.J. & Sharma, O.K. Frontiers in peptide-MHC class II multimer technology. *Nat Immunol* **3**, 887-889 (2002).

176. Ferlin, W., Glaichenhaus, N. & Mougneau, E. Present difficulties and future promise of MHC multimers in autoimmune exploration. *Curr Opin Immunol* **12**, 670-675 (2000).

177. Stemmer, W.P. Rapid evolution of a protein in vitro by DNA shuffling. *Nature* **370**, 389-391 (1994).

178. Stemmer, W.P. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci U S A* **91**, 10747-10751 (1994).

179. Busch, R., Hill, C.M., Hayball, J.D., Lamb, J.R. & Rothbard, J.B. Effect of natural polymorphism at residue 86 of the HLA-DR beta chain on peptide binding. *J Immunol* **147**, 1292-1298 (1991).

180. Cizman, B.B. et al. Identification of DRB1 allele (DRB1*1316) with aspartate at position 86: evolutionary considerations and functional implications. *Tissue Antigens* **47**, 153-154 (1996).

181. Ong, B. et al. Critical role for the Val/Gly86 HLA-DR beta dimorphism in autoantigen presentation to human T cells. *Proc Natl Acad Sci U S A* **88**, 7343-7347 (1991).

182. Smith, K.J., Pyrdol, J., Gauthier, L., Wiley, D.C. & Wucherpfennig, K.W. Crystal structure of HLA-DR2 (DRA*0101, DRB1*1501) complexed with a peptide from human myelin basic protein. *J Exp Med* **188**, 1511-1520 (1998).

183. Kwok, W.W., Domeier, M.E., Johnson, M.L., Nepom, G.T. & Koelle, D.M. HLA-DQB1 codon 57 is critical for peptide binding and recognition. *J Exp Med* **183**, 1253-1258 (1996).

184. Todd, J.A., Bell, J.I. & McDevitt, H.O. HLA-DQ beta gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature* **329**, 599-604 (1987).

185. Weber, D.A., Evavold, B.D. & Jensen, P.E. Enhanced dissociation of HLA-DR-bound peptides in the presence of HLA-DM. *Science* **274**, 618-620 (1996).

186. Narayan, K. et al. HLA-DM targets the hydrogen bond between the histidine at position beta81 and peptide to dissociate HLA-DR-peptide complexes. *Nat Immunol* **8**, 92-100 (2007).

187. Hamilton, S.R. & Gerngross, T.U. Glycosylation engineering in yeast: the advent of fully humanized yeast. *Curr Opin Biotechnol* **18**, 387-392 (2007).

188. Wildt, S. & Gerngross, T.U. The humanization of N-glycosylation pathways in yeast. *Nat Rev Microbiol* **3**, 119-128 (2005).

189. Franzusoff, A., Duke, R.C., King, T.H., Lu, Y. & Rodell, T.C. Yeasts encoding tumour antigens in cancer immunotherapy. *Expert Opin Biol Ther* **5**, 565-575 (2005).

190. Stubbs, A.C. et al. Whole recombinant yeast vaccine activates dendritic cells and elicits protective cell-mediated immunity. *Nat Med* **7**, 625-629 (2001).