Thesis for the Master's degree in Molecular Biosciences
Main field of study in molecular biology

Polymorphic residues of HLA-DQ2.2 and HLA-DQ2.5 that affect CLIP presentation and stability

Axel Berg-Larsen

60 study points

Department of Molecular Biosciences
Faculty of mathematics and natural sciences
UNIVERSITY OF OSLO 10/2009

Abstract

Celiac disease is an autoimmune disease driven by an immune response to gluten peptides. The disease is strongly associated with the MHC class II molecule HLA-DQ2.5. The similar molecule HLA-DQ2.2 is not associated with celiac disease. The two molecules have very similar peptide binding domains and are both able to bind gluten T-cell epitopes. A marked difference in CLIP presentation at the peptide binding groove has been observed between the two HLA molecules. HLA-DQ2.5 presents large amounts of CLIP peptides in the binding groove, while HLA-DQ2.2 does not. There are only 10 polymorphic residues in the membrane distal domains of these HLA- molecules. This thesis is part of a large study using site-directed mutagenesis, flow cytometry and mass spectrometry to clarify which, if any, of these polymorphic residues are responsible for the difference seen in CLIP presentation and stability. This investigation may also elucidate part of the differences seen in celiac disease association. Results from parts of the study have been published by Fallang.et.al.[1] showing that the polymorphism in position α22 is of great importance for the presentation and stability of CLIP and other peptides bound to HLA-DQ2.2 and HLA-DQ2.5.

Findings presented in this thesis show that the polymorphisms in positions α31, α37 and α72 also have an effect on CLIP presentation and stability in HLA-DQ2.5 and HLA-DQ2.2. Their effect appears to be the opposite of the one described in the previous study, and may help explain some of the results presented there.

Acknowledgments

This work has been carried out at the Institute of Immunology, Rikshospitalet.

I would like to thank my supervisor, Ludvig Sollid, for the opportunity to do my master thesis in his group. He has offered wise council and great insight throughout my studies. More than anything, he has been an inspiration and an excellent critic, pushing me to better myself. Lars-Egil Fallang and Elin Bergseng have been my co-supervisors, and they have my profound thanks for all the help and support they have given me in the laboratory and with my thesis. They were always willing to help and answer any question. Their patience with a bumbling master student that sometimes made foolish mistakes has been commendable. Lars-Egil Fallang has helped me relearn all I thought I knew about molecular biology and introduced me to cell culturing of some very delicate cells. Elin Bergseng has helped me discover a new passion for protein chemistry and has shown me the value of an orderly mind both in and out of the laboratory.

My thanks goes out to the other members of the Sollid group whom I have worked closely with for nearly 2 years and who have created a pleasant and goal oriented work environment. In particular, I wish to thank Bjørg Simonsen and Marie Johannesen for their support and help. Their practical knowledge and experience has shown me how true masters of molecular biology work in the laboratory, and without their advice I would never have been able to complete this thesis.

Lastly, I wish to thank my internal supervisor Inger Sandlie. She has always been available for questions regarding my thesis and her experience regarding the intricacies of the master degree process in particular has been an enormous help.

Axel Berg-Larsen, Oslo, October 2009

Abbreviations

αCN	α-cyano-4- hydroxycinnamic acid	**PCR**	Polymerase Chain Reaction
APC	Antigen presenting cell	**PSMF**	Phenylmethanesulphonylfluoride
Az.	Azide (N_3^-)	**RMPI**	Roswell Park Memorial Institute
cDNA	Complementary deoxyribonucleic acid	**RPM**	Revolutions per minute
CD	Cluster of differentiation	**RT**	Reverse transcription
BCR	B-cell receptor	**s**	Second
CLIP	Class II-associated invariant chain peptide	**Tm**	Melting temperature
dNTP	Deoxyribonucleotide triphosphate	**TM**	Transmembrane
DMEM	Dulbecco's Modified Eagle Medium	**TFA**	Trifluoracetic acid
EBV	Epstein-Bahr Virus	**TOF**	Time-of-flight
EDTA	Ethylenediaminetetraacetic acid	**wt**	Wild type
ER	Endoplasmic reticulum		
FCS	Fetal Calf Serum		
FITC	Fluorescein		
Gy	Grey		
h	Hour		
HEPES	*4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid*		
HF	High fidelity		
HLA	Human Leukocyte Antigen		
Ii	Invariant Chain		
LB	Luria-Bertani medium		
MALDI	Matrix-assisted laser desorption		
MHC	Major Histocompatibility Complex		
MQ	Milli-Q		
min.	Minute		
NEB	New England Biolabs		
NP-40	Nonyl-phenoxylpolyethoxylethanol		
ON	Over night		
ONC	Over night culture		
PBS	Phosphate buffered saline		
PE	Phycoerythrin		

Table of contents

1. Introduction

1.1 The immune system

The human immune system is a complex and intricate network of cells, organs, tissues and their products. Its purpose is to protect the body from invasions of pathogenic organisms such as bacteria, helemites, protozoa, fungi and virus. The system also protects the body from malignant growths and neoplasia in the form of cancer [2].

The leukocytes and lymphocytes of the immune system originate from the hematopoetic stem cells in the bone marrow, from which they migrate through the lymphatic system to the peripheral lymph nodes and on to the rest of the body[3]. Two major progenitor lines give rise to the leukocytes of the immune system, the myeloid progenitor and the common lymphoid progenitor.

The myeloid lineage contains most of the cells of the innate immune system, the macrophages, dendritic cells, mast cells, eosinophils, neutrophils and basophils. The common lymphoid lineage includes all the lymphocytes of the adaptive immune system, as well as the natural killer cells (NK-cells)[3].

The immune system itself is characterized by its ability to respond to antigenic (antigen from *anti*body *gen*erating) substances. Dependent on the type of antigen, the immune response can take multiple forms, involving all parts of the immune system, but the ultimate goal of such a response is generally to protect the body from foreign substances. The first part of the immune system to respond is generally the innate part of the system. The first response is generally at the barriers of skin and mucus, then by cells responding to the antigen deeper in the tissues. All these responses are non-specific. The cells of the innate immune system rely on germ-line encoded pattern-recognition receptors to distinguish self from non-self, and they are more often than not sufficient to remove the pathogen and antigenic substance from the body[2].

Throughout an innate immune response, pathogens and antigens will be taken up by antigen presenting cells of the immune system and transported to draining lymph nodes for presentation to the cells of the adaptive immune system.

1.2 The adaptive immune system

The adaptive immune system is a very specific and efficient system of protection against pathogens. While slower to react to new infections than the innate immune system, it is able to adapt to specific pathogens and respond with very specific antibodies and receptors. It is also able to induce immunity, so that a reinfection of the same pathogen will induce an immediate adaptive immune response through immunological memory[3,4]. The adaptive immune system can be divided into humoral and cell-mediated immunity. Humoral immunity consists of antibodies and the B-cells that create them. Cellular immunity covers the CD8+ T-cells and the CD4+ T-cells[2]. T-cells develop in the thymus, where they go through positive and negative selection for reactions to self-antigens by binding to the MHC-molecules of the thymus epithelial cells[5].

The APCs of the innate and adaptive immune system transport antigens to the draining lymph nodes in close proximity to the infected tissue. There the antigens are presented to the T-cells circulating in the lymphatic system. Antigen specific T-cells are retained in the lymph node and are activated to become effector T-cells. B-cells that encounter antigens will take them up through B-cell receptor mediate endocytosis and present it on MHC class II on the cells surface. These cells will migrate to the T-cell area of the lymph node. The activated CD4+ T-cells and B-cells that recognize the same antigen will interact and the T-cell will activate the B-cell while the B-cell provides further survival signals to the T-cell through a variety of interactions[3,6]. The T-cell/B-cell interaction can be self-replicating from this point on, which will perpetuate the immune response outside the lymph node as well[7].

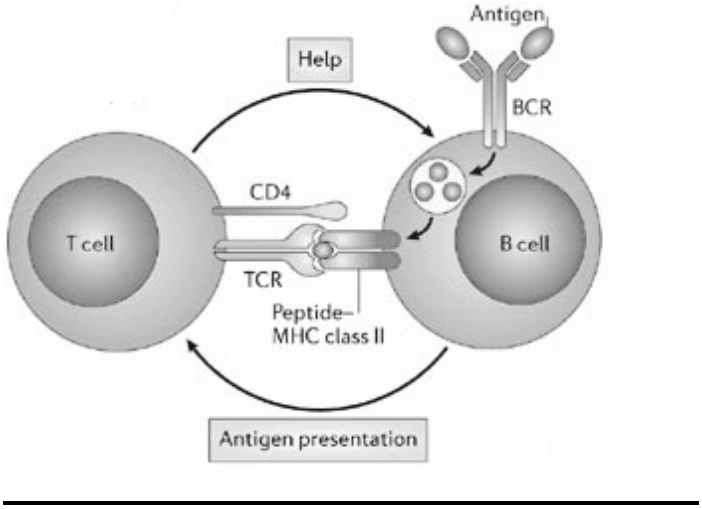


Figure 1: T-cell/B-cell interaction *with T-cell help and B-cell antigen presentation at the initiation of an adaptive immune response. Illustration from Edwards and Cambridge, 2004* [8]

Both the B- and T-cells will start to multiply and create effector cells. The B-cells will congregate into germinal centers and begin to produce a large amount of antibodies against the antigen. The effector T-cells will migrate to the infection site, and will start recruiting more innate immune cells to start combating the infection, including modifying the local cytokine environment, increasing blood flow and destroying aberrant or infected cells[2,3,9]. Once the infection has been eradicated, long lived memory B-cells will migrate to the bone marrow and sustain specific immunity towards the antigen for many years. A second infection will therefore illicit a very swift and specific adaptive immune response facilitated by memory B- and T-cells that retain immunological memory of the antigens[4].

1.3 Antigen presenting cells

To become active and initiate an adaptive immune response, the naive T-cells need to be presented with their specific antigen on the surface of a cell. This is done by use of a Major Histocompatibility Complex molecule (MHC). All normal nucleated cells in the body express MHC class I proteins on their surface, to present peptides from the cells interior to migrating T-cells. The MHC class I molecule is associated with the CD8 co-receptor, so that it is the CD8+ T-cells that respond to peptides presented on these molecules. This ensures that CD8+ T-cells, who are specialized in killing cells infected with intracellular pathogens or are cancerous, are activated to attack the infected cells[10].

MHC class II molecules are the focus of this thesis, and will be explored more in the next section. They are present mostly on professional antigen presenting cells (APCs). These cells are specially equipped to take up and process extra cellular antigens and present them to CD4+ T-cells. These T-cells give activation help to B-cells in the lymph nodes, produce a variety of cytokines and control the development of immune responses by developing into specific subclasses[11]. Professional antigen presenting cells consists of dendritic cells, macrophages and B-cells. Of these, dendritic cells are by far the most effective, but all of them are capable of taking up and presenting antigens[12].

Dendritic cells are divided into three major lineages: follicular, plasmacytoid and myeloid dendritic cells. These designations are given from morphology, location and function. The follicular cells reside in the lymph nodes, where they seem to work as an antigen deposit for activating and stimulating T- and B-cells. The plasmacytoid cells wander the blood and lymph, while the myeloid cells migrate to the tissues of the body and reside there until they pick up antigen and relocate to draining lymph nodes, where they participate in priming the T-cells. Dendritic cells have large amounts of MHC class 2 molecules on their surface and are very efficient at T-cell priming and activation. They all perform pinocytosis, phagocytosis and receptor

mediated endocytose to engulf antigens. The engulfed particles are broken down to peptides and presented at MHC class II molecules on the dendritic cell surface. Myeloid dendritic cells are often associated with a specific tissue, and immune system activations can initially be tissue specific, localized events. Generally, it is the tissue specific myeloid dendritic cells that migrate to the lymph nodes and present their antigens to T-cells to initiate the adaptive immune response[13,14].

Macrophages are an integral part of the innate immune system, and are effective APCs. Their main task is to break down dead cells, pathogens or other refuse in the tissues through phagocytosis. Most of the peptides from the phagosomes will be presented on the surface of the tissue macrophages, but depending on the local cytokine environment and signals the macrophage receive, this may either prevent or escalate inflammatory reactions. Macrophages are also heavily involved in the development of organs and the repair of tissue damage, which prevents a limiting classification of the macrophage as an immune cell only[14,15].

B-cells are APCs, but have a significantly lower amount of MHC class 2 molecules and co-receptors on their surface than dendritic cells. This makes them far less effective as APCs, but B-cells have many other functions, most notably the production of antibodies. B-cells takes up antigens from the environment both by phagocytosis, pinocytosis and receptor mediated endocytosis. Of these, the receptor mediated endocytosis is thought to be the most important. The B-cell receptor is identical to a membrane bound antibody, and receptor mediated endocytosis through the BCR will be specific for that antigen. The BCR mediated endocytosis leads directly to the MHC class II loading compartments, facilitating presentation of the B-cells antigen on its surface. B-cells are able to take up and concentrate small amounts of antigen transported to the lymph node or in the blood this way[16]. The BCRs affinity for the antigen is directly proportional to the B-cells ability to induce CD4+ T-cell proliferation during B-cell/T-cell interaction[17].

1.4 MHC class II molecules

The Major Histocompatibility Complex (MHC) molecules determine much of the immune systems capabilities. The name refers to the molecules being encoded in the Major Histocompatibility Complex region of chromosome 6 in the genome. In humans this region is also called HLA, for Human Leukocyte Antigen. Both terms are used in this thesis when talking of the MHC in humans. The MHC region contain several different MHC class I and II genes, with multiple variants of each gene common in the population, making the MHC region both polygenic and polymorphic[3,18]. Each individual express their MHC genes as haplotypes, a combination of alleles encoded on the same chromosome. MHC class I is encoded in 4 genes, HLA-A, B and C coding for the alpha chain and one gene coding for the β-chain microglobulin. MHC class II is encoded in 7 genes: DR, DQ and DP genes for both an alpha and a beta chain, with DR having 2 beta chain genes instead of one. This means that a typical human can produce 6 different MHC class I and 8 MHC class II molecules[2,18].

MHC class 1 molecules are present on all cells in the body with a few exceptions. They present endogenous peptides produced in the cell to migrating T-cells that can recognize the non-self peptides produced by intracellular parasites, virus and cancerous cells[10].

MHC class II molecules are the primarily focus of this thesis. Unlike MHC class I, MHC class II molecules are present only on antigen presenting cells and some epithelial cells. The proteins is used to present extra cellular molecules taken up by the APC ,through phagocytosis, endocytosis or pinocytosis, on the APC cell surface[19]. The MHC class II protein is made up of an α-and a β- chain in an non-covalent complex, each contributing to the peptide binding groove. Each chain has a membrane distal and a membrane proximal domain, with the membrane distal domains making up the binding groove. The membrane proximal domain is transmembrane, anchoring the MHC protein to the cell surface

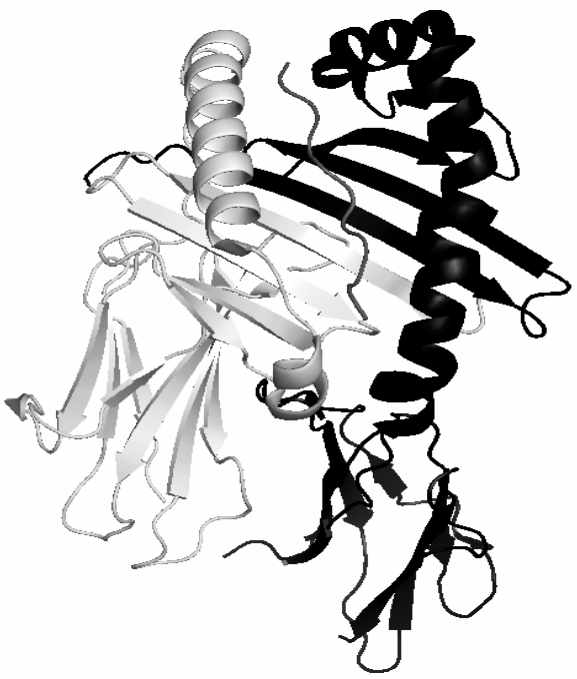


Figure 2: MHC class II *ribbon model created in PyMOL (DeLano Scientific), showing HLA-DQ2.5 with α-chain (black), β-chain (white) and gliadin bound in binding groove (grey).*

MHC class II molecules are synthesized in the rough ER and form a transitory aggregate with the chaperone proteins BiP, calnexin, Erp72 and invariant chain. These help fold and assemble the MHC class II complex correctly. Invariant chain (Ii) is a non-polymorphic transmembrane glycoprotein, but it may vary in length. There are 4 common isoforms of Ii expressed in humans, called p33, p35, p41 and p43, created by variation in initiation sites and differential splicing during Ii translation. The MHC class II complex will collapse without a peptide bound in the peptide binding groove, so the Ii binds to the binding groove of the MHC class II during formation. This also prevents binding of endogenous intracellular molecules to the groove in the ER and Golgi (**figure 3)**. Ii incorporates a variety of sorting signals at the N-terminal end that directs the MHC class II molecule into the MHC class II loading compartments (MIIC), either from the E.R or after internalization from the cell surface[20,21,22,23](**figure 4)**.

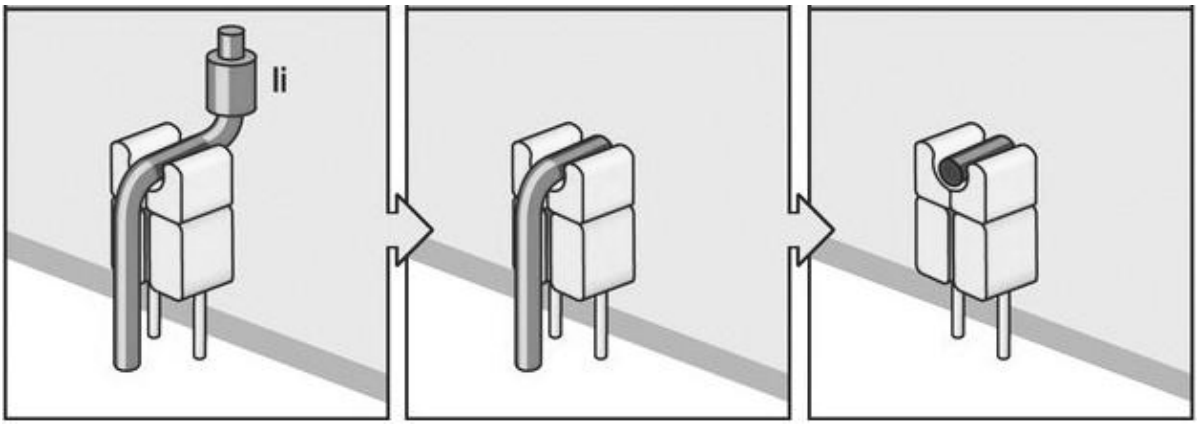


***Figure 3**: **Schematic drawing of the invariant chain interacting with MHC class II**, blocking the binding groove. Ii(light grey), CLIP (dark grey) and MHC class II (white) showing the progressive degradation of Ii as the MHC class II reaches the MIIC compartment. Picture from Immunobiology, 6ed.l 2005[3].*

Once the MHC class II molecule enters the MIIC, Ii and the other chaperones are degraded, leaving only the class II-associated invariant chain peptide (CLIP) blocking the peptide binding site. This peptide is the centre of the invariant chain and the part blocking the binding groove. As with the original invariant chain, the CLIP peptide can vary slightly in length and residue make up, but is in principle non-polymorphic. The CLIP peptide prevents the binding of other peptides to the MHC class II molecule until the CLIP is removed. The removal is accomplished by a third MHC molecule, which in humans is called HLA-DM. HLA-DM is a protein with the same structure as the MHC class II molecule, but it lacks the peptide binding groove. It contains a MHC class II interaction area that interacts with an area in the alpha chain of the MHC molecule. This interaction catalyzes the release of the CLIP peptide from the MHC class II molecule and loading of another peptide into the groove. HLA-DM also catalyze removal of weakly binding peptides to the groove and the loading of strong binders, thus ensuring that only peptides which bind strongly to the MHC class II protein in question can be presented on the cell surface. The deactivation of HLA-DM causes accumulation of CLIP on the MHC class II both in MIIC and on the cell surface[24,25,26].

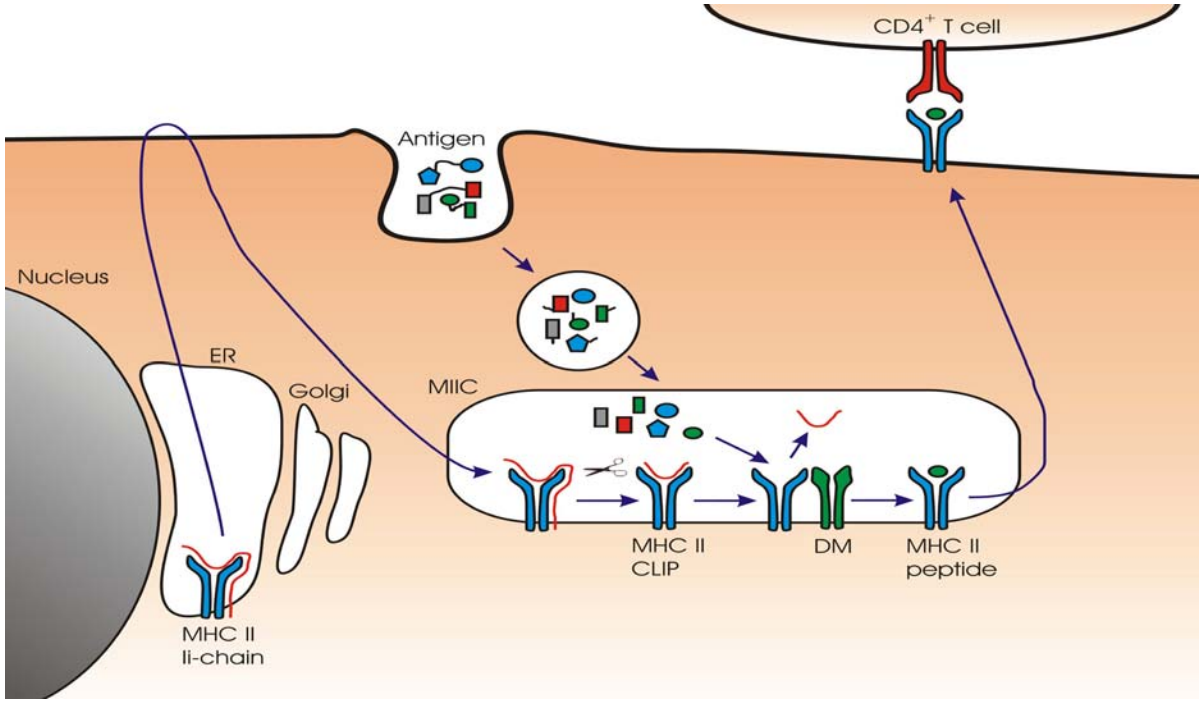


Figure 4: Simplified scheme of MHC class II transport to MIIC and presentation of peptide on the cell surface, created by Lars-Egil Fallang.

In addition to HLA-DM, which all APCs contain, B-cells contain HLA-DO. This non-classical MHC molecule catalyses the deactivation of HLA-DM at high pH by binding to it and preventing interaction with MHC class II[16]. This is thought to facilitate presentation of antigens absorbed through the BCR by the following mechanism. Antigens binding to the BCR moves to the MIIC, but do not detach from the BCR until the pH becomes low because of the stable interaction between the BCR and its antigen. Since HLA-DO prevents CLIP detachment at higher pH, the BCR-bound antigens will be favored for presentation on the B-cell surface[16,26].

1.5 MHC class II peptide binding

The MHC class II binding groove is made up of a “floor” of β-sheets and “walls” of α-helices. The binding groove is made up of residues from the membrane distal domain of both the α- and β- chain, so that the binding properties of the MHC class II molecule is influenced by both chains (**figure 5**). The binding groove is not closed of in each end like the MHC class I, and can accommodate larger peptides and utilize other anchor residues than the class I molecule[27].

Peptides bind to MHC class II in an extended conformation. A pattern of hydrogen bonds between the peptide backbone and the conserved amino acid side chains extend along the length of the binding groove, and this pattern is remarkably conserved between all alleles of the MHCs. Since the hydrogen bonds do not involve the side chains of the bound peptide, the interaction is sequence independent and so explains how a single class II molecule can bind a large number of different peptide sequences[20]. Of the conserved residues responsible for these hydrogen bonds, three are asparagines located at α62, α69 and β82 in positions well suited to make hydrogen bonds to the carbonyl and amide groups of the peptide backbone. Other conserved residues include histidine β81 and tryptophane β61, which are positioned to make hydrogen bonds with carbonyls at the P-1 and P-8 positions[20,28]. This is by no means an exhaustive list. There are many such residues able to form hydrogen bonds to the right peptide in the groove.

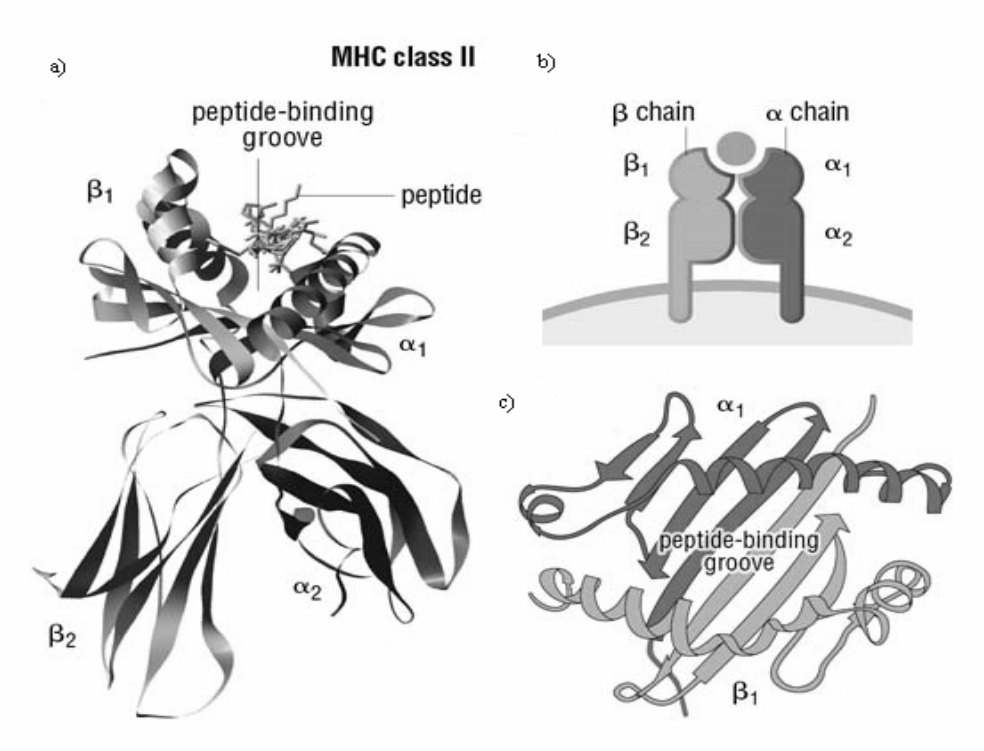


Figure 5: MHC class II *showing a) protein structure, b) schematic drawing, c) binding groove structure. Picture from DeFranco et.al 2007* [27]

The binding groove is lined with polymorphic amino-acids that form a series of pockets. The side chains of the peptide binding to the MHC class II molecule extend into these pockets. Interactions and bonds between the residue forming the pocket and the peptide side-chains increase the specificity of the binding groove. In addition, a substantial amount of surface area can be buried in these pockets and will then be shielded from the solvent, which also increases the specificity of the interaction between the MHC groove and the peptide. The side-chains in the peptide binding pockets differ between the MHC class II alleles, and this explains the difference in peptide binding specificity between them. The pockets are named after the side chains they accept, starting with P1 (see figure 8). In practice, most peptides bind only to the side chains of a few key anchor residues, but these may vary from allele to allele. The most commonly used are P1 and P9, followed by P4 and P6 (**figure 6**). Each allele has a binding motif, and what peptide will bind to the groove can be dependent on only a few critical residues[20,29]. Together, these two mechanisms are thought to decide the binding of peptides in the MHC class II binding groove.

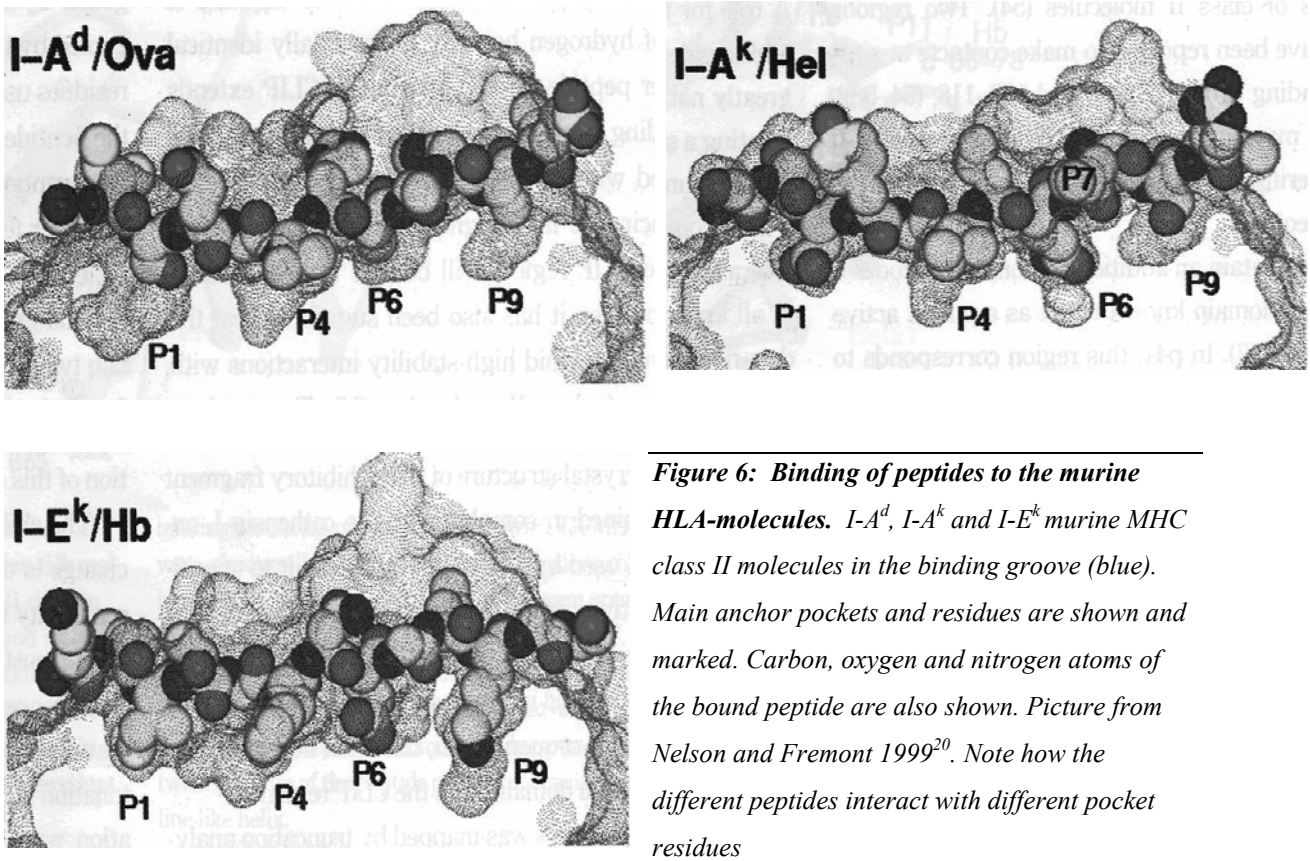


Figure 6: Binding of peptides to the murine HLA-molecules. *I-A^d, I-A^k and I-E^k murine MHC class II molecules in the binding groove (blue). Main anchor pockets and residues are shown and marked. Carbon, oxygen and nitrogen atoms of the bound peptide are also shown. Picture from Nelson and Fremont 1999[20]. Note how the different peptides interact with different pocket residues*

1.6 Celiac disease

Celiac disease is a chronic inflammatory disorder of the intestine causing malabsorption of nutrients and potentially increased mortality and morbidity in afflicted patients. It is an acquired disorder that can develop any time, from childhood to middle age. Celiac disease is classified as an autoimmune disease, but it has a multifactoral cause in the complex interplay of genetic and environmental factors that contribute to the pathogenesis. There is no doubt that celiac disease develops by intolerance and inappropriate immune response to ingested wheat gluten. Gluten is a complex mixture of gliadin and glutenin polypeptides. Of these, gliadins are monomers and are thought to be the main environmental factor triggering the autoimmunity. Glutenins are larger polymeric structures, and believed to be less involved in the development of celiac disease [30].

Celiac disease is an uncommonly good model for chronic inflammatory and autoimmune diseases, since both the environmental factor and the disease associated HLA-genes are know. In addition, gluten is relatively easy to add or withhold in an experimental system, and the affected organ is easily accessible for biopsy and study.

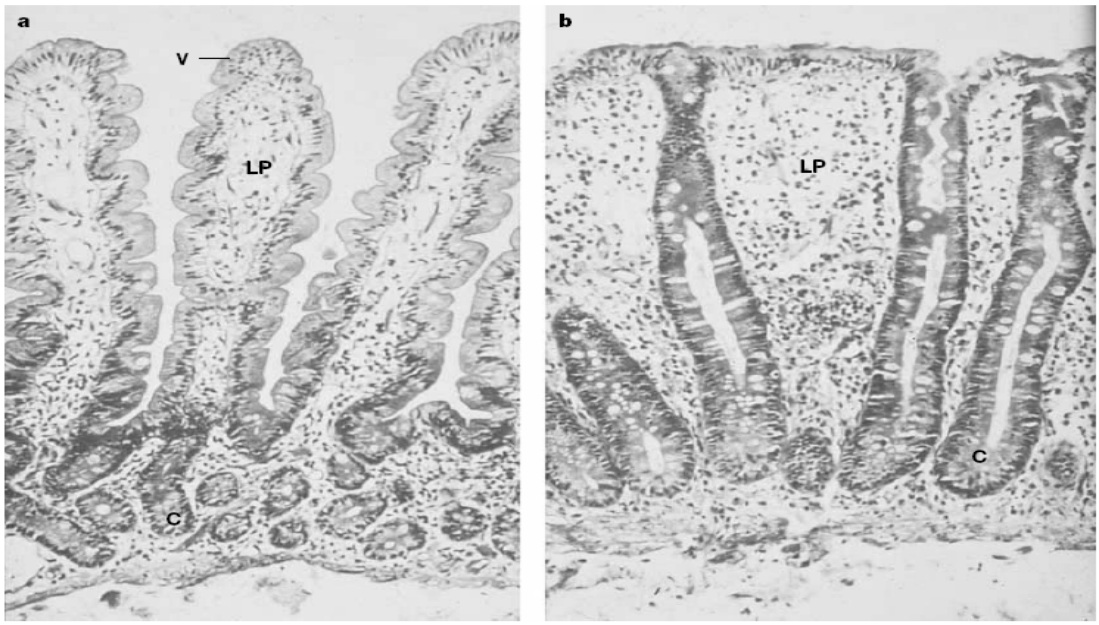


Figure 7: Gut epithelial histology *from a) normal and b) celiac small intestine. Histology shows the villi (v) and crypts. In b) the lamina propria (LP) has swelled out, the villi has atrophied and there is an extensive lymphocyte infiltration evident (black dots). Picture from Sollid 2002* [30].

Celiac disease is a genetic disorder, with a 75% concordance rate between twins[31]. HLA-linked genes contribute a large amount to the genetic susceptibility, but dependent on the model used for the disease non-HLA linked genes may play a larger role. In this thesis, the concern is primarily with the HLA-linked genes. The HLA association in most celiac patients is with DQ2 (DQA1*05/DQB1*02), with a minority associated with DQ8 (DQA1*0301/DQB1*0302) instead. Most patients have the DR3-DQ2 haplotype (DRB1*0301-DQA1*0501-DQB1*0201) or are DR5-DQ7 and DR7-DQ2 heterozygote (DRB1*11/12-DQa1*0505-DQB1*0501 and DRB*07-DQA1*0201-DQB1*0202) [32](**figure 8**). These DQ molecules are also associated with other autoimmune diseases, for instance insulin dependent diabetes mellitus[33], which establishes a broader basis for the study of these MHC class II haplotypes in particular and MHC class II molecular binding mechanisms in general.

The HLA-DQ2 and DQ8 molecules are assumed to present gluten peptides to T-cells in the gut, explaining their association with the disease. Both the HLAs have preference for negatively charged peptides in the peptide backbone, but gluten peptides contain relatively few negative charges. This conundrum is solved by tissue transglutaminase 2 (TG2), a deaminating enzyme that is also the predominant auto-antigen in celiac disease. The enzyme catalyzes transamidation of peptides, and is also the prevalent autoantigen in celiac patients[34,35].This way the enzyme introduces negative charges into the gluten peptides *in situ* and makes them adequate binders to DQ2/DQ8 HLA. The presence of proline at crucial areas of the peptide normally hinders binding to HLAs, with the exception of promoting PPII helix structures. DQ2 and DQ8 allow for proline more readily than most other HLA molecules and are favored for binding proline rich peptides compared to other HLAs. It is known that gluten peptides are often unusually proline rich[36], and it is postulated that the lack of a residue in position α53 is the cause of the preference for proline in the position normally making a hydrogen bond to that residue.

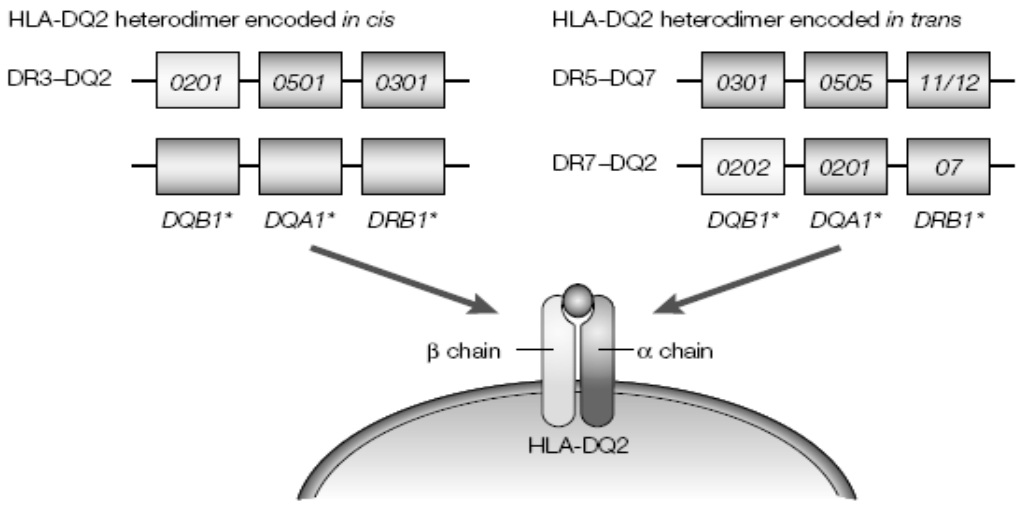


Figure 8: HLA-DQ2.5 germ line encoding. HLA-DQ2.5 *can be encoded in 3 different haplotypes, either in cis or trans, combining to make the complete HLA-DQ2.5 protein after translation. As can be seen, A05 and B02 can be combined to a molecule in both haplotypes. Figure from Sollid 2002* [30]

DQ2 has two major variants, HLA-DQ2.2 and HLA-DQ2.5. Only HLA.DQ2.5 is associated with celiac disease to any degree, with HLA-DQ2.2 only being present in a negligible amount of cases. Their β-chain differs in only one amino acid in the

membrane proximal region and has no influence on peptide binding. The alpha-chains (DQA1*0202 and DQA01*0205) have a 10 polymorphic residues in the membrane distal domain, as well as 6 polymorphisms in the membrane proximal domain. 4 of the membrane distal polymorphisms are located around the antigen binding area, while the remaining 6 may be located in the presumed HLA-DM interaction area of the MHC class II molecules [25] (**figure 9**). Only the membrane distal polymorphisms interact with the bound peptide, and these are therefore the major points of interest when considering the difference between the two alleles.

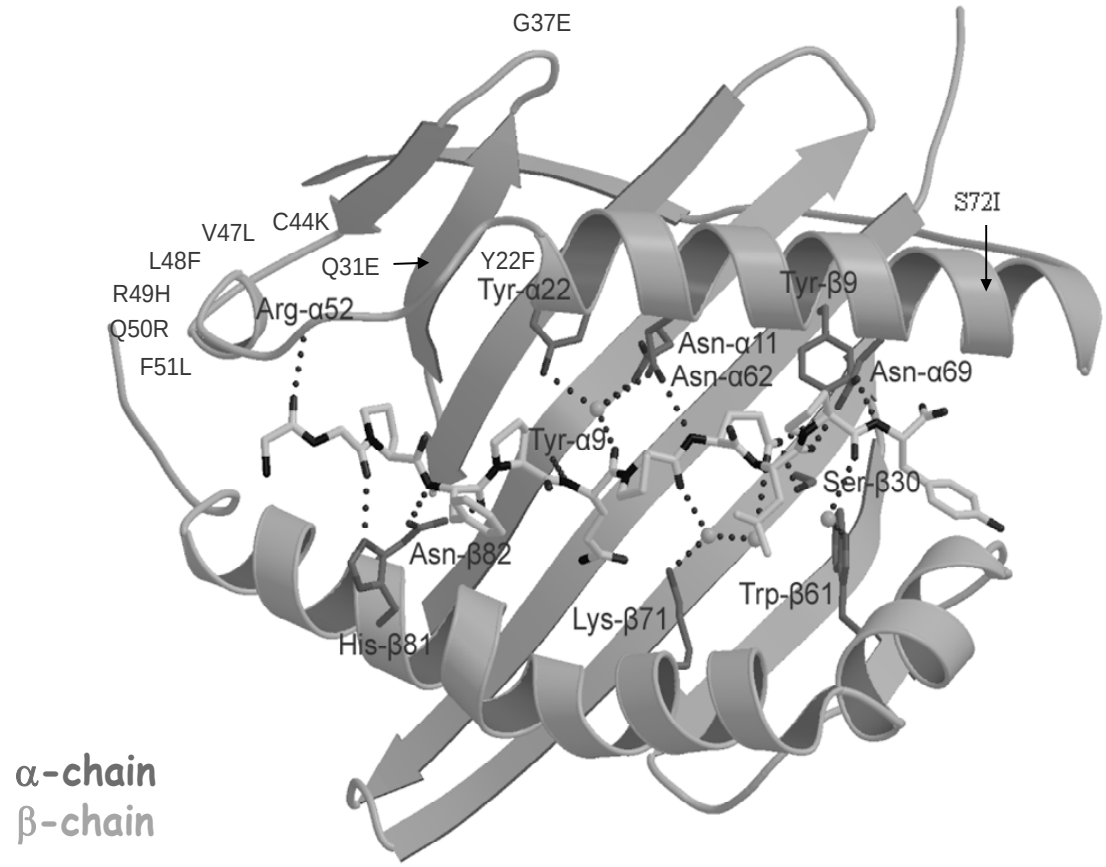


***Figure 9: HLA-DQ2.5 polymorphic sites**. HLA-DQ2.5 is shown with the peptide backbone in the binding groove and hydrogen bonds in dashed lines. Target polymorphisms from HLA-DQ2.5 to HLA-DQ2.2 are marked. Figure modified from Kim.et.al[37].*

It has been shown that DQ2.5 has fairly unusual antigen binding properties, retaining large amounts of CLIP on the cell surface. HLA-DQ2.5 has been shown to be a poor substrate for HLA-DM[38], and as 6 of the polymorphisms are located in the presumed

HLA-DM interaction region, that might be part of the answer. Early results from studies done at the Sollid group indicated that the polymorphisms around the binding groove are more important, especially since it has been indicated that HLA-DQ2.2 is also an equally poor DM substrate[1]. The prime target for finding the amino acids responsible for these differences in the binding properties of the DQ2.2 and DQ2.5 are therefore the 4 variable amino acids clustered around the antigen binding groove more than the 6 amino acids clustered in the HLA-DM interaction area.

Table 1: CLIP-peptides commonly eluted from HLA-DQ2.5[38]

CLIP peptide	***Sequence***	***Length***	***Mass***
Ii81–104	LPKPPKPVSKMRMATPLLMQALPM	24	2674. 5
Ii81–103	LPKPPKPVSKMRMATPLLMQALP	23	2543.4
Ii82–103	PKPPKPVSKMRMATPLLMQALP	22	2430.4
Ii81–101	LPKPPKPVSKMRMATPLLMQA	21	2333.3
Ii93–109	MATPLLMQALPMGALPQ	17	1781. 9
Ii93–108	MATPLLMQALPMGALP	16	1653. 9

1.7 Aim of thesis

The aim of this thesis is to investigate which of the 10 membrane distal polymorphisms in HLA-DQ2.5 and HLA-DQ2.2 may explain the differences seen in CLIP presentation by the two molecules. The results may also help explain part of the celiac disease association differences between them. Site directed mutagenesis of the HLA-DQ2 genes will be used to introduce one polymorphic residue from HLA-DQ2.2 into the corresponding site in HLA-DQ2.5. By eluting and analyzing the CLIP peptides presented by the mutant EBV-transformed B-cells on mutated HLA-DQ2.5 molecules, differences in CLIP presentation can be measured. A difference in the level of CLIP presentation may indicate differences in CLIP stability in the binding groove. If such a difference is observed, the reverse mutation will be attempted by introducing the polymorphic residue from HLA-DQ2.5 into HLA-DQ2.2. The opposite effect on the level CLIP presentation of the HLA-DQ2.2 mutant will confirm the importance of the residue in terms of CLIP stability in the binding groove of HLA-DQ2.

This thesis work is a part of a study by the Sollid group at the Centre for Immune Regulation, Oslo, and will focus only on 3 of the 10 polymorphic residues. The polymorphic residues in position α31, α37 and α72 were chosen as the ones to focus on in this thesis work.

2. Methods and materials

When nothing else is indicated, all solutions and methods mentioned below were done in accordance with the protocols found in Sambrook[39].

2.1 PCR and creation of the target gene with a transmembrane tail

Reagents

cDNA encoding HLA-DQA01*0201(from B-LCL 9047 PLH), HLA-DQA1*0501 and DQB1*0201(from B-LCL CD114), Phusion DNA polymerase (NEB), High Fidelity Buffer (NEB), 10 mM dNTP (NEB)

20X TBE: 216 g Trisma base (Sigma), 110 g boric acid (Merck), 18.6 g EDTA(triplex III) (Sigma), MQ water for total volume 1L

Agarose gel: 1% agarose (Merck), 50 ml 1X TBE, 5 µl ZyberSafe (Promega)

Primers

Table 2: PCR primers*, from the Celiac Group, Centre for Immune Regulation*

Primer 1	TGGGAAGCTTATGATCCTAAACAAAGCTCTG
Primer 2	GGATAAGCTTCACAAGGGCCCTTGGTGTC
Primer 3	CCAGCCCCTATGTCAGAGCTCACAGAG
TM-tail megaprimer	CCAGCCCCTATGTCAGAGCTCACAGAGACTGTGGT CTGCGCCCTGGGATTGTCTGTGGGCCTCGTGGGCAT TGTGGTGGGCACTGTCTTCATCATCCGAGGCCTGCG TTCAGTTGGTGCTTCCAGACACCAAGGGCCCTTGTGTC

The HLA-DQA1*0501 and HLA-DQA1*0201 genes studied in this thesis were earlier created by cDNA synthesis from a B-cells line graciously given to the Sollid group by B.Roep. The plasmids created with this cDNA contain the HLA-genes modified with a leucin-zipper soluble tail, and this must be replaced with the native trans-membrane tail before mutagenesis and retroviral transduction can be done. The TM-tail was already encoded in a pLHCX a different construct. The encoded trans-membrane domain needed to be lifted out of the construct by use of PCR (Polymerase Chain reaction), *in vitro* amplification of DNA through use of a heat-stable DNA polymerase and a programmable heat-block as described by Mullis et.al.[40] and by Sambrook et.al[39]. All PCRs were done on a Veriti 96-well Thermal Cycler (Applied Biosytems). The sequence of all primers used is shown in table 2.

PCR mix: 5 µl HF buffer, 15.5 µl MQ water, 1 µl construct pLHCX construct with TM-tail, 10 nM primer 2, 10 nM primer 3, 1 µl dNTP, 0.5 µl Phusion enzyme.

PCR-program: Hold (1): 98° -30 s., Melt (2): 98° - 10 s, Anneal (3): 57° -10 s, Elongation (4): 72° - 10 s, End (5): 72° - 5 min, 30 cycles

The resulting 145bp DNA fragment was purified on a 1% agarose gel, then used as a megaprimer for the PCR of the target gene (HLA-DQA1*0501/HLA-DQA1*0201), the forward primer being primer 1. These primers introduced HindIII restriction sites in the gene. PCR-program: Hold (1): 98° -30 s., Melt (2): 98° - 10 s, Anneal (3): 57° - 15 s, Elongation (4): 72° - 20 s, End (5): 72° - 5 min, 30 cycles.

The PCR-product was treated with HindIII restriction enzyme and ligated into the pMOS mutagenesis vector in accordance with protocol described in 3.5.

2.2 Ligation of gene into cloning- and expression vector

Reagents

pLHCX expression vector (Clonetech), pMOS Blue cloning vector (GE Healthcare Life Sciences, cDNA encoding HLA-DQA01*0201(from B-LCL 9047 PLH), HLA-DQA1*0501 and DQB1*0201(from B-LCL CD114), T4 ligase (NEB), T4 ligase buffer (NEB).

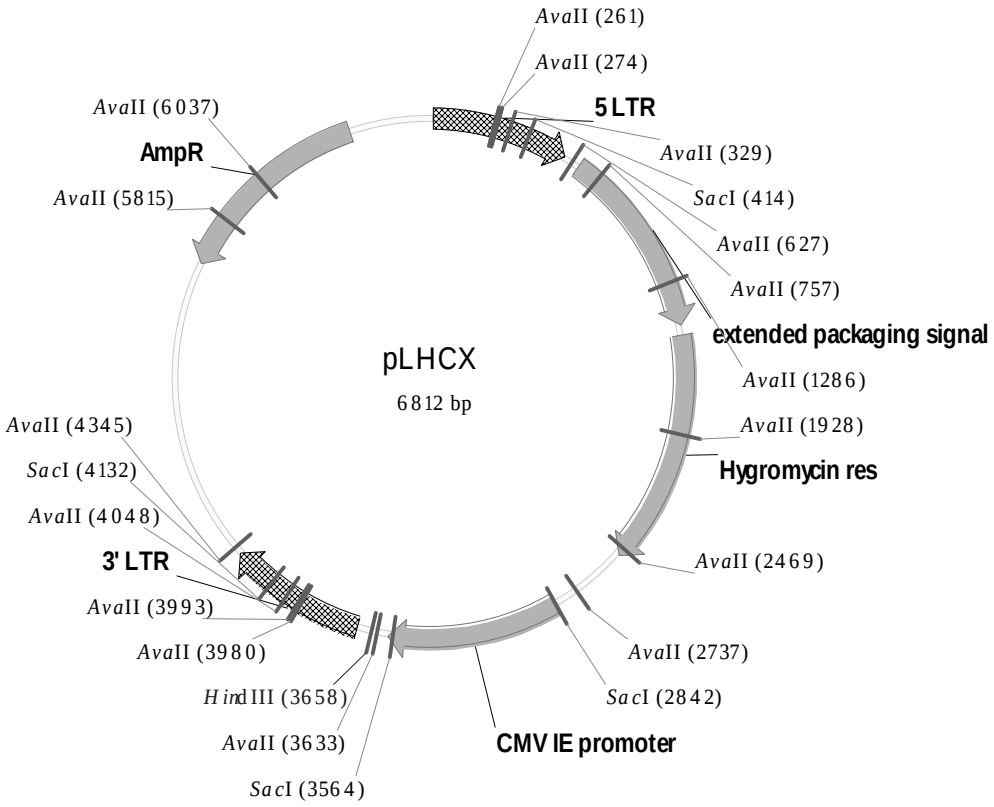


Figure 10: pLHCX expression vector (Vector NTI)

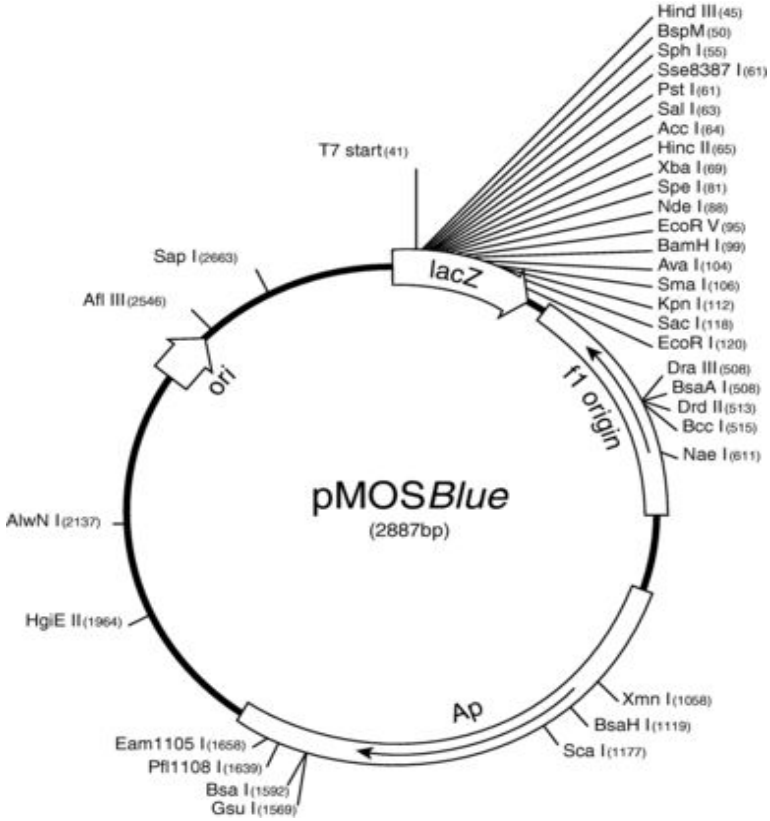


Figure 11: pMOS-blue cloning vector, *multiple cloning site shown in the lacZ region. Ampicilin resistance shown as Ap.*

T4 ligase(NEB) was used for ligation of the target genes into cloning- and expression vectors, using the commercial available protocol[41]. The target genes, produced with TM-tail as described in 3.1 and target vectors were cut with HindIII, purified on an agarose gel (3.4) and ligated with the target gene as described in the NEB protocol.

2.3 TOPO-cloning of target gene

Reagents

TOPO-cloning kit (Invitrogen)

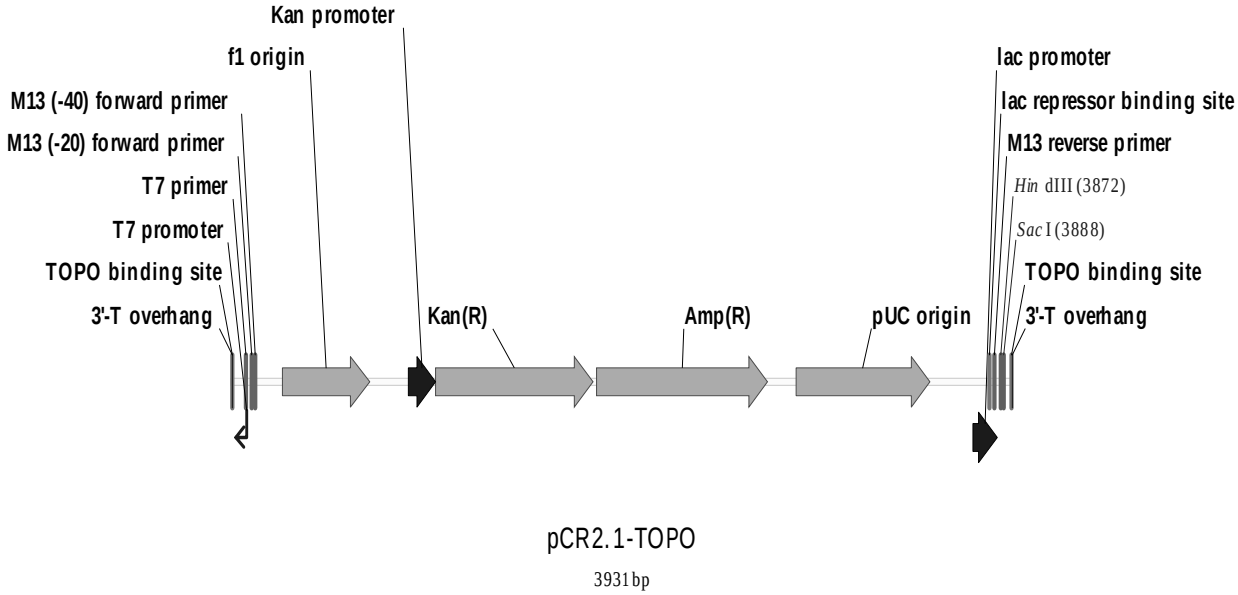


Figure 12: pCR2.1-TOPO cloning vector *(linear). The target gene is inserted between the two terminal 3'T overhangs, creating a circular plasmid.*

TOPO-cloning was done as described in the commercial TOPO-cloning kit protocol(Invitrogen)[42]. This is a commonly used TOPO cloning kit, containing a linearized plasmid (pCR2.1) coupled with topoisomerase for swift ligation of an adenine flanked DNA fragment into the linearized plasmid. TOPO cloning was done to create the pCR2.1 plasmid with HLA-DQA1*0501 since some of the HLA-DQA05 mutagenesis primers, notably primer 5 (see table 3), seemed at one point to anneal at non-target sites in the pMOS plasmid, creating unspecific mutations.

2.4 Transformation of XL-1 bacteria

Reagents

SOC-media: 950 ml MQ water, 20 g bacto-trypton (Merck), 5 g bacto-yeast extract (Merck) 0.5 g NaCl (Merck), 250 mM KCL (Sigma), 5 ml 2 M $MgCl_2$ (Sigma), 20 ml 1 M glucose (Merck).

LB-media: 950 ml MQ water, 10 g Bacto-trypton (Merck), 5 g Bacto-yeast extract (Merck), 10 g NaCl (Merck)

LB-agar: 1 L LB-media, 15 g Bacto-agar (Bectin, Dickinson and Company)

80µl XL-1 E.coli bacteria was thawed on ice and approx. 300 ng of the vector plasmid was added, then incubated 30 min. on ice, 45 s at 42°C and a further 2 min.on ice. 120 µl SOC media was added to the bacteria and the mix was incubated with shaking at 37°C for 1h. Streaked out on a LB-agar plate with ampicillin (50 µg/ml) and incubated at 37°C ON.

Overnight cultures were created by picking a single colony and suspending it in 3ml LB media/ampicillin (50 µg/ml). The cultures were placed at 37°C in an incubator ON.

2.5 DNA-isolation and gel electrophoresis

Reagents

Wizard Miniprep Kit (Promega), Plasmid Midikit (Qiagen), Qiaquick Gel Extraction kit (Qiagen), NEBuffer 1-4 (NEB), HindIII(NEB), SacI(NEB), BglII(NEB), AvaI(NEB), 1x Loading Buffer (Promega), Isopropanol (Arcus Kjemi), 100% Ethanol (Arcus Kjemi)

20X TBE: 216 g Trisma base (Sigma), 110 g boric acid (Merck), 18.6 g EDTA(triplex III), MQ water for total volume 1L

Agarose gel: 1% agarose (Merck), 50 ml 1X TBE, 5 µl ZyberSafe (Promega)

Wizard Miniprep (Promega)[43] and Plasmid Midikit (Qiagen)[44] were used to isolate the DNA from the transformed bacteria grown in ONCs, as described in their respective commercial protocols. The miniprep. kit is a standard kit used by a large amount of research laboratories at the University of Oslo, regularly giving a DNA concentration of 50-150 ng/µl when eluted according to protocol. To control the size and identity of plasmids, and to isolate particular genes from a plasmid, specific restriction enzymes were used to cut out the gene and/or cut the plasmid into linear pieces and analyze them on an agarose gel as described in the commercial protocol (NEB)[45]. Hind III and SacI were used for such control of the pMOS and pLHCX vectors with DQA gene inserts.

50 ml 1% agarose TBE/5µl ZyberSafe gels were used for gel electrophoresis in a plastic electrophoresis chamber. DNA was applied with 1x Loading Buffer (Promega). Standard running time was 40 min at 80V in TBE-buffer. An UV-gel-dock was used to record and identify bands for analysis or DNA gel extraction described. To remove the DNA from the gel after electrophoresis, the Qiaquick Gel extraction kit (Promega) was used, as described in the commercial protocol[46].

2.6 Site-directed mutagenesis

Reagents

Quikchange Multi-site Mutagenesis kit (Stratagene)

Primers

Table 3: Mutagenesis primers*, synthesized at Eurogentech S.A. All mutagenesis primers are phosphorylated at the 5'end. Each primer introduces a specific mutation by a single nucleotide exchange, the target nucleotide is marked by a frame. The mutations has been named in the table as follows: Original residue, residue position, mutated residue, followed by which DQ2 variant the mutation is to be introduced in. Primer 4 is marked Sα72I, DQ2.5, and introduces the mutation serine to isoleucine in position 72 of the HLA-DQ2.5 alpha chain.*

Primer 4 (mutation Sα72I, DQ2.5)	ctaaaacataacttgaacattctgattaaacgctccaac
Primer 5 (mutation Qα31E, DQ2.5)	gatggagatgaggagttctacgtggacc
Primer 6 (mutation Gα37E,DQ2.5)	ctacgtggacctggagaggaaggagactg
Primer 7 (mutation Iα72S,DQ2.2)	ctaaaacataacttgaacagcctgattaaacgctccaac
Primer 8 (mutation Eα31Q,DQ2.2)	gatggagacgagcagttctatgtggacctg
Primer 9 (mutation Eα37G,DQ2.2)	ctatgtggacctgggggaggaaggagactg

Site-directed mutagenesis *in vitro* is a commonly used technique to introduce specific mutations in a target gene. Several approaches are possible, but many include cumbersome use of single-stranded DNA templates and subcloning in a M13-based bacterial vector. The Quikchange Multi-site Mutagenesis Kit (Stratagene)[47] avoids many of these problems, and was therefore chosen for the mutagenesis reaction. Using the Multi-site kit compared to a standard site-directed kit allows the kit to be used with only one mutagenesis primer for each mutation. Primers with the required mutations were designed by using an online tool to calculate Tm. at

(www.stratagene.com). The mutagenesis process is in principle a normal PCR with a primer including the mutated base(s). As template the pMOS-DQA105, pMOS-DQA10201 and pCR2.1 DQA105 plasmids created earlier were used. After the PCR, the PCR-products will incorporate the mutation, as seen below.

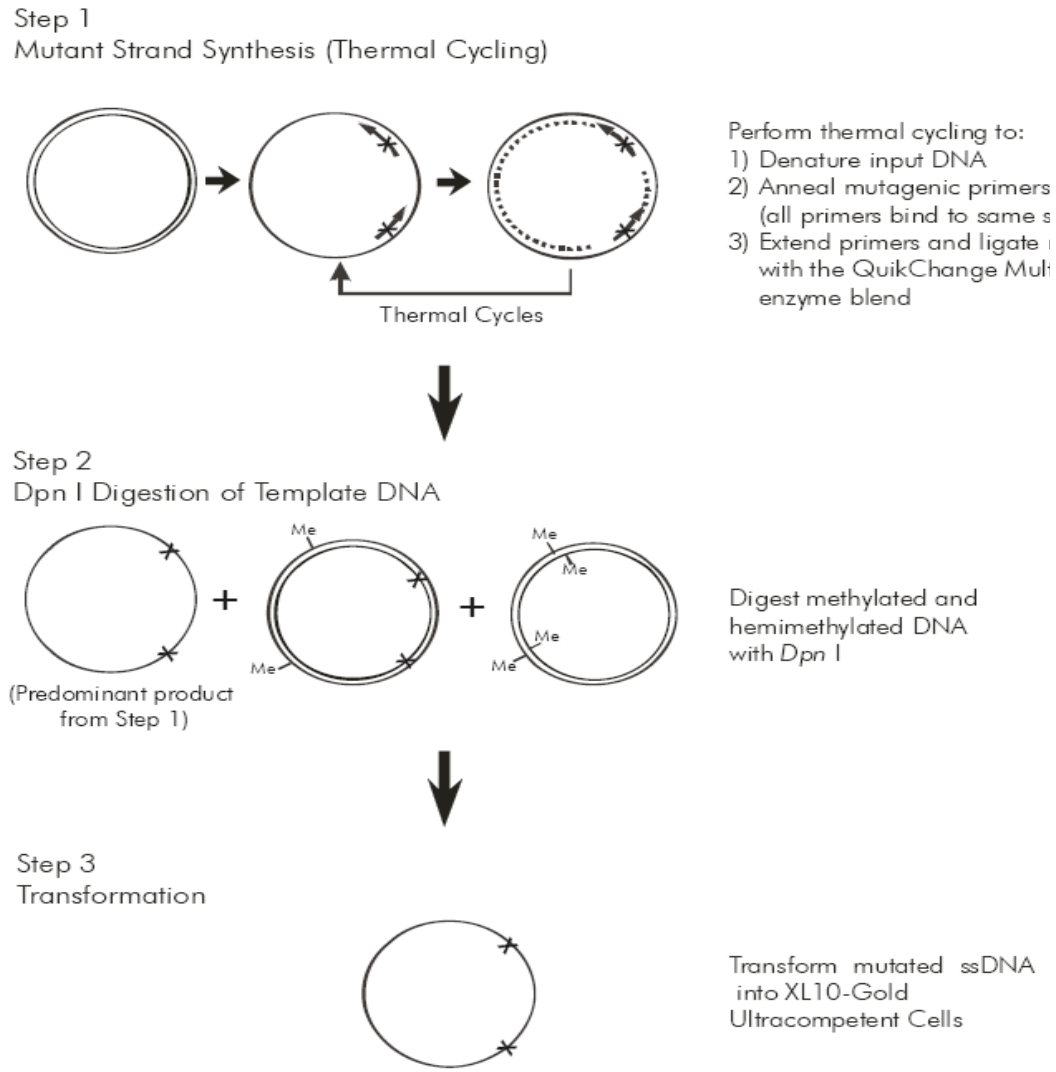


Figure 13: Site-directed mutagenesis, *illustration from the Stratagene Mutagenesis kit [47]. Step 1 shows how the mutagenesis primers anneal and extends the plasmid, creating a copied plasmid that includes the mutation. As the PCR reaction continues, all copies of the original plasmid will include the mutation. Step 2 shows Dpn1 restriction enzyme digestion that removes the non-replicated DNA, which is methylated and not mutated. Step 3 shows the finished plasmid with mutations included, ready for transformation into XL-1 bacterial cells.*

2.7 Sequencing

Sequencing of the mutations in cloning vectors and expression vectors were done at GATC Biotech in Cologne, Germany, according to their own protocols.

2.8 Cell maintenance

Reagents

721.82 EBV transformed B-cell line from B.Roep, GP2-293 cells (Clontech, #641530), DMEM with Na/Pyruvate(Invitrogen), dPBS (Invitrogen), PBS (Invitrogen), Fetal Calf Serum(FCS)(Cambrex), Penicillin (Panpharma), Streptomycin (X-gen Pharmaceuticals), Neomycin (Invitrogen), 1M 1X Hepes (Invitrogen), Trypsin 0,5% EDTA (Invitrogen)

RMPI complete media: 500 ml RMPI 1640(Invitrogen), 10% FCS (Cambrex), 1M 1X Hepes(Invitrogen),

All cell flasks and dishes were grown in a 37 °C, 5% CO_2 cell incubator if nothing else is noted.

The 721.82 B-cell line is EBV transformed to survive in suspension *in vitro*. The particular line used here has also been retroviraly transduced with a plasmid coding for the DQ2 β-chain and resistance against neomycin. A single tube containing approx. $5*10^6$ cells were thawed and suspended in 20 ml RPMI complete media and 1.0 μg/ml neuomycin. The cells were split when confluent, usually every third day, by exchanging 90% of the medium with fresh medium.

The GP2-293 cells are part of the Retroviral Universal Packaging System, batch 631530, from Clonetech. Liquid nitrogen stock was thawed and transferred to T-75 flask (NUNC) with 20 ml DMEM (Invitrogen) and 10% FCS. The cells are adherent

and became confluent after about 72 hours. The cells were split 1/12 by removing the old media, washing with 5 ml PBS, then adding 3 ml trypsin to detach the cells. The cells were suspended and centrifuged at 800rpm for 5 min. The cells were resuspended again in 11 ml fresh media for a 1/12 split, and placed back in the incubator.

2.9 Transfection of GP2-293 cells

Reagents

GP2-293 cells (Clontech, #641530), DMEM with Na/Pyruvate(Invitrogen), PBS (Invitrogen), Fetal Calf Serum(Cambrex), Trypsin 0.5% EDTA (Invitrogen), Poly-L-Lysine (Sigma), Optimem (Invitrogen), Lipofectamin 2000(Invitrogen), pAMPHO plasmid (Clonetech).

The line of EBV transfected B-cells used in these experiments is not transfectable directly by lipofectamin or other methods. As such, the target gene must be transfected into GP2-293 kidney cells, provided by Clonetech as part of their Retro Viral Universal Packaging system[48], and using a pAMPHO plasmid which code for the viral envelope as a co-transfectant. These cells can then stably express the viral coat proteins *gag* and *pol*.

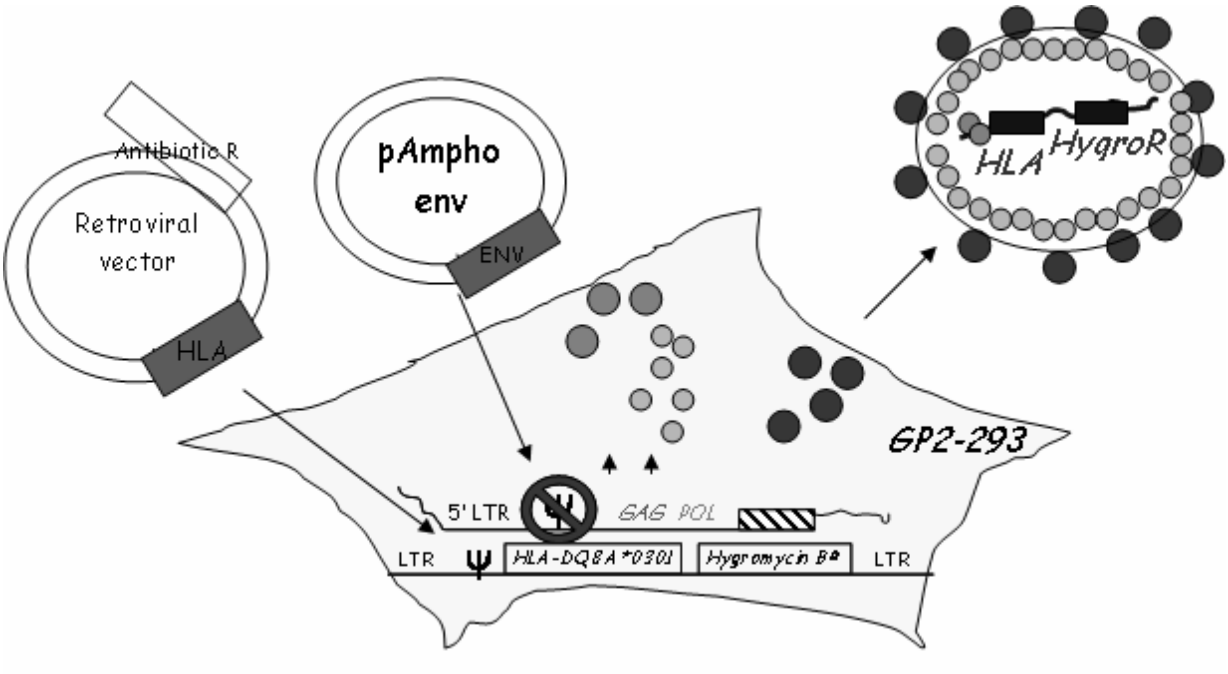


Figure 14: GP2-293 transfection system illustrated, *modified from an illustration by Anna Hayman. GAG,POL and ENV are viral molecules expressed in the cell that forms the viral capsule for further retroviral transduction. Note the packaging signal ψ, directing packing of the target genes instead of the instead of the viral DNA.*

The transfected cells produce virus containing the co-transfected DNA and can then be used to co-culture with 721.82 B-lymphocytes previously infected with the gene coding for the β-chain. The 721.82 cell line has the HLA-DQ and DR region deleted from the genome, and so cannot produce HLA-molecules except the ones coded in the DNA introduced by viral transduction. The viral coat will allow the HLA alpha chain DNA to transduce the B-cells thus introducing the required gene for the alpha chain of the MHC class II molecule, as well as hygromycin resistance to allow for selection of the successfully transduced cells. This accomplishes a transduction by use of a viral vector. The following protocol is the only established one for this kind of viral transduction into B-cells, and is a modified version of the Clonetech protocol [48] by Anna Hayman, after principles explained in Morgenstern[49] and Sekine[50]

10cm diameter dishes were coated with 1:10 diluted poly-l-lysine for 5 minutes to prevent the adherent cells from detaching during transfection, and the dishes was washed three times with PBS. Using cells maintained as described in 2.8, 1 confluent flask of GP2-293 cells for every 3 transduction reaction (including 1 negative control for every 3 transductions) were trypsinated (3 ml trypsin 0.5% EDTA) and washed with PBS . Detached cells were suspended in 8 ml 10%FCS DMEM, spun down for 5min at 800 rpm, then resuspended to a concentration of $4.5*10^5$ cells/ml. 10 ml of suspension added to each dish, incubate ON until 90% confluent, then the media was exchanged to 2%FCS DMEM.

Transfection DNA dilution was done using 10 µg pAMPHO and 10 µg pLHCX expression vector with target gene inserted, mixed with 1.5 ml Optimem for each transfection reaction. A lipofectamin transfection solution for each transfection was created by incubating 50 µl Lipofectamin 2000 with 1.5 ml Optimem for 5 min, then mixing the DNA solutions with the transfection solutions. Note that the negative control mix contains lipofectamin and pAMPHO, but no pLHCX plasmid. After 20min incubation, the lipofectamin/DNA mixture were added drop-wise to the GP2-293 cells, one mix to each dish, and then placed in the incubator.

2.10 Retroviral transduction of 721.82 B-cells in co-culture with GP2-293 cells

Reagents:

721.82 EBV transformed B-cell line, from B.Roep, Fetal Calf Serum (Cambrex), Penicillin (Panpharma), Streptomycin (X-gen Pharmaceuticals), Hygromycin (Invitrogen), 1M HEPES (Invitrogen)

RMPI complete media: 500ml RMPI 1640(Invitrogen), 10% FCS (Cambrex), 1M Hepes(Invitrogen), Penicillin (Panpharma), Streptomycin (X-gen Pharmaceuticals).

24h post-transfection, the transfected GP2-293 cells were irradiated with 12Gy. The target 721.82 EBV-transformed B-cells (described in 2.8) for co-culture retroviral transduction were spun down and resuspended in 50 % used, 50 % fresh complete RMPI media to a concentration of $2.5*10^5$ cells/ml. The 2 % DMEM media of the irradiated GP2-293 cells were discarded, and 8 ml ($2*10^6$ cells in total) were added to each irradiated GP2-293 dish, and incubated for 48 h at 37°C.

After incubation, the dishes were sprayed with a Gibson pipette to detach the GP2-293 cells, then the cells and media from each dish was transferred to a T-25 flask (NUNC). Each dish was washed with 3 ml complete RPMI media, and then added to the same T-25 flasks. After another 24 h of incubation, the cells were spun at 700rpm for 5 min, the media discarded, then resuspended in 8ml complete RMPI media with 0.05mg/ml hygromycin selection antibiotics. The cells were returned to the incubator for another 72 h and then spun as above and resuspended in 8 ml complete RMPI with 0.1 mg/ml hygromycin. The retroviral transduction was now complete, and the pLHCX vector was translated into DQ2-alpha chains and hygromycin resistance inducing proteins. The efficiency of the retroviral transduction is low, so a majority of B-cells will not survive even the mild selection pressure of 0.1 mg/ml hygromycin.

The negative control was used to indicate when the majority of the unsuccessful transductions were dead. Once the transduced cells started to grow, often as much as 14 days after end of co-culture, they were transferred to larger T-75 flasks and expanded. The selection pressure was then increased to 0.3 mg/ml hygromycin to ensure higher levels of DQ2 expression.

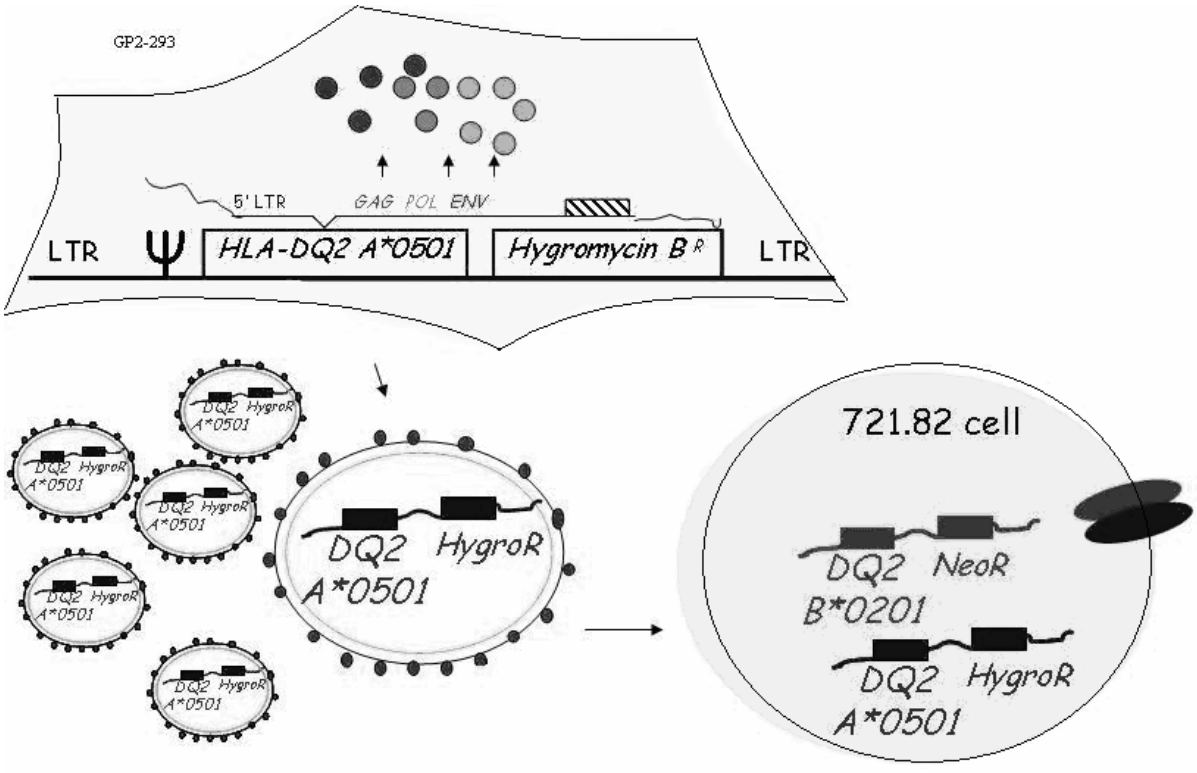


Figure 15: Retroviral transduction of 721.82 B-cells*. GP2-292 cells transfected with HLA-DQ2A*0501 producing viral capsules for transduction into the B-cells already transduced with the β-chain, which results in production of HLA-DQ2 shown on the surface of a 721.82 cell. Illustration modified from an illustration by Anna Hayman.*

2.11 Flow cytometry of transduced B-cells expressing DQ2

Reagents:

Rabbit Anti-Human HLA-DQ2 PE labeled antibody (Diatec) , CerCLIP Rabbit Anti-Human CLIP FITC labeled antibody (BD Pharmingen), PBS (Invitrogen), Fetal Calf Serum (Cambrex)

Cytometry is a process in which the chemical and/or physical properties of a single cell are measured. By extension, flow cytometry is a method where such measurements are done while the cells flow through a measurement apparatus. A suspension of cells is passed through a laser beam, and the refraction from the cells is measured in detector channels[51].

3 attributes of the cells can be measured by light scatter flow cytometry: size (forward scatter), relative complexity or granularity (side scatter), relative intensity of fluorescence. Forward scatter is a measured by looking at the scattering (angle and amount) in a forward direction. Forward scattering is dependent on the surface area of the cell, and the amount and the angle of forward scattering depends on the cell size. Side scatter measures granularity or complexity of the cell by measuring the amount of light that is scattered to the side of the cell when hitting an internal granule or organelle in the cell and not continuing on through the cell to hit the forward detector.

Relative fluorescence is detected by channels FL1-4, dependent on the cytometer. Each channel can detect only one specific wavelength of fluorescence. In these experiments, using the FACSCalibur instrument (BD), both the FL1 and FL2 channels was utilized, capable of generating and detecting fluorescence from FITC (530nm) and PE (585nm) labels[51]. Fluorescence intensity is a measurement of how

many antibodies with conjugated fluorochromes are attached to each cell, and can thus be used to measure the expression of a given protein on that cell.

Both antibodies were used to mark each cell, selecting the population expressing HLA-DQ2 with CLIP bound on the cell surface. Compensation for cross reaction between the two antibodies was done before the measurements were started.

Gating and compensation

Gating establishes which of the cells in a population needs to be analyzed, and which cells should be ignored. Given the high amount of selection pressure the retroviraly transduced cells are under, a large population of dead cells will be present in the samples. These were gated out by selecting a gate that includes cells with low SSC and relatively high FSC, which is the live lymphocyte population. Compensation for cross reactions of the anti-bodies was done by using controls marked only with one anti-body.

Approx. 10^5 cells were used in each sample for the flow cytometry, and added to a well in a 96-well plate. Non-transduced 721.82 cells were included as an isotype control for gating, and two wells for each of transfected 721.82 that express DQ2.5 were included for antibody compensation. The cells were washed by spinning the plates 3X 3 min, 1500rpm , 4 °C, and flicked to remove the liquid between each centrifugation. The cells were resupended in 200 µl 3 % FCS dPBS. 1:10 suspensions in 3 % FCS PBS of the anti-DQ-PE and anti-CLIP-FITC antibodies were created, then 50 µl of each were added to each well. Two wells were only mixed with one antibody each, for antibody compensation. The plate was then incubated in the dark for 30 min on ice. After incubation, the cells were washed twice as described above, leaving each well with 200 µl 3 % FCS dPBS cell suspension. Each cell should now be marked with anti-CLIP antibodies sending a FITC signal and anti-DQ antibodies sending a PE signal. The intensity of these signals will show the amount of each molecule present on the cell surface.

Analysis of the cells was done at a FACSCalibur flow cytometer using CellQuest Pro software suite and WinMDI according to standard operating procedures (BD Bioscience). Graphs and statistical analysis was done using GraphPad (Microsoft).

2.12 Protein extraction and purification of MHC class II molecules

Reagents

Sepharose CL-4B(GE Healthcare), Protein A-sepharose CL-4B (GE Healthcare), PBS (Invitrogen), Sodium azide(Sigma), NP-40(Sigma), Citrate, (Sigma), 2.7µm filter (GE Healthcare Whatman) , 0.7 µm filters (GE Healthcare Whatman), 0.4 µm filters (Millipore), Octylglycoside (Calbiochem), Saturated HCl (Merck), Trisma Base (Sigma)

2X Lysis Buffer: 10 mM Na-orthovandate (Sigma), 50 mM Iodoacetamide (Sigma), 2% NP-40, 2 mM 17.4 µl/ml PMSF serine protease inhibitor (Roche), 15 ml 0.5% Azide PBS

Blind Elution Buffer 1:0.1 M acetate (Sigma), 0.5 M NaCl**,** calibrate to pH 4, filtered

Blind Elution Buffer 2: 0.05 M diethylen (Sigma), 0.15 M NaOH(Sigma) , 0.02 % Azide**,** calibrate to pH11, filtered

Elution Buffer: 0.05 M diethylen, 0.15 M NaOH, 0.02 % Azide ,1 % Octylglycoside (Calbiochem) calibrate to pH 11, filtered

Neutralizing Buffer: 2 M Tris HCl (Sigma), pH 6.3, filtered

To prepare samples for the MALDI-TOF MS, approx.$0.5*10^9$ cells from each retroviral transduction harvested and washed twice with PBS by centrifugation for 20

min at 2200 rpm,. The cells were then resuspended in 50 ml PBS, spun again and frozen at -70°C as a pellet until $1.5*10^9$ cells was harvested from each transduction culture. The cell pellets were resuspended in 15 ml PBS/azide and incubated on ice with periodical mixing for 30min. The suspended cells were centrifuged 2x25 min at 2800rpm to retrieve the supernatant with the peptides, discarding the precipitate. The supernatant was filtered through a succession of 2.7μm, 0.7 μm and finally 0.45 μm filters using water driven vacuum.

The collected lysates from each transduced B-cell culture were run through a prewashed (150 ml PBS/ 0.05% NP-40) column series to capture and purify the MHC class 2 molecules: i: Sepharose CL-3B column, ii: Protein A-sepharose CL4B column, iii: 2.12.E11 rabbit anti-human HLA DQ2 coated-Protein A- sepharose CL4B column created at the Sollid group. The lysate was added at a rate of 2 drops/s, until it had all go onto the columns.

The columns were washed with 50 ml PBS/0.5%NP-40, NP-40 to prevent denaturation of the bound peptides. The protein A column was washed with 50 ml, 0.1M citrate buffer (pH 3) and 25 ml PBS/Az/0.5 % NP-40, then 200 ml PBS/Az. The anti-DQ2 column was washed with 50 ml PBS/0.0% NP-40/Az then with 500 ml PBS/0.1 % SDS/0.5 % NP-40/Az ON, then 50 ml PBS/0.5 % NP-40/Azide, followed by 45ml PBS/0.5% octylglycoside/Az.

Elution of the peptides from columns was accomplished with 25 ml elution buffer (pH11.5) directly into 3 ml of neutralization buffer(pH 6.3)

All the columns were washed with 25 ml PBS/0.5 % NP-40/Az. before blind-eluting the columns with 50 ml blind elution buffer 2, followed by PBS, blind elution buffer 1, PBS, blind elution buffer 2 and finally 200 ml PBS/0.5 % NP-40/Azide. to prepare the columns for the next elution. The eluted DQ2 was then concentrated using Vivaspin 20 MW10000 PES concentration columns (Satorius Stedim Biotech), 3000rpm at 4°C down to 500μl, then washed 2x with 10ml PBS/Az./0.5% octylglucoside and centrifuged down to 100 μl. The concentration of the peptides was

measured using a BCA Protein Concentration Kit(BIO-RAD) and a Victor^3 Multi-counter spectrophotometer.

2.13 Elution of peptides from MHC class II

Reagents

TFA (Sigma), MQ water, 10,000MWCO Microcon filters (Micon)

µg of each of the mutants and wt HLA-DQ2 molecules were used for this. The HLA-DQ2 molecules were added to a 10,000MWCO Microcon filter (Micon), prewashed with 2x500µl MQ water. The filters were then washed through twice with 500µl MQ water, then acid eluted with 75µl 0.5% TFA, pH 2 for 30 min at 37°C. The filters were then spun dry and the filtrate collected.

2.14 Mass spectrometry

Reagents

TFA (Sigma), αCN matrix (Sigma), reversed phased C18 column material (Sigma), R2 column material solution (Sigma)

The basic principles of mass spectrometry are relatively simple. The spectrometer performs 3 functions: it vaporizes the sample compound, produces ions from the resulting gas-phase molecules and then separates these ions according to mass. When producing the gas-phase, the spectrometer invariably shatters the molecule that is being looked out into more than one charged ion [52].

There are many different ways to conduct mass spectrometry, and they are chiefly different in vaporization and mass-measurement technology. The powerful MALDI-TOF technique, which relies on matrix-assisted laser description (MALDI) for vaporization and time-of-flight (TOF) measurement for mass calculation, is often used when dealing with large biomolecules like peptides and proteins as in this thesis.

Matrix-assisted laser desorption uses a nitrogen laser to rapidly sublimate a matrix crystal and the analyte into gas-phase. In the gas-phase proton transfer reactions create charged ions of the peptides, nearly always mono-charged, so there is a 1 to 1 ratio between charge and mass for the analysis[53]. Time-of-flight is measured to calculate the mass of the fragments. Given a constant acceleration voltage, the flight time of a given molecule though the spectrometer is a function of its mass (m)/charge (z). For this to be accurate, the flight path needs to be extensive, often incorporation ion-mirrors that also calibrate the acceleration of the molecules to minimize variables other than m/z. The size of the peptides can then be calculated based on the final TOF [53]. From the size of the peptide fragments measured, it can be deduced which peptides and proteins are present in the sample, to a degree. The size of the CLIP peptides

bound to MHC class II are known, and so sequencing though MS was not needed. The general shape a typical and atypical MS specter eluted from MHC class II is also know, and can be compared with the results from the MALDI-TOF analysis.

The eluted peptides were applied to a ground steel MALDI MS-plate. A Gel-loader peptide tip is plugged with a small amount of C18, and added 5µl R2 material to aid in CLIP retention. The makeshift column was calibrated with 2x12 µl 0.1%TFA, pressed through by a syringe. Then 20 µl eluted peptide (pH2) was applied on to the column and the column was then washed with 4x12 µl 0.1 %TFA. The peptides were then eluted with matrix by applying 1 µl αCN matrix unto the column, then applying the eluate to a spot on the MALDI-plate.

The MALDI-TOF analysis was done on a Bruker Daltonics MALDI-TOF Mass spectrometer at the Proteomics Core Facility, Rikshospitalet, using Bruker Daltonics control and analytic software according to Bruker Daltonics specifications. The heigth of the peaks in the MALDI-TOF mass spectrua produced are not directly comparable to peaks in other such spectra because the strength of the signal from each ion is dependent on many factors that vary between each sample. The peaks can be compared to other peaks within the same spectrum, however. The relative heigth of each peak then indicates the amount of that peptide. Each spectrum can also be compared to other spectra if the peak heigth is ignored. That is why no numbers have been given for the Y-axis of the MS-spectra presented in Results, as the units used are abitrary and of no use for comparissons with other such spectra.

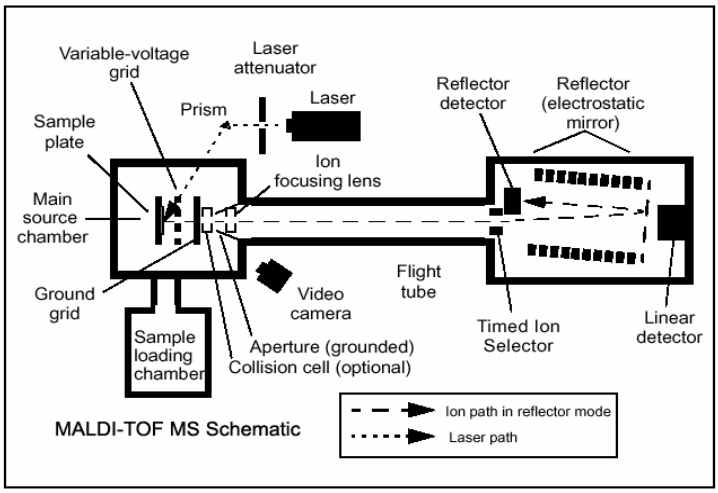


Figure 16: MALDI-TOF MS Schematic *(Applied Biosystems), showing the ions path through the flight tube.*

3. Results

3.1 Mutagenesis

The two MHC class II molecules studied in this thesis, HLA-DQ2.2 and HLA-DQ2.5, differ in only 10 residues in the membrane distal domain of the alpha chain. **Table 4** shows the full sequence of both genes and their polypeptide chains, and where the polymorphisms are located. To study the effect on CLIP presentation these polymorphisms may have, mutations were introduced in position α31 (glutamine to glutamate), α37 (glycin to glutamate) and α72 (serine to isoleucine) in HLA-DQ2.5. In HLA-DQ2.2, mutations were introduced in position α31 (glutamate to glutamine) and α37 (glutamate to glycin). This is shown in **figure 17**. Note that the mutagenesis of HLA-DQ2.5 in position α72 introduces the codon ATT instead of ATC, but they both code for isoleucin. The presence of the C is thus not relevant for the study of the translated protein. The plasmid with HLA-DQ2.2 mutated in position α72 failed when used for retroviral transduction, and that mutation is not shown.

The DQA1*0201 and DQA1*05 DNA sequences have previously been created by cDNA synthesis from B-cells. For mutagenesis, the genes were inserted in the pMOS-Blue vector which was used for bacterial transformation cloning and mutagenesis with the Stratagene Kit described in 2.6. The constructs were subcloned in XL-1 bacteria. The successful mutant genes were sequenced (2.7, **figure 17**) and transferred to the expression vector (2.2).

***Table 4**: **DQ2 α-chain sequences,** wild-type DNA[54 55] and peptide sequences[55 56] of DQA1*05 and DQA1*0201. The DNA sequence in italicized black, aa residues in grey, membrane distal polymorphic residues are marked in bold black.*

DQA1*05

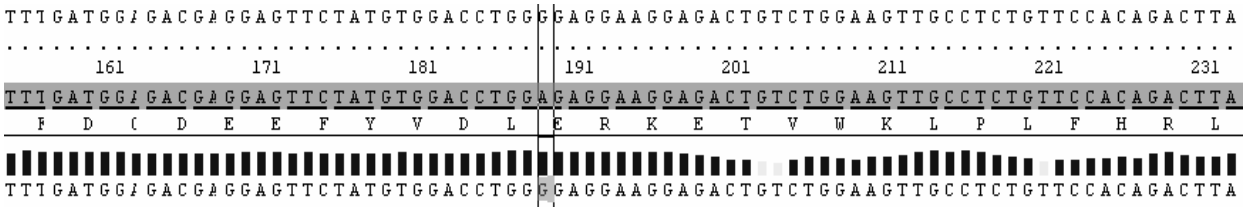
ATGATCCTAAACAAAGCTCTGATGCTGGGGGCCCTTGCCCTGACCACCGTGATGAGCCCC
M I L N K A L M L G A L A L T T V M S P
TGTGGAGGTGAAGACATTGTGGCTGACCACGTCGCCTCTTATGGTGTAAACTTGTACCAG
C G G E D I V A D H V A S Y G V N L Y Q
TCTTACGGTCCCTCTGGCCAGTACACCCATGAATTTGATGGAGATGAGCAGTTCTACGTG
S Y G P S G Q **Y** T H E F D G D E **Q** F Y V
GACCTGGGGAGGAAGGAGACTGTCTGGTGTTTGCCTGTTCTCAGACAATTTAGATTTGAC
D L **G** R K E T V W **C** L P **V L R Q F** R F D
CCGCAATTTGCACTGACAAACATCGCTGTCCTAAAACATAACTTGAACAGTCTGATTAAA
P Q F A L T N I A V L K H N L N **S** L I K
CGCTCCAACTCTACCGCTGCTACCAATGAGGTTCCTGAGGTCACAGTGTTTTCCAAGTCT
R S N S T A A T N E V P E V T V F S K S
CCCGTGACACTGGGTCAGCCCAACATCCTCATCTGTCTTGTGGACAACATCTTTCCTCCT
P V T L G Q P N I L I C L V D N I F P P
GTGGTCAACATCACATGGCTGAGCAATGGGCACTCAGTCACAGAAGGTGTTTCTGAGACC
V V N I T W L S N G H S V T E G V S E T
AGCTTCCTCTCCAAGAGTGATCATTCCTTCTTCAAGATCAGTTACCTCACCCTCCTCCCT
S F L S K S D H S F F K I S Y L T L L P
TCTGCTGAGGAGAGTTATGACTGCAAGGTGGAGCACTGGGGCCTGGACAAGCCTCTTCTG
S A E E S Y D C K V E H W G L D K P L L
AAACACTGGGAGCCTGAGATTCCAGCCCCTATGTCAGAGCTCACAGAGACTGTGGTCTGC
K H W E P E I P A P M S E L T E T V V C
GCCCTGGGATTGTCTGTGGGCCTCGTGGGCATTGTGGTGGGCACTGTCTTCATCATCCGA
A L G L S V G L V G I V V G T V F I I R
GGCCTGCGTTCAGTTGGTGCTTCCAGACACCAAGGGCCCTTGTGA
G L R S V G A S R H Q G P L -

DQA1*0201

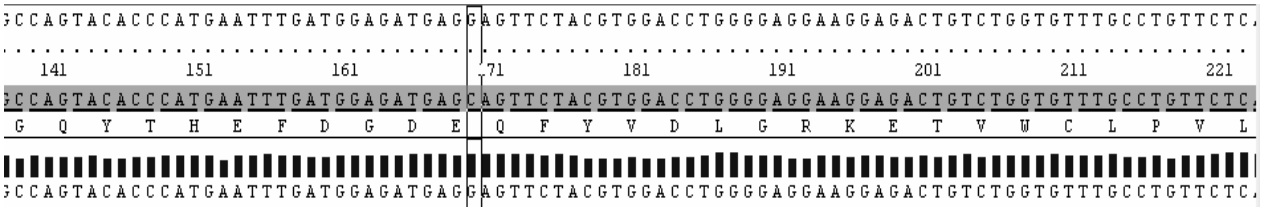
ATGATCCTAAACAAAGCTCTGATGCTGGGGGCCCTCGCCCTGACCACCGTGATGAGCCCT
M I L N K A L M L G A L A L T T V M S P
TGTGGAGGTGAAGACATTGTGGCTGACCACGTTGCCTCTTACGGTGTAAACTTGTACCAG
C G G E D I V A D H V A S Y G V N L Y Q
TCTTACGGTCCCTCTGGCCAGTTCACCCATGAATTTGATGGAGACGAGGAGTTCTATGTG
S Y G P S G Q **F** T H E F D G D E **E** F Y V
GACCTGGAGAGGAAGGAGACTGTCTGGAAGTTGCCTCTGTTCCACAGACTTAGATTTGAC
D L **E** R K E T V W **K** L P **L F H R L** R F D
CCGCAATTTGCACTGACAAACATCGCTGTGCTAAAACATAACTTGAACATCCTGATTAAA
P Q F A L T N I A V L K H N L N **I** L I K
CGCTCCAACTCTACCGCTGCTACCAATGAGGTTCCTGAGGTCACAGTGTTTTCCAAGTCT
R S N S T A A T N E V P E V T V F S K S
CCCGTGACACTGGGTCAGCCCAACACCCTCATCTGTCTTGTGGACAACATCTTTCCTCCT
P V T L G Q P N T L I C L V D N I F P P
GTGGTCAACATCACCTGGCTGAGCAATGGGCACTCAGTCACAGAAGGTGTTTCTGAGACC
V V N I T W L S N G H S V T E G V S E T
AGCTTCCTCTCCAAGAGTGATCATTCCTTCTTCAAGATCAGTTACCTCACCTTCCTCCCT
S F L S K S D H S F F K I S Y L T F L P
TCTGCTGATGAGATTTATGACTGCAAGGTGGAGCACTGGGGCCTGGATGAGCCTCTTCTG
S A D E I Y D C K V E H W G L D E P L L
AAACACTGGGAGCCTGAGATTCCAGCACCTATGTCAGAGCTCACAGAGACTGTGGTCTGT
K H W E P E I P A P M S E L T E T V V C
GCCCTGGGGTTGTCTGTGGGCCTCGTGGGCATTGTGGTGGGGACCGTCTTGATCATCCGA
A L G L S V G L V G I V V G T V L I I R
GGCCTGCGTTCAGTTGGTGCTTCCAGACACCAAGGGCCCTTGTGA
G L R S V G A S R H Q G P L -

DQA1*0201 E31Q (GAG →CAG)

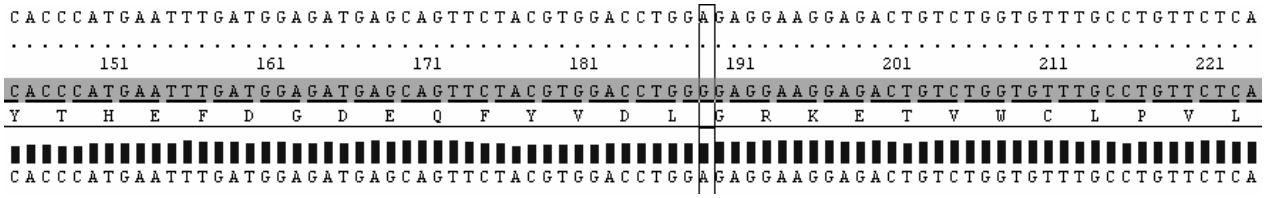
DQA1*0201 E37G (GAG → GGG)



DQA1*05 Q31E (CAG→GAG)



DQA1*05 G37E (GGG→GAG)



DQA1*05 S72I (AGT→ATT)

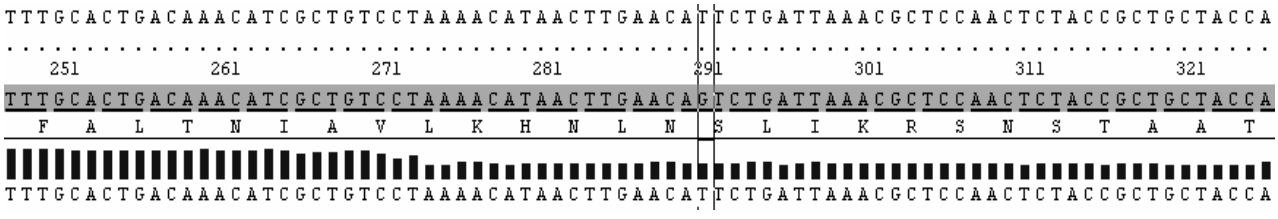


Figure 17: Sequences of mutated polymorphic sites with mutated base-pairs marked with a frame*.*
Consensus sequence is shown on top, then the wt gene and polypeptide chain. The mutated gene is shown on the bottom. The comparison of sequences was done using SecScape2.5 (Applied Biosystems).

3.2 Flow cytometry

The expression vector containing the mutated α –chain genes were retroviraly transduced into EBV-transformed B-cells as described in the methods part of this thesis. These cells were grown to a density of at least 400,000 cells/ml, and then used for flowcytometry on a FACSCalibur flowcytometer (2.11).

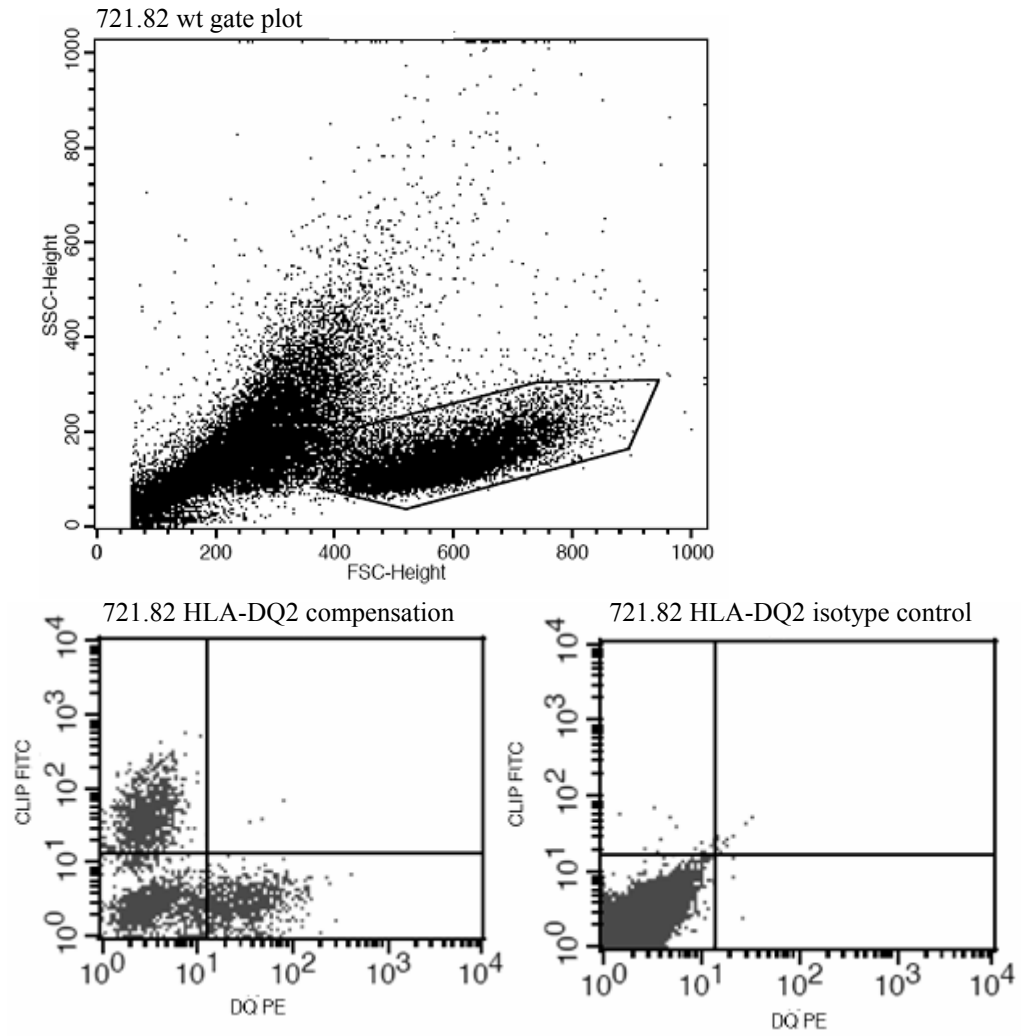
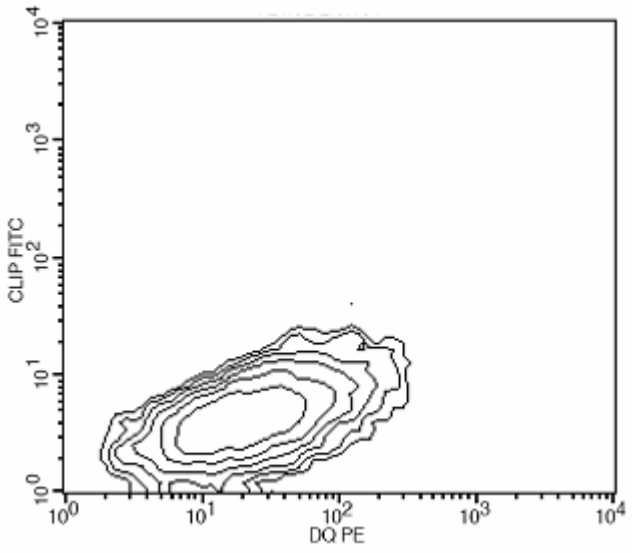
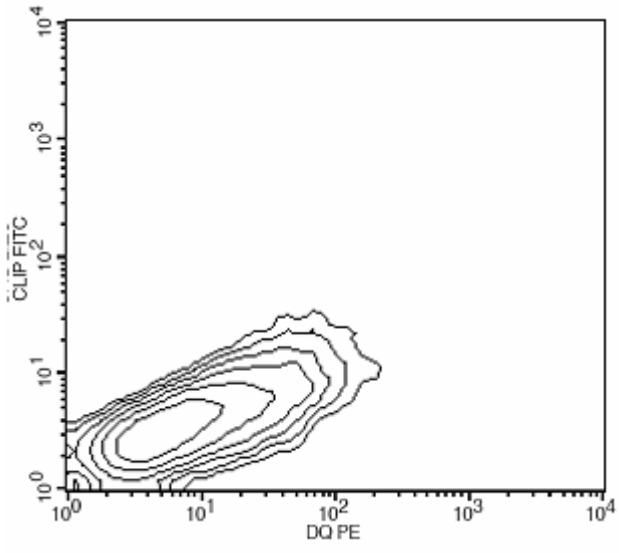


Figure 18: Flow cytometry calibration *a) Gate shown on unmarked 721.82 cells, the gate was used for all cytometry measurements. b) Compensation with 721.82 HLA-DQ2.5 cells, c) Isotype control using 721.82 cells to show that no unspecific reactions are present. See 2.11 for methodology details.*

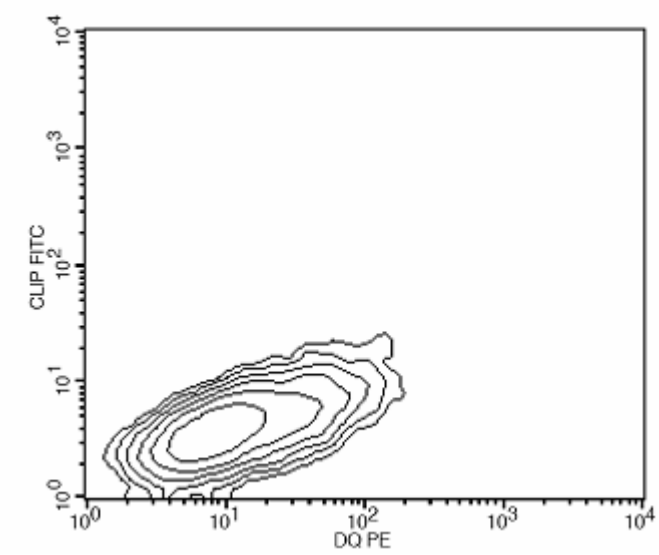
721.82 HLA-DQ2.5wt



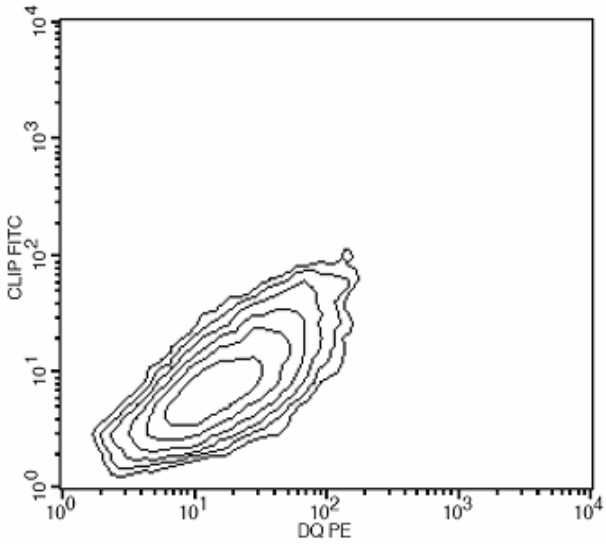
721.82 HLA-DQ2.5Qα31E



721.82 HLA-DQ2.5Gα37E



721.82 HLA-DQ2.5Sα72I



721.82 HLA-DQ2 negative control

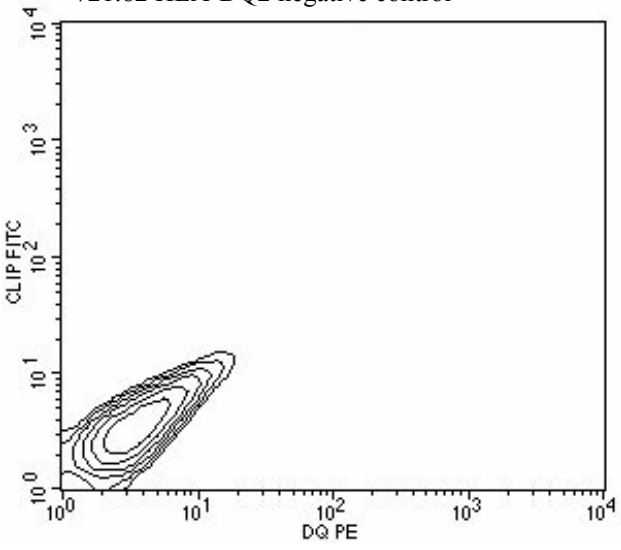


Figure 19: Flow cytometry plots for cells expressing HLA-DQ2.5 mutants. *Signal intensity of HLA-DQ2 (PE) vs. CLIP(FITC). The negative control has not been labeled with any antibodies. Note that the scale is logarithmic on both axes.*

The flow cytometry results from HLA-DQ2.5 are presented in **figure 19**. Contour plots of the wild type (wt) cell population and the three populations expressing mutant HLA-DQ2.5 show that there is a distinct increase in CLIP surface presentation from wt to mutants.

The largest increase is seen with the mutation in position α72, an isoleucine replacing a serine. The α72 plot shows a marked increase in CLIP signal at a given level of HLA-DQ2.5, compared to the wt plot. The plot for HLA-DQ2.5 mutated in position α31, exchanging glutamine for glutamate, shows a more modest increase of CLIP signal. The last mutant, glycin exchanged for glutamate in position α37, shows a small increase of CLIP signal compared to wt, but not as much as the other two mutants. However, the lower overall HLA-DQ2.5 expression in the α37 mutated cell population will increases the effect of HLA-DM. The observed lower CLIP presentation can be partially explained by this higher HLA-DM effect.

The flow cytometry of HLA-DQ2.5 shows that all the mutations in the three target polymorphic sites increase CLIP presentation the binding groove, with the residue at α72 providing the most pronounced increase.

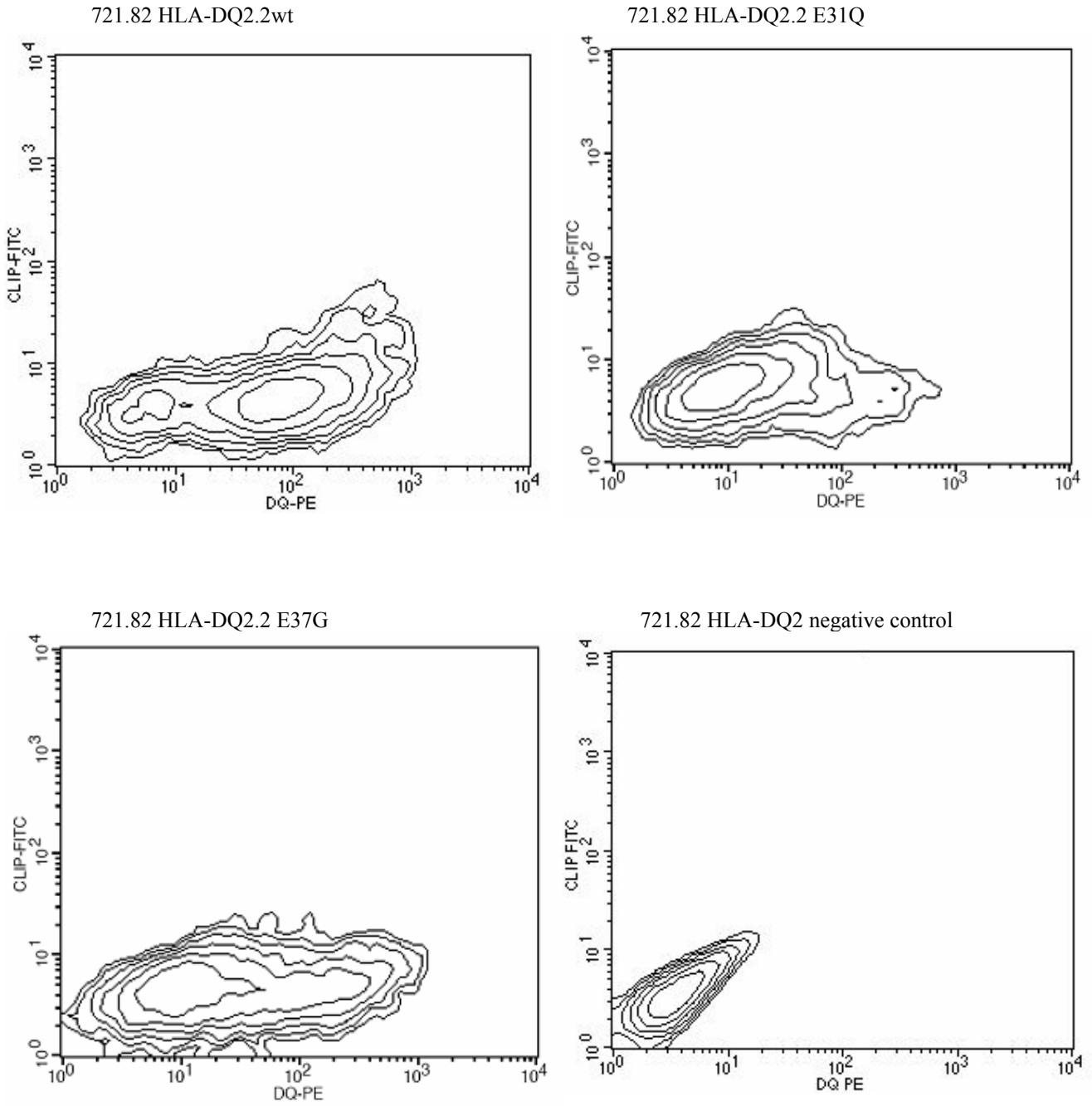


Figure 20: Flow cytometry plots for cells expressing HLA-DQ2.5 mutants. *Signal intensity of HLA-DQ2 (PE) vs. CLIP(FITC). The negative control has not been labeled with any antibodies.*

The flow cytometry results from HLA-DQ2.2 wt and the two successfully mutated HLA-DQ2.2 mutant molecules show that the mutations have little effect on CLIP presentation. These plots are shown in **figure 20**. The same kind of plots has been used for HLA-DQ2.2 as for HLA-DQ2.5, and they can be interpreted the same way. All the plots show a relatively high expression of surface HLA-DQ2.2, but low CLIP signals even with large amounts of HLA-DQ2 expressed. The observed low CLIP at relative high expression levels indicates poor CLIP stability in the binding groove. Interestingly, the wt HLA-DQ2.2 plot has a much higher HLA-DQ2 signal than any of the mutants of HLA-DQ2.5. Yet, even at 10X the DQ2 expression the amount of CLIP is barely equal to the HLA-DQ2.5 α72 mutant.

The cells expressing HLA-DQ2.2 with mutation in α31 and α37 have a lower HLA-DQ2.2 expression than the wild-type, and the plot shows no CLIP presentation compared to the negative control. These results indicate that the introduced mutations in HLA-DQ2.2 lowered the amount of CLIP presented. Even at high HLA-DQ2.2 expression, the CLIP amount is not comparable to HLA-DQ2.5 wt and mutants. This serves to confirm the observations from the HLA-DQ2.5 mutations, since the opposite mutations seem to have the opposite effect.

3.3 Mass spectrometry

Retroviraly transduced 721.82 B-cells were harvested and lysed, and the MHC molecules were purified. The peptides were eluted from the binding groove as described in 2.12 and 2.13. MALDI-TOF mass spectrometry was then performed on the eluted peptides (2.14).

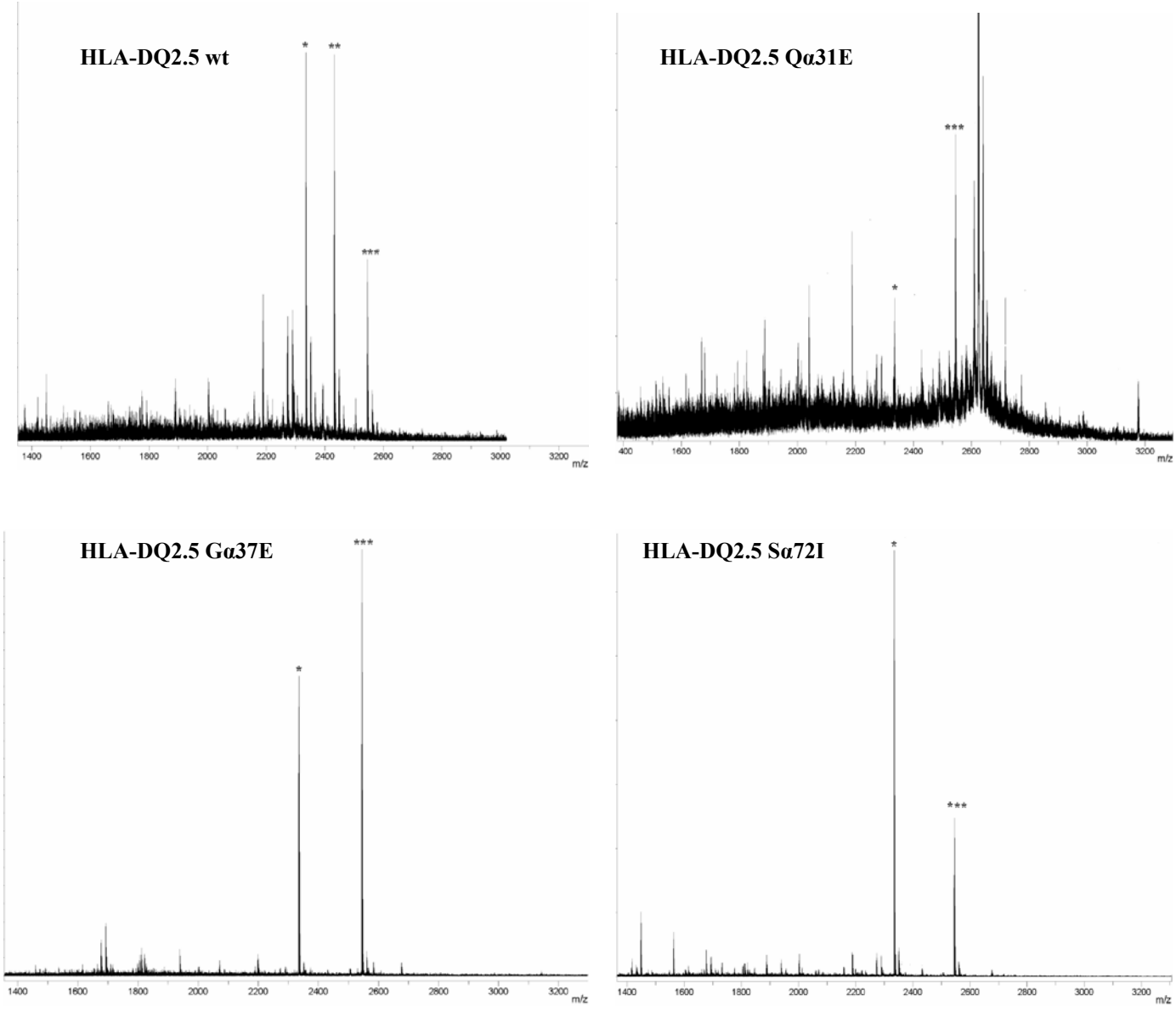


Figure 21: MALDI-TOF spectra of HLA-DQ2.5. *Mass/charge along the X-axis, arbitrary units along the Y axis. Peaks that have been identified as CLIP are marked with stars. *: 2333.3 m/z, **:2430.4m/z, ***:2543.4m/z.*

The spectra from MALDI-TOF of HLA-DQ2.5 wt and mutants are shown in **figure 21**. The wt HLA-DQ2.5 spectrum is dominated by CLIP peptide peaks, with other endogenous peptides also evident. Note that CLIP easily ionize, due to the K and R residues in the peptide. This may make the CLIP peptides appear more aboundant than they acctualy are, and will tend to skew the spectra so that CLIP peptides always show up with a strong signal while some endogenous petides might not show up at all.

The spectrum from HLA-DQ2.5 Qα31E shows as large or larger amounts of CLIP presentation compared to wt, but the spectrum is slightly distorted because of the very strong signal from the oxidized peptide. That makes the results more unreliable from this spectrum. The largest differences can be seen in the spectra for HLA-DQ2.5 Gα37E and Sα72I, both completely dominated by the CLIP peaks. The amount of other peptide is likely larger than can be seen in the spectra, but the CLIP signal for both these samples is so strong that it is clear that the amount of CLIP is considerable. This confirms the flow cytometry results showing that a larger amount of CLIP peptides are presented on the mutant HLA-DQ2.5 compared to wild type in the molecules mutated in position α37 and α72, and to a lesser degree for the molecules mutated in α31.

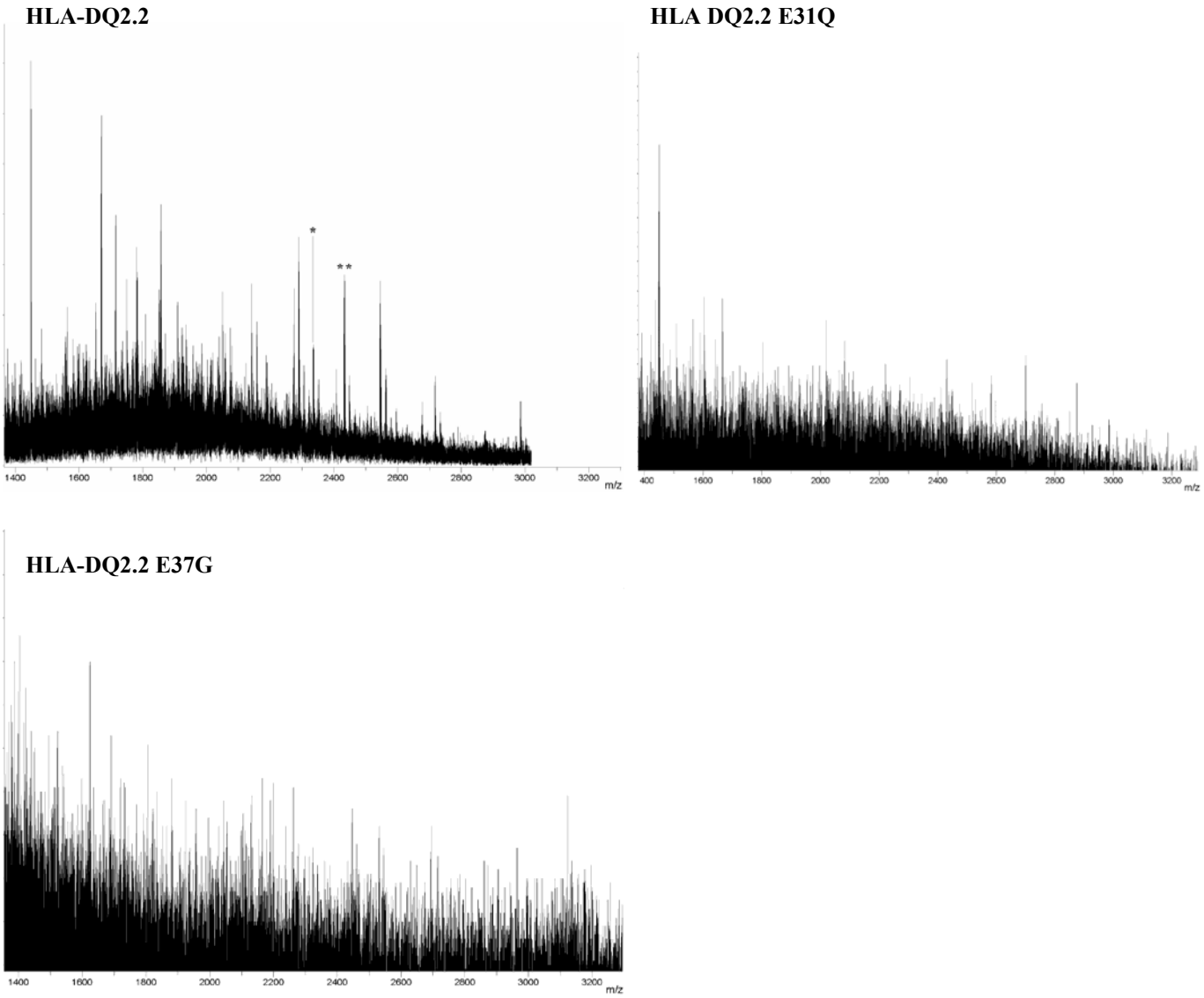


Figure 22: MALDI-TOF spectra of HLA-DQ2.2. *Mass/charge along the X-axis, arbitrary units along the Y axis. Peaks that have been identified as CLIP are marked with stars. *: 2333.3 m/z, **:2430.4*

The spectra for wt HLA-DQ2.2 and mutants are shown in **figure 22.** Two low CLIP peptide peaks in the wt HLA-DQ2.2 spectrum are seen. The peaks are not dominating the spectrum as with HLA-DQ2.5. As CLIP is easily ionziable, this means that the amount of CLIP present is likely very small. Both the mutant spectra show only a large amount of different endogenous peptides, with no single peak dominating the spectra. No distinct CLIP peptide peaks can be found. This confirms the flow cytometry results showing that mutations in HLA-DQ2.2 position α31 and α37 either lower or have no effect on the amount of CLIP presented in the binding groove of HLA-DQ2.2.

4. Discussion

Autoimmune diseases are largely a conundrum even after decades of research on the subject. Especially the molecular and genetic backgrounds of such disorders have been under close scrutiny. Genetic predisposition for autoimmune diseases is tied largely to the MHC region of chromosome 6. Considering the vital part MHC class I and MHC class II peptide presentation play in the activation of the adaptive immune response, this is not surprising. Celiac disease is an example of a condition where a MHC class II gene, in the form of HLA-DQ2.5, is strongly associated with the disease. The very similar HLA-DQ2.2 molecule is not associated with celiac disease. The two molecules are highly similar, and their membrane distal domains contain only 10 polymorphic residues in the alpha-chain. The molecules have similar peptide binding motifs and both bind gluten T-cell epitopes. However, they do show a marked difference in CLIP presentation in the peptide binding groove has been observed between the two HLA molecules. HLA-DQ2.5 is known to present CLIP peptides, while HLA-DQ2.2 is not. Combined with the recent discovery that both HLA-DQ2 molecules appear to be poor substrates for HLA-DM[1,38], there are indications that the differences in CLIP presentation can be explained by investigating the 10 polymorphic residues in their membrane distal domains. The results from part of that investigation are presented in this thesis.

The membrane distal polymorphic sites in the alpha-chain of HLA-DQ2.5 and HLA-DQ2.2 are designated α22, α31, α37, α44, α47, α48, α49, α50, α51 and α72. Of these polymorphisms, α44 and α47-51 are located in a region of HLA-class II thas is suggested to be involved in the HLA-DM interaction mechanism[24,25]. The other 4 polymorphic sites are located in or around the peptide binding groove. The location of all of the polymorphic sites can be seen in **figure 9**.

As described in the introduction, the binding groove of the MHC-class II molecules contain binding pockets that interact with the side chains of the peptide in the groove. Combined with hydrogen bonds to the peptide backbone from invariant anchor

residues, these interactions bind the peptide to the binding groove[57]. The polymorphic sites in the area around the binding groove might be involved in this anchoring, and were thought to be interesting targets for studying the cause of the differences in CLIP presentation mentioned above. This thesis work focused on the effects of the polymorphic residues α31, α37 and α72.

In this thesis, flow cytometry and mass spectrometry data from mutated HLA-DQ2.5 and HLA-DQ2.2 have been presented. The mutated HLA-DQ2.5 molecules have had a residue in one of the 3 target polymorphic sites exchanged for one normally present in the HLA-DQ2.2 molecule. This creates 3 different mutant HLA-DQ2.5 molecules with a single residue from HLA-DQ2.2. The results show that such a mutation in any of the three polymorphic sites investigated increase the amount of CLIP peptide presented by the mutant molecule, when compared to wt HLA-DQ2.5. The greatest effect is seen with a mutation of serine to isoleucine in position α72.

The large amount of CLIP presented by wt HLA-DQ2.5 was earlier reported by Van De Wal et.al [58] and Vartdal et.al[59] investigating the binding of peptides to the DQ2 groove. The results reported in this thesis indicate that the wt residues (Q,G and S) in position α31, α37 and α72 of HLA-DQ2.5 do not increase CLIP presentation. Those residues may in fact have a limiting effect on the CLIP presentation reported in the above papers. This hypothesis was partially confirmed by doing the reverse mutation, introducing the residues from HLA-DQ2.5 into HLA-DQ2.2. The mass spectrometry and flow cytometry data presented here from the HLA-DQ2.2 mutants show that introducing the HLA-DQ2.5 residues in the target polymorphic sites either have no effect or reduce the CLIP presentation of those molecules. Together, these results indicate that the residues in the polymorphic sites α31, α37 and α72 do not explain the high CLIP presentation of HLA-DQ2.5, but on the contrary increase the peptide affinity of HLA-DQ2.2.

It must be mentioned that the reverse mutation experiment did not succeed for the residue in position α72 of HLA-DQ2.2 (isoleucine exchanged with serine). The DNA mutated to express serine in position α72 failed repeatedly to retroviraly transduce the

target B-cells. Retroviral transduction of B-cells is a very delicate process, and only a small amount of cells are successfully transduced and survive the antigenic selection process. The most likely explanation of the failure in this case is that impure plasmid DNA was used for the transfection of the packaging cells. Unstable DNA or the presence of proteins would disrupt the transfection process. If a new expression vector that includes the mutated gene could be constructed, there is no reason the transduction should not work.

The results presented in this thesis are a separate and independent work, and show the influence of the polymorphic residues in position α31, α37 and α72 on the presentation of CLIP by HLA-DQ2.5 and HLA-DQ2.2. The results show that these sites are do not explain for the high CLIP presentation of HLA-DQ2.5, but are involved in increasing peptide affinity in HLA-DQ2.2. The results also indicate that these polymorphic sites have an effect on CLIP presentation in both molecules, and that they are somehow limiting the CLIP presentation seen in wt HLA-DQ2.5. Introduction of HLA-DQ2.2 residues in the polymorphic positions of HLA-DQ2.5 creates even better CLIP presenting molecules, while HLA-DQ2.5 residues introduced in HLA-DQ2.2 remove the meager CLIP presentation shown in the HLA-DQ2.2 wt. How can this be explained?

To understand how these observed effects may be related to CLIP presentation and stability in HLA-DQ2.5, the results from the rest of the study by the Sollid group will have to be included and discussed. Fallang et.al.[1] shows that mutations in the presumed DM-interaction area (α 44 and α 47-51) have little effect on the HLA-DQ2.5 CLIP presentation. The reverse mutations in HLA-DQ2.2 were not done, as little effects on HLA-DQ2.5 were observed. The mutation of position α22 from a tyrosine to a phenylalanine in HLA -DQ2.5 has a much more pronounced effect, which removes CLIP presentation completely from HLA-DQ2.5. The reverse reaction shows an equally pronounced effect that increase the CLIP presentation on the HLA-DQ2.2 to a level far beyond wt HLA-DQ2.5, or even the HLA-DQ2.5 α72 mutant presented in this thesis. These results are shown in **Figure 23**.

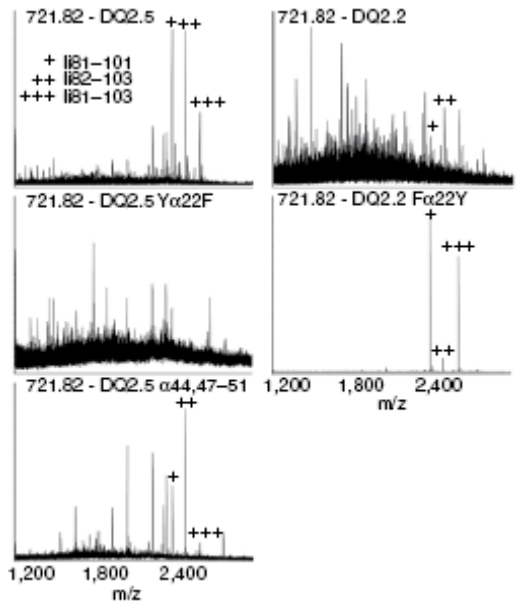


Figure 23: MALDI-TOF results from Fallang et.al, 2009[1]. *The HLA-DQ2.5 α44, α47-51 mutant presents the same amount of CLIP as wt HLA-DQ2.5. Loss of CLIP presentation is clear in the HLA-DQ2.5 mutated in position α22. Very distinct and high CLIP peaks are seen from the HLA-DQ2.2 molecules mutated in position α22.*

The implications from the results presented above are that tyrosine in position α22 is the main factor determining CLIP presentation in the HLA-DQ2.5 binding groove. The polymorphism in position α22 is critical for explaining the difference in CLIP presentation between the two HLA-DQ2 molecules, but the results presented in this thesis makes it clear that the situation is a little more complex.

The mutant HLA-DQ2.2 with tyrosine in position α22 has an extreme amount of CLIP presentation, much higher than wt HLA-DQ2.5. The results obtain in this thesis indicate that the residues in position α31, α37 and α72 of HLA-DQ2.2 (E, E and I) increase CLIP binding when combined with the tyrosine in position α22, as shown in **figure 21,** and allows for CLIP binding even if no tyrosine is present (**figure 22**)**.** Conversely, the residues in position α31, α37 and α72 of HLA-DQ2.5 (Q,G and E) lowers the CLIP binding when combined with tyrosine in α22, and also if combined with a phenylalanine in the same position.

The combination of all of the residues that increase CLIP binding, as in HLA-DQ2.2 with tyrosine introduced in position α22, gives a combined effect of extreme CLIP presentation. The removal of all such residue, as in HLA-DQ2.5 with phenylalanine in position α22, gives no CLIP presentation. While the polymorphic residue in position α22 is the most important residue for explaining the CLIP presentation; the three polymorphic sites studied in this thesis also have a clear effect in combination

with α22. Wild-type HLA-DQ2.5 lacks the glutamates and isoleucine in position α31, α37 and α72 and as such is not as strong a CLIP presenter as it could be. HLA-DQ2.2 has those residues in the polymorphic sites and can thus present a small amount of CLIP even without the tyrosine in α22.

Combined, these results explain the disparity of CLIP presentation between HLA-DQ2.2 and HLA-DQ2.5. They do not explain how and why the presence of a certain residue in these polymorphic sites can influence CLIP presentation on the cell surface, however. Other results indicate that the residues increase or decrease the stability of CLIP binding in the groove of the MHC class II molecules. Binding stability will have an increased importance since it has been indicated that both HLA-DQ2 molecules are poor substrates for HLA-DM[1,38]. Fallang et.al.[1] proposes that the tyrosine in position α22 can engage in a hydrogen bond to the peptide backbone of the bound peptide to increase stability. The importance of such a H-bond between polymorphic residue and the peptide backbone has not been previously studied, not has it been implicated in pathogenesis of a disease (**Figure 24**). How the polymorphic residues in α31, α37 and α72 would influence the stability of CLIP in the binding groove is not known, and more experiments and crystal structure of DQ2.2 would have to be completed before a hypothesis could be presented. The results in this thesis only indicate that all of the mentioned positions are important for CLIP binding stability, α72 in particular.

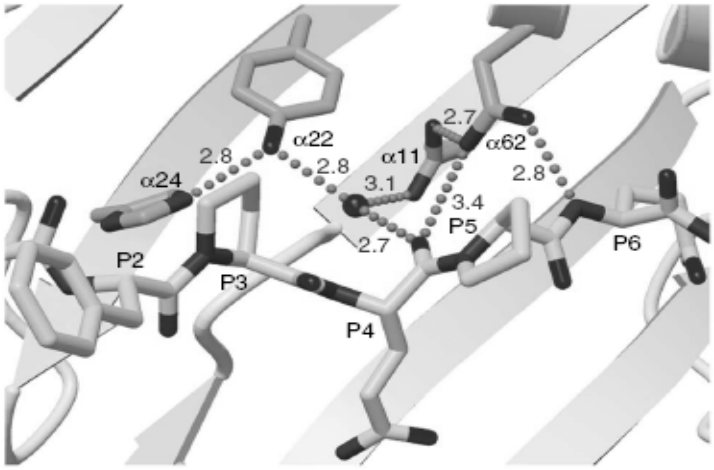


Figure 24: Hydrogen binding network surrounding tyrosine in position α22 in HLA-DQ2.5*, showing how the tyrosine can induce a stronger hydrogen bond network through H_2O. α-I-gliadin bound to the groove (white), oxygen (black), hydrogen bonds shown in dashed lines. Numbers show distances in Å. Illustration from Fallang.et.al.*[1].

CLIP stability in the binding groove should be comparable to the stability of other peptides in the binding groove. Fallang et.al[1] shows that APCs expressing wt HLA-DQ2.5 have protracted antigen presentation after encounter with gliadin peptides, compared to APCs expressing wt HLA-DQ2.2. Gliadin peptides seem to bind to the wt HLA-DQ2.5 with higher stability than wt HLA-DQ2.2, and are able to stimulate T-cells for a longer time. HLA-DQ2.2 which has been mutated in position α22 to express tyrosine instead of phenylalanine shows the same ability for protracted antigen presentation. The reverse mutation introduced in HLA-DQ2.5 removes this ability. There are no such experiments for the polymorphic residues studied in this thesis, but the ones from Fallang.et.al[1] confirm that the high presentation of CLIP in HLA-DQ2.5 is due to increased peptide stability in the groove. They also show that the increases or decreases in stability are influenced by the polymorphic residues, and are not limited to CLIP peptides only. These results support that the conclusion of this thesis; that the effects of the polymorphisms in α31, α37 and α72 on CLIP presentation are due to changes in binding stability.

The main question surrounding the differences between HLA-DQ2.5 and HLA-DQ2.2 is still their different association to celiac disease. Can larger amounts of CLIP

presentation and longer CLIP binding stability explain part of this difference? It is well known that T-cells are negatively and positively selected in the thymus, and that central tolerance to self-peptides is induced there. This is done by the thymus epithelial cells, which express MHC class II and the AIRE transcription factor. Peptides from every tissue are presented on the MHC class II molecules in the thymus and used for negative selection of autoreactive T-cells[5]. HLA-DQ2.5 would present large amount of CLIP peptides on these epithelial cells, which might influence development of self-tolerance in unknown ways, perhaps leading to the development of autoreactive T-cells.

Protracted antigen presentation can play a critical role in the initiation of a T-cell response to an antigen. Following antigen loading, the journey for an APC from tissue to a draining lymph node may last up to 48 h. During that time, the bound peptide can easily be lost if the stability of the binding is low. In addition, T-cells are more easily activated by stabile peptides that can participate in longer T-cell interactions [60,61]. Henrickson et.al.[62] recently showed that T-cell activation is dependent on the antigen dose that is presented by APCs. They also show that there is a threshold of antigens which must be reached before a T-cell response is activated. A MHC class II molecule that binds gluten T-cell epitopes with low stability might be unable to retain enough such antigens to activate a T-cell response. A molecule where such peptides are more stabile would likely be able to reach the threshold for activation and start an immune response to the antigen.

The stability of the binding of gliadin-peptides to HLA-DQ2.5 might be critical for the initiation of an autoimmune response against gluten. The polymorphic sites studied in this thesis seem to work against the influence of the residue in position α22 in wild-type HLA-DQ2.5, lowering the stability of peptide binding. The characterization of this effect may have taken us a small step further in understanding celiac disease and other HLA-associated autoimmune diseases, as that peptide binding stability may be at the heart of the association of HLA-DQ2.5 with such diseases.

5. Further prospects

Some of the results from this thesis have already been part of a publication by the Sollid group[1], but there are still many ways to take the project further. The mutation in position α72 of HLA-DQ2.2 will need to be tested to show that an isoleucine in this position has the predicted effect. Dissociation studies and protracted antigen presentation studies for all 3 polymorphic residues studied here would help confirm the effect of these polymorphic sites on the binding stability of non-CLIP peptides.

A major goal would be to be able to crystallize the CLIP-DQ2.2 complex and do X-ray crystallography of its structure, to measure potential hydrogen bonds and see the effect of the isoleucine in position α22.

Knowing which residues influence gliadin peptide stability in the HLA-DQ2.5 binding groove may also be of interest as target sites for development of potential medication against autoimmune HLA-linked diseases like celiac disease.

Lastly, there are many other autoimmune diseases linked to HLA-molecules similar to HLA-DQ2.5. The Sollid group has already started a project to investigate the tyrosine in position α22 of HLA-DQ8. Position α31, α37 and α72 in HLA-DQ8 might also be of interest if this study finds any difference in HLA-DQ8 peptide stability when position α22 is mutated.

6. References

1 Lars-Egil Fallang, Elin Bergseng, Kinya Hotta et al., *Nat Immunol* **10**, 1096 (2009).

2 Bjarne Bogen and Ludvig Munthe, *Immunologi*, 2 ed. (Universitetsforlaget, 2007).

3 Charles Alderson Janeway, Paul Travers, Mark Walport et al., *Immunobiology*, 6 ed. (Garland Science, 2005).

4 Rolf M.Zinkernagel, *Current Opinion in Immunology* **14** (4), 523 (2002).

5 Bruno Kyewski and Ludger Klein, *Annu Rev Immunol.* **24** (1), 571 (2006).

6 Gail A. Bishop and Bruce S. Hostager, *Current Opinion in Immunology* **13** (3), 278 (2001).

7 Mark J Shlomchik, Joseph Craft, and Mark J. Mamula, *Nat Rev Immunol* **1** (2), 147 (2001).

8 Jonathan C. W. Edwards and Geraldine Cambridge, *Nat Rev Immunol* **6** (5), 394 (2006).

9 Tucker W.LeBien and Thomas F.Tedder, *Blood* **112** (5), 1570 (2008).

10 Brian Flutter and Bin Gao, *Cell Mol Immunol* **1** (1), 22 (2004).

11 Liang Zhou, Mark M. W. Chong, and Dan R. Littman, *Immunity* **30** (5), 646 (2009).

12 Paula Bryant and Hidde Ploegh, *Current Opinion in Immunology* **16** (1), 96 (2004).

13 Li Wu and Yong-Jun Liu, *Immunity* **26** (6), 741 (2007).

14 J. Magarian Blander, *Annals of the Rheumatic Diseases* **67** (Suppl_3), iii44 (2008).

[15] Dmitry A.Ovchinnikov, *Genesis* **46** (9), 447 (2008).
[16] Daniel Rodríguez-Pinto, *Cellular Immunology* **238** (2), 67 (2005).
[17] Facundo D. Batista and Michael S. Neuberger, *Immunity* **8** (6), 751 (1998).
[18] The M. H. C. sequencing consortium, *Nature* **401** (6756), 921 (1999).
[19] Peter Cresswell, *Annu Rev Immunol.* **12** (1), 259 (1994).
[20] C.A. Nelson and D.H. Fremont, *Reviews in Immunogenetics* **1** (1), 47 (1999).
[21] Richard Wubbolts and Jacques Neefjes, *Immunological Reviews* **172** (1), 189 (1999).
[22] Oddmund Bakke and Bernhard Dobberstein, *Cell* **63** (4), 707 (1990).
[23] Colin Watts, *Nat Immunol* **5** (7), 685 (2004).
[24] Achal Pashine, Robert Busch, Michael P. Belmares et al., *Immunity* **19** (2), 183 (2003).
[25] Matthew N. Davies, Abigail Lamikanra, Clare E. Sansom et al., *Molecular Immunology* **45** (4), 1063 (2008).
[26] Harald Kropshofer, Günter J. Hämmerling, and Anne B. Vogt, *Immunological Reviews* **172** (1), 267 (1999).
[27] Anthony L DeFranco, Richard M Locksley, and Miranda Robertson, *Primers in Biology: Immunity*. (New Science Press, 2007).
[28] E. Yvonne Jones, *Current Opinion in Immunology* **9** (1), 75 (1997).
[29] Christopher A.Nelson, Nicholas J.Viner, and Emil R.Unanue, *Immunological Reviews* **151** (1), 81 (1996).

[30] Ludvig M. Sollid, *Nat Rev Immunol* **2** (9), 647 (2002).

[31] L. Greco, R. Romino, I. Coto et al., *Gut* **50** (5), 624 (2002).

[32] Ludvig M.Sollid, Gunnar Markussen, Johan Ek et al., *J Exp Med.* **169** (1), 345 (1989).

[33] S. Yasunaga, A. Kimura, K. Hamaguchi et al., *Tissue Antigens* **47** (1), 37 (1996).

[34] Walburga Dieterich, Tobias Ehnis, Michael Bauer et al., *Nat Med* **3** (7), 797 (1997).

[35] Pål Stenberg, E.Bodil Roth, and Klas Sjöberg, *European Journal of Internal Medicine* **19** (2), 83 (2008).

[36] Shuo-Wang Qiao, Elin Bergseng, Øyvind Molberg et al., *J Immunology* **175** (1), 254 (2005).

[37] Chu-Young Kim, Hanne Quarsten, Elin Bergseng et al., *PNAS* **101** (12), 4175 (2004).

[38] Lars-Egil Fallang, Sujin Roh, Anders Holm et al., *Journal of Immunology* **181** (8), 5451 (2008).

[39] Joseph Sambrook and David Russel, *Molecular Cloning*, 3 ed. (2004).

[40] K. Mullis, F. Faloona, S. Scharf et al., *Biotechnology* **51** (Pt 1), 263 (1986).

[41] New England Biolabs, T4 Ligase, Available at http://www.neb.com/nebecomm/products/productM0202.asp.

[42] Invitrogen, *TOPO TA Cloning Manual*.

[43] Promega, *Wizard® Plus SV Minipreps DNA Purification System* **1** (1) (2008).

[44] Qiagen, *Qiagen Plasmid Plus Purification* **1** (1).

[45] New England Biolabs, Restriction Endonuclease, Available at

http://www.neb.com/nebecomm/products/category1.asp?#2.

46 Qiagen, *Qiagen Spin Handbook* (2008).

47 Stratagene, *QuikChange Multi Site-directed Mutagenesis Kit Manual*. (2005).

48 Clonetech, in *www.clonetech.com* (www.clonetech.com, 2006).

49 Jay P. Morgenstern and Hartmut Land, *Nucleic Acids Research* **18** (12), 3587 (1990).

50 Rieko Sekine, Toshio Kitamura, Takashi Tsuji et al., *International Journal of Hematology* **87** (4), 351 (2008).

51 Howard M. Shapiro, *Practical Flow Cytometry*, 4th ed. (John Wiley & Sons, 2005).

52 Dudley H. Williams and Ian Fleming, *Spectroscopic Methods in Organic Chemistry, 5th ed.* (1995).

53 R. Yates John, *Journal of Mass Spectrometry* **33** (1), 1 (1998).

54 J. A. Todd, J. I. Bell, and H. O. McDevitt, *Nature* **329** (6140), 599 (1987).

55 L. Schenning, D. Larhammar, P. Bill et al., *EMBO J* **3** (2), 447 (1984).

56 H. C. Chang, T. Moriuchi, and J. Silver, *Nature* **305** (5937), 813 (1983).

57 Elin Bergseng, Jiang Xia, Chu-Young Kim et al., *Journal of Biological Chemistry* **280** (23), 21791 (2005).

58 Yvonne van de Wal, Yvonne M. C. Kooy, Jan Woulter Drijfhout et al., *Immunogenetics* **44** (4), 246 (1996).

59 Frode Vartdal, Bente H. Johansen, Thomas Friede et al., *European Journal of Immunology* (1996).

[60] Gilbert J. Kersh, Ellen N. Kersh, Daved H. Fremont et al., *Immunity* **9** (6), 817 (1998).

[61] Frances C. Hall, Joshua D. Rabinowitz, Robert Busch et al., *Eur J Immunol.* **32** (3), 662 (2002).

[62] Sarah E. Henrickson, Thorsten R. Mempel, Irina B. Mazo et al., *Nat Immunol* **9** (3), 282 (2008).