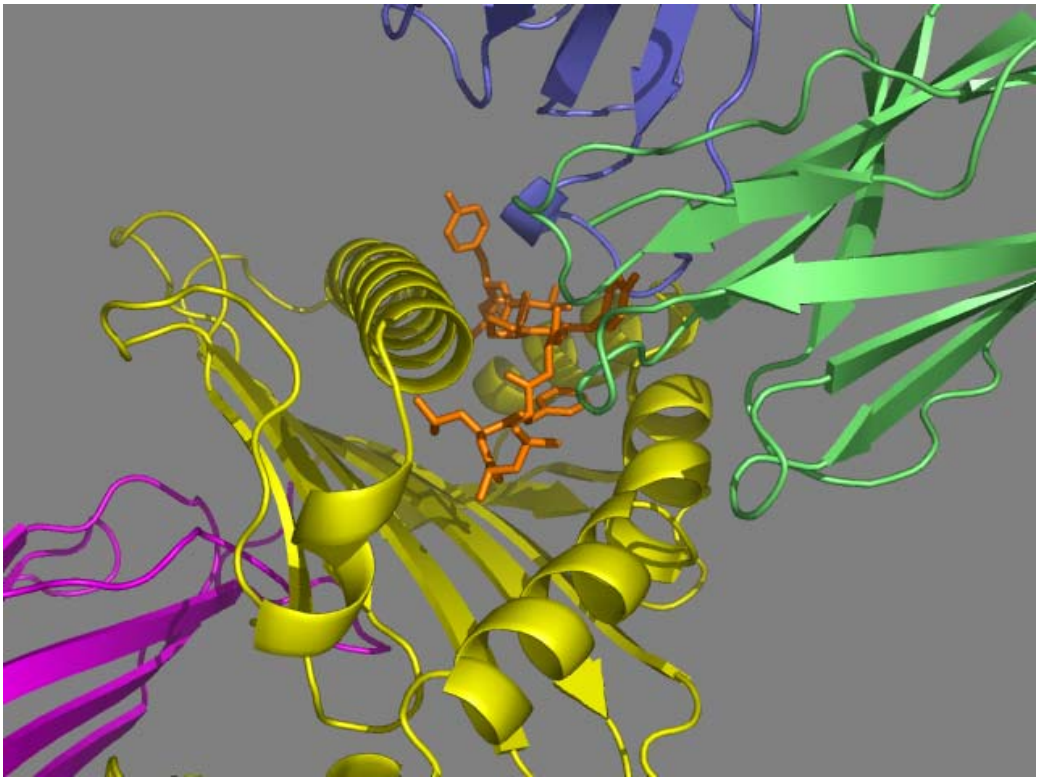
Why do jawed vertebrates have intermediate numbers of MHC molecules? – A modelling approach.



Diplomarbeit

an der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von Benno Wölfing

MAX-PLANCK-INSTITUT FÜR EVOLUTIONSBIOLOGIE, PLÖN

August 2007

Why do jawed vertebrates have intermediate numbers of MHC molecules? – A modelling approach.

Diplomarbeit

an der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von Benno Wölfing

MAX-PLANCK-INSTITUT FÜR EVOLUTIONSBIOLOGIE, PLÖN

August 2007

Referent: Prof. Dr. Manfred Milinski

Max-Planck-Institut für Evolutionsbiologie
August-Thienemann-Str. 2
24306 Plön

Koreferent: Prof. Dr. Thomas Bauer

Zoologisches Institut
Christian-Albrechts-Universität zu Kiel
Am Botanischen Garten 1-9
24118 Kiel

Title picture:

The CDR loops of the T-cell receptor A6 α- (green) and ß-chain (blue) interacting with a peptide-MHC complex. The tax-peptide (orange) lies in the peptide-binding groove formed by two α-helices of the α-chain (yellow) of the MHCI molecule HLA-A2. The ß2-microglobulin is shown in pink (picture was created from structural data in the Protein Data Bank (www.rcsb.org) under accession code 1ao7 using PyMOL version 0.98 software).

Contents

1. Summary

The major histocompatibility complex (MHC) is famous for containing the most polymorphic loci in the genome of jawed vertebrates. Immunological research has centered on the MHC, because a response of the adaptive immune system usually can only be initiated if a MHC molecule binds a peptide derived from a pathogen and if the resulting peptide-MHC complex is recognized by a T-cell. Given that each MHC molecule type can only present peptides that match its peptide-binding groove, it seems advantageous for an individual to express many different MHC molecules so that any foreign peptide can be presented and consequently every pathogen be attacked. Therefore the question arises why each individual only expresses an intermediate number of different MHC molecules that pales into insignificance beside the extensive diversity of MHC alleles at the population level.

In this thesis selection pressures that may set an upper limit to intraindividual MHC diversity are discussed. MHC molecules do not only present foreign peptides but also self-peptides, so that T-cells which recognize self-peptide – MHC complexes have to be eliminated during their maturation in order to avoid autoimmune diseases. Increasing intraindividual MHC diversity is therefore expected to lead to an increased loss of mature T-cells, entailing reduced immunocompetence. While some existing models suggest that this rational can explain why individuals usually do not have more than 20 MHC class I and II alleles, other models conclude that different explanations must be true.

Based on findings of recent experimental studies I develop a new model and conclude that the number of MHC alleles present in individuals may be optimal to balance the advantages of presenting an increased range of peptides and the disadvantages of an increased loss of T-cells.

How the model predictions can be tested on threespined sticklebacks (*Gasterosteus aculeatus*) is shown in the outline of an experimental approach, which may allow to directly measure TCR repertoire diversity before and after negative selection. A better understanding of the stickleback adaptive immune system is a precondition for this experiment. The stickleback thymus – which so far has only been mentioned in a brief communication by Bigaj et al. (1987) – has been clearly identified and characterized in this thesis.

3. Introduction

3.1 The adaptive immune system

3.1.1 An introduction from a theoretical perspective

Systems, whose elements share resources freely and depend on establishing a controlled microenvironment for their survival, are only stable if their integrity is protected from intruders. Protection from intruders in systems as diverse as computer networks, social insect colonies and animal bodies always involves a discrimination task: Harmful elements have to be distinguished from self elements before defence mechanisms can be brought into action. Three different ways to solve the challenge of discriminating between self and foreign elements exist, all of which are employed in the vertebrate immune system (Boehm 2006). First, nonself structures characteristic of groups of pathogens – such as lipopolysaccharides in bacterial membranes – can be detected by "pattern-recognition receptors". Second, self structures are recognized by inhibitory receptors on Natural Killer cells. This ensures that the cytotoxic activity of these cells is only unleashed in the absence of self-ligands. Third, receptors with random specificities can be generated. As they can potentially recognize both nonself and self structures, receptors strongly reactive to self ligands have to be eliminated or functionally inactivated during their maturation. This type of receptor evolved in vertebrates about 500 million years ago after hagfish and lampreys had split from the lineage giving rise to jawed vertebrates (Boehm 2006). Receptors of this type form the mainstay of the adaptive immune system. Their characteristics enable it to mount an immune response against pathogens lacking "typical" molecular patterns, to target its effector mechanisms very specifically at a pathogen and to develop a specific immunological memory (Janeway et al. 2005). Although the adaptive

immune system is a powerful tool to fight pathogens and parasites, its detection system also contains gaps. To illustrate them a brief description of how a response of the adaptive immune system is initiated and how T-cell receptors are screened for self-reactivity during maturation is needed. As research has focused on studies "of mice and men", data pertaining to the mouse immune system is used throughout the modelling part of the thesis unless otherwise stated.

3.1.2 Initiation of a response of the adaptive immune system

The proteasome cuts a subsample of all proteins in the cytosol into peptide fragments which can transit into the endoplasmic reticulum where a subset binds to the peptide-binding groove of major histocompatibility complex (MHC) class I molecules. The peptide-MHC complex is subsequently transported to the cell surface. If a mature CD8+ T-cell recognizes a complex of a foreign peptide and a MHC I molecule and receives costimulatory signals, its cytotoxic activity is unfolded and it starts to proliferate. Similarly extracellular proteins can be taken up by cells in vesicles, be degraded into peptides by cathepsins and finally be loaded onto MHC II molecules. As fighting of extracellular pathogens requires different effector functions, MHCII-peptide complexes are not recognized by CD8+ T-cells but by CD4+ T-cells (Janeway et al. 2005).

3.1.3 αß T-cell maturation

Precursor T-cells arise from hematopoietic stem cells and migrate into the thymus. In the cortex of the thymus the T-cell receptor (TCR) is first expressed after somatic rearrangement of the subgenic fragments at the TCR ß- and TCR α-locus. In a process called positive selection TCRs are tested for general reactivity to the individual's MHC-molecules. Only T-cells with receptors that are functional in the context of the individual's MHC-molecules pass this checkpoint and enter the medulla, which is the

main site for negative selection. During negative selection T-cells, whose receptors react strongly with self-peptide – MHC complexes and which can cause autoimmune diseases if they mature, die by apoptosis, are functionally inactivated or develop into regulatory T-cells that have immunosuppressive activity (Fontenot 2005; Starr et al. 2003; Janeway et al. 2005).

3.1.4 Types of errors in the detection system of the adaptive immune system

Consider a set U of all self and foreign proteins (compare Forrest and Beauchemin 2007). Errors in the discrimination process are unavoidable if a protein is part of the self and foreign subset. If the two subsets are disjoint, errors that are principally avoidable may occur. False positives result if self-reactive T-cells have not been eliminated during negative selection. Since "self" is the default-classification in the adaptive immune system, false negatives may arise due to a variety of reasons: The proteasome or cathepsins may fail to cut out the peptides that would have allowed detection. The individual's MHC molecules may not have binding sites appropriate for the peptides. Even if a nonself peptide is presented by an MHC molecule, it may still stay undetected because no TCR binds to it (either because an appropriate TCR has never been generated or because it was deleted during negative selection). Selection is expected to minimize errors in the detection system. As discussed in the next chapter, there is powerful experimental evidence that selection pressures shape the MHC accordingly.

3.2 Evolutionary Ecology of the MHC

The Major Histocompatibility Complex (MHC) is well known for being the most polymorphic region in vertebrate genomes (e.g. Piertney and Oliver 2006). Variability among alleles at MHC class I and MHC class II

loci is highest in regions that code for amino acids of the peptide binding groove (Janeway et al. 2005). In these regions the rate of nonsynonymous nucleotide substitutions exceeds the rate of synonymous substitutions (Hughes and Nei 1988; Hughes and Nei 1989; metaanalysis by Bernatchez and Landry 2003). This provides strong evidence that the polymorphism among MHC class I and II molecules is not due to neutral drift but is the result of balancing selection. Selection may shape the MHC by favouring individuals whose MHC makes them more resistant to pathogens, less prone to autoimmune diseases and more attractive as mates or proves to be beneficial in the context of kin recognition.

Since MHC molecules play a crucial role in the adaptive immune system – no T-cell dependent immune response can be initiated without presentation of immunogenic peptides by MHC molecules – and since involvement in the immune response is the only postulated function of the MHC which all jawed vertebrates share, the universal selective force operating on the MHC is probably the pressure to maximize the functionality of the immune system (Klein and O'Huigin 1994). The two main theories that seek to explain why MHC diversity at the population level has evolved and is stable – the negative frequency-dependent selection hypothesis and the overdominance hypothesis – are both based on the fact that any given MHC molecule can only present a limited range of peptides and can thus only mediate resistance to a limited number of pathogens (Piertney and Oliver 2006). Indeed many studies demonstrate that resistance to a specific pathogen or parasite is correlated with the presence of certain MHCI and II alleles (Bonneaud et al. 2005, Hill et al. 1991, Langefors et al. 2001, Paterson et al. 1998, Schad et al. 2005, review by Martin and Carrington 2005).

The question arises why the mechanism that ultimately leads to polymorphism at the population level does not generate extensive intraindividual MHC diversity: Given that any MHC molecule can only

mediate resistance to a limited range of pathogens, one would expect selection to favour the expression of many different MHC genes per individual. In reality, however, despite the potential for evolutionary duplication and diversification of individual loci (Lawlor et al. 1990), the intraindividual MHC diversity is surprisingly low compared to the large diversity of MHC alleles at the population level. Individuals from natural populations of polyploid *Xenopus* species do not express genes from more than two major histocompatibility complexes and have silenced – probably by deletion – up to ten major histocompatibility complexes on the other chromosomes, whereas duplicated loci of many other genes have been shown to remain active (Flajnik 1996). Since artificially generated polyploids do express all constituent major histocompatibility complexes, silencing of MHC genes in individuals from natural populations probably occured over an evolutionary time period (Kobel and Du Pasquier 1986). This suggests that selection pressures other than maximizing the presentation probability of pathogens are involved in shaping the MHC and set an upper limit to its diversity.

Ideally experimental studies, that test whether intermediate rather than maximal MHC diversity is optimal, should compare fitness parameters of individuals with different MHC diversities but common (or randomized) genetic background after exposing them to multiple parasites or pathogens. Overrepresentation of certain types of alleles in certain diversity classes can confound studies, especially if individuals are only infected with a single pathogen or parasite species.

Wegner et al. (2003a) determined the number of parasite species found in MHCII ß - typed sticklebacks (*Gasterosteus aculeatus*) belonging to different natural populations. The authors found that residual parasite species numbers of an ANOVA with population as factor were minimal for individuals carrying around five different MHCII ß alleles. To corroborate these observational findings sticklebacks belonging to six

sibships were simultaneously exposed to three common parasite species in a lab study (Wegner et al. 2003b). Parasite load was found to be minimal in individuals carrying about five different MHCII B alleles as well in an analysis that pooled all individuals as in an analysis that considered sibships separately. MHCII-based mate choice in sticklebacks also suggests that intermediate intraindividual MHCII diversity is optimal (Aeschlimann et al. 2003).

Studies on other animals have found associations between fitness parameters of wild-caught individuals and MHC diversity, further supporting the hypothesis that intermediate MHC diversity is optimal. In tropical water pythons (*Liasis fuscus*) individuals of intermediate MHCI diversity had lowest loads of blood parasites and were older and longer (Madsen et al. 2006). Bonneaud et al. (2004) found that female house sparrows (*Passer domesticus*) with an intermediate MHCI diversity laid larger clutches than individuals with lower or higher MHCI allelic diversity. Furthermore the authors tested how individuals responded to an injection of Newcastle disease virus vaccine. Most probably, however, differences in the response to the vaccine are the result of differences in initial body condition and can only be indirectly related to MHCI diversity, since reaction to the vaccine mainly involves MHCII mediated antibody production.

The findings of Harf and Sommer (2004) do not support the hypothesis that intermediate MHC diversity is optimal. Wild-caught hairy-footed gerbils (*Gerbillurus paeba*) with three DRß alleles had significantly higher numbers of helminth eggs in their faeces than individuals with two or four alleles. Other studies (Carrington et al. 1999, Penn et al. 2002, Thursz et al. 1997) support the heterozygote advantage hypothesis.

All in all many studies have found suggestive evidence for an immunogenetic optimum. However, its existence and the physiological

processes that underlie it are still a matter of debate. In my thesis I discuss selection pressures that may set an upper limit to intraindividual MHC diversity. It has been argued that high intraindividual MHC diversity can lead to extensive negative selection of preselection thymocytes resulting in a depletion of the mature T-cell receptor repertoire and reduced immunocompetence (Vidovic and Matzinger 1988). This verbal argument has been formalized by Nowak et al. (1992). Since then numerous models have been published, that either support the depletion hypothesis or suggest that different explanations for the rather low MHC diversity observed in individuals must be considered (Borghans et al. 2003, DeBoer and Perelson 1993, van den Berg and Rand 2003). Based on recent advances in immunology I develop a new model in order to re-evaluate the depletion hypothesis. An outlook on how the model predictions can be tested on threespined sticklebacks (*Gasterosteus aculeatus*) is given.

4. Parameter estimates and scenario to be modelled

4.1 Quantitative estimates of positive selection efficiency

In order to model the efficiency of positive selection in individuals of different MHC diversity, it is necessary to know (1) which proportion of preselection thymocytes an average MHC molecule can positively select and (2) how much the fraction of thymocytes positively selected by an additional MHC molecule overlaps with the fraction of thymocytes that has already been positively selected.

To deal with the first challenge activation markers have been widely (Merkenschlager et al. 1997, Monteiro et al. 2005) and sometimes wildly (Boehm, personal communication) used. The cell surface level of activation markers like e.g. CD69 is known to be transiently upregulated by preselection thymocytes upon interaction with positively selecting peptide-MHC complexes (Hare et al. 1999, Feng et al. 2002), so that positively selected thymocytes can in principal be discriminated from those that die by neglect on the basis of CD69 expression.

To test if CD69 is a reliable marker for thymocytes that are being positively selected, thymuses of (OT-I TCR transgenic/RAG-1°/ß2m°) mice only capable of expressing a single TCR, for which the mice do not naturally express positively selecting ligands, were cut into fragments and cultured. Presentation of defined peptide-MHC complexes was restored by addition of ß2 microglobulin and peptides to the organ culture. While 73.6% of all double positive (DP) thymocytes upregulated CD69 upon interaction with peptide-MHC complexes known to induce negative selection, 51.3% and 28.4% of all DP thymocytes were stimulated to express CD69 following exposure to two different positively selecting ligands and less than 2% of DP thymocytes were grouped as CD69+

when exposed to nonselecting ligands (Merkenschlager et al. 1997). Clearly thymocytes that die of neglect do not express CD69.

Negatively selecting ligands seem to be more potent inducers of CD69 upregulation than positively selecting ligands (this result was also found by Daniels et al. (2006) for low ligand concentrations). This can obviously be due to the fact that (1) negatively selecting ligands induce stronger upregulation of CD69, so that more thymocytes are grouped as CD69+ or that (2) negatively selecting ligands are capable of positively selecting more thymocytes in the organ culture. While option (1) implies that the percentage of CD69+ cells may increase although positive selection efficiency does not, option (2) alludes to another problem of measuring positive selection efficiencies in organ cultures. This is the question if the situation in the organ culture accurately reflects that in nature, that is in this case if the large proportion of DP thymocytes that die by neglect in the organ culture (as judged by CD69) although a suitable positively selecting ligand is present would also die under natural conditions.

Studies of freshly extracted thymuses that only determine the percentage of DP thymocytes that are CD69+ and not absolute thymocyte numbers (Monteiro et al. 2005) cannot demonstrate that the analyzed subsample represents the composition of the thymocyte populations in vivo since thymocytes strongly interacting with stromal cells are more difficult to extract and may therefore be underrepresented in the analysis. The fact that proportions of CD69+ thymocytes and thymocytes positive for other activation markers (e.g. CD5) can deviate considerably in the same experimental setup (e.g. 10.1 ± 2.5 CD69+ n = 15 and 18.5 ± 2.9 CD5+ n = 6 in cultures of MHC naive thymocytes with H-2b stroma; Merkenschlager et al. 1997) illustrates, that the proportion of positively selected cells cannot be simply taken to be equal to the proportion of activation marker positive cells. Nevertheless these studies give an order

of magnitude for positive selection efficiency: About 20% of preselection thymocytes appear to be positively selected in MHC heterozygous mice (Merkenschlager et al. 1997, Monteiro et al. 2005).

A different approach to assess the reactivity of the preselection TCR repertoire is to induce maturation of preselection thymocytes and analyze the reactivity of the corresponding mature T-cells. Zerrahn et al. (1997) used a combination of anti-TCR αß and anti-CD4 monoclonal antibodies to induce maturation of DP thymocytes harvested from mice deficient in MHC I and II. The authors found that 4.4% of the hybridomas obtained from antibody-selected T-cells were reactive against stimulator cells expressing a given MHC II molecule. Thus in mice expressing a single MHC molecule this fraction of the preselection TCR repertoire is thought to be subject to negative selection. Based on this rational the authors estimate that 20 – 30% of preselection thymocytes are deleted during negative selection in MHC heterozygous animals. Since the set of negatively selected thymocytes is a subset of the set of positively selected thymocytes, at least 20 – 30% of preselection thymocytes can be expected to be positively selected.

To estimate the quantitative impact of MHC diversity on positive selection efficiency, it is useful to note that interactions between the TCR and a peptide-MHC complex can formally be classified into interactions with the peptide, contributions by bindings between the TCR and residues unique to an MHC molecule type and contributions by bindings between the TCR and conserved features all MHC molecules share. If interactions with conserved features of MHC molecules play a key role in the process of positive selection, repertoires positively selected by different MHC molecules will be largely overlapping. Non-overlapping repertoires would result if unique features of MHC molecules and bound peptides were decisive.

Studies based on activation markers found evidence for an increase of positive selection efficiency with MHC diversity (Merkenschlager et al. 1997, Monteiro et al. 2005). However, as outlined above, the increase in the fraction of CD69+ cells can simply reflect increasing numbers of strong interactions between self-peptide-MHC complexes and thymocytes and thus an increase in the proportion of thymocytes that succumb to negative selection.

Studies analysing the TCR repertoire of mice expressing a single peptide-MHCII complex found considerable numbers (20 – 50% of wildtype) of polyclonal CD4 single positive thymocytes (that express the full range of Vß – segments) despite the extreme reduction in the diversity of positively selecting ligands (Fung-Leung et al. 1996, Ignatowicz et al. 1996, Martin et al. 1996). This suggests that, although positive selection of some TCRs is known to depend on the presence of special peptide-MHC complexes (e.g. Santori et al. 2002), interaction with conserved features shared by all/most MHC molecules is sufficient to induce positive selection in a large proportion of thymocytes.

A study by Huseby et al. (2005) strongly supports this view. The authors assessed the TCR repertoire prior to negative selection by using a system in which thymocytes were exposed to less stringent negative selection (I A^b - SP mice). TCRs from many mature T-cells in these mice were very crossreactive and some TCRs even interacted with peptide-MHC complexes of both class I and class II MHC molecules. These findings contradict the traditional view that nearly all preselection TCRs are highly specific and that positive selection picks out the few TCRs that are specific to the individual's MHC molecules. Instead the gene segments for the TCR alpha and beta chains, which are encoded in the germline, are hypothesized to give rise to an inherent high affinity of TCRs to MHC molecules (even more than already proposed by Jerne 1971). This high default-affinity is thought to be based on conserved interactions, e.g. with

the backbone atoms of the MHC α-helices flanking the peptide-binding groove. The fundamental contribution of conserved interactions is expected to allow positive selection of a broad spectrum of T-cells by a single peptide-MHC complex (as observed, Huseby et al. 2005). Since positive selection is mainly mediated by conserved interactions, many of these positively selected TCRs are expected to be very crossreactive (as observed, Huseby et al. 2005). In this scenario it is extensive negative selection (Huseby et al. 2003) that ensures peptide and MHC specificity of the mature TCR repertoire by deleting crossreactive T-cells.

All in all the following tentative picture, which serves as a basis for the model, emerges. Around 20% of all preselection thymocytes are positively selected. As interactions with conserved features shared by all/many MHC molecules play a key role in positive selection, a substantial increase in positive selection efficiency with MHC diversity is not expected. Many of the remaining 80% of preselection thymocytes have produced TCR in the semirandom rearrangement process that are not functional in the context of any MHC molecule. These thymocytes die of neglect in all animals irrespective of their MHC diversity.

4.2 Quantifying negative selection

A simple calculation can give a rough estimate of how many clonotypes need to be functionally removed from the reperoire in order to ensure tolerance to self. Since about 10^3 to 10^4 different self peptides can be eluted from a given MHC molecule (Hunt et al. 1992) and since a mature T-cell typically recognizes 1 in 10^5 to 10^6 peptide-MHC complexes (Zinkernagel 1996), we arrive at the following probability for a positively selected thymocyte to be removed from the functional repertoire upon interaction with complexes of one MHC molecule type:

$$P\{\text{“reactive to at least 1 self-peptide”}\} = 1-(1-10^{-5})^{10^4} \approx 1\% \text{ to } 10\%$$

This calculation is quite sensitive to the rough estimates given above and assumes that thymocytes that have not yet passed negative selection have TCRs of the same specificity (10^{-5} to 10^{-6}) as mature T-cells. It is unlikely that this is the case since preselection thymocytes with broad specificities can be expected to be particularly susceptible to negative selection and experimental evidence suggests that thymocytes are more sensitive to negative selection than mature cells are to activation (Davey et al. 1998, Lucas et al. 1999). For decades immunologists have tried to get direct estimates of the percentage of thymocytes that are deleted in the process of negative selection.

In radiation bone marrow chimeras radioresistant thymic epithelial cells (the main mediators of positive selection) express host MHC, whereas the host's radiosensitive antigen presenting cells of hematopoietic origin (responsible for the bulk of negative selection) have been destroyed and are replaced by cells arising from the donor's hematopoietic stem cells. Van Meerwijk et al. (1997) employed C57BL/6 mice (H-2^b haplotype = wildtype) with various MHC deficiencies to construct bone marrow chimeras. The authors demonstrate a 1.7±0.2 - fold increase in the steady-state percentage of CD4 single positive (SP) thymocytes in MHCII° (bone marrow) -> MHCI° (host) chimeras compared to MHCII+ -> MHCI° chimeras. Since the total cell number in thymuses from the chimeras was similar, this reflects an increase in absolute numbers of CD4 SP cells. These findings were corroborated by analysing the kinetics of T-cell development in MHCI°II° -> wildtype versus MHCI° -> wildtype and wildtype -> wildtype chimeras. In chimeras not expressing MHCII on hematopoietic elements the generation of CD4 SP thymocytes is reported to be approximately twofold greater. The authors also claim to see a two- to threefold increase in generation of CD8 SP thymocytes in

mice not expressing MHCI on antigen presenting cells of hematopoietic origin. Since C57BL/6 mice express a single MHCII molecule (I-A^b), it can be concluded that around 50% of the thymocytes positively selected by I-A^b get negatively selected by peptide-MHC complexes composed of this MHC molecule. The kinetic measurements with wildtype individuals (expressing 1 MHCII and 3 MHCI molecules) indicate that a slightly larger percentage of positively selected cells succumb to negative selection in these individuals.

Laufer et al. (1996) found that 5% of the mature CD4 T-cells in MHCII° mice, that were tailored to re-express I-A^b only on thymic cortical epithelium, were markedly reactive to antigen presenting cells presenting self-peptides bound to I-A^b. The authors conclude that at least 5% (but possibly more) of the positively selected thymocytes must undergo negative selection.

Zerrahn et al. (1997) substituted the receptor engagement responsible for normal positive selection with binding of anti-TCR and anti-CD4 antibodies and found that more than 4.4% of the T-cell lineages in the mature repertoire were reactive against a given MHCII molecule. If antibodies induce positive selection in all thymocytes and normal positive selection efficiency by one MHCII molecule is 20% of the preselection repertoire, these results imply that more than 20% of the positively selected thymocytes must succumb to negative selection. The authors estimate that in heterozygous individuals 20 - 30% of preselection thymocytes are negatively selected, that is almost all positively selected thymocytes (provided that the hypothesis that around 20% of preselection thymocytes are positively selected irrespective of MHC diversity holds, see chapter on positive selection).

As outlined in the chapter on positive selection a large polyclonal mature TCR repertoire is found in mice expressing a single peptide-MHCII

complex. Researches studying different mouse models of this type consistently found that an overwhelming fraction (65 – 80%) of CD4+ T-cells that matured under these conditions of reduced negative selection is reactive to the same MHC molecule in complex with the normal diversity of self-peptides (Ignatowicz et al. 1996 demonstrated that 65% of the hybridomas established from the mature T-cells were stimulated to upregulate interleukin-2 production; Surh et al. 1997 found that 75% of the mature CD4+ T-cells were autoreactive by the blood to lymph recirculation method). Moreover in bone marrow chimeras that present a single peptide-MHC complex on cortical thymic epithelial cells but a range of different peptides bound to the same MHC molecule in the medulla, the same proportion died due to negative selection (Tourne et al. 1997, Surh et al. 1997).

This illustrates the power of negative selection, but note that the percentage of positively selected thymocytes that get negatively selected can be different in mice whose TCRs are positively selected by a diversity of peptide-MHC complexes. On the one hand the positively selected TCR repertoire in single peptide-MHC mice is not complete and may well be particularly rich in broadly reactive thymocytes that are particularly susceptible to negative selection. If positively selecting MHC molecules in the cortex are diverse, we may thus expect a lower percentage of positively selected thymocytes that succumb to negative selection by one MHC molecule. On the other hand thymocytes in single-peptide-MHC mice have undergone some – albeit greatly reduced – negative selection. True power of negative selection may thus be even greater. In fact, based on the observation that the repertoire of T-cells reactive to a certain peptide-MHCII complex was less peptide-degenerate if the T-cells had undergone negative selection by a single peptide bound to the same MHCII molecule than if the T-cells had undergone negative selection on allogeneic MHCII molecules, Huseby et al. (2003) estimate that up to 90% of all thymocytes that are positively selected in mice

expressing a single peptide-MHCII complex would be deleted in mice expressing the same MHC molecule in complex with a variety of self-peptides.

Apart from deletion other processes are known that lead to a removal of clonotypes from the functional repertoire. Their effect cannot be assessed by the studies described above and is hard to quantify. Thymocytes with an increased affinity of TCRs for self-peptide-MHCII complexes can develop into regulatory T-cells, which absorb activating cytokines and secrete immunosuppresive cytokines upon binding to peptide-MHC complexes (Kronenberg and Rudensky 2005). Regulatory T-cells are thus not only removed from the functional repertoire themselves, but they can also suppress activation of responder T-cells in the periphery. The fundamental role of regulatory T-cells in controlling immune responses can be seen from the fact that CD4+ T-cell populations from which CD25+ CD4+ regulatory T-cells have been removed cause autoimmune diseases after adoptive transfer into lymphopenic hosts (Sakaguchi et al. 1995). CD25+ CD4+ regulatory T-cells constitute 2 – 10% of total CD4+ T-cells in the periphery (Kronenberg and Rudensky 2005). Moreover, mechanisms of self-tolerance induction in the periphery include deletion and induction of anergy (Steinman and Nussenzweig 2002).

Altogether it is unclear what proportion of positively selected thymocytes is subsequently removed from the TCR repertoire due to self-reactivity. A minimum of 20% (based on the study by Zerrahn et al. 1997) to 25% (based on the observation that more than 50% of all positively selected thymocytes committed to the CD4+ lineage succumb to negative selection by one MHCII molecule, van Meerwijk et al. 1997) seem to be negatively selected by one MHC molecule. However, many studies suggest that the actual percentage may be much higher (up to 90%, see above). As negative selection is highly peptide and MHC specific (see e.g. the large proportion of positively selected thymocytes in single-

peptide-MHC mice, that is negatively selected in mice presenting a variety of self-peptides on the same MHC molecule) its effect is expected to increase with MHC diversity. I model negative selection by several MHC molecules as subsequent independent events.

4.3 Estimates for other parameters

The probability that a given peptide which has been produced by the proteasome / the cathepsins can bind to a given MHC molecule is estimated to be 2% (Kast et al. 1994). One T-cell in 10^5 to 10^6 of the naive repertoire recognizes a given peptide-MHC complex (Zinkernagel 1996). The size of the preselection TCR repertoire that gives rise to the mature repertoire in a young mouse can roughly be estimated from the number of T-cell clones in the C57BL/6 mouse spleen ($2 \cdot 10^6$ clones of about 10 cells each, Casrouge et al. 2000), the total number of T-lymphocytes in mice ($1 \cdot 10^8$, Casrouge et al. 2000) and the fact that in the order of 1% of preselection thymocytes enter the mature repertoire. It follows that the preselection repertoire consists of approximately 10^9 TCRs.

5. The model

5.1 Model development

Let R denote the number of clonotypes in the mature repertoire and $R_0 = 10^9$ be the size of the preselection TCR repertoire that gives rise to R. Since the efficiency of positive selection is assumed to be 20% and independent of MHC diversity, the size of the repertoire after positive selection is given by $0.2 \cdot R_0$. Let n denote the probability that a positively selected T-cell gets negatively selected by a given MHC molecule. In order to enter the mature repertoire R, clones need to avoid negative selection by all of the M MHC-molecules of the individual (negative selection is evaluated on all MHC molecules, because thymocytes are committed to the CD4 or CD8 lineage depending on the class of MHC molecule that mediates their positive selection, Janeway et al. 2005). It follows that

$$R(M) = R_0 \cdot 0.2 \cdot (1 - n)^M \qquad \text{[eq 1]}$$

with $R_0 = 10^9$ and n in the range of 0.25 to 0.9.

To calculate the probability P_i that an immune response against a given pathogen, from which e different peptides are generated, is initiated, a reasoning similar to that of previous models is used (Nowak et al. 1992, Borghans et al. 1993, DeBoer et al. 1993):

$P_i = 1 - P\{\text{“pathogen escapes detection”}\} =$
$= 1 - [P\{\text{“peptide not presented by given MHC molecule”}\} +$
$P\{\text{“peptide presented by given MHC molecule”}\} \cdot$
$P\{\text{“complex not recognized by any TCR”}\}]^{M \cdot e}$

Let q = 0.02 denote the probability that a given peptide is presented by a given MHC molecule type and $r = 10^{-5}$ denote the probability that a given TCR recognizes a given peptide-MHC complex. Then $P_i(M)$ is given by:

$$P_i(M) = 1 - [1 - q + q(1 - r)^R]^{eM} \qquad \text{[eq 2]}$$
$$\text{with } R(M) = R_0 \cdot 0.2 \cdot (1 - n)^M$$

5.2 Evaluation of the model

[eq 1] predicts a rapid decrease of the size of the mature TCR repertoire with increasing MHC diversity (fig. 1). The assumptions underlying the graph for n = 0.9 (positive selection efficiency totally independent of MHC diversity, outcome of negative selection by given MHC molecule is independent of outcome of negative selection by other MHC molecules and very stringent negative selection by every single MHC molecule) imply too stringent negative selection and do not reflect the natural situation accurately, since individuals with seven different MHC molecules obviously still have a polyclonal TCR repertoire.

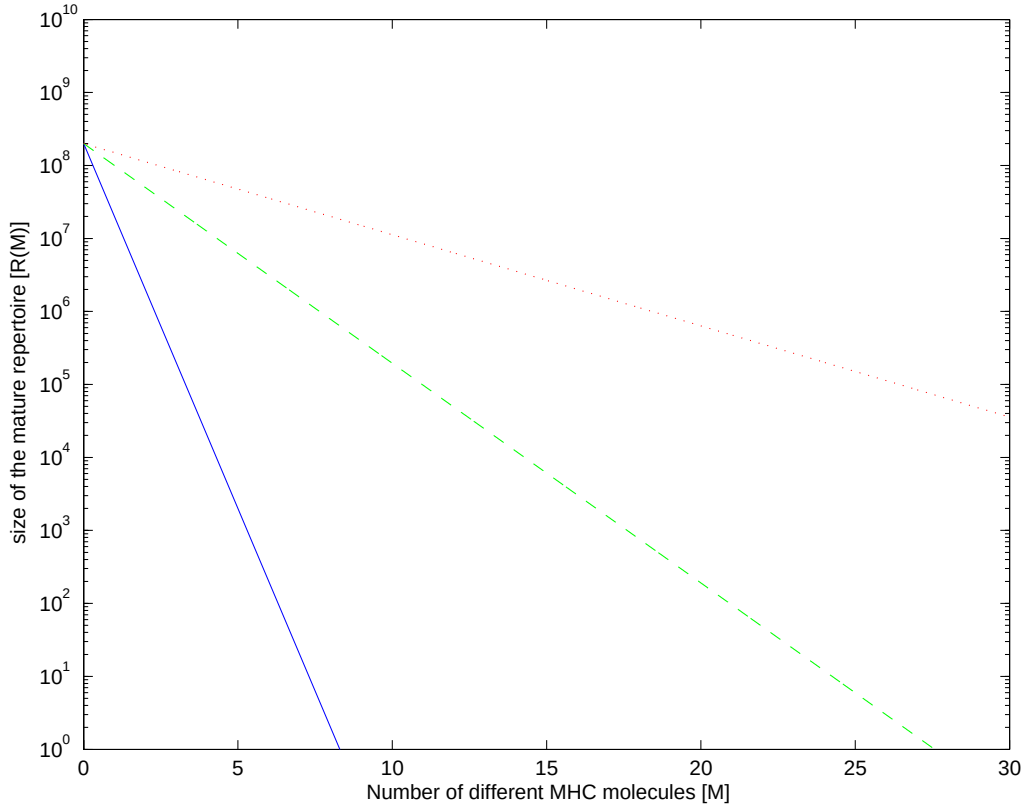


Fig. 1: The exponential decrease of the size of the mature TCR repertoire R (given by [eq 1] with $R_0 = 10^9$) with increasing intraindividual MHC diversity M depicted for three scenarios with different stringencies of negative selection (red, dotted curve: n = 0.25; green, dashed curve: n = 0.5; blue, solid curve: n = 0.9)

Studies on the kinetics of T-cell development can be used to test if the model predictions based on n = 0.5 and n = 0.25 are in the right order of magnitude. Egerton et al. (1990) estimate that in CBA CaH WEHI mice (6 MHC molecules) 3% of the thymocytes produced in the thymus reach the periphery. Given that thymocytes can divide during their maturation, the fraction of clones reaching the periphery lies between f = 3% and f = $2^{-6} \cdot 3\% = 0.047\%$ (Detours et al. 2000) corresponding to $R(6) = 3 \cdot 10^7$ and $R(6) = 5 \cdot 10^5$. Scollay et al. (1980) report that 5% of preselection thymocytes in (BALB/c x C57BL/6)F1 mice entered the periphery. In heterozygous individuals the number of functional MHC molecules is not simply given by the sum of the MHC molecules encoded by the two haplotypes, since MHCII molecules are heterodimers, which can form between alpha and beta chains of different haplotypes and different individual loci within a haplotype. However, not all theoretically possible

combinations form functional heterodimers and some are only expressed at low levels (for the used (BALB/c x C57BL/6)F1 mice see Mineta et al. 1991). Borghans et al. (2003) give an MHC diversity of 16 molecules for (BALB/c x C57BL/6)F1 mice. The estimates correspond to $R(16) = 5\% \cdot 10^9 = 5 \cdot 10^7$ and $R(16) = 5\% \cdot 2^{-6} \cdot 10^9 = 7.8 \cdot 10^5$. The model predictions for n = 0.25 and n = 0.5 do not differ markedly from these rough estimates.

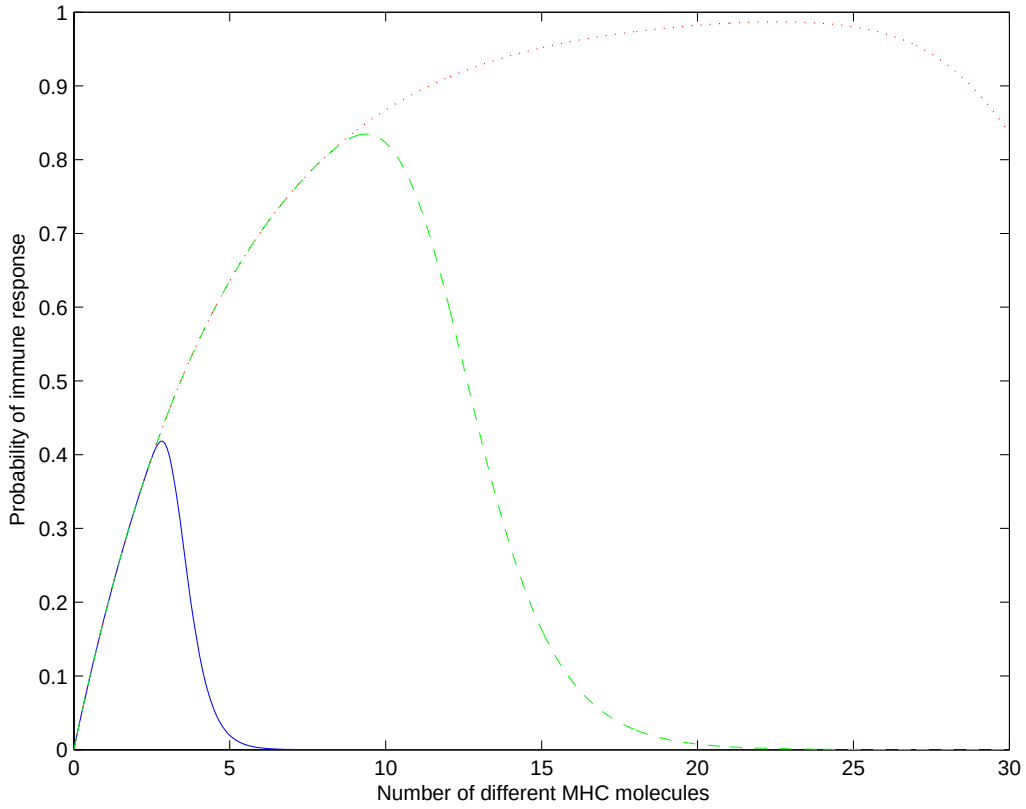


Fig. 2: The probability P_i (given by [eq 2] that a response of the adaptive immune system against a pathogen is initiated plotted as a function of the intraindividual MHC diversity M for three scenarios with different stringencies of negative selection (red, dotted curve: n = 0.25; green, dashed curve: n = 0.5; blue, solid curve: n = 0.9). Parameter values in [eq 2] are q = 0.02; $r = 10^{-5}$; e = 10; $R_0 = 10^9$

For all estimates of n, the probability P_i of mounting an immune response against a given pathogen reaches a maximum at a MHC diversity $M_{opt} < 30$ (fig. 2). The increase of P_i at low MHC diversity is due to the increase in the probability that peptides of a given pathogen get presented by at least one MHC molecule, while TCR repertoire depletion by extensive negative selection accounts for the rapid decrease of P_i at MHC

diversities exceeding M_{opt}. Individuals with an MHC diversity of M_{opt} are predicted to have the highest immunocompetence. Note that the values of M_{opt} for n = 0.25 and n = 0.5 are in the range of 10 to 25 MHC molecules, which is the same order of magnitude as that typically observed in individuals. The model thus suggests that the observed intermediate MHC diversity in individuals is the result of a trade-off between the benefit of a high antigen presentation probability which requires high MHC diversities and the necessity to avoid extensive negative selection.

5.3 Comparison with other models

5.3.1 Model by Nowak et al. (1992)

The only major difference between the model published by Nowak et al. (1992) and the model just presented is that Nowak et al. assume that positive selection increases with MHC diversity. Let h denote the fraction of thymocytes that is positively selected by a single MHC molecule. The other parameters are as described in section 5.1. The size of the mature repertoire R is then given by:

$$R(M) = R_0 \cdot [1 - (1 - h)^M] \cdot (1 - n)^M$$

Borghans et al. (2003) criticized that the conditional probability n that a positively selected thymocyte gets negatively selected by a given MHC molecule (n = P{“negatively selected by given MHC molecule”} / P{“positively selected by any MHC molecule”}) is assumed to be constant (and does not decrease with increasing M and thereby increasing P{“positively selected by any MHC molecule”}). They argue that the assumption implies that a thymocyte that is only positively selected by an additional MHC molecule (and not by the set of original MHC molecules) can nevertheless still be negatively selected by the original MHC molecules, which is physiologically unlikely. In my model I

assume that positive selection efficiency does not increase with MHC diversity. If this assumption is correct, n is constant and the mature repertoire size decreases even more rapidly than in Nowak et al.'s model, because no additional thymocytes get positively selected with increasing MHC diversity in my model.

Nowak et al. conclude that MHC diversities found in individuals may have intermediate values, because intermediate MHC diversity conveys maximal immunocompetence. Thus the general predictions of their model are the same as mine, although they had to base their model on an even scarcer amount of quantitative estimates.

5.3.2 Model by Borghans et al. (2003)

In order to develop a model which avoids conditional probabilities, Borghans et al. (2003) start out from the following rational. The number of clonotypes in the mature repertoire R is given by the size of the repertoire after positive selection minus the number of clonotypes that get negatively selected:

$R = R_0 \cdot P\{\text{“a preselection thymocyte gets positively selected”}\} -$
$R_0 \cdot P\{\text{“a preselection thymocyte gets negatively selected”}\}$

Let p and n* denote the unconditional probabilities that a preselection thymocyte gets positively and negatively selected by one MHC molecule. M signifies the number of MHC molecules. R is then given by:

$$R = R_0\left[\left(1-(1-p)^M\right)-\left(1-(1-n^*)^M\right)\right] = R_0\left[(1-n^*)^M-(1-p)^M\right]$$

The model by Borghans et al. (2003) differs from my approach in two main aspects: (1) Borghans et al. treat positive selection by each MHC

molecule as an event independent of selection outcome on other MHC molecules and (2) they use different estimates for the values of crucial parameters.

Aspect (1) is equivalent to assuming a strong increase in positive selection efficiency with MHC diversity. In their main model the percentage of positively selected thymocytes asymptotically approaches 100% of the preselection repertoire. It is unlikely that all semirandomly generated TCR can interact with MHC molecules even if MHC diversity is great (see chapter 4.1). As argued above (chapter 4.1) a model in which positive selection efficiency is independent of MHC diversity may reflect the situation found in nature more adequately.

The parameter values of n and p strongly influence the position of the maximum of $P_i(M)$, which is found at MHC diversities M > 1500 with the authors' estimates. Their estimates of p and n are based on van Meerwijk et al.'s finding (1997) that more than 50% of all positively selected thymocytes succumb to negative selection and evidence from two kinetic studies (Egerton et al. 1990, Scollay et al. 1980) that 3% and 5% of all preselection thymocytes reach the periphery in mice with M = 6 and M = 16 MHC molecules respectively. Uncertainties of the kinetic estimates have been outlined in chapter 5.2. Based on van Meerwijk et al.'s (1997) estimate, Borghans et al. (2003) deduce, that the proportion of preselection thymocytes that does not survive positive selection is given by $1-2\frac{R}{R_0}$. The value of R/R_0 is then obtained from the kinetic studies, so that the value of p can be solved from $1-2\frac{R}{R_0}=(1-p)^M$. This implies that around 50% of the positively selected thymocytes die during negative selection not only in mice with 4 different MHC molecules (where this has been demonstrated by van Meerwijk et al. 1997) but also in mice with 6 and 16 different MHC molecules. Note that this

approximation rests on the assumption that pM << 1 for M ≤ 16 (so that $\frac{1-(1-n)^M}{1-(1-p)^M} \approx \frac{n \cdot M}{p \cdot M} = const = 50\%$). It is not obvious from experimental data that the assumption pM << 1 for M ≤ 16 holds. The value of n is estimated analogously and rests on the same assumptions.

The combined effect of the assumption that positive selection efficiency increases strongly with MHC diversity and the procedure of obtaining parameter estimates is very low values for p (= 0.01) and n (= 0.005). Many experimental studies found that a single MHC molecule can positively select a percentage of preselection thymocytes far exceeding the estimated p = 1% (Merkenschlager et al. 1997 found that 10% of preselection thymocytes upregulated activation markers when exposed to peptide-MHC complexes composed of the MHC molecule A^b) and n = 0.5% (Zerrahn et al. 1997 found that 4.4% of antibody-selected thymocytes (see chapter 4.1) were reactive against stimulator cells expressing a single MHCII molecule).

The result of the extremely low value of p is, that the first 20 MHC molecules select almost non-overlapping parts of the preselection repertoire. The same reasoning applies to n, so that positive and negative selection efficiencies grow additively with MHC diversity for M ≤ 20.
The fact that positive selection efficiency is strongly underestimated for low MHC diversities and the fact that its increase with MHC diversity is strongly overestimated are responsible for the size of the mature repertoire and $P_i(M)$ reaching their maximum at values that far exceed natural MHC diversities.

The model by Borghans et al. (2003) predicts that the response chance $P_i(M)$ saturates at MHC diversities of about 10 < M < 20 (if e = 25 different peptides are generated from a pathogen). Given that many experimental studies found associations between the presence of specific

MHC alleles and resistance to certain pathogens (see 3.2), it seems unlikely that the model prediction is true.

5.3.3 Model by De Boer and Perelson (1993)

De Boer and Perelson (1993) do not consider positive selection at all. This means that they implicitly assume that positive selection efficiency does not change with MHC diversity. So their basic model is almost identical to the model presented here, although the authors are primarily interested in deriving an equation that gives the size of the TCR repertoire prior to negative selection, that is required to allow detection of pathogens with a certain minimal probability. The authors conclude that there is an optimal number of MHC molecules, if TCR specificity is a fixed parameter (as it is in my model: $r = 10^{-5}$). By contrast no optimal number of MHC molecules exists, if the TCR specificity is allowed to vary, so that it has an optimal value for any MHC diversity. It is likely that physiological constraints do not allow TCR specificity to vary extensively.

6. Other processes influenced by MHC diversity

6.1 Risk of Autoimmune diseases

In humans correlations between susceptibility to certain autoimmune diseases and presence of certain MHC alleles in an individual's genotype are well established (Jones et al. 2006). For example type 1 diabetes is associated with HLA-DQB1*0302 (Todd et al. 1987), more than 90% of patients with severe rheumatoid arthritis have the HLA-DRB1*0401, HLA-DRB1*0404 allele and/or HLA-DRB1*0101 allele (Wordsworth et al. 1989) and HLA-DRB1*1501 predisposes for multiple sclerosis (Oksenberg et al. 2004). Because of the extensive linkage disequilibrium in the MHC it has only recently become clear that in most cases the MHC alleles themselves (and not linked genes) are functionally involved in the increased susceptibility to autoimmune disorders (Jones et al. 2006).

This suggests that individuals with a diverse set of MHC alleles should have a greater risk of suffering from autoimmune diseases (Borghans et al. 2003). However, an increase in MHC diversity may also increase the likelihood that self-reactive T cells are eliminated or inactivated during their maturation. Indeed MHC alleles mediating dominant protection have been identified (Jones et al. 2006). For example humans carrying the MHCII molecule encoded by HLA-DQA1*0102/HLA-DQB1*0602 are almost completely protected against diabetes even in the presence of the predisposing allele HLA-DQB1*0302 (Siebold et al. 2004). Structural studies comparing the predisposing and protecting MHC molecules suggest that differences in the P6 and P9 pockets may lead to the presentation of an expanded peptide repertoire by the protecting MHC molecule. The authors give suggestive evidence that the protecting MHC molecule causes selection of regulatory T cells, which prevent autoimmune-mediated destruction of pancreatic ß-cells.

Assume that the presence of alleles predisposing for a certain autoimmune disease leads to an e.g. 80% chance of getting the disease, if no protecting alleles (which offer full protection) are present. Let a denote the frequency of predisposing alleles and s denote the frequency of protecting alleles. The probability P_a to develop the autoimmune disease can then be written as a function of the number of MHC molecules M:

$$P_a(M) = 0.8 \cdot \sum_{k=1}^{M} \binom{M}{k} \cdot a^k \cdot (1-a-s)^{(M-k)} \qquad \text{[eq3]}$$

Observe that the risk of autoimmune diseases only increases steadily with MHC diversity (in the relevant range of MHC-diversities, that is M < 30) if the frequency of protecting alleles is low (fig. 3). Moreover, in natural populations selection against alleles mediating autoimmune diseases that break out at an early age may be strong, so that the frequencies of these alleles are low and possibly do not constitute a selection pressure strong enough to lead to an MHC optimum by themselves. Alleles of autoimmune diseases that only become manifest after the main phase of reproduction are not strongly selected against, but have no power to select for an MHC optimum either. Altogether, it can not be taken for granted that the risk of autoimmune diseases increases with MHC diversity and selects for an optimum.

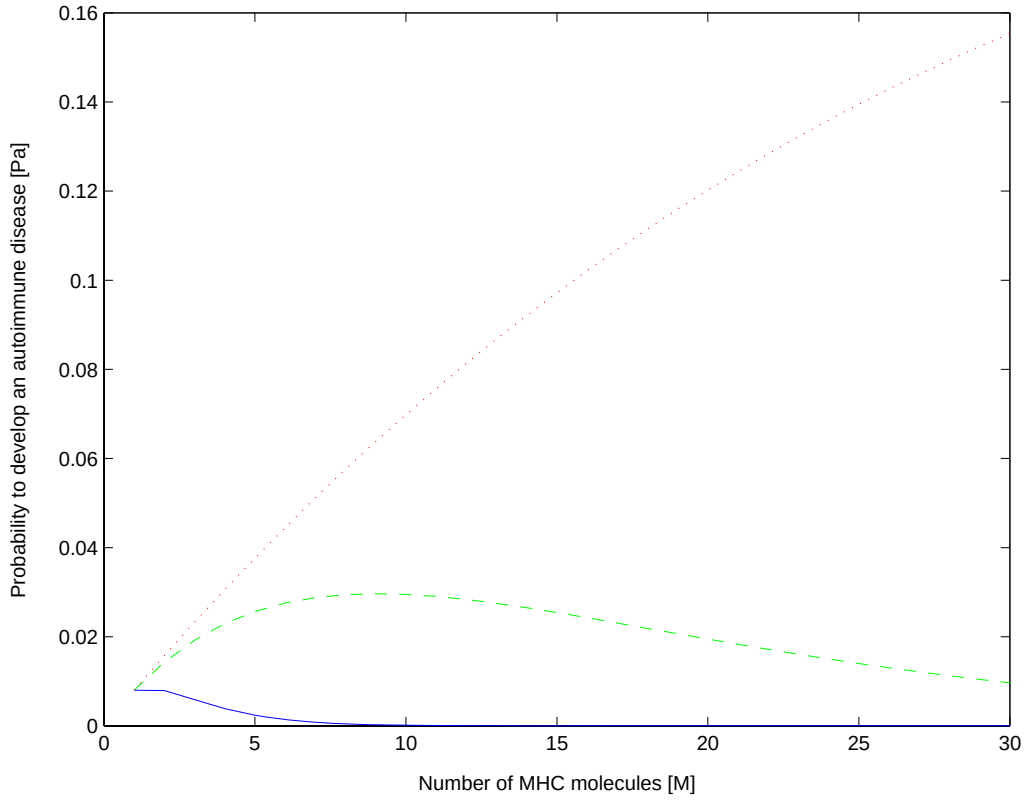


Fig. 3: The probability P_a to develop a certain autoimmune disease is plotted as a function of intraindividual MHC diversity M for three different frequencies of protecting alleles. Parameters in [eq 3] are a = 0.01 and s = 0.5 (blue, solid curve); s = 0.1 (green, dashed curve) and s = 0.01(red, dotted curve)

6.2 Level of antigen presentation

The activation of mature T-cells by antigen presenting cells is a poorly understood and complex process involving the formation of a highly organized and stable contact zone – the immunological synapse - with an outer ring of adhesion molecules and an inner cluster of TCRs with their co-receptors (Janeway et al. 2005). Ultimately activation has to depend on a TCR recognizing - that is binding strongly and for a prolonged time period to - its epitope. Structural analyses have found no differences in the conformation of a TCR that is bound to either agonistic or antagonistic peptide-MHC complexes, whereas correlations between the half-life of the TCR-peptide-MHC complex and T-cell activation have been reported (reviewed in Krogsgaard and Davis 2005). Therefore structural rearrangements are thought to have only a minor influence on

T-cell activation, while the main signal the T-cell receives seems to be the dissociation rates of its TCRs (Krogsgaard and Davis 2005). However, dissociation of the TCR from peptide-MHC complexes is inherently a stochastic process, because even low affinity ligands can bind to TCRs for prolonged time periods albeit with reduced likelihood (George et al. 2005). A T-cell thus faces the challenge to convert the analogue input of all the interactions of its TCRs (with stochastically distributed binding times) into a digital output (activation or ignorance to the encountered antigen presenting cell).

Van den Berg and Rand (2003) developed a model, which is based on the assumption that each TCR is triggered to produce some intracellular signal if it binds to a peptide-MHC complex for a period longer than a certain threshold time. It is supposed that a T-cell is activated if the sum of all the single triggering events per unit of time exceeds a certain threshold. Van den Berg and Rand (2003) argue that a T-cell can only discriminate between the noisy background of self-signals (caused by the stochastic distribution of dissociation times and the fact that the composition of self-peptide-MHC complexes varies between antigen presenting cells) and interaction with foreign peptides (presented by MHC molecules), if the number of TCRs that are triggered by the foreign peptide (presented by an MHC molecule) stands out against the self background. The authors show that this is the case if MHC diversity is low, so that a particular foreign peptide in complex with an MHC molecule can dominate the profile of peptide-MHC molecules of an antigen presenting cell.

However, the question remains whether the model captures the essential features of T-cell activation. In contrast to the model predictions as few as 1 to 10 agonist peptide-MHC complexes against a background of 10^5 to 10^6 irrelevant peptide-MHC complexes were recognized by mature T-cells (reviewed in George et al. 2005, Krogsgaard and Davis 2005).

Selection outcome in the thymus was shown to depend entirely on the affinity between a TCR and peptide-MHC complex (Daniels et al. 2006). Increasing the concentration of a positively selecting ligand could not convert it into a negatively selecting ligand (Daniels et al. 2006).

The fact that MHC expression is upregulated upon infection (e.g. Matsuyama et al. 2007) suggests that increasing the number of MHC molecules (without increasing their diversity) can enhance the effect of an immune response. This could be due to the fact that a foreign peptide that can bind to a certain MHC molecule type of the host still can be outcompeted by self-peptides when the MHC molecule is loaded. A sufficient number of MHC molecules may be needed to present the whole range of peptides capable of binding to the MHC molecule type just after they have been produced. Moreover it can be assumed that – although a single agonist peptide-MHC may be able to activate a T-cell – the probability that a T-cell encounters and reacts to a given peptide-MHC complex type is increased if the complex is present in higher numbers.

These ideas can be incorporated into our model by replacing the probability q that a given peptide is presented by a given MHC molecule type with the probability q'(M) that a given peptide is "sufficiently" presented to be able to activate T-cells effectively. Let q_0 denote the probability that a given peptide can bind to a certain MHC molecule type and T denote the minimal number of peptide-MHC complexes required for sufficient presentation. N signifies the total number of MHC molecules in the contact area between the antigen presenting cell and the T-cell and s is the probability that a representative of a given peptide type that can be presented by an MHC molecule is loaded onto it in a given loading event.

$$q'(M) = q_0 \cdot (1 - \sum_{i=0}^{T} \binom{N/M}{i} \cdot s^i (1-s)^{\frac{N}{M}-i}) \qquad \text{[eq4]}$$

This approach was not further pursued, because experimental data is not sufficient to obtain accurate parameter estimates for s and T. However, the qualitative effect is clear: If “sufficient presentation” is required, the optimal MHC diversity will decrease.

7. Outlook: Experimental test of the model predictions

7.1 Testable hypotheses and outline of experimental approach

The model predicts that (1) an increase in MHC diversity causes a rapid depletion of the mature TCR repertoire because of extensive negative selection and that (2) this depletion of the mature TCR repertoire reduces immunocompetence markedly. The first prediction can be tested by comparing either the mature TCR repertoire diversity or the stringency of negative selection in individuals of different MHC diversity.

There is evidence that positive and negative selection are spatially separated processes in the mammalian thymus. After passing positive selection in the cortex, thymocytes migrate into the medulla, where the bulk of negative selection occurs (Janeway 2005). This anatomical separation of the two selection processes offers the opportunity to isolate cortical thymocytes which have not yet been subjected to negative selection from medullary thymocytes most of which have passed negative selection since many thymocytes that are negatively selected die relatively quickly by apoptosis at the cortico-medullary boundary.

A technical solution to separate cortex and medulla can be to cut out cortical and medullary regions in a series of paraffin sections by a combination of laser microbeam microdissection and laser pressure catapulting (Imamichi et al. 2001). DNA can be extracted from the obtained samples and the sequence diversity of the rearranged TCR ß genes can be assessed by e.g. pyrosequencing (Margulies et al. 2005) or AmpliCot (a technique based on DNA hybridization kinetics; Baum and McCune 2006). While TCR ß sequence complexity in the cortex is not expected to be dependent on the MHC diversity of the individual, TCR ß sequence diversity in the medulla is expected to decrease with increasing MHC diversity.

The strengths of this approach are that it assesses the effect of negative selection by directly measuring TCR diversity prior to and after negative selection. In contrast to studies measuring cell numbers, results can thus not be confounded by proliferation. While earlier studies that estimated negative selection have involved extensive manipulation of the animals' natural physiology (e.g. bone marrow chimeras, transgenics expressing no MHC on cells of hematopoietic lineages; see 4.2), our approach is based on analyzing the thymus of an animal, whose physiology has not been manipulated at all. A possible weakness is that the separation of the cortex and medulla may not allow perfect discrimination between thymocytes that have not yet undergone negative selection and those that have passed it.

As the approach does not depend on cell markers, the production of transgenic animals or manipulation of the thymus, it is also applicable to non-model species. Sticklebacks (fig. 4) are particularly suited for the experiment, because stickleback MHC genotypes can differ extensively in the number of MHC class I and II ß alleles (Reusch et al. 2001). Moreover the existence of an immunogenetic optimum has been most clearly demonstrated in this species and the stickleback MHC has been exceptionally well studied (see chapter 3.2). Other aspects of the stickleback adaptive immune system, however, are still largely unexplored.

Fig. 4: Male stickleback (*Gasterosteus aculeatus*, picture from www.aaltojenalla.fi)

Clear identification and characterization of the stickleback thymus, which so far has only been mentioned in a brief communication on the lymphoid organs of sticklebacks (Bigaj et al. 1987), as well as the identification of genes of the adaptive immune system (especially the TCR ß gene

segments) are preconditions for studying the impact of MHC diversity on negative selection and TCR diversity in sticklebacks. The experimental part of my thesis focuses on these challenges.

7.2 Methods

7.2.1 Sectioning

Four-month-old labbred sticklebacks that showed no obvious signs of infection were killed with an overdose of MS 222 (Sigma; 1.5 mg/ml). Bodies were cut anterior and posterior to the gill chamber and the excised region around the gill chamber was fixed in 4% paraformaldehyde in 0.01M phosphate buffered saline (PBS) overnight. For decalcification samples were immersed in 10% acetic acid overnight (samples to be embedded in resin) or 20% Na-EDTA (pH = 7.4) for two days (samples to be embedded in paraffin). Samples were washed and subsequently dehydrated through a series of incubations in ethanol (30%, 50%, 70%, 90%, 96%; two hours each). Afterwards they were embedded in cold polymerizing resin (Technovit 7100, Kulzer) or paraffin (after preinfiltration steps with Roticlear-ethanol and Roticlear-paraffin mixtures, Roth) following the manufacturers' protocols. Serial cross sections of 3µm were obtained using a Leica RM2165 microtome. While paraffin sections were mounted on slides coated with protein glycerol (Roth), resin sections could be mounted on slides without coating.

7.2.2 Staining

Hematoxylin-eosin (HE) staining of all sections embedded in Technovit 7100 and a subsample of the paraffin sections was carried out following standard protocols.

For analyzing MHC class I and II expression in the stickleback thymus, antisera, which were raised against recombinant fusion proteins of the stickleback MHCI α and the MHCII α and ß chains, were used. These antisera are characterized in Scharsack et al. (in press). The only change to the staining protocols given in this article was that slides were incubated in 0.1M citrate buffer pH 6 for 20 minutes at 80°C instead of 0.01M citrate buffer pH6 for 30 minutes at 120°.

The lectin Concanavalin A stains mature T-cells more strongly than immature thymocytes in the zebrafish thymus (Schorpp et al. 2002). For lectin staining paraffin sections were dewaxed with Roticlear (Roth) and subsequently rehydrated through a series of ethanol solutions (99%, 90%, 70%, 50%, 30%; 1 min. each) and incubated in 0.01M PBS. Afterwards sections to be stained using a conjugate of Concanavalin A with peroxidase were treated with 1% H_2O_2 for 15 minutes. After washing with PBS these sections were incubated with citrate buffer (10mM citric acid, adjusted to pH 6 with NaOH) at 80°C for 20 minutes. All slides were mounted in shandon coverplates (Thermo Electron) and incubated with the corresponding Concanavalin A conjugate, that is either with Concanavalin-FITC (Sigma, diluted to concentrations in the range of 0.1mg/ml to 5µg/ml in PBS) or with Concanavalin-Peroxidase (Sigma, diluted to concentrations in the range of 0.1mg/ml to 0.01mg/ml in PBS) for 1 hour. Afterwards all slides were washed with PBS. Sections treated with Concanavalin-FITC were fixed and counterstained with propidium iodide (1 mg/l) in anti-fading mounting medium (Fluka) for fluorescence microscopic analysis. Sections that had been incubated with the peroxidase conjugate were incubated with DAB solution (0.05% 3,3'-diaminobenzidine tetrahydrochloride, 0.015% H_2O_2 in 0.01M PBS). After washing with H_2O, hematoxylin was used for nuclear counterstain and sections were mounted in glycerinegelatine.

As staining with both conjugates did result in a high background (for higher conjugate concentrations) or very low signals (for lower conjugate concentrations), biotinylated Concanavalin A (Sigma) was used in combination with Streptavidin-HRP (Sigma). It was hoped that the additional amplification step would lead to more specific staining. The protocol given in Schorpp et al. (2002) was followed.

7.2.3 Identification of genes of the adaptive immune system in *Gasterosteus aculeatus*

The stickleback sequences of RAG-1, CD8 α as well as the genomic region containing TCR ß V, D, J, and C-segments were obtained by blasting nucleotide sequences (and their amino acid translations) of homologous genes of other teleostei against the stickleback genome (www.ensembl.org). Hits were checked by blasting sequences of stickleback exons (obtained from ESTs aligned to the genome in ensembl or obtained from gene predictions performed by ensembl) against the nr database of GenBank (www.ncbi.nlm.nih.gov). Putative TCR ß gene segments were extracted by screening the TCR ß genomic region for recombination signal sequences and conserved motifs in the amino acid translation.

7.3 Results and discussion

7.3.1 Histological identification and characterization of the thymus

Apart from the foremost parts of the head kidney, regions dorsolateral of the gill chamber take an intense basophilic stain in HE cross sections (fig. 5). Further analysis (see below) identifies these darkly stained regions as the thymus, which is a paired organ in teleosts (Press and Evensen 1999). The series of sections proves that the thymus is clearly seperated from the head kidney (fig. 5).

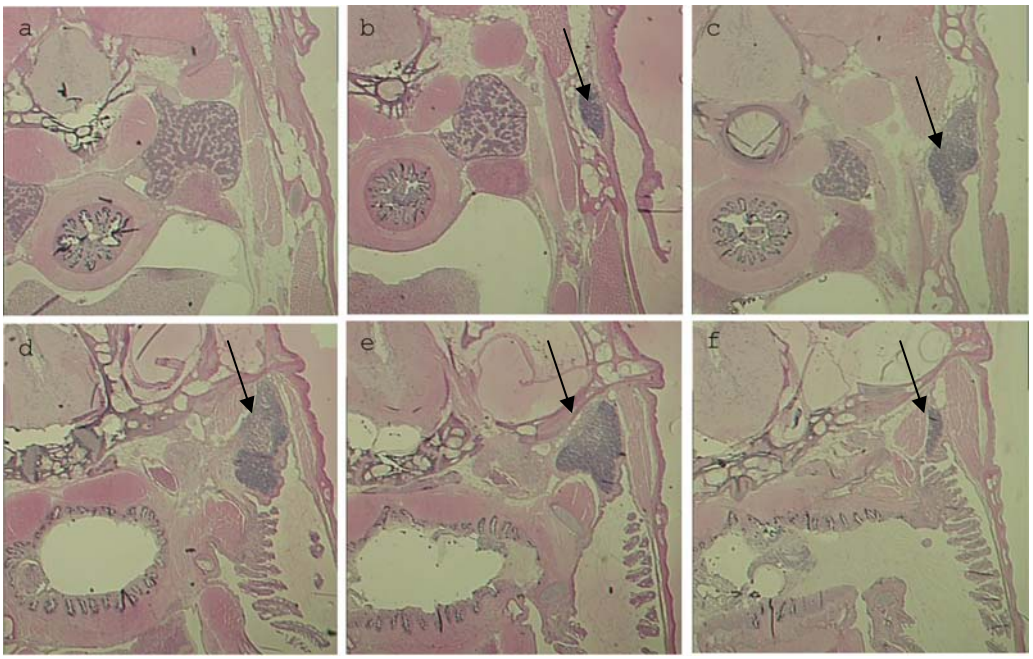


Fig. 5: Series of sections (distance between sections ≈ 100µm) showing the thymus (arrow) as a darkly stained region dorsolateral of the gill chamber (15x)

It is delineated by a connective tissue capsule that does not completely surround the organ, so that part of the thymic parenchyma is in direct contact with the branchial epithelium (fig. 6). This feature has been described in histological studies of the thymus in other teleost species (e.g. zebrafish *Danio rerio*, sharpsnout seabream *Diplodus puntazzo*; Willett et al. 1997, Romano et al. 1999) and has been mentioned in a brief communication on the lymphoid organs of sticklebacks previously (Bigaj et al. 1987). The direct contact with the epithelium of the gill cavity reflects the ontogeny of the thymus, which is formed by epithelial cells that create a threedimensional network in which thymocytes and macrophages are dispersed. Because of this organisation the thymus is referred to as an intraepithelial organ (Press and Evensen 1999).

Stromal cells in the thymus are expected to show high levels of MHC class I as well as MHC class II expression, because these molecules mediate positive and negative selection. Staining with antisera against MHC I α and MHC II α and ß chains resulted in intense labelling of stromal cells in the thymus (fig. 7; compare the strong signal from the

thymus with the staining pattern in the oesophagus where single cells – putative residual macrophages – are stained and the low green fluorescence of the hindbrain indicating low background staining).

Typical small (diameter around 5µm), round, dark-staining thymocytes with a large nucleus are scattered all over the organ (fig. 6). Based on the different densities of thymocytes an inner zone (region farthest from the gill chamber in which lymphoid elements are less numerous and stromal cells more prominent) can be distinguished from an outer zone (fig. 6). Such a differentiation into two zones was observed in many teleost species including *Zoarces viviparus, Ictalurus punctatus, Chelon labrosus* (reviewed in Chilmonczyk

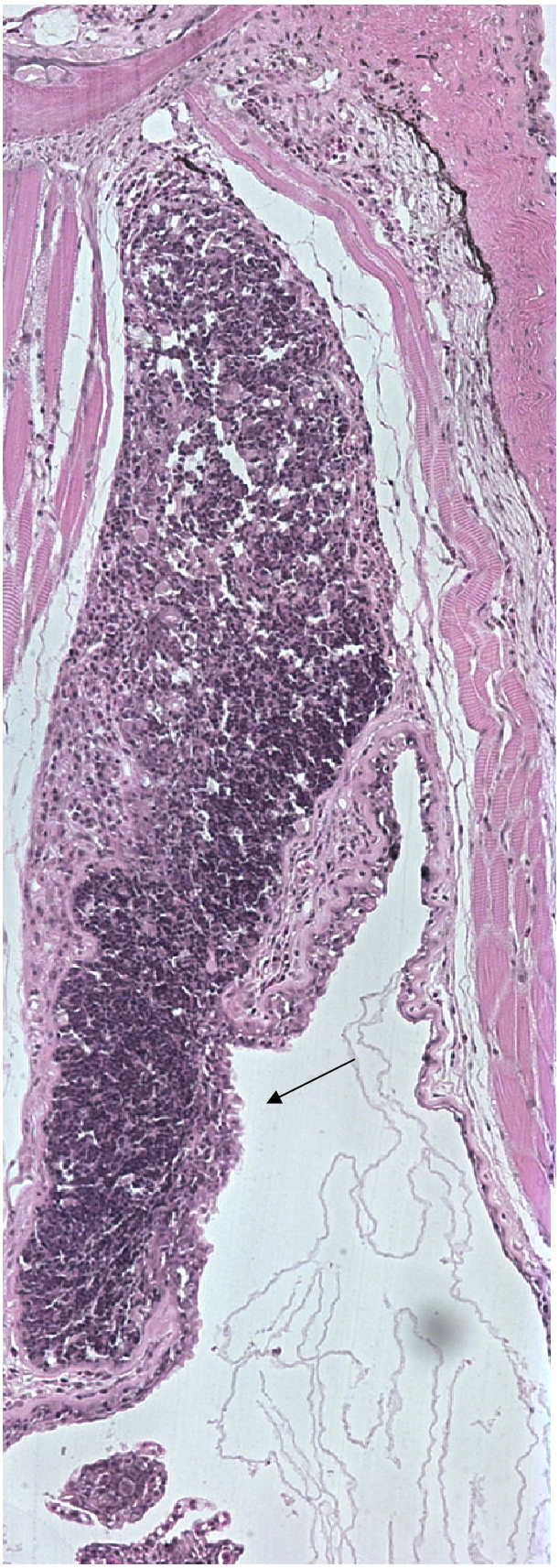


Fig. 6: Stickleback thymus, arrow marks direct contact between the branchial epithelium and the thymic parenchyma (100x)

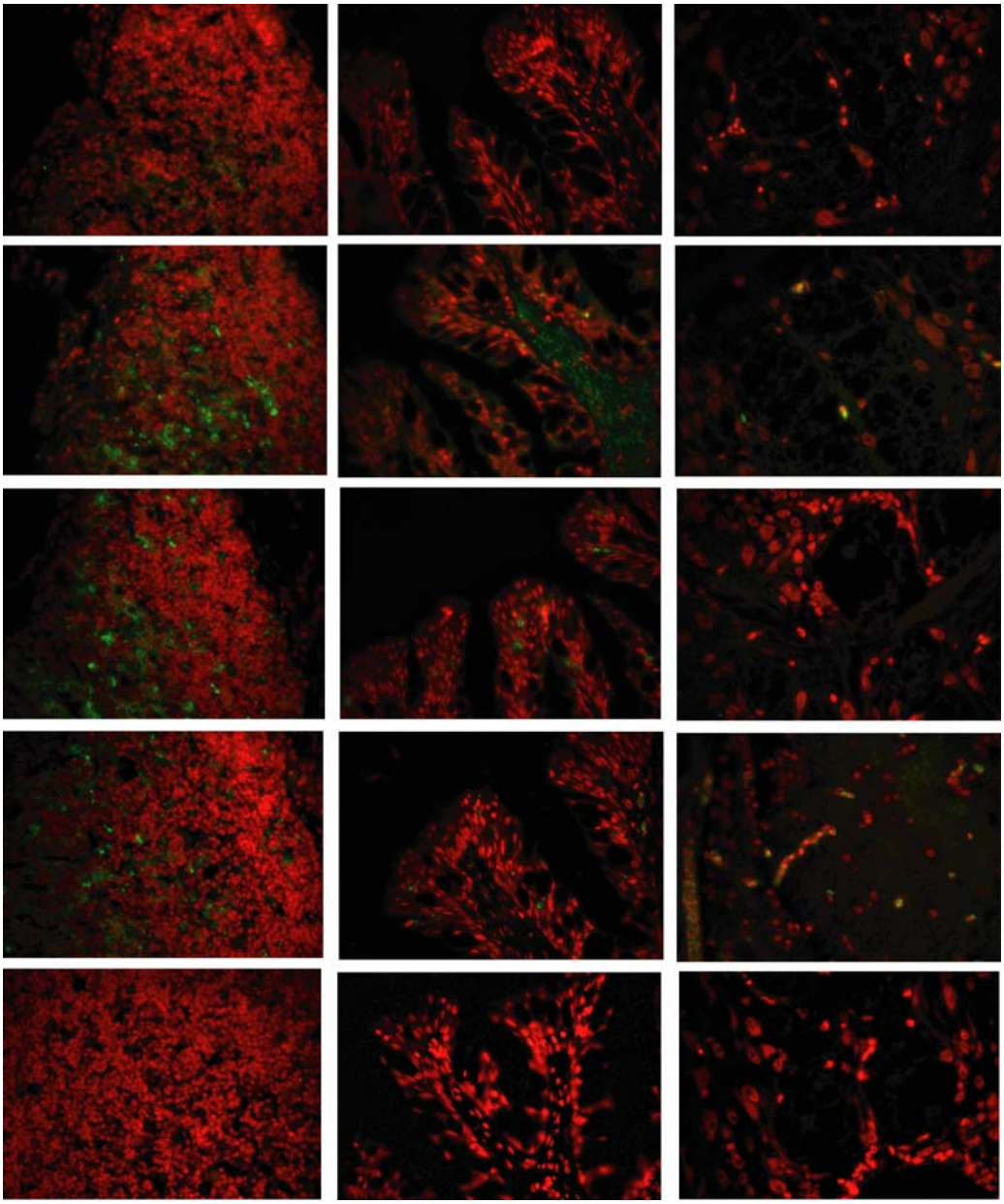


Fig. 7: Staining of thymus (1st column), oesophagus (2nd column) and hindbrain (3rd column) with MHCI (1st row), MHCII α-chain (2nd row) and MHCII ß-chain (3rd and 4th row) antisera. 5th row is control. (130x)

1992), *Diplodus puntazzo* (Romano et al. 1999) and *Danio rerio* (Schorpp et al. 2002). Therefore an organisation into an outer zone and an inner zone seems to be typical of a teleost thymus, although some authors identified more zones (e.g. in *Oncorhynchus mykiss*; reviewed in Chilmonczyk 1992). The situation is reminiscent of the mammalian thymus which also contains two discrete regions termed cortex and medulla. Like the outer zone of the teleost thymus the mammalian cortex

stains darkly in HE and contains numerous lymphoid elements, whereas the mammalian medulla is similar to the inner zone of a teleost thymus in that it contains fewer thymocytes and has more prominent stromal cells.

Further support for the hypothesis that the outer and inner zone in teleost thymuses correspond to the mammalian cortex and medulla in terms of anatomy and function comes from a study on zebrafish. Schorpp et al. (2002) showed that the lectin Concanavalin A which stains mature T-cells more strongly than immature thymocytes predominantly stains T-cells in the inner zone of the zebrafish thymus. Moreover Recombinase Activating Gene - 1 (RAG-1) expression was found to be restricted to the outer zone. The staining pattern of Concanavalin A as well as that of the RAG-1 in situ hybridization are analogous to staining patterns in mammals, where preselection thymocytes in the cortex express RAG-1 (which is required for somatic rearrangement) and mature T-cells that undergo or have passed negative selection are found in the medulla. Staining of the stickleback thymus with three different Concanavalin A – conjugates resulted in rather unspecific staining (but observe that the difference in the staining of thymocytes in the inner and outer zone is also not very marked in zebrafish, Schorpp et al. 2003). The Concanavalin A – FITC conjugate staining is shown as an example (fig. 8). Although Concanavalin A – FITC seems to stain predominantly the inner zone, it is unclear if this is specific staining or just high background staining. Based on these results it is not possible to conclude with certainty that mature thymocytes are localized in the inner zone whereas immature thymocytes are present in the outer zone. To achieve clearer

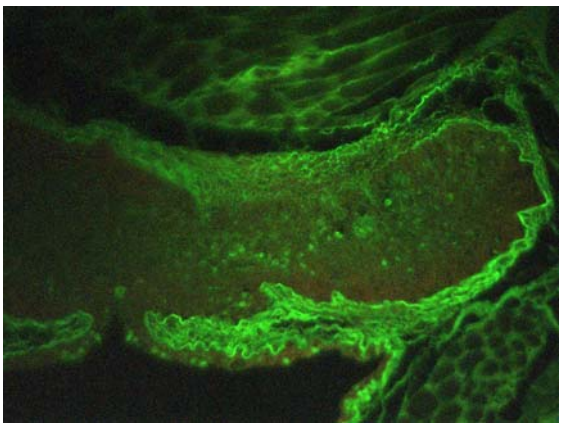


Fig. 8: Thymus stained with Concanavalin – FITC (80x)

staining results it seems to be most promising to focus on staining with biotinylated Concanavalin A (because of the addititonal amplification step in the staining process) and perform experiments with slight variations of the staining protocol (longer incubation time in citric acid, lower concentration of the biotinylated Concanavalin and higher concentrations of Streptavidin-HRP). Alternatively a RAG-1 in situ hybridization, which gave clear results in Schorpp et al. (2003), can be envisioned.

7.3.2 Identification of genes of the adaptive immune system in Gasterosteus aculeatus

The stickleback sequences of RAG-1 (genomic location: contig 1625, 14.493.756 – 14.197.787) and CD8α (genomic location: groupXVII, 8.570.785 - 8.572.274) were obtained. When the genes were blasted against the nr database of GenBank, the most similar genes found were always homologous genes in other vertebrates. V-, D-, J- and C-TCRß gene segments were found in two regions (contig 3794 and contig 4121) of the stickleback genome. Both regions lie on linkage group XV and are separated by a sequence of 10Mbp length. Most probably the separation is due to a mistake in the genome assembly, because such a separation is unknown from TCRß-regions of other species. An alignment of the amino acid sequences of a putative stickleback TCRß (accession code gi|62068636) with a cod TCRß (accession code AJ133848) shows that the stickleback sequence shares conserved amino acid motifs, which are characteristic of TCRß - sequences, with the cod sequence (fig. 9). The stickleback Vß-region contains the conserved residues Glu6, Cys23, Trp34 and Cys92 (marked in red in fig. 9) which are supposed to be important for the convolution of the protein (Wermenstam and Pilström 2001). The J-segment contains the motif Phe-Gly-X-Gly (marked in red in fig. 9), which is highly conserved among TCRß – sequences of all vertebrates. The conserved antigen receptor transmembrane motif

(CART), which is involved in the assembly of the TCR-CD3 complex (Wermenstam and Pilström 2001) is present in the stickleback transmembrane region (2 tyrosin residues, 1 valin residue; marked in red, fig. 9).

```
            L                    ><V
Gas. ac.    ---------- --MLWTAGVS QSVLITQWPR YISSPPNGSA EMTCYQNDTN YDYLYWYQQL
Gad. mor.   MCQFICSLGI IFVCFTHEV- QSVIIKQSSA KIVRKGAKGI QIDCSHDDSS YPLMYWYQRK

                                                                        ><
Gas. ac.    S-GRSFQLI- VSTVGGFSSY EVGFKSGFQS VKFSEKQWSL TVTGLQGNHA AVYLCAASQG
Gad. mor.   DESPSLTLIG FGYESSTQNY EDRFEERLNI TRESVLQGTL VLTEAAESDS AVYFCAASMG

            D ><J            ><C
Gas. ac.    ---ASEAYFG KGTKLTVLEP GLNVTKPKVK VLRPSARECG NPKDKARKTI VCVASGFYPD
Gad. mor.   EGGSEPAFFG KGTKLTVLEP GCIVSPPTVV VLPPSEKECR DRKEQLKKTL VCVASGFYPD

Gas. ac.    HVGILWNIDG QNVIDGQNVT DRRVATDNAA LRVDQYYRIT SRLSVLAEEW FTEGKNFTCI
Gad. mor.   HVGVSWTVNG QSVIKG---- ---VASDHPA LRVDDKYQIT SRLRVEARKW YTGGNIFTCN

                                  ><CPS                  ><TM
Gas. ac.    VSFFNGKNTE SYQSSTIGEK DKRSSX---- --DREKYLMV TQTARFSYII FIFQSSIYGA
Gad. mor.   VSYFNGNDTI YTSAEVYGGG DVRWIKTEPD GETREEFVKV TQTAKLSYIV MIVKNIVYGV

                     ><CYT      >
Gas. ac.    FVXFLVWRL- -QRSAGKQND
Gad. mor.   FVTILAWKLG LGRSHATAKK
```

Fig. 9: Alignment of a stickleback and a cod TCRß sequence. Leader- (L), variable (V), diversity (D), joining (J) and constant (C) region are marked. CPS = connecting peptide sequence, TM = transmembrane region, CYT = cytoplasmic region. Conserved amino acid residues are coloured in red.

8. Conclusion

The predictions of the model outlined in chapter 5.1 clearly show that the depletion hypothesis can potentially explain why jawed vertebrates have intermediate numbers of MHC molecules. The underlying intuition is that there is a trade-off between ensuring that peptides of a pathogen are presented on MHC molecules with a high probability (which requires high MHC diversities) and having a diverse T-cell receptor repertoire (which requires low MHC diversities to reduce the impact of negative selection). This is an interesting result because Borghans et al. (2003) dismissed the hypothesis that extensive negative selection at high MHC diversities can result in intermediate MHC diversities being optimal. The model by Borghans et al. (2003) and the model presented in this thesis reach different conclusions mainly because positive selection efficiency is assumed to be independent of MHC diversity in the latter model, whereas Borghans et al. (2003) assume a strong increase of positive selection efficiency with MHC diversity. Experimental data indicates that – although neither view is completely correct – the new model probably captures the situation found in nature more adequately (chapter 4.1). While it is interesting to see that existing experimental data is not at variance with the depletion hypothesis, this is certainly not a proof that it holds. The outlined test for the model prediction that markedly fewer T-cells mature in individuals of high MHC diversity will either reject or give strong support for the depletion hypothesis.

Apart from the extent of negative selection and the probability that an immunogenic peptide can be presented, other processes are also influenced by MHC diversity. The risk of autoimmune diseases is generally thought to increase steadily with MHC diversity. My results show that this is only the case if protecting alleles are scarce. This effect is caused by the fact that depletion of the mature TCR repertoire by extensive negative selection decreases the risk of autoimmune diseases.

In individuals of high MHC diversity many autoreactive T-cells are deleted or develop into regulatory T-cells. Thus it cannot be taken for granted that autoimmune diseases constitute a selection pressure that sets an upper limit to intraindividual MHC diversity.

Note that almost all remaining uncertainties are due to the fact that we are only beginning to understand how binding of TCRs to their ligands mediates T-cell activation. In order to understand how positive and negative selection efficiencies are affected by MHC diversity, we need to know how important conserved interactions are for TCR binding and which role interactions with MHC- and peptide-specific residues play. Further insights into how triggering of TCRs leads to the activation of mature T-cells will clarify the influence of antigen presentation levels on T-cell activation and help us to understand why some T-cells that have evaded tolerance induction can become activated by self-peptide – MHC complexes in the periphery. Future research will certainly focus on how slight differences in the interaction between TCRs and their ligands can cause such a variety of cellular responses including death by neglect, lineage deviation, deletion, anergy and unleashing of effector functions.

9. Literature

Aeschlimann, P.B., Haberli, M.A., Reusch, T.B.H., Boehm, T., and Milinski, M. (2003). Female sticklebacks Gasterosteus aculeatus use self-reference to optimize MHC allele number during mate selection. Behavioral Ecology and Sociobiology 54, 119-126.

Baum, P.D., and McCune, J.M. (2006). Direct measurement of T-cell receptor repertoire diversity with AmpliCot. Nature Methods 3, 895-901.

Bernatchez, L., and Landry, C. (2003). MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? Journal of Evolutionary Biology 16, 363-377.

Bigaj, J., Dulak, J., and Plytycz, B. (1987). Lymphoid Organs of Gasterosteus-Aculeatus. Journal of Fish Biology 31, 233-234.

Boehm, T. (2006). Quality control in self/nonself discrimination. Cell 125, 845-858.

Bonneaud, C., Mazuc, J., Chastel, O., Westerdahl, H., and Sorci, G. (2004). Terminal investment induced by immune challenge and fitness traits associated with major histocompatibility complex in the house sparrow. Evolution 58, 2823-2830.

Bonneaud, C., Richard, M., Faivre, B., Westerdahl, H., and Sorci, G. (2005). An Mhc class I allele associated to the expression of T-dependent immune response in the house sparrow. Immunogenetics 57, 782-789.

Borghans, J.A.M., Noest, A.J., and De Boer, R.J. (2003). Thymic selection does not limit the individual MHC diversity. European Journal of Immunology 33, 3353-3358.

Carrington, M., Nelson, G.W., Martin, M.P., Kissner, T., Vlahov, D., Goedert, J.J., Kaslow, R., Buchbinder, S., Hoots, K., and O'Brien, S.J. (1999). HLA and HIV-1: Heterozygote advantage and B*35-Cw*04 disadvantage. Science 283, 1748-1752.

Casrouge, A., Beaudoing, E., Dalle, S., Pannetier, C., Kanellopoulos, J., and Kourilsky, P. (2000). Size estimate of the alpha beta TCR repertoire of naive mouse splenocytes. Journal of Immunology 164, 5782-5787.

Chilmonczyk, S. (1992). The thymus in fish: development and possible function in the immune response. Annual Review of Fish Disease 2, 181-200.

Daniels, M.A., Teixeiro, E., Gill, J., Hausmann, B., Roubaty, D., Holmberg, K., Werlen, G., Hollander, G.A., Gascoigne, N.R.J., and Palmer, E. (2006). Thymic selection threshold defined by compartmentalization of Ras/MAPK signalling. Nature 444, 724-729.

Davey, G.M., Schober, S.L., Endrizzi, B.T., Dutcher, A.K., Jameson, S.C., and Hogquist, K.A. (1998). Preselection thymocytes are more sensitive to T cell receptor stimulation than mature T cells. Journal of Experimental Medicine 188, 1867-1874.

Deboer, R.J., and Perelson, A.S. (1993). How Diverse Should the Immune-System Be. Proceedings of the Royal Society of London Series B-Biological Sciences 252, 171-175.

Detours, V., Mehr, R., and Perelson, A.S. (2000). Deriving quantitative constraints on T cell selection from data on the mature T cell repertoire. Journal of Immunology 164, 121-128.

Egerton, M., Scollay, R., and Shortman, K. (1990). Kinetics of Mature T-Cell Development in the Thymus. Proceedings of the National Academy of Sciences of the United States of America 87, 2579-2582.

Feng, C.G., Woodside, K.J., Vance, B.A., El-Khoury, D., Canelles, M., Lee, J., Gress, R., Fowlkes, B.J., Shores, E.W., and Love, P.E. (2002). A potential role for CD69 in thymocyte emigration. International Immunology 14, 535-544.

Flajnik, M.F. (1996). The immune system of ectothermic vertebrates. Veterinary Immunology and Immunopathology 54, 145-150.

Fontenot, J.D., Rasmussen, J.P., Williams, L.M., Dooley, J.L., Farr, A.G., and Rudensky, A.Y. (2005). Regulatory T cell lineage specification by the forkhead transcription factor FoxP3. Immunity 22, 329-341.

Forrest, S., and Beauchemin, C. (2007). Computer immunology. Immunological Reviews 216, 176-197.

FungLeung, W.P., Surh, C.D., Liljedahl, M., Pang, J., Leturcq, D., Peterson, P.A., Webb, S.R., and Karlsson, L. (1996). Antigen presentation and T cell development in H2-M-deficient mice. Science 271, 1278-1281.

George, A.J.T., Stark, J., and Chan, C. (2005). Understanding specificity and sensitivity of T-cell recognition. Trends in Immunology 26, 653-659.

Hare, K.J., Jenkinson, E.J., and Anderson, G. (1999). CD69 expression discriminates MHC-dependent and -independent stages of thymocyte positive selection. Journal of Immunology 162, 3978-3983.

Harf, R., and Sommer, S. (2005). Association between major histocompatibility complex class II DRB alleles and parasite load in the hairy-footed gerbil, Gerbillurus paeba, in the southern Kalahari. Molecular Ecology 14, 85-91.

Hill, A.V.S., Allsopp, C.E.M., Kwiatkowski, D., Anstey, N.M., Twumasi, P., Rowe, P.A., Bennett, S., Brewster, D., McMichael, A.J., and Greenwood, B.M. (1991). Common West African Hla Antigens Are Associated with Protection from Severe Malaria. Nature 352, 595-600.

Hughes, A.L., and Nei, M. (1988). Pattern of Nucleotide Substitution at Major Histocompatibility Complex Class-I Loci Reveals Overdominant Selection. Nature 335, 167-170.

Hughes, A.L., and Nei, M. (1989). Nucleotide Substitution at Major Histocompatibility Complex Class-Ii Loci - Evidence for

Overdominant Selection. Proceedings of the National Academy of Sciences of the United States of America 86, 958-962.

Hunt, D.F., Michel, H., Dickinson, T.A., Shabanowitz, J., Cox, A.L., Sakaguchi, K., Appella, E., Grey, H.M., and Sette, A. (1992). Peptides Presented to the Immune-System by the Murine Class-Ii Major Histocompatibility Complex Molecule-I-a(D). Science 256, 1817-1820.

Huseby, E.S., Crawford, F., White, J., Kappler, J., and Marrack, P. (2003). Negative selection imparts peptide specificity to the mature T cell repertoire. Proceedings of the National Academy of Sciences of the United States of America 100, 11565-11570.

Huseby, E.S., White, J., Crawford, F., Vass, T., Becker, D., Pinilla, C., Marrack, P., and Kappler, J.W. (2005). How the T cell repertoire becomes peptide and MHC specific. Cell 122, 247-260.

Ignatowicz, L., Kappler, J., and Marrack, P. (1996). The repertoire of T cells shaped by a single MHC/peptide ligand. Cell 84, 521-529.

Imamichi, Y., Lahr, G., and Wedlich, D. (2001). Laser-mediated microdissection of paraffin sections from Xenopus embryos allows detection of tissue-specific expressed mRNAs. Development Genes and Evolution 211, 361-366.

Janeway, C.A., Travers, P., Walport, M., Shlomchik, M.J. (2005). Immunobiology: the immune system in health and disease. 6th edition.

Jerne, N.K. (2004). The somatic generation of immune recognition (Reprinted from Eur. J. Immunol. vol 1, pg 1-9, 1971). European Journal of Immunology 34, 1234-1242.

Jones, E.Y., Fugger, L., Strominger, J.L., and Siebold, C. (2006). MHC class II proteins and disease: a structural perspective. Nature Reviews Immunology 6, 271-282.

Kast, W.M., Brandt, R.M.P., Sidney, J., Drijfhout, J.W., Kubo, R.T., Grey, H.M., Melief, C.J.M., and Sette, A. (1994). Role of Hla-a Motifs in Identification of Potential Ctl Epitopes in Human Papillomavirus Type-16 E6 and E7 Proteins. Journal of Immunology 152, 3904-3912.

Klein, J., and Ohuigin, C. (1994). Mhc Polymorphism and Parasites. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences 346, 351-357.

Kobel, H.R., and Dupasquier, L. (1986). Genetics of Polyploid Xenopus. Trends in Genetics 2, 310-315.

Krogsgaard, M., and Davis, M.M. (2005). How T cells 'see' antigen. Nature Immunology 6, 239-245.

Kronenberg, M., and Rudensky, A. (2005). Regulation of immunity by self-reactive T cells. Nature 435, 598-604.

Langefors, A., Lohm, J., Grahn, M., Andersen, O., and von Schantz, T. (2001). Association between major histocompatibility complex class IIB alleles and resistance to Aeromonas salmonicida in

Atlantic salmon. Proceedings of the Royal Society of London Series B-Biological Sciences 268, 479-485.

Laufer, T.M., DeKoning, J., Markowitz, J.S., Lo, D., and Glimcher, L.H. (1996). Unopposed positive selection and autoreactivity in mice expressing class II MHC only on thymic cortex. Nature 383, 81-85.

Lawlor, D.A., Zemmour, J., Ennis, P.D., and Parham, P. (1990). Evolution of Class-I Mhc Genes and Proteins - from Natural-Selection to Thymic Selection. Annual Review of Immunology 8, 23-63.

Lucas, B., Stefanova, I., Yasutomo, K., Dautigny, N., and Germain, R.N. (1999). Divergent changes in the sensitivity of maturing T cells to structurally related ligands underlies formation of a useful T cell repertoire. Immunity 10, 367-376.

Madsen, T., and Ujvari, B. (2006). MHC class I variation associates with parasite resistance and longevity in tropical pythons. Journal of Evolutionary Biology 19, 1973-1978.

Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., Berka, J., Braverman, M.S., Chen, Y.J., Chen, Z.T., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. Nature 437, 376-380.

Martin, M.P., and Carrington, M. (2005). Immunogenetics of viral infections. Current Opinion in Immunology 17, 510-516.

Martin, W.D., Hicks, G.G., Mendiratta, S.K., Leva, H.I., Ruley, H.E., and VanKaer, L. (1996). H2-M mutant mice are defective in the peptide loading of class II molecules, antigen presentation, and T cell repertoire selection. Cell 84, 543-550.

Mason, D. (1998). A very high level of crossreactivity is an essential feature of the T-cell receptor. Immunology Today 19, 395-404.

Matsuyama, T., Fujiwara, A., Nakayasu, C., Kamaishi, T., Oseko, N., Tsutsumi, N., Hirono, I., and Aoki, T. (2007). Microarray analyses of gene expression in Japanese flounder Paralichthys olivaceus leucocytes during monogenean parasite Neoheterobothrium hirame infection. Diseases of Aquatic Organisms 75, 79-83.

Merkenschlager, M., Graf, D., Lovatt, M., Bommhardt, U., Zamoyska, R., and Fisher, A.G. (1997). How many thymocytes audition for selection? Journal of Experimental Medicine 186, 1149-1158.

Mineta, T., Matsunaga, M., Takashima, H., Uno, M., Miyazaki, J., Yamamura, K., and Kimoto, M. (1991). (Balb/Cxc57bl/6)F1 Mice Are Tolerant to Undetectable Mixed Haplotype a-Beta-Da-Alpha-B and Mixed Isotype a-Beta-De-Alpha-D Self Class-Ii Molecules. International Immunology 3, 453-460.

Monteiro, M.C., Couceiro, S., and Penha-Goncalves, C. (2005). The multigenic structure of the MHC locus contributes to positive selection efficiency: A role for MHC class II gene-specific restriction. European Journal of Immunology 35, 3622-3630.

Nowak, M.A., Tarczyhornoch, K., and Austyn, J.M. (1992). The Optimal Number of Major Histocompatibility Complex-Molecules in an Individual. Proceedings of the National Academy of Sciences of the United States of America 89, 10896-10899.

Oksenberg, J.R., Barcellos, L.F., Cree, B.A.C., Baranzini, S.E., Bugawan, T.L., Khan, O., Lincoln, R.R., Swerdlin, A., Mignot, E., Lin, L., et al. (2004). Mapping multiple sclerosis susceptibility to the HLA-DR locus in African Americans. American Journal of Human Genetics 74, 160-167.

Palmer, E. (2003). Negative selection - Clearing out the bad apples from the T-cell repertoire. Nature Reviews Immunology 3, 383-391.

Paterson, S., Wilson, K., and Pemberton, J.M. (1998). Major histocompatibility complex variation associated with juvenile survival and parasite resistance in a large unmanaged ungulate population (Ovis aries L.). Proceedings of the National Academy of Sciences of the United States of America 95, 3714-3719.

Penn, D.J., Damjanovich, K., and Potts, W.K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. Proceedings of the National Academy of Sciences of the United States of America 99, 11260-11264.

Piertney, S.B., and Oliver, M.K. (2006). The evolutionary ecology of the major histocompatibility complex. Heredity 96, 7-21.

Press, C.M., and Evensen, O. (1999). The morphology of the immune system in teleost fishes. Fish & Shellfish Immunology 9, 309-318.

Reusch, T.B.H., Haberli, M.A., Aeschlimann, P.B., and Milinski, M. (2001). Female sticklebacks count alleles in a strategy of sexual selection explaining MHC polymorphism. Nature 414, 300-302.

Romano, N., Fanelli, M., Del Papa, G.M., Scapigliati, G., and Mastrolia, L. (1999). Histological and cytological studies on the developing thymus of sharpsnout seabream, Diplodus puntazzo. Journal of Anatomy 194, 39-50.

Sakaguchi, S., Sakaguchi, N., Asano, M., Itoh, M., and Toda, M. (1995). Immunological Self-Tolerance Maintained by Activated T-Cells Expressing Il-2 Receptor Alpha-Chains (Cd25) - Breakdown of a Single Mechanism of Self-Tolerance Causes Various Autoimmune-Diseases. Journal of Immunology 155, 1151-1164.

Santori, F.R., Kieper, W.C., Brown, S.M., Lu, Y., Neubert, T.A., Johnson, K.L., Naylor, S., Vukmanovic, S., Hogquist, K.A., and Jameson, S.C. (2002). Rare, structurally homologous self-peptides promote thymocyte positive selection. Immunity 17, 131-142.

Schad, J., Ganzhorn, J.U., and Sommer, S. (2005). Parasite burden and constitution of major histocompatibility complex in the malagasy mouse lemur, Microcebus murinus. Evolution 59, 439-450.

Scharsack, J.P., Kalbe, M., Schachl, H. (in press). Characterization of antisera raised against stickleback (Gasterosteus aculeatus) MHC class I and class II molecules.

Schorpp, M., Leicht, M., Nold, E., Hammerschmidt, M., Haas-Assenbaum, A., Wiest, W., and Boehm, T. (2002). A zebrafish orthologue (whnb) of the mouse nude gene is expressed in the epithelial compartment of the embryonic thymic rudiment. Mechanisms of Development 118, 179-185.

Scollay, R.G., Butcher, E.C., and Weissman, I.L. (1980). Thymus-Cell Migration Quantitative Aspects of Cellular Traffic from the Thymus to the Periphery in Mice. European Journal of Immunology 10, 210-218.

Siebold, C., Hansen, B.E., Wyer, J.R., Harlos, K., Esnouf, R.E., Svejgaard, A., Bell, J.I., Strominger, J.L., Jones, E.Y., and Fugger, L. (2004). Crystal structure of HLA-DQ0602 that protects against type 1 diabetes and confers strong susceptibility to narcolepsy. Proceedings of the National Academy of Sciences of the United States of America 101, 1999-2004.

Starr, T.K., Jameson, S.C., and Hogquist, K.A. (2003). Positive and negative selection of T cells. Annual Review of Immunology 21, 139-176.

Steinman, R.M., and Nussenzweig, M.C. (2002). Avoiding horror autotoxicus: The importance of dendritic cells in peripheral T cell tolerance. Proceedings of the National Academy of Sciences of the United States of America 99, 351-358.

Surh, C.D., Lee, D.S., FungLeung, W.P., Karlsson, L., and Sprent, J. (1997). Thymic selection by a single MHC/peptide ligand produces a semidiverse repertoire of CD4(+) T cells. Immunity 7, 209-219.

Thursz, M.R., Thomas, H.C., Greenwood, B.M., and Hill, A.V.S. (1997). Heterozygote advantage for HLA class-II type in hepatitis B virus infection. Nature Genetics 17, 11-12.

Todd, J.A., Bell, J.I., and McDevitt, H.O. (1987). Hla-Dq-Beta Gene Contributes to Susceptibility and Resistance to Insulin-Dependent Diabetes-Mellitus. Nature 329, 599-604.

Tourne, S., Miyazaki, T., Oxenius, A., Klein, L., Fehr, T., Kyewski, B., Benoist, C., and Mathis, D. (1997). Selection of a broad repertoire of CB4(+) T cells in H-2Ma(0/0) mice. Immunity 7, 187-195.

van den Berg, H.A., and Rand, D.A. (2003). Antigen presentation on MHC molecules as a diversity filter that enhances immune efficacy. Journal of Theoretical Biology 224, 249-267.

vanMeerwijk, J.P.M., Marguerat, S., Lees, R.K., Germain, R.N., Fowlkes, B.J., and MacDonald, H.R. (1997). Quantitative impact of thymic clonal deletion on the T cell repertoire. Journal of Experimental Medicine 185, 377-383.

Vidovic, D., and Matzinger, P. (1988). Unresponsiveness to a Foreign Antigen Can Be Caused by Self-Tolerance. Nature 336, 222-225.

Wegner, K.M., Kalbe, M., Kurtz, J., Reusch, T.B.H., and Milinski, M. (2003b). Parasite selection for immunogenetic optimality. Science 301, 1343-1343.

Wegner, K.M., Reusch, T.B.H., and Kalbe, M. (2003a). Multiple parasites are driving major histocompatibility complex polymorphism in the wild. Journal of Evolutionary Biology 16, 224-232.

Wermenstam, N.E., and Pilstrom, L. (2001). T-cell antigen receptors in Atlantic cod (Gadus morhua L.): structure, organisation and expression of TCR alpha and beta genes. Developmental and Comparative Immunology 25, 117-135.

Willett, C.E., Zapata, A.G., Hopkins, N., and Steiner, L.A. (1997). Expression of zebrafish rag genes during early development identifies the thymus. Developmental Biology 182, 331-341.

Wordsworth, B.P., Lanchbury, J.S.S., Sakkas, L.I., Welsh, K.I., Panayi, G.S., and Bell, J.I. (1989). Hla-Dr4 Subtype Frequencies in Rheumatoid-Arthritis Indicate That Drb1 Is the Major Susceptibility Locus within the Hla Class-Ii Region. Proceedings of the National Academy of Sciences of the United States of America 86, 10049-10053.

Zerrahn, J., Held, W., and Raulet, D.H. (1997). The MHC reactivity of the T cell repertoire prior to positive and negative selection. Cell 88, 627-636.

Zinkernagel, R.M. (1996). Immunology taught by viruses. Science 271, 173-178.

10 Danksagung

In dieser Arbeit steckt die Unterstützung durch viele Menschen, ohne die ich an vielen Punkten sicher nicht weiter gekommen wäre.

Manfred Milinski gilt mein Dank für die Idee an dem Thema zu arbeiten, seine stete Unterstützung und seinen Optimismus während Durststrecken. Ganz besonders danke ich ihm auch dafür, dass er mir die Gelegenheit gab Thomas Boehm und Martin Nowak zu besuchen, um von ihnen zu lernen und um mit ihnen Probleme im Zusammenhang mit dieser Arbeit zu diskutieren.

Viele der Schlüsselideen in dieser Arbeit habe ich der Intuition von Thomas Boehm zu verdanken. Für seine freundliche Art zuzuhören und Gedanken abzuwägen, um dann genau festzustellen wo ich „Recht hab“ oder wo er „sich nicht so sicher“ war, möchte ich ihm danken.

Martin Nowak und Arne Traulsen danke ich für eine unheimlich bereichernde Woche. Ganz besonders Arne hat dankenswerterweise auch danach immer wieder Zeit gefunden, mir über mathematische Hürden zu helfen.

Jörn Scharsack hat sein Wissen über histologische Färbetechniken und seine Antiseren mit mir geteilt. Sascha Hibbeler und Tobias Lenz haben mir bei molekularbiologischen Experimenten wertvolle Starthilfe gegeben.

Ohne Helga Luttmann, Anja Hasselmeyer und Verena Büscher wäre der Übergang von der Theorie zur Praxis sicher nicht so sanft verlaufen.

Marko Rohlfs danke ich für seine Freundschaft und so manchen guten Tipp und manche Hilfe seit dem ersten Semester.

Bei Thomas Bauer möchte ich mich besonders dafür bedanken, dass er sich so kurzfristig bereit erklärt hat, die Korrektur der Arbeit zu übernehmen.

Meinen Eltern, meiner Schwester und meinen Freunden danke ich für viele schöne Stunden in den letzten Jahren.

11 Stellungnahme

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbständig angefertigt und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe.

Mit der Einstellung dieser Arbeit in die Fachbibliothek des Biologiezentrums sowie in die Zentralbibliothek der Christian-Albrechts-Universität zu Kiel bin ich einverstanden.

Kiel, _____ _____

Benno Wölfing