Autoimmune Diseases and Peptide Variations

Wataru Honda[1] honda@kuicr.kyoto-u.ac.jp
Shuichi Kawashima[2] shuichi@hgc.jp
Minoru Kanehisa[1] kanehisa@kuicr.kyoto-u.ac.jp

[1] Bioinformatics Center, Institute for Chemical Research, Kyoto University, Gokasho Uji, Kyoto 611-0011, Japan
[2] Human Genome Center, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai Minato-ku, Tokyo, 108-8639, Japan

Abstract

The immune system plays an essential role in the defense of the host against invaders. Enormous numbers of lymphocytes are recruited and immense numbers of antibodies or cytokines are secreted in various kinds of immune response. But the system also has the possibility of being the cause of tissue injury or some kind of diseases. For example, when their functions target their host, an autoimmune disease occurs. Although the pathogenesis of various autoimmune diseases has been scrutinized intensively, there is little evidence as of yet. But it has been reported that in most of the disease subjects, a broad spectrum of antibodies recognizing components of self tissue or circulating self antigens that normally should be ignored are observed. In this study, we come to the conclusion that proteins targeted by these autoreactive antibodies share the same peptides with some kind of proteins of viruses known to infect human. This result supports the fact that viral infection is a speculative cause of the disease in some subjects.

Keywords: autoimmune disease, autoreactive antibody, peptide

1 Introduction

The adaptive immune system serves the indispensable function of host defense against pathogenic microorganisms, but the system is also known to cause damage to self tissue and cause disease. Disease caused by failure of self-tolerance and subsequent immune responses against self, or autologous, antigens are called autoimmune diseases [16]. Among autoimmune diseases, many types of antibody-mediated disorders are well known, but have a poorly understood etiology [6]. As described in Figure 1, the causes of antibody-mediated diseases are either antigen-antibody complexes that form in the circulatory system and deposited in vessel walls or antibodies that recognize antigens in particular cells or extra-cellular tissues [28]. Antibodies opsonize self-cells and may be followed by activation of a complement system, which leads to phagocytosis of the cells via Fc receptors or C3 receptors (Figure 1-1). Self-cells with injured plasma membrane are recognized by autoreactive antibodies targeting nucleus components and forming immune complexes, so called Haematoxylin, that are phagocytosed by neutrophils (LE phenomenon). Antibodies recognizing antigens on self-tissues or self-cells recruit leukocytes by binding to Fc receptors or activating the complement system and thereby releasing by-products that are chemotactic for leukocytes (Figure 1-2). Antibodies may be deposited as immune complexes that are formed in the circulatory system, which leads to complement- and Fc receptors-mediated recruitment and activation of inflammatory cells secreting neutrophil granule enzymes, and reactive oxygen intermediates (Figure 1-3).

Antibodies found in subjects of antibody-mediated diseases are known usually to recognize molecules included mainly in the nucleus. For instance, SLE (Systemic Lupus Erythematosus) is characterized

by antibodies that bind to dsDNA, nucleoproteins and so on [22, 35]. Some antigens associated with antibody-mediated autoimmune diseases have been identified, but a common biological feature lying behind them has not been found yet. We have considered that a possible cause of autoimmune disease is associated with the boundary between self and nonself. The immune system in organisms exists to strictly demarcate self from nonself, whereas the boundary has remained ambiguous. While huge numbers of studies have so far revealed epitopes only on some famous proteins derived from pathogenic micro organisms [8, 9], there are few cases that analyze self-nonself discrimination comprehensively [4, 14]. To make ambiguous boundary of self and nonself distinct is important for profound understanding of immune system and diseases caused by misrecognition, that is, autoimmune diseases. We compared peptide variations on proteins between human and viruses known to infect human in order to investigate a relation between autoimmune diseases and a boundary of self-nonself.

Our results clarify that peptide patterns extracted from human proteins are totally different that of from viral proteins, which means that a viral peptide fragment found on a groove of MHC molecule class I carries enough information to be discriminated from host protein segments. But, though they are very small in number, there are a few problematic peptides that are found on both human proteins and viral proteins, indicating that these shared peptides have possibility to be judged as either self or nonself by killer T-lymphocytes. Following explanation about autoimmune diseases being considered, we concluded that these shared peptides are one of the speculative pathogenesis of autoimmune diseases with serious disease presentation. In this study, we clarified that many proteins recognized by autoreactive antibodies are sharing peptides with viral proteins. This result should be important not only to pharmacology but also to the understanding of the pathogenesis of autoimmune diseases.

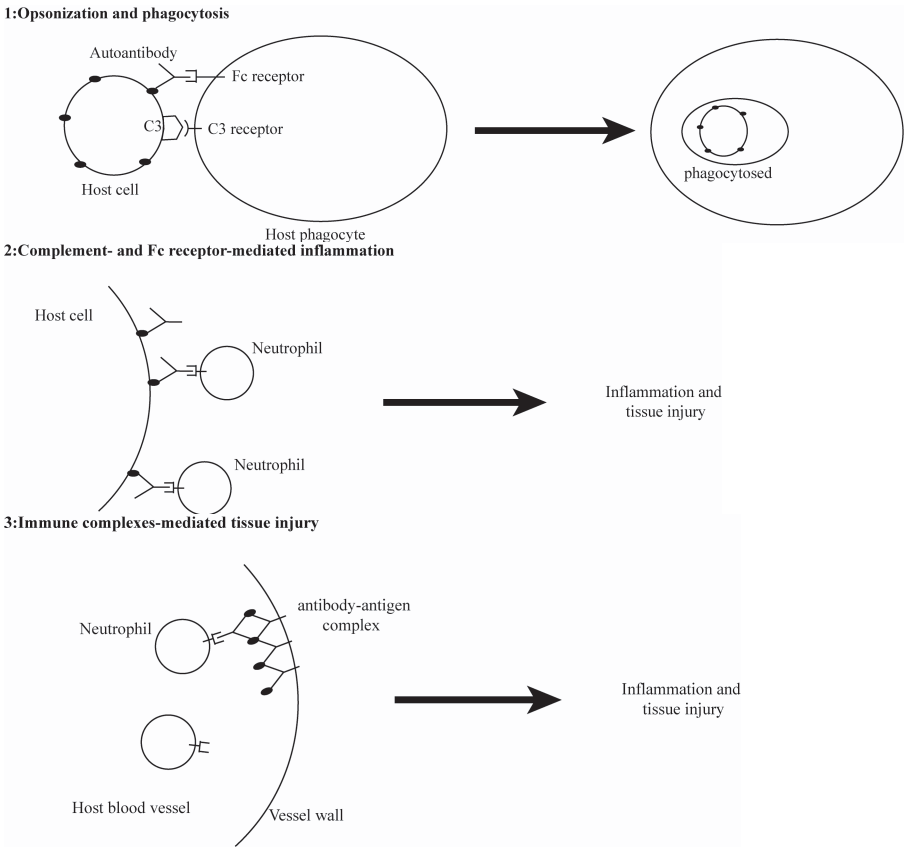

Figure 1: Examples of mechanisms behind autoreactive antibody-mediated disease.

2 Method and Data Set

The KEGG/VGENOME database [19] has 817 virus genomes, and we manually selected 659 genomes of viruses either known to infect human or suspected to infect human. All the human protein sequences and all protein sequences from viruses we selected were obtained from the KEGG/GENES and VGENES databases (Table 1). We extracted all patterns of 4∼10-mer peptides in the virus and human proteins and listed patterns by exact matching observed only in viruses.

Table 1: Number of amino acids and proteins.

	Number of amino acids	Number of proteins
H.sapiens	8,359,195	16,468
viruses	3,842,080	24,745

3 Result

3.1 Peptide Variations in Human and Viruses

In a human immune response, major histocompatibility complex (MHC) class I molecules, antigen-presenting molecules, play a pivotal role. Viral proteins translated in a host cell, so called endogenous antigens, are degraded into peptides composed of 8-10 amino acids via the ubiquitin-proteasome pathway [21]. Many of these degraded peptides are loaded and presented by MHC class I molecules on the host cell surface. On the cell surface, MHC class I molecules with loaded viral derived peptides fragments are recognized by TCR (T-cell receptor) expressed on a Killer T lymphocyte as a token of the viral infected cell [27]. Additionally, these peptides share certain patterns of sequences called MHC class I binding motifs [3]. This well-known fact explains that merely 8∼10 residues of amino acid sequences are enough in length to be discriminated from host protein fragments. Here we analyze all 4∼10-mer peptide patterns in all virus proteins.

Table 2: Number of oligopeptides in viruses.

	8residues	9residues	10residues
All oligopeptides in virus	3,039,755	3,075,199	3,102,211
The number of oligopeptides unique in virus	3,029,380	3,071,667	3,100,029
	99.7%	99.9%	99.9%
The number of motif sharing		887,029	1,532,018
		28.9%	49.4%

As shown in the Figure 2, the degree of coincidence in oligopeptide sequence patterns between *H.sapiens* and viruses decline markedly around 5∼7-mers. Between 8∼10-mers, known as suitable lengths for MHC class I groove, almost all the oligopeptides derived from virus protein sequences are unique for viruses. Next, we examined matching of 9, 10-mer oligopeptides unique for virus and MHC class I binding motifs. Known MHC class I binding motifs were constructed as PROSITE patterns and represented in regular expressions based only on anchor and auxiliary anchor residues data from SYFPEITHI [26]. Anchor and auxiliary anchor residues are defined by the score that is calculated by the frequency of amino acids in the respective position in aligned peptides. That is, motifs we employed include only anchor and auxiliary anchor residues that are often found in presented peptides.

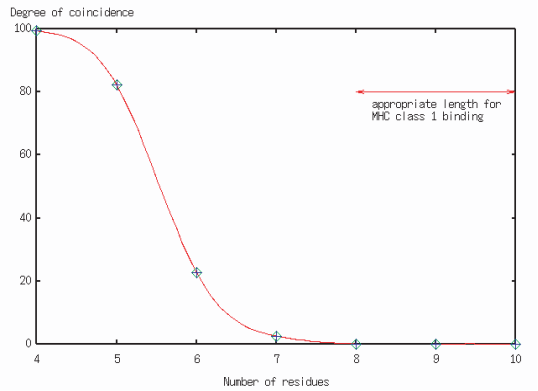

Figure 2: Degree of coincidence.

Table 3: Example of MHC binding motifs.

Type of MHC class I	Motif (PROSITE patterns)	Regular Expressions
HLA-A*31012	x(1)-[LVYF]-[FLYW]-x(2)-[LFVI]-x(2)-R	\w[LVYF][FLYW] \w{2}[LFVI] \w{2}R
HLA-B*39011	x(1)-[RH]-x(3)-[IVL]-x(2)-L	\w[RH] \w{3}[IVL]\w{2}L
HLA-Cw*0401	x(1)-[YPF]-x(3)-[VIL]-x(2)-[LFM]	\w[YPF] \w{3}[VIL] \w{2}[LFM]

In the 120 motifs we used, there were some clear features. For instance, hydrophobic residues are preferred in the carboxy terminus, and aromatic residues are often seen nearby in the amino terminus [5, 11]. These allocations of amino acids in the motifs are suitable for binding to MHC hydrophobic groove. In the case of 10-mers, we found almost half of the peptide sequences observed only in virus proteins shared common features with MHC class I binding motifs. Although several motif patterns have a possibility to overestimate the number of peptide matching motifs because of the low selectivity of our regular expressions based only on anchor and auxiliary anchor residues, this result shows us that viruses are encoded by totally different amino acid sequences compared with human. This difference is remarkably interesting immunologically.

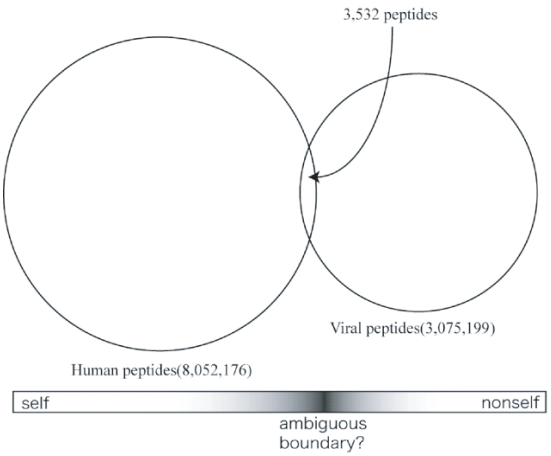

Figure 3: Self-nonself boundary based on 9-mer peptide variations.

As represented in Figure 3 and Table 2, the number of peptides shared by both human proteins

and viral proteins is 3,532, which is only 0.0004% of human peptides and 0.001% of viral peptides. Additionally, these shared peptides are found in 1849 human proteins and 1172 viral proteins. But these proteins encompass all kind of proteins such as cell adhesion molecule, nucleotide polymerase, cytokine, receptor and so on. That is, no outstanding characteristics among these proteins have been identified so far.

3.2 Autoimmune Diseases and Peptide Fragments

Next we investigated whether these shared peptides are included by proteins known to associate with autoreactive antibody mediated diseases. The number of peptides found on proteins known to associate with autoimmune diseases amounts to 193. Table 4 shows a list of proteins (with GENES id) that are identified in well-known autoimmune diseases and examples of sharing sequence. In this list, both systemic diseases and idiosyncratic diseases for specific organs or tissues are included. Listed antigens include cross-reactively recognized proteins. For example, in SLE patients various autoreactive antibodies targeting other proteins are found and it has not been determined whether all the characteristic autoreactive antibodies are pathogenic, suggesting that idiotypic network dysregulation is also one of conjectural pathogenesis [31].

Table 4: Autoimmune disease and autoreactive antibodies.

Disease	Antigen recognized by autoantibody	Sharing sequence
Pernicious anemia [17]	H^+K^+ATPase (hsa:23439)	EEEEAEEEA
Autoimmune hepatitis [33]	Cytochrome P450 (hsa:1579)	SLLILLLLL
CREST syndrome [10]	Centromere proteins (hsa:1059)	EEEEEGEGE
		DEEEEEDEE
Systemic lupus erythematosus [7]	RNA polymerase (has:55703)	PGGYFIVKG
		GEMERDCLI
	Histone (hsa:3008)	APAAPAAPA
	snRNPs (hsa:54433)	GGRGGGGGG
Systemic sclerosis [15]	DNA topoisomerase II (hsa:7155)	GGKDAASPR
		DAASPRYIF
Rheumatic fever [20]	Myosin (hsa:4650)	LDSKSLKLI
		NRIIHRDVK
Type II diabetes mellitus [29]	Insulin receptor (hsa:3643)	SFGVVLWEI
Pemphigus vulgaris [12]	Cadherin (hsa:27253)	GGGTGGGGG
		AVAAVAAAG
Goodpasture's syndrome [18]	Type IV collagen (has:1288)	GAVGPAGPP
Graves' disease [13]	Thyroid hormone receptor (hsa:9967)	SKSRSRSRS

4 Discussion

In this study, we clarified that proteins associated with autoimmune diseases share peptides with viral proteins, suggesting a possibility that these shared peptides are recognized by host immune system as pathogenic. Normally autoreactive T-lymphocytes or B-lymphocytes are thought to be killed in thymus or bone marrow in positive and negative selection process. But now it is known that the processes are not complete and therefore a number of lymphocyte having possibility to attack host are released from such lymphoid tissues [25]. Lymphocytes having possibility to recognize self-components dispatched from lymphoid tissues are in state called anergy [30, 32]. Anergy T-lymphocytes are known to become responsive against self-components when IL-2 molecules stimulate them [2] and

anergy B-lymphocytes are also known to become responsive against self-antigens when CD40-ligands bind to them [24]. Additionally, viral infection-driven IFN-gamma production leads to stimulation and proliferation of B-lymphocytes targeting autoantigens. Therefore autoreactive lymphocytes are considered to become responsive to self-antigens in peripheral blood.

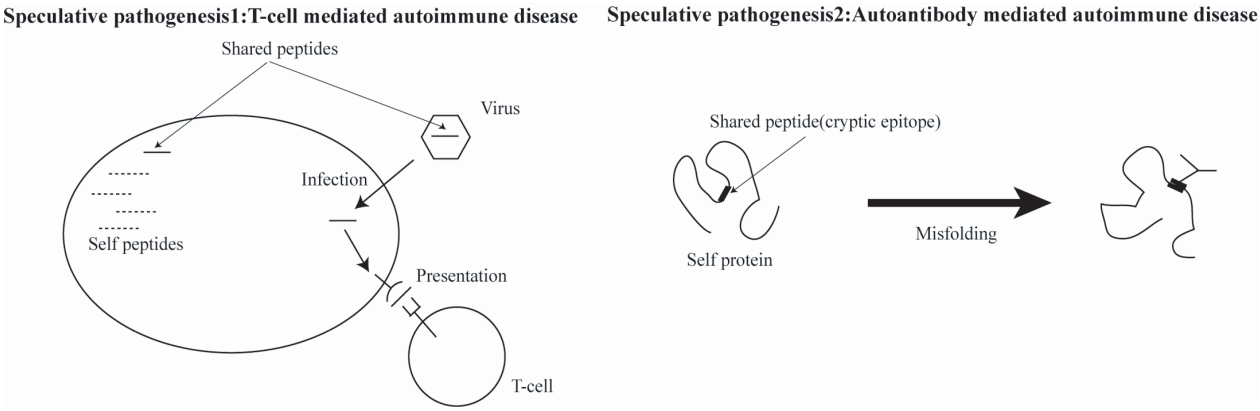

Figure 4: Speculative pathogenesis based on shared peptides.

As represented in Figure 4, we propose two speculative pathogenesis of autoimmune diseases based on shared peptides we observed. Our conjecture is closely related with molecular mimicry hypothesis according to which T-lymphocytes kill self-cells expressing MHC molecules with peptides resembling those excised from viral proteins [23, 36]. That is, our results propose that cross-reactive peptide antigens in molecular mimicry hypothesis are composed of the shared peptides we found here.

T-lymphocyte mediated speculative pathogenesis is shown in left side of the figure 4. In some diseases, killer T-lymphocytes kill host cells and subsequently trigger tissue injury or inflammation. Our result infers that viral peptides sharing the same peptides with human are presented and recognized accidentally by T-cells. Viral infections induce an activation of helper T-lymphocytes, which leads to secretion of interleukin or interferon molecules. As mentioned above, such molecules trigger an emancipation of autoreactive lymphocytes from anergy state and subsequently recruit these cells to injure self-components. Viral infection and the following IFN-gamma production also promote expression of MHC molecules on a host's cell surface. That means density of MHC molecules binding self peptides increases. It would be a cause of autoimmune diseases. Therefore viral infections are considered to be one speculative cause of autoimmune diseases [1].

The picture on the right shows conjectures that shared peptides are cryptic epitopes of self proteins. A couple of autoimmune diseases are reported that viral infection and subsequent tissue injury cause the disease because of exposure of cryptic epitopes. For example, rheumatic fever follows streptococcal infection with some serologic types of beta-hemolytic streptococci. Infection leads to the production of antibodies against a bacterial wall protein (M protein). Some of these antibodies are known to cross-react with myocardial sarcolemmal proteins and myosin and are deposited in the heart and subsequently cause inflammation (carditis) [34]. Our result shows a possibility that these cryptic epitopes are encompassing peptides shared by both human and viruses.

In this study, we investigated a characteristic of human proteins sharing peptides with viral proteins. Analysis of viral proteins and viruses including these peptides are ongoing. Our results will provide clinically important evidence of autoimmune diseases.

Acknowledgments

We would like to express our gratitude to Dr. Kiyoko F. Aoki-Kinoshita for helpful comments and an overall improvement of our manuscript. This work was supported by the grants from the Ministry of Education, Science, Sports and Culture, the Science and Technology Agency, and the Japan Society for the Promotion of Science. The Computational resource was provided by the Bioinformatics Center, Institute for Chemical Research, Kyoto University and the Super Computer System, Human Genome Center, Institute of Medical Science, the University of Tokyo.

References

[1] Atkins, G.J., McQuaid, S., Morris-Downes, M.M., Galbraith, S.E., Amor, S., Cosby, S.L., and Sheahan, B.J., Transient virus infection and multiple sclerosis, *Rev. Med. Virol.*, 10(5):291–303, 2000.

[2] Bendiksen, S. and Rekvig, O.P., Interleukin-2, but not interleukin-15, is required to terminate experimentally induced clonal T-cell anergy, *Scand. J. Immunol.*, 60(1-2):64–73, 2004.

[3] Bouvier, M. and Wiley, D.C., Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules, *Science*, 265(5170):398–402, 1994.

[4] Burroughs, N.J., Boer, R.J., and Kesmir, C., Discriminating self from nonself with short peptides from large proteomes, *Immunogenetics*, 56:311–320, 2004.

[5] Carbone, F.R., Conformational constraints involved in MHC class I restricted antigen presentation, *Int. Rev. Immunol.*, 7(2):129–138, 1991.

[6] Casiano, C.A. and Tan, E.M., Recent developments in the understanding of antinuclear autoantibodies, *Int. Arch. Allergy. Immunol.*, 111(4):308–313, 1996.

[7] Chabre, H., Amoura, Z., Piette, J.C., Godeau, P., Bach, J.F., and Koutouzov, S., Presence of nucleosome-restricted antibodies in patients with systemic lupus erythematosus, *Arthritis Rheum.*, 38(10):1485–1491, 1995.

[8] Engelhard, V.H., Structure of peptides associated with class I and class II MHC molecules, *Annu. Rev. Immunol.*, 12:181–207, 1994.

[9] Engelhard, V.H., Structure of peptides associated with MHC class I molecules, *Curr. Opin. Immunol.*, 6(1):13–23, 1994.

[10] Gonzalez-Buitrago, J.M., Gonzalez, C., Hernando, M., Carrasco, R., Sanchez, A., Navajo, J.A., and Papisch W., Antibodies to centromere antigens measured by an automated enzyme immunoassay, *Clin. Chim. Acta.*, 328(1-2):135–138, 2003.

[11] Grey, H.M., Ruppert, J., Vitiello, A., Sidney, J., Kast, W.M., Kubo, R.T., and Sette, A., Class I MHC-peptide interactions: Structural requirements and functional implications, *Cancer Surv.*, 22:37–49, 1995.

[12] Hertl, M. and Riechers, R., Analysis of the T cells that are potentially involved in autoantibody production in pemphigus vulgaris, *J. Dermatol.*, 26(11):748–752, 1999.

[13] Hirooka, Y., Li, C., Takagi, J., Gotoh, M., Habu, S., Yasaka-Nomura, T., Ishihara, R., Nakasone, Y., Nakamura, R., Morikawa, R., Otake, K., Nogimori, T., and Ishizuki, Y., Comparison of new different assay systems for thyrotropin receptor antibodies with reference to thyroid-stimulating

antibodies and thyroid stimulation-blocking antibodies in Graves' disease, *Int. J. Clin. Pharmacol Res.*, 24(4):111–116, 2004.

[14] Honda, W., Kawashima, S., and Kanehisa, M., Self-nonself discrimination based on incompatibility of amino acid sequences of human and viruses, *Genome Informatics*, 14:432–433, 2003.

[15] Hu, P.Q., Fertig, N., Medsger, T.A. Jr, and Wright, T.M., Molecular recognition patterns of serum anti-DNA topoisomerase I antibody in systemic sclerosis, *J. Immunol.*, 173(4):2834–2841, 2004.

[16] Kalden, J.R., Breedveld, F.C., Burkhardt, H., and Burmester, G.R., Immunological treatment of autoimmune diseases, *Adv. Immunol.*, 68:333–418, 1998.

[17] Kalha, I. and Sellin, J.H., Common variable immunodeficiency and the gastrointestinal tract, *Curr. Gastroenterol Rep.*, 6(5):377–383, 2004.

[18] Kalluri, R., Meyers, K., Mogyorosi, A., Madaio, M.P., and Neilson, E.G., Goodpasture syndrome involving overlap with Wegener's granulomatosis and anti-glomerular basement membrane disease, *J. Am. Soc. Nephrol.*, 8(11):1795–1800, 1997.

[19] Kanehisa, M., Goto, S., Kawashima, S., Okuno, Y., and Hattori, M., The KEGG resource for deciphering the genome, *Nucleic Acids Res.*, 32(Database issue):277–280, 2004.

[20] Khanna, A.K., Nomura, Y., Fischetti, V.A., and Zabriskie, J.B., Antibodies in the sera of acute rheumatic fever patients bind to human cardiac tropomyosin, *J. Autoimmun.*, 10(1):99–106, 1996.

[21] Lauvau, G., Kakimi, K., Niedermann, G., Ostankovitch, M., Yotnda, P., Firat, H., Chisari, F.V., and van Endert, P.M., Human transporters associated with antigen processing (TAPs) select epitope precursor peptides for processing in the endoplasmic reticulum and presentation to T cells, *J. Exp. Med.*, 190(9):1227–1240, 1999.

[22] Lipsky, P.E., Systemic lupus erythematosus: an autoimmune disease of B cell hyperactivity, *Nat. Immunol.*, 2(9):764–766, 2001.

[23] Mackay, I.R. and Rowley, M.J., Autoimmune epitopes: autoepitopes, *Autoimmun. Rev.*, 3(7-8):487–492, 2004.

[24] Martinez-Barnetche, J., Madrid-Marina, V., Flavell, R.A., and Moreno, J., Does CD40 ligation induce B cell negative selection?, *J. Immunol.*, 168(3):1042–1049, 2002.

[25] Nossal, G.J., Negative selection of lymphocytes, *Cell*, 76(2):229–239, 1994.

[26] Rammensee, H., Bachmann, J., Emmerich, N.P., Bachor, O.A., and Stevanovic, S., SYFPEITHI: Database for MHC ligands and peptide motifs, *Immunogenetics*, 50(3-4):213–219, 1999.

[27] Rock, K.L., Gramm, C., Rothstein, L., Clark, K., Stein, R., Dick, L., Hwang, D., and Goldberg, A.L., Inhibitors of the proteasome block the degradation of most cell proteins and the generation of peptides presented on MHC class I molecules, *Cell*, 78(5):761–771, 1994.

[28] Rose, N.R. and Bona, C., Defining criteria for autoimmune diseases (Witebsky's postulates revisited), *Immunol. Today*, 14(9):426–430, 1993.

[29] Sacks, D.B. and McDonald J.M., The pathogenesis of type II diabetes mellitus. A polygenic disease, *Am. J. Clin. Pathol.*, 105(2):149–156, 1996.

[30] Seroogy, C.M. and Fathman, C.G., T-cell anergy: From phenotype to genotype and back, *Immunol. Res.*, 28(3):255–264, 2003.

[31] Shoenfeld, Y., Common infections, idiotypic dysregulation, autoantibody spread and induction of autoimmune diseases, *J. Autoimmun.*, 9(2):235–239, 1996.

[32] Tsubata, T. and Honjo, T., B cell tolerance and autoimmunity, *Rev. Immunogenet.*, 2(1):18–25, 2002.

[33] Vergani, D. and Mieli-Vergani G., Autoimmune hepatitis, *Minerva Gastroenterol. Dietol.*, 50(2):113–123, 2004.

[34] Wu, X., Liu, B., Van der Merwe, P.L., Kalis, N.N., Berney, S.M., and Young, D.C., Myosin-reactive autoantibodies in rheumatic carditis and normal fetus, *Clin. Immunol. Immunopathol.*, 87(2):184–192, 1998.

[35] Yee, W.S. and Weisbart, R.H., The fine specificity of IgG antiguanosine antibodies in systemic lupus erythematosus, *Clin. Immunol. Immunopathol.*, 36(2):161–167, 1985.

[36] Zhao, Z.S., Granucci, F., Yeh, L., Schaffer, P.A., and Cantor, H., Molecular mimicry by herpes simplex virus-type 1: Autoimmune disease after viral infection, *Science*, 279(5355):1344–1347, 1998.