The pathogenicity of self-antigen decreases at high levels of autoantigenicity: a computational approach

Anmar Khadra¹, Pere Santamaria² and Leah Edelstein-Keshet³

¹Laboratory of Biological Modeling, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Building 12A, Room 4007, 12 South Drive MSC 5621, Bethesda, MD 20892-5621, USA
²Julia McFarlane Diabetes Research Centre, Department of Microbiology and Infectious Diseases, Faculty of Medicine, University Calgary, Calgary, Alberta T2N 4N1, Canada
³Department of Mathematics, University of British Columbia, Vancouver, British Columbia V6T 1Z2, Canada

Correspondence to: A. Khadra; E-mail: khadraa@niddk.nih.gov

Transmitting editor: P. Ohashi

Received 23 December 2009, accepted 7 April 2010

Abstract

Recent experimental evidence suggests that antigenic stability facilitates antigen shuttling from target tissue to dendritic cells (DCs), enabling cross-priming of naive T cells. On the other hand, antigenic stability affects the efficiency of peptide–MHC (p-MHC) complex formation, altering a target cell’s susceptibility to killing by the resulting CTLs. Using a mathematical model, we show how antigenic stability and p-MHC production efficiency interplay in autoantigenicity and pathogenic potential of target cell proteins in autoimmune disease. We consider protein allocated to both rapidly degraded versus stable functional pools [fractions \( f_1, f_2 \)], contributing, with relative efficiency \( h_t \), to p-MHC presentation on a target cell, as well as to cross-presentation on a DC; we analyze the combined effect of these parameters. Our results suggest that autoantigenicity and pathogenicity (ability to elicit T cell activation versus target cell lysis) are not equivalent and that pathogenicity peaks at low to moderate levels of autoantigenicity.

Keywords: autoimmunity, CD8⁺ T cells, mathematical model, p-MHC efficiency, protein allocation

Introduction

During growth and metabolism of a normal cell, proteins of all types are synthesized and incorporated into cellular domains but also degraded, processed into peptides by proteasomes and presented on the cell surface as peptide–MHC complexes (p-MHC). The number of p-MHC complexes displayed per protein molecule degraded (generally a small value in the range of \( 10^{-2} \) to \( 10^{-3} \)) is called the efficiency of processing of the given protein into p-MHC (1–5). Here we are concerned with aspects of protein allocation and efficiency that affect whether the given protein could become a pathogenic self-antigen.

In certain autoimmune diseases, normal healthy cells become targets for destruction by autoreactive CD8⁺ CTLs. Activation and recruitment of pathogenetic autoreactive CTLs require two fundamental steps: first, self-p-MHC complexes must be presented to peripheral naive autoreactive CD8⁺ T cells by cross-priming antigen presenting cells (APCs), i.e. autoreagnosten-loaded dendritic cells (DCs) in the target organ’s draining lymph nodes, and second the same p-MHC complexes must be displayed on the target cells’ surface as triggers of the CTLs’ cytolytic machinery. Whereas the first event elicits differentiation of naive autoreactive T cells into CTLs and recruitment of the differentiated CTLs into the target organ through the bloodstream, the second event marks target cells for cognate recognition and destruction by the recruited CTLs.

Type 1 diabetes (T1D) is prototypic organ-specific autoimmune disease that results from selective destruction of pancreatic β cells by autoreactive T cells (6–8). Non-obese diabetic (NOD) mice spontaneously develop a form of T1D, closely resembling human T1D, that is caused by T cells recognizing a growing list of autoantigens (9). Initiation of T1D requires CD4⁺ and CD8⁺ T cells (10, 11).

It has been proposed that autoreactive CD4⁺ and CD8⁺ T cells differentiate into effectors by engaging cognate p-MHC complexes along with co-stimulatory molecules on professional cross-priming APCs, such as DCs, capable of shuttling β cell autoantigens from the pancreas to the pancreatic lymph nodes (PLNs) through afferent lymphatics; that T1D is initiated by CD4⁺ T cells, presumably insulin reactive (12), and that CD8⁺ T cells contribute to T1D by killing β cells (13).

A significant fraction of islet-associated CD8⁺ T cells in NOD mice uses Va17*Ja42* T cell receptors (TCRs), which we refer...
What makes self-antigen pathogenic?

To as 8.3-like (14–17). These cells are already a notable component of the earliest islet CD8+ infiltrates (17–19), are diabetogenic (15, 16) and target a peptide from islet-specific glucose-6-phosphatase catalytic subunit-related protein (IGRP) in the context of the MHC class-I molecule H-2Kd (20).

The CD8+ T cells that recognize IGRP,26-21/Kd are unusually frequent in the circulation (21) and are not subjected to thymic negative selection due to the fact that IGRP is exclusively expressed in β cells (unlike other systemically expressed proteins such as insulin). Although islet-associated CD8+ T cells also recognize epitopes on other antigens (13, 22, 23), IGRP autoreactivity (against multiple different epitopes) accounts for ~40% of the intra-islet CD8+ T cell pool, indicating that IGRP is a prevalent autoantigen in murine T1D (24).

An intriguing question is why certain β cell proteins, such as IGRP, take on a dominant autoantigenic role, whereas many other β cell proteins that are also exclusively expressed in β cells do not. It is also unclear whether autoantigenicity (the ability of a given protein to elicit T cell responses) correlates with pathogenicity (the ability of these T cell responses to trigger target cell death).

Here, we use a mathematical model to explore these issues and test the hypothesis that, unlike autoantigenicity, pathogenicity is inhibited at high levels of protein stability. We describe protein stability by an allocation parameter, defined as the fraction of molecules of a given protein directed to a rapidly degraded pool (a fraction f) versus the fraction (1 − f) that resists degradation by persisting in a stable functional state. An important aspect of these pools is their relative efficiency of processing of the protein into p-MHC, which need not be identical for the two pools, even if the protein is itself the same. (For example, cellular localization could influence that efficiency.)

We argue below, based on current experimental evidence, that both the relative allocation and the relative efficiency of these pools can directly affect the two fundamental principles underlying autoreactive CTL activation, recruitment and target cell cytotoxicity; namely that (i) the rapidly degraded protein pool contributes predominantly to the pool of p-MHC complexes displayed on the surface of target cells, hence pathogenicity, whereas (ii) the stable protein pool contributes most strongly to cross-presentation of p-MHC by DC, hence autoantigenicity. Both slow and rapid protein pools will contribute to p-MHC presentation on the cell surface but generally at different rates and efficiencies. Also, the greater the pathogenic potential of a given autoantigen, the higher the likelihood that recognition of its p-MHC derivatives will foster autoantigenic clonality of target cells to DCs (by eliciting antigen shedding). This, we show, could lead to a positive feedback loop that drives the system autocatalytically.

We base this hypothesis on a growing body of evidence in the literature. First, as established by refs (1–4), a solid link exists between the rate of protein degradation in a normal cell and its rate of (direct) p-MHC presentation. Second, several observations point to the fact that DCs preferentially cross-present p-MHC complexes derived from intact cellular proteins that they take up [see refs (25–27) and references therein]. Proteins that resist degradation are more likely to survive apoptotic and necrotic cell death, hence to access the cross-presentation pathway in DCs.

As both direct and cross-presentation are complex phenomena, verbal and intuitive arguments can fall short of deciphering the link between assumptions and rigorous predictions. For example, one could intuitively argue verbally that moderate levels of allocation (i.e. f ≈ 0.5) would be most consistent with the dual aspects of autoantigenicity and pathogenicity. However, this turns out to be true only in a restricted set of cases. In fact, predictions depend more subtly on the allocation to and relative processing efficiency of the proteins in the two pools. How these two quantities are combined with the many other parameters governing immune response is not transparent and difficult to decipher from simple arguments lacking a formal structure. For this reason, we investigated our hypothesis within the context of a mathematical model, where detailed assumptions are formalized, and predictions (based on objective, detailed analysis) and simulations are made.

Methods

Experimental methods

Experiments on antigen processing and presentation on which our model is based are described in refs (2, 5, 28). For T cell avidity and dynamics, together with β cell and APC mass and turnover, the model uses data available in refs (24, 29, 30). Experimentally, protein synthesis and turnover and p-MHC complex formation and turnover have been measured by infecting target cells (mouse L-Kd) with a panel of recombinant vaccina viruses expressing chimeric proteins containing nucleoprotein infused with green fluorescence protein. T cell avidity is measured with p-MHC tetramers, multimeric complexes of p-MHC labeled with fluorescent marker (31).

Mathematical models

A system of ordinary differential equations is used to model the dynamics of autoantigen processing and presentation in target cells such as β cells. We use data and observations specific for viral proteins described in ref. (2), as well as IGRP, a dominant T1D autoantigen present in β cells. Parameter estimation is performed by using our own experiments and the data available in the literature. We apply reasonable assumptions and utilize common non-linear dynamic techniques to analyze the model theoretically and numerically.

Software

Steady-state analysis, bifurcation analysis as well as numerical integration of the model were carried out using the public domain software package XPPAUT (http://www.cnbc.cmu.edu/~bard/files.html) and MATLAB.

Results

The model

A schematic version of our model is shown in Fig. 1, with a more detailed view of the protein processing shown in Fig. 2. The model tracks the time dynamics of mean protein and p-MHC levels in a single target cell [Q(t); the level of normal protein, R(t); the level of protein targeted for rapid processing and P1(t); the level of p-MHC displayed on the
What makes self-antigen pathogenic?

Fig. 1. A summary of the main ingredients of the model. A given protein is allocated to two pools, \( Q \) and \( R \), each contributing differently to the production of p-MHC \((P_1)\) on the surface of a target cell. (In this diagram, pool \( R \) consists of a rapidly degraded pool of the protein, and its stable variant, \( Q \), is embedded in the endoplasmic reticulum membrane.) Protein in pool \( Q \) survives apoptosis intact, to be taken up by APCs. It is cross-presented as p-MHC \((P_2)\) that can trigger naive autoreactive T cells to be activated and to produce a population of CTLs. These effector T cells recognize and kill target cells displaying similar p-MHC, generating more apoptotic bodies. Under suitable conditions, this feedback amplifies T cell expansion and leads to exponential growth of CTLs, representative of autoimmune disease initiation. Here we investigate the effect of allocation fraction \( f \) and relative p-MHC production efficiency \( \eta \) on that amplification. See also Fig. 2 for other details.

Regarding CTLs, we assumed that the rate of activation of autoreactive CD8\(^+\) T cells depends on the p-MHC displayed on DCs, \( P_2 \), and that CTLs kill target cells at a rate that depends on target cell p-MHC, \( P_1 \). We also assumed, for simplicity, that the number of DCs, \( A \), the number of target cells, \( B \) (e.g. \( \beta \) cells in the case of T1D), and the number of circulating naive CD8\(^+\) T cells are constant over the time span of interest, which is the initiation phase of autoimmunity.

We are concerned in this paper with how peptide processing in a target cell affects whether autoreactive T cells become amplified or depressed during early stages of an autoimmune disease. During early stages of T1D, T cell populations (if amplified) grow exponentially. The resulting death of their targets, the \( \beta \) cells, fuels autoantigen peptide uptake and furthers T cell activation. At the same time, the population of \( \beta \) cells apparently changes more gradually with symptoms of T1D appearing long after a huge expansion of T cell populations has taken place (i.e. many weeks later). This suggests that the decline of \( \beta \) cells is slow relative to the rapid growth of T cells. For this reason, during the early disease stages, it is reasonable to consider \( \beta \) cell populations as roughly constant as an approximation. Related models in which \( \beta \) cell population is dynamic are discussed in ref. (32).

Based on the above assumptions, our model equations are as follows:

\[
\frac{dQ}{dt} = S(1 - f) - a_1 Q, \tag{1a}
\]

\[
\frac{dR}{dt} = S' - a_2 R, \tag{1b}
\]

\[
\frac{dP_1}{dt} = a_2 R + a_4 Q - \mu_1 P_1, \tag{1c}
\]

\[
\frac{dP_2}{dt} = a_5 Q E K(P_1) - \mu_2 P_2, \tag{1d}
\]
Parameter values

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Meaning</th>
<th>Default value</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>Rate of protein synthesis$^a$</td>
<td>$5.33 \times 10^4$ proteins per min</td>
<td>(0–1)</td>
</tr>
<tr>
<td>$f$</td>
<td>Fraction of rapidly degraded protein$^b$</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>$a_1$ and $a_2$</td>
<td>Decay rates of $Q$ and $R$</td>
<td>0.0002, 0.07 min$^{-1}$</td>
<td>(0–1)</td>
</tr>
<tr>
<td>$\mu_1$ and $\mu_2$</td>
<td>p-MHC decay rates (on target and DC cells)</td>
<td>0.01 min$^{-1}$</td>
<td>(0.001–0.01)</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Killing rate per CTL$^c$</td>
<td>$3 \times 10^{-12}$ (T cells $\times$ p-MHC $\times$ min)$^{-1}$</td>
<td>$5 \times (10^{-13} - 10^{-9})$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>CTL expansion rate</td>
<td>$3.16 \times 10^{-5}$</td>
<td>$(10^{-8} - 10^{-4})$</td>
</tr>
<tr>
<td>$\eta_1$, $\eta_2$</td>
<td>Net p-MHC efficiencies$^d$ ($R$ and $Q$)</td>
<td>1/2000, p-MHC/protein, —</td>
<td>(1/3122–1/994), —</td>
</tr>
<tr>
<td>$\eta_3$</td>
<td>Net p-MHC efficiency in DC</td>
<td>1/1406 p-MHC/protein</td>
<td>(1/1750–1/1211)</td>
</tr>
<tr>
<td>$\eta_a$</td>
<td>Efficiency ratio $\eta_a = \eta_1/\eta_3$</td>
<td>$\sim -0.1$</td>
<td>(0–1)</td>
</tr>
<tr>
<td>$a_b$</td>
<td>$= \eta_2 B/A$</td>
<td>$\sim -4.5 \times 10^{-4}$</td>
<td></td>
</tr>
</tbody>
</table>

$^a$ Based on KEKE-NP-GFB (2).

$^b$ The fraction of protein that is rapidly allocated to the processing and presentation pathway.

$^c$ Of target cells, such as $\beta$ cells.

$^d$ Efficiency (for proteins $Q$ and $R$) is the number of p-MHC complexes formed per protein degraded. See text for discussion of how these estimates were made. The symbol $^\sim$ refers to fitted values.

\[
\frac{dE}{dt} = A(P_2) - \gamma E. \tag{1e}
\]

Equations (1a–c) are simple conservation statements for protein and p-MHC in a target cell, stating that both normal ($Q$) and rapidly degraded ($R$) pools of the given protein are synthesized (rate $S$ is the total given protein synthesized per cell per unit time) with proportion $(1 - \eta_1):f$ and contribute to the target cell’s p-MHC level, $P_1$.

Peptide displayed on DCs, $P_2$, derives from intact normal protein (of pool $Q$) from target cells that are killed. Under suitable assumptions (see Supplementary data, available at International Immunology Online), the production of $P_2$ is roughly proportional to $Q$, to the number of CTLs, $E$, and to their killing rate $K(P_1)$ (target cells killed per CTL per unit time). Here we generally assume that the killing rate of target cells by CTLs is simply proportional to the surface display of p-MHC, i.e. $K(P_1) = \kappa P_1$. In Equation (1d), $a_2$ is a dimensionless constant, depending on the ratio of DCs to target cells ($A/B$) and DC-specific p-MHC efficiency, i.e. $a_2 = \eta_2 B/A$ (see Supplementary data for details, available at International Immunology Online).

Effector T cells ($E$) are produced from circulating naive cells at rate $A(P_2)$ and die at rate $\gamma$, as depicted in Equation (1e). Equations (1d and e) are simplifications of more complex compound events that contain the positive feedback link between DC cross-presentation and CTL expansion. The dependence of the killing term in Equation (1d) on $P_1$ forms the key link between autoantigenicity and pathogenicity that we are exploring here.

We considered three possible cases for the way that the T cell activation rate, $A(P_3)$, depends on cross-presented p-MHC on DCs: (I) linear $[A(P_3) = \alpha P_3/k_3]$; (II) saturating and (III) sigmoidal $[A(P_3) = \alpha P_3/k_3^0 + P_3^2]$, with $n = 1$ for (II) and $n = 2$ for (III). Here $\alpha$ is a maximal rate of activation and $k_3$ is the level of p-MHC display on DCs leading to half-maximal activation. The parameter $k_3$ can be considered as the reciprocal of T cell avidity. Variant III is shown in the ‘activation’ panel in Fig. 1. All variants have the property that there is no T cell activation if the relevant p-MHC is not displayed on the surface of DCs ($A(0) = 0$).

In our analysis, we will focus particularly on the allocation fraction $f$ and on the ratio of p-MHC complexes formed to protein degraded, i.e. the so-called efficiency, an aspect of peptide processing that has received attention in the recent literature (1–5). Parameters in our dynamic model are related in a transparent way to efficiency (here represented by the symbol $\eta$, where subscript indicates the given pool from which peptide is derived). That is, the efficiencies of p-MHC produced from, respectively, the stable and the rapidly degraded protein are $\eta_Q = a_2/a_1$ and $\eta_R = a_2/a_2$, where $a_1$ and $a_2$ are decay rates of $Q$ and $R$ (with $a_2 > a_1$, since pool $R$ is much more rapidly degraded than $Q$) and $a_2$ are rates of p-MHC production from $R$ and $Q$, respectively. In general, these two efficiencies need not be equal (2), and evidence in ref. (4) suggests that $\eta_R < \eta_Q$. In fact, the special case in which these are precisely equal turns out to miss a range of possible behavior that occurs when their values are distinct. For this reason, we considered the possibility that $\eta_Q \neq \eta_R$ and defined their ratio, $\eta = \eta_Q/\eta_R$, as a parameter of interest. $\eta = 1$ is the limited special case of equal efficiencies of peptide processing from the two protein pools, whereas $\eta < 1$ follows the suggestion of ref. (4).

**Predicted scenario for spontaneous autoimmunity**

Normal protein production and processing will lead to some equilibration in the levels of $Q$, $R$ and $P_1$ in a target cell. It is easy to see that the model consisting of Equation (1) always has a ‘healthy’ steady state, where the given protein has no p-MHC presentation on DCs and where there are no effector cells with the given peptide specificity ($P_2 = 0$, $E = 0$). We are concerned with the possibility of spontaneous autoimmunity, i.e. we ask whether small immunological perturbations could lead to expansion of $P_2$ and $E$. This amounts to asking under what conditions this healthy steady state is unstable.

For the purpose of answering this question, it suffices to study a reduced model that treats $Q$, $R$, and $P_1$ as static variables (so-called ‘quasi-steady state’) and tracks only

---

**Supplementary data** are available at *International Immunology* Online.
What makes self-antigen pathogenic? 575

A graphical representation of this inequality is given in Fig. 3. The plot in Fig. 3(a) shows a surface whose whose height, \( z(f, \eta) \), represents the value of \((1 - f)(f + \eta(1 - f))\) as a function of the allocation fraction \( f \) and the relative efficiency \( \eta \) in the ranges \( 0 \leq f \leq 1, 0 \leq \eta \leq 2 \) (panel a). Portions of that surface above some height (labeled “threshold” in the plot as an illustrative example) would correspond to a dangerous range for autoimmunity. The level curves of this surface are also shown on the contour plot of Fig. 3(b). Walking “downward” on this plot corresponds to decreasing the quantity \((1 - f)(f + \eta(1 - f))\), making it less likely that inequality (3) is satisfied (and lowering the danger of autoimmunity). In Fig. 3(c), a similar result is shown in a different way, with a sequence of curves, each one for a specific value of \( \eta \). We refer henceforth to these curves as “efficiency clines”, since they are curves of the form \( y(f) = (1 - f)(f + \eta(1 - f)) \) along which the efficiency, \( \eta \), is constant.

We now compare predictions for a set of possible candidate autoantigen proteins with various values of \( \eta \) and \( f \). In particular, we consider what happens as parameters in the ratio \( L \) have settings that make the threshold for immunological response (black line in Fig. 3c) gradually decrease from an initially very high value (resistance to autoimmunity) to lower values (susceptibility to autoimmunity). For example, this decrease of \( L \) could happen due to the possible appearance of a clone of high-avidity autoreactive T cells, which would decrease the average value of \( \kappa_p \). We ask under what conditions, the inequality (3) would first be satisfied, leading to autoimmunization of the immune response.

Suppose some \( \eta > 1 \) is fixed (e.g. red curve in Fig. 3c). This is the case when stable proteins contribute more p-MHC per molecule than unstable proteins within host cells (likely an unrealistic scenario based on the literature). Then the efficiency cline is concave up and always highest at the value \( f = 0 \). Thus, proteins with \( \eta > 1 \) and \( f \approx 0 \) would be most immunogenic, being first to satisfy inequality (3), even when the threshold for autoimmunity is relatively high (large value of \( L \)). Note that the concavity of efficiency clines changes at \( \eta = 1 \) (yellow curve in Fig. 3c). Indeed, for \( \eta \approx 1 \), the curves have no local maxima and attain a global maximum at \( f = 0 \). The value \( \eta = 1 \) (at which efficiencies of both pools are identical) is a straight line, which is a special case. For \( \eta < 1 \), the efficiency cline has a local maximum. Elementary calculus leads to the conclusion that the maxima of the efficiency clines is then at

\[ f_{\text{max}} = \frac{2\eta - 1}{2(\eta - 1)}. \]

Note, in particular, the dependence of the location of this maximum on the value of the relative efficiency.

Now consider \( \eta = 1/2 \) (green curve in Fig. 3c). This is the case when the efficiency of p-MHC production from the rapidly degraded form is twice as large as that of the stable form. As seen from the figure, this case is at a borderline value of \( \eta \) for which the maximal value of \( y(f) \) is still at \( f = 0 \). This implies that proteins with very low \( f \) values (i.e. allocated almost completely to stable forms) are still most likely to be immunogenic. Observe, however, that the threshold value, \( L \), has to be considerably lower (\( L \leq 0.5 \))—i.e.
a greater overall susceptibility to autoimmunity is needed—in order to satisfy inequality (3).

Proteins with $g < 1/2$ no longer behave in the same way: the maximum of the efficiency cline shifts into the interior of the interval $0 < f < 1$. For example, when $g = 1/4$ (blue curve in Fig. 3c), that maximum is at $f_{\text{max}} = 1/3$, and in the extreme case of $g \approx 0$, the maximum approaches $f_{\text{max}} = 1/2$. Within the range $0 \leq g \leq 0.25$, it is no longer true that the greatest danger stems from proteins with allocation fraction $f = 0$. Indeed, in this range, the most dangerous proteins are those with intermediate values of $f$, i.e., within some small range close to $f_{\text{max}}$.

These predictions strongly suggest that, aside from the specificity of autoreactive T cells and aside from details of how the immune system is triggered (whether by viral infection or other means), the way protein is channeled between fast processing versus stable pools together with the relative efficiency of processing from each pool to p-MHC can make that protein’s participation in autoimmunity more or less likely. Proteins that are mostly degraded rapidly ($f$ close to 1) are relatively unlikely to become autoantigens, all else being equal, as shown in Fig. 3. This conclusion is independent of the relative efficiency $g$. However, which proteins are dangerous for low values of $f$ is predicted to depend on $g$. All in all, low to moderate values of $f$ are potentially pathogenic, with the intermediate values of $f$ most problematic when $g$ is very small.

**Model analysis and qualitative dynamics**

Results described above are based on analysis of steady-state existence, behavior and stability of the model. Because the reduced model contains only two variables, namely scaled p-MHC on DCs and scaled CTL population, it is readily analyzed by qualitative techniques. Such analysis indicates that there is always a healthy state ($P_2 = 0, E = 0$). There can be one or two other steady states depending on assumed T cell activation function (Cases I–III) and values of the above ratios. The analysis is carried out in the Supplementary data (available at *International Immunology* Online).
There, we show that when (3) is violated, only the healthy state exists, so that any immune reaction is resolved to a healthy state. If either \( f \) or \( \eta \) have values that satisfy this inequality for a given setting of parameters in the ratio \( L \) then an autoimmune state and a transitional (unstable) state can appear. Very small deviations away from the healthy state will lead to disease in Cases I and II, but not III. An initial state with a somewhat greater level of p-MHC display on DCs or CTL level (e.g. due to prior inflammation or infection) evolves toward autoimmunity in all the Cases (I–III). The range of initial states that does so (the ‘basin of attraction’ of the autoimmune state) is parameter dependent. Such ‘bistable’ dynamics are typical of disease models with sigmoidal functional dependencies [see, for example ref. (32)].

For saturating T cell activation (Case II), there are at most two steady states, and transitions that cause (3) to be satisfied result in the healthy state losing stability in favor of a stable disease state. The behavior of the model for linear T cell activation is less interesting overall, as the dynamics of CTLs are unbounded. However, the stability of the healthy state in that simplest case is identical to Case II of saturated activation.

Up to this point, results are obtained as fully generic predictions of the model, given our simplifying assumptions. That is, these qualitative conclusions such as the condition (3) are independent of the specific protein, providing unequivocal criteria for growth or decay of a CTL response within the framework of the model. To go further and investigate quantitative predictions for model behavior, we turn to numerical simulations. Before doing this, we must estimate reasonable values for the model’s parameters. The usefulness of this next step is that it enables a systematic investigation of how sweeping through a range of values of key parameters affects the predicted dynamics; it also helps to further delineate regions of parameter space in which the disease can occur.

Parameterizing the model

We obtained turnover rates of protein (\( a_1 \) and \( a_2 \)) and p-MHC (\( \mu_1 \) and \( \mu_2 \)) from refs (25, 28) and rates of production of p-MHC (\( a_3 \) and \( a_4 \)) from refs (2); they find that the ‘efficiency’ of p-MHC production (ratio of p-MHC per protein degraded, here denoted by constants labeled \( \eta \)) to be \( \eta \approx 1/2000 \) for rapidly degraded protein. Ref. (4) finds the efficiency for slowly degraded protein molecules (denoted by \( \eta_{3D} \)) to be much lower. However, precise values of \( \eta = \eta_{1D}/\eta_{1} \) are not known with certainty. In particular, whether \( \eta \geq 0.25 \) or \( \eta < 0.25 \), which we have shown to be significant by our analysis, is not yet known for specific proteins. For protein processed and presented by DCs, the average efficiency is \( \eta_{3D} \approx 1/1400 \) p-MHC/protein (2).

Parameters typical for T cell activation and turnover (\( \alpha, k_\rho \) and \( \gamma \)) and for \( \beta \) cell killing (\( k_\beta \)) are based on our previous modeling experience (30, 32) and on ref. (33). All parameter values are concentrated in Table 1, with detailed derivations and explanation in the Supplementary data (available at International Immunology Online). In each case, we examined the behavior of our model for a range of values of the given parameters. Results of the analytical calculations reveal how dynamics depend on parameters, as previously indicated, so we are assured that we have not missed important dynamic effects by restricting attention to a subset of possible parameter choices. This exploration, together with the bifurcation analysis (below), is yet further affirmation that conclusions are robust. (Thus, we can be confident that within the given simplified representation of the biology, we have characterized all possibilities.)

Model predictions and parameter dependence

Bistability in the dynamics. We ran the full model with \( f \approx 0.5 \) and killing rate \( \kappa = 3 \times 10^{-8} \) (T cells × p-MHC × min\(^{-1}\)) while keeping all other parameters at default values shown in Table 1, including the relative efficiency, \( \eta = 0.1 \). As anticipated from previous discussion, we observed the following results: the levels of protein (\( Q \) and \( R \)) and p-MHC (\( P_i \)) on the target cell quickly equilibrate to their steady-state levels, regardless of initial values assumed (data not shown). The levels of \( P_3 \) are shown in Fig. 4 for two initial settings and a given value of the protein allocation fraction. In the case of sigmoidal T cell activation, we note the bistability of the system: when \( f \) takes on values close to 0.5, some initial conditions very close to the healthy state (very small region) will recover to health (panel a), while those with larger immune flare up will lead to autoimmunity (panel b). In the latter case, a steady increase in the size of activated T cell pool is observed, reaching levels consistent with those observed experimentally (~10\(^6\), results not shown). This does not happen when \( f \) is close to 1, where any initial setting will lead to the healthy state (results not shown). The latter result is consistent with the observation made about the inability of rapidly degradable proteins to cross-prime T cells (27).

We also ran the model with saturated T cell activation function (results not shown). This model variant has no bistability. When parameters are consistent with the existence of the autoimmune state, all initial conditions lead to autoimmunity.

Variation of parameters can affect these results in various ways. The tools of bifurcation analysis can be used to succinctly summarize how changing one or another parameter can change the number and stability of steady states. We highlight several such studies here and concentrate others.
What makes self-antigen pathogenic?

T cell avidity affects the dynamics. The full model predicts that increasing T cell avidity (lowering $k_p$ or lowering $k$), which decreases the value of $L$ in Equation 2) promotes autoimmunity in two ways: (i) by making the disease state possible, where it was previously absent, and (ii) by increasing the basin of attraction of the autoimmune state, hence making it more likely, given an arbitrary perturbation away from the healthy state. In the panels of Fig. 5, we show the effect of the protein allocation fraction $f$ (horizontal axes) on the presence of the disease state when T cell avidity is successively increased and the value of relative efficiency is kept constant at $\eta = 0.1$. At low avidity, (panel a), only the healthy steady state is present (solid line at $E = 0$; $P_3$ is not shown), and all initial states evolve toward health. For higher avidity (panel b), a pair of new steady states at elevated $E$ levels exists over some range $f_a < f < f_b$ centered around $f = 0.5$. When $f$ is in this interval, some initial states evolve to health and others to autoimmunity, with the most effectively pathogenic values of $f$ (i.e. values of $f$ corresponding to disease states with the highest level of T cells) close to 0.5. (Solid/ dashed curves represent stable/unstable steady states.) Increasing T cell avidity even more, as in panel (c), broadens the dangerous range $(f_a - f_b)$ so that it occupies most of the interval $(0, 1)$. This makes the disease state more prevalent and also increases its basin of attraction.

Increasing the value of $\eta$ to 0.7 and carrying out the same simulations conducted in Fig. 5 reveal similar results: the range of bistability expands with increasing T cell avidity (decreasing $k$), as shown in the panels of Fig. 6. The only noticeable difference, however, between the two figures is the observed shift of the most pathogenic value of $f$ to the left in panels (b) and (c). In these two cases, $f = 0$ carries the most pathogenic potential, a result already predicted by our previous analysis.

Effect of the efficiency ratio. We also numerically confirmed the predictions of our analysis by varying $\eta$ and $f$ simultaneously, both in the range $(0, 1)$. In Fig. 7, we summarize the behavior of the full model with sigmoidal T cell activation by simultaneously displaying the dependence on $f$ and on $\eta$ (two parameter bifurcation diagram). The possibility of autoimmunity (gray region of the diagram, where the model is bistable) depends on both $\eta$ and $f$. The figure agrees with panel (b) of Fig. 3.

If $f$ is low, meaning that most of the protein is allocated to normal function and slow degradation, then increasing $\eta$ (increasing the efficiency of p-MHC production from the slow protein pool) significantly increases the likelihood of autoimmunity. (This is particularly noticeable close to $f \approx 0$, where increasing $\eta$ from 0.1 to 0.3, for example crosses into the gray ‘autoimmune’ region.) This can be understood intuitively as follows: if most of the protein is channeled to $Q$ and very little to $R$, then the level of p-MHC on the cell, $P_1$, depends more strongly on the efficiency of processing from $Q$. Thus, increasing $\eta$ is needed to produce sufficiently high level of $P_1$ to mark the cell as target for CTL killing.

For allocation fraction $f$ closer to 1, the parameter $\eta$ hardly matters, as previously noted. Indeed, for conditions shown in Fig. 7, the region $f \geq 0.75$ is white at all values of $\eta$, so only the healthy state exists. For this scenario, most protein is allocated to $R$ and rapidly degraded. We would expect that p-MHC on the target cell is plentiful, regardless of the efficiency ratio. However, as little normal protein is produced, hardly any survives for uptake by DCs and triggering of T cells. This prevents the self-amplification and no autoimmunity occurs.

T cell avidity also affects the features of diagrams such as Fig. 7 (results not shown). Increasing avidity results in a larger ‘gray’ autoimmune region, causing the triangle at the lower left corner to shrink closer to the origin and making the stripe near $f = 1$ get narrower. However, in keeping with the qualitative results, the healthy state is globally stable close enough to $f = 1$. As above, this suggests that proteins that are rapidly degraded in the cell (and thus whose $f$ is close to 1) pose little danger of becoming self-antigens in autoimmunity.

Discussion

Here we have examined a biological hypothesis describing the pathogenic potential of self-antigens in the context of

---

Fig. 5. Steady-state levels of the scaled CTL population (vertical axis) versus protein allocation fraction $f$ (horizontal axis) for low, intermediate and high T cell avidity $k$: (a) $k = 10^{-3}$, (b) $k = 8 \times 10^{-3}$ and (c) $k = 3 \times 10^{-3}$ (T cells $\times$ p-MHC $\times$ min)$^{-1}$ for the model with sigmoidal T cell activation function. In this ‘bifurcation diagram’, the stability of the steady state is indicated by solid line for stable and dashed line for unstable steady state and $\eta = 0.1$. Increasing the avidity of T cells (decreasing $k$) increases the range of $f$ over which an elevated CTL population exists with a maximum pathogenic potential occurring at $f \approx 0.5$. 

---

Downloaded from https://academic.oup.com/intimm/article-abstract/22/7/571/849146 by guest on 02 April 2019
Fig. 6. As in Fig. 5 but with η = 0.7. The range of f exhibiting bistability of the two steady states (health and disease) increases with increasing avidity of T cells (decreasing k). Observe here that the most pathogenic values of f occur at f = 0, as predicted by our analysis.

Fig. 7. The region in the $f-\eta$ parameter plane that leads to the possibility of autoimmunity, shown here in gray, as computed numerically. The remaining parameters of the model are kept at their default values listed in Table 1. Horizontal axis represents the protein allocation fraction, $f$. The vertical axis is the efficiency ratio for p-MHC production from normal versus rapidly degraded protein ($\eta = \eta_p/\eta_\eta$). The boundary of the gray region here represents one of the contours obtained in Fig. 3.

**What makes self-antigen pathogenic?**

It could be argued that the value of our model depends on how one rates the importance of direct priming versus cross-priming during the priming phase and effector phase of autoimmune disease, because changing the relative weight of the two processes might change the conclusions. It should be noted, however, that several lines of experimental evidence indicate that activation of naive autoreactive CD8$^+$ T cells requires antigen cross-presentation in the PLNs and that such activation cannot be induced by direct engagement of cognate p-MHC complexes on the surface of pancreatic β cells. First, β cells do not express the type of co-stimulatory molecules that are absolutely essential for T cell activation. In fact, even though NOD-derived NIT-1 insulinoma cells can trigger IFN-γ secretion by differentiated 8.3-CTLs, they are unable to induce the proliferation of naive 8.3-CD8$^+$ T cells or their differentiation into cytolytic effectors. Second, cross-priming in the PLNs has been established as the major pathway through which naive autoreactive CD8$^+$ T cells undergo activation in T1D (34); DCs acquire β cell autoantigens in the pancreas and ferry them to the PLNs, where they present them to cognate autoreactive CD8$^+$ T cells in the context of MHC class-I molecules (35). This is strongly supported by the observation that surgical removal of PLNs at 3 weeks of age almost completely abolished insulitis and diabetes in NOD mice (36), indicating that recognition of antigens in the PLNs is essential for T cell activation. Similarly, both islet-specific CD4$^+$ (37, 38) and CD8$^+$ T cells (39) proliferate specifically in the PLNs upon adoptive transfer into NOD mice and this precedes their homing into pancreatic islets, indicating that β cell autoantigens are presented primarily by APCs residing in the PLNs rather than by antigens displayed on the surface of pancreatic β cells. Third, it has been shown that abrogation of MHC class-I expression on pancreatic cells delayed the recruitment and reduced the retention of IGRP-reactive CD8$^+$ T cells in the islets without abrogating T cell proliferation and activation in the PLNs, demonstrating that cross-presentation of β cell autoantigens can proceed without a prior CD8$^+$ T cell attack on cells (40).

The role of antigenic stability in autoantigenicity in the context of autoimmune diseases has not been explored experimentally. The notion of antigenic stability fostering cross-presentation has been investigated in the context of viral infections and anti-viral CD8$^+$ T cell responses. Our modeling effort seeks to gain insight into how this concept can help us understand why certain proteins are preferred targets of an organ-specific autoimmune response versus others that are expressed at similar or even higher levels. Our model supports the idea that the quality of the antigen and more specifically antigenic stability and p-MHC production efficiency play
a critical role. It provides the rationale behind a novel hypothesis that can now be tested experimentally. For example, we are now testing the validity of the hypothesis suggested by this modeling effort by investigating differences in the ability of stable versus unstable forms of transgenic IGRP, expressed in an IGRP-deficient NOD background, to cross-prime cognate autoreactive CD8+ T cells in vivo. In the absence of this modeling effort, it would be difficult, if not impossible, to appreciate that autoantigenicity and pathogenicity are not equivalent. In fact, it would be safe to say that conventional wisdom would tend to support the opposite view.

As demonstrated in the context of our mathematical model, the proportion, \( f \), of protein allocated to rapid processing into p-MHC versus protected normal function can affect the likelihood that such protein becomes an autoantigen and, most importantly, the likelihood that targeted recognition of epitopes from this protein will have pathogenic significance. Our model reveals that a subtle combination of aspects of protein allocation and processing into p-MHC affect autoantigenicity and pathogenicity.

Our results can be summarized as follows: first, dual allocation to both rapidly degraded pool and stable pool can, under some circumstances, make a protein more immunogenic by contributing both to direct p-MHC presentation on the target cell and to cross-presentation on DCs promoting CTL expansion. This scenario holds in cases where, in the target cell of interest, the number of p-MHC complexes formed per protein molecule degraded is much lower for stable protein than for rapidly degraded protein (i.e., low value of the parameter \( \eta \) which is the ratio of the p-MHC production efficiencies from the two pools). Second, if the p-MHC production efficiency of the stable form of the protein is not too low relative to the rapidly degraded form (specifically, for \( \eta > 0.25 \)), having very low allocation to rapid degradation (making \( f \) closer to 0) makes the protein more immunogenic. Third, proteins that have a very high allocation to rapid degradation (making \( f \) very close to 0, regardless of the efficiency ratio \( \eta \)) are hardly ever immunogenic. The reason is that these do not contribute strongly enough to cross-presentation. Fourth, factors such as low rates of turnover (of proteins, p-MHC and CTLs), high rates of production and/or synthesis are pro-inflammatory and exacerbate the situation, as do such aspects as high values of T cell activation rate, T cell avidity and CTL-mediated target cell killing rate. Fifth, while each of the above rate parameters is in itself important, doubling some while halving others is predicted to increase the rate. Fifth, while each of the above rate parameters is in itself important, doubling some while halving others is predicted to increase the rate. In particular, we have assumed that the level of target cells and DCs are constant, considered the growth of a single mono-specific population of autoreactive CTLs [ignoring competition, (44)] and abbreviated many steps in the process, including effects of cytokines, other cells (such as CD4+ and regulatory T cells), class-II MHC, etc. Initiation of autoimmunity is much more complicated than portrayed here, but keeping the model simple leads to important insight about autoantigen that more realistic and complex models tend to obscure. An attractive feature of simple models is that they can be understood fully using mathematical tools, as shown by the analysis associated with inequality (3) that demonstrated how antigenic stability could influence immunodominance. This comes at the expense of lack of detail. However, within the basic framework of a model like ours and based on insights it provides, rests an ability to address extensions, investigate specific cases for which data are available and make quantitative predictions in the future.

The treatment here is general, but it has been motivated by the immunodominance of IGRP as an autoantigen in T1D. Because IGRP is a transmembrane endoplasmic reticulum protein that is also constitutively expressed in the nuclear membrane, it is likely to survive apoptosis intact. In fact, IGRP segregates into the apoptotic body membrane that envelops fragmented nuclei (A. Shameli and P. Santamaria, unpublished observations). Furthermore, by virtue of its high hydrophobic nature, IGRP is also likely to contribute significantly to a DRiP pool arising from misfolded polypeptide chains. This points to the scenario described in this paper, suggesting that the allocation of IGRP within the cell, along with its own biochemical properties, might explain its ability to trigger pathogenic autoreactive CD8+ T cell responses (15, 16, 18–21, 24). Such paradigm may also apply to other autoreactive T cell specificities that recognize diabetogenic and expressed autoantigens (e.g., dystrophia myotonica kinase).

The model makes a number of predictions that are experimentally testable. For example, manipulating the p-MHC processing pathways so as to increase the [p-MHC produced per stable protein molecule degraded] relative to the other protein pool (i.e., increasing \( \eta \) faster than \( \eta \), which increases their ratio, \( \eta \)) should increase the incidence of autoimmunity sharply for proteins that are mostly found in stable forms in the cell while having little to no effect on proteins that have moderate to low allocation to stable forms (\( f \gtrsim 0.5 \)). (This can be seen from the fact that contours for the level of danger in Fig. 3b are crowded together at the high-\( \eta \) low-\( f \) region but very far apart at the opposite corner of the diagram.) Manipulating the protein allocation in the cell (by ubiquitination or other means) so as to increase the proportion of rapidly degraded protein (increasing \( f \) from low to moderate values) should lower the incidence of autoimmunity.
in most cases, except for proteins with very low values of \( \eta \), where it has the opposite effect. Further increasing \( f \) from moderate to high values (to \( f > 0.5 \)) should restore the anti-immunogenic effect. Such future experiments, together with successive modification and refinement of the model, should allow for a better understanding of the aspects of proteins that combine to create dangerous self-antigens.

### Supplementary data

Supplementary data are available at International Immunology Online.

### Funding

Mathematics of Information Technology and Complex Systems Canada to A.K., L.E.K.; the Juvenile Diabetes Research Foundation to A.K., L.E.K.; Institute for Theoretical Biology, Humboldt University, Berlin to A.K.; Natural Sciences and Engineering Research Council of Canada to P.S., L.E.K.; the Canadian Institutes of Health Research to P.S., B.S.; the National Institutes of Health (R01 GM086882) to L.E.K.; the Juvenile Diabetes Research Foundation (JDRF) to A.K., L.E.K.; the Institute for Theoretical Systems Canada to A.K., L.E.K.; the Canadian Institutes of Health Research to P.S., B.S.; the National Institutes of Health (R01 GM086882) to L.E.K.

### Acknowledgements

We wish to thank Daniel Coombs (University of British Columbia) for discussions at an early phase of this research and J.W. Yewdell (National Institutes of Health) for comments that led to improvement of our paper. The authors gratefully acknowledge funding by the Mathematics of Information Technology and Complex Systems, Canada, and by the Juvenile Diabetes Research Foundation (JDRF). P.S. is a Scientist of the Alberta Heritage Foundation for Medical Research and a JDRF Scholar. L.E.K. has been a 2009–2010 Distinguished Scholar in Residence at the Peter Wall Institute for Advanced Studies (University of British Columbia).

### References