A Mathematical Model of the Pathogenesis, Prevention, and Reversal of Type 2 Diabetes

Joon Ha, Leslie S. Satin, and Arthur S. Sherman

Laboratory of Biological Modeling (J.H., A.S.S.), National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892; and Department of Pharmacology and Brehm Diabetes Research Center (L.S.S.), University of Michigan, Ann Arbor, Michigan 48105

Type 2 diabetes (T2D) is generally thought to result from the combination of 2 metabolic defects, insulin resistance, which increases the level of insulin required to maintain glucose within the normal range, and failure of insulin-secreting pancreatic β-cells to compensate for the increased demand. We build on a mathematical model pioneered by Topp and colleagues to elucidate how compensation succeeds or fails. Their model added a layer of slow negative feedback to the classic insulin-glucose loop in the form of a slow, glucose-dependent birth and death law governing β-cell mass. We add to that model regulation of 2 aspects of β-cell function on intermediate time scales. The model quantifies the relative contributions of insulin action and insulin secretion defects to T2D and explains why prevention is easier than cure. The latter is a consequence of a threshold separating the normoglycemic and diabetic states (bistability), which also underlies the success of bariatric surgery and acute caloric restriction in rapidly reversing T2D. The threshold concept gives new insight into “Starling’s Law of the Pancreas,” whereby insulin secretion is higher for prediabetics and early diabetics than for normal individuals. \textit{(Endocrinology 157: 624–635, 2016)}

Plasma glucose is normally maintained within narrow ranges by a classic negative feedback system. Glucose rises after meals, stimulating secretion of insulin by pancreatic β-cells, which brings glucose back to baseline over several hours by promoting its storage as glycogen and its use as a fuel. However, when the system is chronically challenged by excess nutrients, resulting in obesity, homeostasis can be impaired by insulin resistance, requiring more insulin to achieve the same results. Most insulin-resistant individuals are able to maintain normoglycemia by increasing insulin secretion. Those who are unable to raise insulin levels adequately progress to type 2 diabetes (T2D).

The above scenario, that T2D results from the failure of insulin secretion to compensate for insulin resistance, is widely accepted and may be termed “the standard model.” However, it has come under criticism for 2 main reasons. First, cross-sectional studies of progression to diabetes have shown that during prediabetes and early diabetes plasma insulin concentrations are higher than normal, but glucose is normal or near normal. The onset of diabetes is associated with falling insulin rather than low insulin. These observations have been codified as “Starling’s Law of the Pancreas,” which postulates that insulin is an inverted U-shaped function of glucose: as glucose rises over years, insulin first rises to supernormal levels, then falls, marking the onset of T2D. Insulin can remain elevated for years before finally falling below normal (1). Because glucose does not increase much during the prediabetic phase when insulin levels rise dramatically, it is natural to suppose that high insulin causes high glucose rather than the reverse. It has been suggested that hypersecretion is the primary defect, causing insulin resistance or β-cell failure (2–4).

Second, even if one accepts that insulin resistance is the cause of hyperglycemia: it is not clear how β-cells can
compensate for insulin resistance in the absence of a signal from increased glucose (5, 6).

We address these questions with a mathematical model and show that the standard model is consistent with the observation that hyperinsulinemia precedes hyperglycemia. The model accounts in detail for the onset of diabetes in a classical animal model, the Zucker diabetic fatty (ZDF) rat, which follows the Starling Law. From this case study we extract general principles that can be applied, with quantitative modifications, to humans.

Our model is an update of the pioneering model of Topp et al (7). They proposed that when the normal daily insulin-glucose feedback system is unable to maintain normal glucose levels, such as when it is hindered by insulin resistance, a second slower level of negative feedback comes into play to increase β-cell mass. If mass does not expand adequately or rapidly enough, however, glucose may rise to a tipping point at which glucotoxicity causes mass to contract rather than expand. Negative feedback then gives way to positive feedback, accelerating the progression of diabetes.

We examine whether compensation involves enhanced β-cell function (secretion per cell) as well as increased β-cell mass (number of cells). This is problematic because most studies only measure insulin, which depends on both mass and function and is not easily dissociated into its components. We take advantage of a coherent longitudinal dataset from ZDF rats in which independent measurements of plasma insulin and mass were made and used to infer function (8).

We also address the signal for compensation, showing that if insulin resistance develops slowly, glucose need not rise detectably to provide a signal. Further, we incorporate new data (9), suggesting that, rather than glucose per se, the signal for β-cell replication is β-cell workload (secretion rate per cell). We investigate whether homeostatic regulation of workload leads to homeostatic regulation of β-cell mass, function, and glucose.

All of the above features of the model are explained concisely by a threshold separating normoglycemic and hyperglycemic states. This indicates that the model framework is not sensitive to the quantitative assumptions. We demonstrate the power of this unifying framework by explaining why prevention of diabetes is much easier than cure and why bariatric surgery is more effective than ordinary diets.

Materials and Methods

The model equations were solved using the free software package XPPAUT (10). XPPAUT was also used to create bifurcation diagrams, summaries of how the system behavior depends quantitatively and qualitatively on parameters. Equations and parameter values are described in the Supplemental Materials along with computer code for simulating the models.

Model parameters were determined by fitting output by eye to data from female ZDF rats (8). To simulate humans, adjustments were made to account for their greater β-cell mass and slower progression of insulin resistance and diabetes. Default parameters are listed in Supplemental Tables 1–4, and modifications for particular figures are given in the legends.

Results

Verbal and pictorial summary of model

Here, we describe the model qualitatively; the equations are contained in the Supplemental Materials.

The model of Topp et al (7), our starting point, and its successors (11, 12), have 2 components. The first component governs the rapid meal-to-meal response of insulin to glucose and consequent glucose disappearance (Eqs. 1 and 2). When glucose rises, insulin secretion responds through a sigmoidal dose-response curve (DRC) (Figure 1A). More subtle aspects of insulin secretion that depend on the rate

![Figure 1](image-url)
and duration of the glucose rise were neglected for simplicity, and we follow suit here.

The second component describes the slow compensation of the insulin-glucose system when insulin sensitivity (S_I) or some other parameter is changed. The response consists of an increase in β-cell mass by replication, driven by persistently elevated glucose, which is a signal that the fast response system is not adequate (Figure 1D; Eq. 4).

The data in Ref. 8 show that mass expansion, even in rodents, is insufficient to account for the observed rise in insulin concentrations. We therefore extend the model of Topp et al (7) by making β-cell function dynamic in 2 distinct ways. First, we assume that persistent high glucose shifts the glucose DRC to the left, analogous to reducing the half-maximal concentration (K_m) of an enzymatic reaction (Figure 1B). One way to achieve such a left shift, corresponding to an increase in the variable γ (Eq. 5), is a reduction in the number of ATP-sensitive K^- (KATP) channels in the plasma membrane of the β-cell. Although calcium is not included explicitly in the model, this would result in increased cytosolic calcium, and hence increased insulin secretion, at a given level of glucose. β-Cells can adjust K(ATP) trafficking and calcium levels rapidly, as demonstrated in the experimental companion paper (47), which shows that isolated pancreatic islets in vitro exhibit such changes after overnight culture in high glucose.

Increased glucose sensitivity in response to prolonged high glucose could also result from up-regulation of glucokinase (13) as well as altered expression or activity of many other factors. We neglect these in this debut of the model to keep it simple and focus on key principles rather than aim for completeness.

The second way we increase function in the model is to increase the maximum rate of insulin secretion, analogous to increasing the maximum rate (V_max) of an enzymatic reaction (Figure 1C) and corresponding to the variable σ (Eq. 6). Although increasing γ mainly increases insulin secretion at low glucose, σ mainly increases secretion when glucose is high enough to close most of the KATP channels. It thus resembles the potentiation of glucose-induced secretion by metabolic amplification (14) or elevation of cAMP (15) by the incretins glucagon-like peptide-1 (GLP-1) and glucose-dependent insulinotropic polypeptide (16). These mechanisms all increase the effectiveness of calcium to stimulate exocytosis rather than increasing calcium concentration. Our fits to the ZDF rat data suggest that σ is slower than γ but faster than β. The 4 response mechanisms in Figure 1 are thus ordered by their timescales.

In the model of Topp et al (7), increased β-cell replication is driven directly by glucose. Normally, glucose increases insulin secretion indirectly by increasing β-cell metabolic rate (M), which closes KATP channels and raises calcium. Porat et al (9), however, dissociated these 2 effects, finding that increasing M with a glucokinase activator (GKA) did not increase proliferation if KATP channel closure was prevented by either the channel opener diazoxide or channel mutations. This indicates that a calcium rise and/or insulin secretion is required to stimulate proliferation. In fact, although GKA could amplify the effect of KATP closure, GKA alone, in the absence of a calcium rise or secretion, decreased proliferation. Apoptosis was not measured in that study, but Kassem et al (17) found that a GKA-activating mutation increased both proliferation and apoptosis. The data are complex, but most of the results can be accounted for by 2 simple assumptions. First, proliferation is driven mainly by increased insulin secretion rate (ISR) (Eq. 3), as cells seek to avoid excessive workload. Second, apoptosis is driven mainly by increased β-cell M (Figure 2; Eq. 7). We suggest that normal postprandial increases in insulin raise both ISR and M in parallel, leaving mass unchanged. However, persistent increases in glucose from, for example, reduced S_I, would either tilt the balance in favor of ISR and promote proliferation to compensate, or favor increased M, leading to decompensation (see Supplemental Figures 1–3).

Compensation and failure

Figure 3 simulates the compensation to a step increase in insulin resistance. We use parameters appropriate for rat, and the events accordingly unfold rapidly. The shift in the glucose DRC, γ, is treated as a dynamic variable, but for now, the maximal rate of secretion, σ, is kept fixed to isolate the effect of γ. At time 0, S_I, is stepped down 50%, resulting in a 20% increase in glucose (G) (Figure 3A, green curve). In response, γ increases within a day (Figure 3, lower left), and this left shift in the DRC restores G almost to normal before the end of day 6 and completely normalizes M, as shown in the lower right panel.
Figure 3. Successful β-cell compensation prevents diabetes when $S_I$ decreases from 0.8 to 0.4 mL/μU d. A, Glucose rises rapidly when $S_I$ is decreased but recovers over a week as the glucose DRC shifts left ($γ$ increases). Black curve, slower recovery of $G$ when $γ$ is fixed at the basal level (0). B, $γ$ increases, corresponding to a left shift of the glucose DRC, then recovers as $β$ increases. C, $β$ rises slowly to a new steady state. D, Trajectory in β-cell M-insulin secretion rate or workload (M–ISR) plane. Double arrows, fast motion along glucose DRC raises M and ISR in parallel; single thick arrow, intermediate speed, corresponding to rise of $γ$, lowers M; thin arrow, slow, corresponding to rise of $β$, lowers ISR. $A_{\text{max}} = 5.5$.

Figure 4. Failure of β-cell compensation results in diabetes when $S_I$ is decreased from 0.8 to 0.1 mL/μU d. A, Glucose rises rapidly when $S_I$ is decreased and, after partial recovery, rises slowly. B, Glucose DRC shifts left ($γ$ increases) and cannot recover because glucose remains elevated. C, β-Cell mass $β$ slowly decays. D, Trajectory in M–ISR plane. In contrast to Figure 2, workload (ISR) never recovers. Double arrows, fast motion along glucose DRC raises M and ISR in parallel; single thick arrow, intermediate speed, corresponding to rise of $γ$; thin arrow, slow, corresponding to fall of $β$. $A_{\text{max}} = 5.5$.

Changes in glucose, increasing the sensitivity of the response to hyperglycemia by an order of magnitude.

If $γ$ were not dynamic but constant, as in Ref. 7, $G$ would still decrease (Figure 3A, black curve) because mass would still increase, but the transient change in $G$ would be much larger, and it would take weeks, not days, to return $G$ to baseline (data not shown).

In contrast, Figure 4 shows failure of compensation for insulin resistance when the reduction in $S_I$ is larger. As in Figure 3, the β-cells respond by traveling up the DRC and then, when the high glucose persists, shift the DRC to the left (increased $γ$). However, because the insulin resistance is more severe, this fails to restore $G$ to a normal level, so $γ$ stays high. Although ISR is increased, as in Figure 3, M also increases, and the net effect is a decline in mass. Thus, the larger drop in $S_I$ results in positive feedback rather than negative feedback, and $G$ and ISR climb irreversibly as mass shrinks catastrophically. The trajectory in the M–ISR plane (Figure 4D) moves up and to the right and never returns to the initial healthy state.

The Endocrine Society. Downloaded from press.endocrine.org by [individual/UserDisplayName] on 02 March 2016 at 09:34 For personal use only. No other uses without permission. All rights reserved.
the susceptibility to apoptosis, represented by the parameter $A_{\text{max}}$ in Eq. 10, and how far $S_t$ must fall to cause diabetes. The upward-sloping dark line indicates that the threshold level of $S_t$ for developing diabetes increases with the tendency for $\beta$-cell apoptosis. That is, individuals with high $S_t$ can tolerate all but the highest degree of apoptosis propensity. Equivalently, individuals with low apoptosis propensity, $S_t$ must fall to a low level to trigger diabetes. Alternatively, increasing the capacity to proliferate ($P_{\text{max}}$; Eq. 9) would be protective and make the threshold slant down to the right rather than up to the left.

**Case study: the ZDF rat**

In Figures 3–5, we assumed for simplicity that $S_t$ changed instantaneously, but in reality it changes gradually, as in the data from female ZDF rats (8) we model next. For convenience, the key figures in that paper are assembled in Supplemental Figure 5. The rats did not become diabetic unless fed a high-fat diet (HFD). $S_t$ estimated from the measured $G$ and $I$ using Eq. 1, dropped with age in both the low-fat diet (LFD) and HFD groups but more deeply and rapidly during HFD. As expected from the model, $\beta$-cell mass increased, and in fact, initially increased more in HFD than LFD before declining. Mass did not decline until the animals were already very hyperglycemic, and at the end of the experiment mass was about the same in the LFD and HFD rats even though the glucose levels were very different (10mM vs 30mM). Another feature that requires explanation is that the insulin levels rose earlier and more rapidly than mass (especially in HFD) and also fell earlier and more rapidly. A possible resolution was suggested in Ref. 8 by estimating $\beta$-cell function. This was done by substituting the measured $G$, $I$, and $\beta$ values and the derived $S_t$ values into Eqs. 1–3 and solving for $\sigma$.

We assume those measurements represent mainly $\sigma$ in our model, although they could include effects that we ascribe to $\gamma$. The data show that in the HFD animals $\sigma$ increased before mass in response to elevated $G$ but then fell, and suggest that this triggered the diabetes and the fall of mass.

Our model can simulate all the above features provided $\sigma$ is made the fifth and final dynamic variable in the model (Eq. 6), as shown in Figure 6. Although many characteristics of the HFD rats could have differed from those of the LFD rats, we hypothesized for simplicity that the different outcomes could be explained solely by different rates of fall of $S_t$ (Figure 6A). The LFD rats (green curves) look much like the example of successful compensation in Figure 2, except that the changes in Figure 6 are slower. As $G$ rises, $\gamma$ increases (left shift of the DRC) (Figure 6B) and then recovers as $\beta$ increases (Figure 6F). $\sigma$ increases only modestly (Figure 6C). The net effect is an increase in insulin (Figure 6E) that is adequate to maintain normal $G$ (Figure 6D).

For the HFD rats (red curves in Figure 6), $G$ increases more than for the LFD rats because of the more rapid fall of $S_t$. Initially the increase in glucose is modest but then it becomes rapid, in agreement with the data (Supplemental Figure 5). $\gamma$ increases more than for the LFD rats and, more important, never recovers, because the increases in $\sigma$ and $\beta$, although large, are insufficient to bring $G$ back down. The rise in insulin is much greater than for LFD but is nonetheless insufficient to prevent diabetes. In agreement with the experiments, the increase in $I$ is also larger and faster than the increase in $\beta$ and peaks earlier. The increase in $I$ results mainly from increased $\sigma$ and $\gamma$, with $\beta$ contributing later and less because it is slower.

Note that in the simulation $G$ keeps rising even after $S_t$ stabilizes, as in the experimental data (8). This rise is due to progressive declines in $\sigma$ and eventually mass, that are independent of $S_t$ after the initial sharp drop carries the system over the threshold.

**Preventing and reversing diabetes in humans**

Next, we consider the effects of interventions to alter the course of diabetes in humans. The model has been “humanized” by increasing volume of distribution 500-fold compared with rats, which increases islet mass in the same proportion. In addition mass dynamics were made 200-fold slower to fit data from human cadavers (18), and
function ($\sigma$) was made 23-fold slower. (The results are not sensitive to these choices for the time scales, as long as they are at least an order of magnitude slower than in rodents and mass is much slower than function.)

Figure 7 shows the effects of increasing $S_I$ at 3 different time points: prediabetes (black curves), early diabetes (green curves), and late diabetes (red curves). If $S_I$ is raised early and rapidly enough after the onset of diabetes (Figure 7A), normal glucose is restored (Figure 7D, green curve). The system crosses back below threshold, as indicated by the fact that $G$ starts to fall and $\sigma$ starts to rise. The glucose DRC remains left-shifted ($\gamma$ remains positive) (Figure 7B), but it eventually normalizes if the simulation is continued (not shown). In contrast, if the intervention is too late (eg, 46 mo) (Figure 7A, red curve), $G$ continues to rise after a brief decline (Figure 7D, red curve). Note that $I$ decreases both when the intervention successfully reverses diabetes (Figure 7E, green curve) and when it fails (Figure 7E, red curve), but for 2 distinct reasons. In the successful intervention, $I$ decreases because $G$ decreases and high $I$ is no longer needed. In the unsuccessful interventions, $I$ continues to decrease even after $G$ starts to rise again because function and mass fail ($\sigma$ and $\beta$ decrease) (Figure 7, C and F). The most interesting case, however, is the black curves: if $S_I$ is stabilized early enough during prediabetes, progression is halted. As long as the system remains below threshold, it is not necessary to restore $S_I$ to normal, just to prevent it from decreasing further. Once threshold is crossed, recovery requires an increase in $S_I$ to return the system below threshold.

These simulations offer an explanation for the findings that a program of weight loss and exercise, which by design was moderate and gradual to make it sustainable, was effective at preventing (19) but not reversing diabetes (20). They also help explain why it is difficult to reverse diabetes of longer duration (21). The study in Ref. 19 also considered metformin as a prevention treatment. Metformin likely acts mainly on hepatic glucose production rather than peripheral insulin resistance, but the results in the model are similar. If applied early enough it can be effective for prevention, but it cannot reverse established T2D.

Bariatric surgery, in contrast, can produce lasting reversal of diabetes on a mass scale (22), albeit not in all subjects and with a tendency towards relapse (23). Acute caloric restriction has been reported to have similar efficacy (24) and to rapidly normalize hepatic glucose production (25). In both interventions, the improvement occurs almost immediately, long before weight loss.

The mechanism underlying the metabolic benefits of bariatric surgery is currently unknown and hotly debated,

![Figure 6. Simulation of female ZDF rats on LFD (green) and HFD (red). The only difference between the 2 cases is the rate at which $S_I$ falls (Eq. 17). In the LF case, $\beta$-cell mass grows rapidly enough to compensate for the mild rate of drop in $S_I$ ($S_0 = 0.23 \text{ mL/}[\mu\text{U d}], \tau_I = 16.2 \text{ d})$. In the HF case, the drop in $S_I$ is too deep and rapid ($S_0 = 0.022 \text{ mL/}[\mu\text{U d}], \tau_I = 5.1 \text{ d})$ to be compensated by the combined effects of increased mass and function, resulting in diabetes. Compare with data in Supplemental Figure 5, taken from Ref. 8 (Figures 2–4). Note that their $G$ values, which are nonfasted, are about twice ours, which are daily average values and dominated by the fasting levels. Their parameter for $\beta$-cell function ($\beta_{sc}$) corresponds to a combination of our $\sigma$ and $\gamma$. $A_{\text{max}} = 3.1$.](doi: 10.1210/en.2015-1564)
but guided by data (26, 27) showing that hepatic glucose production is rapidly reduced (within the first week) after Roux-en-Y gastric bypass (RYGB), we assume a rapid reduction of the term $R_0$ in Eq. 1. Another study (28) showed a rapid normalization of fasting plasma glucose and insulin after RYGB, further evidence of reduced hepatic glucose production.

That study also showed a trend toward partial improvement in $\beta$-cell function in the first week after surgery. The improved function may be at least partly due to GLP-1, which was also increased during the first week (28). Putting these pieces together, we simulate in Figure 8 the effects of RYGB as a lowering of $R_0$ (Eq. 1) and an increase in $\sigma$. We did not alter $S_I$, which also showed a trend toward partial improvement (28), but that would only improve the result.

In contrast to Figure 7, where slow improvement in $\sigma$ was a consequence of lowered glucose, in Figure 8, we assume that the surgery directly and rapidly improved $\beta$-cell function (in Figure 8, step increase at arrow in C). Consequently, even late intervention (>30 mo) is able to restore normoglycemia. As in the successful reversal of T2D in Figure 7, $I$ falls because of the drop in $G$. In contrast to the hypothesis that diabetes is reversed because $I$ is reduced (3), the model suggests that $I$ is reduced because the diabetes is reversed.

The simulated intervention can fail if it is made too late or $G$ is higher (data not shown), consistent with reports that earlier treatment and less severe hyperglycemia are associated with higher likelihood of success of bariatric surgery (23).

Central role of bistability

We have shown that the model accounts for successful compensation for insulin resistance and failure that results in diabetes, and delineates the conditions for interventions to reverse established diabetes. The model shows further that these features result from the combination of negative feedback (homeostasis) for small perturbations from normal and positive feedback (failure) for large perturbations. Such characteristics are ubiquitous in dynamical systems and lead naturally to a threshold separating 2 coexisting steady states (bistability), here health and disease (29). This is illustrated in Figure 9.

The threshold, indicated by the dashed blue line, corresponds to 125-mg/dL $G$, the diagnostic level of diabetes. Below threshold, $I$ increases and $G$ decreases toward normal (downward, straight blue arrow) because of increased mass and/or function. Above threshold, $I$ decreases and $G$
increases toward diabetes (upward, straight blue arrow) because of decreased mass and/or function. The system also defends against hypoglycemia. If $G$ is below normal, mass and/or function decrease to drive $G$ back up.

For an individual with normal $G$ who loses SI due to weight gain, $G$ would rise unless opposed by increased insulin. If the decrease in $S_I$ is slow enough for mass and function to compensate, $G$ recovers (blue, upside-down U-shaped arrow in Figure 9). However, if the decrease in $S_I$ is too rapid, or mass and function respond sluggishly, $G$ rises in spite of increased $I$ (upward red arrow). Once the threshold is crossed, compensation is overwhelmed by glucotoxicity and secretion decreases, raising $G$ further. The threshold represents the point at which the net effect of mass and function changes on insulin reverses, corresponding to the peak of the Starling curve. Once above this point, typical diet and exercise interventions are ineffective because they cannot overcome the cellular processes driving down mass and function. Analogously, below threshold, increasing $G$ is like trying to go up a down escalator, whereas above threshold, increases in $G$ become easy, because the escalator switches to the up direction. Acute caloric restriction, whether by reduced intake or bariatric surgery (light blue, downward arrow in Figure 9), can overcome this challenge. A more precise and mathematical analysis is given in Supplemental Figure 4.

**Discussion**

**Core model features and predictions**

Building upon a previous model of $\beta$-cell mass regulation of (7), we have developed a mathematical model for the pathogenesis, prevention and reversal of T2D. The model is based on the “standard model” whereby T2D develops in response to insulin resistance unless prevented by increased $\beta$-cell mass and function (objections to this view are discussed briefly below). We revised the mass regulation component in Ref. 7 by making $\beta$-cell proliferation responsive to secretory workload (secretion per unit mass) rather than directly to glucose, in line with (9). The toxic effect of glucose on $\beta$-cell mass has been incorporated as an apoptosis rate that increases with $\beta$-cell M. Under normal conditions, postprandial increases in glucose are assumed to increase both proliferation and apoptosis, increasing $\beta$-cell turnover but maintaining mass constant (Figure 2). When those signals become persistently unbalanced, as in insulin resistance, mass either increases to restore normal workload (negative feedback) or fails to respond adequately, leading to T2D (positive feed-
We have shown that the homeostatic regulation of workload based on these principles naturally leads to regulation of both glucose and \( G \) and \( \beta \)-cell mass. Moreover, workload is a more sensitive signal than glucose because the dynamic range of ISR (Figure 3) is larger. This is a natural feature of the \( G-I \) system, which allows \( I \) to vary much more than \( G \).

In addition, we have incorporated 2 forms of functional compensation, both more rapid than mass. The faster form of the 2 is a left shift in the glucose DRC (increased \( G \), analogous to decreased \( K_m \) in enzyme kinetics), which could be achieved by reducing \( \beta \)-cell KATP channel trafficking to the plasma membrane, thereby up-regulating calcium-dependent triggering of insulin secretion (14). The second, slower response is an increased maximum rate of insulin secretion (\( \sigma \), analogous to \( V_{max} \) in enzyme kinetics), corresponding to the metabolic (14) and hormonal (15) amplification pathways of insulin secretion. The need for function to change dynamically was suggested by a longitudinal study of diabetes progression in ZDF rats (8), which showed that the measured increase of \( \beta \)-cell mass was insufficient to account for increased plasma insulin.

The extended model thus proposes a 4-level hierarchy of responses (Figure 1). Ordinary postprandial rises in glucose are met by traversing the glucose DRC, which increases workload and \( M \). Persistent high postprandial glucose over a period of days shifts the DRC to the left (increased \( G \)), reducing metabolic stress but increasing cytosolic calcium and leaving workload elevated. Chronic high workload over weeks in rodents (months in humans) enhances the amplifying pathways (increased \( \sigma \)). This relieves the stress of high calcium by making calcium more effective in stimulating exocytosis, but leaves workload elevated. Ultimately, mass increases, reducing workload and relieving the stress of chronically high insulin secretion, including endoplasmic reticulum (ER) stress. If the increase in mass is adequate, normal glucose and workload are restored, at the expense of hyperinsulinemia.

The predicted left shift in the DRC has been tested and confirmed experimentally in the companion paper (47). We further predict that the left shift should only be transient when insulin resistance is well compensated, as illustrated in our simulation of ZDF female rats fed a LFD. A persistent left shift, as in our simulation of ZDF rats on HFD, is a marker of failure to increase function (\( \sigma \)) or mass sufficiently to overcome insulin resistance. This is seen in \( db/db \) mice (30) and is predicted for other mouse models of diabetes and humans with T2D.

Thus, although decline of \( \sigma \) is a marker of \( \beta \)-cell failure (Figure 6C), decline of \( G \) likely indicates a successful hand-off of responsibility for compensation to \( I \). This difference stems from our assumption that \( \sigma \) is subject to glucotoxicity (Supplemental Figure 2), but \( G \) is not. This remains to be tested experimentally.

Preexisting, genetic defects in \( G \) do play in important role in T2D pathogenesis. Gain of function mutations in KATP channels are prominent risk factors in genome wide association studies. Model simulations (not shown) confirm that mild gain of function mutations lead to T2D in the presence of moderate insulin resistance.

The 2 function variables in the model, \( G \) and \( \sigma \), correspond approximately to the 2 major classes of insulin secretagogues, KATP channel blockers and incretin mimetics, respectively. Supplemental Figure 6 illustrates how they affect the glucose DRC. Supplemental Figure 6A
shows that increasing glucose sensitivity shifts the curve to the left and has its biggest relative effect at low glucose (ie, fasting glucose), whereas increasing maximum capacity has its biggest effect at high glucose (ie, postprandial glucose).

Supplemental Figure 6B shows another important difference between the 2 therapeutic targets. If a KATP blocker is applied after $\sigma$ has already substantially declined, or if it is applied early but does not control $G$ sufficiently to prevent $\sigma$ from declining, then it becomes ineffective. This veto power of $\sigma$ over $\gamma$ is one possible cause of sulfonylurea failure. Another possibility is that $\beta$-cells do not lose secretory capacity but simply stop responding to the drug, either because calcium no longer rises sufficiently or because the exocytotic response to calcium is down-regulated. A study with implanted glibenclamide pellets suggests this type of failure, because secretion recovered after 24-hour drug washout (31).

**Objections to the standard model**

The assumption that increased $\beta$-cell workload drives compensation for insulin resistance by both mass and function helps answer a key objection to the standard model, namely, that there is no obvious rise in $G$ to stimulate compensation. By amplifying the glucose signal, workload allows the response to be sensitive to small changes in glucose.

A further response to this objection is provided by the simulation of the progression to T2D in ZDF rats (Figure 6). When the decline in $S_I$ is slow, as in the natural history of T2D, the response of mass and function can buffer it successfully, continuously correcting for miniscule deviations in glucose.

It may nonetheless be the case that hyperinsulinemia begets further insulin resistance, as proposed in Refs. 2–4. Work in progress suggests that a secondary, frictional effect of hyperinsulinemia to induce insulin resistance (32) can be added to the model without changing the result that T2D arises from failure of compensation. For another model and point of view see Ref. 33.

**The centrality of threshold**

The response to insulin resistance is ultimately all-or-none, it is either compensated or leads to diabetes. The system is very good at buffering against small changes in $S_I$, but when it fails, it fails catastrophically, producing a rapid and inexorable rise in glucose. According to the Starling Law, diabetes commences when insulin starts to decrease, not when the level of insulin falls below normal. This is evident in the ZDF rat data (Supplemental Figure 5) and in our ZDF simulations (Figure 6). Moreover, the period of most rapid rise of $G$ coincides with the peak of $I$, indicative of a threshold crossing. Over time, the clinical definition of T2D has been adjusted empirically based on the potential for future complications (34), but, interestingly, it now essentially coincides with the peak of the Starling curve and thus to crossing the threshold. Further evidence for a threshold is provided by longitudinal data showing that glucose rises very slowly for years before suddenly rising rapidly in both humans (35) and monkeys (36).

The model shows further that the degree of insulin resistance that results in crossing the threshold depends on the ability of the $\beta$-cells to respond (Figure 5). The less susceptible the $\beta$-cells are to apoptosis, the greater the loss of $S_I$ that can be tolerated. Figure 5 thus includes 2 distinct pathways to T2D, one primarily involving insulin resistance (Figure 5, lower left part of diabetes region) and the other primarily involving defective insulin secretion (Figure 5, upper right part of diabetes region), but suggests these are extreme cases of a continuum. If one defect is greater, less is required of the other to cause T2D. Figure 5 also implies that mild $\beta$-cell insufficiency can be a silent preexisting genetic trait that is only unmasked when insulin resistance arises.

The threshold explains why prevention of diabetes is much easier than reversal. Before an individual crosses threshold, the $G$–$I$ system exerts negative feedback and tends to restore normal glucose. At this stage it is sufficient to halt the decline of $S_I$ by diet and exercise, which allows the system to heal itself over time (Figure 7). After crossing threshold, however, the dynamics switch to positive feedback, which reduces $\beta$ and $\sigma$ and drives glucose even higher. More drastic interventions are then required to return below threshold. The only ways currently known to achieve this are bariatric surgery and acute caloric restriction.

The model predicts that the threshold arises from bistability between normoglycemic and diabetic states (Figure 9 and Supplemental Figure 4). Direct evidence for bistability is provided by the bimodal outcome of partial pancreatectomy in rats (37, 38). We suggest that small differences in either the fraction of pancreas removed or innate ability to expand $\beta$-cell mass were amplified into well-separated normal and diabetic outcomes.

The rapid reversal of diabetes after bariatric surgery and acute caloric restriction also strongly suggests coexisting healthy and diseased states. Although we had to make assumptions about mechanism to simulate the metabolic consequences of bariatric surgery (Figure 8), the specific assumptions are not critical. The model merely requires 1 or more parameters (hepatic glucose production, $S_I$, $\beta$-cell mass or function, insulin clearance, or insulin-independent glucose uptake) to change rapidly and...
sufficiently to drive the system back below threshold. The model also predicts that if bariatric surgery were reversed, diabetes would not return, provided the individual did not regain weight. This could be tested in rodents, whose dietary intake can be controlled.

**β-Cell mass in humans and rodents**

We initially parametrized the model to account for data from ZDF rats (8), but found that simply increasing the volume of distribution and making β-cell mass and the component of function represented by σ much slower was sufficient to simulate human diabetes (Figures 7 and 8). The model suggests that function plays a bigger role than mass in rodents and even more so in humans, because it responds more rapidly.

It has been argued that β-cell proliferation is virtually nonexistent in adult humans because it is difficult to find markers of proliferation in islets from human cadavers (39, 40). However, mass was increased 30%–60% in pancreases from obese nondiabetic humans and decreased 30% in pancreases from diabetic humans (40–42). It is beyond the scope of this article to reconcile these findings, but we note that insulin resistance develops much more slowly in humans. Therefore, a much slower rate of replication is sufficient to produce the observed increases in mass, which are an order of magnitude smaller than in rodents. Also, β-cell mass may increase not by replication, but by transdifferentiation of α-cells (43, 44), or there may be dormant β-cells that do not secrete until needed (45, 46). Thus, β-cell mass may not increase but functional β-cell mass may do so. The model can work equally well with that assumption.

Further progress will require new experimental data. The dependence of proliferation and apoptosis on ISR and M in the model, although guided by experiment, is of a rudimentary, phenomenological nature because we lack quantitative data to constrain the auxiliary functions. New data are needed in particular to separate the effects of increased calcium and insulin secretion and to assess the possible roles of ER stress and reactive oxygen species. Further research on the long-term regulation of β-cell function is critical, as this is likely to play a more important role in humans than mass. Finally, the field has been hampered by the need to rely on pancreases from cadavers because there is no noninvasive, longitudinal way to measure β-cell mass or function. We hope the model will provide incentive to develop such methods and give guidance on interpreting the data they produce.

**Acknowledgments**

We thank Richard Bertram for careful reading of the manuscript and numerous poster attendees for helpful comments.

Address all correspondence and requests for reprints to: Arthur S. Sherman, Laboratory of Biological Modeling, National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases, 12A South Drive, Room 4007, Bethesda, MD 20892-5621. E-mail: arthurs@mail.nih.gov.

This work was supported by the Intramural Research Program of the National Institutes of Health (National Institute of Diabetes and Kidney Diseases) (A.S.S. and J.H.). L.S.S. was supported by the National Institute of Diabetes and Digestive and Kidney Diseases Grant RO1DK46409.

Disclosure Summary: The authors have nothing to disclose.

**References**


35. Mason CC, Hanson RL, Knowler WC. Progression to type 2 diabetes characterized by moderate then rapid glucose increases. Diabetes. 2007;56(8):2054–2061.


