An Enzyme Mechanism Language for the Mathematical Modeling of Metabolic Pathways

Chin-Rang Yang[1,3], Bruce, E. Shapiro[4]

Eric D. Mjolsness[2,3] and G. Wesley Hatfield[1,3]

[1] Department of Microbiology and Molecular Genetics, College of Medicine,
[2]School of Information and Computer Science,
[3]Institute for Genomics and Bioinformatics,
University of California-Irvine,
Irvine, California 92697-4025

[4]Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109
U.S.A.

Correspondence should be addressed to G.W.H. (gwhatfie@uci.edu)

ABSTRACT

We have developed an enzyme mechanism language for the mathematical modeling of biochemical pathways. This language, kMech, is a comprehensive collection of single and multiple substrate enzyme reactions and regulatory mechanisms that extends Cellerator function for the mathematical modeling of enzyme reactions. Each mechanism has been codified to generate a set of elementary reactions that can be translated by Cellerator into ordinary differential equations (ODEs) solvable by Mathematica. We also present methods that use common kinetic measurements to estimate physical constants required to solve these ODEs. Because kMech assembles fundamental modular association-dissociation reaction mechanisms to describe complex enzyme mechanisms, the kMech/Cellerator dynamic modeling system is more flexible, and easily extended, than commonly used simulation systems based on Michaelis-Menton kinetic equations. We use this system to model branched chain amino acid biosynthesis in *Escherichia coli.* Simulations of metabolic perturbations and genetic mutations predict results reported in the literature.

INTRODUCTION

Systems biology may be broadly defined as the integration of diverse data into useful biological models that allow scientists to easily observe complex cellular behaviors and to predict the outcomes of metabolic and genetic perturbations. Current and emerging high-throughput genomic and proteomic technologies now allow us to study the simultaneous behavior of thousands of cellular components in response to such perturbations. In order to integrate the large amounts of data these high-throughput technologies produce, we must develop computational methods to simulate complex biological systems.

As a first step towards the elucidation of the systems biology of the model organism, *Escherichia coli*, we have elected to use a dynamic modeling approach to describe fundamental molecular interactions with differential equations. We have also elected to confine our initial efforts to the complex but well-studied metabolic pathways for the biosynthesis of the branched chain amino acids (L-isoleucine, L-valine, and L-leucine) [1]. However, before we can mathematically simulate carbon flow through these pathways, we must have knowledge of the enzyme mechanism of each step in the pathway, we must know the values of kinetic and physical parameters that describe these mechanisms, and we must be aware of the metabolic regulatory circuits that control these pathways. In the case of *E. coli*, an organism that has been studied for over fifty years, much of this information is available in the literature.

In this report we present a Cellerator[2] language extension, kMech, that describes a suite of enzyme mechanisms suitable for the mathematical modeling of metabolic pathways[*]. Each mechanism has been codified to generate a set of elementary reactions that can be translated by Cellerator into ordinary differential equations (ODEs) solvable by Mathematica. Alternatively, Cellerator can generate ODEs in System Biology Markup Language (SBML) for other simulators [3]. We also present methods that use common kinetic measurements to estimate rate constants required to solve these differential equations. Because kMech assembles fundamental association-dissociation reaction mechanisms to describe complex enzyme mechanisms, the kMech/Cellerator dynamic

[*] A Mathematica™ executable, kMech.m, file and a Mathamatica™ notebook file for the simulation of branched chain amino acid biosynthesis, may be downloaded from the University of California , Irvine, Institute for Genomics and Bioinformatics website at www.igb.uci.edu/tools.htm.

modeling system is more flexible, and more easily extended, than commonly used simulation systems based on Michaelis-Menton kinetic equations [4-7]. Our model incorporates all of the forward and reverse enzyme reactions and regulatory circuits of the branched chain amino acid biosynthesis pathways including: single and multiple substrate (Ping Pong and Bi Bi) enzyme kinetic reactions; feedback inhibition (allosteric, competitive, and non-competitive) mechanisms; channeling of metabolic flow through isozymes and transamination reactions; and active transport mechanisms.

L-threonine deaminase, the first enzyme specific for the biosynthesis of L-isoleucine, is end-product inhibited by L-isoleucine, and α-isopropylmalate synthase, the first enzyme specific for the biosynthesis of L-leucine, is end-product inhibited by L-leucine (Fig. 1). However, because the parallel pathways for L-valine and L-isoleucine biosynthesis are catalyzed by a set of bi-functional enzymes, L-valine inhibition of the first enzyme specific for its biosynthesis might compromise the cell for L-isoleucine biosynthesis. This type of a regulatory problem is often solved by using multiple isozymes that are differentially regulated by multiple end-products. In this case, there are three α-acetohydroxy acid synthase (AHAS) isozymes that catalyze the first step of the L-valine pathway, which is also the second step of the L-isoleucine pathway. AHAS I has a substrate preference for the condensation of two pyruvate molecules required for L-valine, and is end-product inhibited by L-valine. AHAS II, which has a substrate preference for the condensation of pyruvate and α-ketobutyrate required for L-isoleucine biosynthesis, is not inhibited by any of the branched chain amino acids. The third isozyme, AHAS III shows no preference for pyruvate or α-ketobutyrate but is inhibited by L-valine.

L-threonine deaminase (TDA) is an allosteric enzyme that exists in an active (R) state and an inactive (T) state. The fraction of active enzyme in the R state is determined by the concentrations and relative affinities of substrate (L-threonine), inhibitor (L-isoleucine), and activator (L-valine) for the R and T states. The kinetic behavior of this enzyme is accurately captured by the Monod, Wyman, Changeux (MWC) concerted allosteric transition model [8] described in the Methods section.

In addition to these regulatory circuits, the intracellular levels of the branched chain amino acids are influenced by the reversible transamination reactions of each

pathway. When the intracellular levels of any of these end-product amino acids become high, reverse reactions to their cognate ketoacids are favored. For example, high concentrations of L-valine can be converted to α-ketoisovalarate to supplement L-leucine production. The intracellular amino acid levels can be affected by the active transport of these amino acids from the extracellular growth medium. Therefore, enzymes required for the active transport of the branched chain amino acids into the cell against a concentration gradient are included in our simulations.

RESULTS

Computational Modeling of the Dynamics of Carbon Flow through the Branched Chain Amino Acid Biosynthesis Pathways of *E. coli* K12

The three interacting metabolic pathways simulated here consist of eleven enzymes, eighteen metabolic intermediates, and three enzyme cofactors. The mathematical model for this metabolic system consists of 105 ordinary differential equations (ODEs), with 110 association and dissociation rate constants, and 52 catalytic rate constants. The enzymes of these interacting pathways employ three distinct enzyme mechanisms (simple, Bi Bi, and Ping Pong Bi Bi) that are regulated by allosteric, competitive, or noncomptetitive inhibition mechanisms. As described in the Methods section, the physical parameters for these models have been obtained directly from the literature, calculated from data in the literature, or estimated by fitting experimental data. Relative intracellular enzyme levels have been inferred from DNA microarray data [9]. The time dependent approach to steady state for the thirteen pathway intermediates and end-products are shown in Figure 2. Steady-state enzyme activity levels have been optimized to properly channel the metabolic flow of intermediates through these pathways and to match reported *in vivo* levels of pathway intermediates and end-products [10, 11]. The detailed kMech inputs, corresponding ODEs, kinetic rate constants, and initial conditions for solving the ODEs are presented in Supplementary Figure 1 online.

Allosteric Regulation of L-Threonine Deaminase

The allosteric regulatory mechanism of L-threonine deaminase was simulated with the MWC model employing physical parameters based on the literature and optimized as described in the Methods section. The data in Figure 2 show that L-threonine deaminase produces α-ketobutyrate at a steady state level comparable to that observed *in vivo* [10, 11], and the data in Figure 3 illustrate the response of this allosteric enzyme to changes in the levels of its effector ligands, L-isoleucine and L-valine. Since the K_i for L-isoleucine (15 μM) is much less than the K_a for L-valine (550 μM), an initial decrease in the fraction of active TDA as L-isoleucine accumulates is followed by an increase to a final steady level that accompanies the slower accumulation of L-valine (Figs. 2 and 3).

Regulation of the a **-Acetohydroxyacid Synthase (AHAS) Isozymes**

The two-substrate, two-product, AHAS isozymes I and III employ a Ping Pong Bi Bi enzyme mechanism described in the Methods section (the AHAS II isozyme is inactive in *E. coli* K12; [12]) The L-valine inhibition of AHAS I and III is noncompetitive and, in the case of AHAS III, is incomplete since 15-20% of the activity attained at saturating substrate concentrations (V_{max}) remains in the presence of saturating L-valine concentrations [13]. The data in Figure 2 show that the production of α-acetolactate (αAL) produced by AHAS isozymes I and III decreases as L-valine accumulates. These data also show that α-aceto-α-hydroxybutyrate (αAHB), primarily produced by AHAS isozyme III, decreases to a steady state level as its end-product inhibitor (L-valine) accumulates, and as its substrate, α-ketobutyrate (αKB), decreases because L-isoleucine accumulates and inhibits L-threonine deaminase (Fig. 1).

Responses to Metabolic and Genetic Perturbations

L-valine growth inhibition of Escherichia coli K12 is due to a-ketobutyrate accumulation, not L-isoleucine starvation.

It is well known that adding L-valine at a final concentration of 1 mM to the medium of a growing culture of *E. coli* K12 cells inhibits their growth, and that this L-valine inhibition can be reversed by L-isoleucine [14]. Since the AHAS I and AHAS III isozymes of *E. coli* K12 strains are inhibited by L-valine, and since the *ilvG* gene for

AHAS II in *E. coli* K12 strains contains a frameshift mutation that destroys AHAS II activity [12], it was assumed that L-valine inhibition of AHAS I and AHAS III might inhibit growth by interfering with L-isoleucine biosynthesis. However, later studies demonstrated that the intracellular L-isoleucine level is not suppressed by L-valine because its biosynthesis is sustained even at saturating L-valine concentrations, by AHAS III that remains 15-20% active [13, 15] and by the activation of L-threonine deaminase that shuttles more substrate into the L-isoleucine pathway. Indeed, the simulation in Figure 4A shows that in the presence of extra-cellular L-valine, the intra-cellular L-isoleucine level in fact accumulates nearly nine-fold (Fig. 4A). At the same time, the pathway precursor of L-isoleucine, α-ketobutyrate (αKB), increases about three-fold (Fig. 4B). This build-up of α-ketobutyrate (αKB) is caused by L-valine activation of L-threonine deaminase (Fig. 4C) which increases its production, and L-valine inhibition of the AHAS I and AHAS III isozymes, which reduces its consumption. It is now known that this α-ketobutyrate accumulation is toxic to cells because of its ability to inhibit the glucose PTS transport system [16, 17]. Thus, as reproduced by our simulations, L-valine growth inhibition of *Escherichia coli* K12 is not a consequence of L-isoleucine starvation.

The simulation results in Figure 4B show that the growth inhibiting effects of L-valine induced α-ketobutyrate accumulation can be reversed by L- isoleucine by its ability to inhibit L-threonine deaminase activity. This simulation shows that, in the presence of 1 mM L-valine, the level of α-ketobutyrate increases around three-fold; and that in the presence of 500 μM L-isoleucine, α-ketobutyrate levels are reduced to the control level observed in the absence of L-valine. The simulation results in Figure 4C show that, concomitant with the rise in α-ketobutyrate observed in the presence of 1 mM L-valine, nearly 18% of the cellular L-threonine deaminase is converted to the active R state. However, concomitant with the decrease in α-ketobutyrate observed in the presence of 500 μM L-isoleucine, the cellular L-threonine deaminase in the active R state is reversed to the control level observed in the absence of L-valine. It is worth note that the results of these mathematical models are verified by experimental results accumulated from multiple laboratories over a three decade period [14, 16, 17].

Metabolic Engineering L-Isoleucine Over-Production.

An obvious goal of modeling biological systems is to facilitate metabolic engineering for the commercial production of specialty chemicals such as amino acids. In the past, this has been largely accomplished by genetic manipulation and selection methods. For example, a common strategy to over-produce an amino acid has been to isolate a strain with a feedback resistant mutation in the gene for the first enzyme for the biosynthesis of that amino acid. Here we use our model to determine the effects of a feedback resistant L-threonine deaminase for the over-production of L-isoleucine. We can simulate a L-threonine deaminase resistant mutant strain by increasing the K_i for L-isoleucine to a large number, (e.g. 100,000 μM). The simulation in Figure 5A shows that in the absence of L-isoleucine inhibition, the activator and substrate ligands drive nearly 100% of cellular feedback resistant L-threonine deaminase (TDA^R) to the active R state compared to the wild type enzyme that is only 6% present in the active R state. However, in spite of this increased L-threonine deaminase activity, the data in Figure 5B show that the amount of L-isoleucine produced in a feedback resistant *E. coli* K12 strain is only about 5 to 6 fold more than that produced by a wild type strain. On the other hand, the steady state level of α-ketobutyrate is increased about forty-fold (Fig. 5C). This is because *E. coli* K12 does not have an active AHAS II isozyme that favors the condensation of pyruvate and α-ketobutyrate for L-isoleucine production; thus, α-ketobutyrate accumulates. These simulation results suggest that in order to over-produce L-isoleucine we must remove this bottle-neck. The results in Figure 5D show that restoring a wild type AHAS II isozyme and simulating an attenuator mutation that elevates the levels of all of the enzymes of the L-isoleucine and L-valine parallel pathways 11-fold [18] both avoids buildup of subsequent pathway intermediates (Fig. 5C) and results in a forty-fold increase in L-isoleucine production. These simulated results that high level overproduction of L- isoleucine in *E. coli* requires a functional AHAS II isozyme and a de-attenuated genetic background (*ilvGMEDA-att*$^-$) agree with experiments performed by Hashiguchi *et al.* at the Ajinomoto Co., Tokyo, Japan [19].

Excess L-Valine Supplements L-Leucine Synthesis.

An *Escherichia coli* K12 *ilvC* strain lacking acetohydroxyacid isomeroreductase (IR) activity cannot produce α,β-dihydroxy-isovalerate (αDHIV) and α,β-dihydroxy-β-methylvalerate (αDMV), intermediates of the common pathway for the biosynthesis of all three branched chain amino acids, L-isoleucine, L-valine, and L-leucine (Figure 1). However, acetohydroxyacid isomeroreductase deficient strains can grow in the presence of only L-isoleucine and L-valine. They do not need L-leucine because L-valine can be transaminated to α-ketoisovalerate (αKIV), a precursor of L-leucine biosynthesis, by the reverse reactions of transaminase B and C. The simulation results in Figure 6 show that in the extra-cellular presence of 500 μM L-valine and L-isoleucine enough L-leucine can be produced to support the needs of an *ilvC* strain.

DISCUSSION

Because kinetic rate constants and *in vivo* enzyme concentrations are usually unavailable, existing models often employ Michaelis-Menton pseudo steady state assumptions. These models use experimentally measured K_m ($\frac{k_r + k_{cat}}{k_f}$), and V_{max} ($E_T \cdot k_{cat}$) values to derive differential equations that describe one-way carbon flux reactions [4-7]. However, while K_m and V_{max} values are often available for most enzymes, these methods do not allow the modeling of many biologically important enzyme parameters. A major achievement of the methods reported here is the development of kMech, a Cellerator language extension that describes a collection of complex enzyme reactions and metabolic regulatory models. Because kMech describes complete enzyme reaction mechanisms it allows the incorporation of, among other parameters, both free enzyme and enzyme-substrate complex concentrations as well as forward (k_f), reverse (k_r) reaction and catalytic (k_{cat}) rate constants. This advantage is illustrated by the fact that the Michaelis-Menton kinetic models do not allow accurate modeling of channeling substrate flow among the AHAS isozymes. In this case, it was important to use kMech models with complete enzyme mechanisms that account for the fact that each enzyme is partitioned between a substrate intermediate and a catalytic form, and that each enzyme

has different end-product inhibition properties. The Michaelis-Menton kinetic models also do not allow accurate modeling of carbon flow through the reversible transaminase reactions. In this case, the same enzyme, TB, catalyzes the reversible transamination of all three branched chain amino acids. Thus, the relative affinities and concentrations of these amino acids and their cognate ketoacids for the pyridoxal (TB) and pyridoxamine (TBNH2) enzyme intermediates defined by the kMech models determine the rates of transamination of each end-product amino acid.

In spite of the fact that *E. coli* genetic and metabolic regulation mechanisms have been studied for more than fifty years, not all the physical parameters required for the metabolic simulation of the branched chain amino acid biosynthetic pathways, or any pathways, are contained in the literature. However, to use the kMech reaction models we must know the kinetic rate and catalytic constants for each enzyme mechanism. Thus, it was necessary to develop methods to obtain robust estimates of these parameters from available kinetic data. In the Methods section we describe a Lambda (Λ) approximation method to estimate kinetic constants, k_f and k_r, from experimentally measured enzyme turnover numbers (k_{cat}) and K_m values. We also describe an Omega (Ω) approximation method to estimate inhibition constants, k_{fi} and k_{ri}, from measured K_m, K_i, and estimated k_f and k_{cat} values. Intracellular enzyme levels can be inferred from the specific activities of purified enzymes and DNA microarray measurements. For the simulations reported here, each of these experimentally constrained parameters for each enzyme reaction was optimized to predict experimentally-measured, intracellular, substrate, intermediate, and end-product concentrations.

The methods that we have developed to simulate the metabolic regulation of branched chain amino acid metabolism in *E. coli* represent the first level of a bottom-up approach to an elucidation of the systems biology of this model organism. The metabolic pathways described here are further regulated by hierarchical levels of genetic regulatory mechanisms that respond to chemical and environmental signals. These hierarchical levels of control are: (i) global control by chromosome structure, (ii) global control of stimulons and regulons composed of many operons, and (iii) operon-specific controls [20]. The first, or highest, level of control is exemplified by DNA-topology-dependent mechanisms that coordinate basal level expression of all of the genes of the cell

(independent of operon-specific controls). This level is mediated by DNA architectural proteins and the actions of topoisomerases in response to nutritional and environmental growth conditions [21]. The second level of control is mediated by site-specific DNA binding proteins, which, in cooperation with operon-specific controls, regulate often overlapping groups of metabolically related operons in response to environmental or metabolic transitions or stress conditions. The third level of control is mediated by less abundant regulatory proteins that respond to operon-specific signals and bind in a site-specific manner to one or a few DNA sites to regulate single operons. All of these regulatory levels participate in the regulation of the production of cellular proteins required for growth and intermediary metabolism. Each of these levels of control impacts metabolic regulation by influencing enzyme levels. Thus, a complete model of branched chain amino acid biosynthesis in *E. coli* must include these higher levels of gene regulation. To this end we are currently working on another software package, gMech, that will describe genetic regulatory mechanisms such as attenuation, activation and repression, and DNA topological mechanisms assembled from the fundamental association-dissociation reactions as in kMech. Ultimately, these kMech and gMech models will be integrated with built-in Cellerator signal transduction cascade and cell development models. These efforts will be appropriate for mathematically modeling the metabolic, genetic, and developmental systems of any organism.

METHODS

Lambda approximation method for enzyme rate constants:

In contrast to other models that simulate metabolic flux with Michaelis-Menton based rate equations [4-7], we have employed the kMech/Cellerator package to generate ordinary differential equations (ODEs) that describe detailed interactions between substrates, products, enzymes, enzyme intermediates, and coenzymes. However, in order to solve these ODEs, we need to obtain values for forward and reverse kinetic rate constants k_f, k_r, that are often not available. To deal with this problem, we developed the Lambda (Λ) approximation method described in detail below. Briefly, this method

calculates these constants from the easily determined, and usually available, enzyme parameters, K_m (Michaelis-Menten constant) and k_{cat} (catalytic constant or enzyme turnover number). This approach results in simple mathematical relations that can be used to estimate rate constants of differential equations for enzyme reactions. For example, a single substrate-single product enzyme reaction can be represented as:

$$A + En \underset{k_r}{\overset{k_f}{\Leftrightarrow}} EnA \xrightarrow{k_{cat}} En + C \quad \text{.....(1)}$$

where A is the substrate, En is the free enzyme, EnA is the enzyme-substrate complex, C is the product, k_f is the forward (association) rate constant, k_r is the reverse (dissociation) rate constant, and k_{cat} is the catalytic rate constant (enzyme turnover number). The Michaelis-Menton pseudo-steady state hypothesis assumes that, in the biological enzyme reaction, En, A, and EnA reach equilibrium rapidly relative to the rate of catalysis. Thus, according to the Law of Mass Action,

$$k_f [A][En] \rangle\rangle k_{cat} [EnA] \text{.....(2)}$$

In this case, the ratio of $\frac{k_f [A][En]}{k_{cat} [EnA]}$ represents a large number Q that varies with time as the reaction approaches steady state,

$$Q = \frac{k_f [A][En]}{k_{cat} [EnA]} \text{.....(3)}$$

In the special condition (denoted by *) where $[A]^* = K_m$ (the Michaelis-Menton constant), 50% of En is saturated with A at steady-state equilibrium; In other words, the amounts of free and substrate-bound enzyme are equal,

$$K_m = \frac{k_r + k_{cat}}{k_f} = [A]^* \text{,.....(4)}$$

$$[En]^* = [EnA]^*.$$

Consequently, the large quantity Q becomes a constant in this condition,

$$Q^* = \frac{k_f}{k_{cat}} \frac{[A]^* [En]^*}{[EnA]^*} = \frac{k_f \cdot K_m}{k_{cat}} \text{.....(5)}$$

Therefore, we define Λ as a special case of Q above,

$$\Lambda = \frac{k_f \cdot K_m}{k_{cat}} \text{ ……..……...(6)}$$

$$k_f [A]^* [En]^* = \Lambda \cdot k_{cat} [EnA]^* \text{ ……..(7)}$$

where Λ is a large *time-invariant* constant number, and the larger Λ is, the faster *EnA* reaches steady state equilibrium.

By rearrangement of (6), we can approximate k_f as,

$$k_f = \frac{\Lambda \cdot k_{cat}}{K_m} \text{. ……..(8)}$$

Now, by substitution of (8) into (4), we can approximate k_r as

$$k_r = (\Lambda - 1) \cdot k_{cat} \text{ ……..(9)}$$

By substitution of equations (3) and (9) into the differential equation describing the state of *EnA*, we obtain,

$$\begin{aligned} \frac{d[EnA]}{dt} &= k_f [A][En] - k_r [EnA] - k_{cat} [EnA] \\ &= k_f [A][En] - (\Lambda - 1) \cdot k_{cat} [EnA] - k_{cat} [EnA] \\ &= k_f [A][En] - \Lambda \cdot k_{cat} [EnA] \\ &= k_f [A][En] - \Lambda \cdot \frac{k_f [A][En]}{Q} \\ &= (1 - \frac{\Lambda}{Q}) \cdot k_f [A][En] \text{ ……...(10)} \end{aligned}$$

If the value of Λ is large enough and close to Q ($\frac{\Lambda}{Q} \approx 1$), the equation (10) ≈ 0, and *EnA* reaches steady state equilibrium after a short time (equation (7)), which is the same as the Michaelis-Menton pseudo-steady state assumption[22]. To determine the value of Λ, we varied Λ from 10 to 1,000,000 in simulations with no significant changes in the steady levels of intermediates and end-products. Consequently, in simulations reported here, the value of Λ is always set to 100.

K_m values are obtained from the literature, and k_{cat} (mole/min/mole) values are calculated from Specific Activities (μmole/min/mg) and molecular weights of purified enzymes. However, because of uncertainties about the percentages of purified enzymes that are active, k_{cat} values often require adjustments to fit experimental data.

The same Λ approximation methods, (7) and (8) are also used for multiple substrate and product enzymes that bind and release reactants in a Ping Pong fashion. However, in the general case of two-substrate, two-product, enzymes that bind and release reactants in a Bi Bi fashion, k_f and k_r are approximated according to equations (10) and (11).

$$k_f = \frac{\Lambda \cdot k_{cat}}{K_m^{\ A} \cdot K_m^{\ B}} \quad \text{……..(10)}$$

$$k_r = (\Lambda - 1) \cdot k_{cat} \quad \text{……..……(11)}$$

where $K_m^{\ A}$ is the K_m for substrate A, and $K_m^{\ B}$ is the K_m for substrate B.

Omega approximation method for enzyme inhibition rate constants:

In addition to needing methods to approximate forward and reverse reaction rates of enzyme substrate binding reactions, k_f, k_r, we need to be able to estimate the forward and reverse rate constants for enzyme-inhibitor binding reactions, k_{fi}, k_{ri}. To accomplish this we define a value Ω that approximates the rate that an enzyme binds to its substrate relative to its inhibitor. Like the derivation of Λ, the derivation of Ω is based on the Michaelis-Menten pseudo-steady state assumption. The k_{fi} is approximated by

$$k_{fi} = \Omega \cdot k_f$$

, and k_{ri} is approximated by

$$k_{ri} = \Omega \cdot k_f \cdot K_i .$$

As shown in the detailed derivation provided as Supplementary Figure 2 online, value of Ω can be an arbitrary number if free enzyme, enzyme-inhibitor complex, and enzyme-substrate complex are in the steady state equilibrium. Consequently, in simulations reported here, the value of Ω is always set to 1.

A list of reported and optimized enzyme kinetic and physical parameters needed to solve the above equations can be found in Supplementary Table 1 online.

Approximation of Intracellular Enzyme Concentrations from Microarray analysis

With few exceptions, intracellular enzyme concentrations are not available. However, with careful experimental documentation, these concentrations can be

approximated from the yields and specific activities of purified enzymes. For example, calculations based on purification tables in the literature suggest that the intracellular concentration of L-threonine deaminase (TDA) is 4 μM [23]. Furthermore, recent experiments have shown a positive correlation between mRNA levels measured with DNA arrays and protein abundance in both *E. coli* [24] and yeast cells [25]. Thus, the intracellular levels of the remaining enzymes of the branched chain amino acid biosynthetic pathway can be inferred from the calculated intracellular level of L-threonine deaminase and the relative mRNA levels of the other branched chain amino acid biosynthetic enzymes. The data in Supplementary Table 1 online demonstrate that this is a reasonable method. Indeed, simulations using intracellular enzyme concentrations inferred in this manner using DNA microarray data [9] produce experimentally observed steady state pathway intermediate, and end-product levels [10, 11] within two-fold to one-half adjustments of their inferred values.

Simulation of L-Threonine Deaminase Activity and Regulation

The concerted transition allosteric model of Monod, Wyman, and Changeux, the MWC model, is described by two equations [8]. The first equation describes the fraction of the enzyme in the or catalytically active state (R) as a function of substrate and effector concentrations.

$$\mathrm{R} = \frac{(1+a)^n}{L(1+ca)^n + (1+a)^n} \qquad (1)$$

The second equation describes the fractional saturation (Y_f) of the enzyme occupied by substrate as a function of substrate and effector concentrations.

$$Y_f = \frac{Lca(1+ca)^{n-1} + a(1+a)^{n-1}}{L(1+ca)^n + (1+a)^n} \qquad (2)$$

where $L = L_0 \frac{(1+b)^n}{(1+g)^n}, a = \frac{S}{K_m}, b = \frac{I}{K_i}, g = \frac{A}{K_a}$

In order to simulate the MWC model several parameters such as substrate (S), activator (A), and inhibitor (I) concentrations and their respective dissociation constants K_m, K_a, and K_i must be known:. With these values, usually available in the literature, the values of α, β, and γ are calculated. The value of n, the number of substrate and effector ligand binding sites is also, usually available (n = 4 for *E. coli* L-threonine deaminase); however, with few exceptions the values of c, the ratio of the affinity of the substrate for the catalytically active R state and the inhibited T state, and L_0, the equilibrium constant (allosteric constant) for the R and T states in the absence of ligands, are unavailable. However, these values can be readily derived by fitting generally available substrate saturation curves generated in the presence of several inhibitor concentrations as originally described by Hatfield [26]. The values of c (0.013) and L_0 (1.05), used for the simulations reported here were obtained by finding the minimum sum of squared differences between theoretical data and experimental data with the non-linear programming Mathematica function, FindMinimum (see Supplementary Figure 3 online).

The Development of kMech Models for the Simulation of Enzyme Reaction Mechanisms

Ping Pong Bi Bi Model The α-acetohydroxyacid synthase (AHAS) isozymes catalyze the condensation of either two molecules of pyruvate (pyr) to form one molecule of the L-valine and L-leucine precursor α-acetolactate (αAL), or one molecule of pyruvate and one molecule of α-ketobutyrate (αKB) to form one molecule of the L-isoleucine precursor, α-aceto-α-hydroxybutyrate (αAHB) (Fig. 1). Each of these isozymes catalyzes their reactions with a Ping Pong Bi Bi enzyme mechanism [27]. It is an ordered Bi Bi mechanism because a pyruvate molecule must bind to the enzyme, react with a thiamine pyrophosphate cofactor to form an active acetaldehyde group (CH3CO), and release the first product, CO_2, before the second substrate, pyruvate or α-ketobutyrate can bind to the enzyme. Next, the enzyme bound active acetoaldehyde group, must be transferred to the second substrate, pyruvate or α-ketobutyrate, to form the second product, α-acetolactate or α-aceto-α-hydroxybutyrate and to release the free enzyme. It is a Ping Pong

mechanism because the enzyme shuttles between a free and a substrate-modified intermediate state.

These enzyme mechanisms can be represented in a reaction-like notation for input into kMech and Cellerator, and subsequent computer simulation. The kMech input representation for the Ping Pong Bi Bi model is

```
        En, Enx
Enz[{A, B} ⇔ {C, F}, PingPong[kf1, kr1, kcat1, kf2, kr2, kcat2]]
```

where A and B are the substrates; C and F are the products; En is the free enzyme; and Enx is the modified enzyme intermediate; kf1 and kf2 are the rate constants of the enzyme- substrate binding for A and B, respectively; kr1 and kr2 are the dissociation rate constants for the enzyme substrates A and B, respectively; kcat1 and kcat2 are the catalytic rate constants for the formation of products C and F, respectively; and the prefix, Enz[…] is the name of a new formal function defining additional capabilities provided to Cellerator by kMech .

kMech interprets this input and converts it to more elementary reactions (in this case, two single-substrate single-product reactions) defined in Cellerator as follows:

```
{A + En ⇔ $Complex$A$En$, kf1, kr1},
{$Complex$A$En$ ⟶ C + Enx, kcat1},
{B+ Enx ⇔ $Complex$B$Enx$, kf2, kr2},
{$Complex$B$Enx$ ⟶ F + En , kcat2}
```

The first of these symbolic reactions represents the formation of the enzyme complex with substrate A (which is given the new name "$Complex$AEn"). This reaction is reversible and has forward rate kf1 and reverse rate kr1. The second reaction represents the release of product C from the complex $Complex$AEn, and the formation of the enzyme intermediate Enx. It is irreversible and has rate kcat1. The third and fourth reactions of the kMech output represent the formation of the enzyme complex with substrate B ($Complex$BEnx from B and Enx) with rates of kf2 and kr2, and the release of product F and free enzyme En with rate kcat2.

Given these reactions, the Cellerator "interpret" function generates the following Mathematica formatted differential equations:

A'[t] = = -kf1 A[t] En[t] + kr1 $Complex$AEn[t],

B'[t] = = -kf2 B[t] Enx[t] + kr2 $Complex$BEnx[t],

C'[t] = =kcat1 $Complex$AEn[t],

F'[t] = = kcat2 $Complex$BEnx[t],

En'[t]= = -kf1 A[t] En[t] + kr1 $Complex$AEn[t] + kcat2 $Complex$BEnx[t],

Enx'[t]= = -kf2 B[t] Enx[t] + kcat1 $Complex$AEn [t] + kr2 $Complex$BEnx[t],

$Complex$AEn'[t] = = kf1 A[t] En[t] – kcat1 $Complex$AEn[t] – kr1 $Complex$AEn[t],

$Complex$BEnx'[t] = = kf2 B[t] Enx[t] – kcat2 $Complex$BEnx[t] – kr2 $Complex$BEnx[t]

These differential equations and variable definitions are passed to Mathematica where they are solved by the numeric solver (NDSolve) function and time plots are generated. Alternatively, Cellerator can output a model in Systems Biology Markup Language (SBML) for input into another cell simulation environment [3].

Carbon Flow Channeling The above differential equations accurately model the kinetic behavior of AHAS II, which is not inhibited by L-valine. However, since AHAS isozymes are controllers of carbon flow distribution to either L-isoleucine or L-valine and L-leucine biosynthesis by selectivity (K_m) of their second substrate [13], it was necessary to combine the "Union" operator from Mathematica with kMech inputs as shown below:

```
Union[
        AHASII, AHASIICH3CO
Enz[{Pyr, Pyr}⇔ {CO2, aAL}, PingPong[kfAHASIIPyr, krAHASIIPyr, kcat$AHASII$Pyr, kfAHASIIPyr2,
krAHASIIPyr2, kcat$AHASII$Pyr2]],

        AHASII, AHASIICH3CO
Enz[{Pyr, aKB}⇔ {CO2, aAHB}, PingPong[kfAHASIIPyr, krAHASIIPyr, kcat$AHASII$Pyr, kfAHASIIaKB,
krAHASIIaKB, kcat$AHASII$aKB]]
]
```

(The first reaction is for the L-valine & L-leucine pathways, and the second reaction is for the L-isoleucine pathway.)

This union operator is used to eliminate the duplicated first substrate reaction, and allows the following reactions that represent a branch-point for carbon flow into the L-valine and L-leucine or L-isoleucine pathways [28]. This channeling of carbon flow is controlled by the active AHAS II isozyme intermediate(AHASIICH3CO):

+ Pyr ⟺ $Complex$Pyr$AHASIICH3CO$b ⟶ aAL

Pyr + AHASII ⟺ $Complex$Pyr$AHASII$ ⟶ CO2 + AHASIICH3CO

+ aKB ⟺ $Complex$aKB$AHASIICH3CO$ ⟶ aAHB

(aAL is the precursor of the L-valine & L-leucine pathways; aAHB is the precursor of the L-isoleucine pathway)

Reversible Transamination Mechanism The reversible mechanism of the transaminase B and C enzymes (Fig. 1) are described by a Ping Pong model in which the enzyme shuttles between pyridoxal phosphate (TB) and pyridoxamine phosphate (TBNH2) bound states. Thus, the kMech inputs of the transaminase B reaction for L-isoleucine synthesis are as follows:

```
                TB, TBNH2
Enz[{Glu, aKMV} ⟺ {aKG, Ile}, PingPong[fkfTBGlu, fkrTBGlu, fkcat$TB$Glu, fkfTBaKMV, fkrTBaKMV,
fkcat$TB$aKMV]]                          (forward reaction)

                TB, TBNH2
Enz[{Ile, aKG} ⟺ {aKMV, Glu}, PingPong[rkfTBIle, rkrTBIle, rkcat$TB$Ile, rkfTBaKG, rkrTBaKG,
rkcat$TB$aKG]]                           (reverse reaction)
```

Transaminase B catalyzes the final step of the biosynthetic pathways of all three of the branched chain amino acids, L-isoleucine, L-valine, and L-leucine. Each of these transamination reactions use glutamate as an amino donor to form the pyridoxamine intermediate (TBNH2) that can transaminate the a-ketoacids of each pathway. The kMech input for these reactions are

```
Union[
                TB, TBNH2
Enz[{Glu, aKMV} ⟺ {aKG, Ile}, PingPong[fkfTBGlu, fkrTBGlu, fkcat$TB$Glu, fkfTBaKMV, fkrTBaKMV,
fkcat$TB$aKMV]]

                TB, TBNH2
Enz[{Glu, aKIV} ⟺ {aKG, Val}, PingPong[fkfTBGlu, fkrTBGlu, fkcat$TB$Glu, fkfTBaKIV, fkrTBaKIV,
fkcat$TB$aKIV]]

                TB, TBNH2
Enz[{Glu, aKIC}⟺ {aKG, Leu}, PingPong[fkfTBGlu, fkrTBGlu, fkcat$TB$Glu, fkfTBaKIC, fkrTBaKIC,
fkcat$TB$aKIC]]
]
```

Since each of these pathway reactions share the glutamate substrate reaction, the Union operator is used to eliminate the duplicated first substrate reactions.

$$\begin{array}{lll} & + \text{aKMV} \Leftrightarrow \$\text{Complex}\$\text{aKMV}\$\text{TBNH2}\$ \rightarrow \text{Ile} \\ \text{Glu} + \text{TB} \Leftrightarrow \$\text{Complex}\$\text{Glu}\$\text{TB}\$ \rightarrow \text{aKG} + \underline{\text{TBNH2}} & + \text{aKIV} \Leftrightarrow \$\text{Complex}\$\text{aKIV}\$\text{TBNH2}\$ \rightarrow \text{Val} \\ & + \text{aKIC} \Leftrightarrow \$\text{Complex}\$\text{aKIC}\$\text{TBNH2}\$ \rightarrow \text{Leu} \end{array}$$

In this case, the channeling of carbon flow to end-products is controlled by the enzyme-pyridoxamine intermediate (TBNH2).

A similar approach was applied on the reverse reaction to eliminate the duplicated second substrate reaction

$$\begin{array}{l} \text{Ile} + \text{TB} \Leftrightarrow \$\text{Complex}\$\text{Ile}\$\text{TB}\$ \rightarrow \text{aKMV} + \\ \text{Val} + \text{TB} \Leftrightarrow \$\text{Complex}\$\text{Val}\$\text{TB}\$ \rightarrow \text{aKIV} + \underline{\text{TBNH2}} + \text{aKG} \Leftrightarrow \$\text{Complex}\$\text{aKG}\$\text{TBNH2}\$ \rightarrow \text{TB} + \text{Glu} \\ \text{Leu} + \text{TB} \Leftrightarrow \$\text{Complex}\$\text{Leu}\$\text{TB}\$ \rightarrow \text{aKIC} + \end{array}$$

Extensions of Enzyme Reaction Models with Regulatory Circuits

The AHAS I and III isozymes are controlled by a noncompetitive inhibition mechanism [13]. This mechanism can be described by the following kMech input :

$$\text{Enz}\left[\{\text{A, B}\} \overset{\text{En, Enx}}{\Leftrightarrow} \{\text{C, F}\}, \text{PingPong}[\text{kf1, kr1, kcat1, kf2, kr2, kcat2}], \text{NCI}[\text{inh, kfi1, kri1, kfi2, kri2, residualRate}]\right]$$

Here, inh is the inhibitor name, kfi1, kri1, and kfi2, kri2 are inhibitor association and dissociation rate constants for the first and second substrate reactions, respectively, and NCI stands for non-competitive inhibition. residualRate is the fraction of maximal enzyme activity (V_{max}) remaining at saturating inhibitor concentrations. The detailed reactions and equations that include these noncompetitive inhibition and residual activity parameters can be found as Supplementary Figure 4 online.

In the case of α-isopropylmalate synthase (IPMS), the binding of the first substrate acetyl-coA, is competitively inhibited by L-leucine; and the second substrate reaction with α-ketoisovalerate is non-competitively inhibited by L-leucine [29]. The kMech input for this model is

$$\text{En, Enx}$$

Enz[{A, B}⇔ {C, F}, PingPong[kf1, kr1, kcat1, kf2, kr2, kcat2], CI[inh, kfi1, kri1], NCI[inh, kfi2, kri2]]

Here CI stands for competitive inhibition. Other notations are as described above. The detailed reactions and equations for competitive and noncompetitive reactions are included in the Supplementary Figure 4 online.

Bi Bi Model (General Case)

Unlike the Ping Pong Bi Bi mechanisms described above where two substrates must bind in an ordered fashion, the general Bi Bi describes a two-substrate two-product reaction where there is no required order for the binding of substrates to the enzyme, or the release of products, and product formation occurs only after the formation of the enzyme-two substrates complex. The kMech input for this Bi Bi mechanism is

$$\mathrm{Enz}[\{A, B\} \overset{\mathrm{En}}{\Leftrightarrow} \{C, F\}, \text{Bi Bi}[kf1, kr1, kcat1]]$$

The resulting Cellerator reactions and Mathematica formatted differential equations for this enzyme reaction are shown in the Supplementary Figure 4 online. This kMech model describes the acetohydroxyacid isomeroreductase (IR) with substrates α-acetolactate (αAL) or α-aceto-α-hydroxybutyrate (αAHB) and NADPH, as well as the reaction of β-isopropylmalate dehydrogenase (IPMDH) with its substrate β-isopropylmalate (βIPM) and NAD.

Simple Catalytic Model

The dihydroxyacid dehydrase (DAD) enzyme catalyzes a simple single substrate, single product reaction that involves no cofactors or enzyme intermediates. This is a basic Cellerator reaction called 'catalytic' represented as

$$\{A \overset{\mathrm{En}}{\Leftrightarrow} F, kf, kr, kcat\}$$

This model also describes the α- isopropylmalate isomerase (IPMI), and the active transport systems of branch chain amino acid transport enzymes catalyzed by the LIV I (for L-leucine, L-isoleucine, and L-valine), and LS (L-leucine specific) transport

systems[30]. These transport systems were incorporated to simulate the uptake of extracellular branched chain amino acid additions required for perturbation experiments described above.

ACKNOWLEDGEMENTS

This work was supported in part by by a National Institutes of Health grant GM55073 (GWH). CY is the recipient of a postdoctoral fellowship form National Research Service Award 5 T15 LM00744 from the National Library of Medicine. We are appreciative of helpful advice from Donald F. Senear.

REFERENCES

1. Umbarger, H.E. Biosyhthesis of the Branched-Chain Amino Acids, Vol. I, Edn. 2nd. (ASM Press, Washington, D.C.; 1996).
2. Shapiro, B.E., Levchenko, A., Meyerowitz, E.M., Wold, B.J. & Mjolsness, E.D. Cellerator: extending a computer algebra system to include biochemical arrows for signal transduction simulations. *Bioinformatics* 19, 677-678 (2003).
3. Hucka, M. et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* 19, 524-531 (2003).
4. Chassagnole, C., Rais, B., Quentin, E., Fell, D.A. & Mazat, J.P. An integrated study of threonine-pathway enzyme kinetics in Escherichia coli. *Biochem J* 356, 415-423 (2001).
5. Chassagnole, C., Fell, D.A., Rais, B., Kudla, B. & Mazat, J.P. Control of the threonine-synthesis pathway in Escherichia coli: a theoretical and experimental approach. *Biochem J* 356, 433-444 (2001).
6. Mendes, P. Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3. *Trends Biochem Sci* 22, 361-363 (1997).
7. JWS Online: Online Cellular Systems Modelling. *http://jjj.biochem.sun.ac.za/database/index.html.*
8. Monod, J., Wyman, J. & Changeux, J.P. On the nature of allosteric transitions: A Plausible Model. *J Mol Biol* 12, 88-118 (1965).
9. Hung, S.P., Baldi, P. & Hatfield, G.W. Global gene expression profiling in Escherichia coli K12. The effects of leucine-responsive regulatory protein. *J Biol Chem* 277, 40309-40323 (2002).
10. Epelbaum, S. et al. Branched-chain amino acid biosynthesis in Salmonella typhimurium: a quantitative analysis. *J Bacteriol* 180, 4056-4067 (1998).
11. Quay, S.C., Dick, T.E. & Oxender, D.L. Role of transport systems in amino acid metabolism: leucine toxicity and the branched-chain amino acid transport systems. *J Bacteriol* 129, 1257-1265 (1977).
12. Lawther, R.P. et al. DNA sequence fine-structure analysis of ilvG (IlvG+) mutations of Escherichia coli K-12. *J Bacteriol* 149, 294-298 (1982).
13. Herring, P.A., McKnight, B.L. & Jackson, J.H. Channeling behavior and activity models for Escherichia coli K-12 acetohydroxy acid synthases at physiological substrate levels. *Biochem Biophys Res Commun* 207, 48-54 (1995).
14. Umbarger, H.E. & Brown, B. *J Bacteriol* 70, 241-248 (1955).
15. Jackson, J.H., Herring, P.A., Patterson, E.B. & Blatt, J.M. A mechanism for valine-resistant growth of Escherichia coli K-12 supported by the valine-sensitive acetohydroxy acid synthase IV activity from ilvJ662. *Biochimie* 75, 759-765 (1993).
16. Daniel, J., Dondon, L. & Danchin, A. 2-Ketobutyrate: a putative alarmone of Escherichia coli. *Mol Gen Genet* 190, 452-458 (1983).
17. Danchin, A., Dondon, L. & Daniel, J. Metabolic alterations mediated by 2-ketobutyrate in Escherichia coli K12. *Mol Gen Genet* 193, 473-478 (1984).

18. Adams, C.W., Rosenberg, M. & Hatfield, G.W. Analysis of in vivo RNA transcription products of the ilvGEDA attenuator region of Escherichia coli K12. *J Biol Chem* 260, 8538-8544 (1985).
19. Hashiguchi, K., Takesada, H., Suzuki, E. & Matsui, H. Construction of an L-isoleucine overproducing strain of Escherichia coli K-12. *Biosci Biotechnol Biochem* 63, 672-679 (1999).
20. Neidhardt, F.C. & Curtiss, R. Regulation of Gene Expression, Edn. 2nd. (ASM Press, Washington, D.C.; 1996).
21. Hatfield, G.W. & Benham, C.J. DNA topology-mediated control of global gene expression in Escherichia coli. *Annu Rev Genet* 36, 175-203 (2002).
22. Murray, J.D. Michaelis-Menton Theory: Detailed Analysis and the Pseudo-Steady State Hypothesis. *Mathematical Biology* 5.2, 116-117 (1993).
23. Calhoun, D.H., Rimerman, R.A. & Hatfield, G.W. Threonine deaminase from Escherichia coli. I. Purification and properties. *J Biol Chem* 248, 3511-3516 (1973).
24. Arfin, S.M. et al. Global gene expression profiling in Escherichia coli K12. The effects of integration host factor. *J Biol Chem* 275, 29672-29684 (2000).
25. Futcher, B., Latter, G.I., Monardo, P., McLaughlin, C.S. & Garrels, J.I. A sampling of the yeast proteome. *Mol Cell Biol* 19, 7357-7368 (1999).
26. Hatfield, G.W. The Regulation of L-Threonine Deaminase in Bacillus Subtilis by Repression and Endproduct Inhibition. *Ph.D. Thesis, Biochemistry, Purdue University*, 103-112 (1970).
27. Segel, I.H. Enzyme kinetics : behavior and analysis of rapid equilibrium and steady state enzyme systems, Edn. Wiley Classics Library. (Wiley, New York; 1993).
28. Gollop, N., Damri, B., Barak, Z. & Chipman, D.M. Kinetics and mechanism of acetohydroxy acid synthase isozyme III from Escherichia coli. *Biochemistry* 28, 6310-6317 (1989).
29. Kohlhaw, G., Leary, T.R. & Umbarger, H.E. Alpha-isopropylmalate synthase from Salmonella typhimurium. Purification and properties. *J Biol Chem* 244, 2218-2225 (1969).
30. Wood, J.M. Leucine transport in Escherichia coli. The resolution of multiple transport systems and their coupling to metabolic energy. *J Biol Chem* 250, 4477-4485 (1975).

Figure Legends:

Figure 1. The Biosynthetic Pathways for the Branched Chain Amino Acids, L-Isoleucine, L-Valine, and L-Leucine. The bi-functional enzymes involved in the common pathway for branched chain amino acid biosynthesis are abbreviated as follows: TDA, L-threonine deaminase (EC 4.3.1.19); AHAS, acetohydroxyacid synthase (EC 4.1.3.18); IR, acetohydroxyacid isomeroreductase (EC 1.1.1.86); DAD, dihydroxyacid dehydrase (EC 4.2.1.9); TB, transaminase B (EC 2.6.1.42); TC, transaminase C (EC 2.6.1.66); IPMS, α-isopropylmalate synthase (EC 4.1.3.12); IPMI, α-isopropylmalate isomerase (EC 4.2.1.33); IPMDH, β-isopropylmalate dehydrogenase (EC 1.1.1.85). Gene names for each enzyme are italicized. Feedback inhibition patterns are indicated by dashed lines. The amino donor for TB is L-glutamate. The amino donor for TC is L-alanine.

Figure 2. Simulated Flow of Carbon Through the Branched Chain Amino Acid Biosynthetic Pathways of *Escherichia coli* K12. The graphical insets show the approach (minutes) to steady state (μM) synthesis and utilization of the intermediates and end-products of the pathways. The intermediates are abbreviated as follows: Pyr, pyruvate; αKB, α-ketobutyrate; αAL, α-acetolactate; αAHB, α-aceto-α–hydroxybutyrate; αDHIV, α,β-dihydroxy-isovalerate; αDMV, α, β-dihydroxy-β-methylvalerate; αKIV, α-ketoisovalerate; αKMV, α-keto-β-methylvalerate; αIPM, α-isopropylmalate; βIPM, β-isopropylmalate; αKIC, α-ketoisocaproate. The starting substrates L-threonine and pyruvate are supplied at rates to maintain constant levels of 520 μM and 1000 μM, respectively. For the transamination reactions, L-glutamate (TB) and L-alanine (TC) are supplied at rates to maintain a constant level of 2000 μM each. For the isomeroreductase (IR) reaction, NADPH is supplied at a rate to maintain a constant level of 1000 μM. For the isopropylmalate synthase (IPMS) reaction, acetyl-coA is supplied at a rate to maintain a constant level of 1000 μM. The beginning substrates (L-threonine and pyruvate) levels, as well as the end-product (L-isoleucine, L-valine, and

L-leucine) levels, agree with measured values [10, 11]. Where available, the ranges of reported values for pathway intermediate and end-product levels in cells growing in a glucose minimal salts medium are shown in parentheses (μM) in the inset graphs.

Figure 3. Allosteric Regulation of L-Threonine Deaminase (TDA). (A) The Fraction of TDA in the Active R State. At time = 0, and an initial L-threonine concentration of 520 μM, about 0.65 of the TDA enzyme is in the active R state. As L-isoleucine accumulates, TDA is end-product inhibited and as L-valine accumulates this inhibition is slowly countered until at steady state only about 0.05 of the total enzyme is in the active R state. (B) The Fractional saturation of TDA with L-threonine (v_0/Vmax). At time = 0, and an initial L-threonine concentration of 520 μM, 0.08 of the total enzyme is saturated with L-threonine. At a final steady state level of end-product synthesis, it is only 0.012 saturated with L-threonine.

Figure 4. Metabolic Effects of Excess L-Valine on Branched Chain Amino Acid Biosynthesis in *Escherichia coli* K12. The simulation conditions described in Figure 2 were used for the simulations presented here except excess extra-cellular L-valine was added at a rate sufficient to be maintained at a concentration of 1 mM. The data in panel (A) show that, as described in the text, excess L-valine increases rather than inhibits L-isoleucine biosynthesis. The data in Panel (B) show that excess L-valine causes a three-fold increase in the intracellular accumulation of α-ketobutyrate (αKB), that is restored to control levels by the extra-cellular addition of 500 μM L-isoleucine. The data in panel (C) show that the accumulation of α-ketobutyrate observed in the presence of excess L-valine coincides with the conversion of nearly 18% of the cellular L-threonine deaminase to a catalytically active R state; and, that the subsequent extracellular addition of 500 μM L-isoleucine reverses this transition to the control level.

Figure 5. Metabolic Engineering of an *Escherichia coli* K12 Strain that Overproduces L-Isoleucine. The simulation conditions described in Figure 2 were used for the simulations presented here except that the L-threonine deaminase feedback resistant mutant (TDA^R) was simulated by increasing its K_i for L-isoleucine to 100,000 μM; and the *ilvGMEDA* operon attenuator mutant (*ilvGMEDA-att*$^-$) was simulated by increasing TDA, AHAS II, DAD and TB total enzyme levels 11 fold [18]. The simulation in panel (A) shows that the effect of the feedback resistant TDA mutant (TDA^R) is to allow the positive effector ligands, L-threonine and L-valine to transition nearly 100% of the TDA enzyme to the active R state. The simulation results in panel (B) show that L-isoleucine production in the TDA^R mutant is 5 to 6-fold increased. The simulation in panel (C) demonstrates that in a TDA^R K12 mutant, the intermediate, α-ketobutyrate (αKB) accumulates to a level 40-fold higher than in a wild type K12 strain; however, when the AHAS II isozyme is restored, and the bi-functional enzymes of the L-isoleucine and L-valine pathways are genetically de-repressed 11-fold (*ilvGMEDA-att*$^-$), α-ketobutyrate accumulation is relieved (panel C), and L-isoleucine synthesis is increased more than 40-fold over the wild-type K12 level (panel D).

Figure 6. An Acetohydroxyacid Isomeroreductase (IR) mutant (*ilvC*) *Escherichia coli* K12 Strain is auxotrophic for L-Isoleucine and L-Valine, but not L-Leucine. The simulation conditions described in Figure 2 were used for the simulations presented here except that the initial concentration of isomeroreductase (IR) were set to zero to simulate an *ilvC* mutation, and extra-cellular L-valine and L-isoleucine were supplied at a level of 500 μM each. The results show that αKIV (panel A) and L-leucine (panel B) are produced in an *ilvC* strain in the presence of extra-cellular L-valine and L-isoleucine.

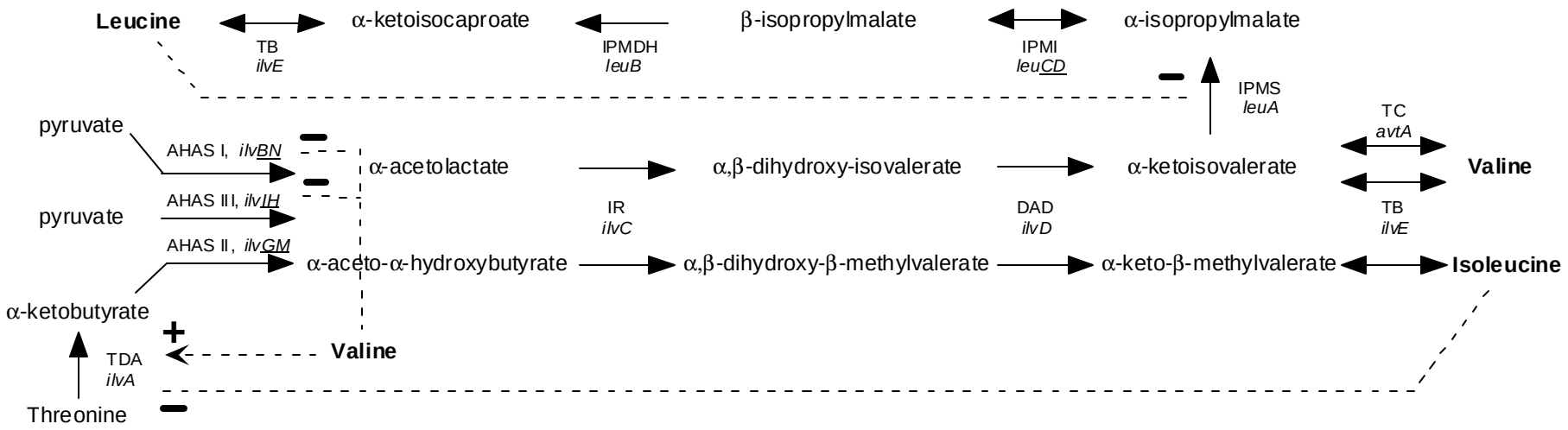


Figure 1 (Hatfield)

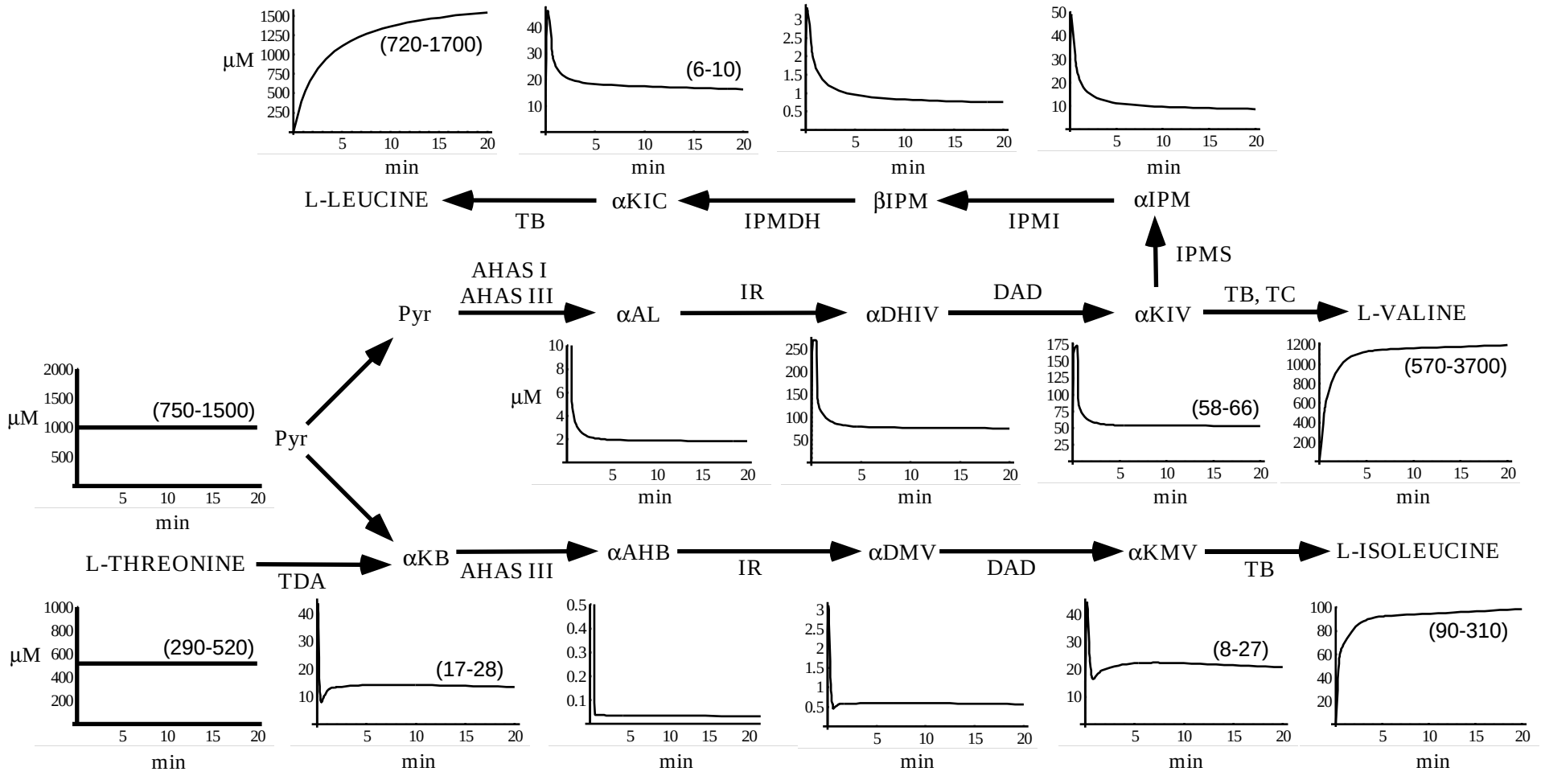


Figure 2 (Hatfield)

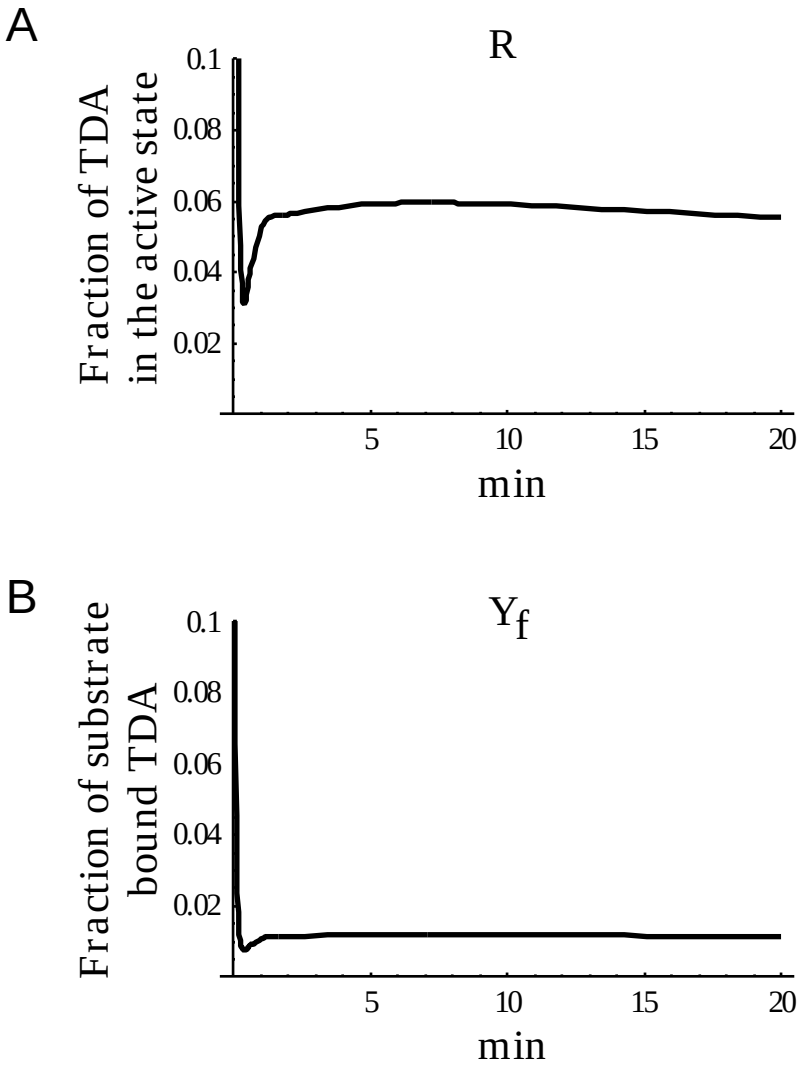
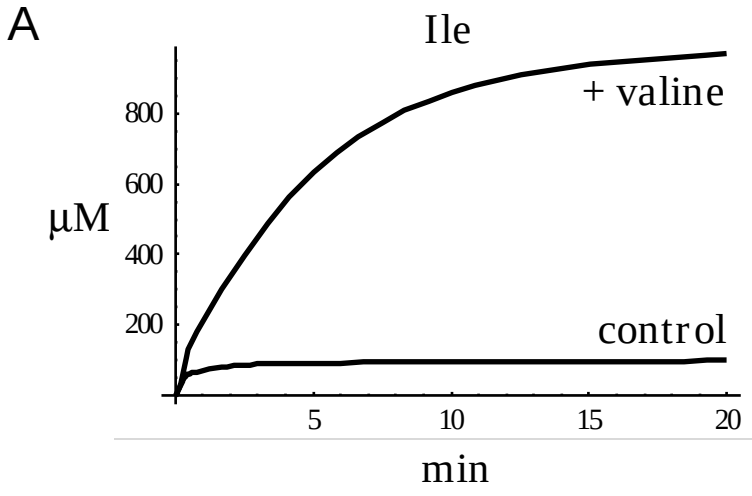
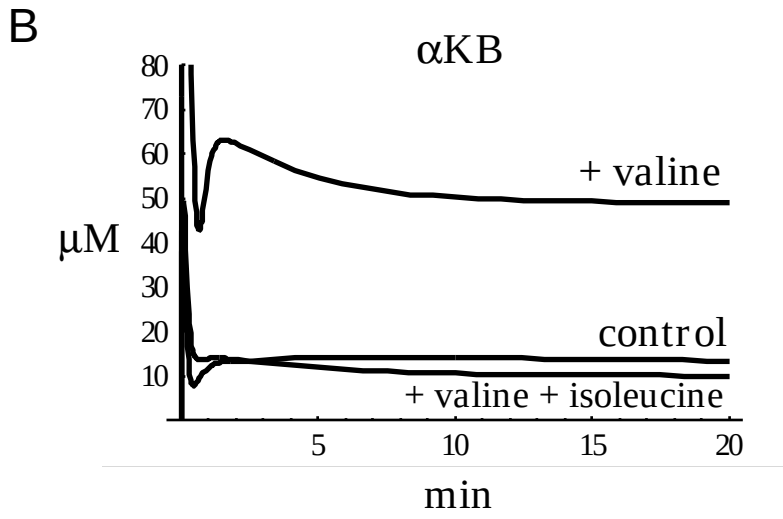


Figure 3 (Hatfield)





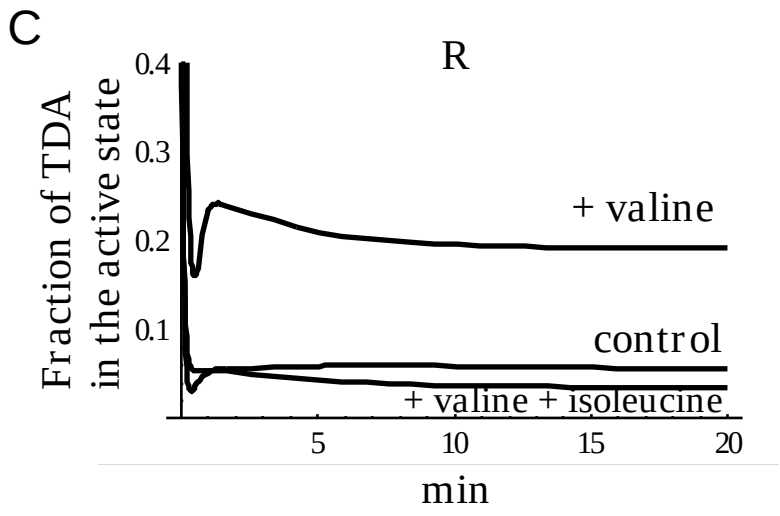


Figure 4 (Hatfield)

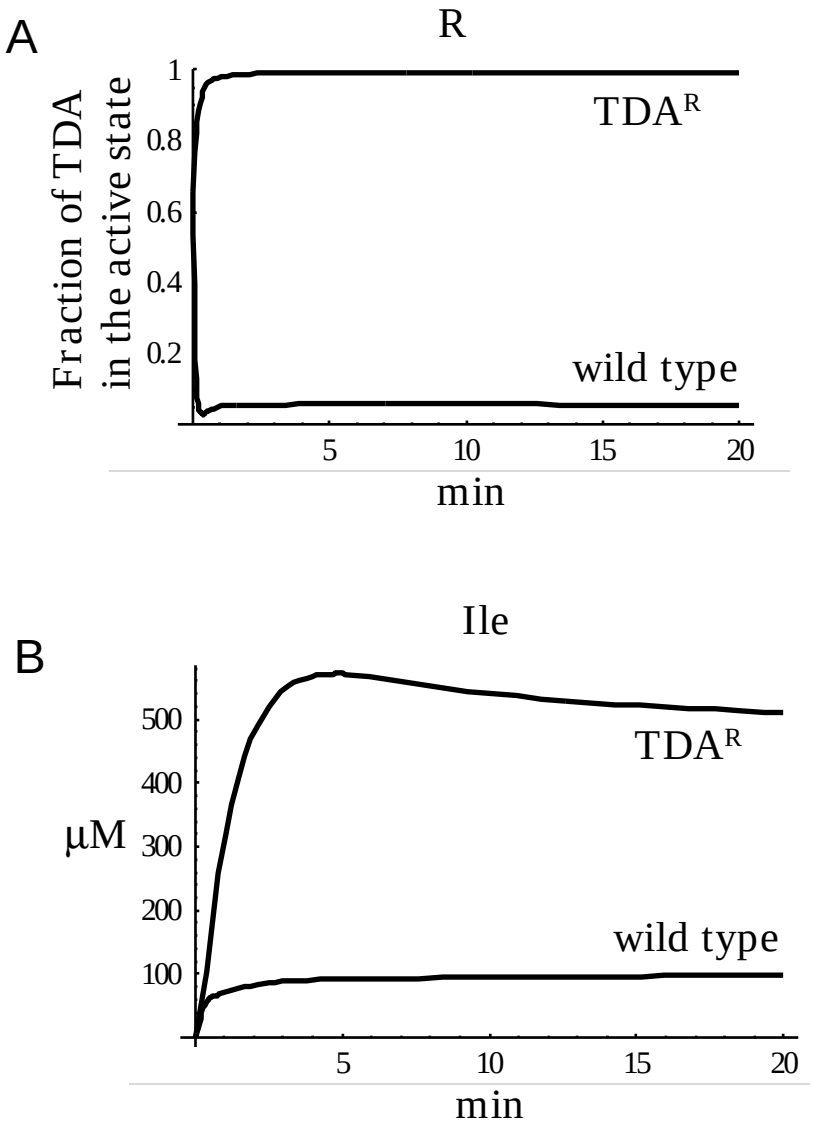
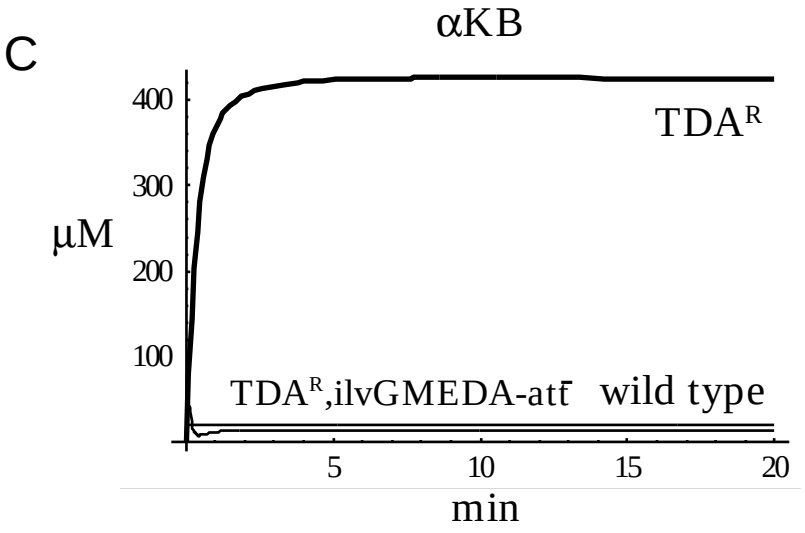


Figure 5 (Hatfield)



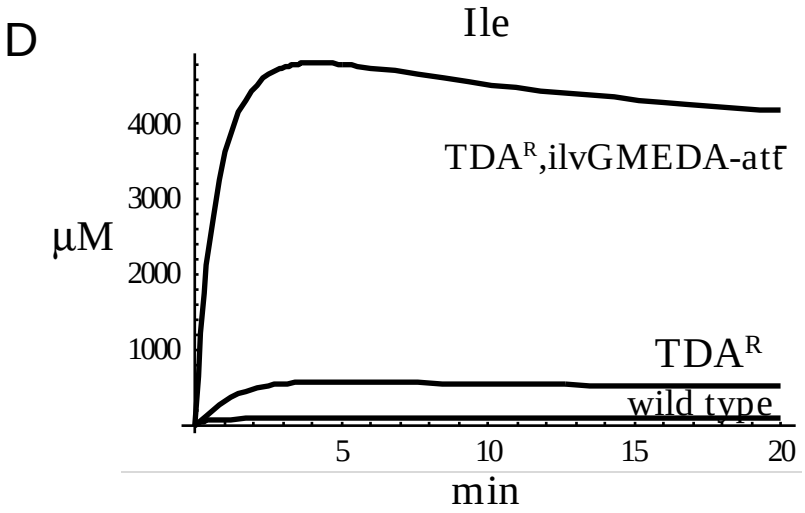


Figure 5 (Hatfield)

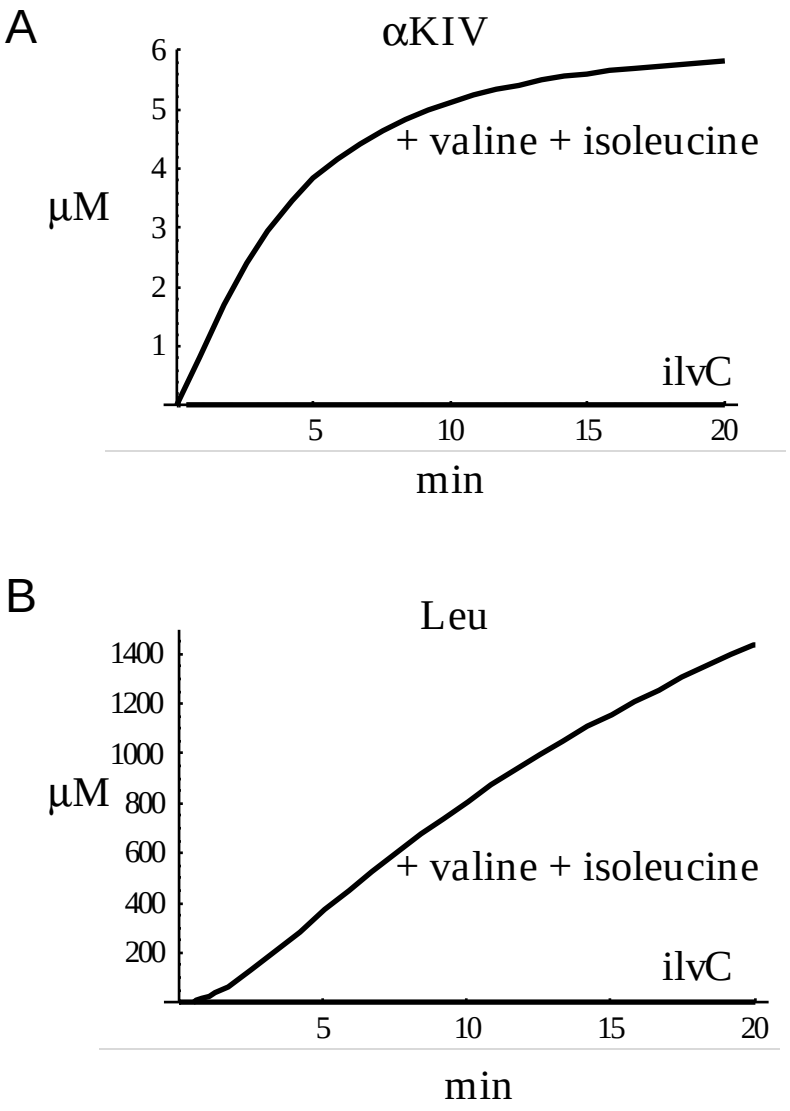


Figure 6 (Hatfield)