Supporting Information

Cornelius et al. 10.1073/pnas.1009772108

SI Text

A metabolic network of m metabolites and n reactions is represented by an $m \times n$ matrix $\boldsymbol{S} = (S_{ij})$, where S_{ij} is the stoichiometric coefficient of metabolite i in reaction j. The state of the metabolic network is represented by a vector of reaction fluxes $\boldsymbol{\nu} = (\nu_j)$, where ν_j is the flux through reaction j. In our analysis, the steady-state solutions of the system are determined by mass balance constraints,

$$\boldsymbol{S} \cdot \boldsymbol{\nu} = 0, \qquad \textbf{[S1]}$$

and additional constraints that limit the range of the individual fluxes,

$$\boldsymbol{\nu_{\min}} \leq \boldsymbol{\nu} \leq \boldsymbol{\nu}_{\max}, \qquad \textbf{[S2]}$$

where the inequalities in this notation are assumed to apply to each component individually. The bounds on individual fluxes are determined by substrate availability in the given medium, the ATP maintenance requirement, and thermodynamic constraints that limit the reversibility of the corresponding reaction. The knockout of the genes associated with enzymes catalyzing reaction j corresponds to the additional constraint $\nu_j = 0$. The exchange fluxes and the biomass flux are excluded from our implementation of the minimization of metabolic adjustment (MOMA) and regulatory on/off minimization (ROOM) objective functions.

Flux Balance Analysis (FBA). FBA (1) is used to identify flux vectors $\boldsymbol{\nu}$ satisfying **S1** and **S2** that maximize biomass production, which is represented by an additional reaction b that drains biomass components. The problem is implemented as a linear program:

$$\begin{aligned} \max \quad & \nu_b \\ \text{s.t.} \quad & \boldsymbol{S} \cdot \boldsymbol{\nu} = 0 \\ & \boldsymbol{\nu}_{\min} \leq \boldsymbol{\nu} \leq \boldsymbol{\nu}_{\max}. \end{aligned} \qquad \textbf{[S3]}$$

In general, the FBA solution is not unique, and the results in this paper were obtained by selecting the optimal solution provided by the simplex algorithm. For a discussion of the effects of choosing alternate FBA solutions, see *Sensitivity to Alternate Optima* below.

Minimization of Metabolic Adjustment. MOMA (2) selects a suboptimal growth state $\boldsymbol{\nu}$ by minimizing the Euclidean distance in flux space from the wild-type optimal growth state, $\boldsymbol{w}$. This is implemented as a quadratic programming problem:

$$\begin{aligned} \min \quad & (\boldsymbol{\nu} - \boldsymbol{w})^T \cdot (\boldsymbol{\nu} - \boldsymbol{w}) \\ \text{s.t.} \quad & \boldsymbol{S} \cdot \boldsymbol{\nu} = 0 \\ & \boldsymbol{\nu}_{\min} \leq \boldsymbol{\nu} \leq \boldsymbol{\nu}_{\max} \\ & \nu_j = 0, j \in A, \end{aligned} \qquad \textbf{[S4]}$$

where A is the set of indices corresponding to reactions deactivated by gene knockouts.

Regulatory On/Off Minimization. Unlike MOMA, which favors a potentially large number of small-magnitude flux changes, ROOM (3) chooses a suboptimal growth solution $\boldsymbol{\nu}$ with a minimal number of "significant" flux changes from the original state. This is usually implemented as a mixed-integer programming problem (integer ROOM):

$$\begin{aligned} \min \quad & \sum_{j=1}^{n} y_j \\ \text{s.t.} \quad & \boldsymbol{S} \cdot \boldsymbol{\nu} = 0 \\ & \boldsymbol{\nu}_{\min} \leq \boldsymbol{\nu} \leq \boldsymbol{\nu}_{\max} \\ & \nu_k = 0, k \in A \\ \text{for } j = 1,\ldots,n: & \\ & \nu_j - y_j(\nu_{\max,j} - w_j^u) \leq w_j^u \\ & \nu_j - y_j(\nu_{\min,j} - w_j^l) \geq w_j^l \\ & w_j^u = w_j + \delta|w_j| + \varepsilon \\ & w_j^l = w_j - \delta|w_j| - \varepsilon \\ & y_j \in \{0,1\}, \end{aligned} \qquad \textbf{[S5]}$$

where $\boldsymbol{w}$ and A are as for MOMA, and δ and ε express tolerances for relative and absolute change from the original state, respectively.

In our numerical experiments we chose to use a linear programming variant of the method (linear ROOM) obtained by allowing the above binary constraints to be continuous in the interval $0 \leq y_j \leq 1$ and setting $\delta = \varepsilon = 0$. Linear ROOM is biologically well-motivated, given that gene activity is best described by a continuous variable, and has the advantage of being computationally inexpensive for all mutants. Table S1 shows a comparison of the growth impacts predicted by integer ROOM and the linear variant upon latent pathway removal. For most mutants, the two methods show similar change in growth rate (increase, decrease, or an insignificant change) when the latent pathways are removed, even though the exact growth rate predictions may differ.

For the integer variant, we followed Shlomi et al. (3) in choosing the values $\delta = 0.03$ and $\varepsilon = 0.001$, which yielded reasonable running times for most knockout strains. Even for larger values of these tolerance values, there are a handful of cases for which no optimal solution is found in any reasonable amount of time. For these cases, in the comparison of Table S1, we take the best solution found after 1 h of computation on a 3.4-GHz CPU. The integer ROOM solutions were further constrained to minimize the aggregate flux change from the wild-type optimal state, $\sum |\nu_j - w_j|$.

Sensitivity to Alternate Optima. In general, the optimal flux distribution given by FBA is not unique, neither before nor after a given gene knockout (4, 5). The implications of this nonuniqueness must be considered for two reasons. First, the set of transiently active reactions is defined with respect to the exact optimal flux distribution before and after the knockout perturbation, as well as the suboptimal flux distribution after the perturbation (see *Effects of Nonoptimal Reference States* below for an analysis of the effects of choosing suboptimal reference states in this definition of latent pathways). Second, the postperturbation flux distribution is itself dependent on a particular choice for the original optimal state, because both MOMA and ROOM operate by minimizing some distance to a reference flux state.

To test the sensitivity of our results to a particular choice of FBA solutions, we repeated our simulations for numerous combinations of wild-type and optimal mutant flux distributions. We sampled the available FBA-predicted states by fixing the corresponding growth rate (either wild type or optimal mutant) and maximizing or minimizing each of the reaction fluxes allowed to vary under this additional constraint. After choosing a particular pair of optimal flux distributions in this way, we determined the corresponding set of latent pathways according to MOMA and ROOM, and the associated suboptimal growth rates before and after the removal of these latent pathways. Fig. S1 shows the results of this analysis for the *ppk*- and *tpiA*-knockout perturbations. The distributions of growth rate predictions for the strains with (green) and without (blue) latent pathways do not overlap. A less extensive sampling for each of the other 50 knockout mutants considered in our study shows similar behavior. Altogether, these results suggest that our predictions about the growth effect of latent pathway availability are robust with respect to alternate optimal flux distributions.

Regulatory Constraints. All results for the full *Escherichia coli* *i*AF1260 model presented in the main text were computed under the uptake and steady-state constraints listed in the *Materials and Methods* section. These constraints do not take into account regulatory effects, which may limit the set of genes that can be transcribed under the nutrient conditions we consider. Regulatory effects are thus expected to constrain the set of steady-state metabolic states predicted by FBA, MOMA, and ROOM that can actually be realized in vivo (6). To address the possible impact of regulatory constraints on our predictions, we have repeated all calculations in the paper for a modified version of the *i*AF1260 model, following Feist et al. (7) in disabling a set of 152 reactions beforehand. There is evidence that, due to regulatory constraints, these reactions are inactive in the aerobic glucose conditions we simulate. Under these additional constraints, there are 46 single-gene knockouts that change the original flux distribution but nonetheless allow nonzero growth in the resulting knockout mutants according to FBA, which is the same criterion we used to select knockout perturbations in the main text for the unmodified *i*AF1260 model.

Table S2 summarizes the predicted growth impact of disabling latent pathways in the modified *i*AF1260 model. Regardless of the approach used to simulate the response (MOMA, ROOM, or hit-and-run sampling), all 46 knockout mutants in this model show nearly equal (within $\pm 1\%$ of the wild type) or improved growth in the short term following the perturbation when the latent pathways are disabled. Moreover, the average growth increases and the numbers of transiently activated reactions are comparable to those presented in Table 1 of the main text for the unmodified *i*AF1260 model. Our results are therefore not dependent on having the full complement of metabolic reactions in the network available for activation in the organisms' initial response to perturbations. Indeed, the predicted adverse growth effect of latent pathway activation is expected to hold even under additional constraints that may reflect other limitations of the metabolic response in vivo.

Effects of Nonoptimal Reference States. The phenomenological models used in the main text predict a suboptimal metabolic state by minimizing the value of a distance metric with respect to the preperturbation reference state, which is assumed to be growth-maximizing as predicted by FBA. The postperturbation reference state was also assumed to be growth-maximizing after adaptive evolution. These simplifications overlook the possibility that in many natural environments the metabolic state may be nonoptimal even before and long after an external perturbation. This situation could arise, for example, under time-varying conditions, where environmental changes prevent the organisms from approaching optimal growth states. In this case, some of the pathways we have classified as latent may fail to qualify in vivo because they carry nonzero flux in one or both of the reference states.

Geometrically, this scenario corresponds to the appropriate (nonoptimal) reference states lying in the interior of the feasible metabolic solution space, where many more reactions are active, as opposed to the boundary of the space, where reaction activity is limited by irreversibility constraints (8). This situation can be accommodated with a variant of our modeling approach. We have systematically tested the growth impact of reaction *upregulation* relative to nonoptimal reference states, which is the natural generalization of the latent reaction activation considered in our original analysis. Given two optimal states $\boldsymbol{\nu}_1$ (wild type) and $\boldsymbol{\nu}_2$ (evolved mutant) and their associated biomass fluxes, $\nu_{b,1}^{\text{opt}}$ and $\nu_{b,2}^{\text{opt}}$, we now limit the biomass flux to some fraction of these optimal values. This is imposed within our in silico models by the additional constraints $\nu_{b,1} = \lambda\nu_{b,1}^{\text{opt}}$ and $\nu_{b,2} = \lambda\nu_{b,2}^{\text{opt}}$, where $0 < \lambda < 1$. The nonoptimal reference states are then defined by replacing the optimal states before and long after a knockout with the closest feasible metabolic states that satisfy these growth constraints. We performed this analysis for $\lambda = 0.4$ and 0.7.

With respect to the new choices of nonoptimal reference states, the short-term response of the metabolic network to single-gene knockouts still exhibits a transient burst of reaction activity. According to MOMA, the average and standard deviation of the number of fluxes with larger magnitude than in both reference states is 260 ± 83 ($\lambda = 0.4$) and 263 ± 82 ($\lambda = 0.7$). These numbers are comparable to the 291 ± 83 reactions that are transiently activated by MOMA with respect to optimal reference states (main text, Table 1). Fig. S2 shows the MOMA-predicted difference in growth rate when these transiently upregulated fluxes are constrained to not exceed the reference states in magnitude. In all cases, downregulation of the transiently upregulated pathways is predicted to improve growth in the short term following a knockout perturbation, by an average of 6.0% ($\lambda = 0.4$) and 10.4% ($\lambda = 0.7$) of the optimal wild-type growth rate. This analysis is therefore in agreement with the prediction presented in the main text for optimal reference states; namely, that the transient activation (or otherwise upregulation) of latent pathways generally inhibits growth in the short term following a perturbation.

Elementary Mode (EM) Analysis. The results of Fig. 3 (main text) suggest that the primary effect of latent pathway removal is to favor the availability of high-growth metabolic states by preferentially eliminating low-growth states following a genetic perturbation. This is accomplished by eliminating reactions that are silent in the optimal states we consider before and after the knockout, thereby increasing the likelihood that initial metabolic response will activate pathways associated with higher growth. It should be noted, however, that a given reaction can in general be active in many metabolic states, spanning both low- and high-growth phenotypes. It is therefore likely that many high-growth states will by eliminated by disabling latent pathways as well.

To systematically examine this connection between high-growth states and latent pathways, we have used EM analysis, which is an approach for analyzing metabolic networks in terms of interconnected sets of reactions (9). An EM is defined as a unique set of active reactions in the network (represented by a vector $\boldsymbol{\nu}$ in flux space) that (*i*) satisfies the steady-state constraints $\boldsymbol{S} \cdot \boldsymbol{\nu} = 0$, (*ii*) obeys all reversibility constraints (negative entries in $\boldsymbol{\nu}$ must correspond to reversible reactions), and (*iii*) is minimal in the sense that no reaction may be removed from the set while still satisfying (*i*) and (*ii*) (10). Any steady-state flux distribution can be represented as a linear combination of EMs with nonnegative coefficients.

The number of EMs grows combinatorially with the size of the metabolic network, making their calculation computationally infeasible for the full *E. coli* *i*AF1260 model. Therefore, we focused on the subnetwork comprising *E. coli's* central metabolism, which was also used for the volume calculation (main text, *Model-Independent Analysis*). Using the program METATOOL (11), we obtained the full set of 18,656 EMs available on glucose. We have classified each mode as "biomass-producing" or "nonbiomass-producing," based on whether it has a positive or zero entry corresponding to the (irreversible) biomass flux, respectively. It follows that a linear combination of EMs representing a general zero or low-growth metabolic state will be composed primarily of nonbiomass-producing EMs, with only a small aggregate contribution from the EMs that produce biomass. Fig. S3 shows the effect of latent pathway removal on these two types of EMs for the *cyoA* and *lpd* mutants. A significant fraction of the original 18,656 EMs are eliminated by the knockout perturbation (namely, those that involve a disabled reaction). Further EMs will be eliminated when the latent pathways are disabled. As expected from the solution space volume calculation and hit-and-run analysis in the main text (*Model-Independent Analysis* section), the effect of latent pathway removal is to skew the distribution of EMs toward those that produce biomass (Fig. S3*A*). But, surprisingly, a large number of biomass-producing modes are sacrificed as well, and in fact comprise the majority of EMs disabled in this process (Fig. S3 *A* and *B*).

Additional insight comes from analyzing the growth capabilities of the metabolism when one eliminates every reaction that is silent in the corresponding FBA-predicted optimal state of the knockout mutant. This corresponds to an average of $1{,}967 \pm 6$ reactions across all 52 knockout mutants, roughly 7 and 16 times larger than the number of latent pathways removed for MOMA and ROOM, respectively (main text, Table 1). Fig. S4 shows the range of growth rates that can be realized by each mutant in this scenario. For the majority of the mutants, metabolic states with very low growth (<10% of the optimal wild-type growth rate) exist, even when this large set of optimally silent reactions is disabled. This is particularly significant given that, with few exceptions (five according to MOMA, one according to ROOM), the unadapted states of the 52 tested knockout mutants exceed this growth threshold in our models when the latent pathways are enabled.

Taken together, these effects confirm that latent pathways cannot be considered in isolation from biomass production, particularly in optimal growth states.

Comparison with *i*JR904 *E. coli* Model. We have repeated our calculations for the extensively curated *i*JR904 reconstruction of the *E. coli* metabolic network (12), which has been previously analyzed in great detail in connection with synthetic rescue interactions. This network consists of 931 reactions, 904 enzyme-coding genes, 618 metabolites, 143 exchange fluxes, and the biomass flux. Within this model, there are 36 single-gene knockout strains that are compatible with growth but for which the original growth-maximizing metabolic state becomes infeasible after the knockout. For this set of genes, the average and standard deviation changes in growth rate are +12.0 (15.8)% (MOMA) and +1.1 (3.1)% (ROOM) and +65.6 (11.4)% (random) for the removal of the latent reactions associated with the individual knockout perturbations. The results are therefore consistent with those presented in Table 1 (main text) for the *i*AF1260 model.

1. Edwards JS, Palsson BØ (2000) The *Escherichia coli* MG1655 *in silico* metabolic genotype: Its definition, characteristics, and capabilities. *Proc Natl Acad Sci USA* 97:5528–5533.
2. Segre D, Vitkup D, Church GM (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci USA* 99:15112–15117.
3. Shlomi T, Berkman O, Ruppin E (2005) Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc Natl Acad Sci USA* 102:7695–7700.
4. Mahadevan R, Schilling CH (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab Eng* 5:264–276.
5. Reed JL, Palsson BØ (2004) Genome-scale *in silico* models of *E. coli* have multiple equivalent phenotypic states: Assessment of correlated reaction subsets that comprise network states. *Genome Res* 14:1797–1805.
6. Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BØ (2004) Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 429:92–96.
7. Feist AM, et al. (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3:121.
8. Nishikawa T, Gulbahce N, Motter AE (2008) Spontaneous reaction silencing in metabolic optimization. *PLoS Comput Biol* 4:e1000236.
9. Trinh CT, Wlaschin A, Sriene F (2009) Elementary mode analysis: A useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl Microbiol Biotechnol* 81:813–826.
10. Schuster S, Dandekar T, Fell DA (1999) Detection of elementary flux modes in biochemical networks: A promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol* 17:53–60.
11. von Kamp A, Schuster S (2006) Metatool 5.0: Fast and flexible elementary modes analysis. *Bioinformatics* 22:1930–1931.
12. Reed JL, Vo TD, Schilling CH, Palsson BØ (2003) An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol* 4:R54.

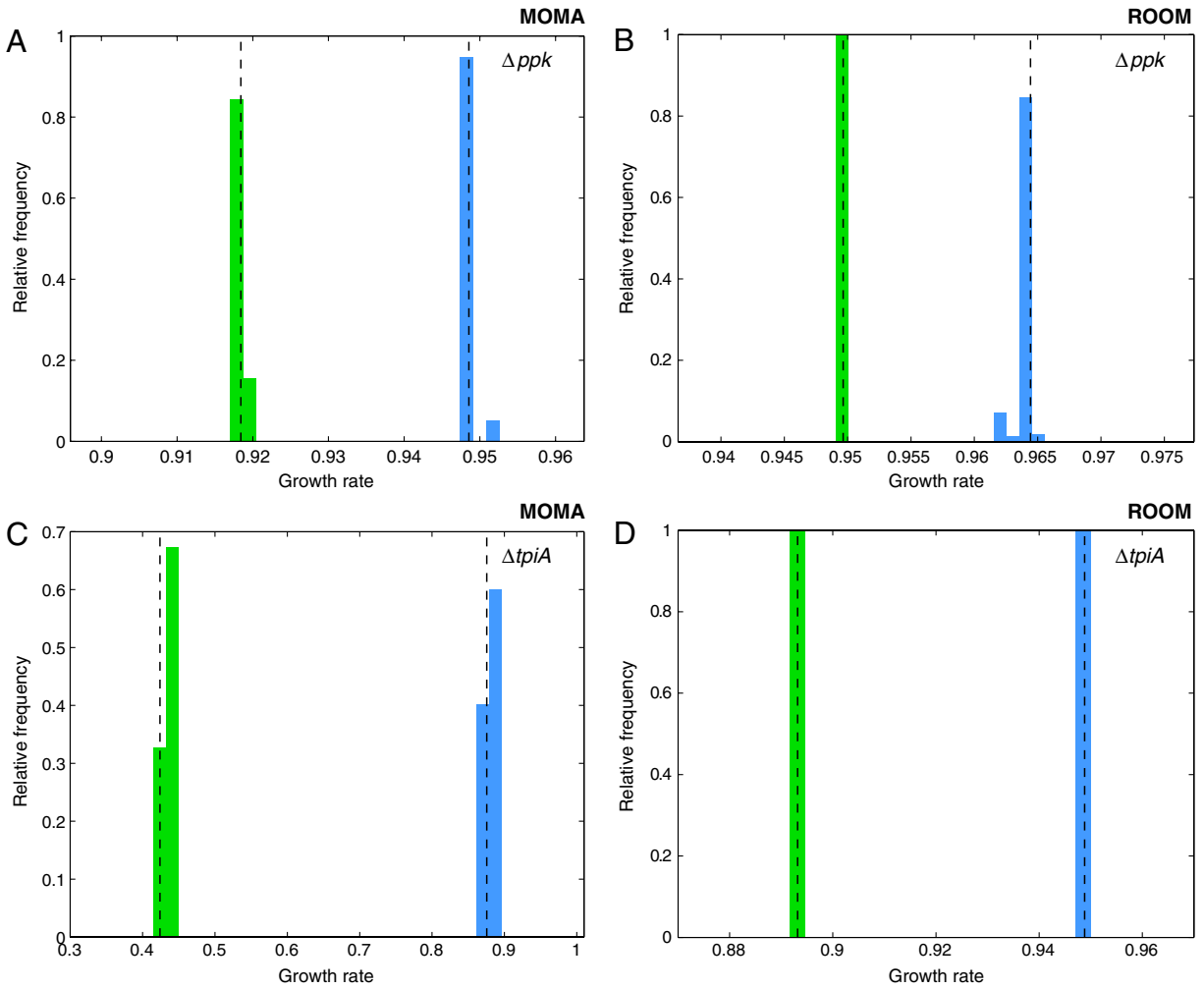

Fig. S1. Impact of alternate optima. (*A* and *B*) Sensitivity to alternate optima for the *ppk*-knockout perturbation according to MOMA (*A*) and ROOM (*B*) (based on 12,760 different pairs of optimal states). (*C* and *D*) Sensitivity to alternate optima for the *tpiA*-knockout perturbation according to MOMA (*C*) and ROOM (*D*) (based on 20,880 different pairs of optimal states). Each panel shows the probability distributions of the predicted early post-perturbation growth rates for strains in which the latent pathways are available (green) and disabled (blue). All growth rates are normalized by the optimal growth rate of the wild type. Dashed lines indicate the growth rate predictions presented in the results of the main text. In all cases, the growth rates are tightly clustered around these values, and the two distributions do not overlap.

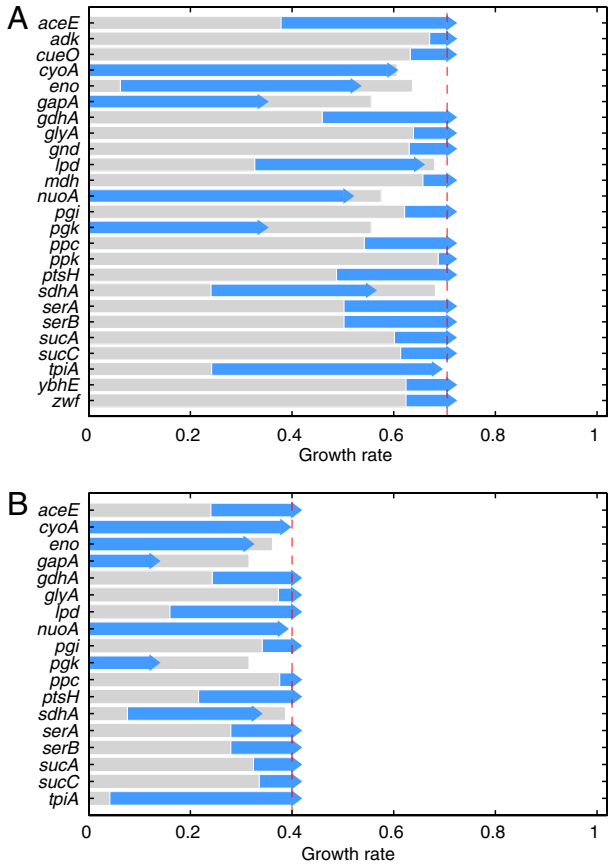

Fig. S2. Impact of nonoptimal reference states. (*A* and *B*) Effect of latent pathway upregulation using reference metabolic states (wild type and "adapted" mutant) whose growth rates are limited to 70% (*A*) and 40% (*B*) of the respective optima. For each knockout, the panels show the change in the MOMA-predicted growth rate after the perturbation when we limit the magnitude of the flux through each latent reaction to the corresponding highest value between the two reference states. The dotted lines and gray bars indicate the maximum allowed growth rates for the wild type and knockout mutants, respectively. All values are normalized as a fraction of the optimal wild-type growth rate. We only display cases that show a significant difference in growth rate (>1% of the wild-type optimum) when the latent pathways are downregulated.

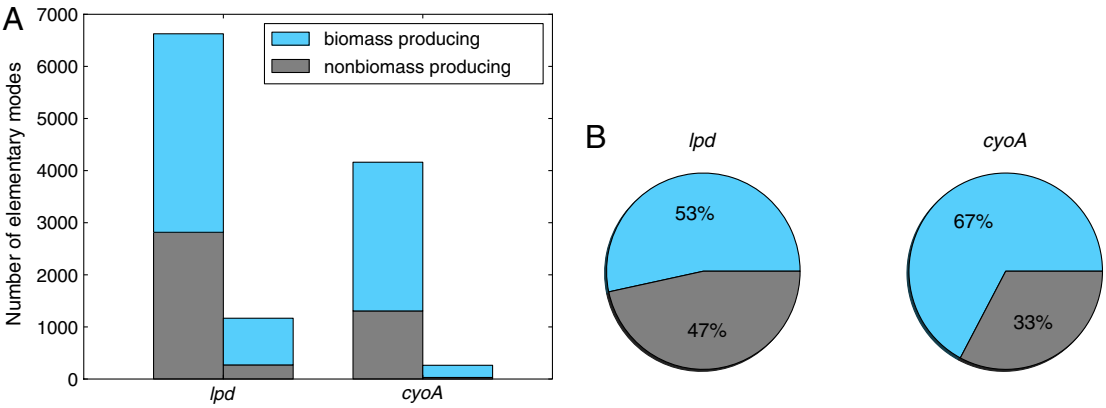

Fig. S3. Effect of latent pathway removal on elementary modes of *E. coli*'s central metabolism. These simulations are based on the *lpd*- and *cyoA* -knockout mutants. (*A*) Number of elementary modes having a zero (gray) and positive (blue) biomass component, before (*Left* bar, each mutant) and after (*Right* bar, each mutant) the removal of latent pathways predicted by MOMA. (*B*) Percentage of elementary modes of each type disabled by latent pathway removal. Latent pathway removal renders a large number of elementary modes unavailable for use. Strikingly, the majority of the eliminated modes involve biomass production.

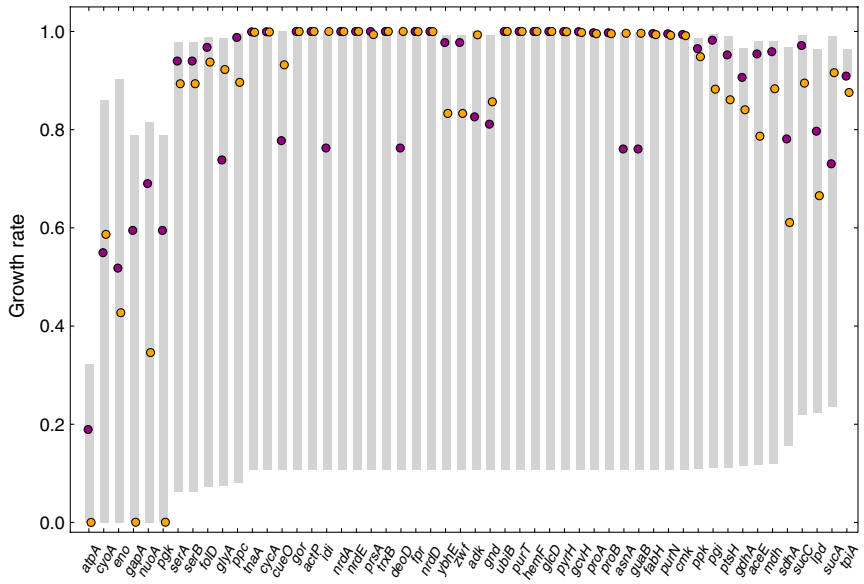

Fig. S4. Range of possible growth rates when all optimally inactive reactions are disabled. The gray bars correspond to the range of realizable growth rates in each of the 52 knockout mutants when all reactions not active in the corresponding growth-maximizing state of the mutant are removed. Purple circles correspond to the MOMA-predicted growth rate in response to each knockout in this case. Orange circles correspond to the MOMA-predicted growth when only the latent pathways are removed, as presented in the main text. Although this extreme pathway removal is also predicted to increase growth following a knockout perturbation, low-growth states remain available for the large majority of mutants. All growth rates are normalized by that of the optimal wild type.

Table S1. Comparison between linear ROOM and integer ROOM

	Integer ROOM		
	+	=	–
Linear ROOM			
+	12	0	0
=	12	27	1
–	0	0	0

The rows and columns indicate the number of single-gene knockout perturbations (out of 52 total) for which the early post-perturbation growth-rate is larger (+), smaller (–), or negligibly different (=) when the latent pathways are disabled beforehand compared to the case in which they are available. As in the main text, we define a change to be negligible if its magnitude is less than 1% of the wild-type growth rate. The majority of cases are on the diagonal, indicating that the two methods predict the same growth impact of latent pathway removal. There are no mutants for which the two methods predict a significant change in opposite directions.

Table S2. Summary of the predicted impact of latent pathways under nutrient-specific regulatory constraints

	MOMA	ROOM	Random
Latent reactions for individual perturbations:			
All knockout perturbations	+10.5 (13.2)%	+1.5 (3.8)%	+61.0 (9.4)%
Significant differences*	+18.4 (12.8)%	+5.5 (5.9)%	+61.0 (9.4)%
Number of reactions removed	267 (82)	96 (45)	902 (5)
Simultaneously nonessential latent reactions:			
All knockout perturbations	+8.8 (11.2)%	+1.5 (3.8)%	
Significant differences†	+16.7 (10.4)%	+5.8 (5.9)%	
Number of reactions removed	235 (76)	82 (38)	

Corresponds to Table 1 of the main text for 46 knockout mutants in a modified *i*AF1260 *E. coli* model in which 152 reactions have been disabled to account for known regulatory constraints in aerobic glucose medium conditions.
*By more than 1% of the wild-type growth rate: 57% (MOMA), 26% (ROOM), and 100% (random) of the perturbations.
†By more than 1% of the wild-type growth rate: 52% (MOMA) and 24% (ROOM) of the perturbations.