Closing the Flux Cone: imposition of maximal capacities

Bernhard Palsson

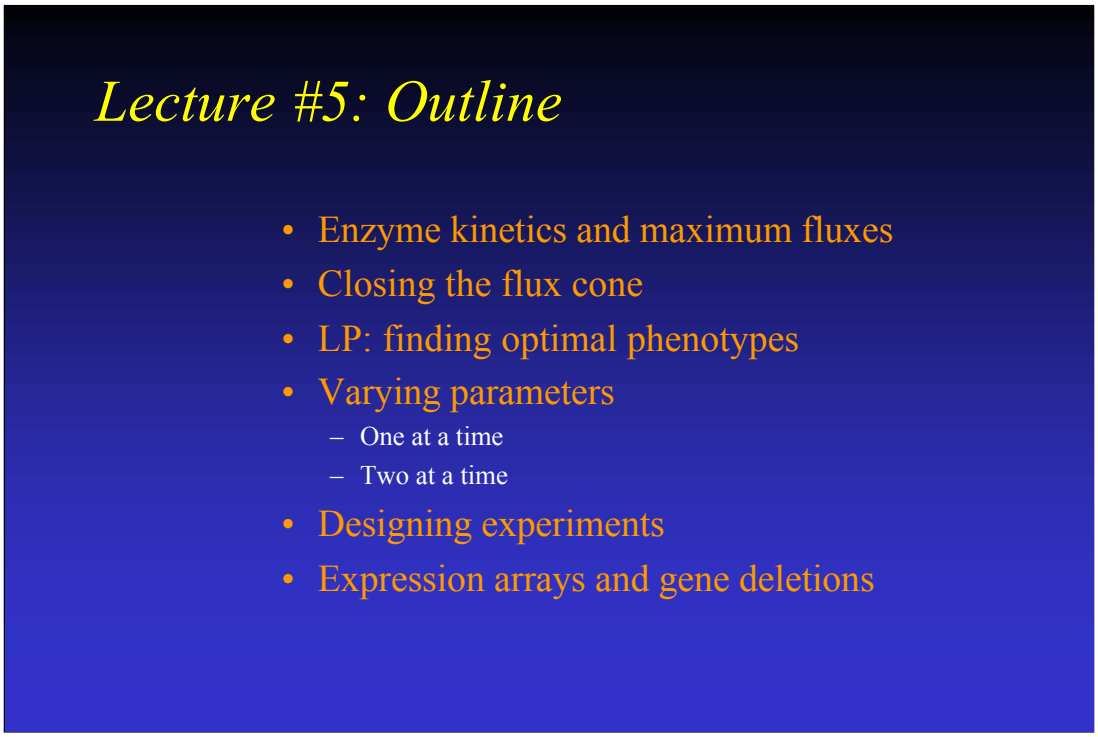
Hougen Lecture #5

Nov 21th, 2000

INTRODUCTION

In the previous lecture we looked at the combined stoichiometric and thermodynamic constraints that cells must obey. These led to the formation of a conically shaped solution space--called the flux cone. The edges are vectors that in a positive linear combination span the cone. These edges where shown to be extreme pathways. The flux through these pathways is limited by a maximum value.

Such maximal constraints close the solution space. In this lecture we explore the characteristics of the closed space.



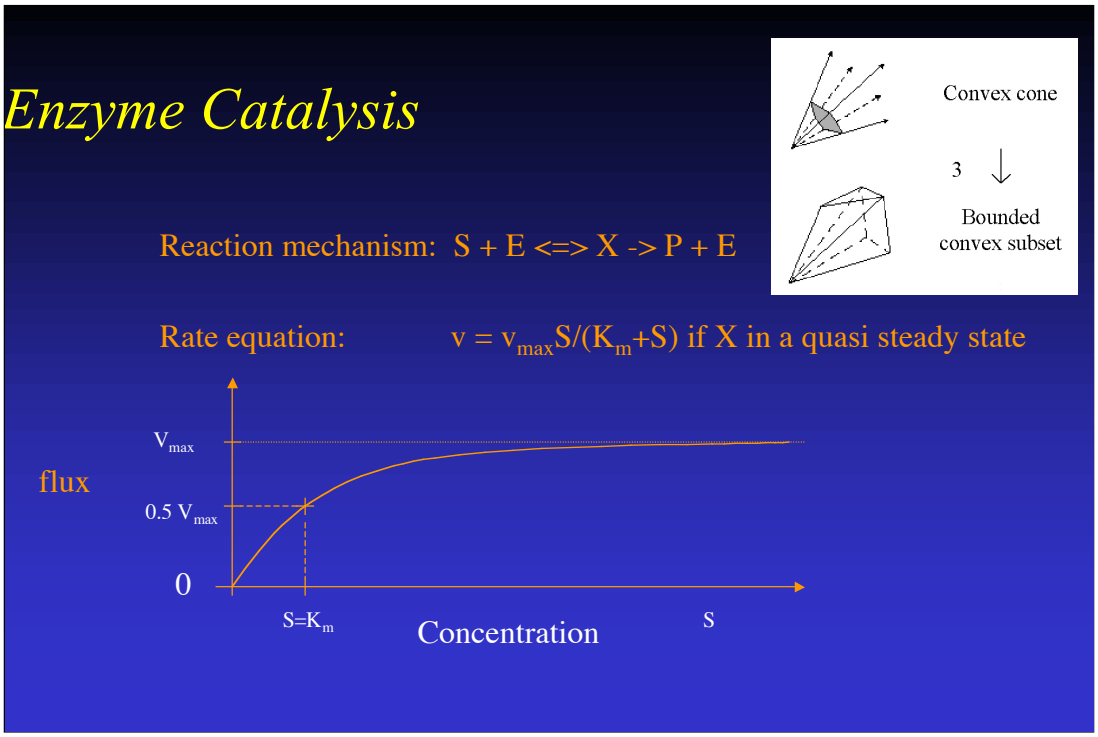
LECTURE #5

This lecture begins with an introduction to the origin of the maximal fluxes that are achievable through an enzymatic reaction and how these limitations cap off and close the flux cone. Although there are infinitely many possible flux distributions found within this closed solutions space, if an objective is stated one can find the 'best' solution by that criteria within this solution space. Linear programming or optimization is used to find this solution. The optimal solution will always lie at the edge of the cone or on one of its surfaces.

A single solution is rarely of interest. We thus explore the optimal solution as a function of an environmentally varying parameter. There are 'kinks' found in the piece-wise linear solutions. At these discontinuities we discover that the shadow price structure of the basal solution changes. These changes will thus correspond to a change in the phenotype. Thus there are a limited number (a discrete number) of phenotypes found within the solution space.

We then explore the simultaneous variation of two environmental variables and introduce the concept of a phase plane. These phase planes can then be used to design insightful experiments.

Finally, we show how flux-balance analysis can be used to interpret and predict the consequences of gene deletions and metabolic shifts as measured by expression arrays.



CONSTRAINTS ON METABOLIC FLUXES

Linear spaces are characterized by a basis set where any linear combination of the basis vectors is found in the space, i.e.;

$$\mathbf{v} = \Sigma_i w_i \mathbf{p_i}$$

Where $\mathbf{p_i}$ are the conical basis vectors, as introduced in the last lecture. The weights, w_i, used to multiply the basis vectors in the summation are positive.

Since the individual reaction steps (v_i) in a pathway vector are carried out by an enzyme there are limitations placed on the numerical values that w_i can take in a real system:

•Minimum: the reactions are irreversible, thus the weights are positive

•Maximum: there is maximum flux through an enzymatic reaction, thus there are maximum weights; thus

$$0 < w_i < w_{max}$$

Since a pathway vector is comprised of a series of individual reactions, the step with the lowest capacity will limit the flux through a linear pathway.

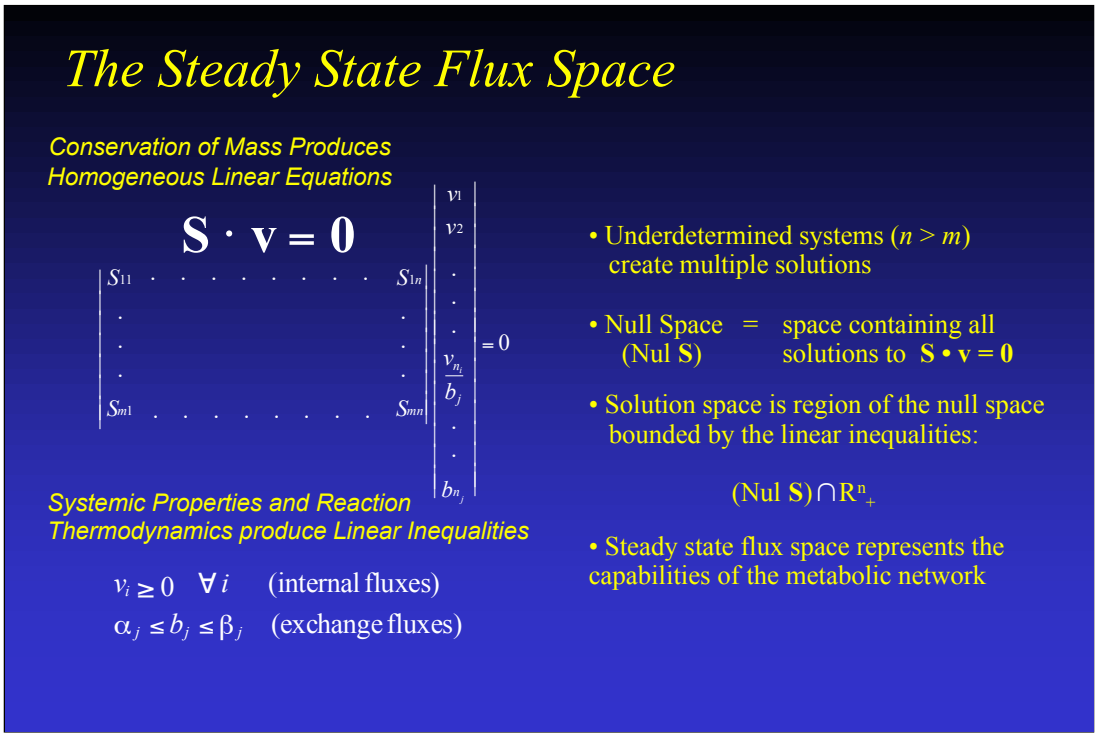
If a reaction is reversible we will write each direction as a separate irreversible reaction.

Estimation of maximal fluxes

- Using typical numerical values for:
  - concentrations for enzymes (4μM) and
  - metabolites (100μM), and
  - theoretical maximal bimolecular association rate constants and
  - data on enzyme turnover numbers, we estimate that:

  V_{max} to be one million molecules per cubic micron per second

- The maximal measured fluxes are about half that value



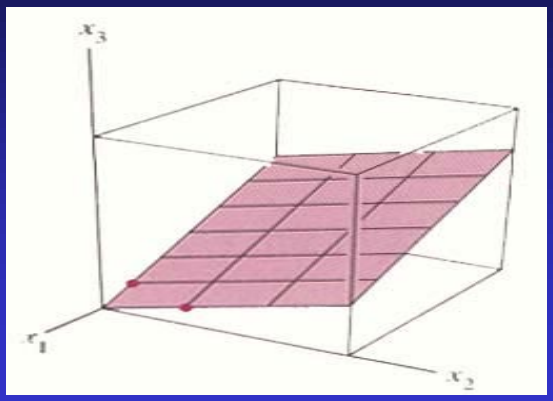
THE CONSTRAINED FLUX SOLUTION SPACE

With the stoichiometric matrix constructed, how do we determine metabolic pathways and analyze them?

As we have seen in previous lectures, the principles of conservation of mass produce a system of homogeneous linear equations, **Sv = 0**. Additionally there are constraints placed on the direction of flow under which each flux can operate creating a set of linear inequalities, $0 < w_i < w_{max}$.

This defines our conditions which in most cases creates an underdetermined system. This means that there are more fluxes operating within the system than there are metabolites which leads to multiple solutions or flux distributions which satisfy all of the stoichiometric constraints, and all the capacity constraints.

A geometric representation of the null space and constraints imposed through inequalities: it is the intersection of the null space and the positive orthant in the n-dimensional space: (Nul **S**) $\cap \mathbf{R}^n_+ \cap \mathbf{V}_{max}$

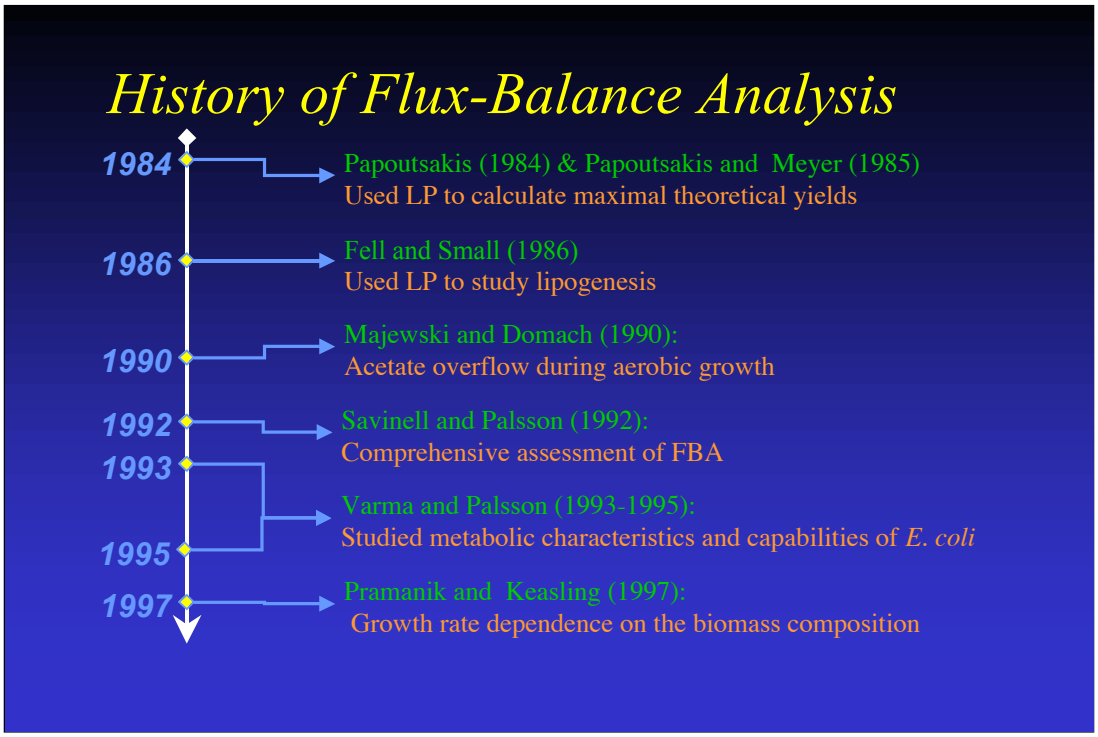


THE CONFINED SOLUTION SPACE AS AN INTERSECTION

Nul S $\cap$ $\mathbf{R}^n_+ \cap \mathbf{V}_{max}$

In linear algebra the term null space is used to describe the space which contains all of the solutions to a system of homogeneous linear equations. The solution space of interest to us is actually the intersection of this null space with the region bounded by the inequalities placed on the weights. This space represents and defines the boundaries and capabilities of a metabolic genotype describing all of the possible flux distributions and routes which can theoretically operate through the system, clearly defining what an organism can and cannot do.

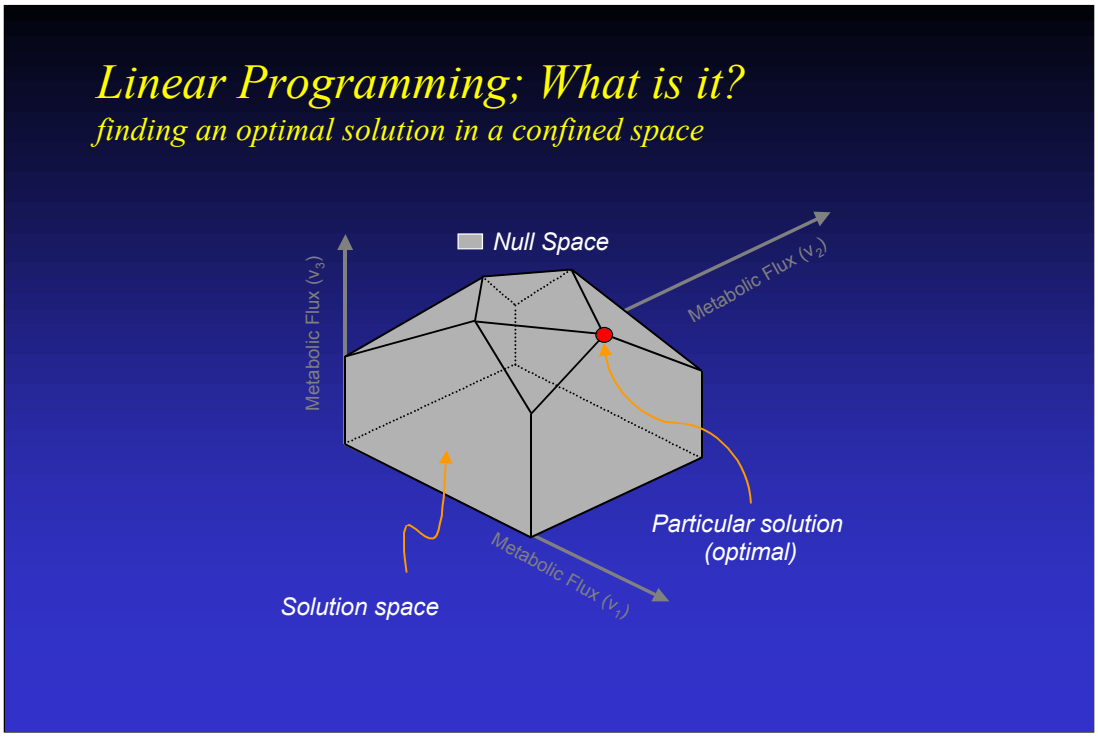
In the solution space we can find the answers to any and all of our questions which pertain to the structure and production capabilities of an organism.



SOME HISTORICAL EVENTS IN THE DEVELOPMENT OF FBA

This slides shows some of the historical events in the development of FBA of under-determined systems. A detailed historical review is found in:

Edwards, et al *Metabolic flux balance analysis* in Metabolic Engineering, Lee and Papoutsakis Editors



LP: What is it?

This diagram depicts a bounded polytope in 3 dimensions. Imagine that it is the space of possible solutions to a set of linear equalities with constraints, such as the flux balance equations and the capacity constraints. Each point in this space satisfies these conditions. However, the nature of the solutions differs. We can choose a particular solution in this space that is the 'best' in some sense.

This idea underlies LP. We state an objective function that measures what we are interested in. Then we try to find the best value for this objective function under the given constraints. The best value normally means the maximum value. Minimization can be performed by simply finding the maximum of the negative of the objective function.

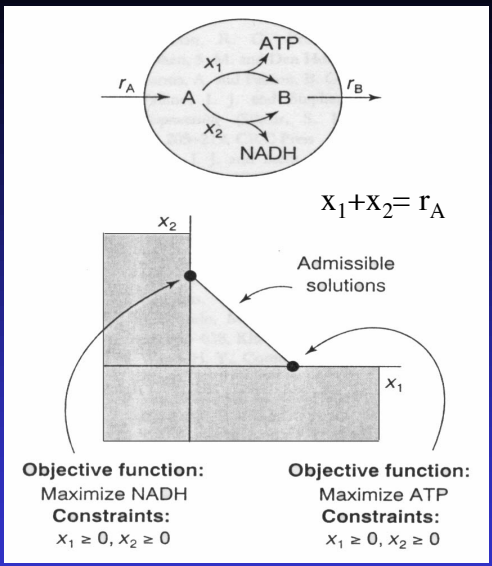
The optimal solution normally lies in a corner of the polytope. Occasionally the objective function has the same value along a whole edge and all the points on that edge are optimal values. In this rare case the objective function is 'parallel' to the edge of the polytope.

How does LP work? A very simple example

The solution space is the line of admissible in the positive orthant.

If we maximize ATP production the solution lies on the x-axis where all the flux would be through reaction x_1. Conversely, maximizing NADH production would give the point at the y-axis, where only reaction x_2 is active.

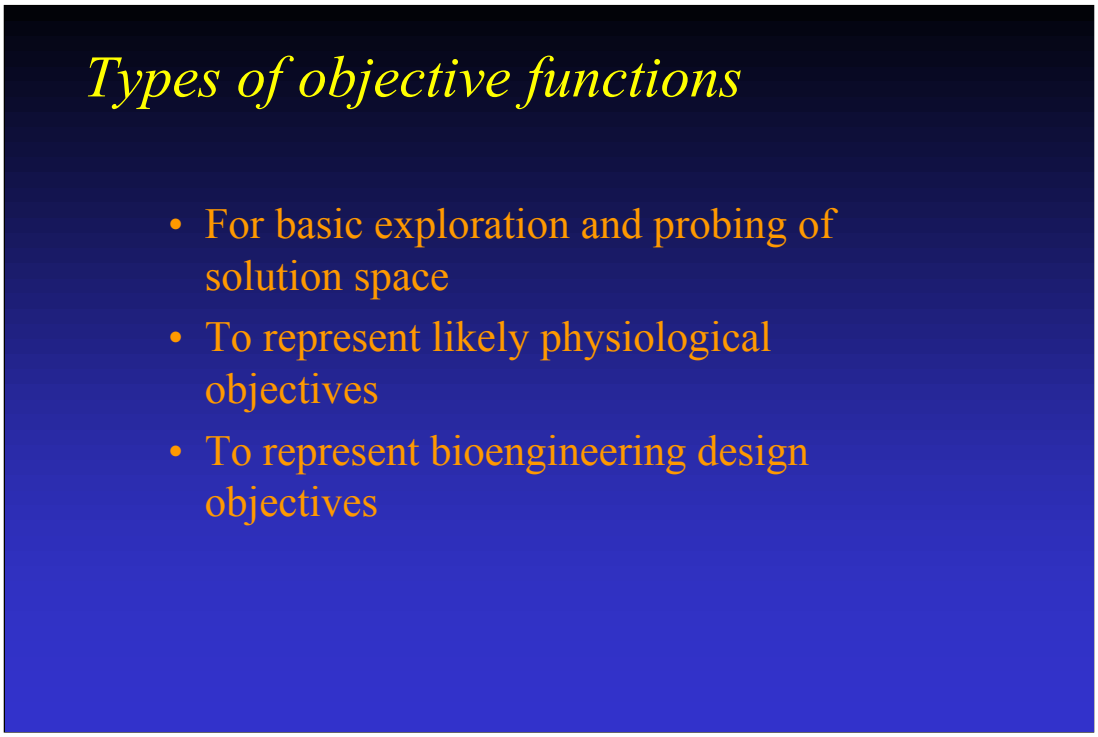
Note that the optimal solutions lie at the boundary of the admissible space.



Bonarius,et al TIBTECH vol 15:308 (1997)

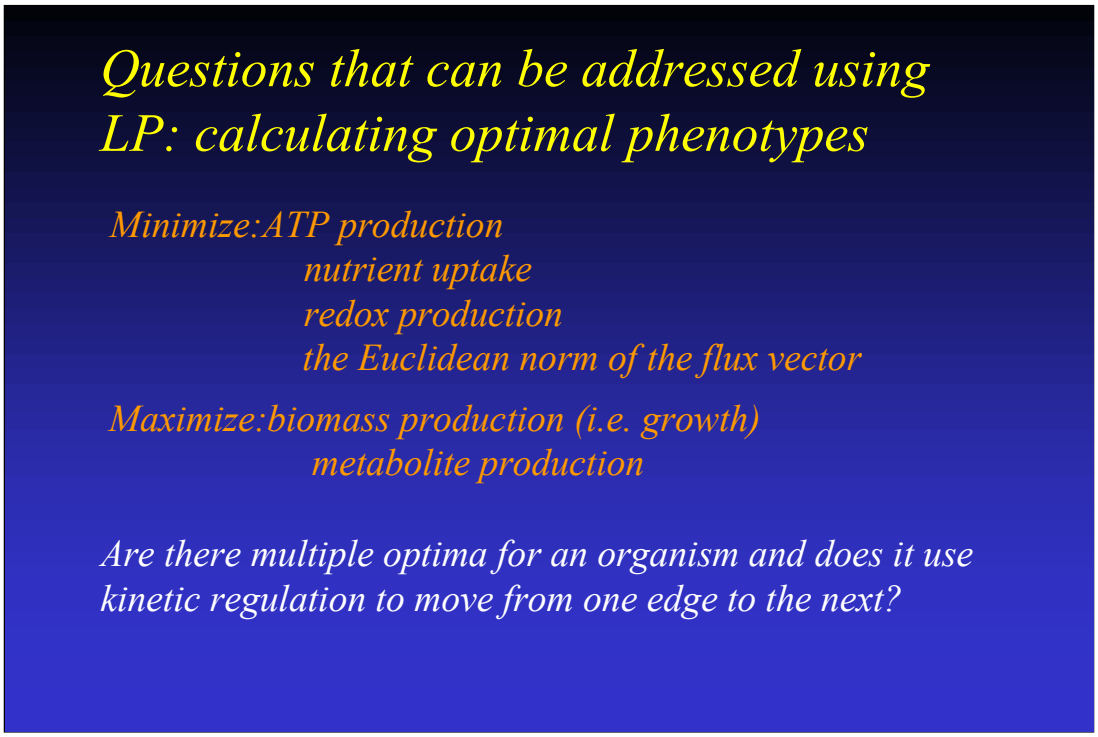
This readily understandable example shows a space of admissible solutions and the optimal phenotypes lying at the edges of this space.

Q: What happens if you optimize x_1+x_2 ?



The Objective Function

Within the solution space defined by the connectivity and capacity constraints, we can search for the best solution using linear optimization. What we search for is determined by the objective function stated. There are several types of objective functions that can be used. First, we can use objective functions to explore the properties of the solution space, and the capabilities of an organism. These objective functions include things like maximizing the ATP from a given substrate, or maximizing the amount of an amino acid produced from a given substrate. These types of objective functions are non-physiological, but can be used to probe the properties of a network. A second class of objective functions would represent objectives that we believe are physiologically relevant. For microbial cells, the belief is that they maximize their growth rate given the constraints under which they operate. In this case, and as shown, below the objective is the balanced exit from the network of all the precursors needed for the synthesis of the cellular mass. The third type of objective function may relate to an intentional engineering objective of a metabolic system. We may wish to maximize a product like Lysine, for instance, and try to figure out what the best flux maps are that lead to the production of Lysine. We can add or delete reactions from the network to determine how those changes affect the yield of the desired product.



OPTIMAL PHENOTYPES

A number of different objective functions have been used for metabolic analysis, these include

Minimize ATP production: This objective is stated to determine conditions of optimal metabolic energy efficiency.

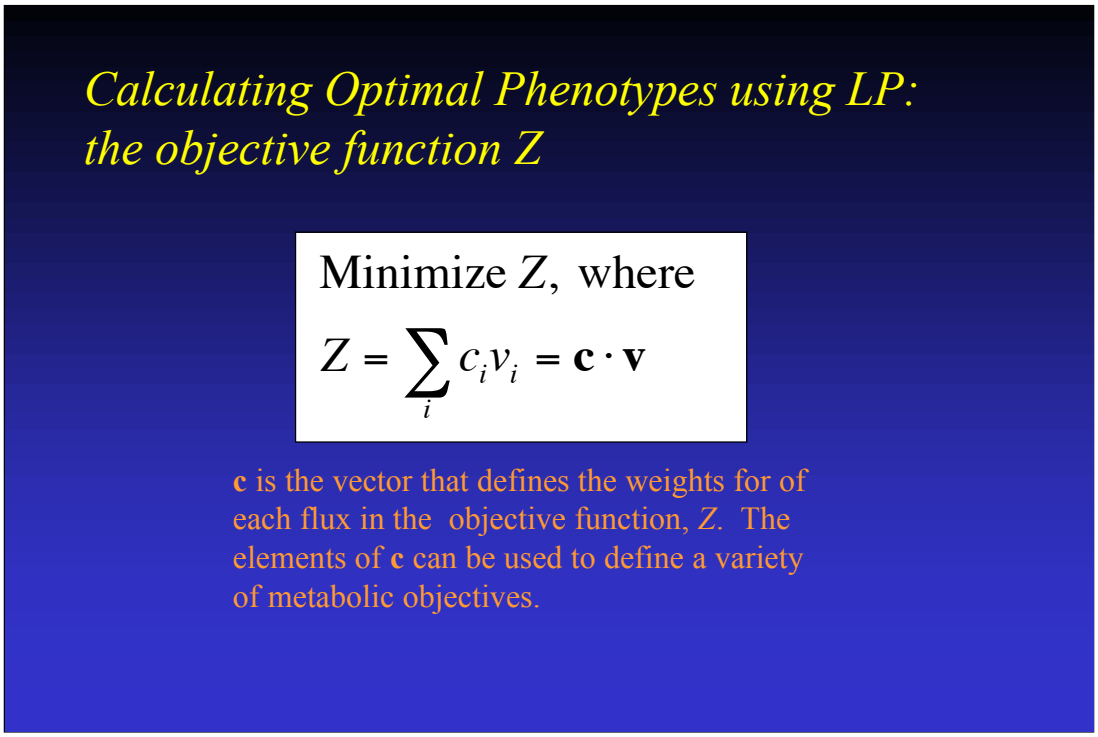
Minimize nutrient uptake: This objective function is used to determine the conditions under which the cell will perform its metabolic functions while consuming the minimum amount of available nutrients.

Minimize redox production: This objective function finds conditions where the cells operate to generate the minimum amount of redox potential.

Minimize the Euclidean norm: This objective has been applied to satisfy the strategy of a cell to minimize the sum of the flux values, or to channel the metabolites as efficiently as possible through the metabolic pathways.

Maximize metabolite production: This objective function has been used to determine the biochemical production capabilities of *Escherichia coli*. In this analysis the objective function was defined to maximize the production of a chosen metabolite (i.e. lysine or phenylalanine).

Maximize biomass and metabolite production: By weighing these two conflicting objectives appropriately, one can explore the tradeoff between cell growth and forced metabolite production in a producing strain.

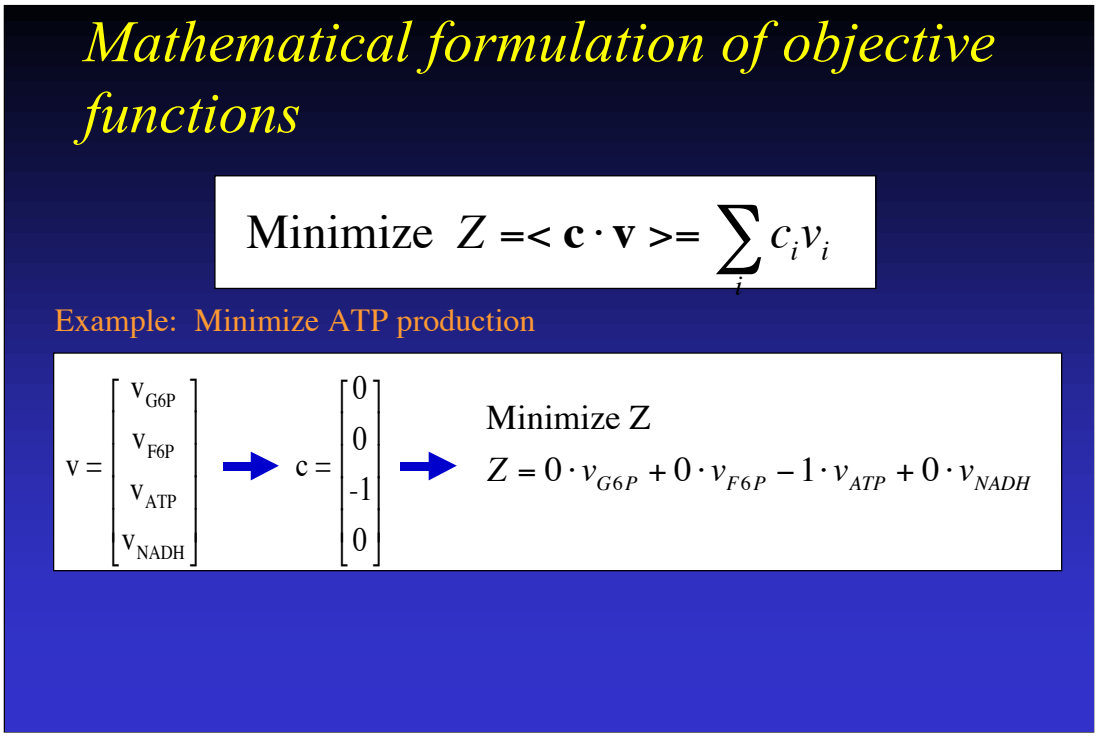


THE OBJECTIVE FUNCTION

Numerous questions about metabolic capabilities can be answered using LP. The stoichiometric and capacity constraints define a range of allowable behavior. We can then find the best value within these constraints. Biologically, we have defined the space of all phenotypes (that is particular solutions) that can be derived from a genotype. We can calculate the best phenotype from a particular standpoint. For instance we can calculate the maximum number of ATP molecules that can be generated from a particular substrate.

The next slide lists a number of important phenotypic behaviors that can be calculated using LP. The maximum growth function is perhaps the one of greatest interests from an evolutionary standpoint.

This general representation of *Z* enables the formulation of a number of diverse objectives. These objectives can be design objectives for a strain, exploitation of the metabolic capabilities of a genotype, or physiologically meaningful objective functions, such as maximum cellular growth.



MATHEMATICAL FORMULATION OF OBJECTIVE FUNCTIONS

This slide illustrates the formation of the objective function using a simple example. In the example there are 4 metabolite fluxes. The objective is to minimize ATP production therefore the **c** matrix has a zero "weight" on all fluxes except v_{ATP} which has a -1 The coefficient on the ATP flux is negative since it is being minimized.

The growth requirements

Metabolic demands of precursors and cofactors required for 1 g of biomass of *E. coli.*

These precursors are removed from the metabolic network in the corresponding ratios.

Thus, the objective function is:

$$Z = 41.2570\, v_{ATP} - 3.547 v_{NADH} + 18.225 v_{NADPH} + \ldots.$$

Metabolite	Demand (mmol)
ATP	41.2570
NADH	-3.5470
NADPH	18.2250
G6P	0.2050
F6P	0.0709
R5P	0.8977
E4P	0.3610
T3P	0.1290
3PG	1.4960
PEP	0.5191
PYR	2.8328
AcCoA	3.7478
OAA	1.7867
AKG	1.0789

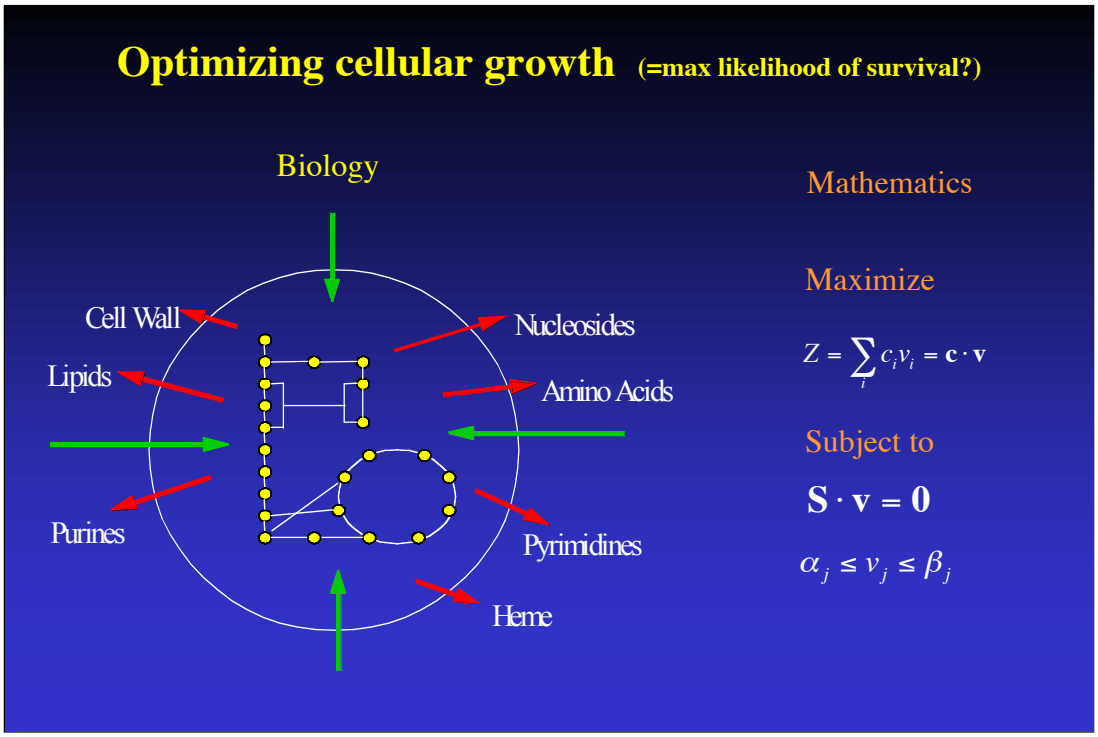
THE GROWTH FUNCTION

This table shows the requirements for making one gram of *E. coli*. This means that for the cell to grow all these components must be provided in these amounts. Thus, a balanced set of metabolic demands makes up the growth objective function:

$$Z = 41.257 v_{ATP} - 3.547 v_{NADH} + 18.225 v_{NADPH} + 0.205 v_{G6P} + 0.0709 v_{F6P} + 0.8977 v_{R5P} + 0.361 v_{E4P} + 0.129 v_{T3P} + 1.496 v_{3PG} + 0.5191 v_{PEP} + 2.8328 v_{PYR} + 3.7478 v_{AcCoA} + 1.7867 v_{OAA} + 1.0789 v_{AKG}$$

The biomass composition thus serves to define the weight vector **c.**

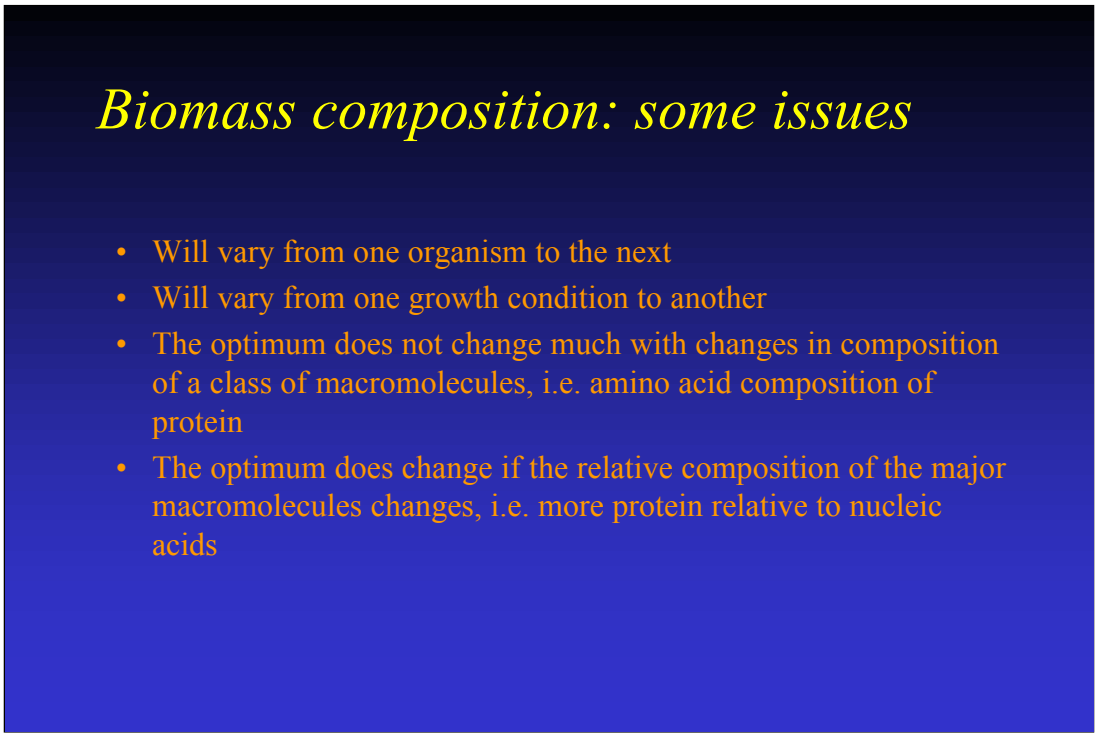
The full growth function for *E. coli* is more complicated than the one given above, since various maintenance functions need to be considered.



THE MAXIMAZATION OF BIOMASS FORMATION

This slide shows schematically on the left the idea of maximizing biomass formation. There can be one or more inputs (the green arrows) and a balanced (linked) output that corresponds to the biomass composition.

On the right we show the mathematical formulation of the problem. We wish to maximize the objective function under the stated constraints. These constraints form a closed cone as explained earlier.



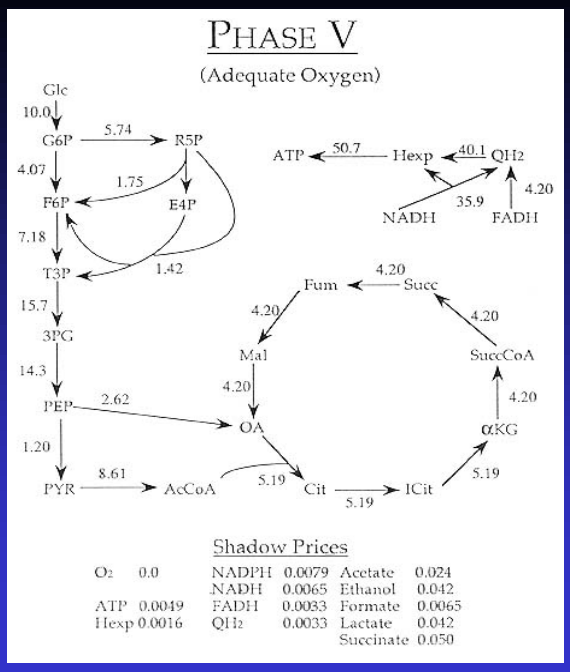
Biomass Composition

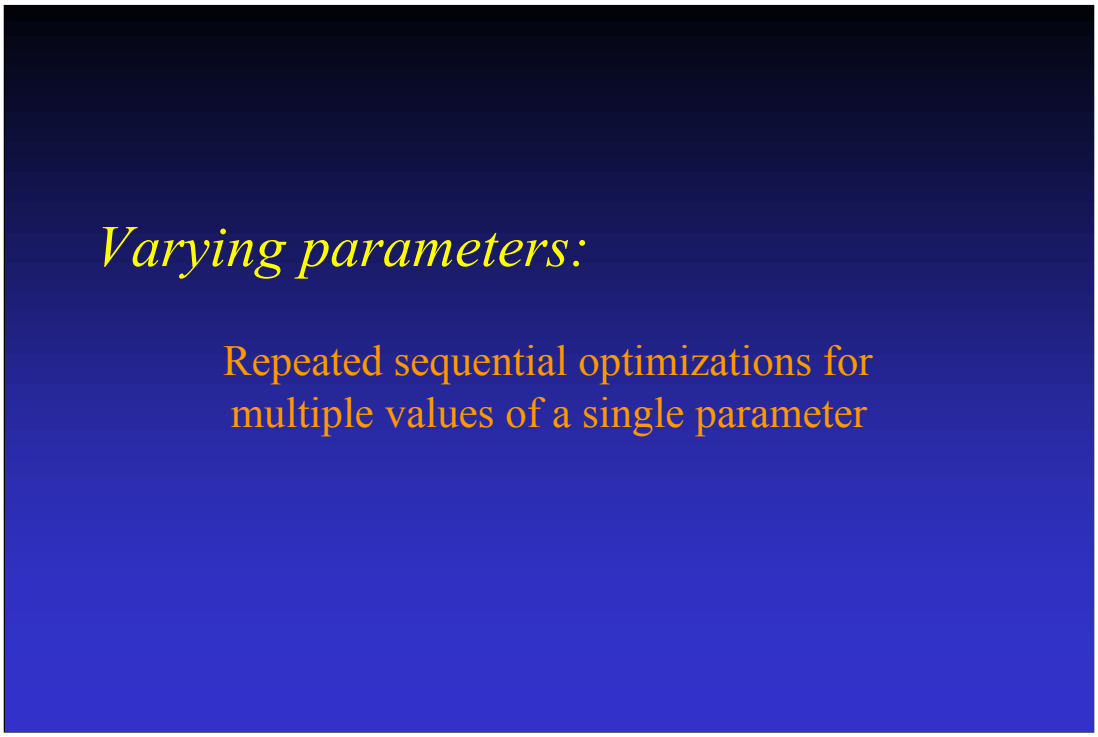
The physiologically interesting objective that we wish to study throughout these notes is the maximization of biomass yield. The definition of the solution space has few ambiguities associated with it, but the statement of the objectives has a few uncertainties built into it. First, the biomass composition is variable. It is different from one organism to another. It varies from one growth condition to another, and both of those may potentially be important issues and change the predicted optimum behavior. Legacy databases of biomass composition are needed.

The limited calculations that have been performed show that the optimum solutions do not change significantly with the monomeric composition of the major macromolecules. For instance, if the Valine to Alanine ratio is varied in the protein of a cell, the optimal growth rate does not significantly change. Conversely, if the protein relative to lipid composition in a cell changes, the optimum solution tends to be affected.

As will be shown, one can invert this problem and look at an edge of the solution space and then calculate all the objective functions that are maximized under those conditions. This might give better insight into the objectives that cells are trying to accomplish.

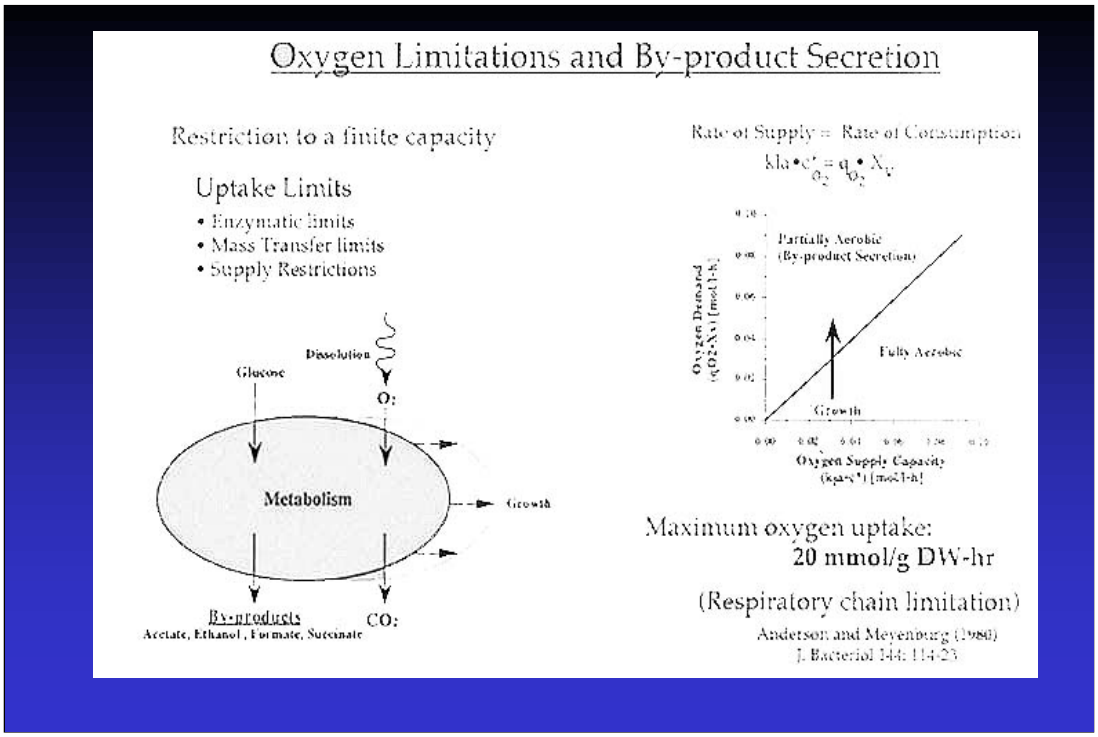
The solution displayed as a flux map: example, aerobic growth on glucose





PARAMETER VARIATION

We looked at one optimal flux map for different substrates and for constraints on several internal fluxes. These are calculations for a discrete set of conditions. We may however be interested in the a range of numerical values for a particular parameter. Thus, we can calculate a series of optimal solutions for small incremental changes in a parameter in the system. If the increments are small enough, we effectively get a continuous variation in the parameter of interest.



EXAMPLE: REDUCING OXYGEN AVAILABILITY

When cells grow in the laboratory with an abundance of substrate they grow into high densities eventually outstripping the ability for oxygen to be supplied rapidly enough to support fully aerobic growth. As oxygen becomes limiting, the cells must partially oxidize their substrate and secrete a metabolic by-product.

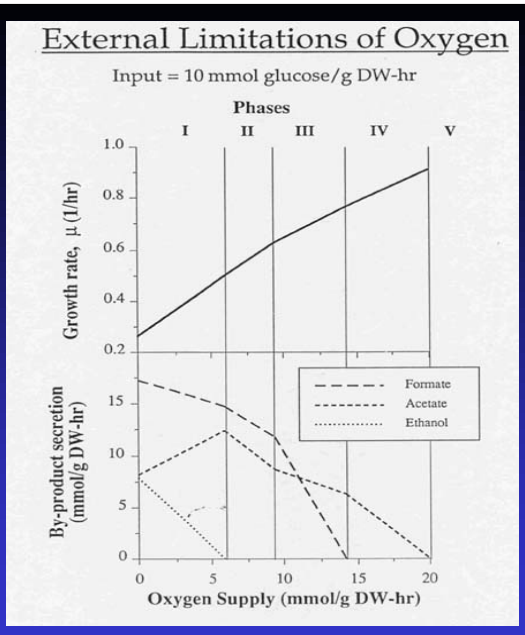
The panel on the left illustrates this problem at the cellular level. On the right this problem is illustrated from a bioprocess viewpoint.

The following slides were prepared with a reduced E. coli model in 1993 (Varma, A&EM), but it illustrates how parameter variations can be used to study problems of fundamental physiological relevance, and those that are of practical importance.

Example:

In this example we vary the maximum allowable uptake rate of oxygen. The whole range of oxygenation is shown, from fully aerobic conditions to fully anaerobic conditions.

The growth rate is graphed in the upper panel and the by-product secretion rates in the lower.



anaerobic ⟶ aerobic

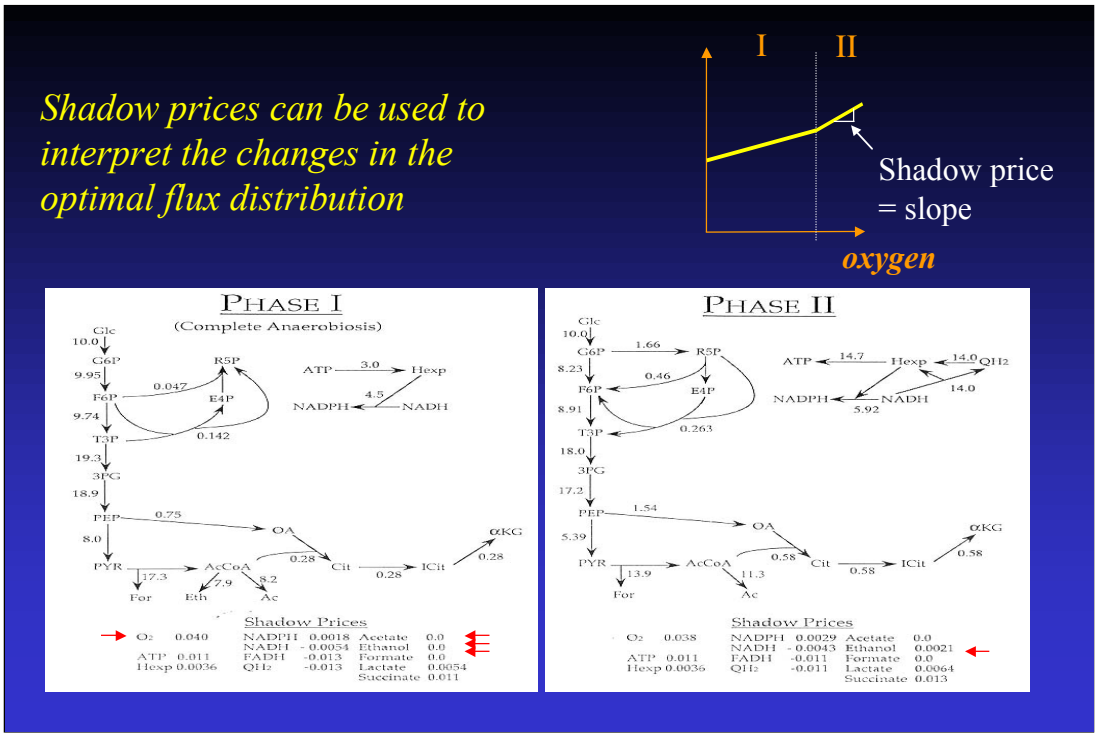
VARYING OXYGEN AVAILABILITY

As the dissolution of oxygen cannot keep up with the high volumetric consumption rates at high cell density, the amount available per cell is reduced. Computationally this is represented by lowering the capacity constraints on the oxygen uptake rate.

The results from a series of LP calculations with varying b_{O2} is shown in this slide. The optimal growth rate drops as the oxygen uptake rate is reduced, as shown in the upper panel. It does so in piece-wise linear fashion where changes in the slope occur at well defined oxygen uptake rates. This feature naturally divided the range of oxygen uptake rates into distinct phases.

The lower panels shows the secretion rates of metabolic by-products; formate, ethanol and acetate. Each one of these by-products is secreted in a fundamentally different way in each phase. As oxygen is reduced, incomplete oxidation of glucose takes place and metabolic by-products are secreted; acetate is first secreted, then formate followed by ethanol.

The LP solution in each phase is fundamentally different and the transition from one to another can be interpreted using shadow prices.

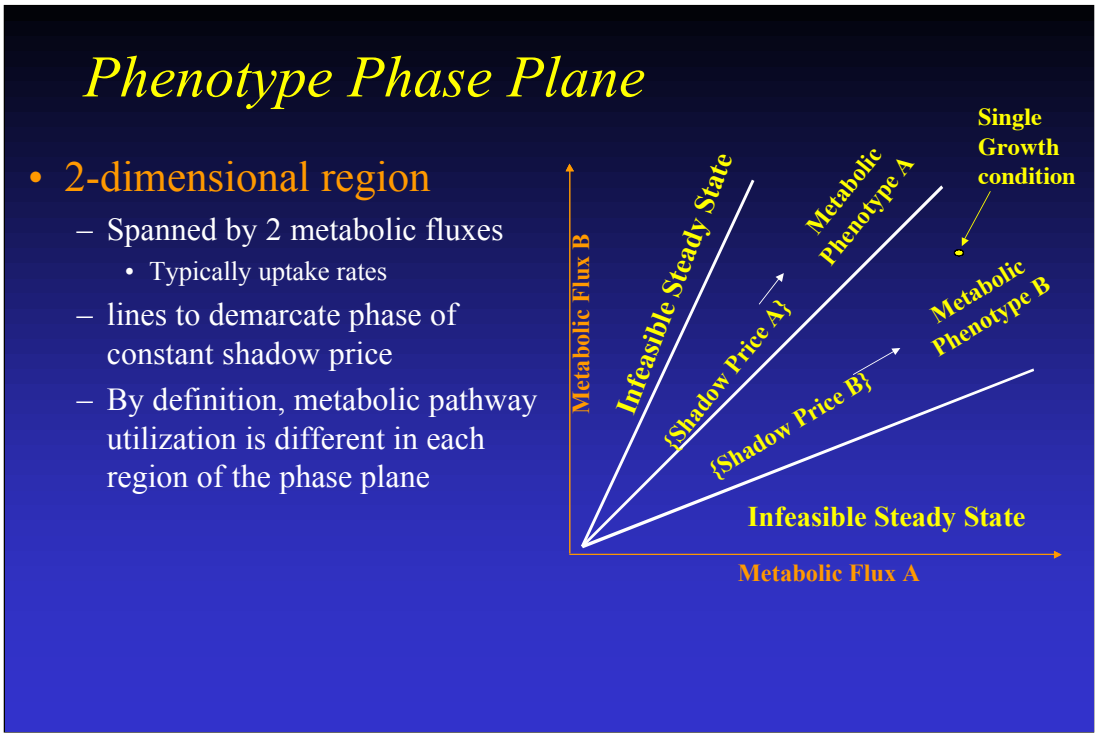


CHANGES IN SHADOW PRICES AT PHASE BOUNDARIES

The shadow price changes discontinuously at the boundary from one phase to the next. In fact the change in the shadow price defines the boundary between the phases. The shadow prices basically tell us how the governing constraints on the objective function change and how the base optimal LP solution changes. This change is reflected in a shift in the flux map.

Phase I shown above is for completely anaerobic growth. The shadow price for oxygen and ATP is positive, indicating that these are constraining factors, since the objective function would increase if more of these compounds were provided to the cell. Some of the redox carriers have negative shadow prices indicating that the cell has a problem with excess redox potential. The latter is characteristic of anaerobic metabolism.

In Phase I, acetate, ethanol, and formate, all have zero shadow prices, indicating that these intermediates are useless to the cell. Thus they are secreted. Notice that in Phase II, ethanol has a positive shadow price. It thus has value to the cell and is not secreted. In fact the defining difference between the optimal flux maps in phase I and II is the secretion of ethanol. The shadow prices are thus key in interpreting the optimal flux maps and changes in the maps as parameters vary.



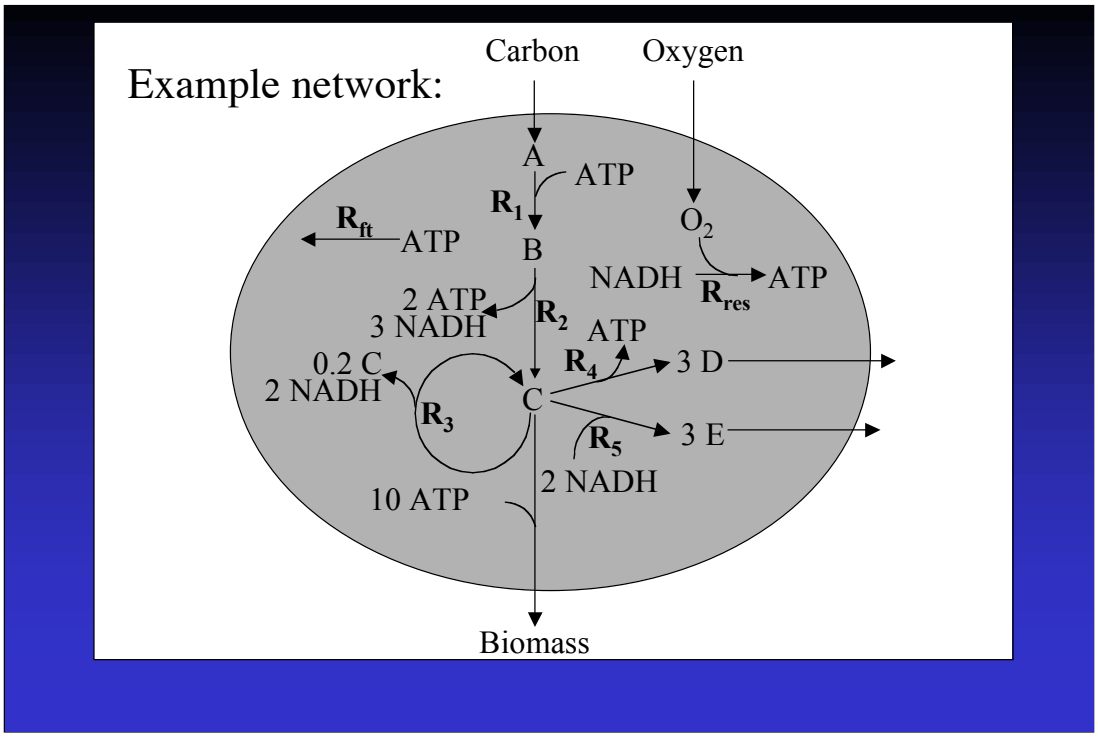
VARYING TWO PARAMETERS: THE PHENOTYPIC PHASE PLANE

A phase plane is a two dimensional region that is spanned by 2 metabolic fluxes. These fluxes are often uptake rates, but this isn't required. The shadow prices for the metabolites are calculated for all the points within this space, and lines are drawn to demarcate regions of constant shadow prices.

The shadow prices are constant within each region and are different in the other regions.

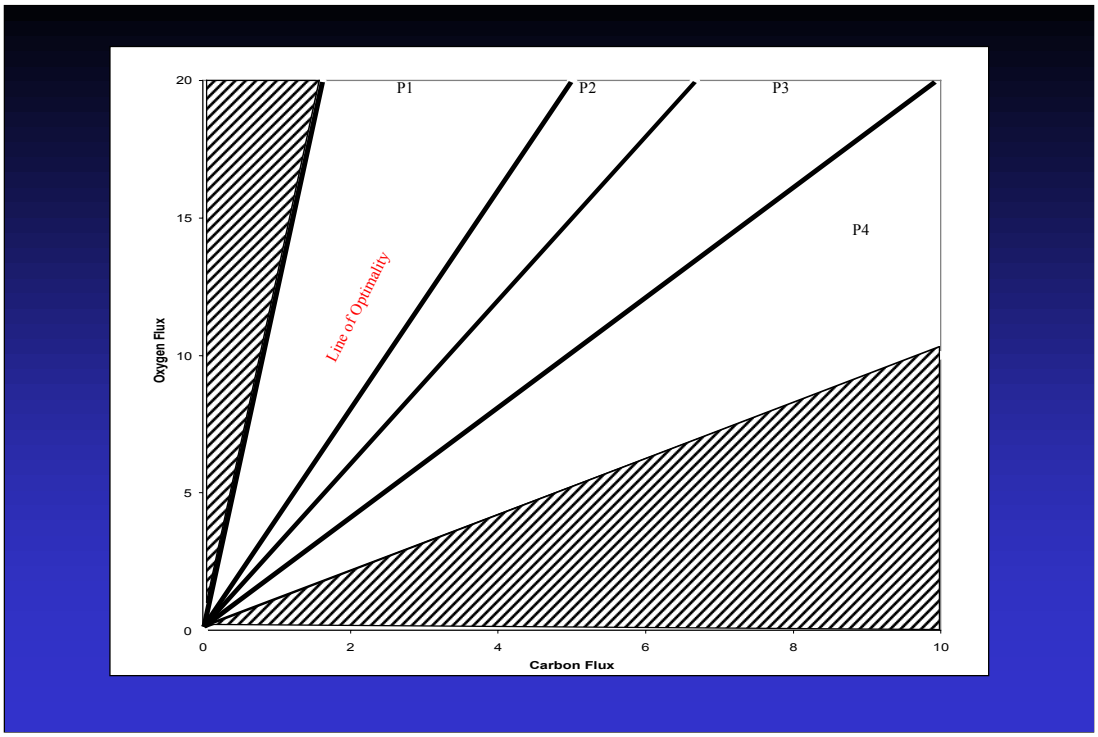
Each region refers to a different basis solution, which implies a different utilization of the metabolic pathways.

Thus, the utilization of the metabolic pathways will be qualitatively different depending on the region of operation within the phase plane.

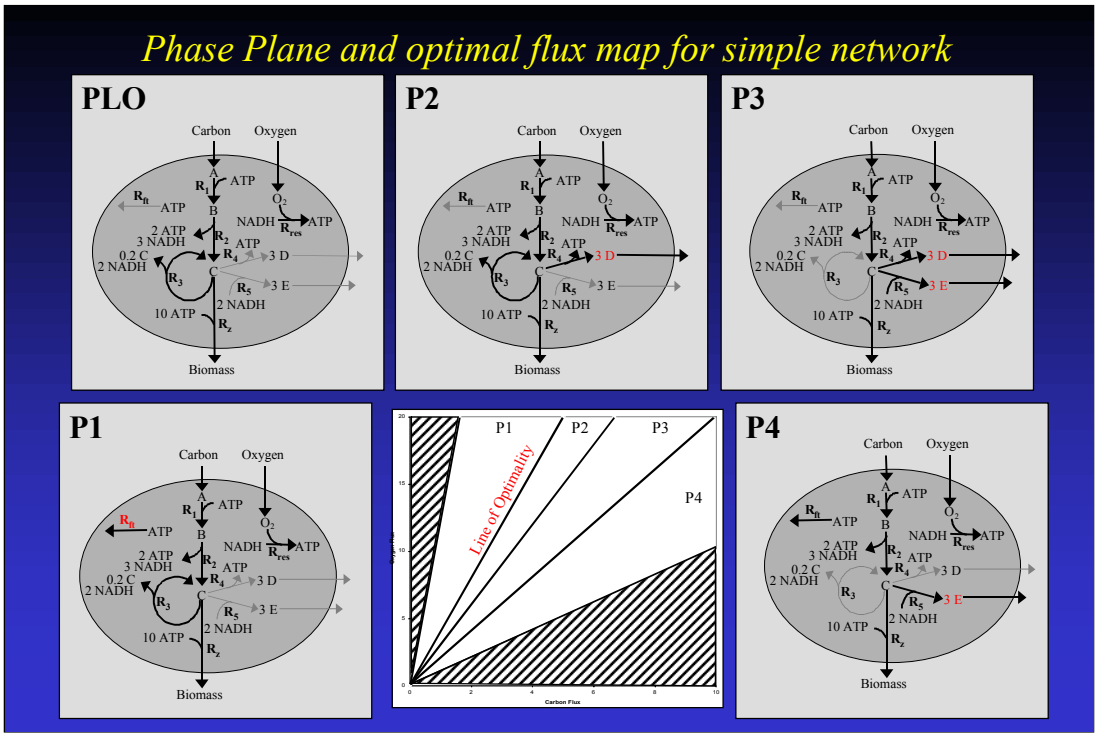


EXAMPLE

To illustrate these concepts, we now present an example of a hypothetical metabolic system. This network utilizes a single carbon source, which it metabolizes to a single biosynthetic precursor, **C**. This precursor is converted into biomass, via $\mathbf{R}_z$ (the objective function), and to two different metabolic by-products, **D** and **E**. An electron acceptor, oxygen, is also included in this example. This electron acceptor can be used to convert redox potential into high-energy phosphate bonds, $\mathbf{R}_{res}$. Additionally, there is a reaction, $\mathbf{R}_3$, which consumers 0.2 **C** to generate NADH. Finally, one reaction, $\mathbf{R}_{ft}$, represents futile cycles that hydrolyzes ATP.



The methods presented in the previous slides were used to calculate the PhPP for this hypothetical metabolic system. The PhPP and the qualitative flux maps for each phase are shown the next slide. P1 is the futile region where the electron acceptor is provided in excess. The metabolic network dissipates the excess electron acceptor taken up by the cell by increasing the flux in $\mathbf{R_3}$, which generates NADH but also oxidizes the precursor, **C**. Additionally, the futile cycle reaction $\mathbf{R_{ft}}$ is utilized to eliminate the excess ATP produced. The upper limit of P1 occurs when the entire biosynthetic precursor produced is oxidized to eliminate the excess electron acceptor, and thus no biomass can be generated.



The metabolic flux map of this system is also shown for conditions on the line of optimality (LO). The LO is a special case of P1, this is the point where the electron acceptor is no longer in excess and the futile cycle flux is zero (Table 1). The LO represents the optimal utilization of this example metabolic network to produce biomass. The qualitative flux map indicates that under conditions defined by the LO there is no metabolic by-product production and futile cycle flux equals zero.

The next distinct flux map for this hypothetical metabolic network is found in region P2. In P2 a reduced metabolic by-product (**D**) is secreted from the cell. The shadow price for the metabolite **D** in this system is zero in region P2, and the utilization of the metabolic pathways in this region is fundamentally different than in P1, Plo, P3, and P4. The metabolic pathway for the production and secretion ($\mathbf{R_4}$) of **D** is turned on under the conditions defined in this region, and the excess redox potential is eliminated through the secretion of **D**.

The utilization of the metabolic network in P3 is fundamentally different than in P2. In P3, the cyclic reaction $\mathbf{R_3}$ is not utilized, and thus redox potential production is reduced. Both of the reduced metabolic by-products are secreted (**D** and **E)** as sinks for redox potential. Thus, in this region, both of these metabolites will have a shadow price equal to zero.

Shadow prices for simple network

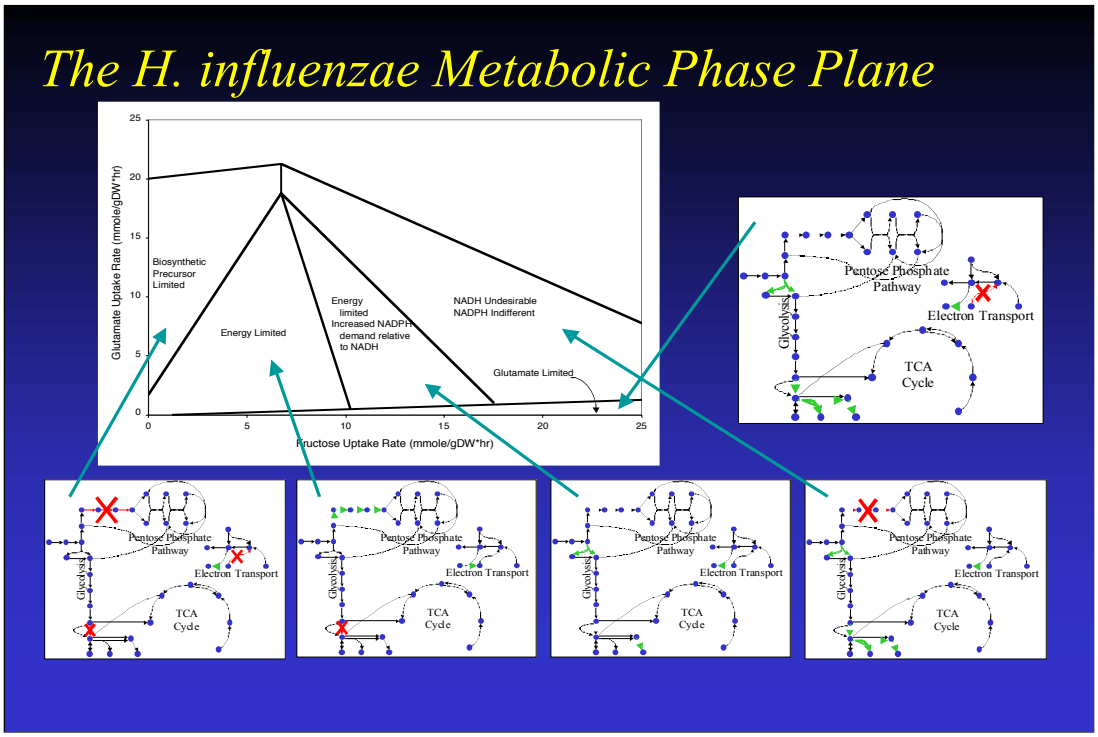
Table 1: Shadow price of the metabolites from the example shown in Figure 1.

	Carbon	A	B	C	D	E	O_2	NADH	ATP
P1	-1.30	-1.30	-1.30	-1.00	-0.33	-0.40	**0.10**	-0.10	
Plo	-0.90	-0.90	-0.93	-0.67	-0.21	-0.27		-0.07	-0.03
P2	-0.21	-0.21	-0.30	-0.09		-0.04	-0.17	-0.01	-0.09
P3	-0.05	-0.05	-0.14	-0.09			-0.23	**0.05**	-0.09
P4	**0.50**	**0.50**	**0.50**	-1.00	-0.33		-0.50	**0.50**	

Finally, in P4, the futile cycle reaction is utilized, and all the metabolic by-product formation is directed toward the formation of the more reduced by-product, **E**. When the oxygen uptake and the carbon uptake define a point on the lower boundary of P4, all the carbon source is directed toward the formation of metabolite **E**, and no biomass is generated. Thus, below this line (the crosshatched region) is another region of unobtainable steady states of the metabolic network.

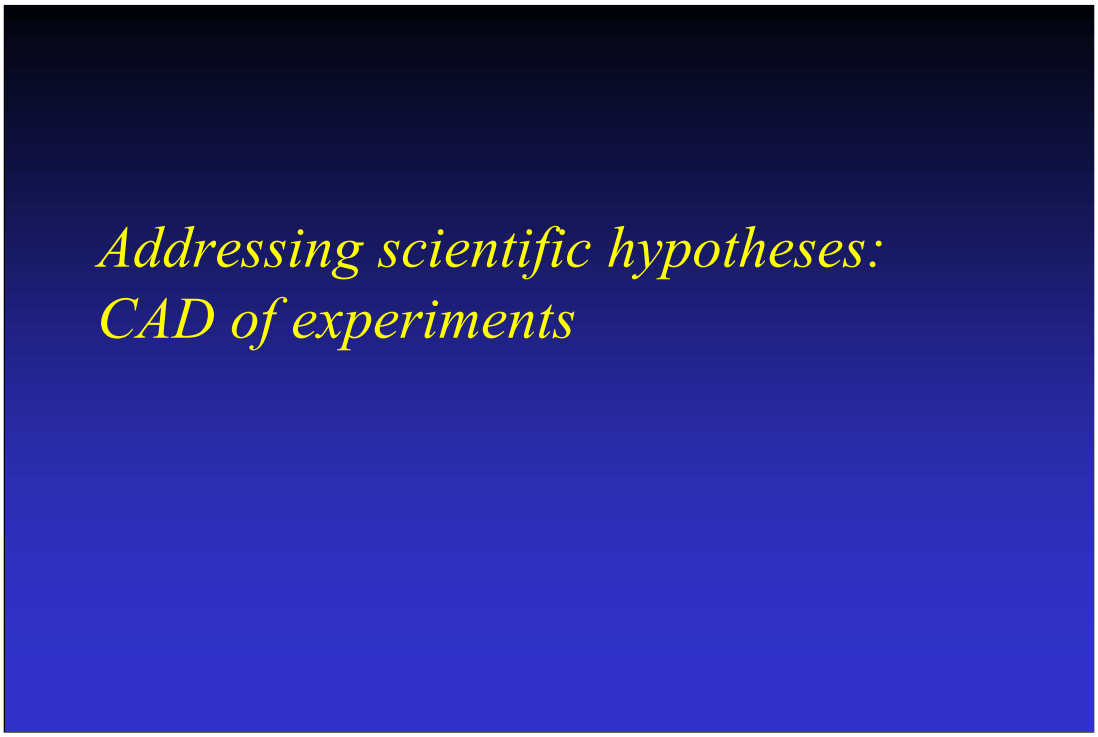
This simple example illustrates the utility of the PhPP in the interpretation of the metabolic physiology of the system. It clearly shows that the optimal phenotypes are condition dependent, and that a finite number of qualitatively different optimal phenotypes can be derived from a single genotype.

End of example

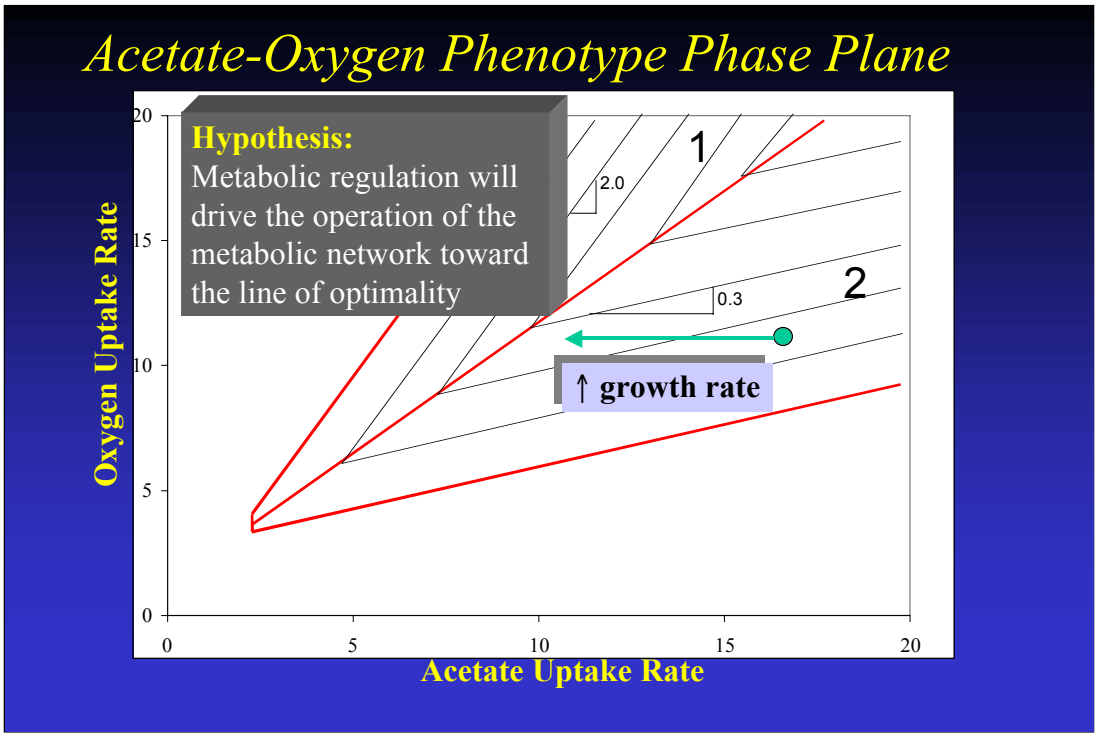


An example of a phase plane for a genome scale metabolic map.

From J.S. Edwards and B.O. Palsson (1999), "Systems Properties of the *Haemophilus influenzae* Rd Metabolic Genotype," *The Journal of Biological Chemistry*, **274**: 17410-17416.



Perhaps the most useful application of in silico strains is to design meaningful experiments. Agreement confirms the model, while failure indicates that the model is missing features. Therefore we like failure, so that the model can be continually improved.



INTERPRETING THE PHASE PLANE:

Using isoclines

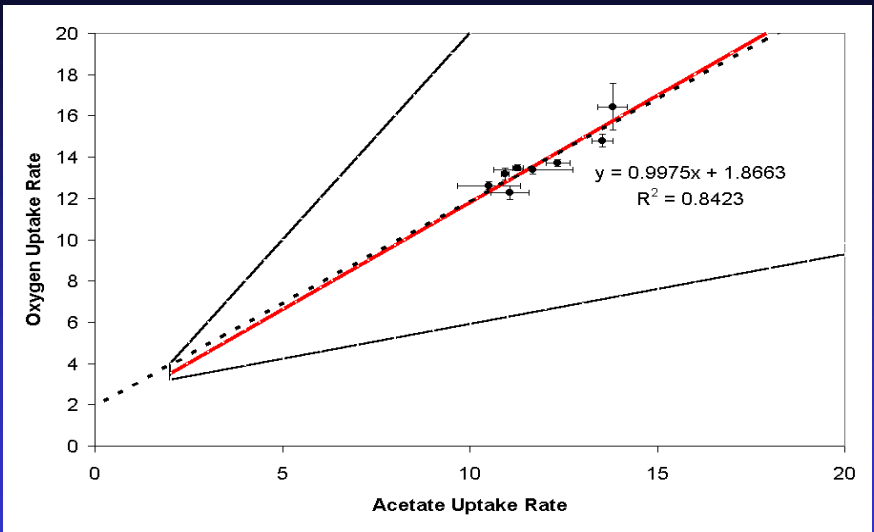
This slide describes the acetate-oxygen phenotype phase plane for E. coli.

It can be seen that there are 2 distinct regions. We have also drawn the isoclines on this figure, and it can be seen that the isoclines have a positive slope in both regions. This means that they are unstable -- it is advantageous for the organism to move the the edge of the region

The optimal growth occurs at the line separation the two phases, the so-called line of optimality.

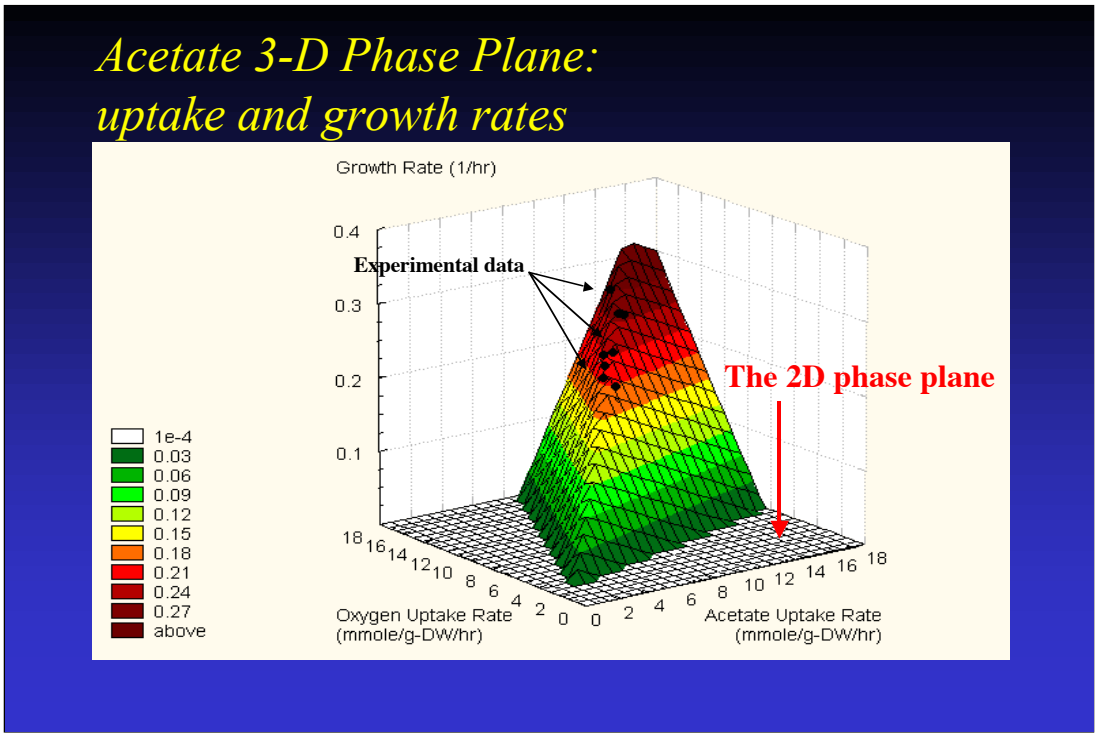
The thinner lines in each feasible phase plane are called isoclines. They denote a constant growth rate.

Acetate Phase Plane:Experimental Data



THE EXPERIMENTAL DATA:

Right on the line!



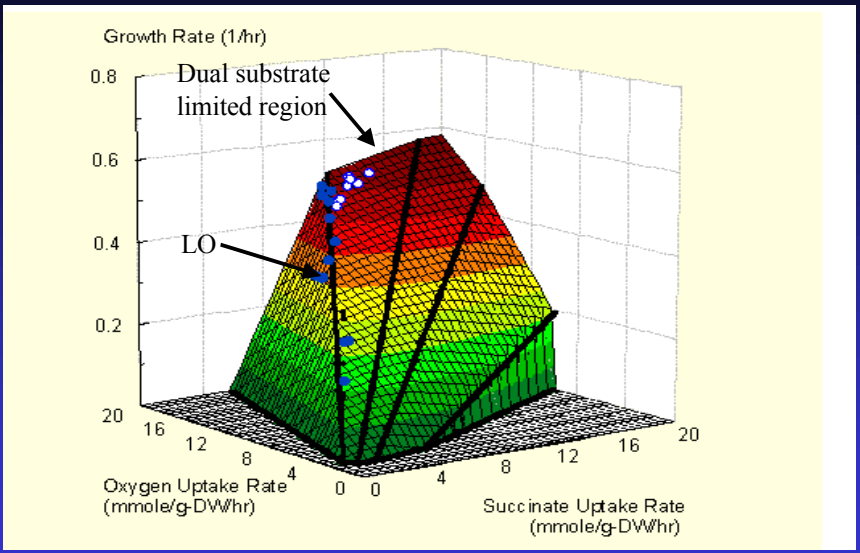
3D REPRESENTATION:

Including growth rate as a dependent variable

This slide shows how the maximal growth rates can be graphed above the phenotypic phase plane. We see the outline of a cone. For a given maximal uptake rate of either acetate or oxygen, the best (highest growth rate) solution is on the edge of the cone.

The experimental data falls there, indicating that the E. coli strain has optimized its growth rate on acetate.

Succinate 3-D Phenotype Phase Plane



This also works for other substrates!

The case of succinate

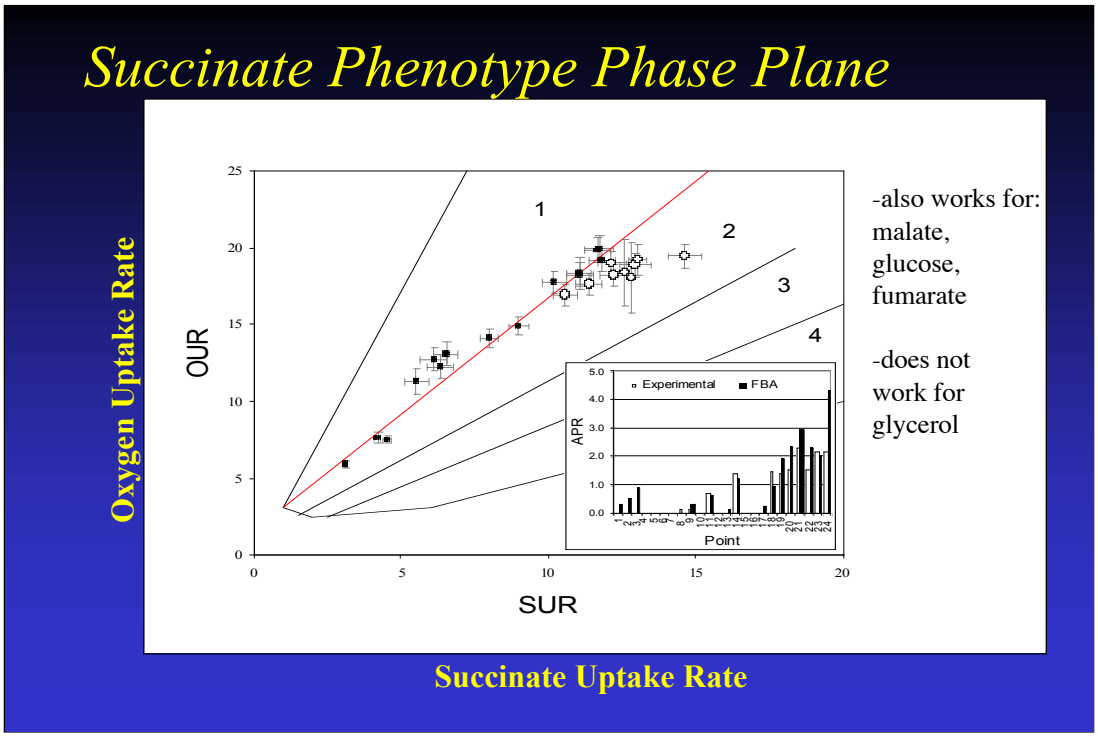
This figure, shows the succinate-oxygen PhPP in three dimensions.

•The formalism is similar to the 3-D acetate PhPP

•Here the effect of the carbon source on the structure of the PhPP can be seen.

•The LO is shown here, and the data points with reduced succinate uptake rates all lie on (or near) the LO,

•However, when the succinate uptake rate was increased, the experimental data followed the LO until the oxygen mass transfer constraint was reached. At this point, the growth rate and the succinate uptake were increased by moving into region 2 of the phase plane (white data points).



Succinate in 2D

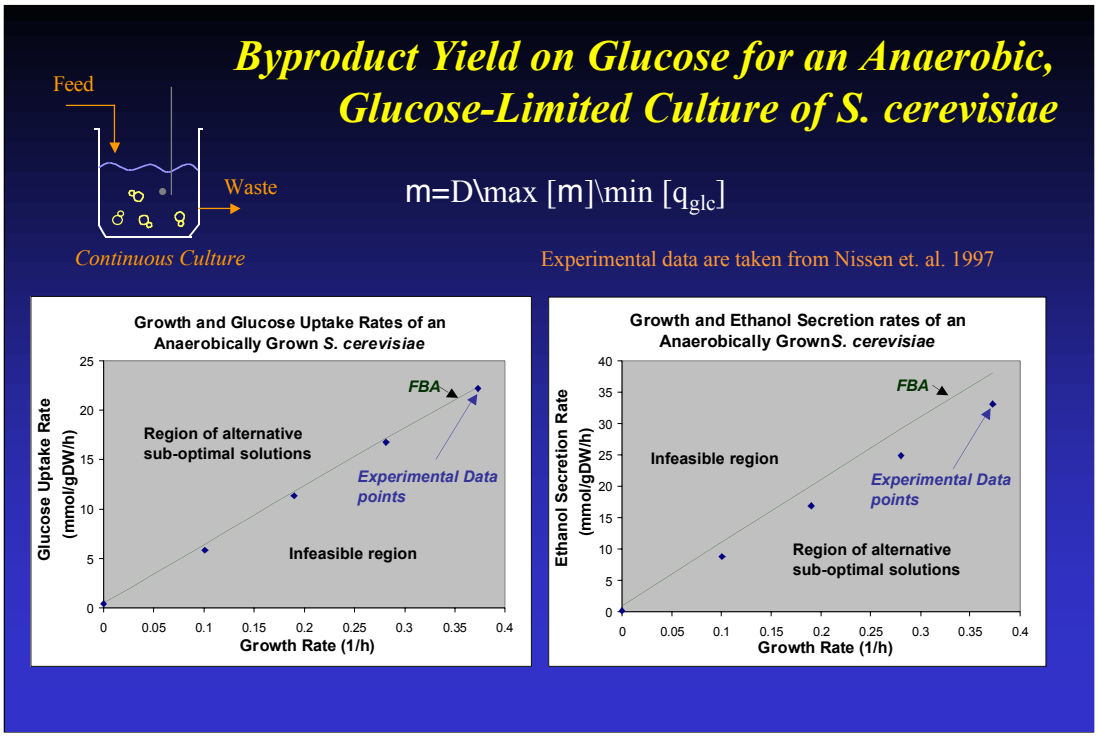
All the experimental data points were plotted onto the succinate-oxygen PhPP. And the results are shown on this slide.

Consistent with the maximal growth hypothesis, all the data points were constrained to region 2 of the PhPP.

•Within region 2, all the points were restricted to two different regions.

- •either they were on the LO, or
- •they were at a maximal oxygen uptake rate with the succinate uptake rate defining points within region 2.

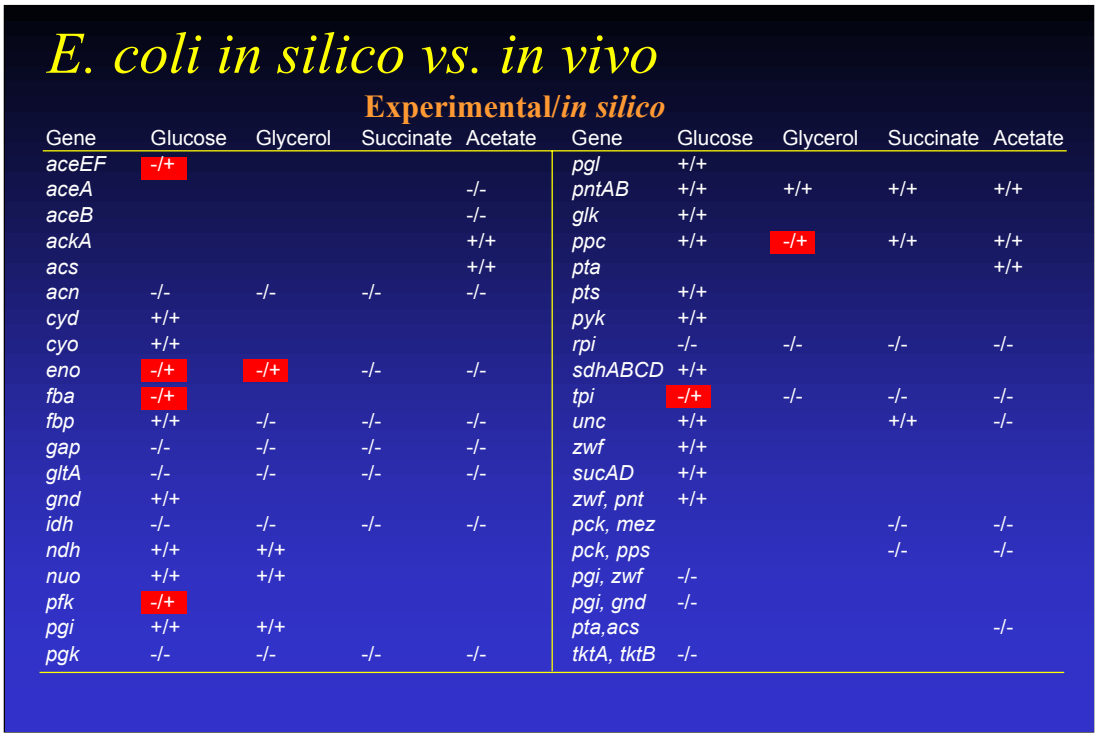
•The insert shows the calculated and measured acetate secretion rate in within region 2



CONTINUOUS CULTURE OF YEAST

In continuous culture the dilution rate specifies the growth rate. In the previous slides we maximized the growth rate (output) for a given uptake rate (input). Here in contrast, we fix the growth rate (output) and thus minimized the input (the uptake rate). The in silico solution and the measured uptake rates are shown and they agree reasonably well. Yeast seems to operate close to the edge of its allowable solution cone.

This formalism can be used to examine changes in the genotype. Genes can be added or deleted and the consequences on the ability to grow, or to generate other phenotypes, can be calculated and compared to the wild type in silico strain.



DELETION STUDY

•An important question arises as to how well these in silico predictions represent the actual metabolic behavior.

•The plus/minus nomenclature represents the ability of the respective mutant cell to grow. The first being the experimental determination, and the second being the in silico prediction.

•We have compared our in silico results to the growth of mutants in about 80 different conditions reported in the literature, and the results are summarized on this slide. The in silico strain correctly predicted the ability to grow in all but 7 cases.

•The inaccuracies are highlighted here by the red boxes.

EMBRACING AND ANALYZING FAILURE

•There are 7 inaccuracies, and they are explained by 2 basic reasons.

- •1. toxic intermediate production
- •2. metabolic regulation

•For example, when these gene products are removed, it is thought that the cell produces a toxic intermediate, and this prevents the cell from growing. This can not be predicted using the methodology that I have introduced.

•Also, when the enolase gene is removed from the system, the experimental data suggests that this cell is unable to grow, whereas the *in silico* cell is able to grow, and upon further examination, it is seen that the *in silico* cell is able to grow by synthesizing and degrading an amino acid, something that the cell is unlikely to do.

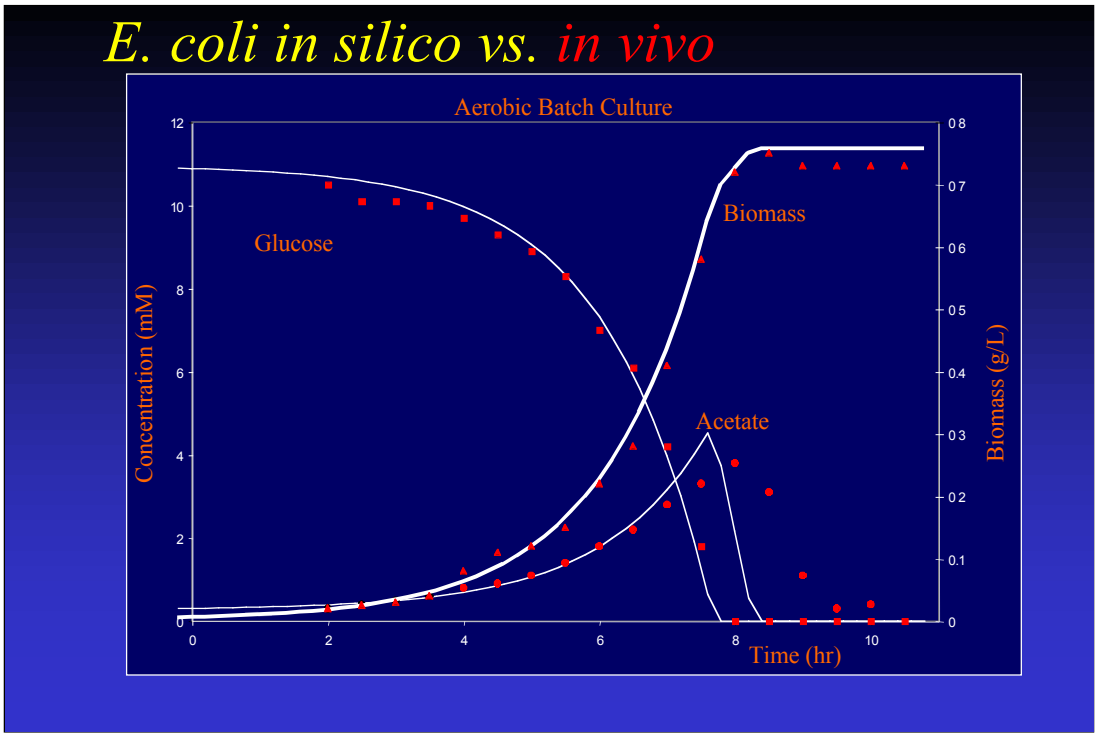
•However, it has been observed that revertants can spontaneously arise with altered expression.

•For example, ATPase mutants have been shown not to grow on succinate, however, this metabolic model predicts that they theoretically can. It was recently reported a couple of months ago that the ATPase deletion strains were unable to grow due to a transport deficiency, and revertants arose after about a week that do grow on succinate, at yields near the theoretical maximum.

Predicting expression arrays

Metabolic maps show the phenotype. Expression arrays also show the phenotypes. One is a flux phenotype whereas the other is the expression phenotype. The two cannot be directly and quantitatively compared.

However, the two can be qualitatively compared for a transition from one state to another. Pathways need to be up and down regulated. The patterns of the two can be compared qualitatively, i.e. in an off/on sense,



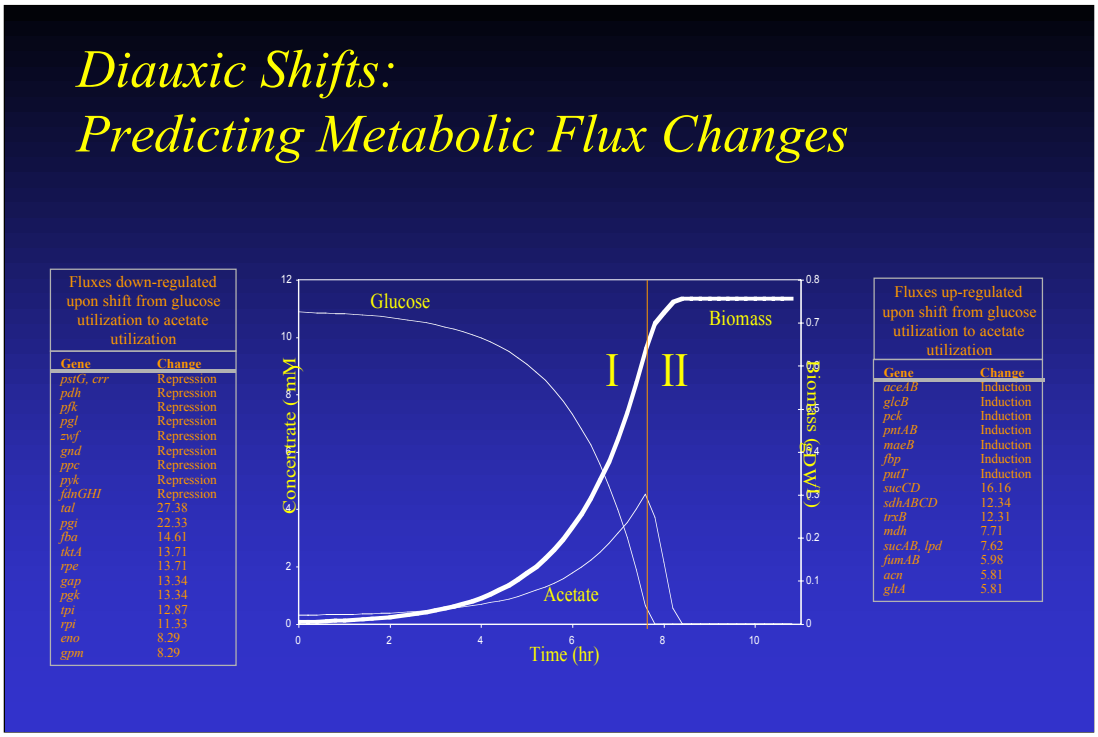
METABOLIC SHIFTS

This slide demonstrates an aerobic batch culture in glucose minimal media. The lines are the FBA predictions from a quasi-steady state simulation in a batch culture, and the points are experimentally determined.

This line represents the glucose concentration in the media, and it can be seen, as the glucose is utilized, the cells grow, and produce acetate. At this point, the glucose is completely utilized from the media, and the simulation predicts the reutilization of the acetate, and this is also experimentally observed.

However, it is at this point that the in silico predictions deviate from the experimental data. Due to the steady state assumption, the in silico strain is able to immediately reutilize the acetate. However, the experimental data lags behind by about 40 minutes.

This lag is due to the time required to to adjust the metabolic network for acetate utilization.

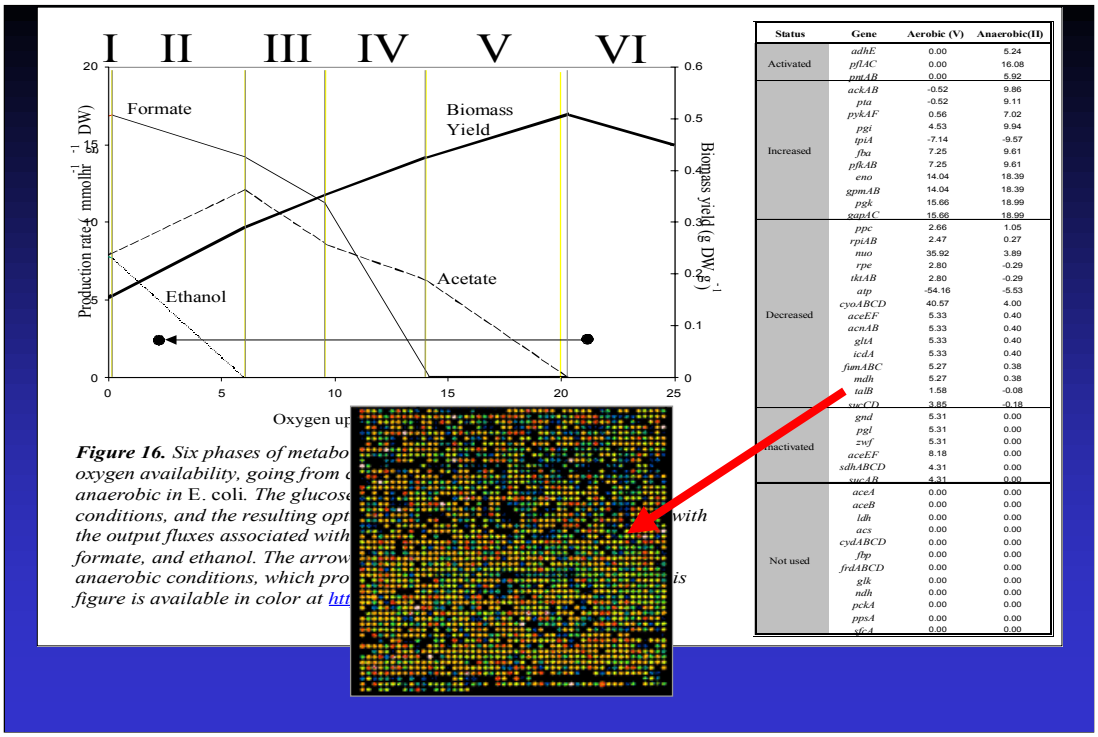


THE DIAUXIC SHIFT

This figure reproduces the data from an earlier slide, where batch growth on glucose was observed with the secretion of acetate. Then the acetate was re-consumed. The flux maps for growth on acetate and glucose are quite different. The relative flux levels through all the steps can be compared. Based on such comparisons relative fluxes through the different metabolic steps can be estimated.

If the expression levels are proportional to the needed flux levels then the indicated (predicted) up- and down-regulation of genes should be observed.

This result is a testable experimental hypothesis.



Status	Gene	Aerobic (V)	Anaerobic(II)
Activated	*adhE*	0.00	5.24
	pflAC	0.00	16.08
	pntAB	0.00	5.92
Increased	*ackAB*	-0.52	9.86
	pta	-0.52	9.11
	pykAF	0.56	7.02
	pgi	4.53	9.94
	tpiA	-7.14	-9.57
	fba	7.25	9.61
	pfkAB	7.25	9.61
	eno	14.04	18.39
	gpmAB	14.04	18.39
	pgk	15.66	18.99
	gapAC	15.66	18.99
Decreased	*ppc*	2.66	1.05
	rpiAB	2.47	0.27
	nuo	35.92	3.89
	rpe	2.80	-0.29
	tktAB	2.80	-0.29
	atp	-54.16	-5.53
	cyoABCD	40.57	4.00
	aceEF	5.33	0.40
	acnAB	5.33	0.40
	gltA	5.33	0.40
	icdA	5.33	0.40
	fumABC	5.27	0.38
	mdh	5.27	0.38
	talB	1.58	-0.08
	sucCD	3.85	-0.18
Inactivated	*gnd*	5.31	0.00
	pgl	5.31	0.00
	zwf	5.31	0.00
	aceEF	8.18	0.00
	sdhABCD	4.31	0.00
	sucAB	4.31	0.00
Not used	*aceA*	0.00	0.00
	aceB	0.00	0.00
	ldh	0.00	0.00
	acs	0.00	0.00
	cydABCD	0.00	0.00
	fbp	0.00	0.00
	frdABCD	0.00	0.00
	glk	0.00	0.00
	ndh	0.00	0.00
	pckA	0.00	0.00
	ppsA	0.00	0.00
	sfcA	0.00	0.00

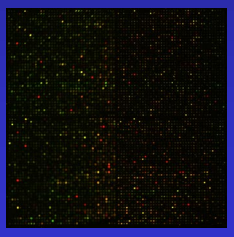
Figure 16. *Six phases of metabo*
oxygen availability, going from a
anaerobic in E. coli. *The glucose*
conditions, and the resulting opt *with*
the output fluxes associated with
formate, and ethanol. The arrow
anaerobic conditions, which pro *is*
figure is available in color at htt

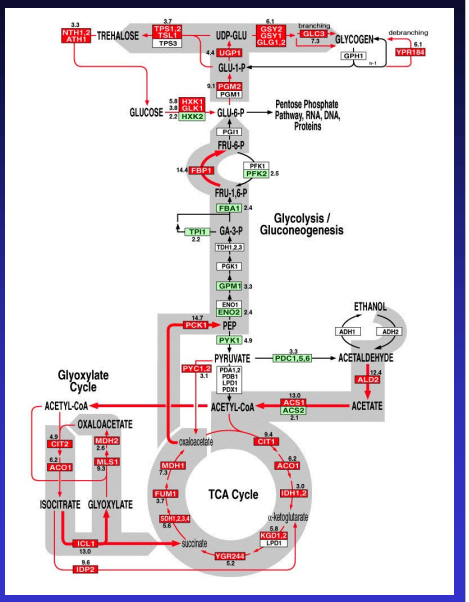
The relative *in silico* calculated fluxes can be compared to the relative expression levels under the two conditions considered. Only qualitative comparisons can be made since the flux is not proportional to the expression levels.

Gene expression on a genomic scale

- Analyzing the gene expression patterns during growth in different conditions
  - oxygenation, carbon sources
- Comparing the gene expression patterns to the FBA predictions - wild-type & knockout strains
- Goal is to relate gene expression patterns to metabolic pathway utilization

DeRisi et.al.,
Science, 278:680-686, 1997





DIAUXIC SHIFT IN YEAST FOR GROWTH ON GLUCOSE

It has been shown by Patrick Brown's group at Stanford, that the shift in metabolic pathway utilization can be determined from genomic scale measurements of gene expression.

They have generated cDNA micro-arrays with probes for virtually every gene in the yeast genome, and used these micro-arrays to study the changes in gene expression on a genome scale during a diauxic shift from glucose to ethanol utilization.

Shifts in expression levels that correspond to pathway usage were observed.

Summary

- Maximum capacity constraints close the flux cone
- LP can be used to find optimal solutions in the so formed closed solution space
- There are many types of objectives that can be studied; perhaps the maximal growth rate is the most appropriate
- Methods can be developed to show all optimal solutions as a function of environmental parameters
- The phase plane analysis shows that there is a finite number of optimal phenotypes
- This analysis can be used to interpret and predict the consequences of losing genes and the expression changes during shifts from one growth condition to another

References

- Papoutsakis, E.T., "Equations and calculations for fermentations of butyric acid bacteria," *Biotechnol Bioeng*, **26**: 174-187 (1984).
- Papoutsakis, E. and Meyer, C.,"Equations and calculations of product yields and preferred pathways for butanediol and mixed-acid fermentations," *Biotechnol Bioeng*, **27**: 50-66 (1985).
- Fell, D.A. and Small, J.A., "Fat synthesis in adipose tissue. An examination of stoichiometric constraints," *J. Biochem*, **238**: 781-786 (1986).
- Majewski, R.A., and Domach, M.M., "Simple constrained optimization view of acetate overflow in *E. coli*," *Biotechnol Bioeng,* **35**: 732-738 (1990).
- Savinell, J.M., and Palsson, B.O., "Network analysis of intermediary metabolism using linear optimization. II. Interpretation of hybridoma cell metabolism," *J Theor Biol*, **154**: 455-473 (1992).
- Varma, A. and Palsson, B.O., "Metabolic capabilities of *Escherichia coli*. I. Synthesis of biosynthetic precursors and cofactors," *J Theor Biol*, **165**:477-502 (1993).
- Varma, A. and Palsson, B.O., "Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110," *Appl Environ Microbiol*, **60**: 3724-3731 (1994).

References

- Varma, A. and Palsson, B.O., "Predictions for oxygen supply control to enhance population stability of engineered production strains," *Biotechnol Bioeng*, **43**: 275-285 (1994).
- Varma, A. and Palsson, B.O., "Parametric sensitivity of stoichiometric flux balance models applied to wild-type *Escherichia coli* metabolism," *Biotechnol Bioeng*, **45**: 69-79 (1995).
- Pramanik, J. and Keasling, J.D., "Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and and mechanistic energy requirements," *Biotechnol Bioeng*, **56**: 398-421 (1997).
- Bonarius, H. P. J., Schmid, G. & Tramper, J. (1997) *Trends in Biotechnology* **15,** 308-314.
- J.S. Edwards, R. Ramakrishna, C.H. Schilling, and B.O. Palsson, "Metabolic Flux Balance Analysis," *Metabolic Engineering* **2**: 13-57, (1999), S.Y. Lee, Papoutsakis, E.T., Eds; Springer-Verlag: New York.
- J.S. Edwards and B.O. Palsson, "Systems Properties of the *Haemophilus influenzae* Rd Metabolic Genotype," *The Journal of Biological Chemistry*, **274**: 17410-17416 (1999).
- J.S. Edwards and B.O. Palsson, "The *Escherichia coli* MG1655 *in silico* metabolic genotype; Its definition, characteristics, and capabilities," *Proc. Natl Acad Sci (USA)*, **97**: 5528-5523 (2000).