1 4 9 5

UNIVERSITY OF ABERDEEN

HEINRICH HEINE
UNIVERSITÄT DÜSSELDORF

CEPLAS
Cluster of Excellence on Plant Sciences

Differential equation-based models of metabolic networks

Oliver Ebenhöh

AccliPhot
www.accliphot.eu

Rigi Workshop, January 19, 2015

Why do we need mathematical models?

- Simplified representation of reality
- Reduction to the essentials

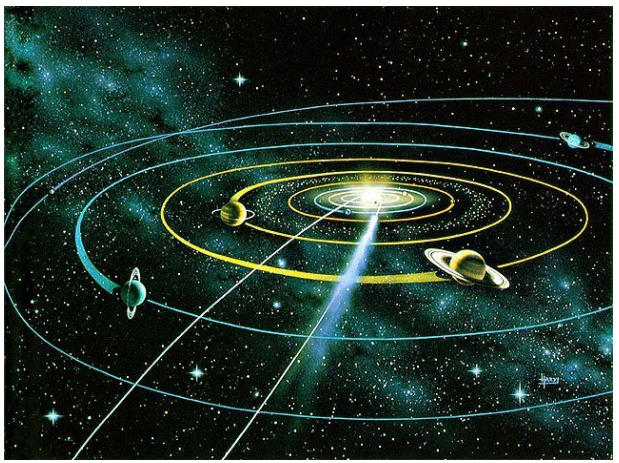
"Simplicity is the ultimate sophistication"

(Leonardo da Vinci)

⇨ Models help to discover general principles!

Example from physics:

$$\vec{F} = m \cdot \vec{a}$$

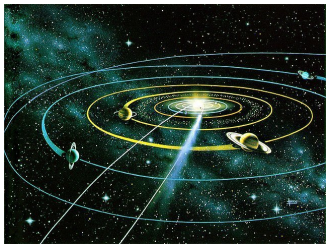


www.thehungryandfoolish.com

www.hh.schule.de

www.welt.de

How does one find principles (theory building)?



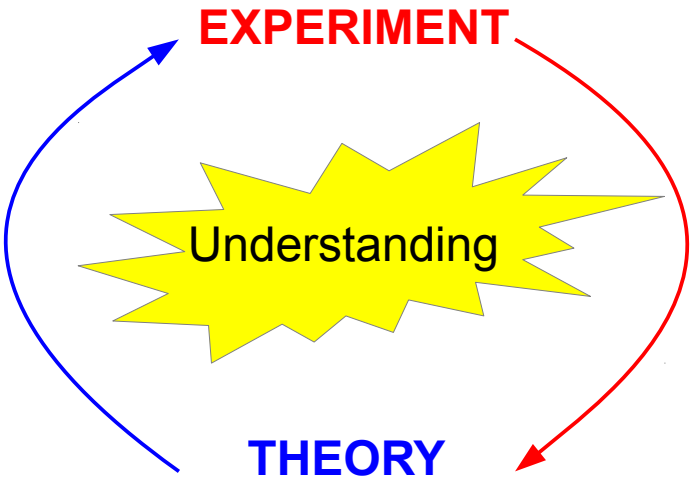
$$\vec{F} = m \cdot \vec{a}$$

Intuition

???

Every model is a small step on this path

- Model predictions / new hypotheses
- Suggestions for new experiments
- Improvement of experimental design

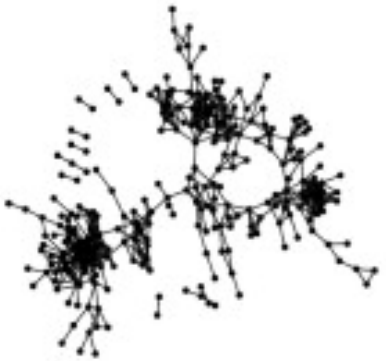


- Initial model formulation
- Confirmation / falsification of predictions
- New model assumptions

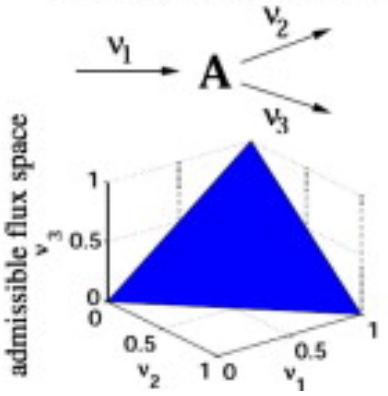
The Systems biology principle

Modelling techniques - overview

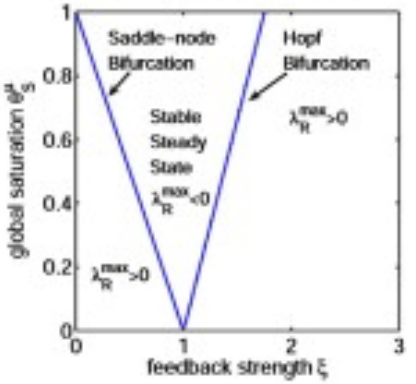
Network Analysis
- Static description
- No kinetic parameters
- Topological properties



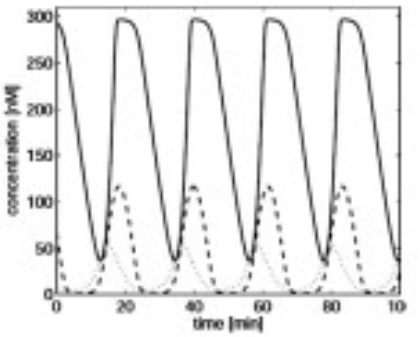
Stoichiometric Analysis
- Static description
- No kinetic parameters
- Quantitative predictions



Structural Kinetic Models
- Dynamic description
- No kinetic parameters
- Bifurcation structure



Kinetic Models
- Dynamic description
- Kinetic parameters
- Differential equations



(Steuer, 2007)

What are kinetic models?

"kinetics" is derived from greek (*κίνησις*): **movement**

In general: the study of quantities which change in time

Chemical kinetics: The study of reaction rates

Methods:

- Differential equations
- Stability analysis
- Control theory

Application to biology

- Population dynamics (long tradition)
- Biochemical reaction networks
- Gene-regulatory networks
- Signal transduction pathways
- … and many more...

Population dynamics

Let *x* denote a population (of anything, e.g. animals, bacteria, molecules, etc.)

Then, $x(t)$ describes the population as a *function* of time *t*

$\dot{x}(t) = \frac{dx}{dt}$ denotes the temporal *change* of population *x*

How the population evolves in time in general depends on the population size:

$$\dot{x} = f(x)$$

(the dependence on population size is described by a function $f(x)$)

Population dynamics

Let *x* denote a population (of anything, e.g. animals, bacteria, molecules, etc.)

Then, $x(t)$ describes the population as a *function* of time *t*

$\dot{x}(t) = \frac{dx}{dt}$ denotes the temporal *change* of population *x*

How the population evolves in time in general depends on the population size:

$$\dot{x} = f(x)$$

(the dependence on population size is described by a function $f(x)$)

ordinary differential equation of first order

Population dynamics

Let *x* denote a population (of anything, e.g. animals, bacteria, molecules, etc.)

Then, $x(t)$ describes the population as a *function* of time *t*

$\dot{x}(t) = \frac{dx}{dt}$ denotes the temporal *change* of population *x*

How the population evolves in time in general depends on the population size:

$$\dot{x} = f(x)$$

(the dependence on population size is described by a function $f(x)$)

ordinary differential equation of first order

In general, more than one variable: $x_1, \ldots, x_n$

System of differential equations: $\dot{x}_1 = f_1(x_1, \ldots, x_n)$ or $\dot{\vec{x}} = \vec{f}(\vec{x})$

$$\vdots$$

$$\dot{x}_n = f_n(x_1, \ldots, x_n)$$

(Bio-)chemical reaction networks

Chemical species: $X_1, X_2, \ldots, X_n$

General description by a system of ordinary differential equations:

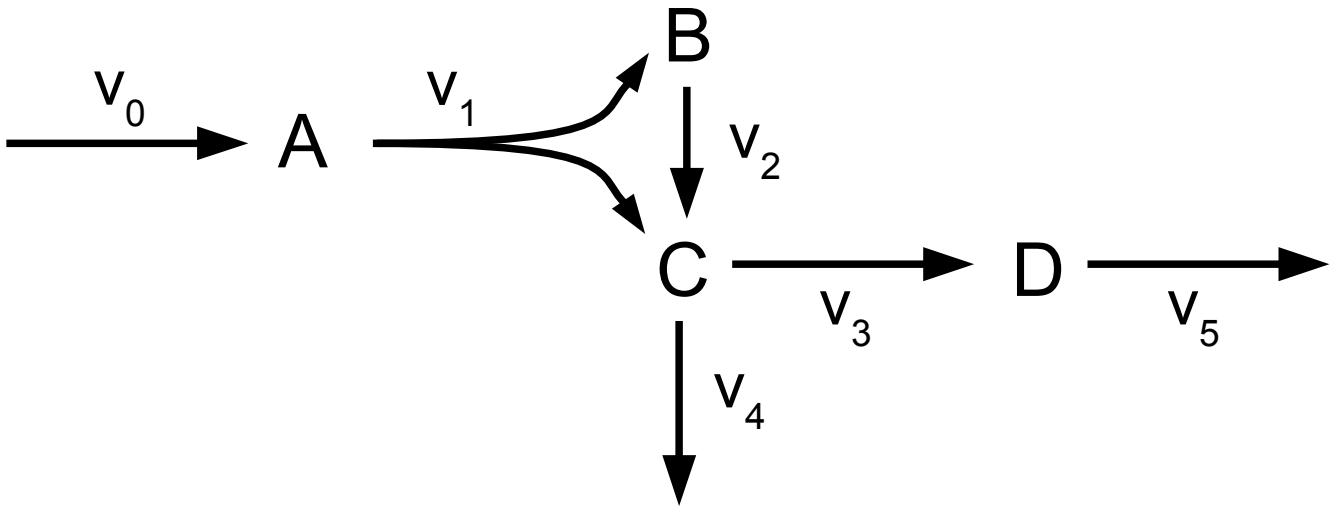
$$\frac{dX_i}{dt} = f_i\left(X_1, \ldots, X_n\right)$$

The functions f_i depend on the rate laws describing the reaction rates

Rate laws

Rate laws describe the rate of a single enzyme in dependence on the concentrations of all participating chemical species (substrates and products) and possibly affectors (activators / inhibitors)

Constructing a pathway model



- Get/derive/guess rate laws for each enzymatic reaction
- Construct a differential equation for each chemical species

$$\frac{dA}{dt} = v_0 - v_1$$

producing reaction

consuming reaction

$$\frac{dB}{dt} = v_1 - v_2$$

$$\frac{dC}{dt} = v_1 + v_2 - v_3 - v_4$$

$$\frac{dD}{dt} = v_3 - v_5$$

...and replace the v_i by the appropriate rate laws

Mass-action rate law

For (uncatalysed) chemical reactions, the rate *v* is proportional to all substrate concentrations – to the power of their stoichiometric coefficient

Examples: $A \underset{k^-}{\overset{k^+}{\rightleftharpoons}} B$ $\quad v = k^+ A - k^- B$

$$\frac{dA}{dt} = -k^+ A + k^- B; \quad \frac{dB}{dt} = -\frac{dA}{dt}$$

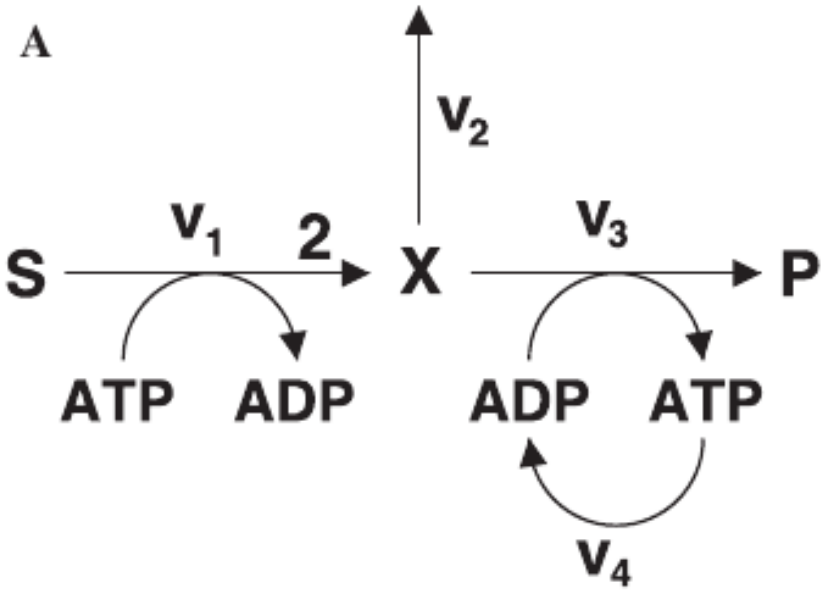
$2X+Y \xrightarrow{k} Z$ $\quad v = k X^2 Y$

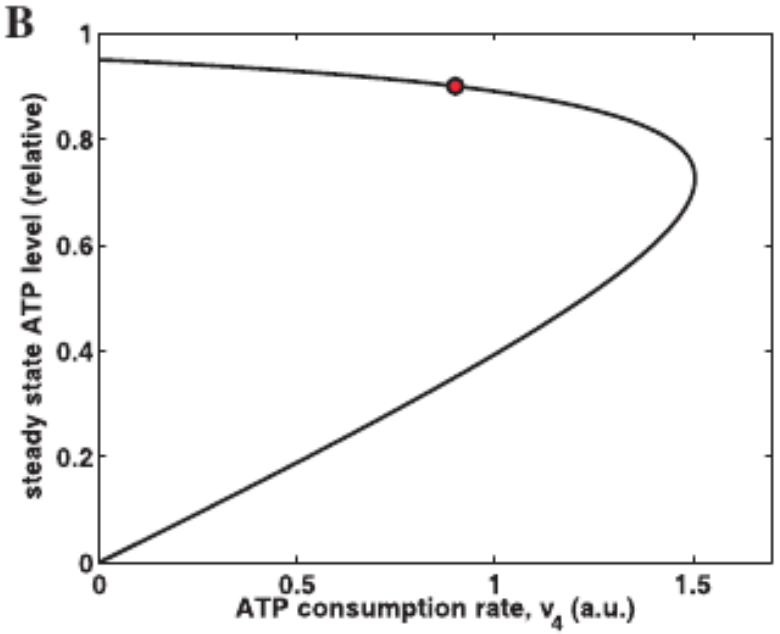
$$\frac{dY}{dt} = -k X^2 Y; \quad \frac{dX}{dt} = 2\frac{dY}{dt}; \frac{dZ}{dt} = -\frac{dY}{dt}$$

In general: $\nu_1 S_1 + \ldots + \nu_k S_k \underset{k^-}{\overset{k^+}{\rightleftharpoons}} \nu_{n+1} S_{n+1} + \ldots + \nu_k S_k$

$$v = k^+ \prod_{j=1}^{n} S_j^{\nu_j} - k^- \prod_{j=n+1}^{k} S_j^{\nu_j}$$

A simple model of glycolysis





C

Stoichiometry matrix:

$$\mathbf{N} = \begin{array}{c} \begin{array}{cccc} R_1 & R_2 & R_3 & R_4 \end{array} \\ \begin{pmatrix} +2 & -1 & -1 & 0 \\ -1 & 0 & +1 & -1 \\ +1 & 0 & -1 & +1 \end{pmatrix} \end{array} \begin{array}{l} X \\ ATP \\ ADP \end{array}$$

Dynamic equations:

$$\frac{d}{dt}\begin{pmatrix} X \\ ATP \\ ADP \end{pmatrix} = \mathbf{N} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \end{pmatrix}$$

Conserved moieties: $\mathbf{C} \cdot \mathbf{N} = \mathbf{0}$

$$\mathbf{C} = (0\,1\,1) \Leftrightarrow \frac{d}{dt}(ATP + ADP) = 0$$
$$\Leftrightarrow ATP + ADP = A = \text{const.}$$

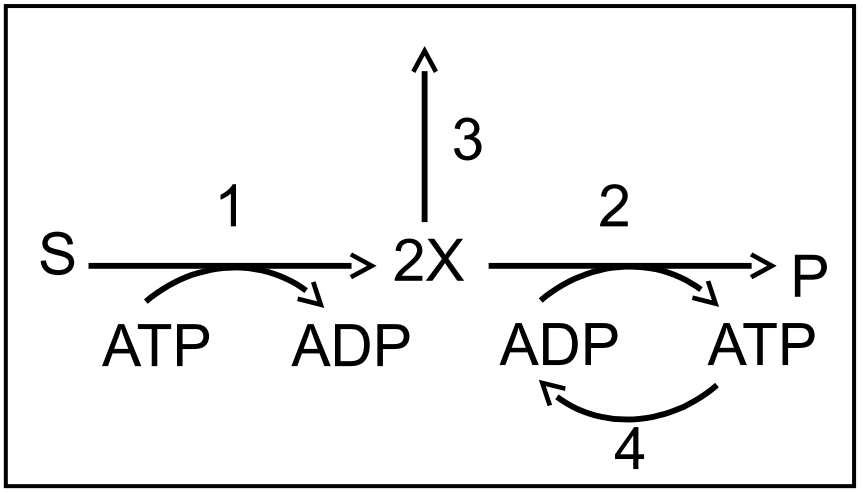
Steady state flux distributions: $\mathbf{N} \cdot \mathbf{K} = 0$

$$\mathbf{K} = (\mathbf{k_1}\,\mathbf{k_2}) \quad \mathbf{k_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{k_2} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}$$

(Pfau et al, 2011, Brief Func Genomics)

Model analysis

J.J. Selkov



Glycolysis

$$\frac{dX}{dt} = 2v_1 - v_2 - v_3$$

$$\frac{dATP}{dt} = -v_1 + v_2 - v_4$$

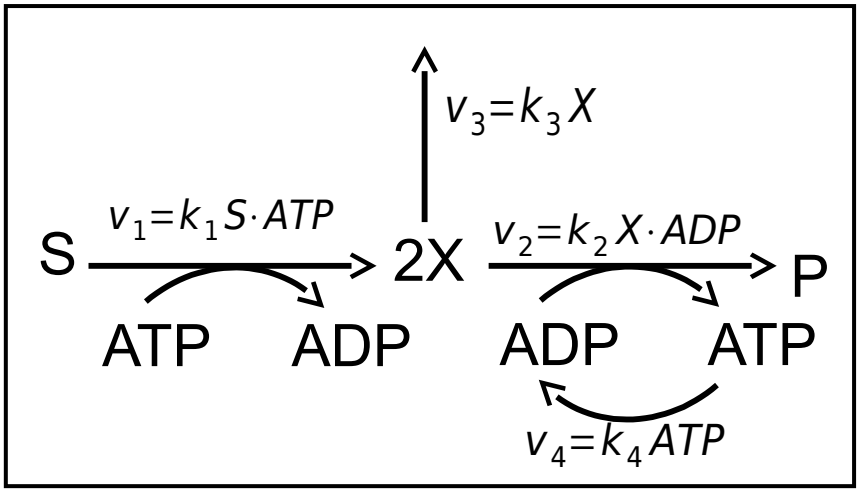
$$\frac{dADP}{dt} = -\frac{dATP}{dt}$$

$$\begin{pmatrix} \frac{dX}{dt} \\ \frac{dA_3}{dt} \end{pmatrix} = \begin{pmatrix} 2 & -1 & -1 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

stoichiometric matrix

$$ADP + ATP = A = const$$

The steady state of a pathway



A metabolite is said to be in **steady state** if its concentration remains constant in time:

$$0=\frac{dX}{dt}=2v_1-v_2-v_3$$

$$0=\frac{dATP}{dt}=-v_1+v_2-v_4$$

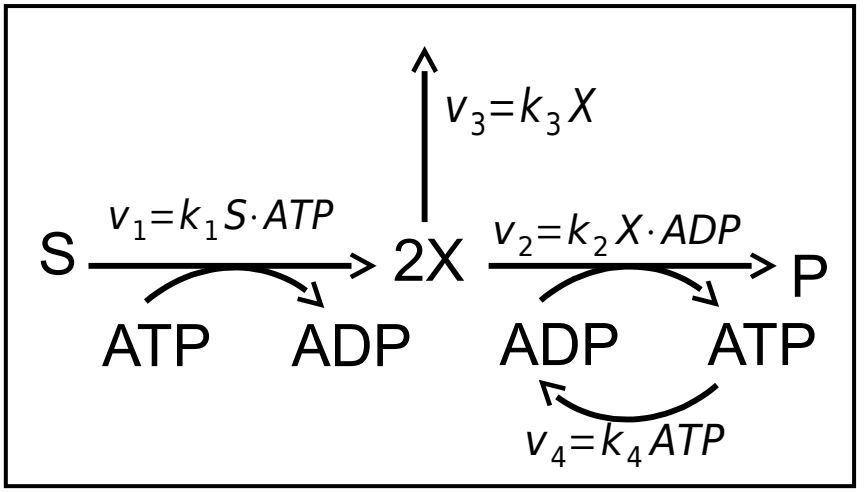
This means that a metabolite is produced and consumed with the same rate!

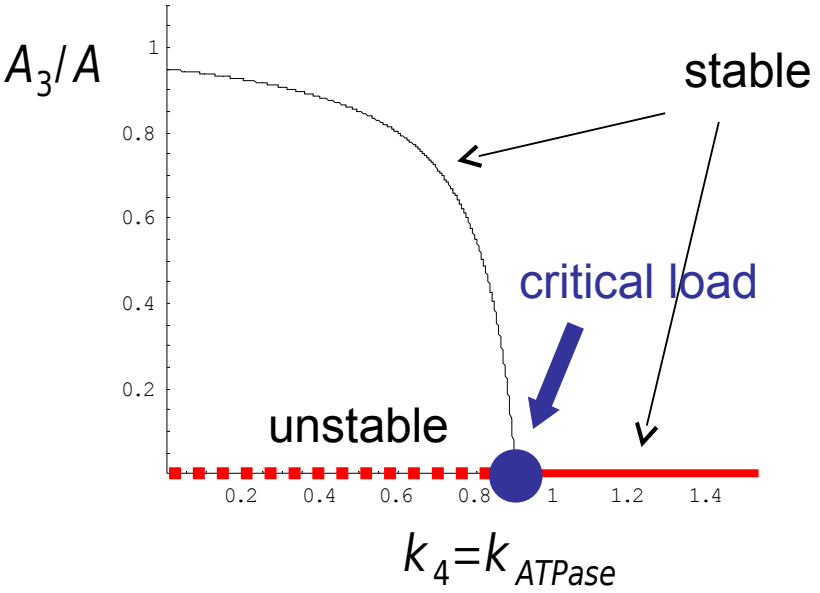
This leads to the following possible stationary ATP concentrations:

1. $ATP=0$

2. $ATP=A-\frac{k_3}{k_2}\frac{k_1S+k_4}{k_1S-k_4}$

Analytic solutions



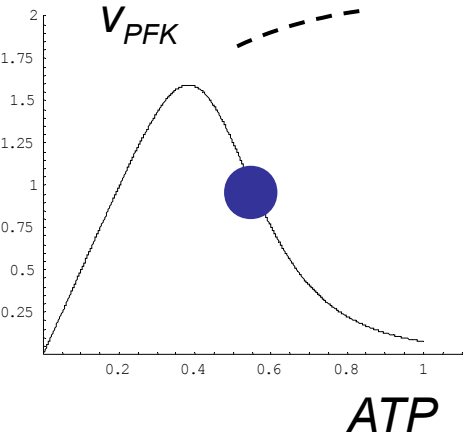


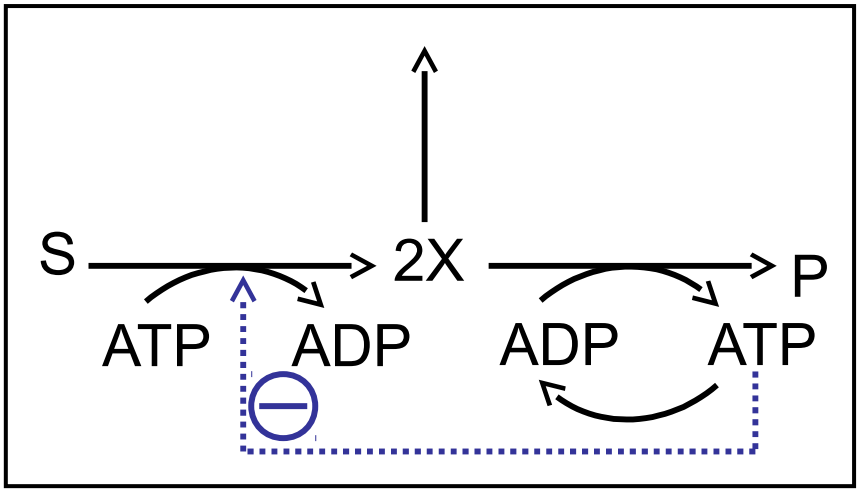
stationary ATP concentrations

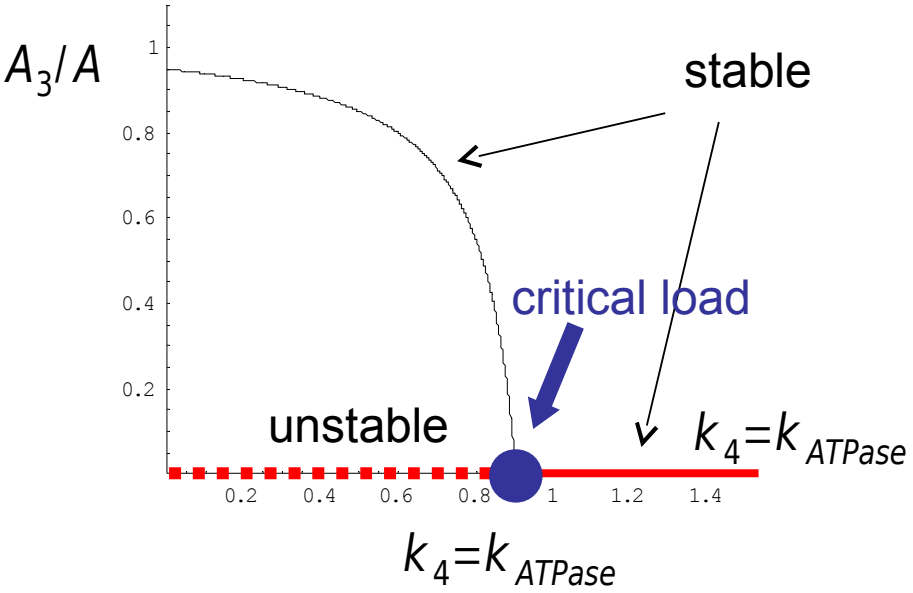
1. $ATP=0$

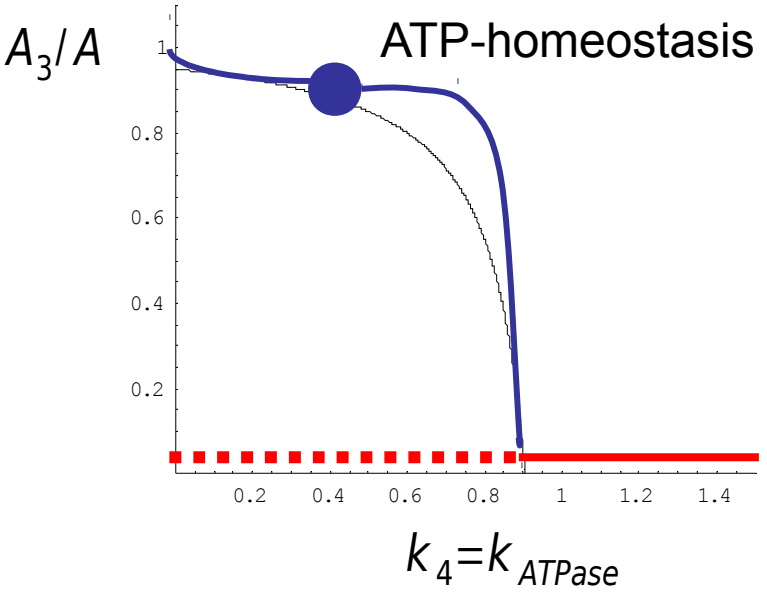
2. $ATP=A-\frac{k_3}{k_2}\frac{k_1S+k_4}{k_1S-k_4}$

Including important details

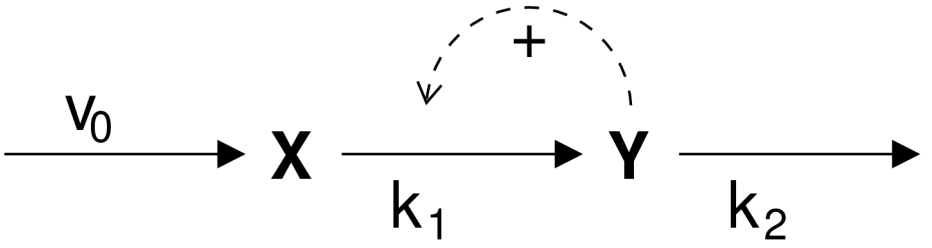




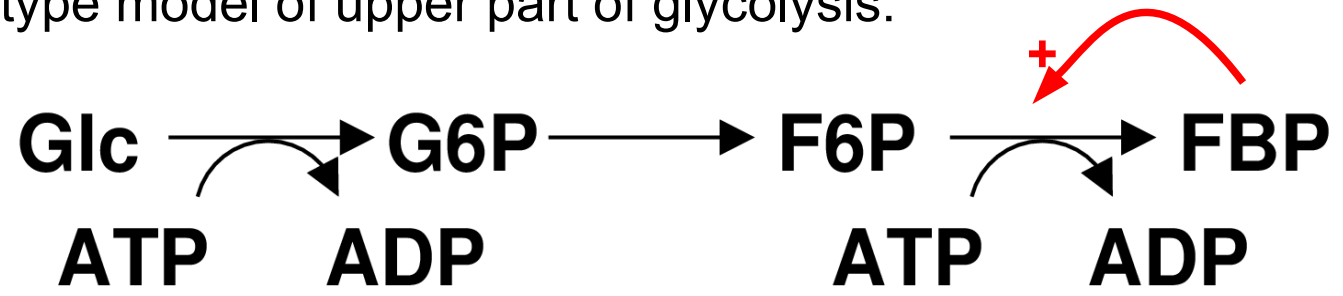




Higgins-Sel'kov oscillator



prototype model of upper part of glycolysis:



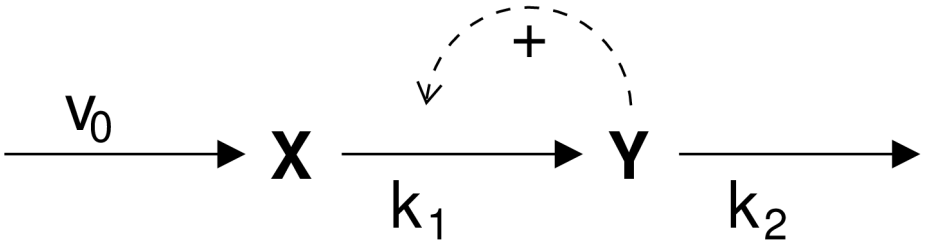
Simplest description: $v_0 = \text{const.}\,;\quad v_1 = k_1 \cdot X \cdot Y^2\,;\quad v_2 = k_2 \cdot Y$

$$\dot{X} = v_0 - k_1 \cdot X \cdot Y^2$$
$$\dot{Y} = k_1 \cdot X \cdot Y^2 - k_2 \cdot Y$$

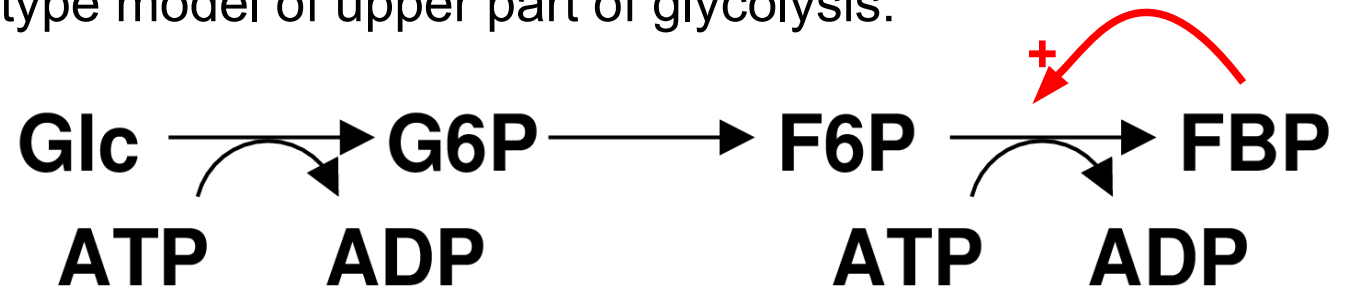
nullclines: $\dot{X} = 0 \;\Leftrightarrow\; X = \dfrac{v_0}{k_1 \cdot Y^2}\,;\qquad \dot{Y} = 0 \;\Leftrightarrow\; Y = \dfrac{k_2}{k_1 \cdot X}$

steady-state: $\dot{X} = 0 \;\text{ and }\; \dot{Y} = 0 \;\Leftrightarrow\; \bar{Y} = \dfrac{v_0}{k_2}\,;\quad \bar{X} = \dfrac{v_0 \cdot k_2^{\,2}}{k_1}$

Higgins-Sel'kov oscillator



prototype model of upper part of glycolysis:



rate law expressing activation by Y

Simplest description: $v_0 = \text{const.}\,;\quad v_1 = k_1 \cdot X \cdot Y^2\,;\quad v_2 = k_2 \cdot Y$

$$\dot{X} = v_0 - k_1 \cdot X \cdot Y^2$$
$$\dot{Y} = k_1 \cdot X \cdot Y^2 - k_2 \cdot Y$$

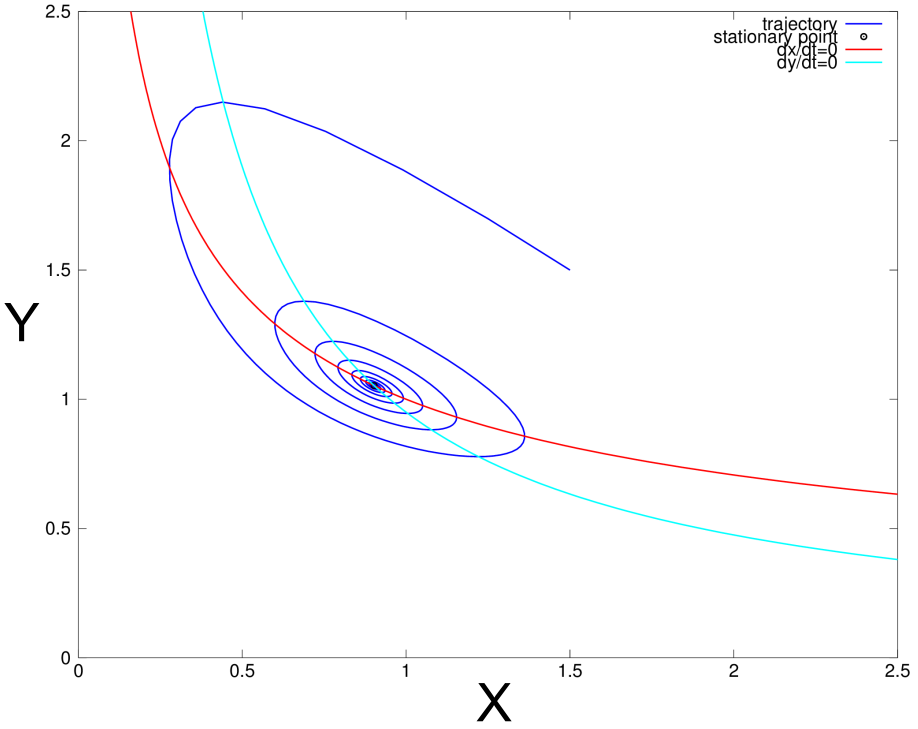
nullclines: $\dot{X} = 0 \;\Leftrightarrow\; X = \dfrac{v_0}{k_1 \cdot Y^2}\,;\qquad \dot{Y} = 0 \;\Leftrightarrow\; Y = \dfrac{k_2}{k_1 \cdot X}$

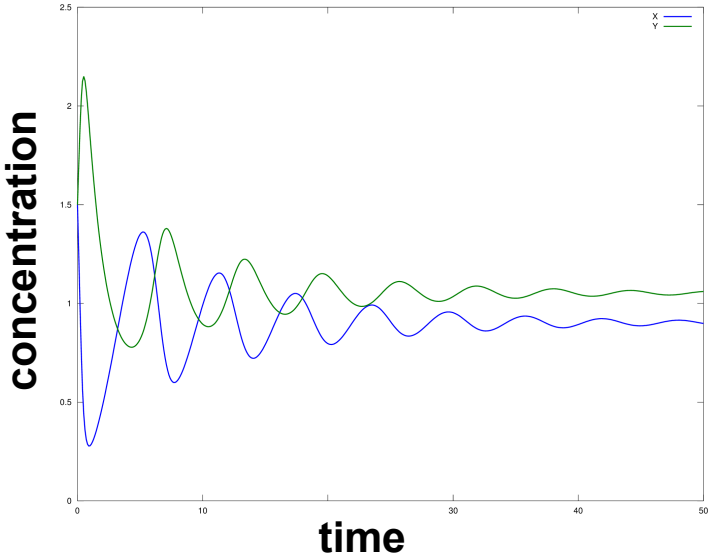
steady-state: $\dot{X} = 0 \;\text{ and }\; \dot{Y} = 0 \;\Leftrightarrow\; \bar{Y} = \dfrac{v_0}{k_2}\,;\quad \bar{X} = \dfrac{v_0 \cdot k_2^{\,2}}{k_1}$

$$\dot{X} = v_0 - k_1 \cdot X \cdot Y^2$$
$$\dot{Y} = k_1 \cdot X \cdot Y^2 - k_2 \cdot Y$$

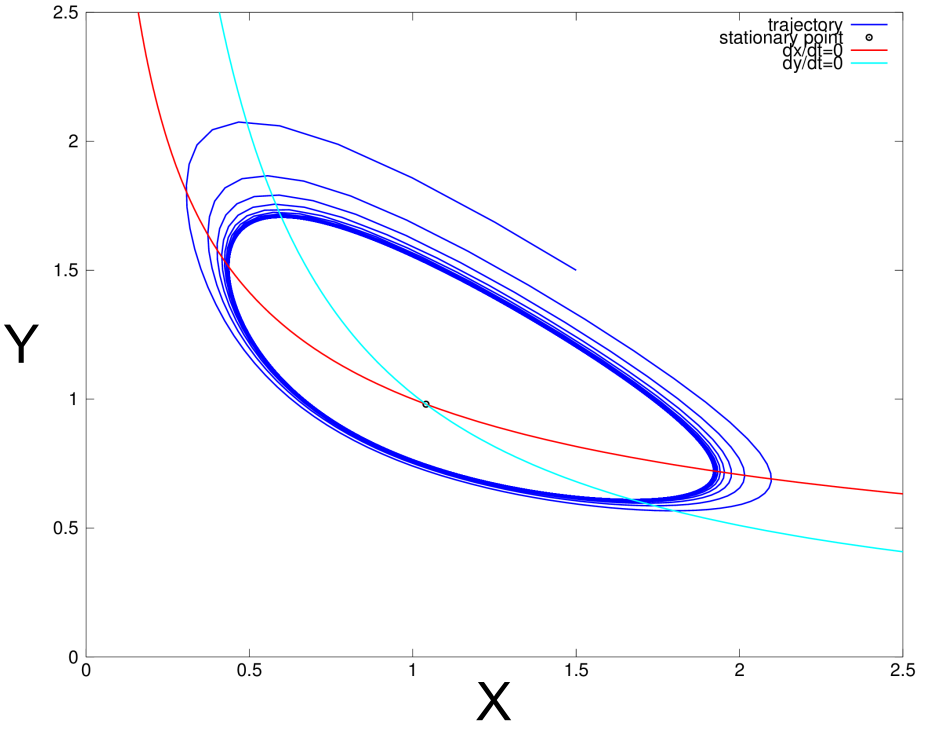
Higgins-Sel'kov oscillator

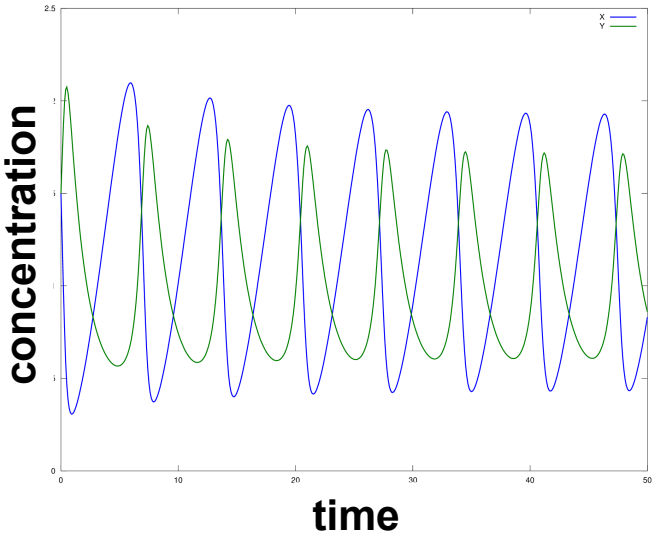
$v_0 = 1;\ k_1 = 1;\ k_2 = 0.95$





$v_0 = 1;\ k_1 = 1;\ k_2 = 1.02$

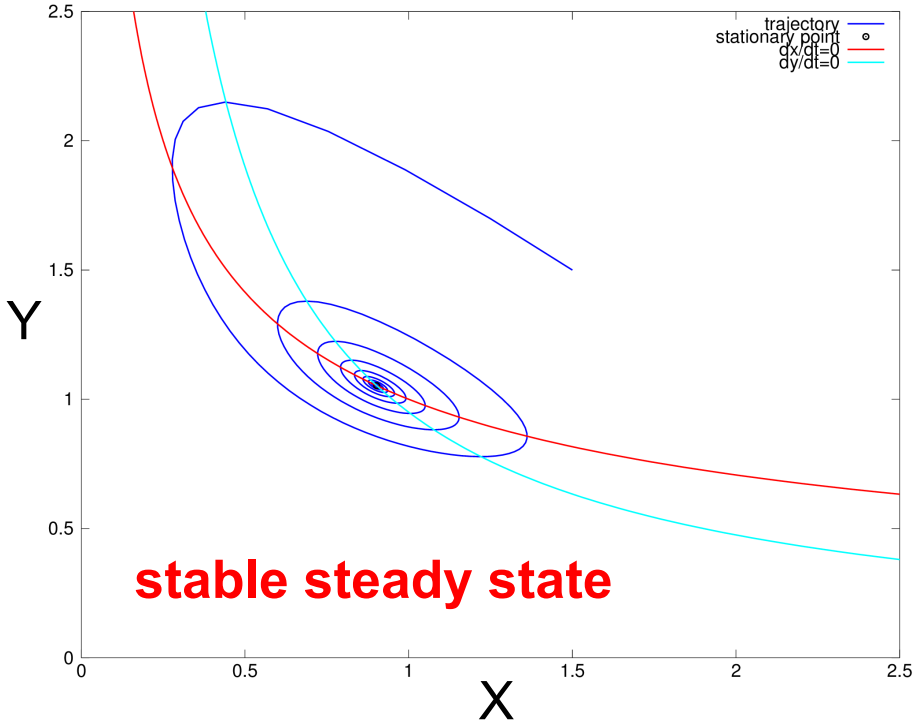




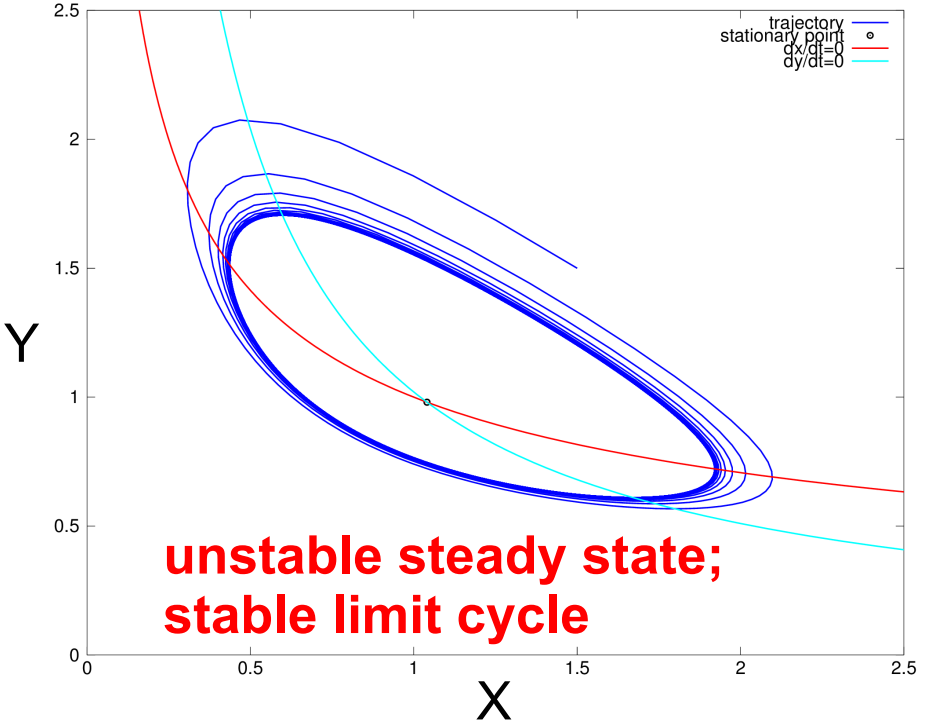
Higgins-Sel'kov oscillator

$$\dot{X} = v_0 - k_1 \cdot X \cdot Y^2$$
$$\dot{Y} = k_1 \cdot X \cdot Y^2 - k_2 \cdot Y$$

$v_0 = 1;\ \ k_1 = 1;\ \ k_2 = 0.95$



$v_0 = 1;\ \ k_1 = 1;\ \ k_2 = 1.02$



nullclines: $\dot{X} = 0 \ \Leftrightarrow\ X = \frac{v_0}{k_1 \cdot Y^2};\quad \dot{Y} = 0 \ \Leftrightarrow\ Y = \frac{k_2}{k_1 \cdot X}$

steady-state: $\dot{X} = 0 \ \text{ and } \ \dot{Y} = 0 \ \Leftrightarrow\ \bar{Y} = \frac{v_0}{k_2};\ \ \bar{X} = \frac{v_0 \cdot {k_2}^2}{k_1}$

Enzyme kinetics

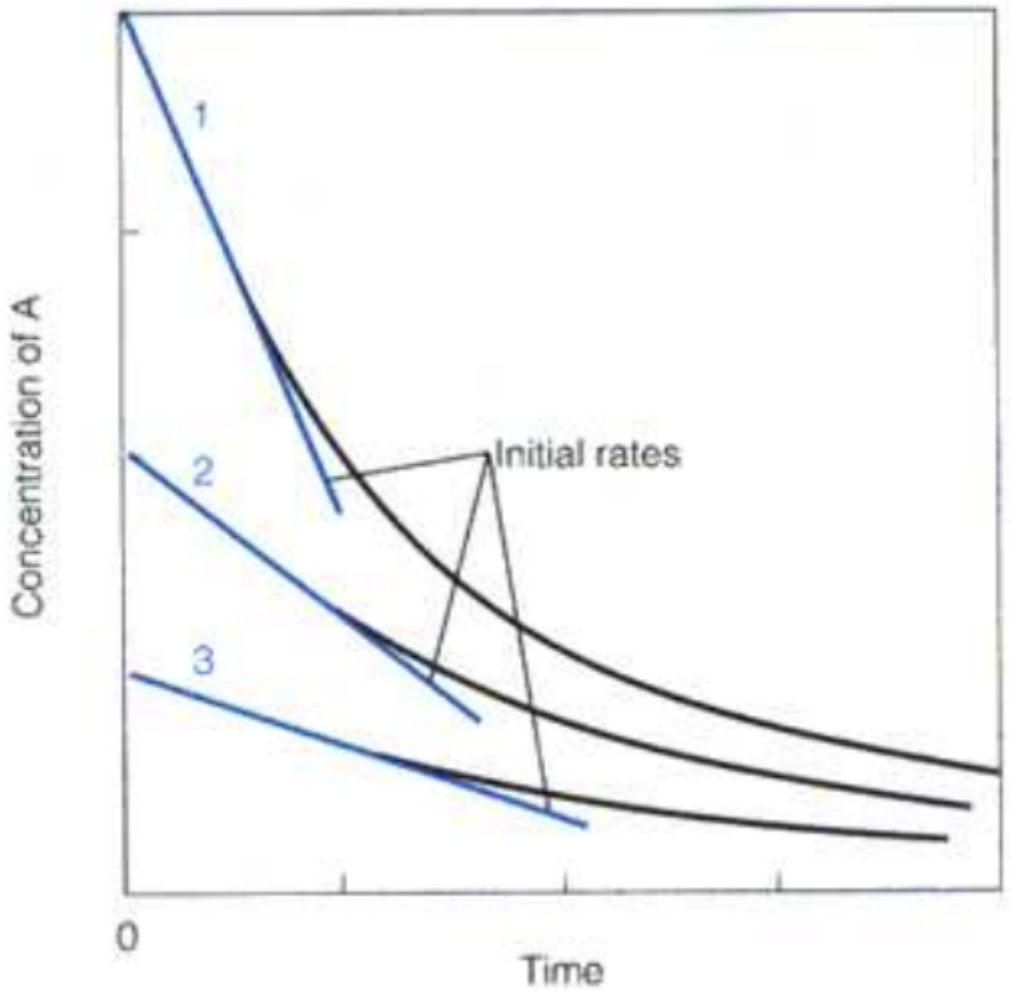
The rate of an enzyme reaction, *v*, is now defined as the change in concentration of product in unit time. Formerly it was defined as the amount, in moles or μmoles, formed per unit time (which is now termed the rate of conversion and used as a basis for the unit of enzyme catalytic activity).

Since there is usually 100% conversion of substrates to products, rates can usually also be measured by use of substrate.

Unless otherwise stated, rates refer to *initial rates*, the instantaneous rate for known concentrations of substrates in the absence of products.

Initial rate measurements

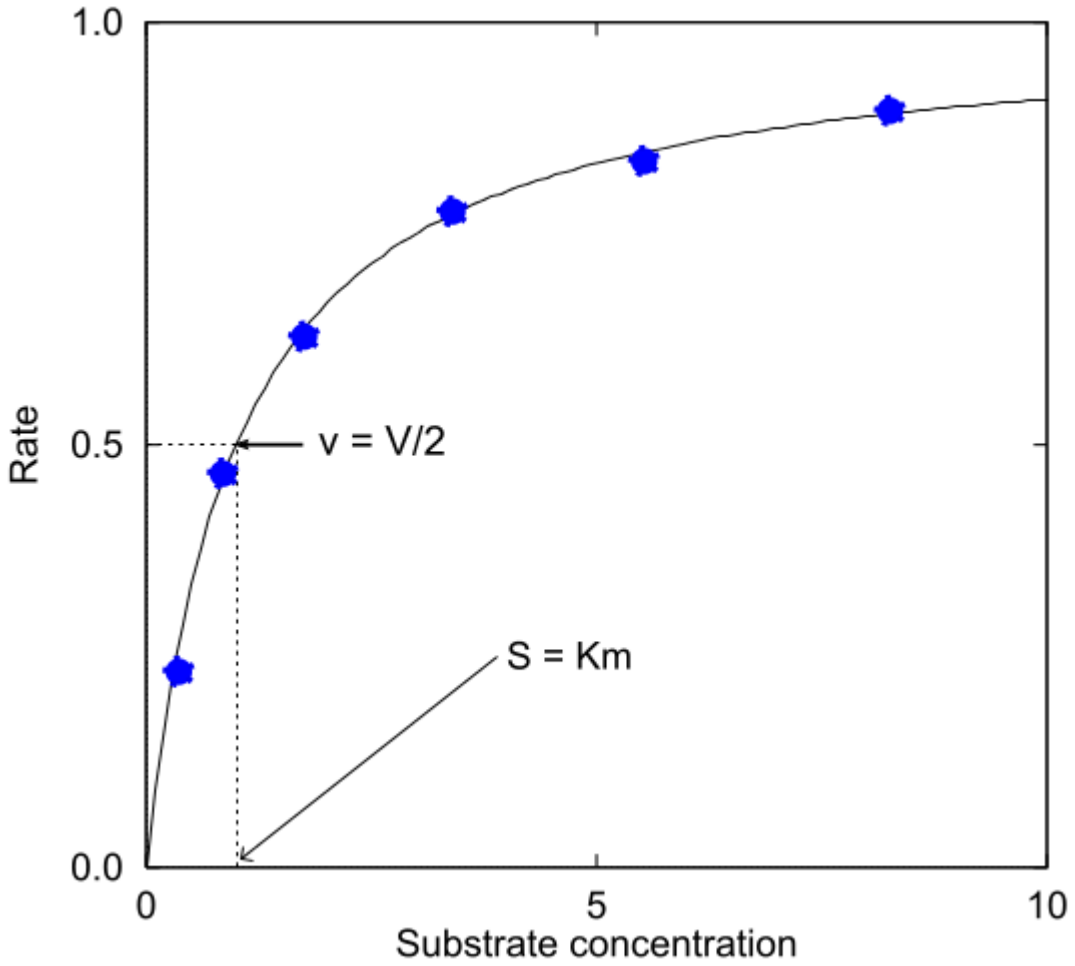
For a reaction S → P , started at different concentrations of S and zero P:



Initial rate measurement is easier with continuous rather than intermittent or spot measurement.

The Michaelis-Menten rate law

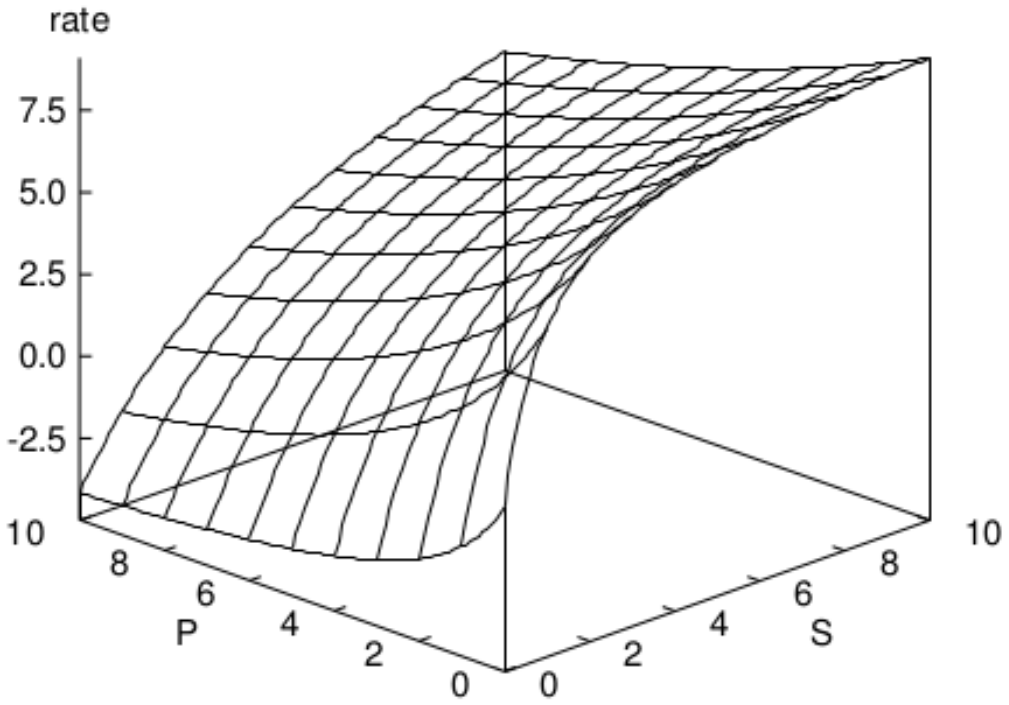
$$v = f(S) = \frac{S \cdot V}{S + K_m}$$



The K_m and V have arbitarily been set to 1, where V is the *limiting rate* (or maximum velocity, V_m) and K_m is the *Michaelis constant*.

The reversible Michaelis-Menten equation

$$v_{\text{net}} = \frac{(V_f / K_{m,S})(S - P/K_{eq})}{1 + S/K_{m,S} + P/K_{m,P}} = f(S, P)$$



Simultaneous dependence of enzyme rate on both substrate and product. The parameters have been set to: $K_{m,S}$ = 1; $V_{m,f}$ = 10; $K_{m,P}$ = 2, and K_{eq} = 4.

Taking the equation apart

The equation is actually composed of two parts:

$$v_{\mathrm{f}} = \frac{\left(V_{\mathrm{f}}/K_{\mathrm{m,S}}\right)\left(S\right)}{1+S/K_{\mathrm{m,S}}+P/K_{\mathrm{m,P}}}$$

and:

$$v_{\mathrm{r}} = \frac{\left(V_{\mathrm{f}}/K_{\mathrm{m,S}}\right)\left(-P/K_{\mathrm{eq}}\right)}{1+S/K_{\mathrm{m,S}}+P/K_{\mathrm{m,P}}}$$

and

$$v_{\mathrm{net}} = v_{\mathrm{f}}+v_{\mathrm{r}}$$

so it is the numerator term that contains the effect of the reverse reaction, whilst the denominator is common.

Taking the equation apart

Looking at the forward term only:

$$v_{\mathrm{f}} = \frac{\left(V_{\mathrm{f}} / K_{\mathrm{m,S}}\right)(S)}{1 + S/K_{\mathrm{m,S}} + P/K_{\mathrm{m,P}}}$$

the equation still contains a term in the product concentration *P*.

This reflects the product inhibition that exists because of its binding at the active site, even when the K_{eq} is so large that the reverse reaction rate v_r is very small.

Taking the equation apart

Considering the reverse term only:

$$v_r = \frac{(V_f / K_{m,S})(-P / K_{eq})}{1 + S / K_{m,S} + P / K_{m,P}}$$

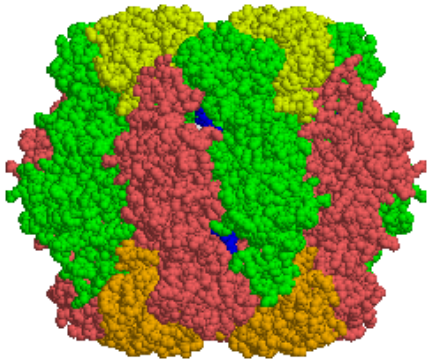
This could also be written as the forward component of the equation written for P → S:

$$v_r = - \frac{(V_r / K_{m,P})(P)}{1 + S / K_{m,S} + P / K_{m,P}}$$

which shows that:

$$K_{eq} = \frac{V_f}{K_{m,S}} \cdot \frac{K_{m,P}}{V_r}$$

This is the *Haldane relationship*, showing that it suffices to know three of the four parameters provided the K_{eq} is known.



EC 4.1.1.39

Enzyme kinetics of RuBisCO

$$RuBP + CO_2 \longrightarrow 2\ PGA$$

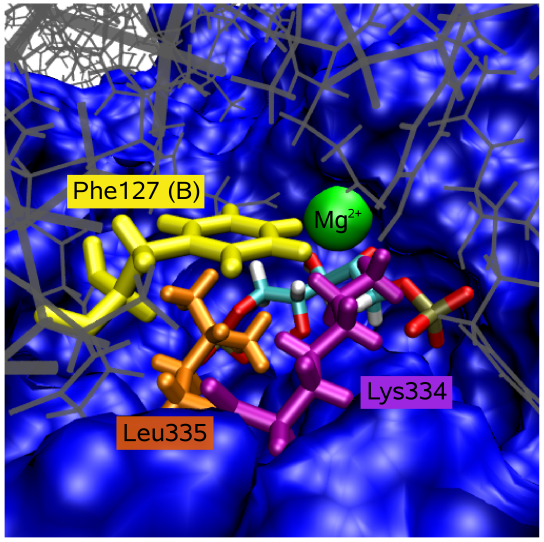
- Very important
- Very big (8L8S subunits, total ~540kDa)
- Very slow (k_{cat} ~ 3/s)
- Very strange

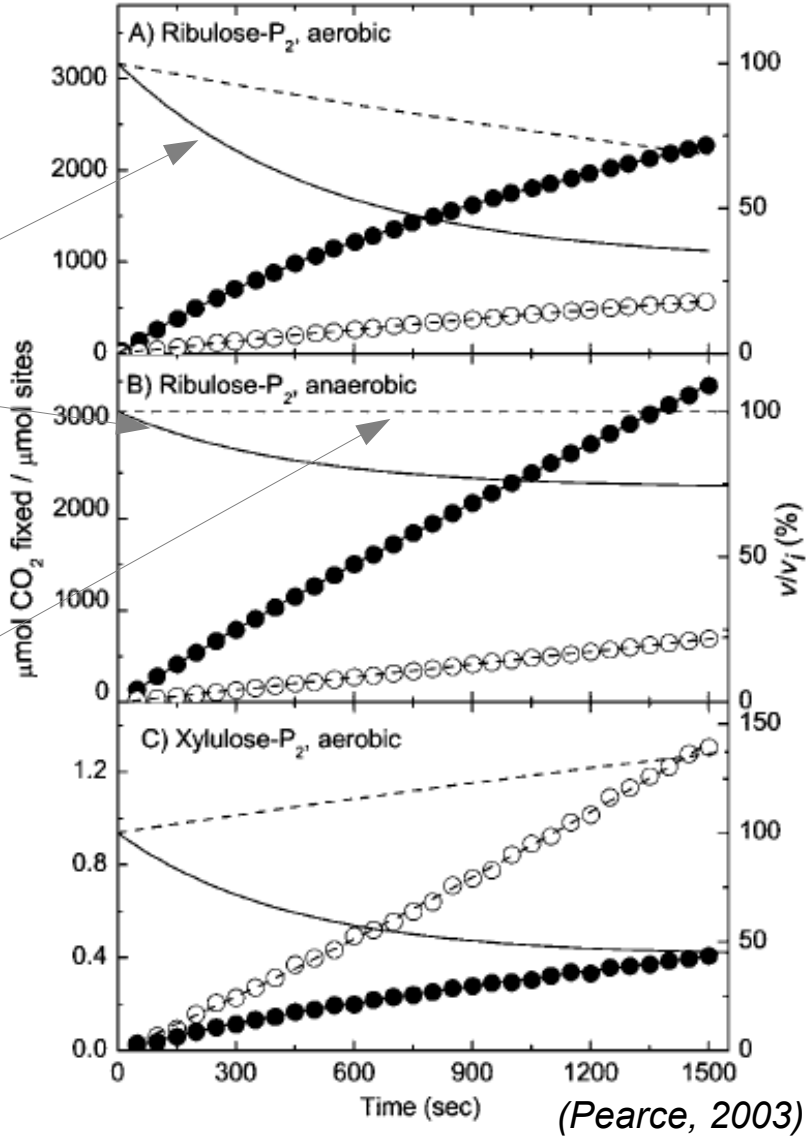
FALLOVER

activity decline by self-formed inhibitor

mutant (L335V) does not display fallover

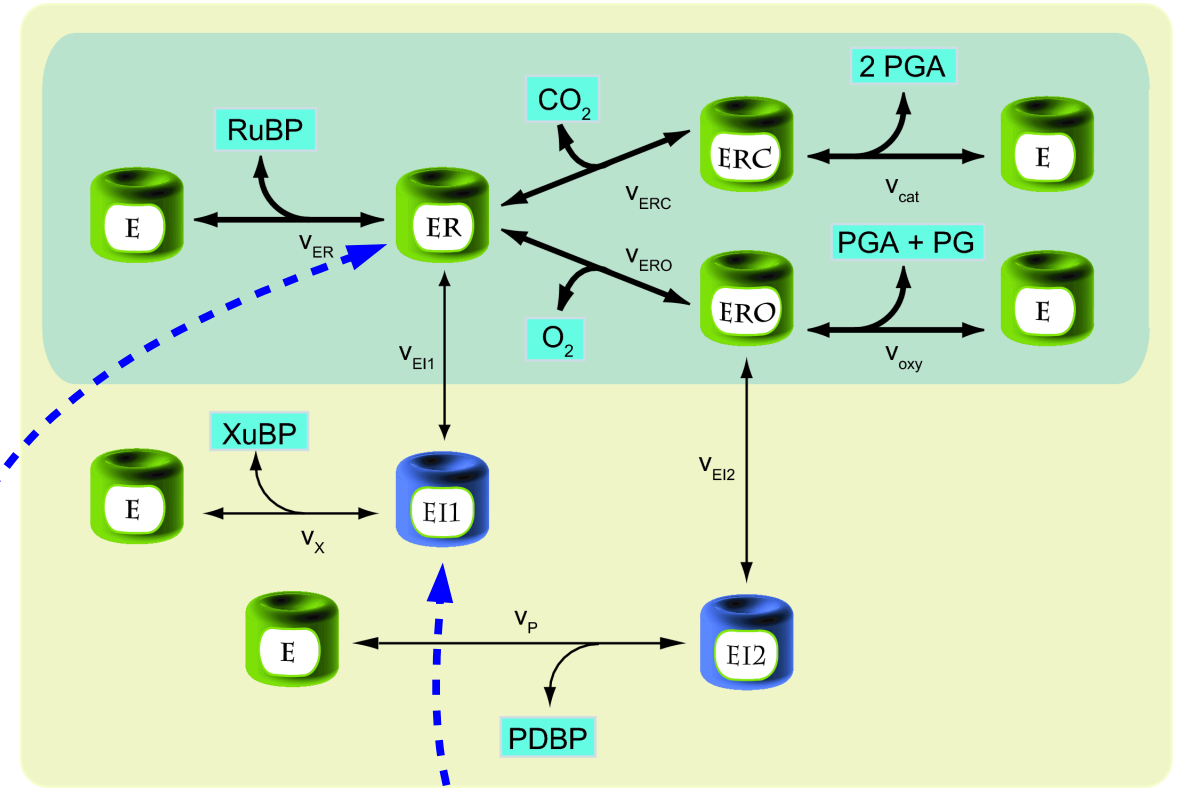
XuBP is poor substrate with strange kinetics

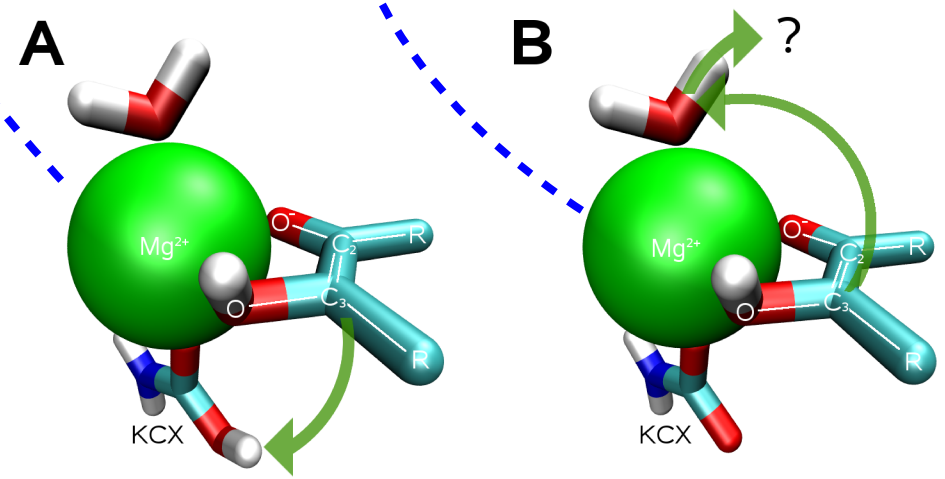




(Pearce, 2003)

Modelling RuBisCO

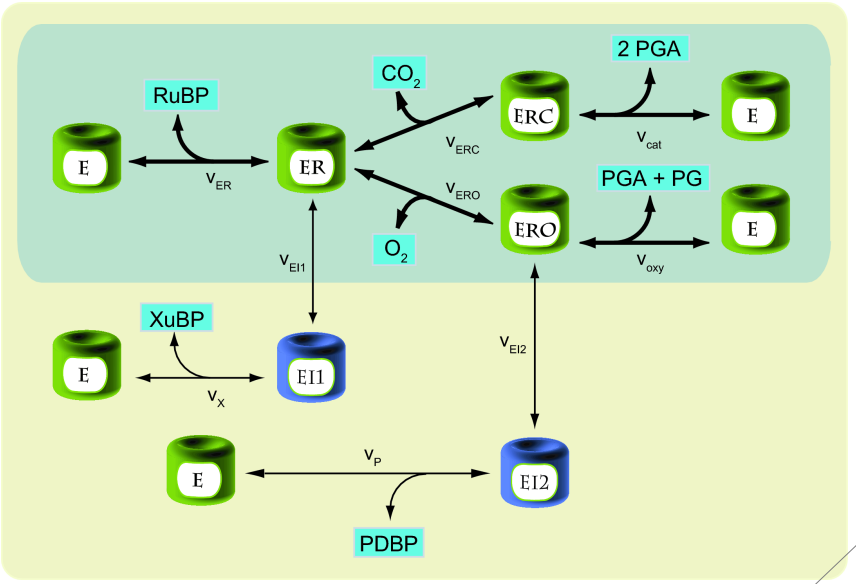




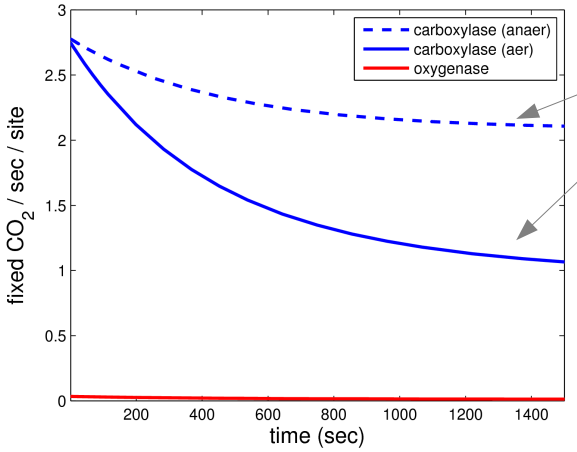
(Witzel et al, 2010, FEBS J)

Modelling RuBisCO

'loop 6' mutant



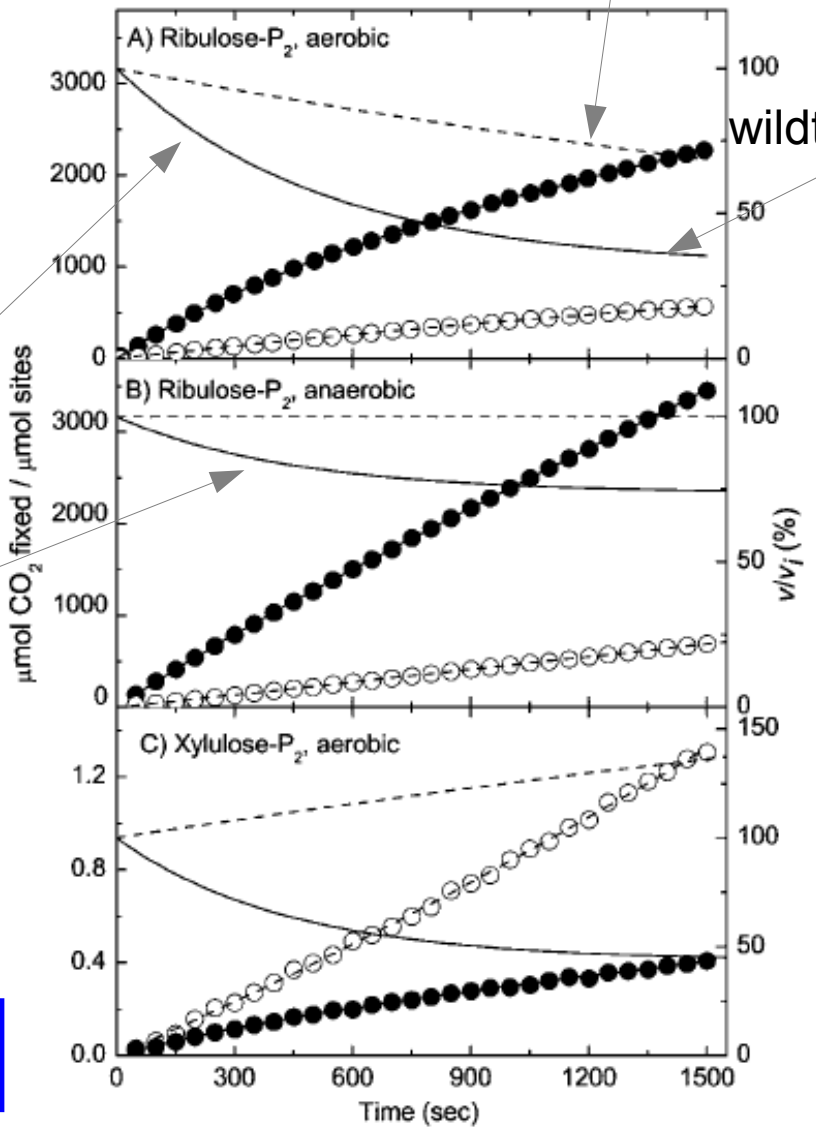
simulation



comparison with experiments

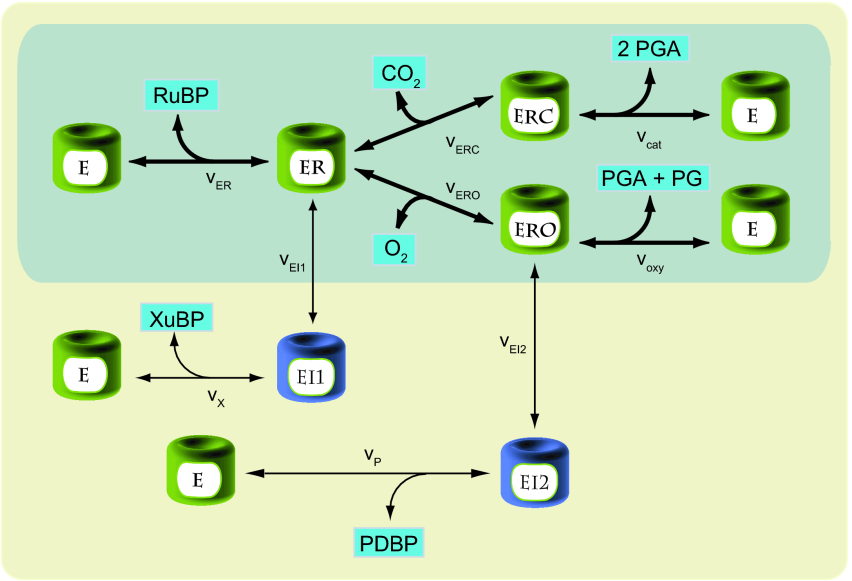
OK FOR WILDTYPE ON RuBP

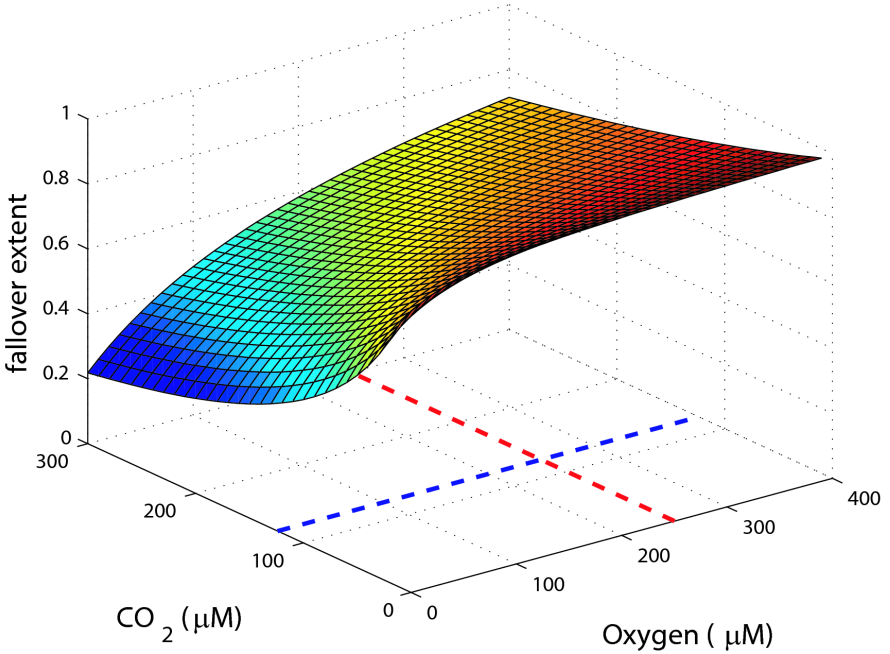
wildtype



(Pearce, 2003)

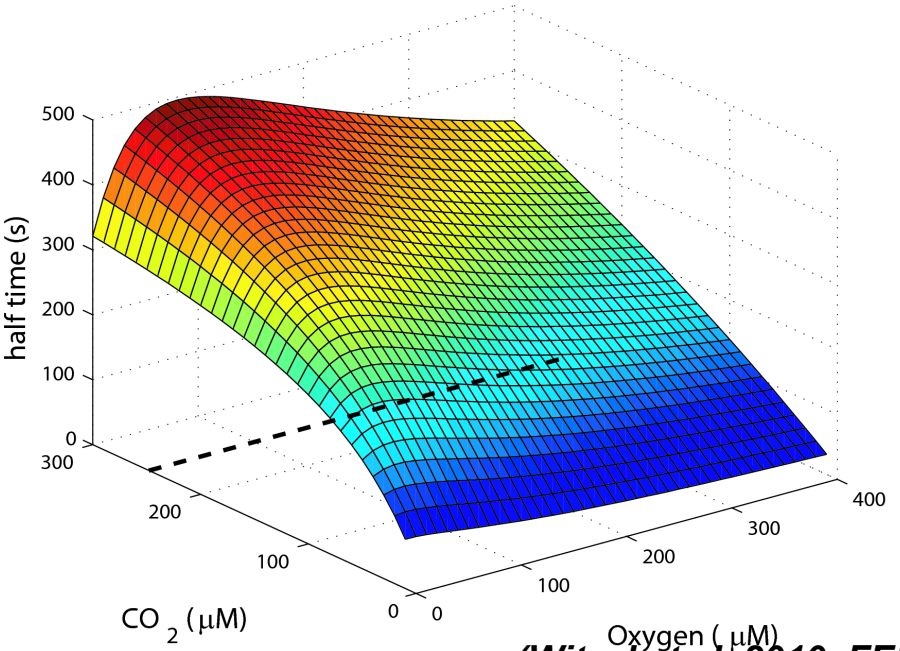
Modelling RuBisCO





A model provides a general description

Simulation of a wide range of conditions possible

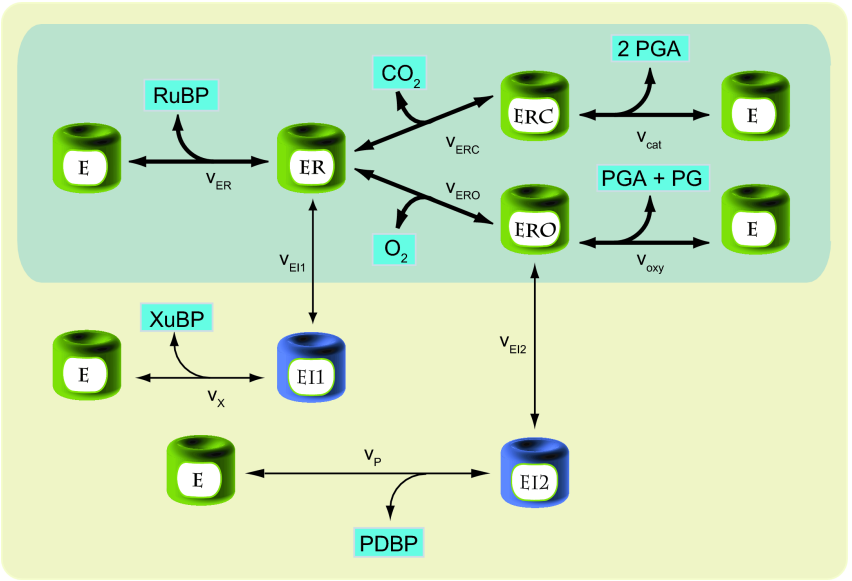


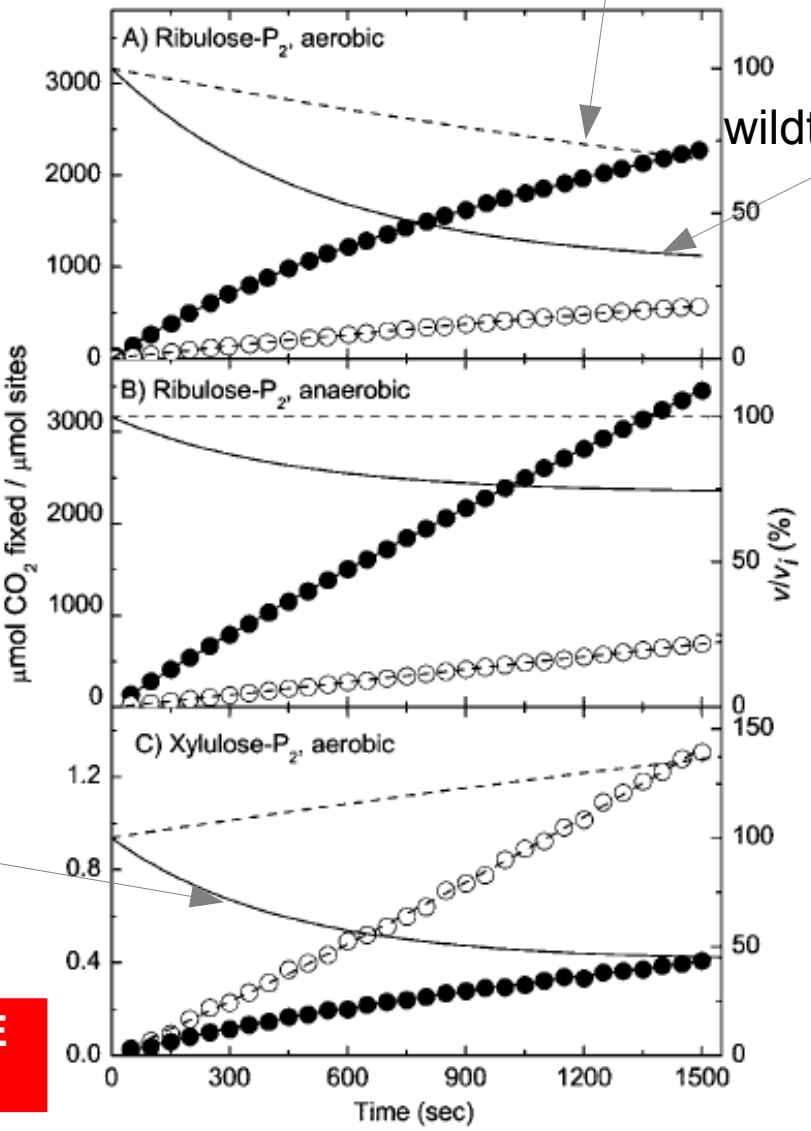
(Witzel et al. 2010, FEBS J)

Modelling RuBisCO

'loop 6' mutant

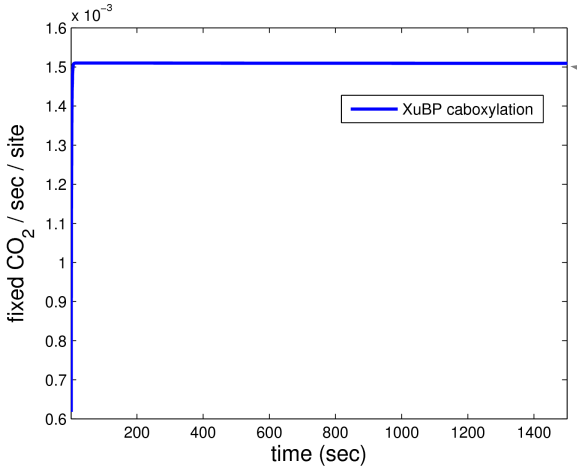
wildtype





(Pearce, 2003)

simulation

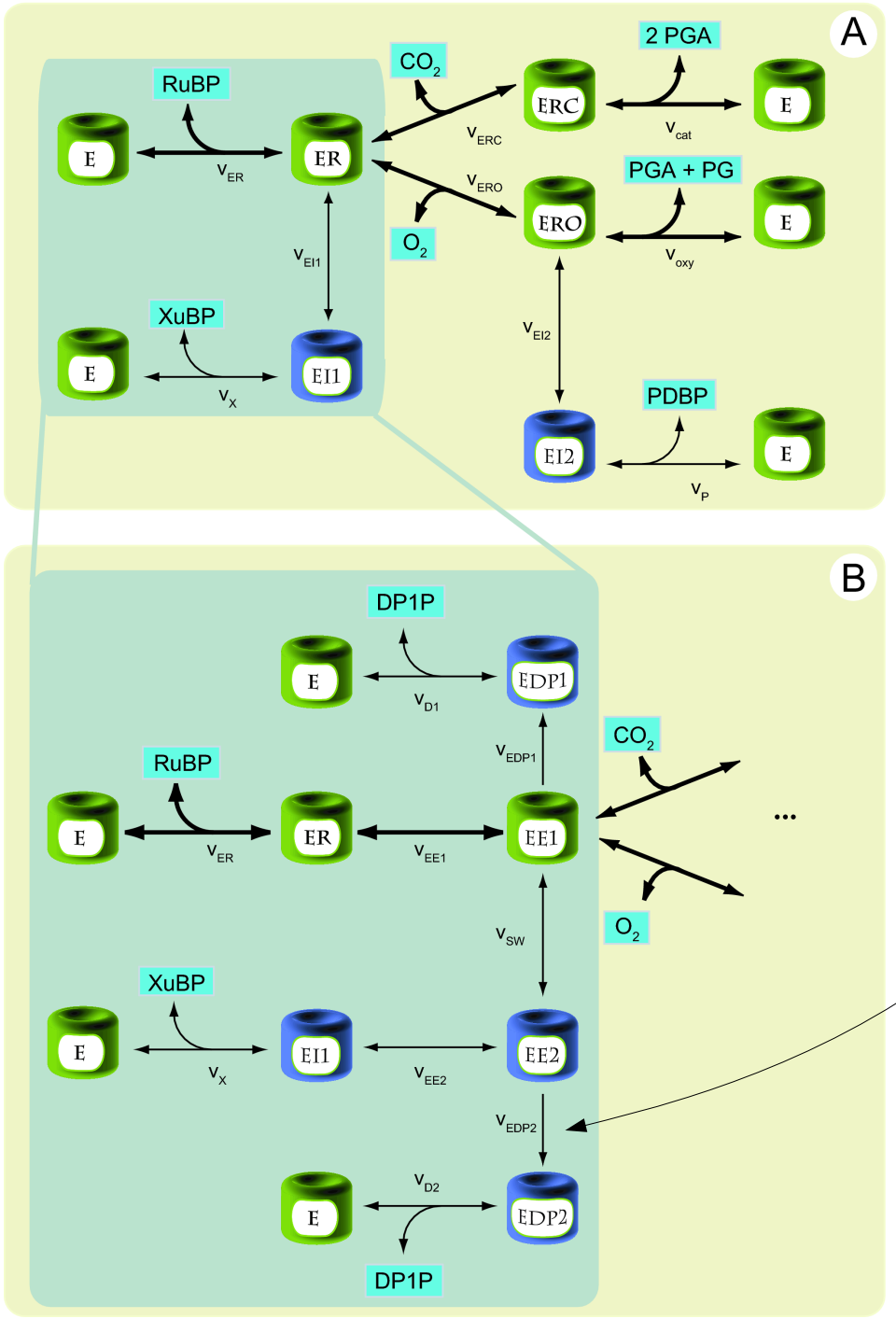


comparison with experiments

NOT OK FOR WILDTYPE ON XuBP

⇨ Improvement needed!

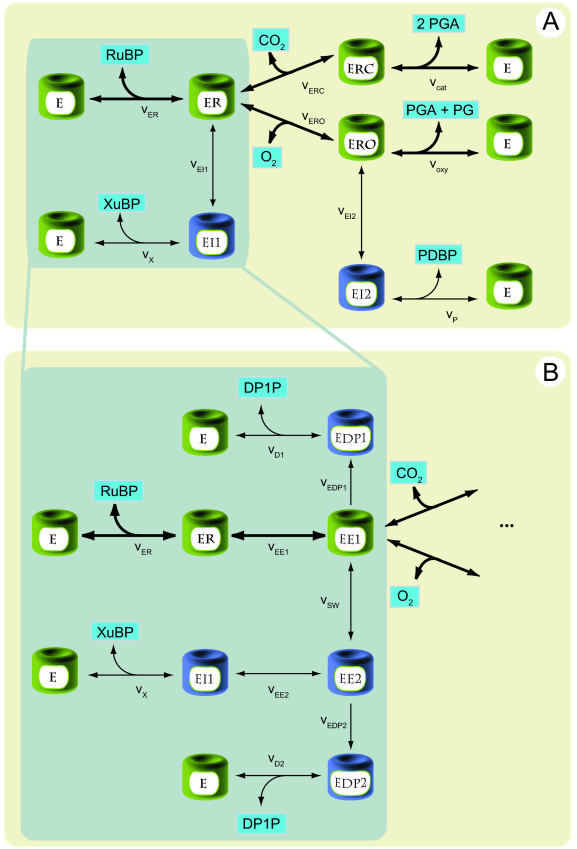
Adding more detail

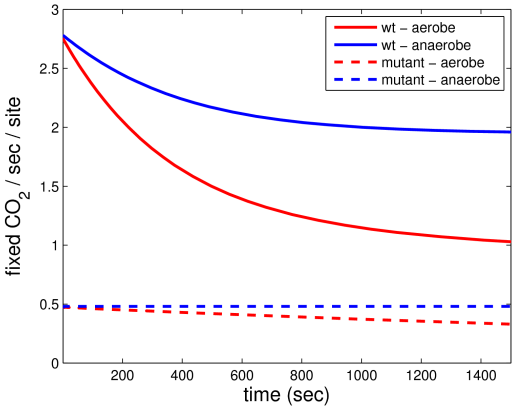


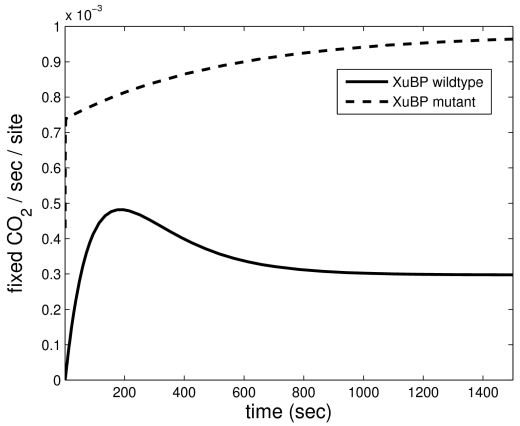
another postulated inhibitor (DP1P)

must be formed from
the XuBP branch

Modelling RuBisCO

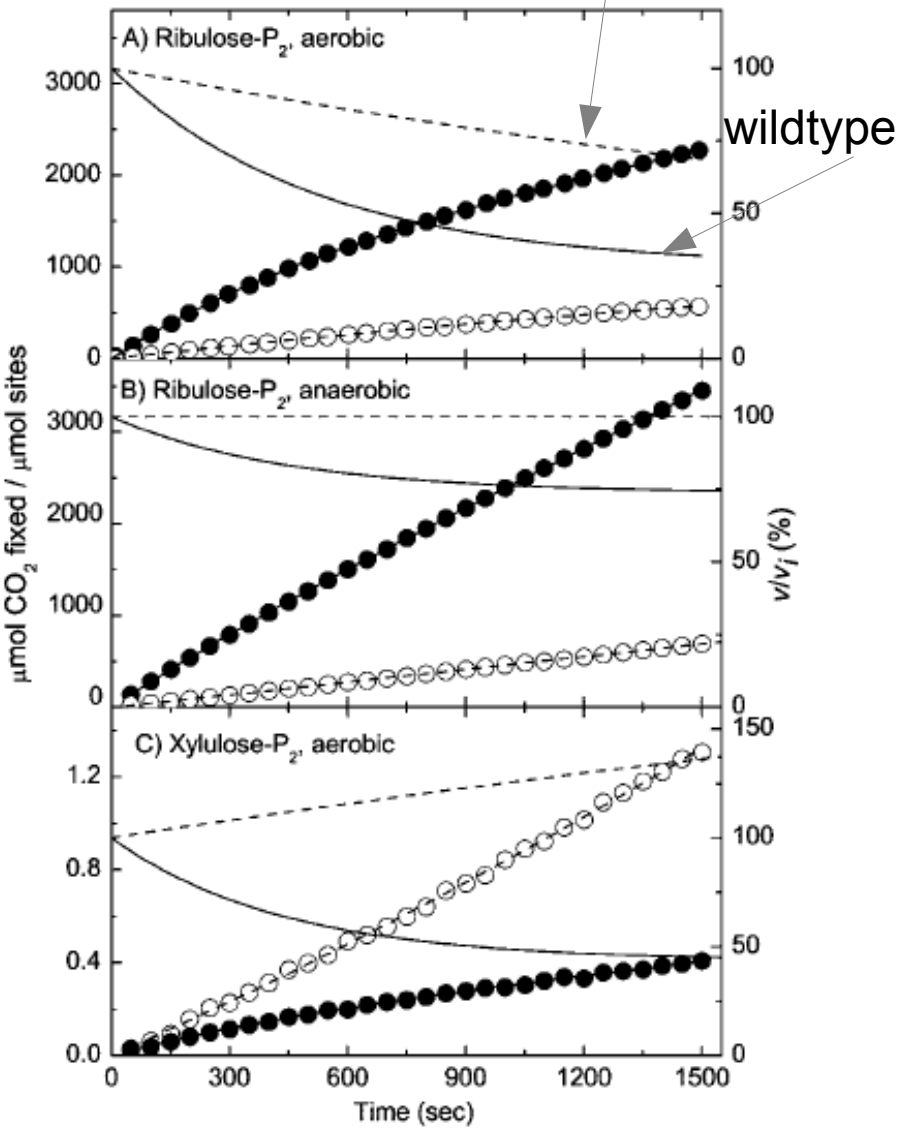






'loop 6' mutant

wildtype



(Pearce, 2003)

Elaborated model explains wildtype and mutant behaviour on both substrates (RuBP and XuBP)

Modeling RuBisCO – Summary

- Model provides a unifying theory for RuBisCO dynamics
- can quantitatively describe the fallover phenomenon
- can describe various types of RuBisCO
- associates specific parameters with observed quantities

Open questions

- Are RuBisCOs optimally adapted to their specific intracellular environment?
  (as suggested by Tcherkez et al., 2006)
- What is the specific role of RuBisCO activase?

Goals

- Find realistic parameters for different types of RuBisCO
- Understand which pressures have led to the evolution of these parameter sets
- Find advantages / disadvantages of RuBisCO / RuBisCO activase system

Open questions

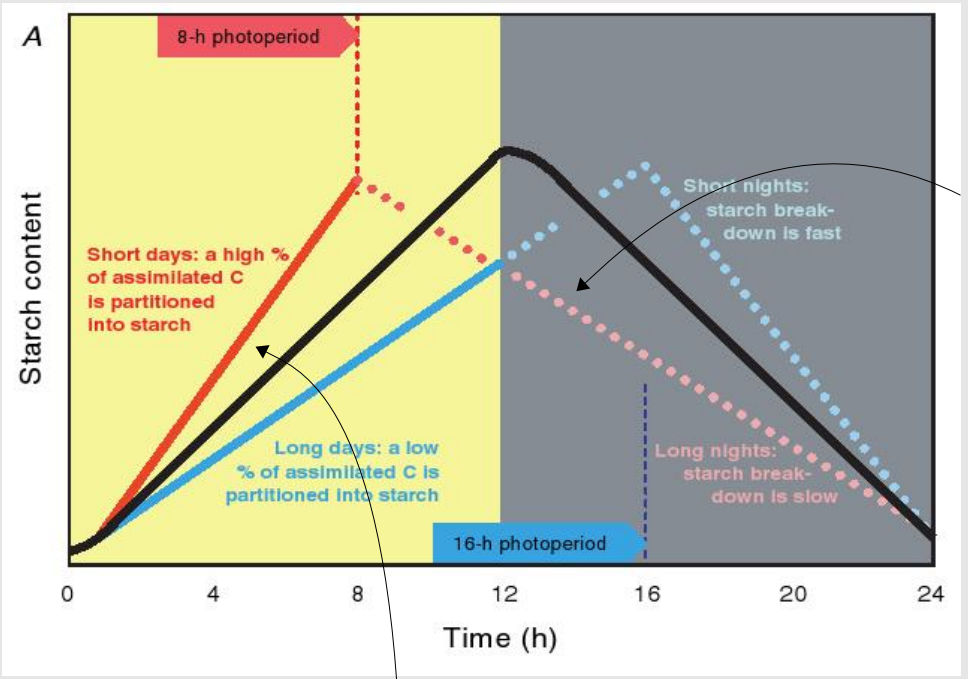
?

?

How does the clock 'tell' expected length of day/night?

What measures the starch content?

?



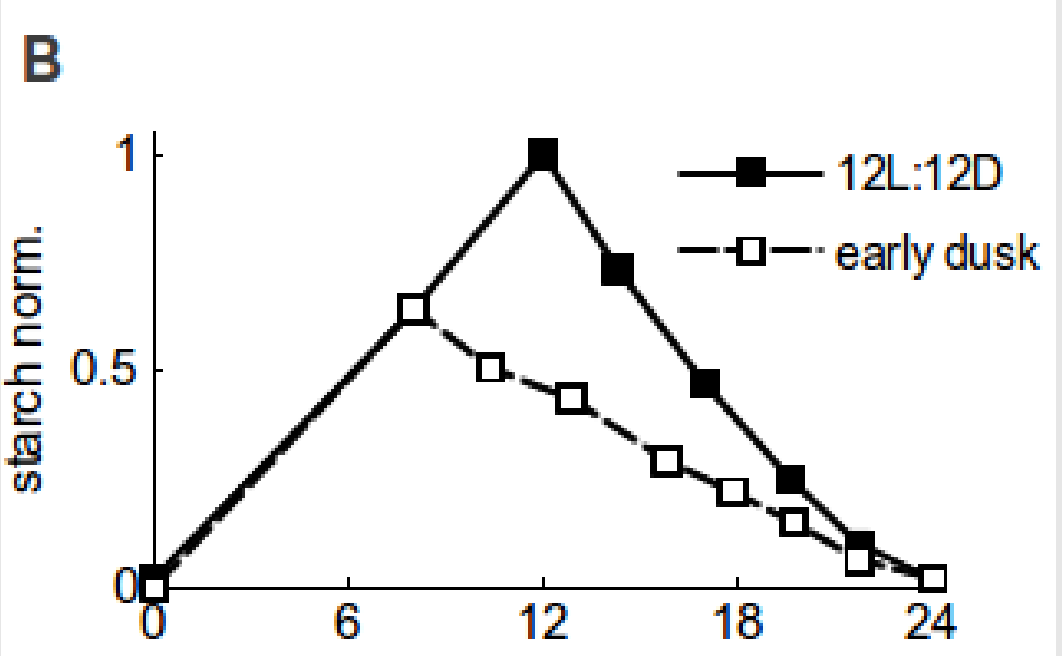
How is the correct breakdown rate 'calculated'?

How is carbon partitioning controlled?

TiMet

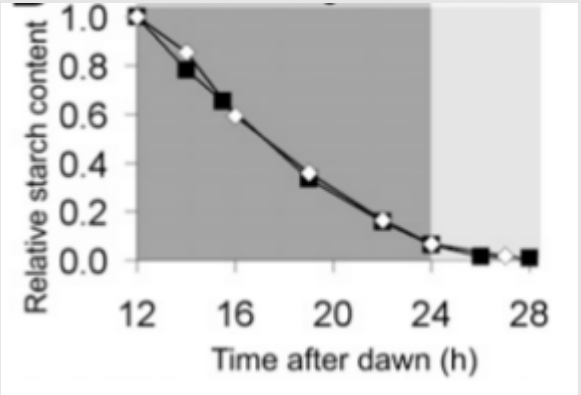
...even more mysteries...

The 'early dusk' experiment by Alexander Graf,
(Graf et al 2010, PNAS)



Even when 'surprised' by a 4 hour shorter day, plants 'know' what to do!

The circadian clock is apparently important, because:



Plants cannot adapt to T-cycles different than 24h!

Building a mathematical model

Known:

- Metabolism
- Circadian clock

Unknown:

- Regulation of starch synthesis
- Regulation of starch breakdown
- How is starch content measured?

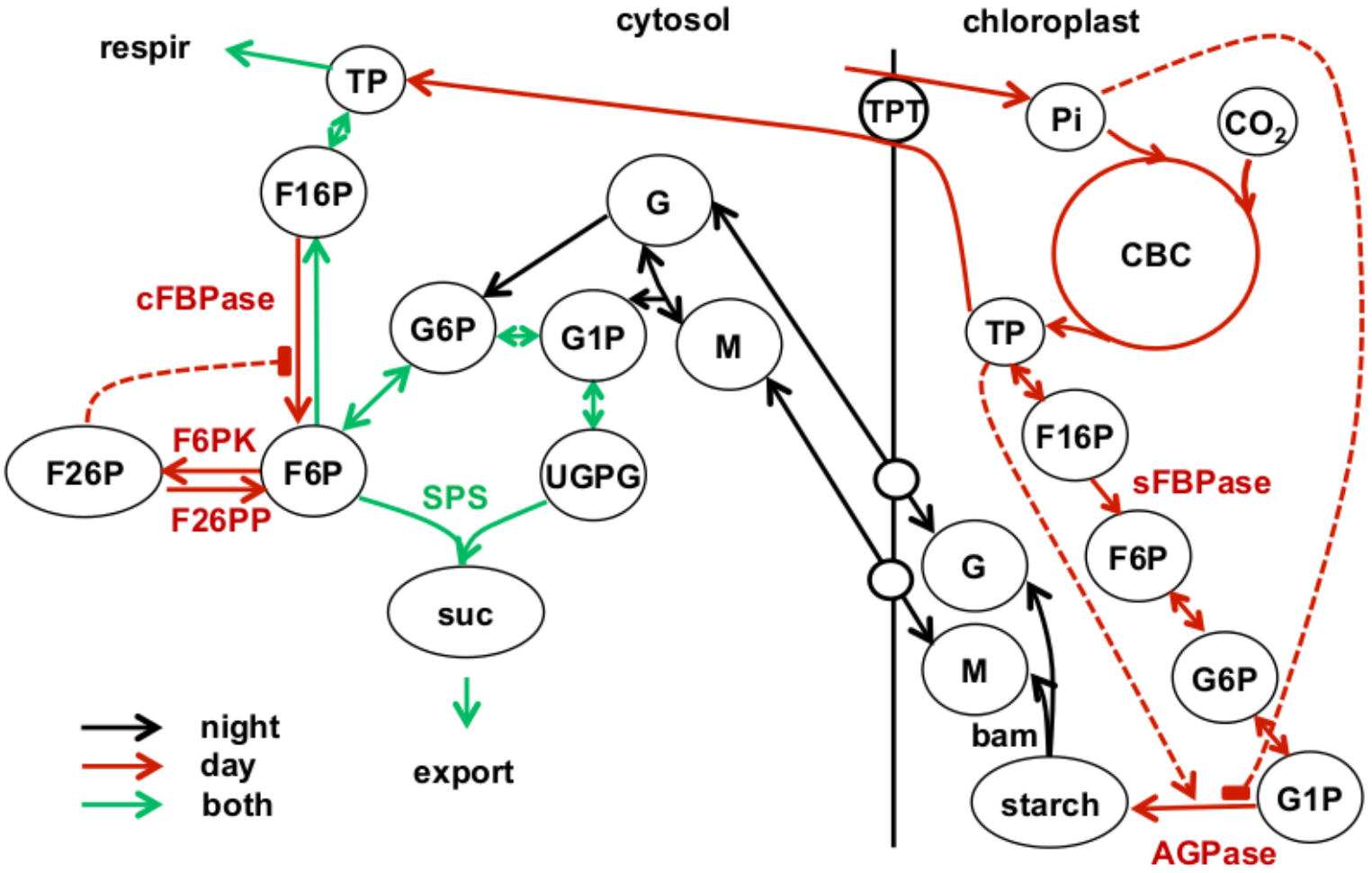
<u>Challenges:</u>

1. The model must combine known systems with plausible, but hypothesised regulatory mechanisms

2. To keep the model tractable, we need to find a compromise between detailedness and simplification

Pokhilko et al, 2014, *Mol BioSyst*

Key metabolic pathways in 'source' leafs



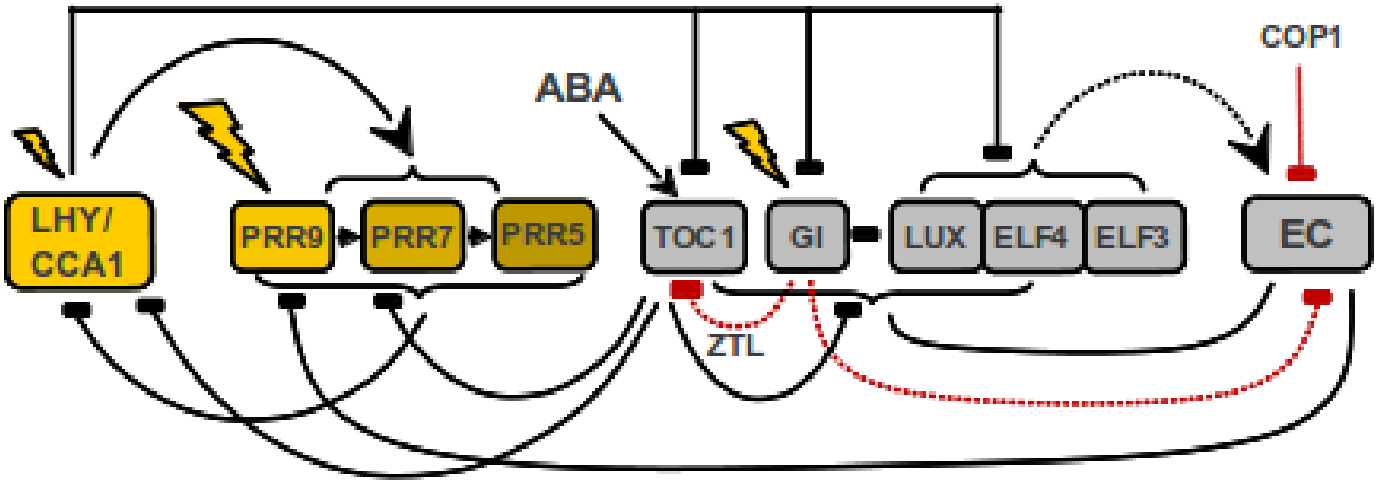
- Carbon fixation
- Starch synthesis
- Starch breakdown
- Sucrose synthesis
- Sucrose export

Include key steps but simplify pathways!

The circadian clock

Done before!

⇨ Use published and validated model!



Pokhilko et al., 2011, Mol Syst Biol

Model assumptions (postulates)

1. Key sensors:

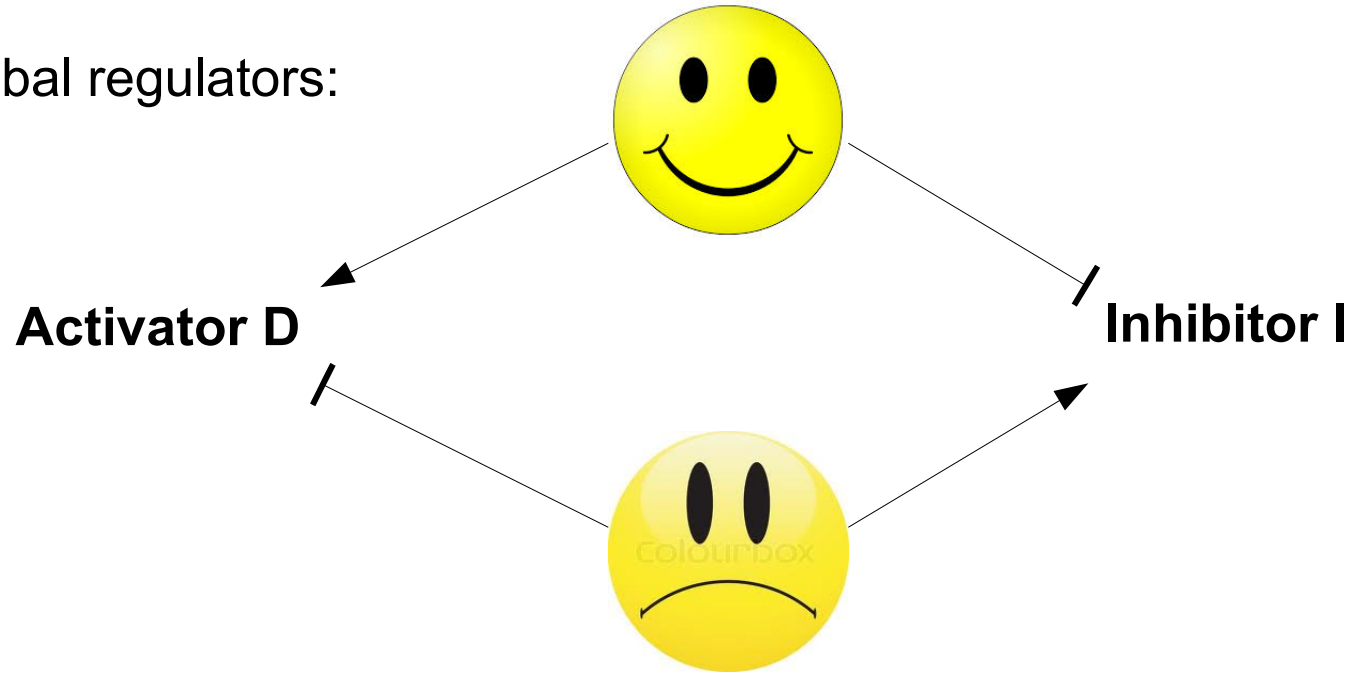
Timer α

time-to-dawn

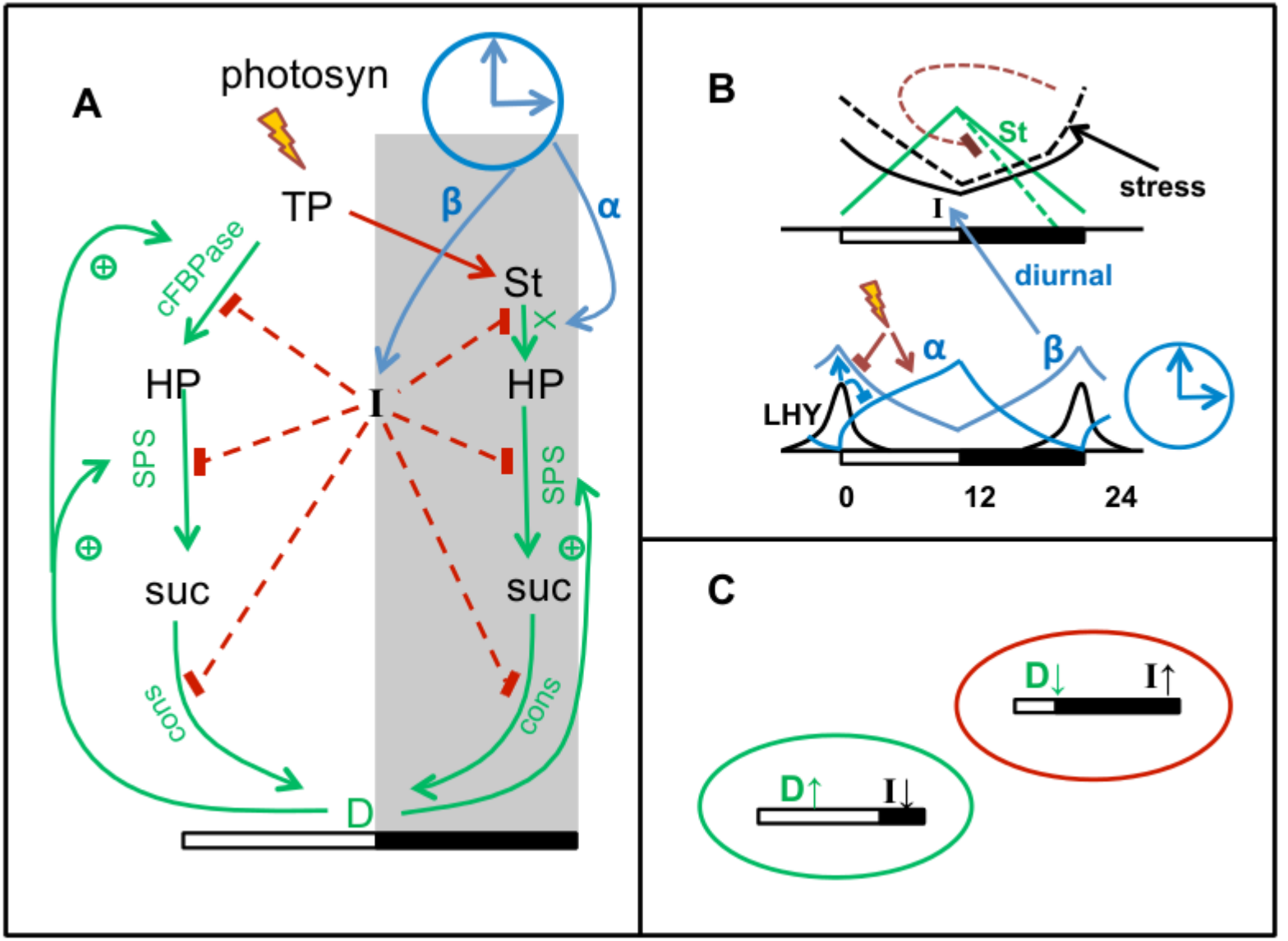
dark sensor β

carbon limitation

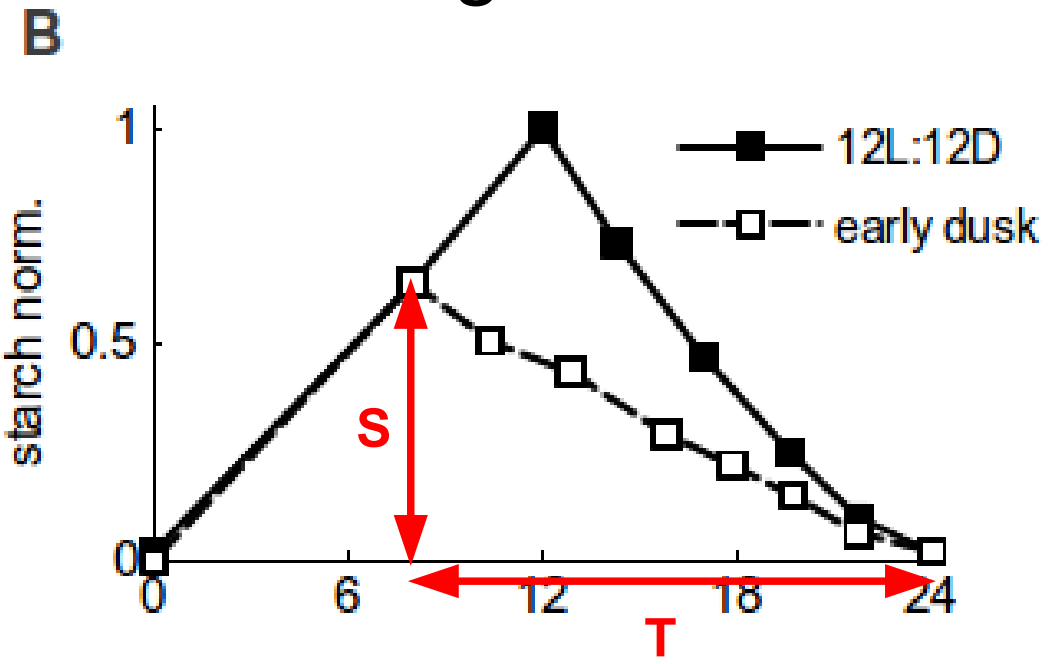
2. Global regulators:



Regulatory principles



Regulation of starch degradation



Arithmetic division

$$v = \frac{S}{T}$$

Simplest solution:

Auxiliary compound *X* (e.g. active form of starch degrading enzyme):

$$\frac{dX}{dt} = k_1 S - k_2 X T$$

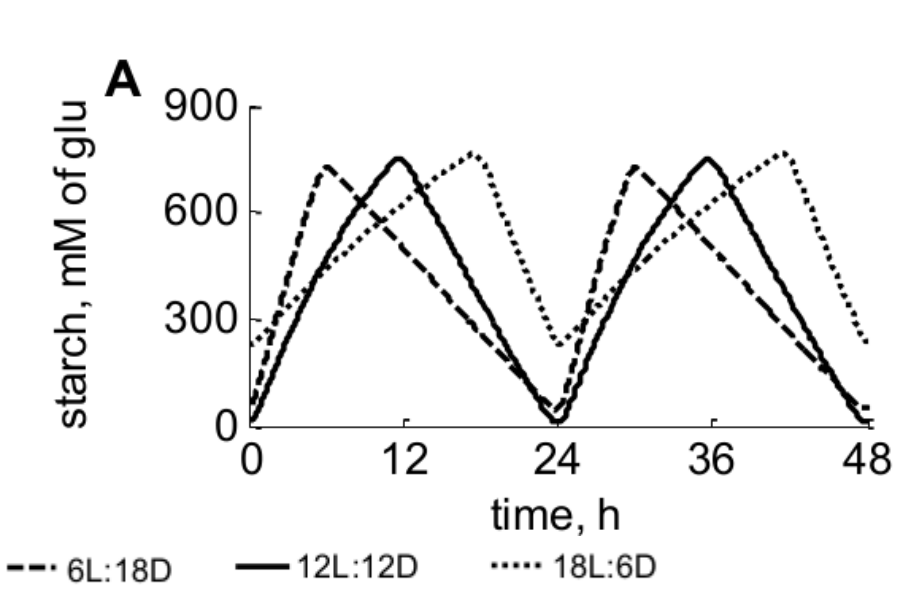
Rapid activation/deactivation: $\frac{dX}{dt} = 0 \Leftrightarrow X = \frac{k_1}{k_2} \cdot \frac{S}{T}$

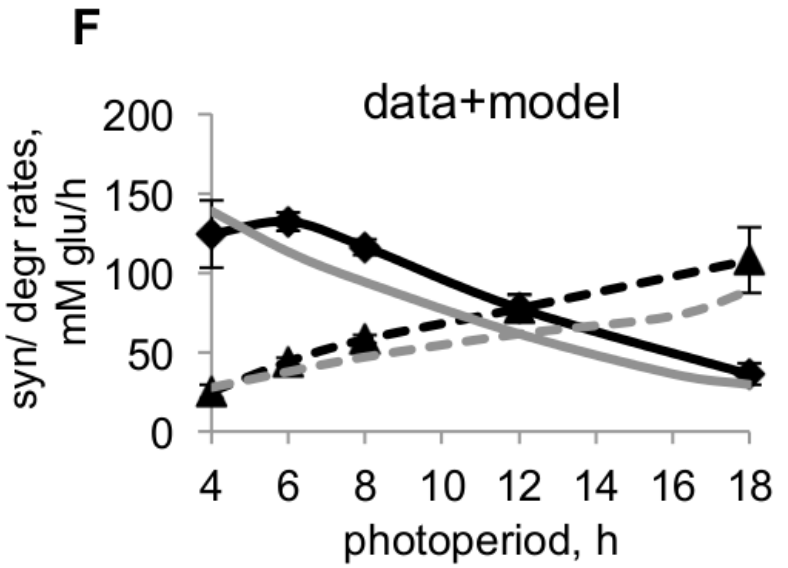
Scaldione et al (2013), eLife: **Arabidopsis plants perform arithmetic division to prevent starvation at night**

Seaton et al (2013), J Roy Soc Interface:
Regulatory principles and experimental approaches to the circadian control of starch turnover

Simulations wild-type

Regulatory principles allow to explain wild-type starch turnover under various photoperiods

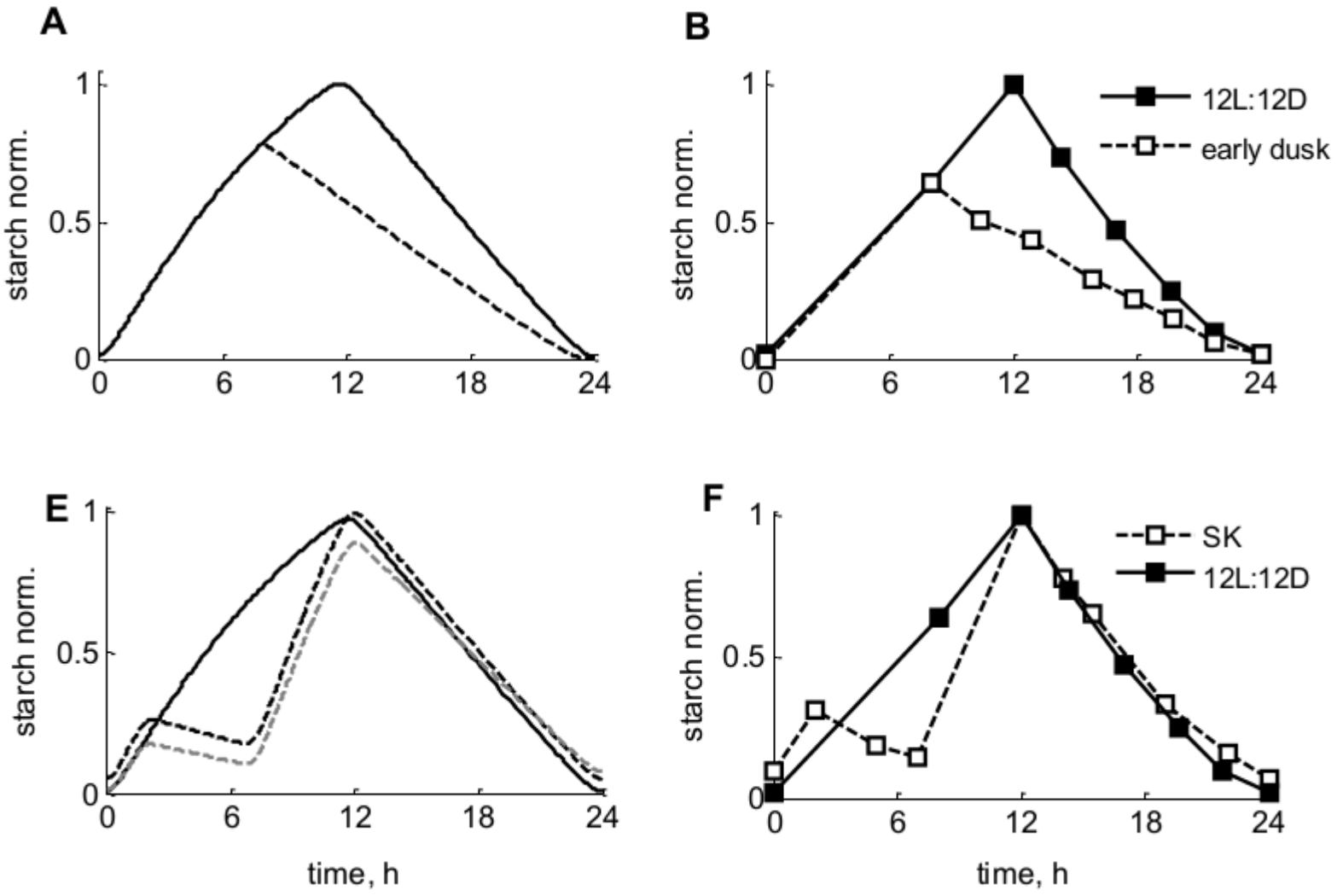




Other light protocols

Early dusk

Skeleton

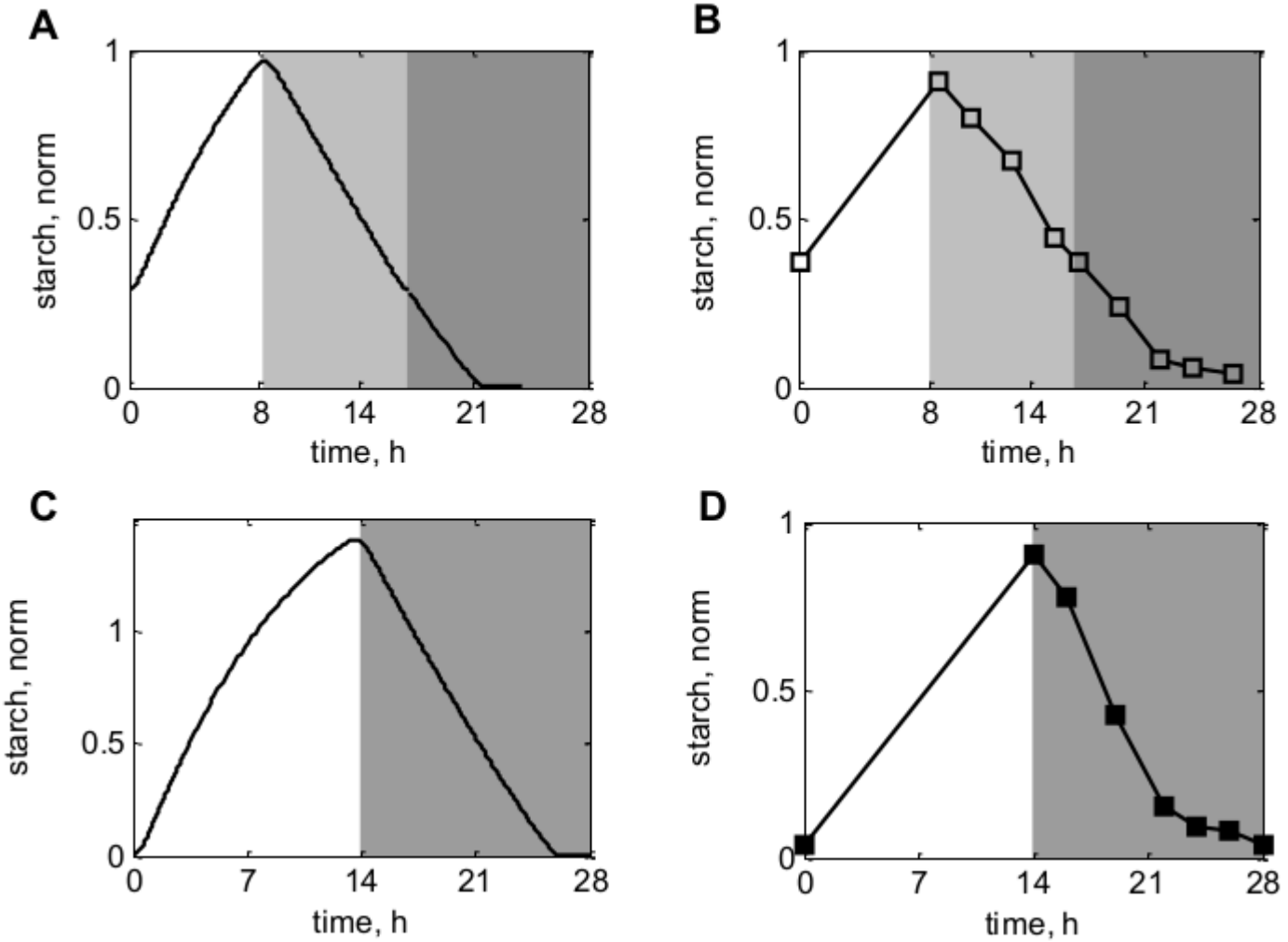


Good quantitative agreement between simulation and experiment!

Other light protocols

17h T-cycle

28h T-cycle



Good quantitative agreement between simulation and experiment!

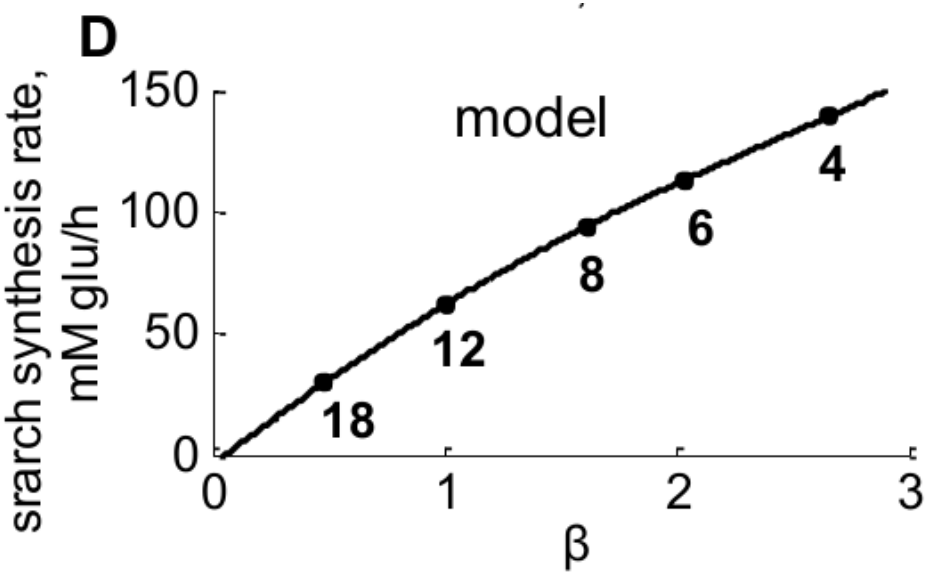
What are the unknown components?

Model allows to make predictions of their behaviour

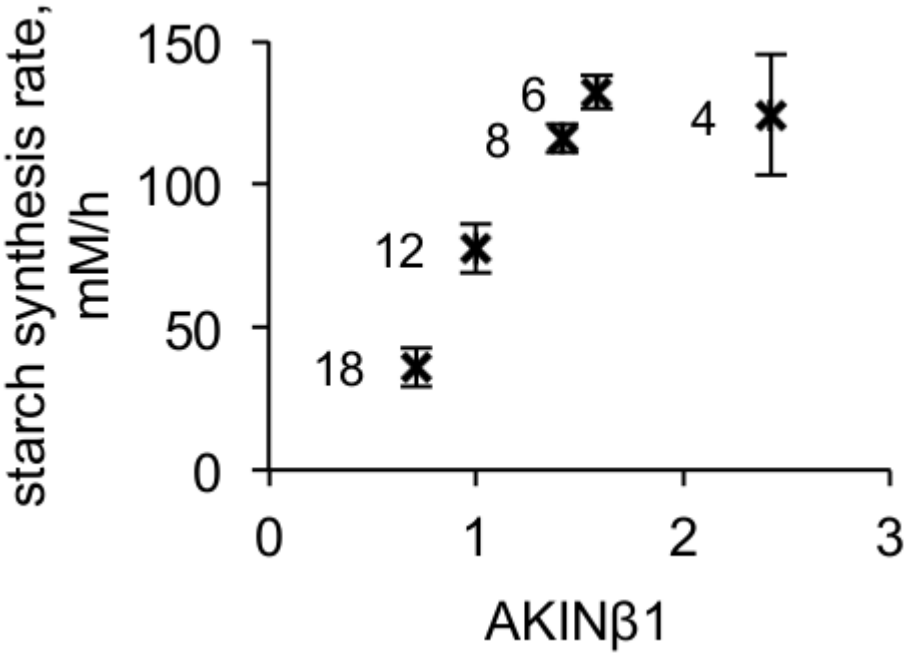
⇨ Helps to identify candidates from expression / proteomics data

For example, the component β:

Predicted peak-levels at dawn



Microarray data for β-subunit of SNRK1



Promotor structure also supports AKINβ1 as good candidate for β

In the tutorial we will play with a model of the Calvin Cycle