Numerical Enzyme Kinetics

using **DynaFit** software

Petr Kuzmič, Ph.D.
BioKin, Ltd.

BioKin L t d.

Statement of the problem

There are **no traditional (algebraic) rate equations** for many important cases:

- **Time-dependent** inhibition in the general case
  - substrate depletion
  - enzyme deactivation
- **Tight binding** inhibition in the general case
  - impurities in inhibitors
  - dissociative enzymes
- **Auto-activation** inhibition
  - e.g., protein kinases
- *Many other practically useful situations.*

Solution

- Abandon traditional algebraic formalism of enzyme kinetics
- Deploy **numerical (iterative) fitting modes** instead

This is approach is *not new*:	*SOFTWARE*
D. Garfinkel (1960's - 1970's)	BIOSYM
C. Frieden (1980's - 1990's)	KINSIM
P. Kuzmic (2000's - present)	DynaFit ←
K. Johnson (2010's - present)	Kinetic Explorer

Introduction: A bit of theory

Numerical vs. algebraic mathematical models

FROM A VARIETY OF ALGEBRAIC EQUATIONS TO A UNIFORM SYSTEM OF **DIFFERENTIAL EQUATIONS**

EXAMPLE: Determine the rate constant k_1 and k_{-1} for **A + B** $\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}}$ **AB**

ALGEBRAIC EQUATIONS	DIFFERENTIAL EQUATIONS
$AB_t = AB_{max}(1 - e^{-k_{obs}t})$ $k_{obs} = k_1[B]_i + k_{-1}$	$d[A]/dt = -k_1[A][B] + k_{-1}[AB]$ $d[B]/dt = -k_1[A][B] + k_{-1}[AB]$ $d[AB]/dt = +k_1[A][B] - k_{-1}[AB]$
Applies only when [B] >> [A]	Applies under all conditions

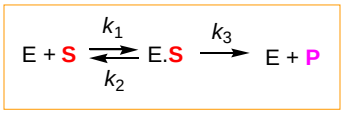
Advantages and disadvantages of numerical models

THERE IS NO SUCH THING AS A FREE LUNCH

ADVANTAGE	ALGEBRAIC MODEL	DIFFERENTIAL MODEL
can be derived for any molecular mechanism	-	+
can be derived automatically by computer	-	+
can be applied under any experimental conditions	-	+
can be evaluated without specialized software	+	-
requires very little computation time	+	-
does not *always* require an initial estimate	+	-
is resistant to truncation and round-off errors	+	-
has a long tradition: many papers published	+	-

A "Kinetic Compiler"

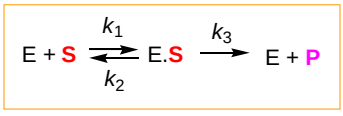
HOW DYNAFIT PROCESSES YOUR BIOCHEMICAL EQUATIONS



Input (plain text file):	Rate terms:	Rate equations:
E + S ---> ES : k1	$k_1 \times [E] \times [S]$	d[***E***] / dt = - $k_1 \times [E] \times [S]$ + $k_2 \times [ES]$ + $k_3 \times [ES]$
ES ---> E + S : k2	$k_2 \times [ES]$	d[***ES***] / dt = + $k_1 \times [E] \times [S]$ - $k_2 \times [ES]$ - $k_3 \times [ES]$
ES ---> E + P : k3	$k_3 \times [ES]$	Similarly for other species...

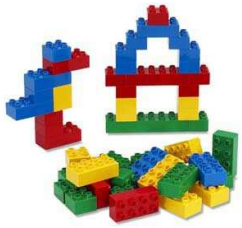
System of Simple, Simultaneous Equations

HOW DYNAFIT PROCESSES YOUR BIOCHEMICAL EQUATIONS



"The LEGO method" of deriving rate equations

Input (plain text file):	Rate terms:	Rate equations:
E + S ---> ES : k1	$k_1 \times [E] \times [S]$	
ES ---> E + S : k2	$k_2 \times [ES]$	
ES ---> E + P : k3	$k_3 \times [ES]$	



DynaFit can analyze many types of experiments

MASS ACTION LAW AND MASS CONSERVATION LAW IS APPLIED IN THE SAME WAY

chemistry
biophysics
enzymology

EXPERIMENT	DYNAFIT DERIVES A SYSTEM OF …
Kinetics (time-course)	Ordinary differential equations (ODE)
Equilibrium binding	Nonlinear algebraic equations
Initial reaction rates	Nonlinear algebraic equations

Example 1: Inhibition of HIV protease

"Tight binding" inhibition constant from initial rates

Use K_i^{app} values, not IC_{50}'s

Wha't wrong with IC_{50}'s ?

Measures of inhibitory potency

INTRINSIC MEASURE OF POTENCY:

$\Delta G = -RT \log K_i$

DEPENDENCE ON EXPERIMENTAL CONDITIONS	Depends on [S]	Depends on [E]	*Example:* Competitive inhibitor
1. Inhibition constant	**NO**	**NO**	K_i
2. Apparent K_i	**YES**	**NO**	$K_i^* = K_i (1 + [S]/K_M)$
3. IC_{50}	**YES**	**YES**	$IC_{50} = K_i (1 + [S]/K_M) + [E]/2$

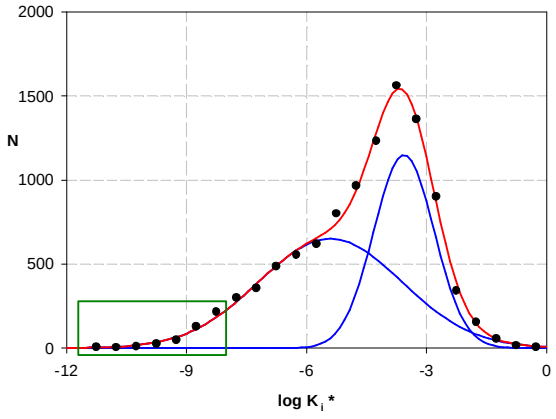
"CLASSICAL" INHIBITORS: $[E] \ll K_i$: $IC_{50} \approx K_i^*$

"TIGHT BINDING" INHIBITORS: $[E] \approx K_i$: $IC_{50} \neq K_i^*$

Tight binding inhibitors : $[E] \approx K_i$

HOW PREVALENT IS "TIGHT BINDING"?

A typical data set: ~ 10,000 **compounds**
Completely inactive: ~ 1,100 ... NOT SHOWN
Tight binding: ~ 400

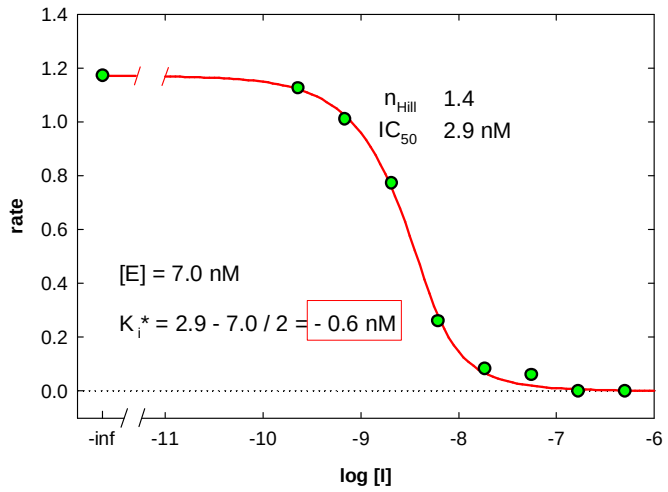


Data courtesy of Celera Genomics

Problem: Negative K_i from IC_{50}

FIT TO FOUR-PARAMETER LOGISTIC:

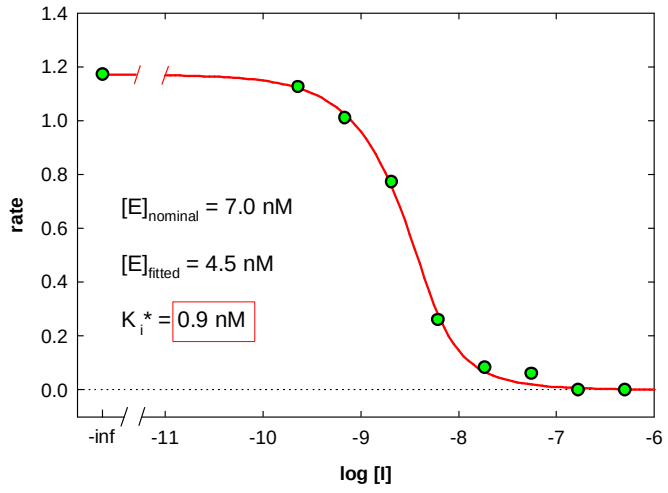
$K_i^* = IC_{50} - [E] / 2$



Data courtesy of Celera Genomics

Solution: Do not use four-parameter logistic

FIT TO MODIFIED MORRISON EQUATION: P. Kuzmic *et al.* (2000) *Anal. Biochem.* 281, 62-67.
P. Kuzmic *et al.* (2000) *Anal. Biochem.* 286, 45-50.



Data courtesy of Celera Genomics

live demo

DynaFit in K_i^{app} determination

- fitting model is very simple to understand:

```
E + I <==> EI
```


Comparison of results

Fitting model	Result
$v = P_{max} - \frac{P_{max} - P_0}{1 + ([I]/IC_{50})^n}$	IC_{50} = (**1.3 ± 0.13**) nM
	10 x !
$v = V_b + V_0 \frac{[E] - [I] - K_i^* + \sqrt{([E] - [I] - K_i^*)^2 + 4[E]K_i^*}}{2[E]}$	K_i^* = (**0.10 ± 0.05**) nM
`E + I <==> EI : Ki*`	K_i^* = (**0.10 ± 0.05**) nM

Fitting models for enzyme inhibition: Summary

MEASURE OF INHIBITORY POTENCY

- Apparent inhibition constant K_i^* is preferred over IC_{50}

MATHEMATICAL MODELS

- Modified Morrison equation is preferred over four-parameter logistic equation:

$$v = V_b + V_0 \frac{[E]-[I]-K_i^* + \sqrt{\left([E]-[I]-K_i^*\right)^2 + 4[E]K_i^*}}{2[E]}$$

- A **symbolic** model (DynaFit) is equivalent and more convenient:

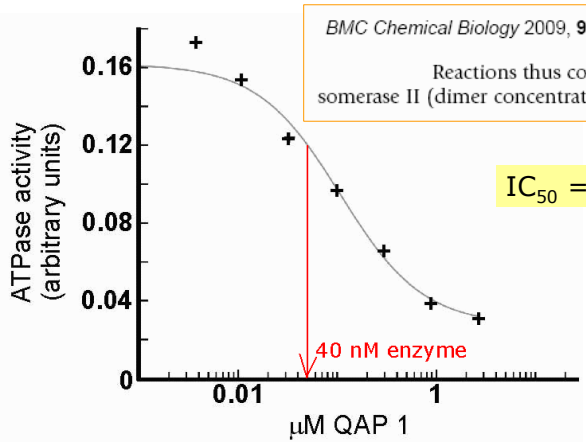
```
E + I <==> EI    :     Ki*
```


Recent IC_{50} work from Novartis, Basel

Patrick Chène *et al.* (2009) "Catalytic inhibition of topoisomerase II by a novel rationally designed ATP-competitive purine analogue" *BMC Chemical Biology* **9**:1

BMC Chemical Biology 2009, **9**:1

Reactions thus contained 10 – 40 nM topoisomerase II (dimer concentration) in reaction buffer (10



$IC_{50} = [E]/2 + K_i^*$

Challenges of moving from IC_{50} to K_i^{app}

- Legitimate need for continuity
  structure of existing corporate databases
- Simple inertia ("fear of the unknown")
- Lack of awareness

POSSIBLE SOLUTIONS:

- Gradual transition (report both IC_{50} and K_i^* for a period of time)
- Re-compute historical data: $\mathbf{K_i^* = IC_{50} - [E]/2}$

Finer points of K_i^{app} determination

Sometimes [E] **must** be optimized, but sometimes it **must not** be:

Kuzmic, P., *et al.* (**2000**) "High-throughput screening of enzyme inhibitors: Simultaneous determination of tight-binding inhibition constants and enzyme concentration" *Anal. Biochem.* **286**, 45-50

"Robust regression" analysis (exclusion of outliers):

Kuzmic, P. (**2004**) "Practical robust fit of enzyme inhibition data" *Meth. Enzymol.* **383**, 366-381

Serial dilution is **not** always the best:

Kuzmic, P. (**2011**) "Optimal design for the dose–response screening of tight-binding enzyme inhibitors" *Anal. Biochem.* **419**, 117-122

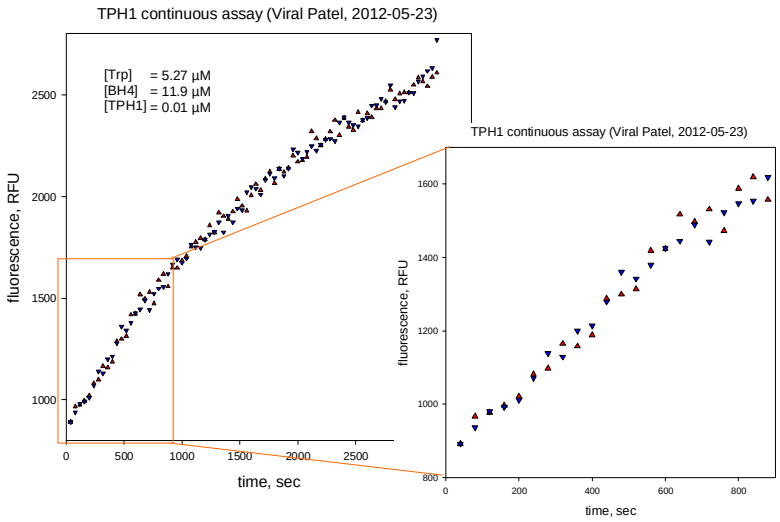
Example 1: TPH1 inhibition

Determination of initial reaction rates

from *nonlinear* progress curves

First look at raw experimental data

NOTE THE EXCELLENT REPRODUCIBILITY: THESE ARE **TWO REPLICATES** ON TOP OF EACH OTHER



The text-book recipe: Linear fit

STANDARD APPROACH: FIT A **STRAIGHT LINE** TO THE **"INITIAL PORTION"** OF EACH CURVE

To calculate the initial velocity from a reaction progress curve, the investigator first needs to examine the curve to ascertain the portion of it that represents the steady-state region. One needs to be sure that in the selected region, pre-steady-state phenomena are fully resolved, and data representing greater than 10% substrate turnover are excluded. Once the steady-state region of the data set is isolated, linear regression is performed to calculate the slope.

Stein, R. "Kinetics of Enzyme Action" (2011), sect. 2.5.4

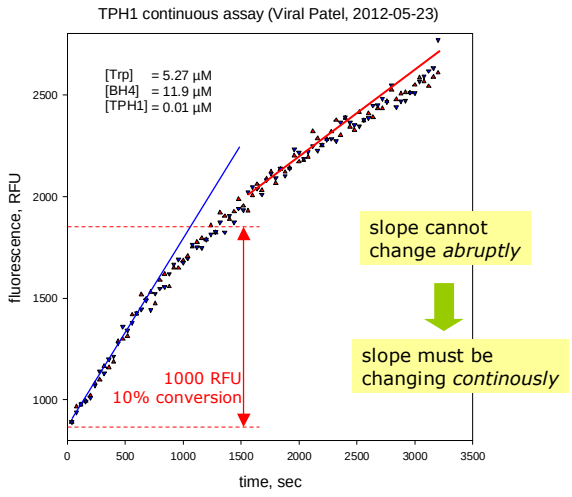
Standard curve → **1 µM** of **product** (5-HT) corresponds to an increase in fluorescence of **1919.9 RFU**

↓

Fluorescence change expected at **10% conversion**: ~ **1000 RFU**

Looking for linearity at less than 10% conversion

IS THIS A "STRAIGHT LINE"?



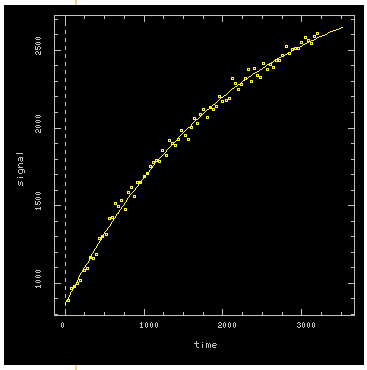
Numerical analysis: Fit to a full system of differential equations

A **NONLINEAR** MODEL OF COMPLETE REACTION PROGRESS

```
[task]
   data = progress
   task = fit

[mechanism]
   E + S <==> ES     :     kas    kds
   ES ---> E + P     :     kdp

[constants]
   kas =   1000
   kds = 100000 ?
   kdp =      1 ?

[concentrations]
   E = 0.01

[responses]
   P = 100 ?

[data]
   sheet B01.txt
   column    2   | offset auto ? | concentration

[output]
   directory .../output/fit-B01
   rate-file .../data/rates/B01v.txt
```

DYNAFIT INPUT

DYNAFIT OUTPUT



DynaFit auto-generated fitting model

A **NONLINEAR** MODEL OF COMPLETE REACTION PROGRESS

```
[mechanism]
   E + S <==> ES     :     kas    kds
   ES ---> E + P     :     kdp
```

$d[E]/dt = - k_{as}[E][S] + k_{ds}[ES] + k_{dp}[ES]$

$d[S]/dt = - k_{as}[E][S] + k_{ds}[ES]$

$d[ES]/dt = + k_{as}[E][S] - k_{ds}[ES] - k_{dp}[ES]$

$d[P]/dt = + k_{dp}[ES]$

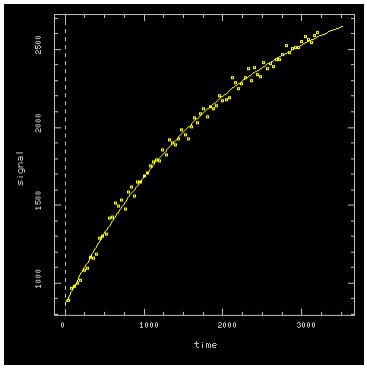
Back to Index...

Program **DynaFit** ver. 4.05.004

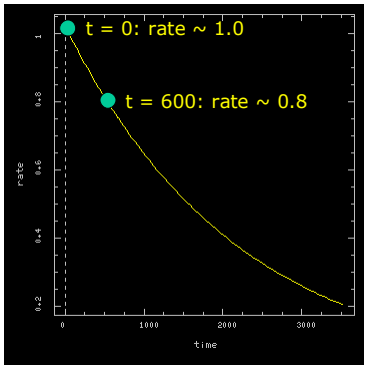
Instantaneous rate plot

THE SLOPE OF THE PROGRESS CURVE **DOES** CHANGE ALL THE TIME

fluorescence:



rate of change in fluorescence:



20 % drop in reaction rate over the first 10 minutes (less than 10% substrate conversion)

live demo

DynaFit in "full automation" mode

- suitable for routine processing of many compounds
- mechanism is assumed to be known independently

Example 2: Inhibition of 5α-ketosteroid reductase

"Slow, tight" binding

Model discrimination analysis

Possible molecular mechanisms of time-dependent inhibition

Slow binding proper

$$E + I \overset{slow}{\rightleftharpoons} E.I$$

Rearrangement of initial enzyme-inhibitor complex

$$E + I \overset{fast}{\rightleftharpoons} E.I \overset{slow}{\rightleftharpoons} E.I^*$$

etc. (several other possibilities)

Model discrimination analysis in DynaFit

ANY NUMBER OF ALTERNATE KINETIC MODELS CAN BE COMPARED IN A SINGLE RUN

DYNAFIT INPUT **SCRIPT FILE**:

```
[task]
   data  = progress   |   task  = fit  |  model = one-step ?

[mechanism]
   E + S ---> ES      :          kaS
   ES ---> E + P      :          kdP
   E + I <==> EI      :          kaI        kdI
...
```

```
[task]
   data  = progress   |   task  = fit  |  model = two-step ?

[mechanism]

   E + S ---> ES      :          kaS
   ES ---> E + P      :          kdP
   E + I <==> EI      :          kaI        kdI
   EI <==> EJ         :          kIJ        kJI
...
```


live demo

DynaFit in "model selection" mode

- model selection criteria (AIC, BIC)
- residual analysis
- use common sense to check results

Summary and conclusions

Benefits of using DynaFit in the study of enzyme inhibition:

- No mathematical models, only symbolic models (E + I <==> EI)

  Everybody can understand this.
  This prevents making mistakes and facilitates "transfer of knowledge".

- Automation

  Can be used to process 1000's of compounds in a single run.
  Automatic model selection (Bayesian Information Criterion).

- No restrictions on experimental design

  Not necessary to have large excess of inhibitor ("tight binding")

- No restrictions on reaction mechanism

  Any number of interactions and molecular species

Possible deployment at Novartis / Horsham

Free support from BioKin Ltd included with site license:

- Periodic data review via email

  Send your raw data, get results back in 72 hours (in most cases).

- Phone support

  Call +1.617.209.4242 any time during US (EST) business hours.

- Periodic on-site "DynaFit course"

  Once a year (either in the Fall or Spring); one-day workshop format.

- Free upgrades

  DynaFit continues to evolve (e.g., "Optimal Experimental Design")