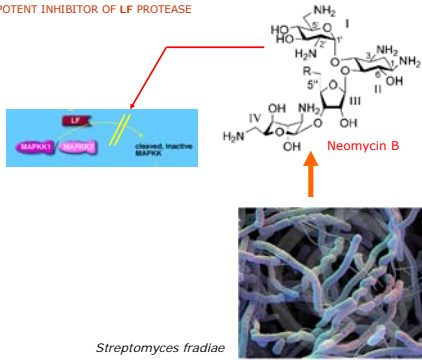


Neomycin B: an aminoglycoside inhibitor

A POTENT INHIBITOR OF LF PROTEASE



Supermodel Evolution in Bio/Chemical Kinetics

7

Neomycin B mechanism: mixed-type noncompetitive

NEOMYCIN B INHIBITION FOLLOWS A COMPLEX MOLECULAR MECHANISM

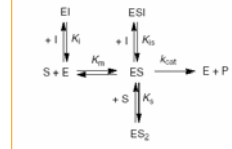
FEBS Journal

Mixed-type noncompetitive inhibition of anthrax lethal factor protease by aminoglycosides

Petr Kuzmic¹, Lynne Cregar², Sherri Z. Millis² and Mark Goldman^{2*}

¹ BioKin Ltd, Pullman, WA, USA
² Hawaii Biotech Inc., Ala, HI, USA

P. Kuzmic et al.



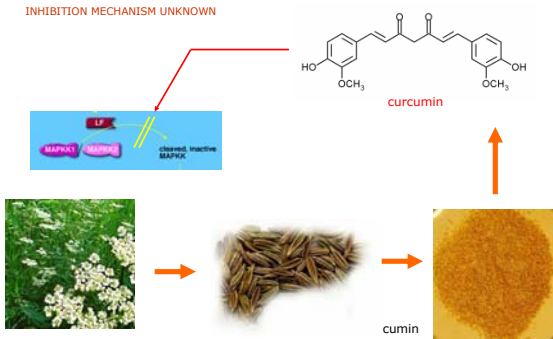
Kuzmic et al. (2006) FEBS J. 273, 3054-3062.

Supermodel Evolution in Bio/Chemical Kinetics

8

Curcumin: an natural product inhibitor

INHIBITION MECHANISM UNKNOWN

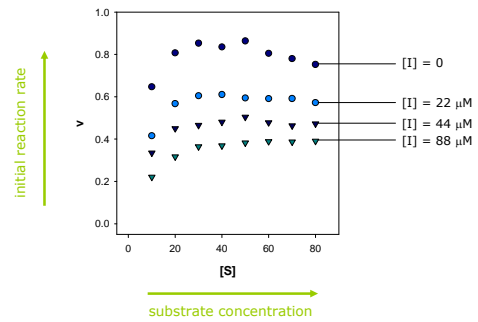


Supermodel Evolution in Bio/Chemical Kinetics

9

LF protease inhibition by curcumin: raw data

SUBSTRATE INHIBITION (MAXIMUM ON SUBSTRATE SATURATION CURVE)



Supermodel Evolution in Bio/Chemical Kinetics

10

Two separate problems to solve

A PREREQUISITE FOR MODEL DISCRIMINATION = FITTING INDIVIDUAL CANDIDATE MODELS

1. Focus on a *single* reaction mechanism:

Given a model (rate equation), find the best-fit kinetic constants

2. Focus on *multiple* reaction mechanisms:

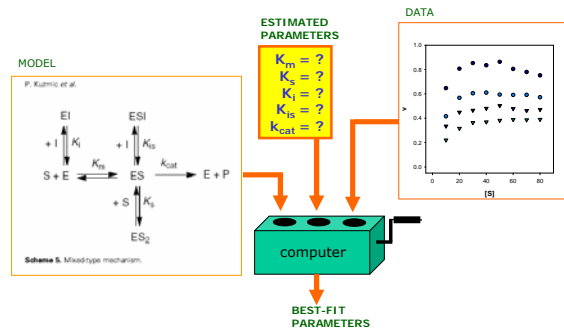
- Repeat 1. for all candidate models (mechanisms)
- Select the most plausible model

Supermodel Evolution in Bio/Chemical Kinetics

11

The mixed-type inhibition model

CONTAINS FIVE KINETIC CONSTANTS

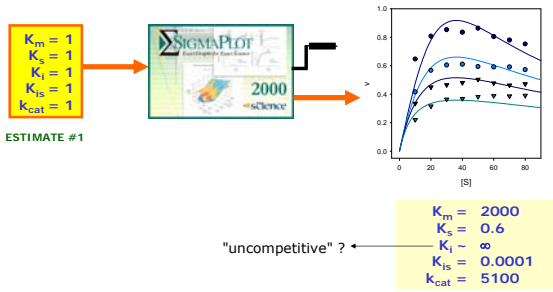


Supermodel Evolution in Bio/Chemical Kinetics

12

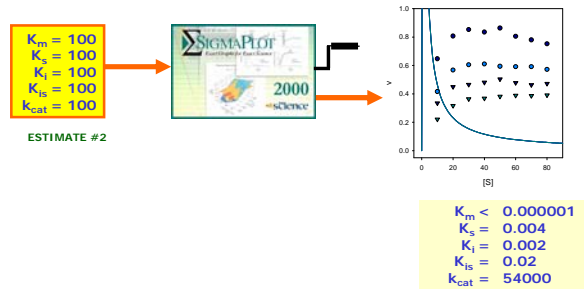
First major difficulty: Sensitivity to initial estimates

TRADITIONAL DATA-FITTING: RESULTS DEPEND ON THE INITIAL GUESS



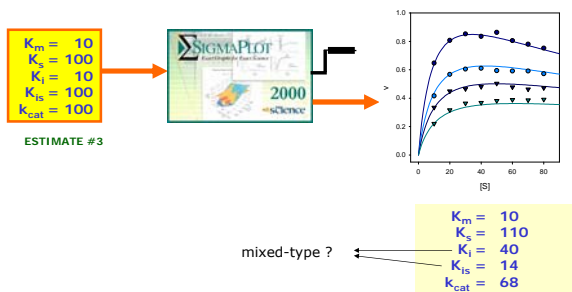
First major difficulty: Sensitivity to initial estimates

TRADITIONAL DATA-FITTING: RESULTS DEPEND ON THE INITIAL GUESS



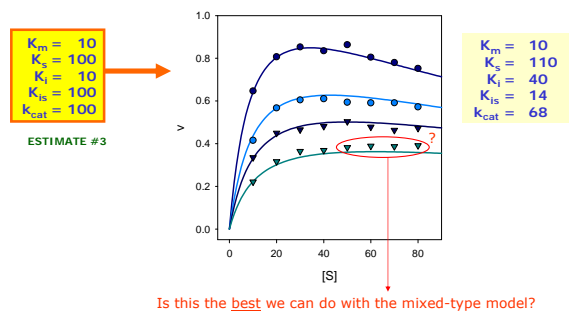
First major difficulty: Sensitivity to initial estimates

TRADITIONAL DATA-FITTING: RESULTS DEPEND ON THE INITIAL GUESS

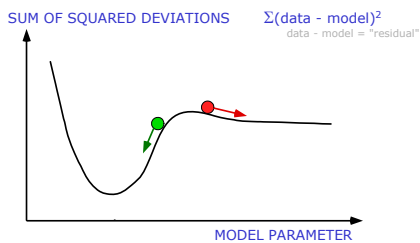


First major difficulty: Sensitivity to initial estimates

TRADITIONAL DATA-FITTING: RESULTS DEPEND ON THE INITIAL GUESS



The crux of the problem: Finding global minima



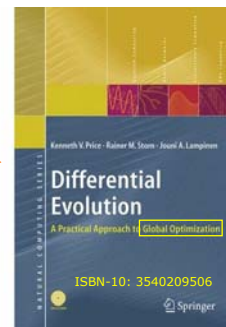
- Least-squares fitting **only** goes "downhill"
- **How do we know where to start?**

Charles Darwin to the rescue

BIOLOGICAL EVOLUTION IMITATED IN "DE"



Charles Darwin (1809-1882)



Biological metaphor: "Gene, allele"

BIOLOGY	COMPUTER
<p>gene</p> <p>...AAGTCG...GTAACCGG...</p> <p>"keratin"</p> <p>four-letter alphabet variable length</p>	<ul style="list-style-type: none"> • sequence of bits representing a number <p>...01110011000001101101110011...</p> <p>"K_M" "K_{cat}"</p> <ul style="list-style-type: none"> • two letter alphabet • fixed length (16 or 32 bits)

BioKin 19 Supermodel Evolution in Bio/Chemical Kinetics

"Chromosome, genotype, phenotype"

BIOLOGY	COMPUTER
<p>genotype</p> <p>...AAGTCGGTTCcGAAGTCGGTTA...</p> <p>keratin oncoprotein</p> <p>phenotype</p>	<ul style="list-style-type: none"> • particular combination of all model parameters <p>011010110110011100110100111101101</p> <p>V_{max} = 1.23 K_M = 4.56 K_{is} = 78.9</p> <p>full set of parameters</p> $v = V_{\max} \frac{[S]/K_M}{1 + [S]/K_M + [S]^2 / K_M K_{is}}$

BioKin 20 Supermodel Evolution in Bio/Chemical Kinetics

"Organism, fitness"

BIOLOGY	COMPUTER
<p>genotype</p> <p>...AAGTCGGTTCcGAAGTCGGTTA...</p> <p>keratin oncoprotein</p> <p>FITNESS: "agreement" with the environment</p>	<ul style="list-style-type: none"> • FITNESS: agreement between the data and the model <p>V_{max} = 1.3 K_M = 9.1 K_{is} = 137.8</p>

BioKin 21 Supermodel Evolution in Bio/Chemical Kinetics

"Population"

BIOLOGY	COMPUTER
<p>high fitness</p> <p>medium fitness</p> <p>low fitness</p>	

BioKin 22 Supermodel Evolution in Bio/Chemical Kinetics

DE Population size in DynaFit

```

DynaFit: settings.txt
File Edit View Help
Input Output
{DifferentialEvolution}
Strategy = 2
PopulationSizePerParameter = 50
PopulationSizePerOrderOfMag = 5
MaximumGenerationsPerParameter = 100
CombineGenerations = 1
RandomSeed = 1345
HistogramBins = 10
ReportFrequency = 1
Weight = 0.9
Crossover = 0.9
StopParameterRange = 0.0001
  
```

number of population members per optimized model parameter

number of population members per order of magnitude

BioKin 23 Supermodel Evolution in Bio/Chemical Kinetics

"Sexual reproduction, crossover"

BIOLOGY	COMPUTER
<p>mother</p> <p>father</p> <p>child</p>	<p>random crossover point</p> <p>01101011011001111001101 00011111011</p> <p>01101011011001111001101 11100011011</p> <p>"sexual mating" probability p_{cross}</p> <p>01101011011001111001101 11100011011</p> <p>V_{max} K_M K_{is}</p>

BioKin 24 Supermodel Evolution in Bio/Chemical Kinetics

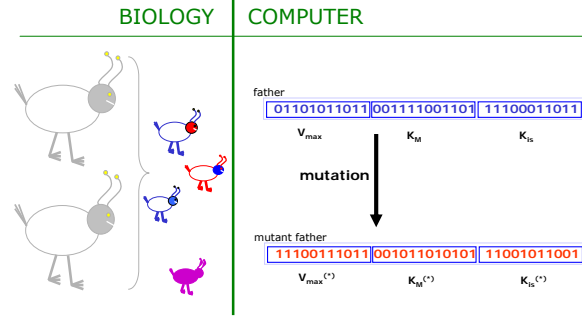
DE Crossover probability in DynaFit

```
DynaFit: settings.txt
File Edit View Help
Input Output

(DifferentialEvolution)
Strategy = 2
PopulationSizePerParameter = 50
PopulationSizePerOrderOfMag = 5
MaximumGenerationsPerParameter = 100
CombineGenerations = 1
RandomSeed = 1345
HistogramBins = 10
ReportFrequency = 1
Weight = 0.9
Crossover = 0.9
StopParameterRange = 0.0001
```

probability that *child* inherits *father's* genes, not *mother's* genes

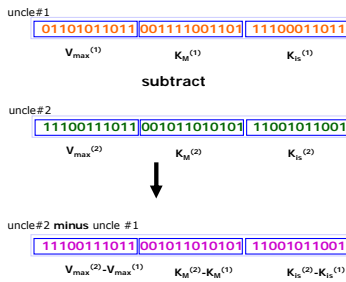
"Mutation, genetic diversity"



"Mutation, genetic diversity"

THE "DIFFERENTIAL" IN DIFFERENTIAL EVOLUTION ALGORITHM - STEP 1

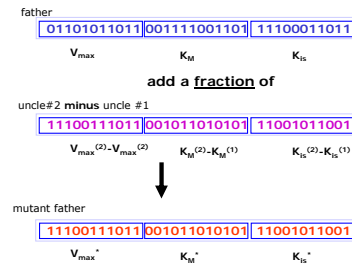
Compute difference between two randomly chosen "uncle" phenotypes



"Mutation, genetic diversity"

THE "DIFFERENTIAL" IN DIFFERENTIAL EVOLUTION ALGORITHM - STEP 2

Add weighted difference between two "uncle" phenotypes to "father"



"Mutation, genetic diversity"

THE "DIFFERENTIAL" IN DIFFERENTIAL EVOLUTION ALGORITHM

EXAMPLE: Michaelis-Menten equation $v = V_{max} \frac{[S]}{[S] + K_M}$

"mutant father" $\rightarrow K_M^* = K_M + F \times (K_M^{(1)} - K_M^{(2)})$

↑
weight (fraction)
mutation rate

DE Mutation rate in DynaFit

```
DynaFit: settings.txt
File Edit View Help
Input Output

(DifferentialEvolution)
Strategy = 2
PopulationSizePerParameter = 50
PopulationSizePerOrderOfMag = 5
MaximumGenerationsPerParameter = 100
CombineGenerations = 1
RandomSeed = 1345
HistogramBins = 10
ReportFrequency = 1
Weight = 0.9
Crossover = 0.9
StopParameterRange = 0.0001
```

fractional difference used in mutations

$$K_M^* = K_M + F \times (K_M^{(1)} - K_M^{(2)})$$

DE Mutation strategies in DynaFit

```

DynaFit: settings.txt
File Edit View Help
Input | Output |

(DifferentialEvolution)
Strategy = 2
PopulationSizePerParameter = 50
PopulationSizePerOrderOfMag = 5
MaximumGenerationsPerParameter = 100
CombineGenerations = 1
RandomSeed = 1345
HistogramBins = 10
ReportFrequency = 1
Weight = 0.9
Crossover = 0.9
StopParameterRange = 0.0001
    
```

Six different mutation strategies (1, 2, ... 6)

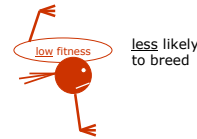
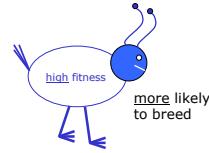
details in the book:



"Selection"

BIOLOGY

COMPUTER



low sum of squares
 $01101011011001111001100001111011$
 $V_{max} \quad K_M \quad K_{is}$
 more likely to be carried to the next generation

high sum of squares
 $00000000001111111111110000000000$
 $V_{max} \quad K_M \quad K_{is}$
 less likely to be carried to the next generation

Basic Differential Evolution Algorithm - Summary

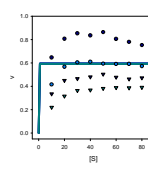
- 1 Randomly create the initial population (size N)
- Repeat until almost all population members have very high fitness:
 - 2 Evaluate fitness: sum of squares for all population members
 - 3 Mutation: random gene modification (mutate *father*, weight E)
 - 4 Sexual reproduction: random crossover with probability P_{cross}
 - 5 Natural selection: keep *child* in gene pool if more fit than *mother*

Application to curcumin: Mixed-type mechanism

THREE EXAMPLES OF POPULATION MEMBERS (POPULATION SIZE $n = 845$)

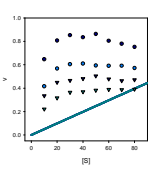
POPUL. MEMBER #1

Km	3.3355e-11
Ks	7868500
Ki	106.99
Kis	4.0737e-9
kcat	29.783



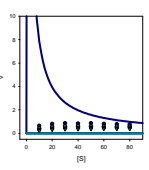
POPUL. MEMBER #2

Km	3713600
Ks	60218
Ki	427880
Kis	153.55
kcat	916170



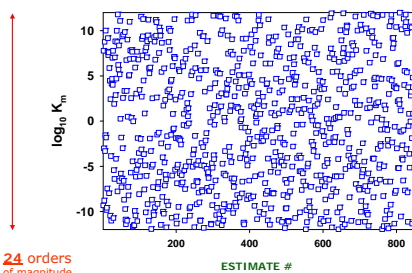
POPUL. MEMBER #642

Km	65.7630
Ks	1.0488e-3
Ki	8.7928e-10
Kis	8.1374e+20
kcat	3760400



Application to curcumin: Mixed-type mechanism

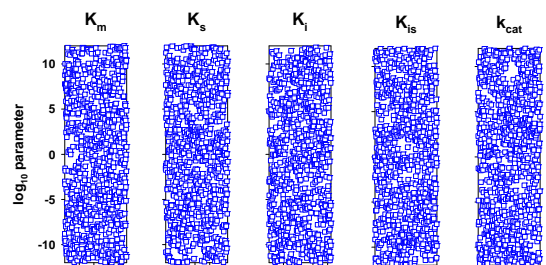
INITIAL DISTRIBUTION OF THE MICHAELIS-CONSTANT K_M



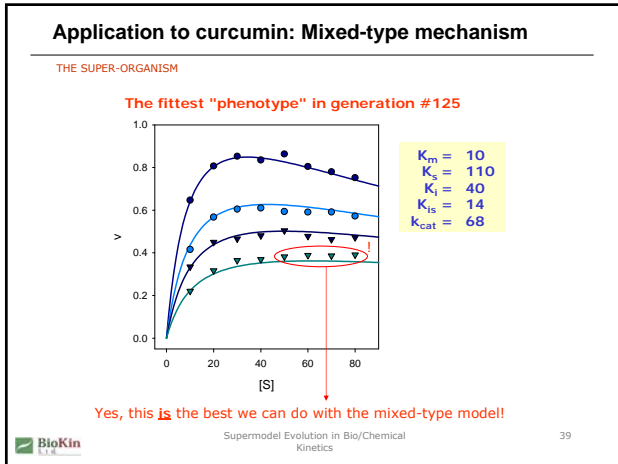
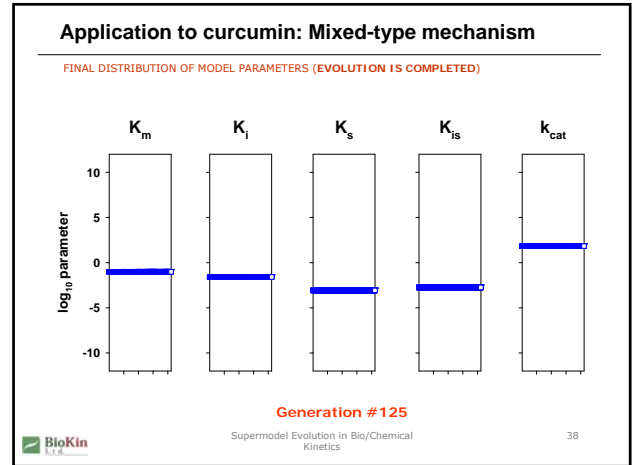
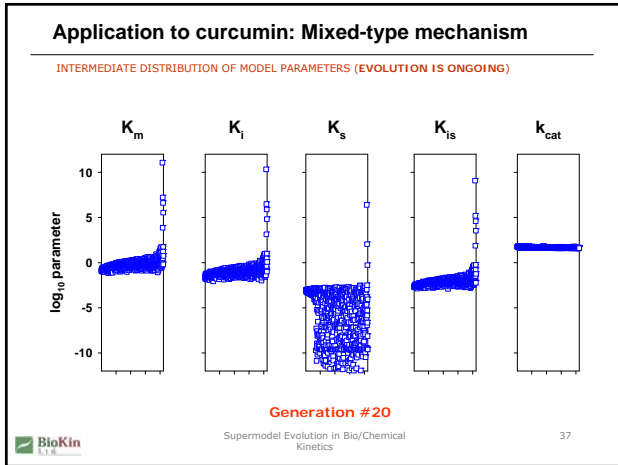
24 orders of magnitude

Application to curcumin: Mixed-type mechanism

INITIAL DISTRIBUTION OF ALL MODEL PARAMETERS



Generation #1



The model selection problem remains

A PREREQUISITE FOR MODEL DISCRIMINATION = FITTING INDIVIDUAL CANDIDATE MODELS

1. Focus on a *single* reaction mechanism:
 - Given a model (rate equation), find the best-fit kinetic constants
2. Focus on *multiple* reaction mechanisms:
 - a. Repeat 1. for **all** candidate models (mechanisms)
 - b. Select the most plausible model

?

Supermodel Evolution in Bio/Chemical Kinetics

BioKin

Model proliferation: CYP450 / reductase example

COMBINATORIAL PROLIFERATION OF POSSIBLE MECHANISMS

EXAMPLE

Global Analysis of Protein-Protein Interactions Reveals Multiple Cytochrome P450 2E1 / Reductase Complexes

Arvind P. Jamakhandi¹, Petr Kuzmič², Daniel E. Sanders³, and Grover P. Miller⁴

Journal: *Biochemistry* Ms# BI-2006-003476.R2 IN PRESS

Supermodel Evolution in Bio/Chemical Kinetics

BioKin

Model proliferation: CYP450 / reductase example

COMBINATORIAL PROLIFERATION OF POSSIBLE MECHANISMS

P = cytochrome P450 (2E1)
R = cytochrome reductase

Model#	Complex				
	PR	P2R	PR2	P2R2	P2
1	A				
2	A	A			
3	A		A		
4	A	A	A		
5	A			A	
6	A	A		A	
7	A		A	A	
8	A	A	A	A	N
9	A				N
10	A	A			N
11	A		A		N
12	A	A	A		N
13	A			A	N
14	A	A		A	N
15	A		A	A	N
16	A	A	A	A	N

A = catalytically active
N = inactive

42 separate mechanisms were examined

Supermodel Evolution in Bio/Chemical Kinetics

BioKin

The "Supermodel" approach

CREATE AN AGGREGATE MODEL ENCOMPASSING ALL POSSIBLE INTERACTIONS

Model#	Complex				
	PR	P2R	P2S	P2S2	P2
1	A				
2	A	A			
3	A	A	A		
4	A	A	A	A	
5	A	A	A	A	A
6	A	A	A	A	A
7	A	A	A	A	A
8	A	A	A	A	N
9	A	A	A	A	N
10	A	A	A	A	N
11	A	A	A	A	N
12	A	A	A	A	N
13	A	A	A	A	N
14	A	A	A	A	N
15	A	A	A	A	N
16	A	A	A	A	N

the most complex (realistic) model

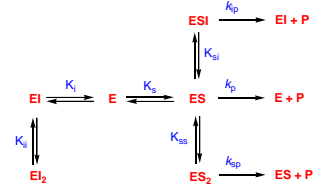
1. Consider only the **most complex** model
2. Evolve all parameters using "DE"
3. Identify redundant parameters by examining the final distribution of fittest phenotypes
4. Eliminate redundant parameters thereby reducing the model ("small is beautiful")

The "Supermodel" for LF inhibition by curcumin

COMPILE AN AGGREGATE OF ALL POSSIBLE MOLECULAR INTERACTIONS

ASSUMPTIONS

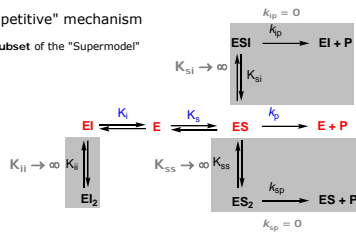
- Substrate can bind with 1:1 or 2:1 stoichiometry
- Inhibitor can bind with 1:1 or 2:1 stoichiometry
- Substrate and inhibitor can bind at the same time
- Any enzyme-substrate complex can have catalytic activity



The "Supermodel" includes all standard mechanisms

STANDARD MECHANISMS DIFFER ONLY BY VALUES OF KINETIC CONSTANTS IN THE "SUPERMODEL"

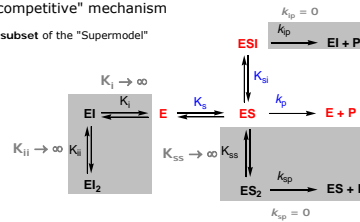
"Competitive" mechanism
is a subset of the "Supermodel"



The "Supermodel" includes all standard mechanisms

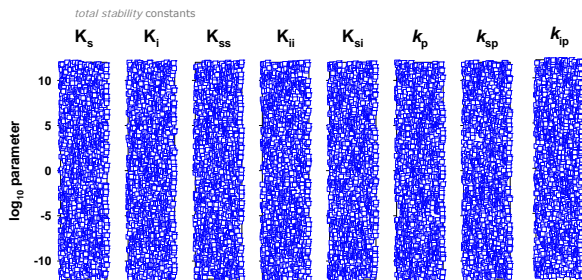
STANDARD MECHANISMS DIFFER ONLY BY VALUES OF KINETIC CONSTANTS IN THE "SUPERMODEL"

"Uncompetitive" mechanism
is a subset of the "Supermodel"



"Supermodel" evolution: Curcumin inhibition of LF

INITIAL DISTRIBUTION OF MODEL PARAMETERS (POPULATION SIZE = 1355)



"Supermodel" evolution: Curcumin inhibition of LF

FINAL DISTRIBUTION OF MODEL PARAMETERS

