MASARYKOVA UNIVERZITA
INNOVATION LECTURES (INNOIEC)
www.muni.cz

Binding and Kinetics for Experimental Biologists

Lecture 3

Equilibrium binding: Theory

Petr Kuzmič, Ph.D.
BioKin, Ltd.

WATERTOWN, MASSACHUSETTS, U.S.A.

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.

EVROPSKÁ UNIE esf MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY OP Vzdělávání pro konkurenceschopnost UNIVERSITAS MASARYKIANA BRUNENSIS

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

BioKin L t d.

Lecture outline

- Theory of equilibrium binding analysis

  - **matrix representation** of simultaneous equilibria

    stoichiometric matrix
    formula matrix
    stability matrix

  - **composition** of complex biochemical mixtures: numerical method
  - **thermodynamic cycles**: inconsistent equilibrium mechanisms
  - **multiple equivalent sites**: statistical factors
  - **nonspecific binding**: representation as a single "weak" binding site

- Applications in DynaFit scripting

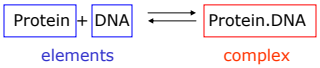
  - Representing equilibrium binding mechanisms in DynaFit

Matrix representation of complex equilibria

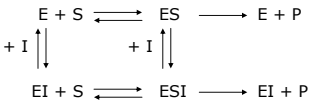
"Complexes" and "elements"

TWO DISTINCT TYPES OF MOLECULAR SPECIES

Example 1: **protein / DNA interactions**



Example 2: **enzyme kinetics (partial noncompetitive inhibition)**



elements: E, S, I, P

complexes: ES, EI, ESI

number of **species** = number of **elements** **+** number of **complexes**

BioKin Ltd.

Illustrative example: Arginine kinase

PHYSIOLOGICAL IMPORTANCE

"**Phosphagen kinases** catalyze the reversible transfer of a high-energy phosphoryl group from phosphorylated guanidino storage compounds known as phosphagens to ADP in the following general reaction: phosphagen + MgADP + $H^+ \rightarrow$ guanidino acceptor + MgATP. These reactions are typically found in **cells that display high and variable rates of energy turnover, such as muscle fibers, neurons, spermatozoa and transport epithelia**"

Journal of Experimental Biology **206**, 1545-1556 (2003)
Compaan DM, Ellington WR

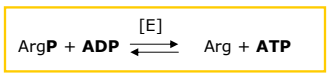
"The kinetic mechanism and evaluation of several **potential inhibitors** of purified arginine kinase from the cockroach (Periplanta americana) were investigated. This monomeric phosphagen kinase is important in **maintaining ATP levels during the rapid energy demands of muscle required for contraction and motility**. [...] Arginine kinase could be a useful chemotherapeutic target for the **control of cockroach proliferation**."

Archives of Insect Biochemistry and Physiology **57**, 166-77 (2004)
Brown AE, Grossman SH

Formula matrix, *F*

DESCRIBES THE COMPOSITION OF COMPLEXES IN TERMS OF ELEMENTS

Example: **Arginine Kinase** *J. Exper. Biol.* **206**, 1545-1556 (2003)



$$\mathbf{F} = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & \cdot & 1 & 1 & 1 & 1 & 1 \\ \cdot & 1 & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & 1 \\ \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot \end{pmatrix}$$

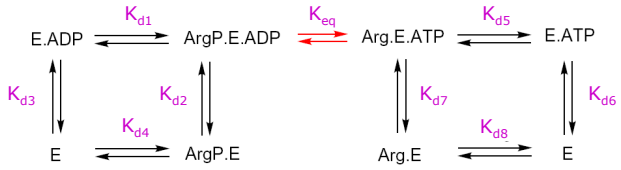
… E
… ArgP
… ADP
… Arg
… ATP

		E	ArgP	ADP	Arg	ATP
elements	E	1	.	.	.	.
	ArgP	.	1	.	.	.
	ADP	.	.	1	.	.
	Arg	.	.	.	1	.
	ATP	.	.	.	.	1
complexes	E·ArgP	1	1	.	.	.
	E·ADP	1	.	1	.	.
	E·ArgP·ADP	1	1	1	.	.
	E·Arg·ATP	1	.	.	1	1
	E·Arg	1	.	.	1	.
	E·ATP	1	.	.	.	1

Stoichiometric matrix, *S*

DESCRIBES INDIVIDUAL **REACTIONS** IN THE MECHANISM

Example: **Arginine Kinase** *Arch. Insect. Biochem. Physiol.* **57**, 166-77 (2004)



	K_{d1}	K_{d2}	K_{d3}	K_{d4}	K_{eq}	K_{d5}	K_{d6}	K_{d7}	K_{d8}
E	.	.	+1	+1	.	.	+1	.	+1
ArgP	+1	.	.	+1	.	.	.	.	.
ADP	.	+1	+1	.	.	.	.	.	.
Arg	.	.	.	.	.	+1	.	.	+1
ATP	.	.	.	.	.	.	+1	+1	.
E·ArgP	.	+1	.	- 1	.	.	.	.	.
E·ADP	+1	.	- 1	.	.	.	.	.	.
E·ArgP·ADP	- 1	- 1	.	.	- 1	.	.	.	.
E·Arg·ATP	.	.	.	.	+1	- 1	.	- 1	.
E·Arg	.	.	.	.	.	.	.	+1	.
E·ATP	.	.	.	.	.	+1	- 1	.	- 1

reactants

products

Definition: Complex formation constants

OVERALL EQUILIBRIUM CONSTANT = PRODUCT OF SEQUENTIAL EQUILIBRIUM CONSTANTS

Two sequential reaction with known K_d's:

$$A{\cdot}B \xrightleftharpoons{K_{d1}} A + B \qquad K_{d1} = [A][B]/[AB]$$

$$A{\cdot}B{\cdot}C \xrightleftharpoons{K_{d2}} A{\cdot}B + C \qquad K_{d2} = [A{\cdot}B][C]/[A{\cdot}B{\cdot}C]$$

Total dissociation constant: $K_{dT} = K_{d1} \times K_{d2}$

$$A{\cdot}B{\cdot}C \xrightleftharpoons{K_{dT}} A + B + C \qquad K_{dT} = [A][B][C]/[A{\cdot}B{\cdot}C]$$

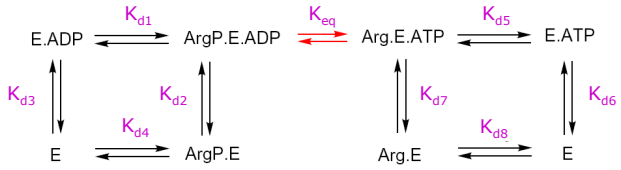
Complex stability constant: $K_A = 1 / (K_{d1} \times K_{d2})$

$$A + B + C \xrightleftharpoons{K_{aT}} A{\cdot}B{\cdot}C \qquad K_{aT} = [A{\cdot}B{\cdot}C]/[A][B][C]$$

Complex stability matrix, *B*

DESCRIBES OVERALL STABILITY CONSTANTS OF COMPLEXES IN TERMS OF BINARY K_d's

Example: **Arginine Kinase** *Arch. Insect. Biochem. Physiol.* **57**, 166-77 (2004)



	K_{d1}	K_{d2}	K_{d3}	K_{d4}	K_{eq}	K_{d5}	K_{d6}	K_{d7}	K_{d8}
E·ArgP	.	.	.	-1	.	.	.	.	.
E·ADP	.	.	-1	.	.	.	.	.	.
E·ArgP·ADP	-1	.	-1	.	.	.	.	.	.
E·Arg·ATP	.	.	.	.	.	.	.	-1	-1
E·Arg	.	.	.	.	.	.	.	.	-1
E·ATP	.	.	.	.	.	.	-1	.	.

$$K_{A,i}^{(tot)} = \prod_{j=1}^{M} K_j^{B_{i,j}}$$

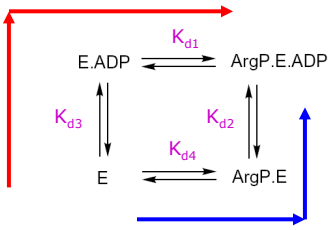
$K_A = 1/ (K_{d7} \times K_{d8})$

Can you spot anything suspicious about this stability matrix?

Complex stability matrix: Ambiguity in case of cycles

TWO WAYS TO DEFINE THE STABILITY OF A TERNARY COMPLEX

Example: **Arginine Kinase** *Arch. Insect. Biochem. Physiol.* **57**, 166-77 (2004)



$K_A = 1 / (K_{d3} \times K_{d1})$

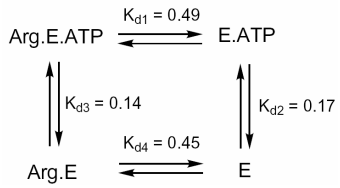
$K_A = 1 / (K_{d4} \times K_{d2})$

Thermodynamic cycles

Thermodynamic "box"

OVERALL EQUILIBRIUM CONSTANTS AROUND A **CYCLE** MUST BE EQUAL TO **1.00**

Numerical values of binary dissociation constants (mM):
Arch. Insect. Biochem. Physiol. **57**, 166-77 (2004)



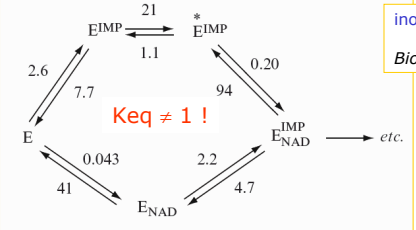
This thermodynamic box is not quite closed!

$$\frac{K_{d1} \times K_{d2}}{K_{d3} \times K_{d4}} = 1.32 \neq 1.00$$

BioKin Ltd.

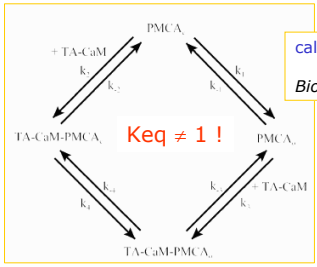
Thermodynamic boxes in the literature

THIS THIS NOT A "MADE UP" ISSUE: **ERRONEOUSLY CLOSED BOXES** ARE IN FACT FOUND!



inosine 5'-monophosphate dehydrogenase

Biochemistry **38**, 2295 (1999)

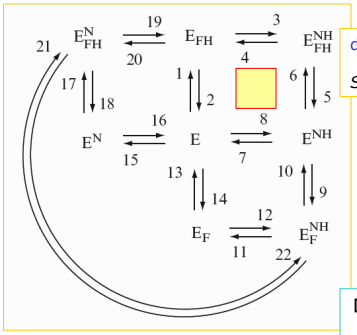


calmodulin / plasma membrane Ca^{2+} pump

Biochemistry **42**, 12115 (2003)

DynaFit notation to "close the boxes" properly

COMES UP IN THE ANALYSIS OF **KINETIC DATA**



dihydrofolate reductase

Science **239**, 1105 (1988)

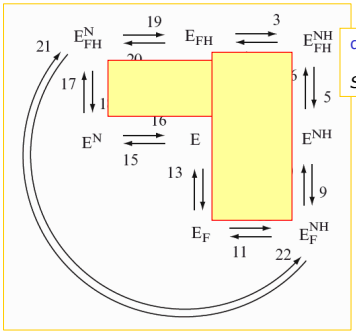
$k_2 = k_1 \times k_3 \times k_5 \times k_7 \,/\, k_4 \times k_6 \times k_8$

```
[constants]

   k1 = ...
   k2 = (k1 k3 k5 k7) / (k4 k6 k8)
   k3 = ...
   k4 = ...
   ...
```

Multiple thermodynamic cycles

COME UP IN MORE COMPLEX MECHANISMS

How many cycles do you see?



dihydrofolate reductase

Science **239**, 1105 (1988)

number of *constraints*
on rate / equilibrium constants

=

number of *cycles*
in the reaction mechanism.

Equilibrium composition of complex mixtures

Composition of complex mixtures at equilibrium

MATRIX FORMALISM - FREE ENERGY MINIMIZATION

Find a vector of species concentrations, **c**, that minimize the total Gibbs free energy

$$G = \sum_{i=1}^{N} c_i \left(\mu_i^{\circ} + \mathbf{R}T \ln c_i \right)$$

where $\mu_i^{\circ} = -RT \ln K_{A,i}^{(tot)}$ $\quad K_{A,i}^{(tot)} = \prod_{j=1}^{M} K_j^{B_{i,j}}$ $\quad K_j = \prod_{i=1}^{N} c_i^{S_{i,j}}$

subject to the mass-conservation constraints

$$\sum_{k=1}^{N} F_{k,j} c_i = c_i^{(tot)}$$

DETAILS:

Royer, C.A.; Smith, W.R.; and Beechem, J.M. (1990)
"Analysis of binding in macromolecular complexes: A generalized numerical approach"
Anal. Biochem., **191**, 287-294.

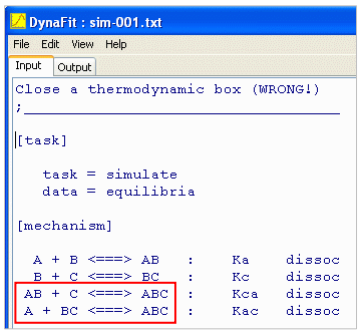
Royer, C.A. and Beechem, J.M. (1992)
"Numerical analysis of binding data: advantages, practical aspects, and implications"
Methods Enzymol. **210**, 481-505.

Equilibrium matrices in DynaFit

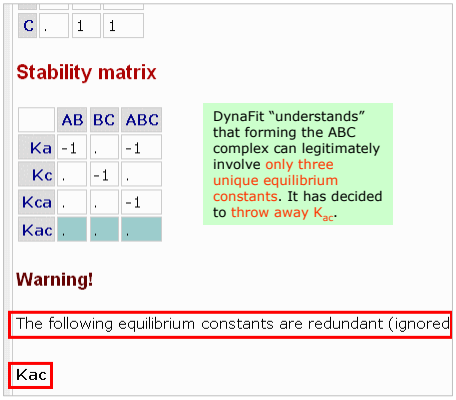
Why should you care about matrices in DynaFit ?

AT LEAST TO UNDERSTAND WARNING MESSAGES WHEN **YOUR MECHANISM IS WRONG**

Example of a "wrong" mechanism: A closed thermodynamic box



invalid mechanism!



Stoichiometric matrices in DynaFit

"TABULAR" TRANSCRIPT OF THE REACTION **MECHANISM**

DynaFit input:

```
[mechanism]

  A + B <===> AB    :    Ka    dissoc
  B + C <===> BC    :    Kc    dissoc
 AB + C <===> ABC   :    Kca   dissoc
 A + BC <===> ABC   :    Kac   dissoc
```

DynaFit output:

Stoichiometric matrix

	A	B	C	AB	BC	ABC
Ka	1	1	.	-1	.	.
Kc	.	1	1	.	-1	.
Kca	.	.	1	1	.	-1
Kac	1	.	.	.	1	-1

Stoichiometric matrix:

- as many rows as there are *reactions*
- as many columns as there are *species*
- entries reflect the *stoichiometry* of reactions
- negative entries represent "reactants"
- positive entries represent "products"

What it means for you:
Always check at least the species names (top row) and equilibrium constant names (leftmost column).
This will assure that DynaFit "understood" what you meant.

Formula matrices in DynaFit

REACTING SPECIES ARE CLASSIFIED INTO **COMPONENTS** AND **COMPLEXES**

DynaFit input:

```
[mechanism]

  A + B <===> AB    :    Ka    dissoc
  B + C <===> BC    :    Kc    dissoc
 AB + C <===> ABC   :    Kca   dissoc
  A + BC <===> ABC  :    Kac   dissoc
```

DynaFit output:

Formula matrix

	AB	BC	ABC
A	1	.	1
B	1	1	1
C	.	1	1

Formula matrix:

- as many rows as there are *components*
- as many columns as there are *complexes*
- entries reflect the *composition ("formula")* of complexes

What it means for you:
Always check the "Model" link in the output.
Make sure that the formula matrix is derived correctly.
This will assure that DynaFit "understood" what you meant.

Formula matrices: Sometimes not completely obvious

DYNAFIT IS FAIRLY GOOD AT "UNDERSTANDING" WHAT IS A COMPLEX OR COMPONENT SPECIES

DynaFit input:

```
[mechanism]

  M1 + M2 <===> M3   :    K1    dissoc
  M4 + M5 <===> M3   :    K2    dissoc
  M4 + M6 <===> M1   :    K3    dissoc
  M6 + M2 <===> M5   :    K4    dissoc
```

Components: ? M6, M2, M4
Complexes: ? M3, M5, M1
↑
"ABC"

DynaFit output:

Formula matrix

	M3	M5	M1
M6	1	1	1
M2	1	1	.
M4	1	.	1

What it means for you:
Always name your species "sensibly".
This will assure that the formula matrix can be easily checked.

Stability matrices in DynaFit

HERE IS WHERE DYNAFIT PUTS ANY WARNINGS ABOUT **INCONSISTENT MECHANISMS**

DynaFit input:

```
[mechanism]

  A + B <===> AB    :     Ka    dissoc
  B + C <===> BC    :     Kc    dissoc
 AB + C <===> ABC   :     Kca   dissoc
 A + BC <===> ABC   :     Kac   dissoc
```

DynaFit output:

Stability matrix

	AB	BC	ABC
Ka	-1	.	-1
Kc	.	-1	.
Kca	.	.	-1
Kac	.	.	.

Stability matrix:

$K_{ABC} = K_a^{-1} K_{ca}^{-1}$

- as many rows as there are *equilibrium constants* in the mechanism
- as many columns as there are *complexes*
- entries reflect the *total stability constants* of complexes
- an empty row means a *redundant step* in the mechanism (e.g., a closed "box")

What it means for you:
Always check the stability matrix for the presence of empty (shaded) rows.
If any are found, see which redundant equilibrium constant was ignored by DynaFit.
If you prefer to keep this constant, **delete some other redundant step**.

Formula matrix derivation: how *does* DynaFit do it?

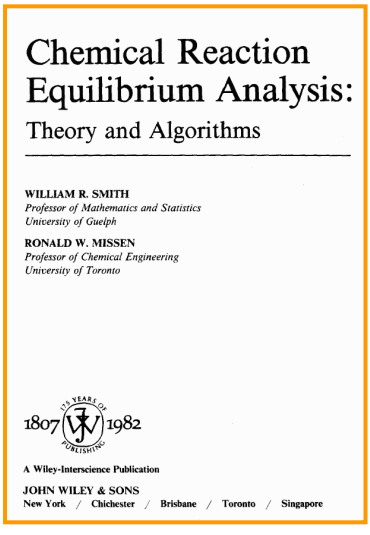
DYNAFIT "DOES" A LOT OF LINEAR ALGEBRA AND "UNDERSTANDS" A VERY NICE **THEORY BOOK**

W.R. Smith and R.W. Missen (1991)

Chemical Reaction Equilibrium Analysis: Theory and Algorithms

2nd Edition, Krieger Publishing
Malabar, Florida

ISBN-10: 0894645846




```
[mechanism]

 M1 + M2 <===> M3   :     K1    dissoc
 M4 + M5 <===> M3   :     K2    dissoc
 M4 + M6 <===> M1   :     K3    dissoc
 M6 + M2 <===> M5   :     K4    dissoc
```

components

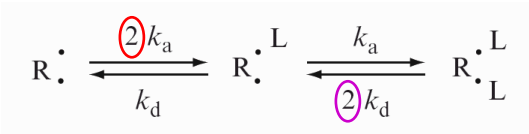
Equivalent binding sites & “statistical factors”

Equivalent binding sites & “statistical factors”

“**INTRINSIC**” VS. “**MACROSCOPIC**” RATE CONSTANTS

Two identical, non-interacting binding sites:

two equivalent ways to come together



two equivalent ways to fall apart

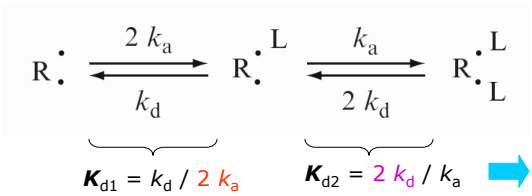
R	receptor
L	ligand
•	binding site
k_a	*intrinsic* association rate constant
k_d	*intrinsic* dissociation rate constant

BioKin Ltd.

2 Equivalent binding sites: Equilibrium constants

"**INTRINSIC**" VS. "**MACROSCOPIC**" EQUILIBRIUM CONSTANTS

Two identical, non-interacting binding sites:



$K_{d1} = k_d / 2\ k_a$ $K_{d2} = 2\ k_d / k_a$ ➡ $K_{d2} = 4\ K_{d1}$

DynaFit notation:

```
[mechanism]
    R + L <==> R.L    :  K1 dissoc
  R.L + L <==> R.L.L  :  K2 dissoc

[constants]
  K2 = 4 * K1
  K1 = 1.2345
```


3 Equivalent binding sites: Equilibrium constants

A DYNAFIT SCRIPT TO **DISTINGUISH** BETWEEN **INDEPENDENT AND INTERACTING** SITES

```
[task]
  data  = equilibria
  model = interacting sites ?

[mechanism]
      P + L <===> P.L     : K1 dissoc
    P.L + L <===> P.L.L   : K2 dissoc
  P.L.L + L <===> P.L.L.L : K3 dissoc

[constants] ; vary independently
   K3 = 1.23 ?
   K2 = 4.56 ?
   K1 = 7.89 ?
...
[task]
  data  = equilibria
  model = independent sites ?

[constants] ; link via statistical factors
  K3 = 1 ?
  K2 = 3 * K3
  K1 = 9 * K3
...
```

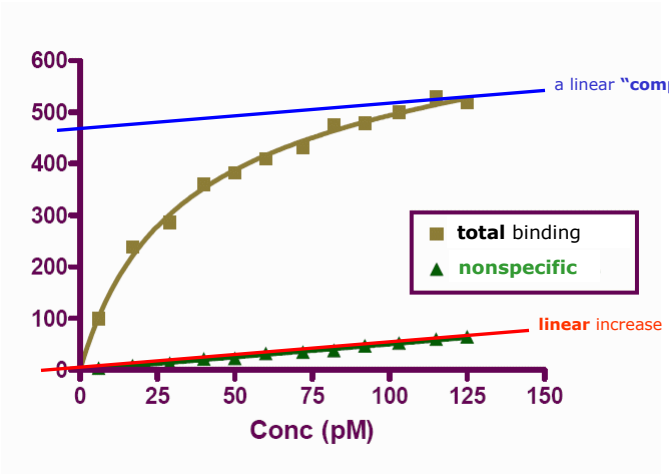
exercise:

derive statistical relationships between equilibrium constants for **four** independent sites

Nonspecific binding interactions

Nonspecific binding: How to identify it

HOW TO RECOGNIZE NONSPECIFIC BINDING IN ACTUAL EXPERIMENTAL DATA



www.graphpad.com/prism/learn/binding%20analysis.pdf

BioKin Ltd.

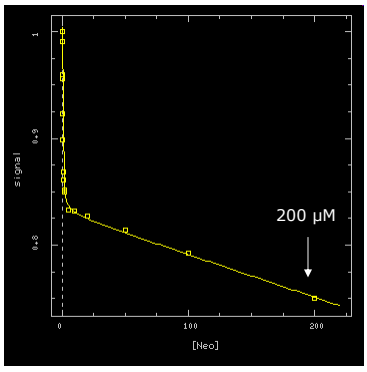
Nonspecific binding: How to analyze it in DynaFit

DEFINE AN **EXTREMELY WEAK** BINDING SITE WITH A **HUGE RESPONSE** FACTOR

1. introduce a second "specific" binding site
   this will represent many very weak (nonspecific) sites
2. assign to it an extremely weak binding affinity
   many orders of magnitude higher K_d
3. treat the weak nonspecific K_d as a fixed constant
   optimize only the *specific* binding constants
4. assign a very large response factor
   corresponding to "many simultaneous sites"
5. treat the response factor as an adjustable parameter
   this represents the "number of weak sites"

Nonspecific binding: Example - data

HIV-1 Rev RESPONSIVE ELEMENT (LABELED RNA FRAGMENT) / NEOMYCIN B



[Neomycin], µM

For the RRE-72AP construct, a further 10% linear decrease in fluorescence was observed as the neomycin concentration was increased from 10 to 200 μM. This decrease is most likely attributable to binding to very weak sites with binding constants in the low millimolar range, and was not analyzed further.

Lacourciere *et al.* (2000) *Biochemistry* **39**, 5630-41

Nonspecific binding: Example – input script

HIV-1 Rev RESPONSIVE ELEMENT (LABELED RNA FRAGMENT) / NEOMYCIN B

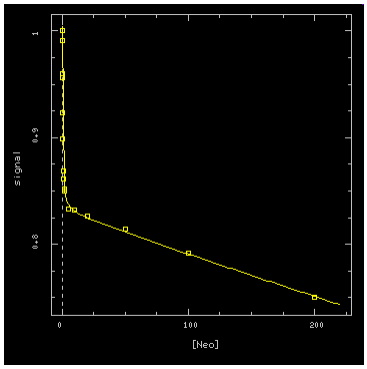
```
[mechanism]

   R72 + Neo <===> R72.Neo            :   K1   dissoc
   R72.Neo + Neo <===> Neo.R72.Neo    :   K2   dissoc

[constants] ; uM

   K1 =   0.1 ?
   K2 = 10000

[concentrations] ; uM

   R72 = 0.1

[responses]

   R72 = 10 ?
   R72.Neo = 5 ?
   Neo.R72.Neo = -5 ?
```

K_{d2} = 10 mM — extremely weak binding

"straight line" binding curve continues to **high negative values**

Nonspecific binding: Example – results of fit

HIV-1 Rev RESPONSIVE ELEMENT (LABELED RNA FRAGMENT) / NEOMYCIN B



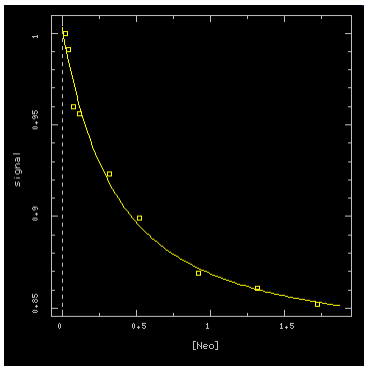
K_1 = (**266 ± 30**) nM

specific binding component

No.	Par#Set	Initial	Final	Std. Error	CV (%)	Note
#1	K1	0.1	0.265656	0.0302723	11.40	
#2	r(R72)	10	10.0441	0.0388884	0.39	
#3	r(R72.Neo)	5	8.30487	0.0286772	0.35	
#4	r(Neo.R72.Neo)	-5	-32.2529	3.01752	9.36	

Nonspecific binding: Sometimes it *can* be ignored

HIV-1 Rev RESPONSIVE ELEMENT (LABELED RNA FRAGMENT) / NEOMYCIN B



K_1 = (**290 ± 70**) nM

specific binding only

No.	Par#Set	Initial	Final	Std. Error	CV (%)
#1	K	0.3	**0.289936**	0.0728402	25.12
#2	r(R72)	10	**10.0329**	0.0556981	0.56
#3	r(R72.Neo)	8.5	**8.2671**	0.104526	1.26

Summary and conclusions

1. We need to understand the matrix representation of complex equilibria
   at least enough to be able to read DynaFit warning messages.
2. DynaFit will happily work even with inconsistent mechanisms,
   but it will ignore redundant equilibrium constants ("closed thermodynamic box")
3. Equivalent sites require a proper use of statistical factors.
4. Nonspecific binding looks like a "linear component"
   superimposed onto the usual nonlinear (hyperbolic) binding curve.
5. Nonspecific binding is modeled in DynaFit as "specific" binding
   characterized by very high dissociation constant "fixed" in the regression model.
6. DynaFit is not a "silver bullet": You must still use your brain a lot.