Determination of Binding Affinities
and Molecular Mechanisms

Biochemical Society
Advancing Molecular Bioscience

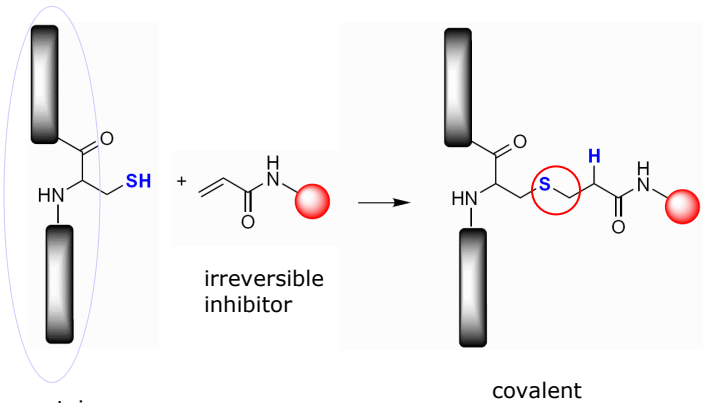
Training Day
May 2, 2014 (London)

Part 3: Case study – Covalent Enzyme Inhibitors

Petr Kuzmič
BioKin, Ltd.

Covalent inhibitors of cancer-related enzymes

COVALENT INHIBITION WORKS BY MAKING A **PERMANENT** **CHEMICAL BOND** WITH THE ENZYME



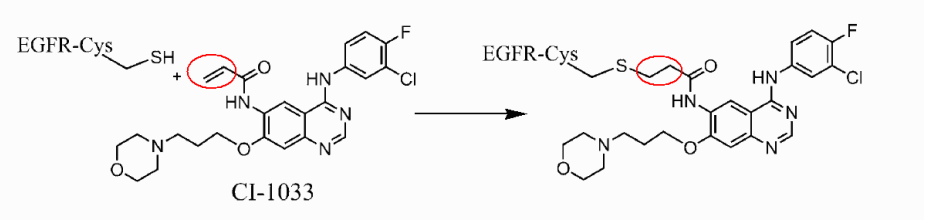
EGFR inhibition by covalent drugs

Schwartz, P.; Kuzmic, P. *et al.* (2014)

"Covalent EGFR inhibitor analysis reveals
importance of reversible interactions
to potency and mechanisms of drug resistance"

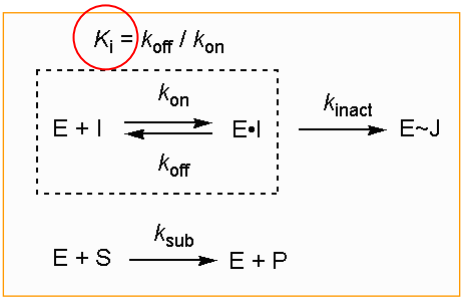
Proc. Natl. Acad. Sci. USA. **111**, 173-178. Issue 1, January 7

EXAMPLE:



Two steps: 1. non-covalent binding, 2. inactivation

equilibrium binding constant



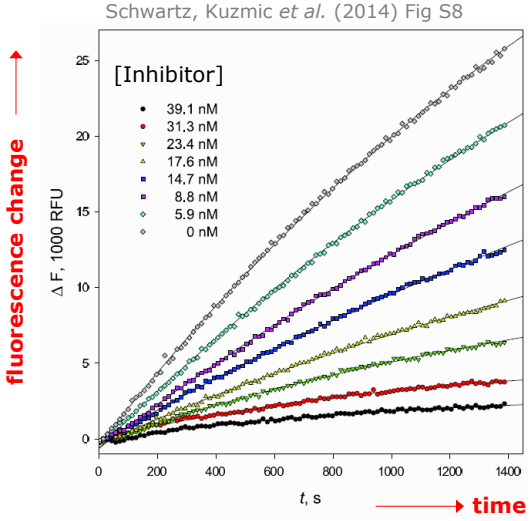
***Kinetic* assay:**

1. Mix enzyme (**E**) + substrate (**S**) + inhibitor (**I**)
2. Continuously observe the appearance of product (**P**) over time
3. Analyze the [Product] vs. time data to determine K_i and k_{inact}

Example experimental data: Neratinib

NERATINIB VS. EFGR T790M / L858R DOUBLE MUTANT

Schwartz, Kuzmic *et al.* (2014) Fig S8



"Textbook method" of data analysis: The Recipe

THIS METHOD ASSUMES THAT THE INHIBITION INDEED PROCEEDS IN **TWO CONSECUTIVE STEPS**

THE TRADITIONAL RECIPE:

1. Fit fluorescence vs. time to an ***exponential*** equation → **"k_{obs}"** values

$$F = F_0 + r_P[P]$$

F … fluorescence signal at time *t*
F_0 … instrument baseline
r_P … concentration-to-signal scaling parameter
[P] … product concentration at time *t*

$$[P] = \frac{v_i}{k_{obs}}\left[1 - \exp(-k_{obs}\ t)\right]$$

t … time
v_i … initial reaction rate
k_{obs} … first-order rate constant

2. Fit "k_{obs}" vs. [Inhibitor] to a ***hyperbolic*** equation → k_{inact} and K_i

$$k_{obs} = k_{inact}\frac{[I]_0}{[I]_0 + K_i}$$

Copeland R. A. (2013) "Evaluation of Enzyme Inhibitors in Drug Discovery", 2nd Ed., Eq. (9.1)(9.2)

"Textbook method" of data analysis: Assumptions

THE "TRADITIONAL RECIPE" RELIES ON SEVERAL IMPORTANT **THEORETICAL** ASSUMPTIONS

1. Control progress curve ([I] = 0) is **strictly linear**

   - Negligibly small substrate depletion over entire time course
   - Absence of product inhibition, enzyme deactivation, etc.

2. Negligibly small **inhibitor depletion**

   - Inhibitor concentrations are very much larger than [Enzyme]
   - Inhibitor concentrations are very much larger than K_i

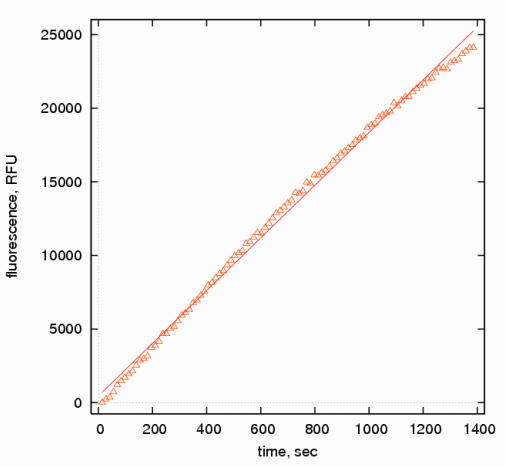
   In other words:

   - The inhibitor is **not** "tight binding".

   - The mole fraction of **unbound inhibitor does not change** over time.

Check linearity: Entire control curve

DynaFit script: ./published/Schw1473/Nera/01-fit-control-R1.txt



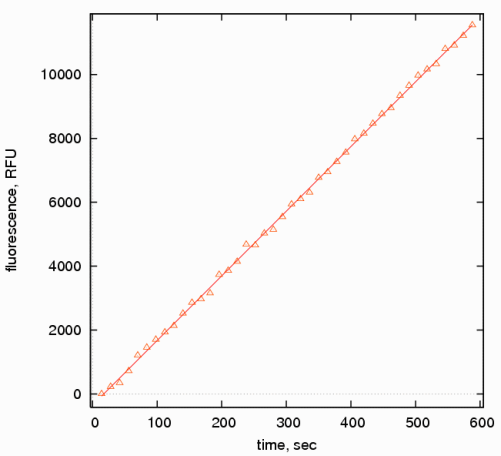
Is this "linear"?

Is this "sufficiently" linear?

Opinions will vary...

Check linearity: The first ten minutes only

DynaFit script: ./published/Schw1473/Nera/02-fit-control-R1.txt



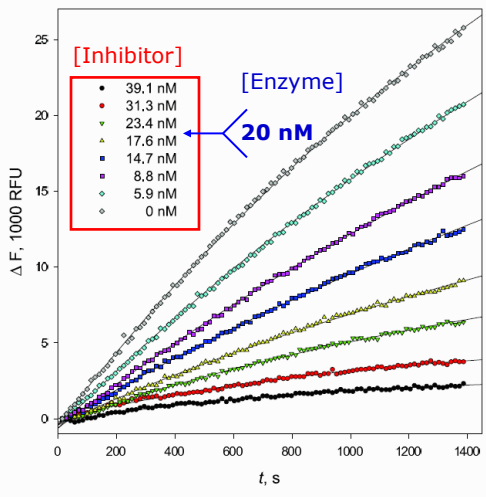
Most definitely linear.

But it's only 50% of the data trace...

Still OK to proceed?

Opinions will vary...

Check concentrations: "Tight binding" or not?



The assumption that [Inhibitor] >> [Enzyme] clearly does not hold.

We have **"tight binding"**,

which is not supposed to happen...

Summary of checks on theoretical assumptions

1. Linearity of control progress curve:

   - There is clearly **some** nonlinearity.

   - However, it is hard to say whether it will be significant or minor.

2. Zero inhibitor depletion:

   - This assumption is most definitely **violated**

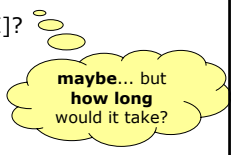
   - However, there is no published study that would clearly demonstrate the impact of this particular violation on K_i and k_{inact} determined by the traditional method.

What are we supposed to do?

Three possible ways to proceed

1. Change the assay: throw out this data and go back to the lab

-**linearity**: can we find conditions where [S] >> K_M?
-**tight binding**: can we find conditions where [E] << [I]?



2. Change the method of data analysis:

-use a **differential equation** model instead of simple algebra
-that method makes no simplifying assumptions of any kind

3. Cross your fingers and hope for a meaningful result

-use the traditional **algebraic** method anyway
-this is the path of least resistance (no changes required)

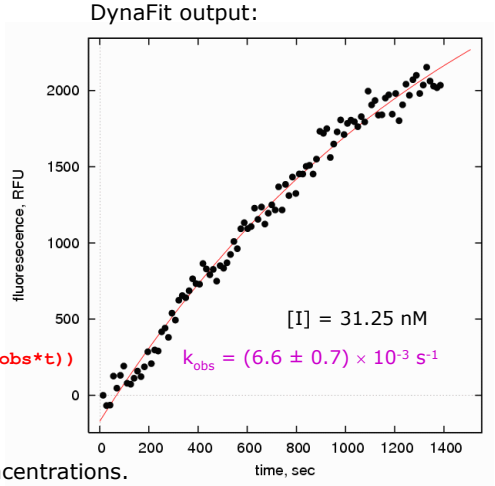
Traditional method of analysis: Determine k_{obs}

DynaFit script: ./published/Schw1473/Nera/04-determine-kobs-R1, -R2, -R3.txt

DynaFit input:

```
[task]
   data = generic
   task = fit

[parameters]
   t, vi, kobs, Fo, rP

[model]
   kobs = 0.001 ?
   vi = 0.001 ?
   Fo = 0 ?
   rP = 5000

   P = (vi/kobs) * (1 - exp(-kobs*t))
   F = Fo + rP*P
```

DynaFit output:



Repeat this for all inhibitor concentrations.

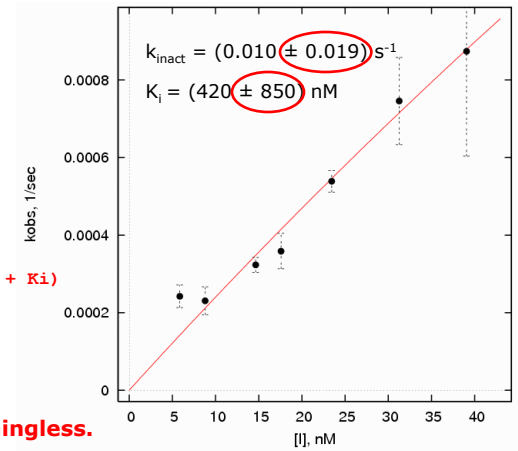
Traditional method of analysis: Determine k_{inact}, K_i

DynaFit script: ./published/Schw1473/Nera/05-fit-kobs-average.txt

DynaFit input:

```
[task]
   data = generic
   task = fit

[parameters]
   I, kinact, Ki

[model]
   kinact = 1 ?
   Ki = 1 ?

   kobs = kinact * I / (I + Ki)
```

DynaFit output:



These results are meaningless.

Determine k_{inact}, K_i from separate replicates

DynaFit script: ./published/Schw1473/Nera/05-fit-kobs-R1, -R2, -R3.txt

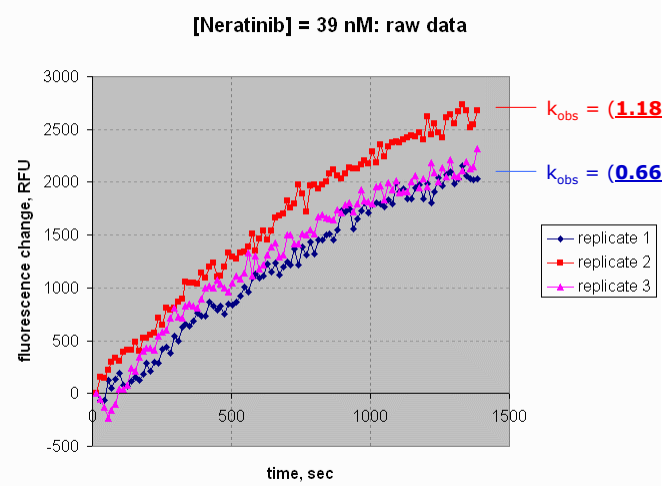
There were three separate experiments (three 96-well plates).
Let's try to average the three best-fit values of k_{inact} and K_i, one from each plate:

	1000× k_{inact}, s^{-1}	K_i, nM
replicate #1	1.6	53
replicate #2	6.1	253
replicate #3	27.7	> 1000000

This is *not* helping...

Is the problem in the poor quality of the raw data?

THIS NERATINIB CONCENTRATION PRODUCED THE "WORST" REPLICATE OF K_{OBS}



k_{obs} = (**1.18** ± 0.07) × 10^{-3} s-1

k_{obs} = (**0.66** ± 0.07) × 10^{-3} s-1

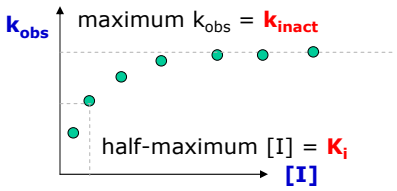
Data are "as good as they will ever get" with this particular assay.
However, the k_{obs} estimates vary up to 100%. This isn't good.

Is the problem with the assumed mechanism?

K_{OBS} PLOTS ARE EITHER **HYPERBOLIC** OR **LINEAR**

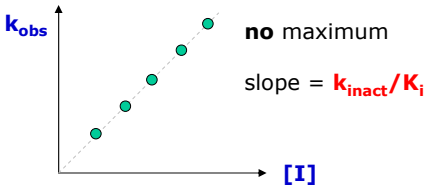
Mechanism "**A**": two steps

$$E + I \underset{}{\overset{K_i}{\rightleftharpoons}} E{\cdot}I \qquad E{\cdot}I \xrightarrow{k_{inact}} E{\sim}I$$

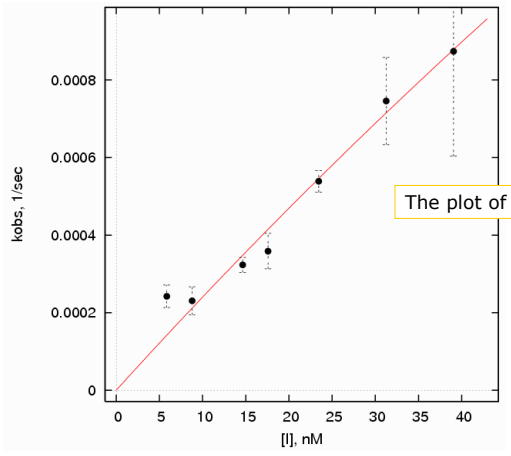


Mechanism "**B**": single step

$$E + I \xrightarrow{k_1} E{\sim}I$$



The k_{obs} plot suggests a single-step mechanism

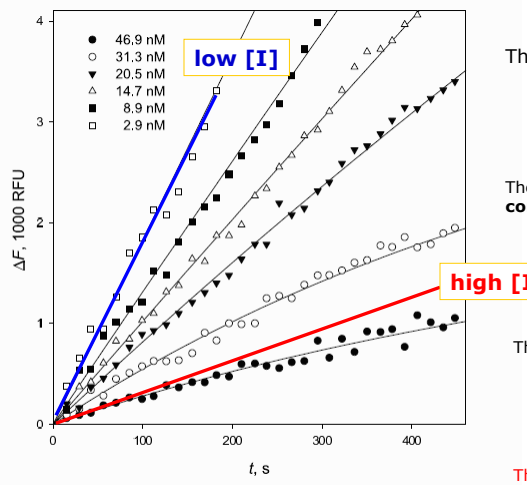


The plot of k_{obs} vs. [I] is essentially **linear**.

Should we just report k_{inact}/K_i (i.e., slope) and be done?

Initial rate data suggests a two-step mechanism

Schwartz, Kuzmic *et al.* (2014) Fig S5



The **initial slope** depends on [I].

There ***must*** be a **non-covalent complex** formed during mixing time.

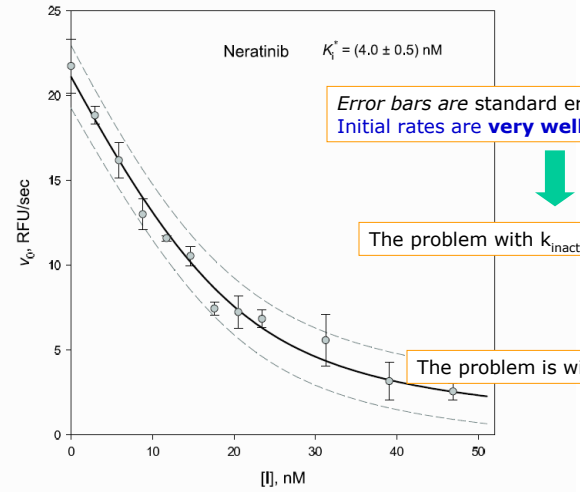
This is a **two-step** mechanism

The "linear" k_{obs} plot makes no sense.

Fig. S 5: Initial rate kinetics of neratinib (one of three replicates):

Binding affinity of initial complex from initial rates

Schwartz, Kuzmic *et al.* (2014) Fig S6



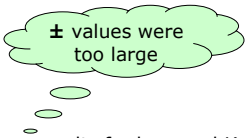
Error bars are standard error from triplicates:
Initial rates are **very well** reproduced (±10%).

The problem with k_{inact} is *not* in data quality.

The problem is with the **fitting model.**

Fig. S 6: Initial rate kinetics of neratinib: *Symbols* - initial reaction rates

Traditional method of kinetic analysis: Summary



1. k_{inact} and K_i could not be determined,

   -the fit of *averaged* k_{obs} values produced **meaningless** results for k_{inact} and K_i
   -the data is good, but the **reproducibility** of k_{inact}, K_i across replicates is very poor

2. The diagnostic plot of k_{obs} vs. [I] contradicts the plot of v_i vs. [I]

   - k_{obs} plot seems to suggest a **single-step** binding mechanism
   - v_i plot (initial rates) strongly suggests a **two-step** mechanism

We can't use these results.

Traditional method of analysis: *Post mortem*

Why did the "classical" algebraic method fail?

1. **Inhibitor depletion**

   Significant inhibition is seen at [Inhibitor] comparable with [Enzyme]

2. **Nonlinear control** progress curve ([I] = 0)

   The nonlinearity is "slight" but, as it happens, has a significant impact.

Two more possible ways to proceed

1. Change the assay: throw out this data and go back to the lab

-**linearity**: can we find conditions where [S] >> K_M?
-**tight binding**: can we find conditions where [E] << [I]?

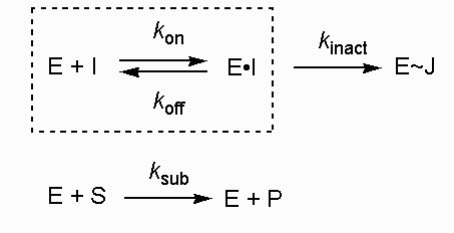
2. Change the method of data analysis:

-use a **differential equation** model instead of simple algebra
-that method makes no simplifying assumptions of any kind

3. Cross your fingers and hope for a meaningful result

-use the traditional **algebraic** method anyway
-this is the path of least resistance (no changes required)

The differential equation model of covalent inhibition

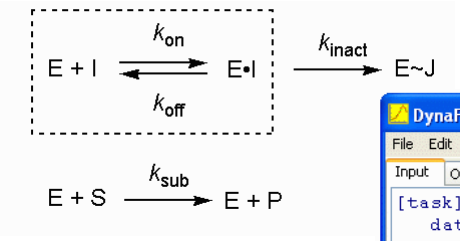


$$d[E]/dt = -k_{sub}[E][S] + k_{sub}[E][S] - k_{on}[E][I] + k_{off}[E.I]$$
$$d[S]/dt = -k_{sub}[E][S]$$
$$d[P]/dt = +k_{sub}[E][S]$$
$$d[I]/dt = -k_{on}[E][I] + k_{off}[E.I]$$
$$d[E.I]/dt = +k_{on}[E][I] - k_{off}[E.I] - k_{inact}[E.I]$$
$$d[E.J]/dt = +k_{inact}[E.I]$$

This model is "integrated numerically".
Whatever that means.

Model of covalent inhibition in DynaFit

DynaFit script: ./published/Schw1473/Nera/06-global-R1.txt



DynaFit input "script":

DynaFit : 06-global-R1.txt

File Edit View Help

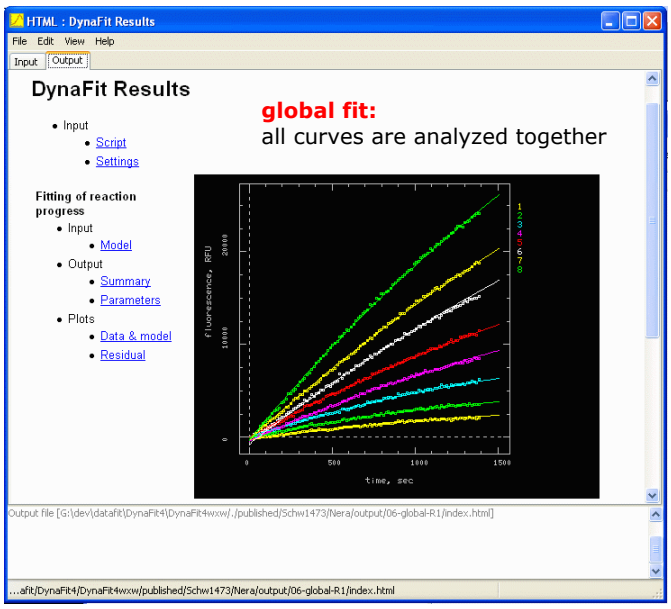
Input Output

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

[mechanism]
   E + S --->  E + P : ksub
   E + I <==> E.I    : kon    koff
   E.I ---> E.J      : kinact

[constants]
   ksub    =  0.02 ?
   kon     =   100
   koff    =     1 ?
   kinact  =     1 ?
```

fixed constant:

"rapid-equilibrium approximation"

Covalent inhibition in DynaFit: Data / model overlay



Covalent inhibition in DynaFit: Model parameters

DynaFit output window:

Optimized Parameters

No.	Par#Set	Initial	Final	Std. Error	CV (%)
#1	ksub	0.02	**0.0141339**	0.000414818	2.93
#2	koff	1	**0.341161**	0.0125877	3.69
#3	kinact	1	**0.000862683**	5.67528e-005	6.58

How do we get K_i out of this?

- We have chosen *micromolar* units throughout this analysis.
- Recall that k_{on} was arbitrarily fixed at **100 µM**-1s-1 ("rapid equilibrium")
- Recall that K_i is defined as k_{off}/k_{on}

$K_i = k_{off}/k_{on} = 0.341 / 100 = 0.00341\ \mu M = 3.4\ nM$

Compare with binding affinity from initial rates

Schwartz, Kuzmic *et al.* (2014) Fig S6

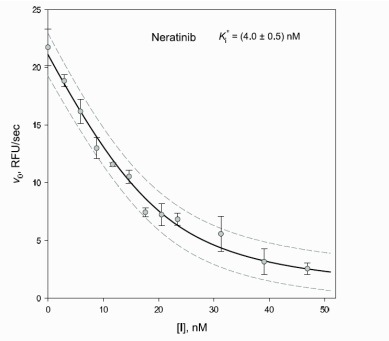
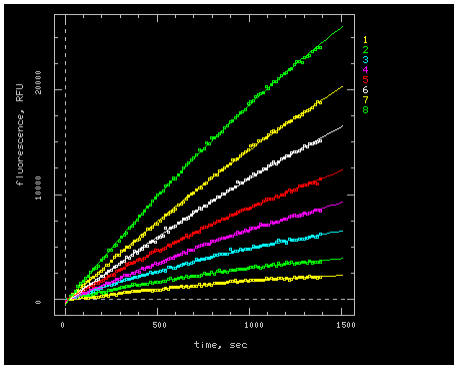


Fig. S 6: Initial rate kinetics of neratinib: *Symbols* - initial reaction rates

Schwartz, Kuzmic *et al.* (2014) Fig S8



Initial rate method:

K_i = **(4.0 ± 0.5)** nM

Global fit, Replicate #1:

K_i = **(3.4 ± 0.1)** nM

Satisfactory agreement between results from two completely independent methods.

Reproducibility of K_i & k_{inact} determinations: Numerical method

NUMERICAL method (differential equations):

Neratinib vs. EGFR double mutant

	K_i, nM	1000 × k_{inact}, s^{-1}
replicate #1	3.4 ± 0.1	0.86 ± 0.05
replicate #2	3.6 ± 0.1	0.85 ± 0.05
replicate #3	3.3 ± 0.1	0.96 ± 0.07

About **10%** reproducibility plate to plate.

Notes:

- Three independent experiments (**three separate 96-well plates** run in sequence).
- Raw data and DynaFit scripts are distributed with the program.

DynaFit script files: ./published/Schw1473/Nera/06-global-R1, -R2, -R3.txt

- Published results (PNAS, 2014) are very slightly different: concentrations were optimized.

Compare results with the classical algebraic method

ALGEBRAIC method ("Traditional Recipe") – using the same data:

Neratinib vs. EGFR double mutant

	K_i, nM	1000 × k_{inact}, s^{-1}
replicate #1	53 ± 22	1.6 ± 0.4
replicate #2	~~254 ± 442~~	~~6.1 ± 9.5~~
replicate #3	~~> 1 000 000~~	~~27 ± 162000~~
average k_{obs}	~~420 ± 850~~	~~10 ± 19~~

nonsense values (± too large)

But what can you do with this?

Bringing in differential equations *can* be expensive:

-**time** to learn and train
-**money** for new software

Why not just stick with k_{inact}/K_I, like most people do?

K_i and k_{inact} as *distinct* determinants of cellular potency

Schwartz, Kuzmic, *et al.* (2014) Fig S10

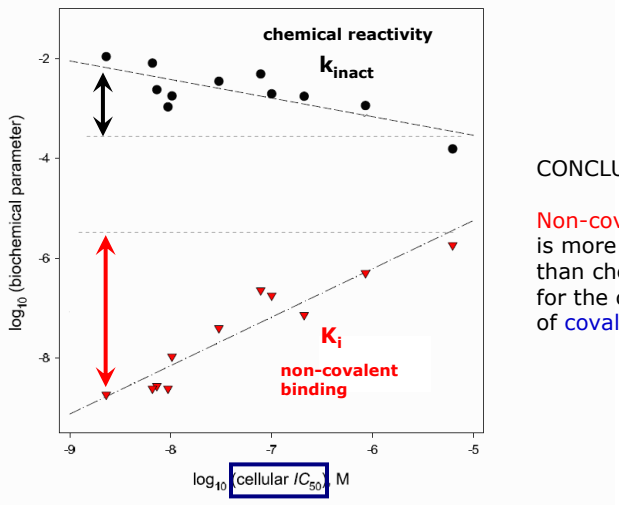


Fig. S 10: Correlation of covalent inhibitor kinetic constants toward EGFR-L858R/T790M with cellular potency (inhibition of EGFR-L858R/T790M autophosphorylation in H1975 tumor cells)

CONCLUSIONS:

Non-covalent initial binding
is more important
than chemical reactivity
for the cellular potency
of covalent anticancer drugs.

K_i and k_{inact} as determinants of cellular potency: Panel of 154

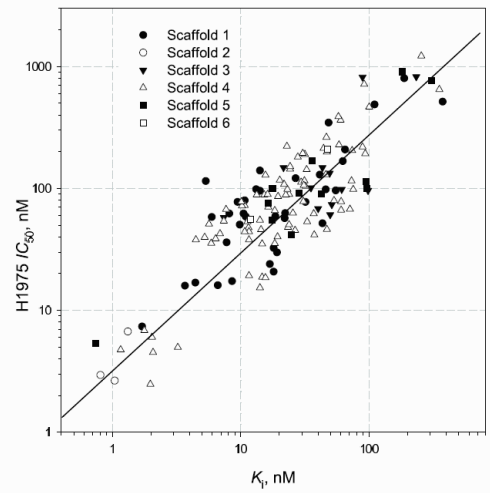


Fig. S 11: Correlation of dissociation constants of the initial non-covalent enzyme/inhibitor complex, as measured by the K_i values for EGFR-L858R/T790M double mutant, with cellular potency (inhibition of EGFR-L858R/T790M autophosphorylation in H1975 tumor cells) for 154 compounds spanning six structural classes.

Schwartz, Kuzmic, *et al.* (2014) Fig S11

Non-covalent K_d
vs.
Cellular IC_{50}

strong correlation
for a larger panel

Summary and conclusions

The **traditional method** to analyze covalent inhibition data breaks down if:

- the "control" progress curve ([I] = 0) is nonlinear
- there is inhibitor depletion (i.e. when [E] is comparable with [I])
- or if both situations apply at the same time.

The nonlinearity in the control curve ([I] = 0) can very "slight",
by subjective standards, and yet the detrimental effect is profound.

There is no need to redesign the assay
to force the experimental data into conformity with the classic algebraic model.
There is nothing "wrong" with the data. It just cannot be analyzed in the usual way.

Instead one must use a differential equation model (DynaFit).

Using this general numerical approach (DynaFit), we were able
to measure the **initial binding constants of covalent drugs** and
determine the impact of K_d on **drug potency in cellular assays**.

Acknowledgments

- Brion Murray – Pfizer
  Leader on the PNAS paper, and in other ways
- Art Wittwer – Confluence Technologies (formerly Pfizer)
  PK/PD initial scripts (and many other ideas)
- Phillip Schwartz – Takeda (formerly Pfizer)
  Data collection for EGFR inhibitors
- Jim Solowiej - Pfizer
  Data collection for EGFR inhibitors

Questions ?

http://www.biokin.com