

Time-dependent inhibition of 5 α -steroid reductase

A DynaFit Tutorial

BioKin Technical Note TN-2015-06

Rev. 1.02 (Nov 24, 2015)

Petr Kuzmic

BioKin Ltd.

CONTENTS

Abstract.....	2
Introduction	3
DynaFit: Step-by-step tutorial	7
1. Export raw data in ASCII (text) format.....	7
2. Create DynaFit script	7
Insert a distributed template script.....	8
Modify the default script template	9
3. Perform model discrimination analysis	15
4. Estimate confidence intervals for top-ranked models.....	18
Confidence intervals for the general two-step model	18
Confidence intervals for the rapid-equilibrium model.....	23
Discussion.....	25
References.....	28

Abstract

Experimental data initially published by Moss, Kuzmic, *et al.* (*Biochemistry* **35**, 3457, 1996) are re-analyzed here using a variety of techniques that were introduced into the DynaFit software (*Anal. Biochem.* **237**, 260, 1996) since the original publications first appeared. The experimental system involves "slow, tight" inhibition of 5α -steroid reductase by a heterocyclic testosterone analog inhibitor. The results revealed a two-step kinetic mechanism of inhibition, involving essentially instantaneous equilibration to form initially a "loose" enzyme-inhibitor complex. The initial complex is characterized by equilibrium dissociation constant $K_d^{(ini)} = 5$ nM. The initial complex rearranges over time to form the final "tight" complex, characterized by the total equilibrium dissociation constant $K_d^{(tot)} = 0.17$ nM. The reversible rearrangement occurs within approximately four minutes. A detailed, step-by-step tutorial is presented, which shows (a) how to construct the appropriate DynaFit input file and (b) how best to interpret the numerical results.

Introduction

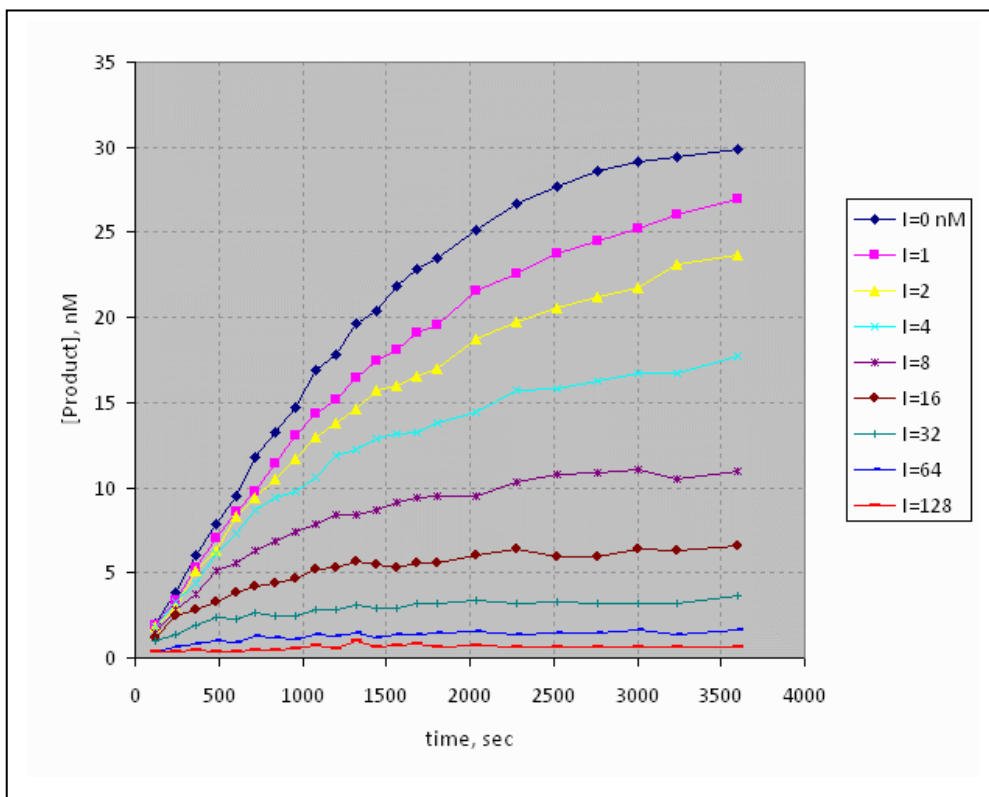
Background

The purpose of this DynaFit [1, 2] tutorial is to present detailed, step-by-step method to determine the kinetic mechanism of "slow, tight" inhibition [3] and the requisite values of microscopic rate constants. The experimental data are taken from ref. [4], Figures 2 and 7. The originally published data (percent product formation) were converted to absolute product concentrations, as shown in *Table 1*. The entries in *Table 1* represent nanomolar concentrations of dihydrotestosterone (DHT) as the final reaction product. The initial concentration of testosterone (T) was 31 nM.

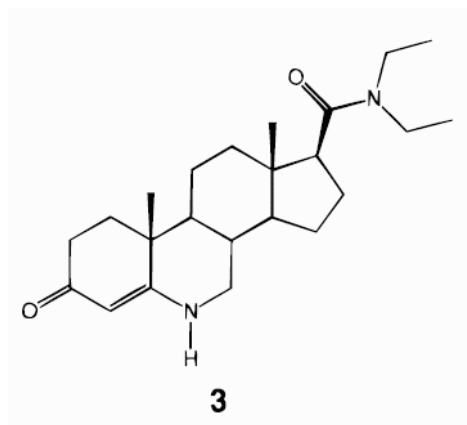
t, sec	I=0 nM	I=1	I=2	I=4	I=8	I=16	I=32	I=64	I=128
120	1.99	1.89	1.84	1.82	1.49	1.22	0.98	0.37	0.33
240	3.86	3.36	2.93	3.33	2.79	2.51	1.41	0.68	0.40
360	6.05	5.26	5.14	4.36	3.78	2.80	1.92	0.82	0.47
480	7.90	7.08	6.29	6.08	5.15	3.30	2.34	1.00	0.35
600	9.53	8.63	8.31	7.30	5.53	3.87	2.32	0.95	0.37
720	11.77	9.78	9.44	8.67	6.29	4.21	2.67	1.25	0.42
840	13.22	11.43	10.54	9.42	6.85	4.42	2.44	1.22	0.42
960	14.70	13.08	11.73	9.75	7.44	4.68	2.49	1.10	0.56
1080	16.87	14.36	12.99	10.64	7.85	5.19	2.81	1.34	0.73
1200	17.80	15.17	13.79	11.90	8.45	5.30	2.83	1.32	0.58
1320	19.65	16.48	14.64	12.23	8.44	5.64	3.15	1.43	0.98
1440	20.35	17.48	15.70	12.90	8.67	5.44	2.95	1.22	0.60
1560	21.88	18.13	16.00	13.14	9.13	5.31	2.95	1.34	0.69
1680	22.84	19.08	16.50	13.26	9.38	5.56	3.23	1.35	0.83
1800	23.53	19.53	17.04	13.80	9.53	5.57	3.17	1.45	0.64
2040	25.09	21.53	18.76	14.47	9.48	6.06	3.37	1.52	0.74
2280	26.67	22.60	19.73	15.68	10.34	6.36	3.17	1.33	0.68
2520	27.66	23.72	20.59	15.81	10.75	5.97	3.31	1.46	0.61
2760	28.63	24.50	21.17	16.23	10.90	5.98	3.21	1.48	0.65
3000	29.12	25.21	21.74	16.75	11.09	6.38	3.21	1.65	0.60
3240	29.42	26.01	23.11	16.73	10.55	6.31	3.21	1.41	0.65
3600	29.91	26.99	23.67	17.75	10.93	6.62	3.63	1.64	0.61

Table 1

Data from Moss *et al.*, ref. [4], Figures 2 and 7, converted from percent product formation to product concentration, nM. The two lowest concentration curves ($[I] = 1$ nM and 2 nM) are previously unpublished data, from the same series of experiments. The original publication [4] mistakenly assigned inhibitor concentration 68 nM to the progress curve collected at $[I] = 64$ nM.

**Figure 1**

Plot of raw experimental data listed in *Table 1*. See text for details.

**Figure 2**

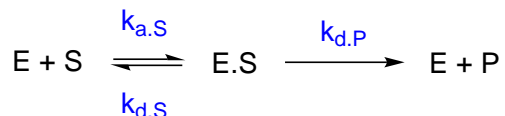
Structure of 5a-reductase inhibitor **3** [4] utilized in this study.

Experimental Methods

Several Eppendorf tubes were prepared containing 31 nM [1,2,6,7-³H(N)]testosterone in 1.04 mL of the reaction buffer (17.6 mM imidazole, 17.6 mM diethanolamine, 0.3 M KCl, 13.2 mM succinic acid, 1 mM DTT at pH 7.0, 37 °C). The tubes contained various concentrations of the inhibitor, between 1 and 128 nM. To each tube, 60 μL of enzyme (50 μM final concentration) was added to initiate the reaction. Aliquots were removed at various reaction times and quenched with excess ethanol. Product formation was monitored by HPLC (separating tritiated testosterone, T, and tritiated dihydrotestosterone, DHT) with radiometric detection, as described in ref. [5]. The T/DHT peak area ratios were utilized to compute the percent conversion of substrate to product at each reaction time. The raw data were subsequently converted to absolute product concentrations, in nM units (see *Table 1*). For additional details see [4].

Substrate kinetic constants

In this kinetic analysis will deploy the "divide and conquer" strategy, in the sense that *substrate* kinetic parameters will be determined separately, and subsequently treated as *fixed* model parameters in the analysis of inhibition data. In assigning numerical values of the substrate kinetic constants, we will utilize results previously obtained in BioKin Technical Note TN-2015-05 [6].



Scheme 1

Specifically, for the Michaelis-Menten mechanism shown in *Scheme 1*, we will assume the following *fixed* values of the microscopic rate constants:

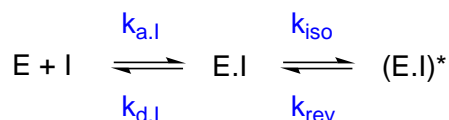
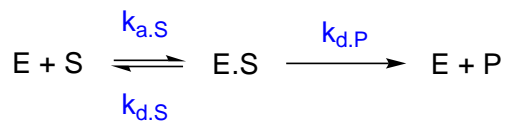
Constant	Unit	Value
$k_{a,S}$	$10^9 \text{ M}^{-1}\text{s}^{-1}$	1
$k_{d,S}$	s^{-1}	19.9
k_{cat}	s^{-1}	0.611

Table 2

Assumed values of substrate kinetic constants. For details see ref. [6].

Assumed inhibition mechanism

In this DynaFit tutorial we will examine seven different variants of the general two-step reversible inhibition mechanism shown in *Scheme 2*.

**Scheme 2**

The seven variations on this most general mechanism are summarized in *Table 2*.

#	model	description	$k_{a,S}$	$k_{d,S}$	k_{iso}	k_{rev}
1	2S	two-step				
2	2S-R	two-step, rapid equilibrium	1			
3	2S-C	two-step, covalent				0
4	2S-RC	two-step, rapid equilibrium, covalent	1			0
5	1S	one-step			0	0
6	1S-R	one-step, rapid-equilibrium	1		0	0
7	1S-C	one-step, covalent		0	0	0

Table 2

Classification of "slow, tight" inhibition models investigated in this study. For details see text.

In Table 2, the entry "1" under a particular rate constant value means that this rate constants is held fixed in the data fitting at a value presumably approaching or even surpassing the diffusion control limit (in this case $1.0 \text{ nM}^{-1}\text{s}^{-1}$ i.e. $10^9 \text{ M}^{-1}\text{s}^{-1}$). The entry "0" signifies that the particular step is not present in the reduced kinetic mechanism. No entry means that the given rate constant is allowed to attain an arbitrary value in data fitting.

DynaFit: Step-by-step tutorial

1. Export raw data in ASCII (text) format

Experimental data for DynaFit processing needs to be converted to the ASCII (plain text) format, either tab-delimited, space-delimited, or comma-delimited.

In this step we will export an Excel data sheet to a comma-delimited .CSV ("Comma Separated Values") plain text file, which will be subsequently analyzed by DynaFit.

PROCEDURE

1. Navigate to the main DynaFit installation directory [**DynaFit4**]. From now on this directory will be abbreviated as **"/**".
2. Navigate to subdirectory **./technotes/2015/06/data/excel**
3. Open file **Moss-1996-Fig7.xls**
4. Go to sheet **"data"**
5. Select menu **"File ... Save As"**
6. Navigate to subdirectory **./technotes/2015/06/data**
This directory is one level up from the subdirectory **excel**.
7. Click on **"Save as Type"** drop down menu
8. Select **CSV (Comma delimited)**
9. Type **data** into the **File name** field.
The default file extension **.csv** will be added automatically.

The final layout of files and directories is shown schematically below:

```
[DynaFit4]
|
+---- [technotes]
      |
      +---- [2015]
            |
            +---- [06]
                  |
                  +---- [data]
                          |
                          +---- data.csv
                          |
                          +---- [excel]
                                  |
                                  +---- Moss-1996-Fig7.xls
```

2. Create DynaFit script

DynaFit scripts can be created in several different ways:

- From scratch (by modifying the built-in standard template).

- By modifying one of distributed template fragments (subdirectory **templates**).
- By modifying one of the complete example scripts.

In this step we will create a DynaFit script file by reusing one of the distributed templates. The script file will be subsequently used to determine substrate kinetic parameters from progress curve data.

Insert a distributed template script

PROCEDURE

1. Start DynaFit. This will display the text editor window containing the start-up minimal template.
2. Press "**Ctrl+N**" or equivalently select menu "**File ... New**".
3. Press **N** or equivalently click on the [No] (do not save changes) button.
4. Press "**Ctrl+S**" or equivalently select menu "**File ... Save**".
5. Navigate to directory `./technotes/2015/06`, where "." is the main DynaFit installation directory.
6. Type **models** into the **File name** field. The file name extension **.txt** will be added automatically.
7. Press "**Alt+E ... I ... F ... F [Enter]**" or equivalently select menu "**Edit ... Insert ... File...**".
8. Navigate to directory `./templates/enzyme-kinetics/inhibition/progress`.
9. Double click to select file **all-models.txt**.
10. Press "**Ctrl+S**" to save the modified script file.

The final layout of files and directories is shown schematically below:

```
[DynaFit4]
|
+--- [technotes]
    |
    +--- [2015]
        |
        +--- [06]
            |
            +--- models.txt
            |
            +--- [data]
                |
                +--- data.csv
                |
                +--- [excel]
                    |
                    +--- Moss-1996-Fig7.xls
```

Our goal in the next several steps is to modify (edit) the default template text.

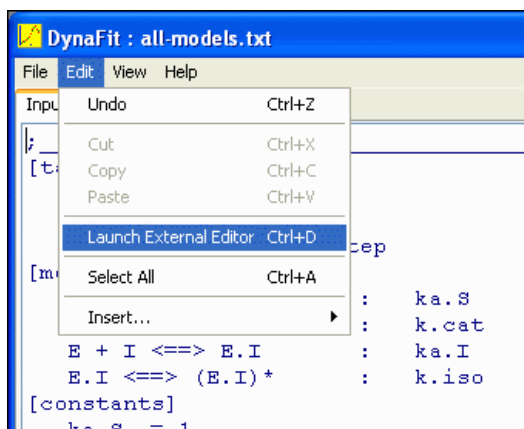
Modify the default script template

Modify the *[constants]* section

The present version of the DynaFit text editor has only limited capabilities. For example, it cannot simultaneously find and replace *multiple* occurrences of a particular text fragment. For this reason, we will utilize the default external text editor installed on the user's computer. In most cases it will be Notepad under Microsoft Windows, or a corresponding simple text editor under Mac OS.

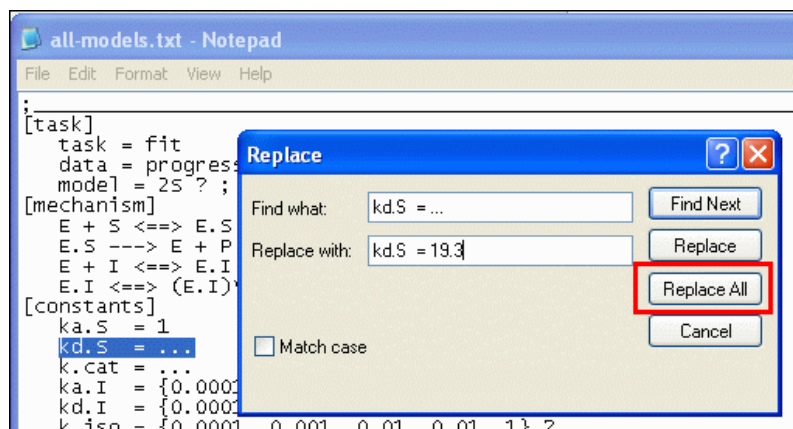
PROCEDURE

1. Press **Ctrl+D** or equivalently select menu "**Edit ... Launch External Editor**"



In the following it will be assumed that the default text editor is Notepad under Microsoft Windows.

2. Scroll to the top of the file in the editor.
3. Highlight the text fragment **kd.S = ...** in the [constants] section.
4. Press **Ctrl+C** to copy the text fragment into the system clipboard.
5. Press **Ctrl+H** or equivalently select menu "**Edit ... Replace**"
6. Press **Ctrl+V** to past the copied text fragment
7. Press [**Tab**] or click on the "Replace with:" text area
8. Press **Ctrl+V** to paste the copied text fragment
9. Delete "...", replace with **19.9** (no quotation marks), press [**Replace All**] button.



10. Use the same procedure to replace the k_{cat} value everywhere:

ORIGINAL	<code>kcat = ...</code>
MODIFIED	<code>kcat = 0.611</code>

11. Save the modified script file: **Ctrl+S**.
12. Close the external text editor (subsequent editing can be conveniently performed with DynaFit).
13. Reload the modified script file into DynaFit:
 - a. Select menu "**File ... Open**"
 - b. Navigate to directory **./technotes/2015/06/**
 - c. Select file **models.txt** [OK]

Modify the [concentrations] section

The purpose of this section is to assign numerical values of initial concentrations.

ORIGINAL

```
[concentrations]
E = ...
S = ...
```

MODIFIED

```
[concentrations]
E = 0.05
S = 31
```

NOTES

- The concentration unit is **nanomolar**. Thus the above notation represents the initial concentrations $[E]_0 = 50 \text{ pM}$, $[S]_0 = 31 \text{ nM}$.
- The concentration units must be identical throughout the entire DynaFit script and must correspond to the concentrations units employed in the experimental data files.

Modify the [responses] section

The purpose of this section is to assign numerical values of physical constants that linearly transform reacting species concentrations to the observed experimental signal.

ORIGINAL

[responses]
P = ...

MODIFIED

[responses]
P = 1 ?

NOTES

- The raw experimental data file in this particular case contains nanomolar concentrations of the reaction product, as determined by HPLC. Therefore, the proportionality constant linking the data values to concentration values is simply unity.
- The question mark means that we wish to determine from the experimental data any possible non-unit value of the instrument scaling factor (i.e., essentially a systematic error of the HPLC peak integrator).

Modify the [data] section

This section specifies the location of experimental data file(s) located on the disk.

In preparation for modifying the [data] template:

- Navigate to directory **./technotes/2015/06/data/excel**
- Open the Excel file **Moss-1996-Fig7.xls**
- Switch to the sheet named **script**

Note that most of the numerical values in this sheet (all grayed out cells) are auto-generated by simple Excel formulas, representing the 1:2 dilution series:

	A	B	C	D	E	F
1	[data]					
2	sheet ...					
3	offset 0 ?					
4	column 2	concentration I =	0	label I =	0	
5	column 3	concentration I =	1	label I =	1	
6	column 4	concentration I =	2	label I =	2	
7	column 5	concentration I =	4	label I =	4	
8	column 6	concentration I =	8	label I =	8	
9	column 7	concentration I =	16	label I =	16	
10	column 8	concentration I =	32	label I =	32	
11	column 9	concentration I =	64	label I =	64	
12	column 10	concentration I =	128	label I =	128	

In the modification procedure below, we will utilize rows 5 through 11 above, representing columns No. 2 through 9 of the raw data sheet.

ORIGINAL

```
[data]
  sheet    ...
  offset  auto ?
  monitor  E, E.S, E.I, (E.I)*
  column   ...
```

PROCEDURE

1. Delete the "." placeholder following the key word **sheet**.
2. Press "**Alt+E ... I ... F [Enter]**" or equivalently select menu "**Edit ... Insert ... File Name...**"
3. Navigate to directory **./technotes/2015/06/data**.
4. Select **CSV files (*.csv)** from the **File types** drop-down menu.
5. Double-click on file **data.csv** to select it.
6. **Delete the line containing the text column ...**
Keep the cursor at the beginning of the deleted line
7. Switch to the Excel file **Moss-1996-Fig7.xls**
8. Highlight the lines representing **column 3** through **column 9**
9. Press **Ctrl+C** to copy the highlighted text into the system clipboard.
10. Switch to DynaFit input window
11. Press **Ctrl+V** to paste the text on the line following the **monitor** keyword.
See below for the desired appearance of the [data] section of the script:

```
...
[data]
  sheet    ./technotes/2015/06/data/data.csv
  offset  auto ?
  monitor  E, E.S, E.I, (E.I)*
  column   3          | concentration I =      1          | label I = 1
  column   4          | concentration I =      2          | label I = 2
  column   5          | concentration I =      4          | label I = 4
  column   6          | concentration I =      8          | label I = 8
  column   7          | concentration I =     16          | label I = 16
  column   8          | concentration I =     32          | label I = 32
  column   9          | concentration I =     64          | label I = 64
[output]
...
```

Modify the [output] section

This section specifies the location of the output files to be generated by DynaFit.

ORIGINAL

```
[output]
```

```
directory ...
```

MODIFIED

[output]

```
directory ./technotes/2015/06/output/models
```

PRODEDURE

1. Delete the placeholder notation ("...") following the keyword **directory**.
2. Type "**Alt+E ... I ... D ... D ... [Enter]**" or equivalently select menu "**Edit ... Insert ... Directory name**"
3. Navigate to subdirectory **./technotes/2015/06** [OK]
4. Type the following at the end of the inserted line:
/output/models

ALTERNATE PROCEDURE

1. Copy the text **./technotes/2015/05** from the [data] section.
2. Paste it to replace the placeholder ("...") following the keyword **directory**.
3. Type the following at the end of the inserted line:
/output/models

Save ("**Ctrl+S**") the final edited version of the script file.

Final form of the modified template text

After all modifications described above, the DynaFit script file **models.txt** should read as follows:

```

;_____
[task]
  task = fit
  data = progress
  model = 2S ? ; two-step
[mechanism]
  E + S <=> E.S      :   ka.S      kd.S
  E.S ---> E + P     :   kd.P
  E + I <=> E.I      :   ka.I      kd.I
  E.I <=> (E.I)*    :   k.iso     k.rev
[constants]
  ka.S = 1
  kd.S = 19.9
  kd.P = 0.611
  ka.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
  kd.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
  k.iso = {0.0001, 0.001, 0.01, 0.01, 1} ?
  k.rev = {0.0001, 0.001, 0.01, 0.01, 1} ?
[concentrations]
  E = 0.05
  S = 31
[responses]
  P = 1 ?

```

```

[data]
  sheet    ./technotes/2015/06/data/data.csv
  offset   auto ?
  monitor  E, E.S, E.I, (E.I)*
  column   3      | concentration I =      1      | label I = 1
  column   4      | concentration I =      2      | label I = 2
  column   5      | concentration I =      4      | label I = 4
  column   6      | concentration I =      8      | label I = 8
  column   7      | concentration I =     16      | label I = 16
  column   8      | concentration I =     32      | label I = 32
  column   9      | concentration I =     64      | label I = 64
[output]
  directory ./technotes/2015/06/output/models
;
[task]
  task = fit
  data = progress
  model = 2S-R ? ; two-step, rapid equilibrium
[constants]
  ka.S = 1
  kd.S = 19.9
  kd.P = 0.611
  ka.I = 1
  kd.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
  k.iso = {0.0001, 0.001, 0.01, 0.01, 1} ?
  k.rev = {0.0001, 0.001, 0.01, 0.01, 1} ?
;
[task]
  task = fit
  data = progress
  model = 2S-C ? ; two-step, covalent
[mechanism]
  E + S <=> E.S      :   ka.S      kd.S
  E.S ---> E + P      :   kd.P
  E + I <=> E.I      :   ka.I      kd.I
  E.I ---> (E.I)*    :   k.cov
[constants]
  ka.S = 1
  kd.S = 19.9
  kd.P = 0.611
  ka.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
  kd.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
  k.cov = {0.0001, 0.001, 0.01, 0.01, 1} ?
;
[task]
  task = fit
  data = progress
  model = 2S-RC ? ; two-step, rapid equilibrium, covalent
[constants]
  ka.S = 1
  kd.S = 19.9
  kd.P = 0.611
  ka.I = 1
  kd.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
  k.cov = {0.0001, 0.001, 0.01, 0.01, 1} ?
;
[task]
  task = fit
  data = progress
  model = 1S ? ; one-step
[mechanism]
  E + S <=> E.S      :   ka.S      kd.S
  E.S ---> E + P      :   kd.P

```

```

E + I <=> E.I      :   ka.I      kd.I
[constants]
ka.S = 1
kd.S = 19.9
kd.P = 0.611
ka.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
kd.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
;
[task]
task = fit
data = progress
model = 1S-R ? ; one-step, rapid equilibrium
[constants]
ka.S = 1
kd.S = 19.9
kd.P = 0.611
ka.I = 1
kd.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
;
[task]
task = fit
data = progress
model = 1S-C ? ; one-step, covalent
[mechanism]
E + S <=> E.S      :   ka.S      kd.S
E.S ---> E + P     :   kd.P
E + I ---> E.I     :   k.cov
[constants]
ka.S = 1
kd.S = 19.9
kd.P = 0.611
k.cov = {0.0001, 0.001, 0.01, 0.01, 1} ?
[end]
;
;

```

3. Perform model discrimination analysis

PROCEDURE

1. Press **Ctrl+U** or equivalently select menu "**File ... Run**".

The complete analysis will take approximately 5 minutes using a typical middle-grade computer equipped with a 2 GHz microprocessor.

2. Press **Ctrl+J** or select menu "**View ... Results in External Browser**".
3. Navigate in the various parts of the auto-generated output

NOTES

- Upon execution, DynaFit automatically writes a collection of various output files (HTML, GIF, EPS, TXT) into a newly created directory named in the **[output]** section.
- The main output file is named **index.html**. It is automatically displayed in the DynaFit main output windows.

- The output files can be navigated using a rudimentary HTML web browser built into DynaFit. However, the DynaFit web browser is *not* equipped to display output in highly stylized form. Instead, it is advantageous to view results of fit using the default web browser installed on the user's computer.
- It is advantageous to use the following shortcuts either in DynaFit or in the external web browser:

Alt+Left Arrow go to the previous page
Alt+Right Arrow go to the next page

Examine results of model discrimination analysis

The main output file **index.html** contains a summary of model discrimination analysis as shown in *Table 3*, where:

- n_D is the number of data points;
- n_p is the number of adjustable model parameters;
- SS_{rel} is the relative sum of squares, where the value 1.000 is by definition assigned to the model that produced the lowest sum of squared deviations;
- AIC_c is the second-order Akaike Information Criterion (ref. [7], p. 66);
- $w^{(AIC)}$ is the corresponding Akaike weight (ref. [7], p. 75);
- BIC_c is the second-order Bayesian Information Criterion (ref. [8, 9]); and
- $w^{(BIC)}$ is the corresponding Bayesian weight.

model	n_D	n_p	SS_{rel}	AIC_c	$w^{(AIC)}$	BIC_c	$w^{(BIC)}$
2S	154	6	1.000	-435.3	0.503	-414.1	0.182
2S-R	154	5	1.022	-434.1	0.278	-415.9	0.459
2S-C	154	5	1.025	-433.7	0.218	-415.4	0.360
2S-RC	154	4	1.218	-409.4	0	-394.2	0
1S	154	4	1.411	-386.6	0	-371.5	0
1S-R	154	3	29.745	80.7	0	92.8	0
1S-C	154	3	3.507	-248.6	0	-236.4	0

Table 3

Results of model discrimination analysis. For details see text.

The sum of squares differences (SSQ_{rel}) in *Table 3* can be safely ignored, because the models differ in the number of adjustable parameters. A fitting model with a larger number of adjustable parameters will always produce a lower sum of squares, whether or not this decrease is statistically significant. Instead, the important components of *Table 3* are the columns listing AIC and BIC values, as well as the corresponding statistical weights.

In particular, if a particular model were to be declared as *significantly* more plausible than certain another models, we need to see the following two characteristics satisfied:

1. The preferred model must be associated with **AIC** and **BIC** values that are at least **10 units** lower than any of the less favored models.
2. The preferred model must be associated with Akaike and Bayesian **weights** higher than **0.95** (95% percent model probability).

After applying the above empirical rules for model selection, we see that the only four models that can be safely excluded from consideration are the three one-step mechanisms plus model "2S-RC".

In contrast, it is *impossible to decide* conclusively in favor of either of three inhibition mechanisms that are *statistically tied for first place* in Table 3 (namely, the tree two-step models labeled "2S", "2S-R", and "2S-C").

However, the two-step, covalent inhibition mechanism "2S-C" can be excluded based on independent experiments [4]. The enzyme was preincubated with excess inhibitor for extended period of time. Subsequently, the enzyme-inhibitor mixture was highly diluted to see if any residual enzyme activity could be recovered. This was indeed the case, which is a direct experimental proof of reversibility of inhibitor binding.

This leaves us with only two kinetic mechanisms of inhibition that need to be ranked by some suitable method, namely, the fitting models labeled "2S" and "2S-R". The only difference between the two kinetic mechanism is that in one case ("2S") the association rate constants k_{aI} is treated as an *optimized* model parameter, whereas in the second case ("2S-R") the the association rate constants k_{aI} is treated as a *fixed* model parameter, se to an appropriate minimum value.

The ultimate model-selection decision can be made be using the likelihood profile method as described in the next section.

4. Estimate confidence intervals for top-ranked models

In this task we will utilize the best-fit parameter values obtained by the model discrimination analysis above. We will focus exclusively on two alternate models, namely "2S" and "2S-R".

Confidence intervals for the general two-step model

Clone and edit the model-discrimination script

PROCEDURE

1. Open the script file **models.txt** (see above) if it isn't already open.
2. Select menu "**File ... Save As...**"
3. Type **confid-2S** into the **File name** field. The file name extension **.txt** will be added automatically.
4. Edit the inhibitor rate constant values:

CURRENT:

```
ka.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
kd.I = {0.0001, 0.001, 0.01, 0.01, 1} ?
k.iso = {0.0001, 0.001, 0.01, 0.01, 1} ?
k.rev = {0.0001, 0.001, 0.01, 0.01, 1} ?
```

MODIFIED:

```
ka.I = 0.001 ??
kd.I = 0.01 ??
k.iso = 0.001 ??
k.rev = 0.0001 ??
```

These initial estimates (within an order of magnitude) are based on the best-fit values obtained for model "2S" by using the script file **models.txt** above.¹

Note the presence of *two question marks (??)*. This notation signifies that we wish to determine the nonsymmetrical confidence intervals using the "profile-*t*" method [10-12].

5. Modify the output directory name:

CURRENT:

¹ Please carefully note the **order of magnitude** of each rate constant being estimated. It is the nature of nonlinear least-squares regression that the initial estimate often needs to be reasonably close to the "true" value (which is of course unknown at the beginning of the investigation).

```
[output]
  directory ./technotes/2015/06/output/models
```

MODIFIED:

```
[output]
  directory ./technotes/2015/06/output/confid-2S
```

6. Delete all **[task]** sections following the first one.
Make sure that the last line of the modified script does contain **[end]**.
7. Delete the (now) superfluous **model** line in the **[task]** section.
8. Press "**Ctrl+S**" to save the modified script file.

Final form of the modified script

```

;
[task]
  task = fit
  data = progress
[mechanism]
  E + S <=> E.S      :   ka.S      kd.S
  E.S ---> E + P     :   kd.P
  E + I <=> E.I      :   ka.I      kd.I
  E.I <=> (E.I)*     :   k.iso     k.rev
[constants]
  ka.S = 1
  kd.S = 19.9
  kd.P = 0.611
  ka.I = 0.001 ??
  kd.I = 0.01 ??
  k.iso = 0.001 ??
  k.rev = 0.0001 ??
[concentrations]
  E = 0.05
  S = 31
[responses]
  P = 1 ?
[data]
  sheet  ./technotes/2015/06/data/data.csv
  offset auto ?
  monitor E, E.S, E.I, (E.I)*
  column 3      | concentration I =      1      | label I = 1
  column 4      | concentration I =      2      | label I = 2
  column 5      | concentration I =      4      | label I = 4
  column 6      | concentration I =      8      | label I = 8
  column 7      | concentration I =     16      | label I = 16
  column 8      | concentration I =     32      | label I = 32
  column 9      | concentration I =     64      | label I = 64
[output]
  directory ./technotes/2015/06/output/confid-2S
[end]
;
;

```

Perform confidence interval estimation

PROCEDURE

1. Select menu "**File ... Run**" or equivalently press **Ctrl+U**
2. Select menu "**View ... Results in External Browser**" or press **Ctrl+J**
3. Click on the "Parameters" link

Examine the confidence intervals:

The summary of results displayed in the "Parameter" page of the DynaFit output is shown in *Table 4*.

Param	Initial	Final	Std. Error	CV (%)	Low ^(a)	High ^(a)
ka.I	0.001	0.00115	0.00037	32.2	0.000736	> 69.3
kd.I	0.01	0.0057	0.0043	75.9	0.0015	> 2400
k.iso	0.001	0.0032	0.0011	35.3	0.0014	0.0076
k.rev	0.0001	0.00011	3.70E-05	33.7	< 2.83e-010	0.000195
r(P)	1	0.9593	0.0033	0.3		
offset	1.89	-0.293	0.045	15.4		

^(a) Confidence intervals at **95 %** probability level.

Table 4

Results of confidence interval estimation using the DynaFit script "**confid-2S.txt**". For details see text.

In *Table 4*, the column labeled "Final" lists the best-fit parameter values. The column labeled "Std. Error" lists the associated formal standard errors. Thus, at that level of detail we could summarize the results of fit as follows:²

- $k_{a.I} = (1.15 \pm 0.37) \times 10^6 \text{ M}^{-1}\text{s}^{-1}$
- $k_{d.I} = (0.0057 \pm 0.0043) \text{ s}^{-1}$
- $k_{iso} = (0.0032 \pm 0.0011) \text{ s}^{-1}$
- $k_{rev} = (0.00011 \pm 0.00004) \text{ s}^{-1}$

This corresponds to the equilibrium dissociation constant of the initial complex

² In converting the best-fit value of the association rate constant $k_{a.S}$ to standard **molar** units, recall that all concentrations throughout the DynaFit script are **nanomolar**. Thus, for example, $0.001 \text{ nM}^{-1}\text{s}^{-1} = 10^6 \text{ M}^{-1}\text{s}^{-1}$.

$$K_d^{(ini)} = \frac{k_{d.I}}{k_{a.I}} = \frac{0.0057}{0.00115} = 5.0 \text{ nM} \quad .$$

The overall dissociation equilibrium complex of the final tightly bound complex is

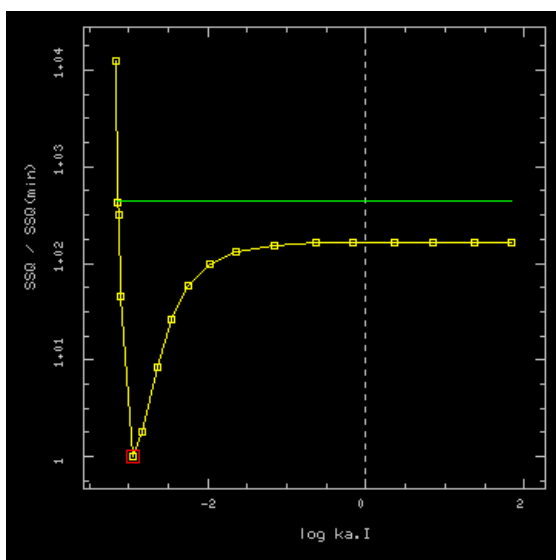
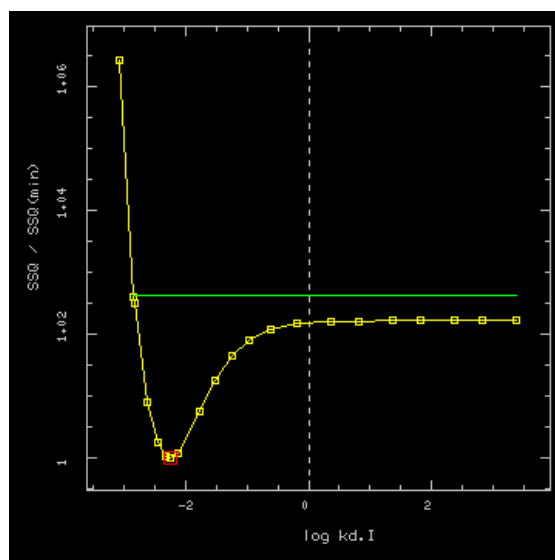
$$K_d^{(tot)} = K_d^{(ini)} \frac{k_{rev}}{k_{iso}} = 5.0 \frac{0.00011}{0.0032} = 0.17 \text{ nM} \quad .$$

It is conceptually difficult to assign a single numerical value that would appropriately characterize the "slowness" of converting the initial (loosely bound) complex to the final (tightly bound) complex, because the dynamics of this "slow, tight" binding event is fully captured by *four* rate constants listed above. However, if we were to focus on the unimolecular isomerization alone, the half-life of isomerization might be calculated as

$$t_{1/2} = \frac{\ln(2)}{k_{iso} + k_{rev}} = \frac{0.693}{0.0032 + 0.00011} = 209 \text{ sec} = \mathbf{3.5 \text{ min}}$$

However, it is very important to note that the three of the four rate constants appearing in Table 4 are characterized by *undefined* values of either the upper limit ($k_{a,S}$, $k_{d,S}$) or the lower limit (k_{rev}) of the 95% confidence level interval. The nonsymmetrical confidence intervals were computed by the "profile-*t*" method of Bates & Watts [10-12].

The fact that the upper limits of the confidence intervals, at 95% confidence level, are undefined for $k_{a,S}$ and $k_{d,S}$ is graphically illustrated in the *likelihood profile* [13] graphs also shown on the "Parameters" output page, see *Figure 3*. In *Figure 3*, the green horizontal line represents the *critical* value of the relative sum of squares corresponding to the given confidence level. If the likelihood profile does not rise to this level at either the lower or the upper end of the confidence interval, it means that the given parameter is statistically indistinguishable from zero or from infinity, respectively.

Confidence interval for **ka.I**Confidence interval for **kd.I****Figure 3**

Half-open confidence intervals for the association and dissociation rate constants, $k_{a,s}$ and $k_{d,s}$, respectively. For details see text.

In this case we can see in *Figure 3* that only the lower limits are defined for the association and dissociation rate constants, whereas only the upper limit is defined for the reverse isomerization rate constants k_{rev} . Thus, at this higher level of detail we could summarize the results of fit as follows:

- $k_{a,I} = (1.15 \pm 0.37) \times 10^6$ $[0.74, \infty] \text{ M}^{-1}\text{s}^{-1}$
- $k_{d,I} = (0.0057 \pm 0.0043)$ $[0.0015, \infty] \text{ s}^{-1}$
- $k_{iso} = (0.0032 \pm 0.0011)$ $[0.0014, 0.0076] \text{ s}^{-1}$
- $k_{rev} = (0.00011 \pm 0.00004)$ $[0, 0.00019] \text{ s}^{-1}$

The values in square brackets are the lower and upper limits of the 95% confidence level intervals. When these numerical limits are translated into plain language, we could say that the inhibitor must be associating very rapidly, certainly with an association rate constant higher than $7.4 \times 10^5 \text{ M}^{-1}\text{s}^{-1}$ (at the 95% confidence level) but we simply do not know just how much higher the association rate constant really is, compared to this limiting value. For all practical purposes the association rate constant is statistically indistinguishable from infinity.

Similarly, in the case of the reverse isomerization rate constant k_{rev} , all we can say is that this rate constant cannot possibly be higher than 0.00019 s^{-1} (at the 95% confidence level) but we simply do not know just how much lower the rate constant could be,

compared to its limiting value. For all practical purposes the reverse isomerization rate constant is statistically indistinguishable from zero, signifying essentially "covalent" inactivation.³

Confidence intervals for the rapid-equilibrium model

Modify the confidence interval estimation script

PROCEDURE

1. Open the script file **confid-2S.txt** (see above) if it isn't already open.
2. Select menu "**File ... Save As...**"
3. Type **confid-2S-R** into the **File name** field. The file name extension **.txt** will be added automatically.
4. Edit the value of the association rate constant and treat it as a fixed parameter:

CURRENT:

```
ka.I = 0.001 ??
```

MODIFIED:

```
ka.I = 0.00115
```

This is the *best-fit* value $k_{a,I}$ resulting from the model discrimination script **models.txt**. Note that the dimension are nanomolar so the notation represents

$$k_{a,I} = 0.00115 \text{ nM}^{-1}\text{s}^{-1} = 1.15 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$$

Given the molecular mass of the inhibitor and enzyme molecules involved in this interaction, an association rate constant equal to approximately $10^6 \text{ M}^{-1}\text{s}^{-1}$ is very likely to be approaching the *diffusion limit* ("rapid equilibrium" approximation).

5. Modify the output directory name:

CURRENT:

```
[output]
directory ./technotes/2015/06/output/confid-2S
```

MODIFIED:

³ However, note that irreversible inhibition was excluded by a high-dilution experiment (Moss *at al.*, 1996).

[output]
 directory ./technotes/2015/06/output/**confid-2S-R**

6. Delete all **[task]** sections following the first one.
7. Delete the (now) superfluous **model** line in the **[task]** section.
8. Press "**Ctrl+S**" to save the modified script file.

Perform confidence interval estimation

PROCEDURE

4. Select menu "**File ... Run**" or equivalently press **Ctrl+U**
5. Select menu "**View ... Results in External Browser**" or press **Ctrl+J**
6. Click on the "Parameters" link

Examine the confidence intervals:

The summary of results displayed in the "Parameter" page of the DynaFit output is shown in *Table 5*.

Param	Initial	Final	Std. Error	CV (%)	Low ^(a)	High ^(a)
kd.I	0.01	0.00567	0.00075	13.3	0.00445	0.00766
k.iso	0.001	0.0032	0.00057	17.8	0.00228	0.00471
k.rev	0.0001	0.00011	0.000037	33.6	0.0000352	0.000184
r(P)	1	0.9593	0.0032	0.3		
offset	1.89	-0.293	0.041	14		

^(a) Confidence intervals at **95 %** probability level.

Table 5

Results of confidence interval estimation using the DynaFit script "**confid-2S-R1.txt**". Note that all three confidence intervals are closed from both above and below. For details see text.

Thus, the results of fit could be summarized as follows:

- $k_{a.I} = 1.15 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ (fixec)
- $k_{d.I} = (0.0057 \pm 0.0008) \quad [0.0045, 0.0077] \text{ s}^{-1}$
- $k_{iso} = (0.0032 \pm 0.0006) \quad [0.0023, 0.0047] \text{ s}^{-1}$
- $k_{rev} = (0.00011 \pm 0.00004) \quad [0.00004, 0.00018] \text{ s}^{-1}$

This corresponds to the equilibrium dissociation constant of the initial complex

$$K_d^{(ini)} = \frac{k_{d.I}}{k_{a.I}} = \frac{0.0057}{0.00115} = 5.0 \text{ nM} \quad .$$

The overall dissociation equilibrium complex of the final tightly bound complex is

$$K_d^{(tot)} = K_d^{(ini)} \frac{k_{rev}}{k_{iso}} = 5.0 \frac{0.00011}{0.0032} = 0.17 \text{ nM} \quad .$$

The approximate half-time of isomerization is

$$t_{1/2} = \frac{\ln(2)}{k_{iso} + k_{rev}} = \frac{0.693}{0.0032 + 0.00011} = 209 \text{ sec} = \mathbf{3.5 \text{ min}}$$

Discussion

We have previously published [4] a kinetic analysis of 5α -steroid reductase inhibition by a heterocyclic testosterone analog shown in *Figure 2*. At that time the analysis was performed using an early version of the software package DynaFit [1]. Here we present a revised kinetic analysis using a more advanced version of DynaFit [2]. The features and capabilities that were added to the software since the original publications are as follows:

- **Automatic scan of initial estimates for rate constants**

In the initial publication, we utilized a tedious and subjective trial-and-error procedure to generate initial estimates of rate constants, required by nonlinear least-squares regression. Here we used an automated procedure examining several hundred initial estimates in a systematic manner. The major advantage of the automated procedure is that the data fitting algorithm can avoid falling into a "false minimum" and ultimately produce incorrect results.

- **Model discrimination using information-theoretic criteria (AIC, BIC)**

In the initial publication, we relied solely on Fisher's F-test [14] in trying to statistically discriminate between rival kinetic models. Here we additionally deployed the second-order (i.e. sample-size corrected) Akaike Information Criterion [7] and the corresponding Bayesian Information Criterion [8, 9]. The major advantage is that the AIC and BIC criteria apply both to nested and to non-nested fitting models, whereas the *F*-test model criterion applies only to nested models.

- **Wider array of candidate models**

In this report we expanded the list of candidate kinetic models from six to seven, not including several variants of the single-step binding mechanism. Most importantly, we also examined a specialized version of the two-step reversible mechanism that involved essentially instantaneous initial equilibration. This model has not been considered in the initial report [4], and yet it turns out to be the optimal kinetic mechanism in this particular case.

- **Improved algorithm form confidence interval estimation**

In the original report, we utilized a minor variation of the search algorithm originally published by Bates & Watts (ref. [10], pp. 302-303, section A.3.5.1). Unfortunately our experience shows that this rudimentary algorithm (less than 20 lines in S-language pseudocode [10]) frequently fails on highly redundant models, with "half-opened" confidence intervals, such as those shown in *Figure 3*. The present version of DynaFit utilizes a more robust search algorithm. The major advantage of the modified algorithm is that poorly defined fitting models (e.g., model "2S" above) can be studied more reliably that was possible before.

- **Additional experimental data**

In this analysis we utilized experimental data (namely, progress curves obtained at relatively low inhibitor concentrations) that were available but were purposely ignored at the time of the original publication [4].⁴ Including the additional data increased the model discrimination power.

A comparison of the previously published results [4] and the results obtained here is shown in **Table 6**. The results are essentially identical but could now be more easily obtained even by a non-expert user because of (a) fully automatic initial estimates and (b) a more robust model-discrimination analysis.

Parameter	Published		This work		Note
	Fitted	Std. Error	Fitted	Std. Error	
$k_{a,i}$, $\text{nM}^{-1}\text{s}^{-1}$	0.005	--	0.00115	--	Fixed parameter
$k_{d,i}$, s^{-1}	0.033	0.008	0.00567	0.00075	
k_{iso} , s^{-1}	0.0035	0.0007	0.0032	0.00057	
k_{rev} , s^{-1}	0.000076	0.000019	0.00011	0.000037	
$K_d^{(\text{ini})}$, nM	6.6		4.9		$= k_{d,i} / k_{a,i}$
$K_{\text{eq}}^{(\text{iso})}$	46.1		29.1		$= k_{\text{iso}} / k_{\text{rev}}$
$K_d^{(\text{tot})}$, nM	0.14		0.17		$= K_d^{(\text{ini})} / K_{\text{eq}}^{(\text{iso})}$
$t_{1/2}$, min	3.2		3.5		$= \ln(2) / [k_{\text{iso}} + k_{\text{rev}}]$

⁴ In the original paper the low concentration data were excluded because we focused on comparing the performance of the classic, algebraic approach to the kinetic analysis of "slow, tight" inhibition with the differential-equation formalism. The algebraic method could not possibly utilize the low inhibitor concentration curves because of unacceptably high substrate conversion.

Table 6

Comparison of results: This work vs. previously published report [4] utilizing the same experimental data.

Perhaps a more informative comparison should be made between the results obtained in this report and the corresponding results based on the traditional, essentially algebraic method of kinetic analysis [3] as described in the original report [4]. The comparison is shown in *Table 7*.

Parameter	Algebraic		This work	
	Fitted	Std. Error	Fitted	Std. Error
$k_{a,i}$, $\text{nM}^{-1}\text{s}^{-1}$	--	--	0.00115	--
$k_{d,i}$, s^{-1}	--	--	0.00567	0.00075
k_{iso} , s^{-1}	0.0035	0.0001	0.0032	0.00057
k_{rev} , s^{-1}	0.00025	0.00006	0.00011	0.000037
$K_d^{(\text{ini})}$, nM	7.7	0.5	4.9	0.7
$K_{\text{eq}}^{(\text{iso})}$	14.0		29.1	
$K_d^{(\text{tot})}$, nM	0.55	0.02	0.17	0.04
$t_{1/2}$, min	3.1		3.5	

Table 7

Comparison of results: This work vs. previously published report [4] utilizing the same experimental data and the "classic" algebraic method of kinetic analysis [3].

The results in *Table 7* illustrate that, in this particular case, the "classic" method of kinetic analysis underestimates the binding affinity in the initial enzyme-inhibitor complex approximately two-fold. Similarly, the overall binding affinity is underestimated by approximately 300%.

It is important to note that, in the original report [4], the "classic" method of kinetic analysis could be utilized at all (as one of the two methods being compared) because "tight binding" [3] is essentially negligible in this case. Indeed the enzyme active site concentration (50 pM) is at least three-fold lower than the overall dissociation constant (170 pM) of the final enzyme-inhibitor complex. Under such experimental conditions "tight binding" phenomena can be neglected. It has been shown experimentally [15] that "tight binding" manifests prominently only when the enzyme concentration is actually *higher* than the overall dissociation constant of the enzyme-inhibitor complex.

However, the main advantage of the numerical (differential equation) formalism utilized in this report is that it is *equally* applicable to *all* experimental situations *irrespective* of the active enzyme concentration relative to the overall dissociation constant of the enzyme-inhibitor complex.

References

1. Kuzmic, P., Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV proteinase, *Anal. Biochem.* 237 (1996) 260-273.
2. Kuzmic, P., DynaFit - A software package for enzymology, *Meth. Enzymol.* 467 (2009) 247-280.
3. Szedlacsek, S., and Duggleby, R. G., Kinetics of slow and tight-binding inhibitors, *Meth. Enzymol.* 249 (1995) 144-180.
4. Moss, M. L., Kuzmic, P., Stuart, J. D., Tian, G., Peranteau, A. G., Frye, S. V., Kadwell, S. H., Kost, T. A., Overton, L. K., and Patel, I. R., Inhibition of human steroid 5-alpha reductases type I and II by 6-aza-steroids: Structural determinants of one-step vs. two-step mechanism, *Biochemistry* 35 (1996) 3457-3464.
5. Tian, G., Stuart, D., Moss, M., Domanico, P., Bramson, N., Patel, I. R., Kadwell, S., Overton, L. K., Kost, T. A., Mook, R., Frye, S. V., Batchelor, K. W., and Wiseman, J., 17-beta-(N-tert-butylcarbonyl)-4-aza-5-alpha-androstan-1-en-3-one is an active site-directed slow time-dependent inhibitor of human steroid 5-alpha-reductase 1, *Biochemistry* 33 (1993) 2291-2296.
6. Kuzmic, P. (2015) "Determination of substrate kinetic parameters from progress curve data: A DynaFit tutorial", BioKin Ltd., www.biokin.com/TN/2015/05.
7. Burnham, K. B., and Anderson, D. R., *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd Ed., Springer-Verlag, New York, 2002.
8. Myung, J. I., and Pitt, M. A., Model comparison methods, *Meth. Enzymol.* 383 (2004) 351-366.
9. Myung, J. I., Tang, Y., and Pitt, M. A., Evaluation and comparison of computational models, *Meth. Enzymol.* 454 (2009) 287-304.
10. Bates, D. M., and Watts, D. G., *Nonlinear Regression Analysis and its Applications*, Wiley, New York, 1988.
11. Brooks, I., Watts, D. G., Soneson, K. K., and Hensley, P., Determining confidence intervals for parameters derived from analysis of equilibrium analytical ultracentrifugation data, *Meth. Enzymol.* 240 (1994) 459-78.
12. Watts, D. G., Parameter estimation from nonlinear models, *Methods Enzymol.* 240 (1994) 24-36.
13. Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J., Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood, *Bioinformatics* 25 (2009) 1923-1929.
14. Mannervik, B., Regression analysis, experimental error, and statistical criteria in the design and analysis of experiments for discrimination between rival kinetic models, *Meth. Enzymol.* 87 (1982) 370-390.
15. Kuzmic, P., Elrod, K. C., Cregar, L. M., Sideris, S., Rai, R., and Janc, J. W., High-throughput screening of enzyme inhibitors: Simultaneous determination of tight-binding inhibition constants and enzyme concentration, *Anal. Biochem.* 286 (2000) 45-50.