

**Author:** Dr. Pall Thordarson, The University of New South Wales. Email:  
p.thordarson@unsw.edu.au

Copyright © 2010.

**Updated (and original) versions of this program will be available** (AFTER PAPER HAS BEEN REVIEWED) on this website:

[www.chem.unsw.edu.au/research/groups/thordarson/fittingprogram](http://www.chem.unsw.edu.au/research/groups/thordarson/fittingprogram)

**First published version:** 18<sup>th</sup> October 2010

**Current version:** 18<sup>th</sup> October 2010

**Licence to use:** The program and the associated data file are free of use but users are kindly asked to cite this publication in their work when using this program or modified versions of this program.

Suggestions for improvements are welcome and can be emailed to the authors.

## **Instructions:**

**Overview:** The zip-file consists of a library of Matlab files to fit 1:1, 1:2 and 2:1 equilibria with up to 4 data sets simultaneously. Sample / Template Excel data files are also included as well as executable Matlab files to import data from these files and, after fitting, export the key results back to these Excel files. A few simulation program are also provided.

**Assumed knowledge/resources:** The program requires Matlab which must be installed on the computer used. The bulk of the code was written for version 7 (R2006) and should run smoothly on any subsequent versions of the program. The Excel files were generated in Excel 2010 and should be readable in Excel 2007. Only the most basic familiarity with Matlab is required – for users completely unfamiliar with this program, the first few pages from internet resources such as:

[http://maxwell.me.gu.edu.au/spl/matlab-page/matlab\\_tony.pdf](http://maxwell.me.gu.edu.au/spl/matlab-page/matlab_tony.pdf)

<http://www.maths.dundee.ac.uk/~ftp/na-reports/MatlabNotes.pdf>

should be more enough to get users going with using this program.

Any commands/program that you need to run from the command line in Matlab are shown in *italics*

**Installation:** Unzip the enclosed zip-file. This will reveal a directory with four subdirectories that include all the files needed. Make sure to include all these four directories in path directory in Matlab (go to “File”, select “Set Path” and then select “Add with Subfolders” to select the folder “ThordarsonFittingProgram” and include all subfolders, then press “save” and then “close”). You can now run the different program from any directory you wish in Matlab. To use the sample Excel / data files, go to the folder “Creating\_input\_including\_examples”

### **Naming system used for fitting programs:**

*fittingXtoYnmrZ*      Fitting program to fit X to Y equilibria to NMR data for Z data sets:

If X and Y = 1: *fitting1to1nmrZ* – for fitting 1:1 equilibria to NMR data

e.g., *fitting1to1nmr4* for fitting 4 data sets globally to 1:1 equilibria.

If X = 1 and Y = 2: *fitting1to2nmrZ* – for fitting 1:2 equilibria to NMR data

e.g., *fitting1to2nmr2* for fitting 2 data sets globally to 1:1 equilibria

If X = 2 and Y = 1: *fitting2to1nmrZ* – for fitting 2:1 equilibria to NMR data

e.g., *fitting2to1nmr1* for fitting 1 data sets **locally** to 1:1 equilibria

*fittingXtoYuvZ*      Fitting program to fit X to Y equilibria to UV data for Z data sets:

If X and Y = 1: *fitting1to1uvZ* – for fitting 1:1 equilibria to UV data

e.g., *fitting1to1uv3* for fitting 3 data sets globally to 1:1 equilibria.

If X = 1 and Y = 2: *fitting1to2uvZ* – for fitting 1:2 equilibria to UV data

e.g., *fitting1to2uv1* for fitting 1 data sets **locally** to 1:1 equilibria

If X = 2 and Y = 1: *fitting2to1uvZ* – for fitting 2:1 equilibria to UV data

e.g., *fitting2to1uv4* for fitting 4 data sets globally to 1:1 equilibria

### **Naming system used for sample data files:**

N.b. You can reuse and rename the excel file as long as you leave the header row unchanged (you can change the headings for columns C onwards but DON'T use numbers alone to name these!). Copy the appropriate data in the same order in the columns below (host concentration, then ligand concentration then 1-4 data sets).

NMRdataforXtoYwithZ      Sample NMR data for X to Y equilibria for Z data sets, e.g.

*NMRdatafor1to1with1*      Sample NMR data for 1:1 equilibria with 1 data set.

UVdataforXtoYwithZ      Sample UV data for X to Y equilibria for Z data sets, e.g.

*UVdatafor2to1with4*      Sample UV data for 2:1 equilibria with 4 data set.

### **Step 1 – prepare your input data file:**

Start with the Excel sample files in the folder “Creating\_input\_including\_examples”

Run the program “*inputfilemaker*” from the Matlab command line and follow the instructions.

If you are Excel program that doesn't generate “xlsx” extension (e.g., Excel 2003 = “xls”), you need to modify the code slightly by opening the file *importfilemaker* and replace “xlsx” with “xls” (or another suffix).

## **Step 2 – Run the appropriate fitting program:**

Choose one of the fitting programs (*fittingXtoYuvZ* or *fittingXtoYnmrZ*)

You will then be prompted to answer (prompts differ slightly depending on fitting program):

Enter name of input file = the input file generated in step 1 above

Enter factor to change htot (usually 1) = here  $htot=[H]_0$  – this allows you to change the concentration of  $[H]_0$ , e.g., if you suspect that it had been overestimated. Normally you wouldn't do this and simply enter 1 here.

Number of iteration to perform (ca. 200-1000) = The higher the number, the longer the fitting process takes. N.b., putting a large number (> 1000) is not necessarily better. Better to run say one fitting “cycle” with 300 iterations, inspect the results and re-enter those results as the initial guess for the next cycle of iterations.

Enter guess for binding constant(s) = when you have no idea where to start pick, e.g., something around  $1/[H]_0$  for 1:1. For 1:2, a good start might be also be  $1/[H]_0$  for  $K_1$  and  $1/10$  of that number for  $K_2$ . These number should always be positive.

Enter guess for DeltaXX = this is the  $\delta_{\Delta\delta H_i}$  in case of NMR or  $\epsilon_{\Delta H G_i}$  for UV-Vis. Try picking some sensible numbers. N.b. these numbers can both be negative or positive.

Notes:

The “Z” (number of data sets) has to match the number of data sets in your input file, otherwise it will crash

Matlab will not stop you from fitting NMR data with a UV program but the answer will be nonsense.

You can fit a data file to any combination of X and Y you wish (1:1, 1:2, 2:1). However, if the initial guesses are bad and/or the data fits that equilibria poorly, the fitting program will probably crash or give nonsense answer (e.g., negative value for  $K_a$ ).

Pick your initial guess carefully. If the program runs OK, run it again using numbers similar to the results as a starting point. Getting the same results again is a good indicator that the results are “stable”.

Experienced Matlab users can of course modify the plots generated. A useful command to convert the figure to a tiff file is:

```
print filename -dtiff -r300
```

(where filename is your name for that figure, and r300 specifies RGB tiff with 300 dpi resolution).

All the programs starting with *nmrXtoYpara...* or *uvXtoYpara* (e.g., *nmr1to2para4bnn*) are executable files that the “*fittingXtoY...*” programs use. Do not edit these unless you understand clearly what you are doing. The program “*uv1to2bb*” in the folder “Creating\_input\_including\_examples” solves the key cubic equation for 1:2 and 2:1 equilibria. Do not edit it and make sure that this directory is included in the Matlab paths (see above).

### **Step 3 – Export back to Excel:**

Once you are satisfied with the results from fitting (after couple of iterations) you can export the results back to the same Excel file you started from.

Run the program “*Exportfilemaker*” from the Matlab command line and follow the instructions.

Notes:

The program exports the **LAST** results for the 1:1, 1:2 or 2:1 fitting progress regardless of quality.

The program exports the following (see also headers in Excel outfile):

Fitted data points (n.b. **ONLY** the last one done even if you fitted more than one type of equilibria (e.g., 1:2 and 1:1).

Results for  $K_a$ ,  $K_1$  and  $K_2$  for 1:1, 1:2 and 2:1

All other parameters fitted (e.g.,  $\delta_{\text{AHG}}$ ...)

Confidence intervals for these parameters (in % from Asymptotic error).

The program will not overwrite your initial raw data, however, if you decided to re-export your data to the same file after performing additional experiments you should use a different name for the file when exporting, otherwise your old exported data will be overwritten (unless that is what you want to do!).

This program can be quite slow so be patient!

If you have performed a fitting procedure that give non-sense results in one of the iteration cycle you may have difficulties here with error messages saying something along the line “NaN is not a logical”. You may have to start the iteration process

again under a new name, or if you are an experienced matlab user, open the result file, delete the offending rows in the results vector (e.g., *oneone*) and try exporting again.

If you are Excel program that doesn't generate "xlsx" extension (e.g., Excel 2003 = "xls"), you need to modify the code slightly by opening the file *Exportfilemaker* and replace "xlsx" with "xls" (or another suffix).

#### Step 4 – Simulation programs:

A few simulation programs are also provided. They all have names *simXtoY...* and come in two types:

- 1) Basic simulation, i.e., where you specify all the parameters (e.g.,  $K_a$ ,  $\delta_{\Delta HG}$ ...), the concentrations and uncertainties and the program will simulate data based on these. The names of these programs are *simXtoYnmr1a* and *simXtoYuv1a* for to simulating X to Y equilibria for NMR and UV experiments, respectively (e.g., *sim1to2nmr1a* to simulate an NMR experiment with 1:2 equilibria). By choosing to **cycle** (repeat) the process x-times (one of the initial parameters) you can get Monte Carlo type information.
- 2) Monte Carlo type simulation using existing data as an input. Please note you still have to enter the parameters (e.g.,  $K_a$ ,  $\delta_{\Delta HG}$ ...) by hand so it advised to write these first down from a previous "good" fitting progress (in the future I will try to automate this process). Again, you can add uncertainties as needed and the program will then simulate series data based on these. The names of these programs are *simXtoYnmrwithinput1* and *simXtoYuvwithinput1a* for to simulating X to Y equilibria for NMR and UV experiments, respectively (e.g., *sim1to1uvwithinput1* to do a Monte Carlo type simulation on an UV experiment with 1:1 equilibria).
- 3) At the moment the simulation program can ONLY handle inputs with one data sets (i.e.,  $Z=1$  in the above terminology).
- 4) The simulation program will report 95% confidence interval for the results obtained. Obviously you need perform at least 20-30 cycles to get meaningful confidence intervals but this can be a bit time consuming.
- 5) All the simulation program use a random number generator to manipulate the data – therefore different runs of these with the same input parameters (and input files if used) will generate slightly different results.