Fitting the MSSV Data of the Interaction of GST-VCA and Arp2/3. *Supplemental Protocol to Padrick & Brautigam, Methods, 2011*

Chad A. Brautigam
Shae B. Padrick
The University of Texas
Southwestern Medical Center at Dallas

Background: The fusion protein GST-VCA (a dimer of about 72,000 Da) interacts with the Arp2/3 complex (a 224,000 Da heteroheptamer). Three experiments were performed
in a single run of the ultracentrifuge.

Cell 1: GST-VCA alone
Cell 2: Arp2/3 alone
Cell 3: A mixture of the two.

The run was performed at 20° C and 42,000 rpm.

The goal: Determine the molar ratio of the complex between the two proteins.

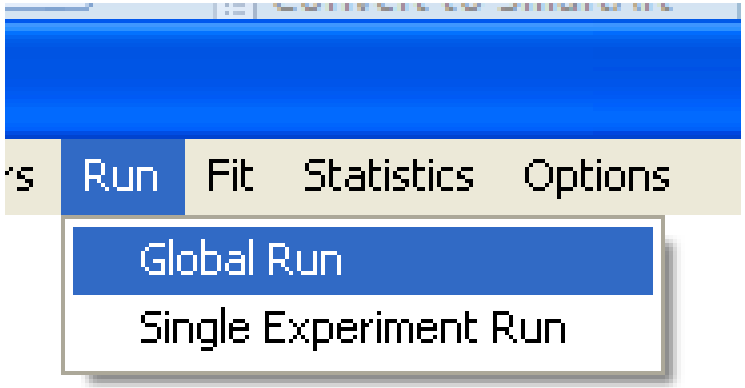
The software: SEDPHAT v. 8.1
(http://www.analyticalultracentrifugation.com/sedphat/default.htm)

Summoning Parameter Windows in SEDPHAT

When Global Parameters need to be summoned, click here.

When Experimental Parameters need to be summoned, click here.

How to “Run” and “Fit”



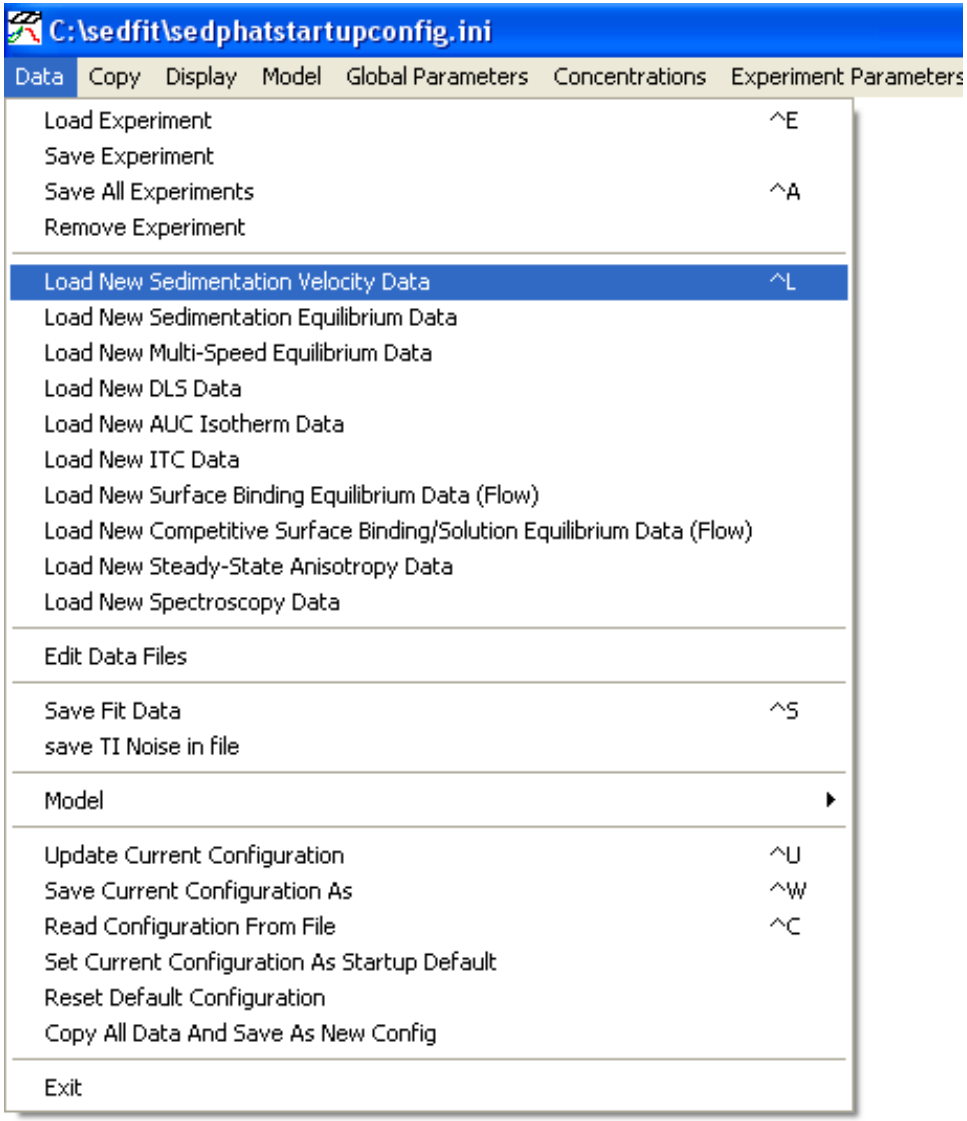
To perform a “Global Run”
(only linear parameters fit)

To perform a “Global Fit”
(linear and non-linear parameters fit)

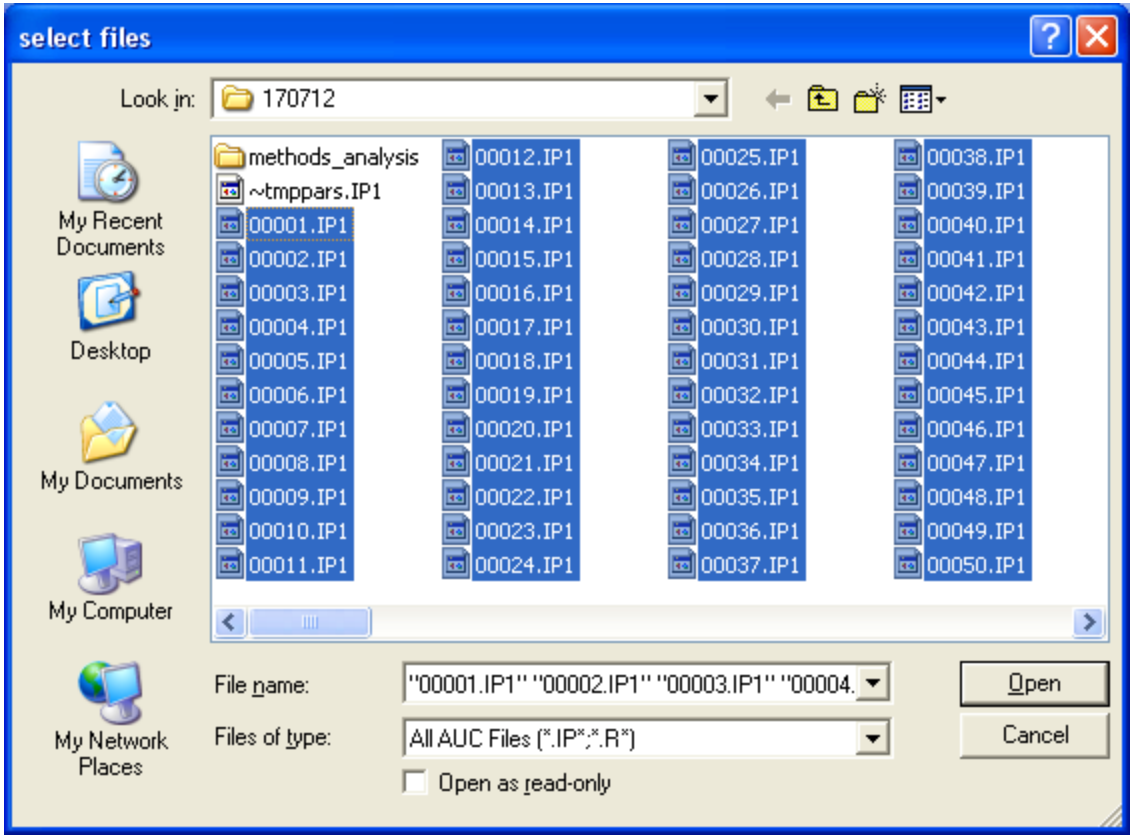
Step 1

The IF data acquired from the GST-VCA-alone sample (Cell 1) were loaded into SEDPHAT. This task is initiated by selecting “Load New Sedimentation Velocity Data” from the “Data” menu of SEDPHAT. All of the “*.IP1” files were displayed, and files 1-101 were selected. After pressing “OK,” the user is prompted by a dialog box to “load every n’th file”. The numeral “2” was entered; thus every 2nd IP1 file between 1 and 101 will be loaded. See the next 2 pages.

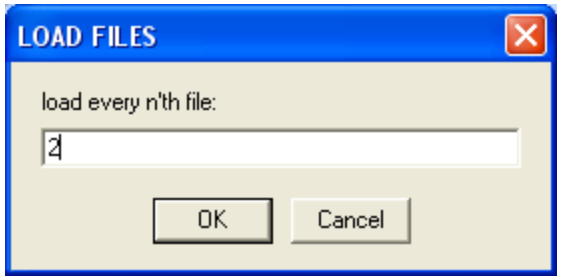
Load New Data



Load the Interference Data First



Interference files 1-101 were selected



Every other file was loaded

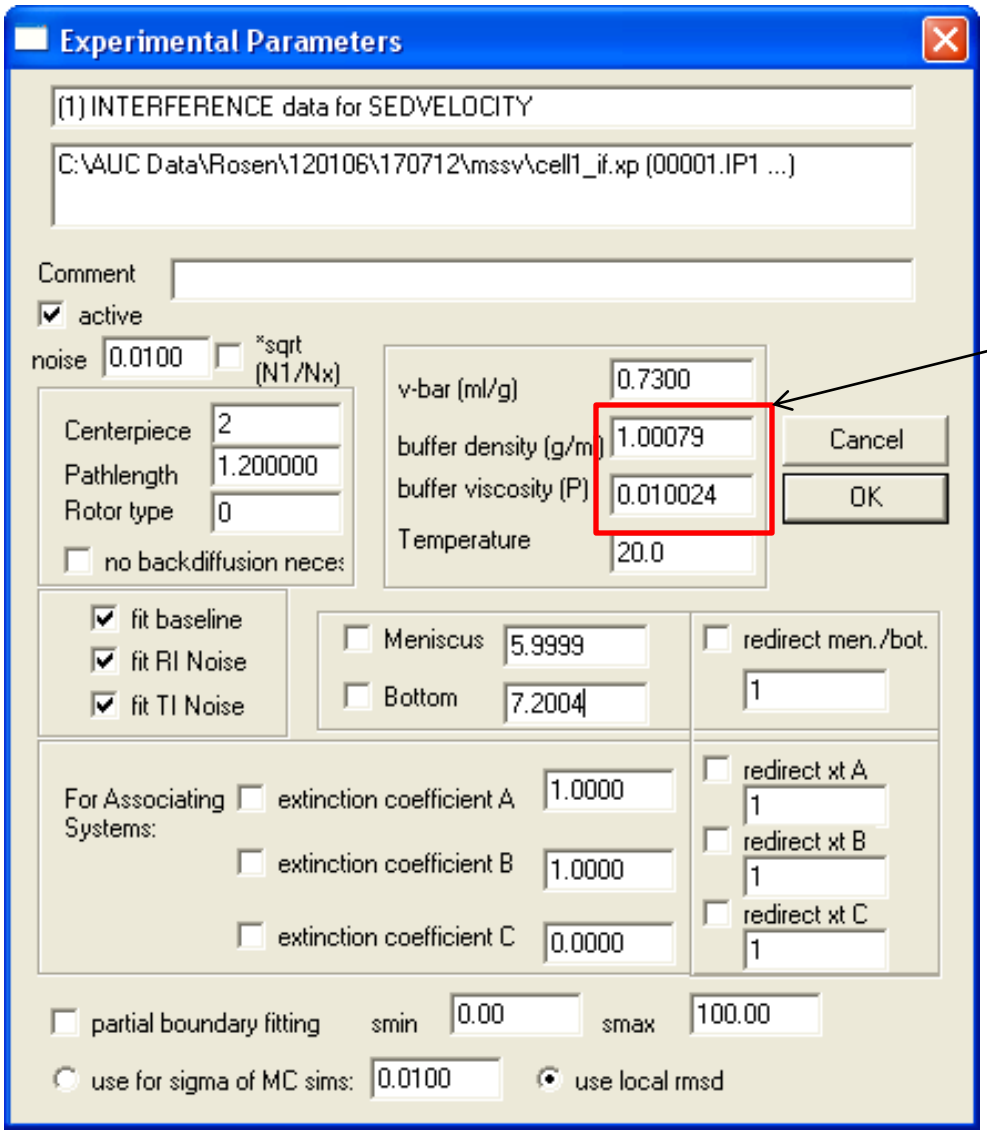
Step 2

A dialog box appears prompting the user for Experimental Parameters (Table 3). SEDPHAT considers each data set loaded as one "experiment." The first data set loaded is deemed Experiment 1, the second Experiment 2, etc. The buffer parameters, i.e. density (ρ) and viscosity (η) were entered. Their values were 1.00079 g/cm^3, and 0.010024 Poise, respectively. Ordinarily, the protein's vbar would be entered here as well. However, the default value of 0.73 cm^3/g was accepted here. The vbar does not strongly bear on the goal of this analysis, which is the molar ratio of the proteins in the GST-VCA:Arp2/3 complex. Therefore, a value that is sensible for most proteins was chosen. By default for IF SV data, the RI noise, TI noise, and baseline checkboxes are activated, which turns on their optimization. This is appropriate. Another parameter to note at this point is the "noise" box. This is the expected root-mean-square error for the experiment; it is by default 0.01 for SV data. This default was accepted in this case. Also, for SV data, dual-sectored centerpieces with a 1.2-cm optical path length are assumed and were accepted in this case.

Step 2 (continued)

Another area of the dialog box accepts the meniscus and bottom parameters. The default values were accepted, and the checkboxes allowing for the refinement of these values are left unchecked for now. The reasons for this choice are both practical and lazy. The meniscus is better chosen graphically, not entered numerically, and there is no opportunity to choose it before this dialog appears. Further, if the meniscus checkbox were activated (as it will be later), the program would ask for fitting limits based on the current, default position of the meniscus. Because the default limits that SEDPHAT sets are the current meniscus ±0.03 cm, the defaults will be unacceptable, and would have to be input manually. See Step 11 for easy ways around these problems. There is also a section for the extinction properties of the proteins. Because these are both vital to the analysis and global to all experiments in MSSV, they are input later in the Global Parameters dialog (see Step 8). See the next page.

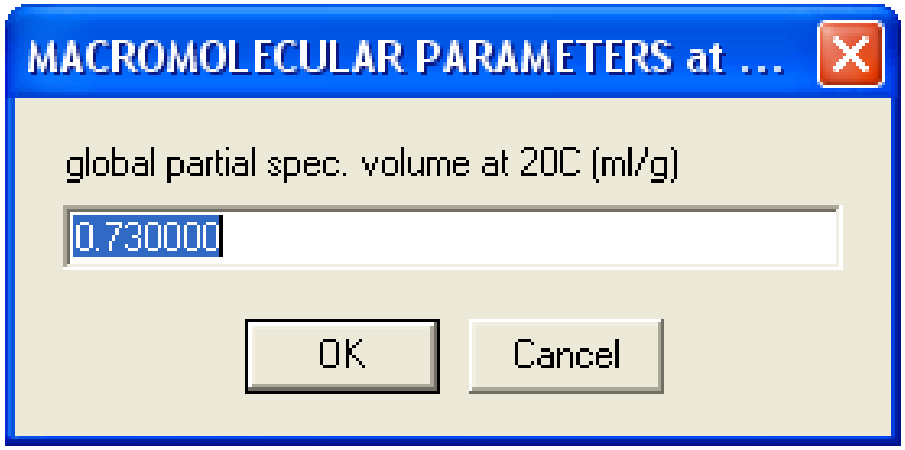
Change solution parameters



Throughout, the boxes show parameters that are altered in the course of the step.

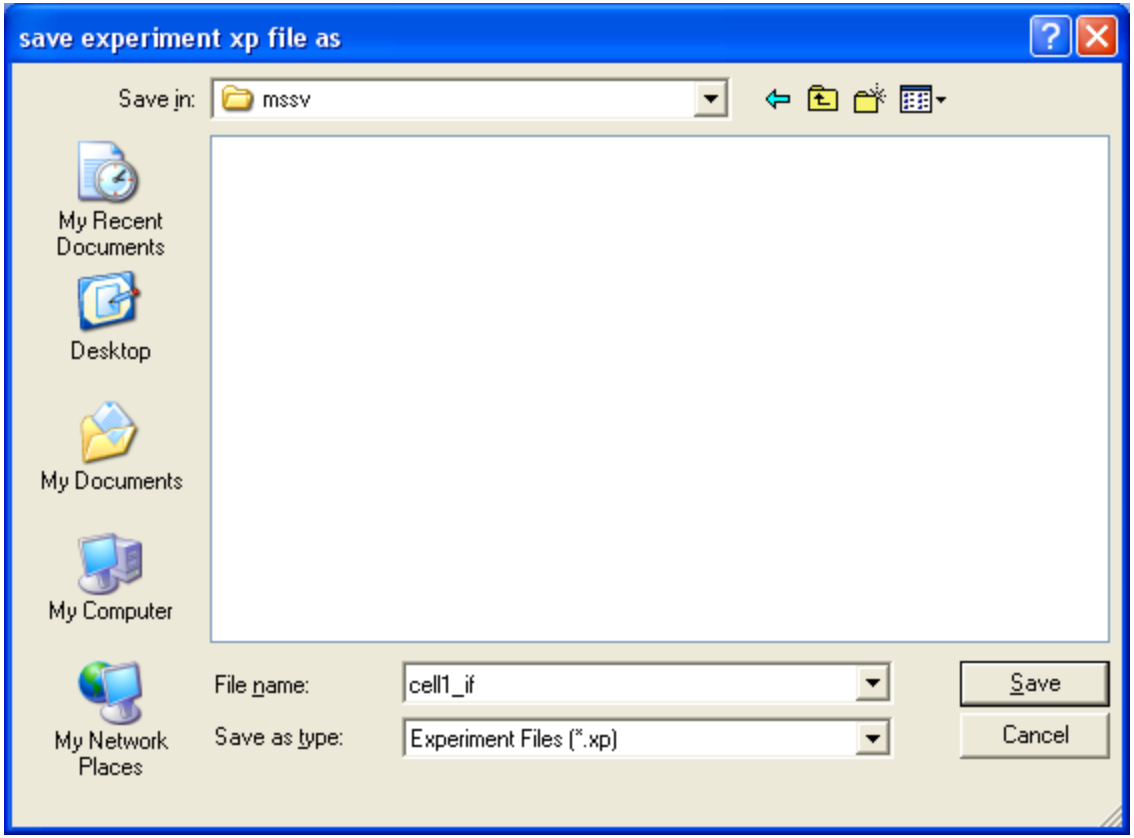
Step 3

The user was prompted to input a global vbar.
The default of 0.73 is accepted.



Step 4

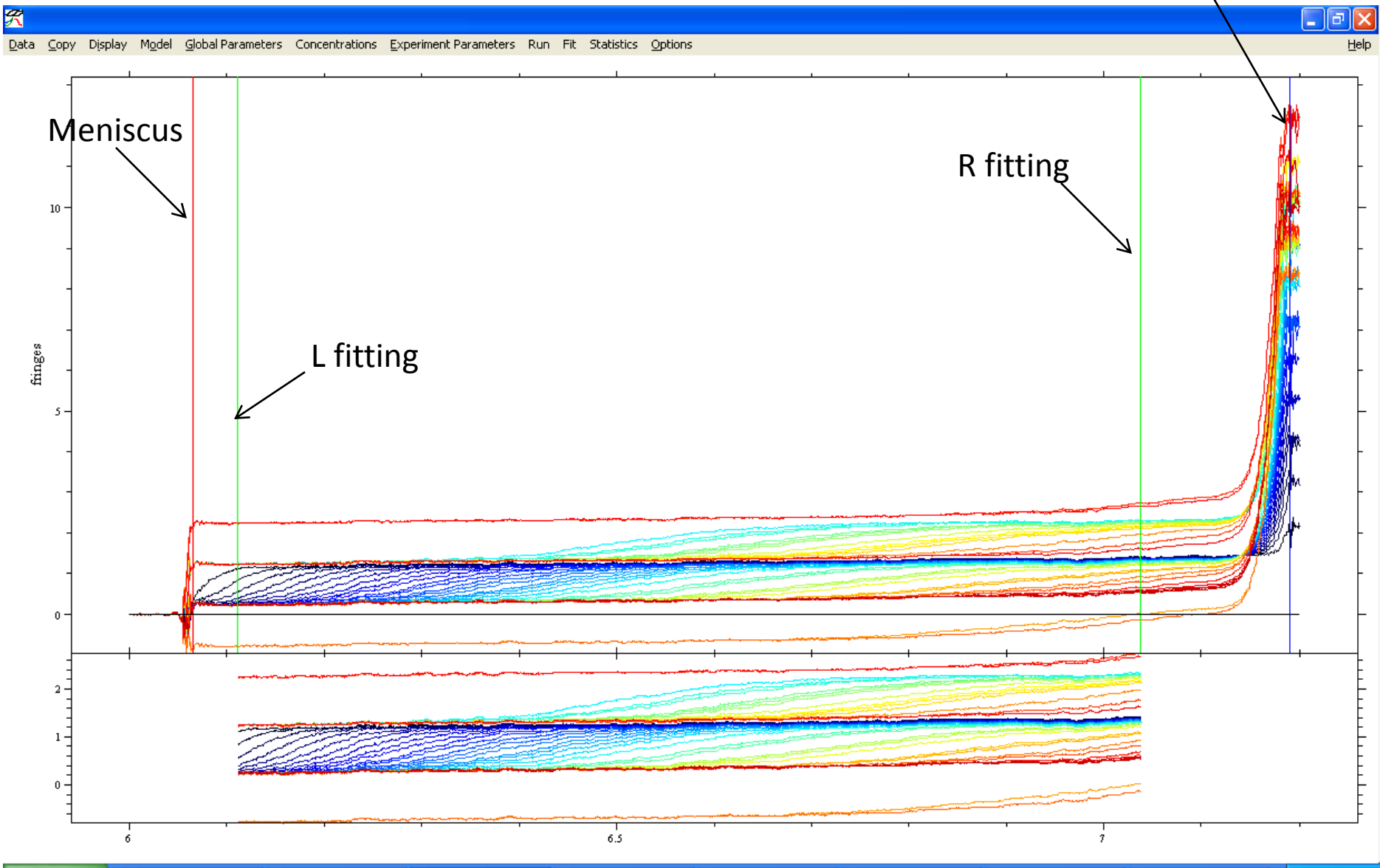
A dialog opened for the purpose of saving the experiment file (file extension .xp). This is a text file that contains information about the location of the data files and all of the parameters input in the Experimental Parameters dialog.



Step 5

The data were displayed by the program, and the meniscus (red line), bottom (blue line), left and right fitting limits (green lines) were set graphically therein. These values can be set by dragging them from their default positions.
See the next page.

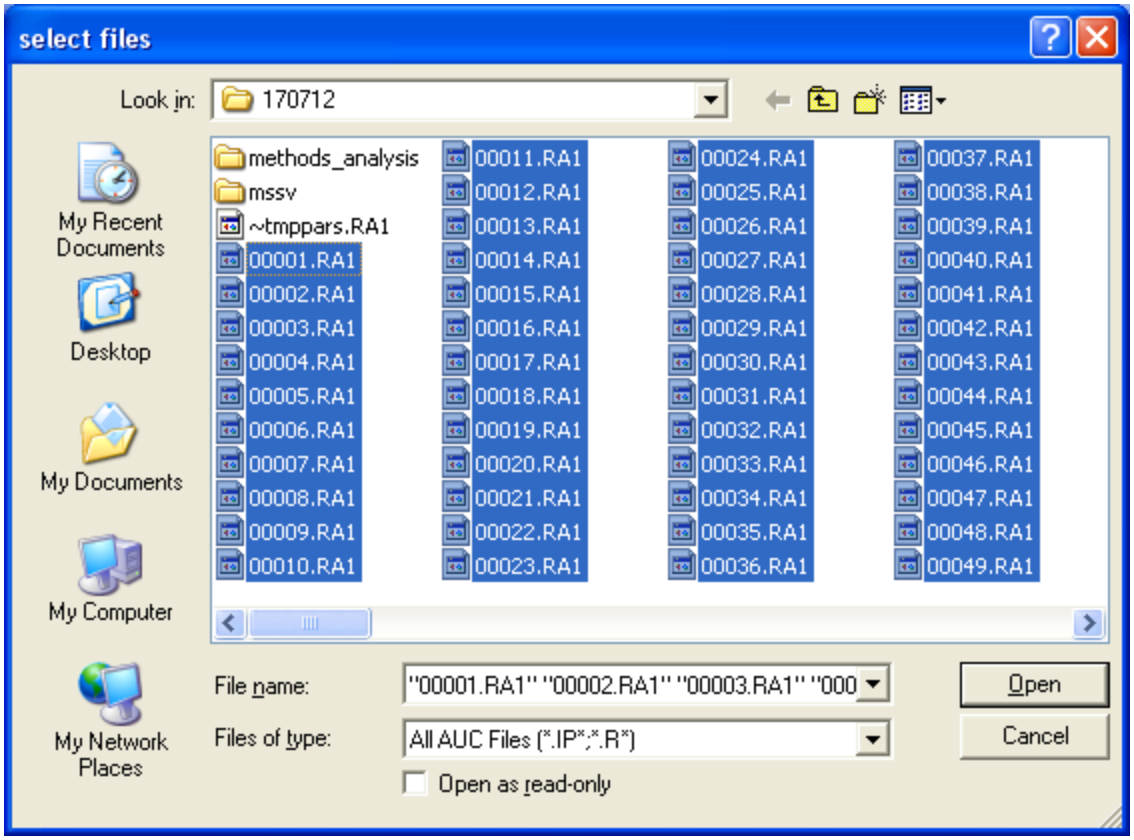
Set the Meniscus, Bottom, and Fitting Limits

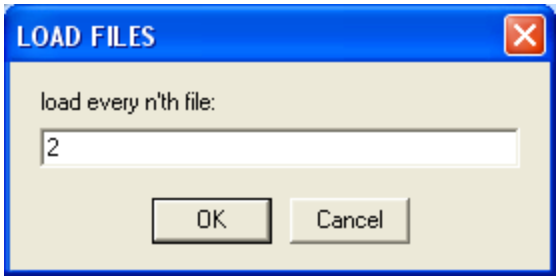


Step 6

Steps 1-5 were repeated for the "*.RA1" data, i.e. the absorbance data acquired from Cell 1 (the same data range was used). There were a few differences (see pg. 17). In the Experimental Parameters box, the RI noise and TI noise boxes are unchecked for absorbance data by default. The TI noise box was therefore activated by the user; time-invariant noise correction is routinely applied to absorbance data to compensate for optical imperfections that manifest in such data. RI noise is inappropriate for absorbance data unless time-dependent baseline changes are observed. Also, the checkbox near to the "noise" input box labeled "*sqrt (N1/Nx)" was activated. The calculation of the global reduced chi-squared (see Text for definition) allows data sets with larger numbers of data points to dominate parameter refinement. Checking this box compensates for this imbalance. The compensation is necessary because the IF experiment had approximately three times as many data points as the absorbance experiment. These substeps are described on the next four pages.

Load the ABS280 Data

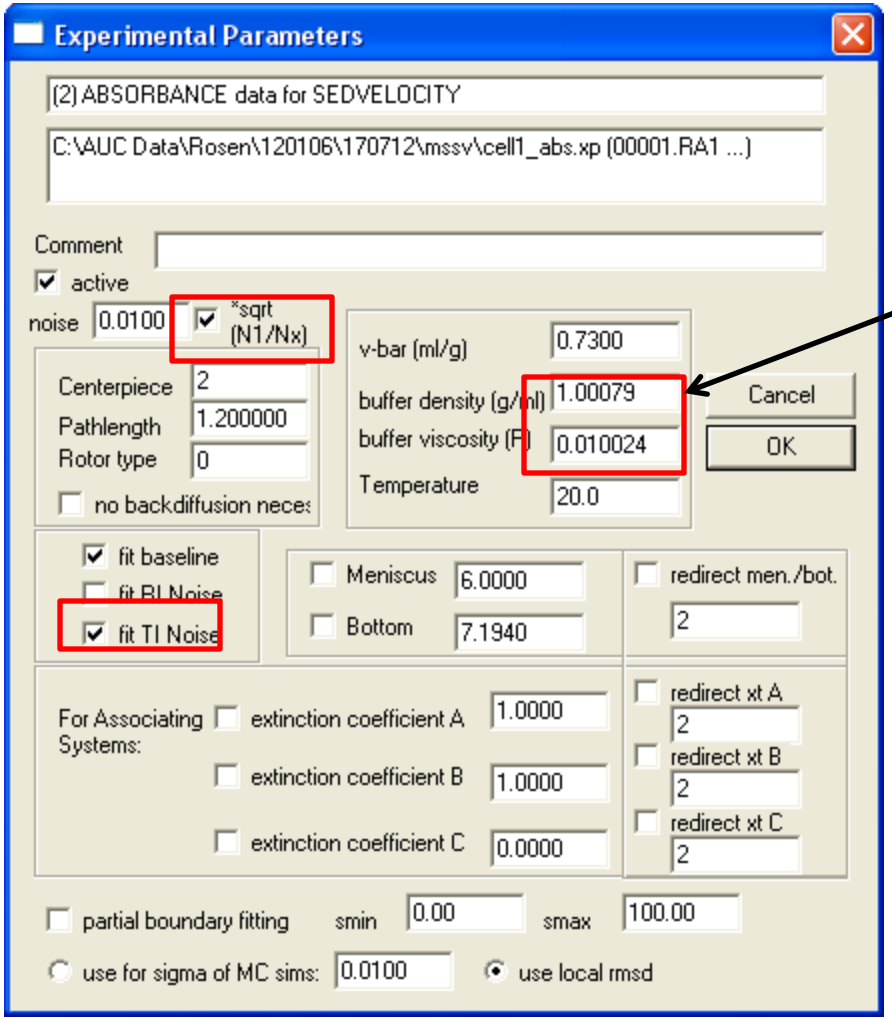




…every other file was loaded.

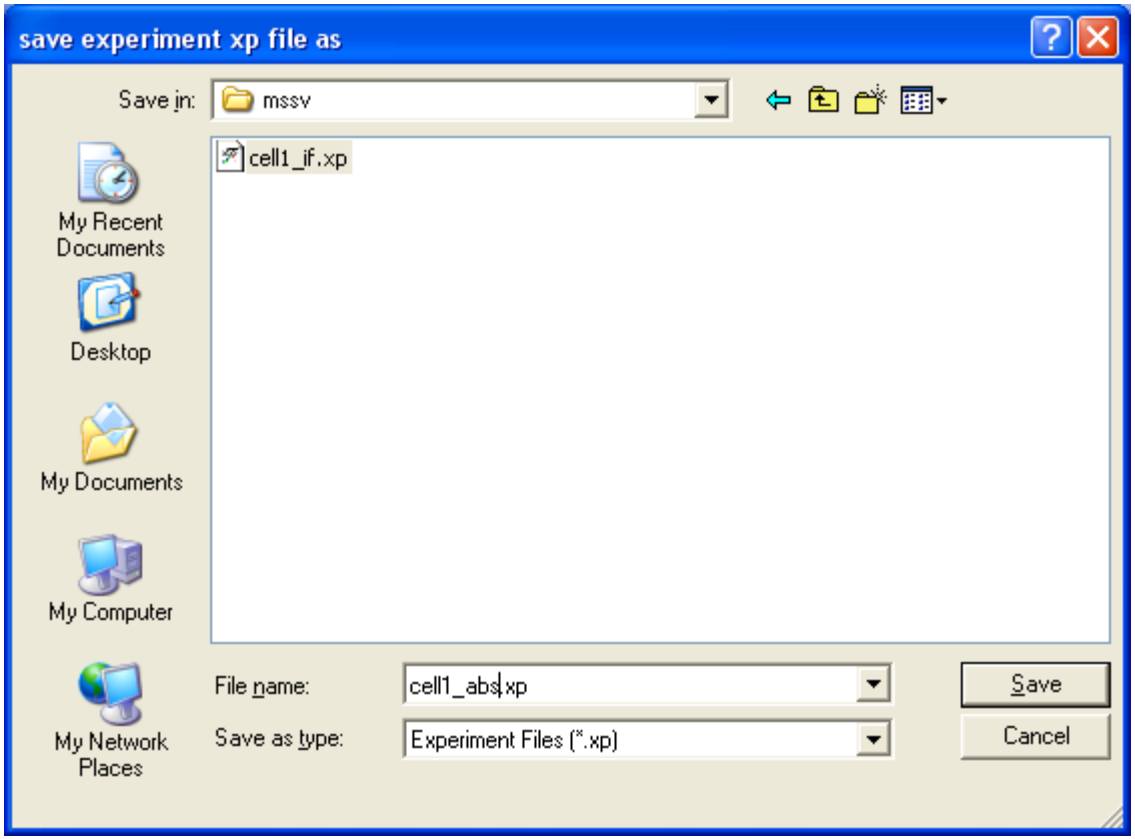
Absorbance files 1-101 are selected…

Experimental Parameters #2

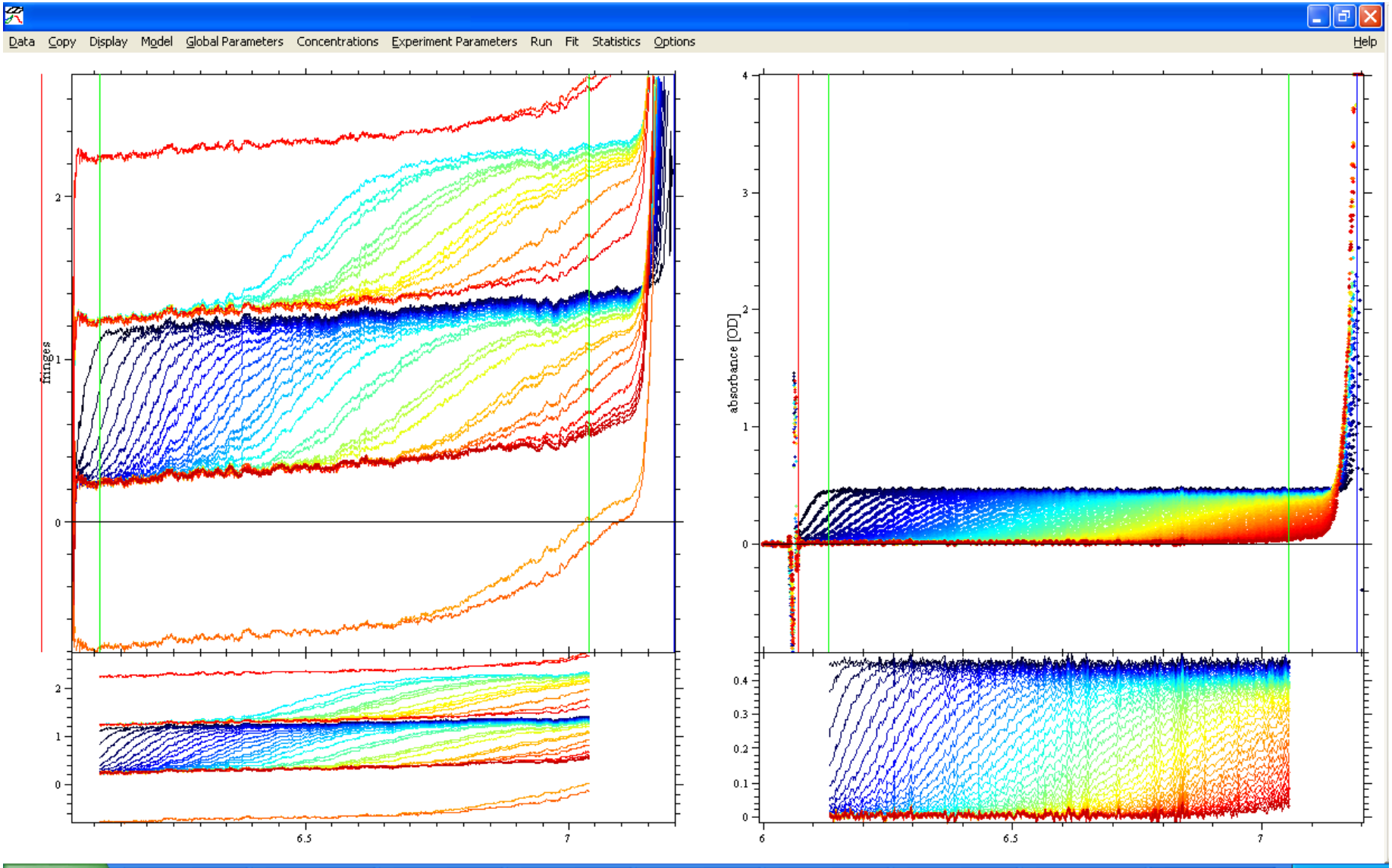


These values should already be input for you by SEDPHAT.

Save the Experiment



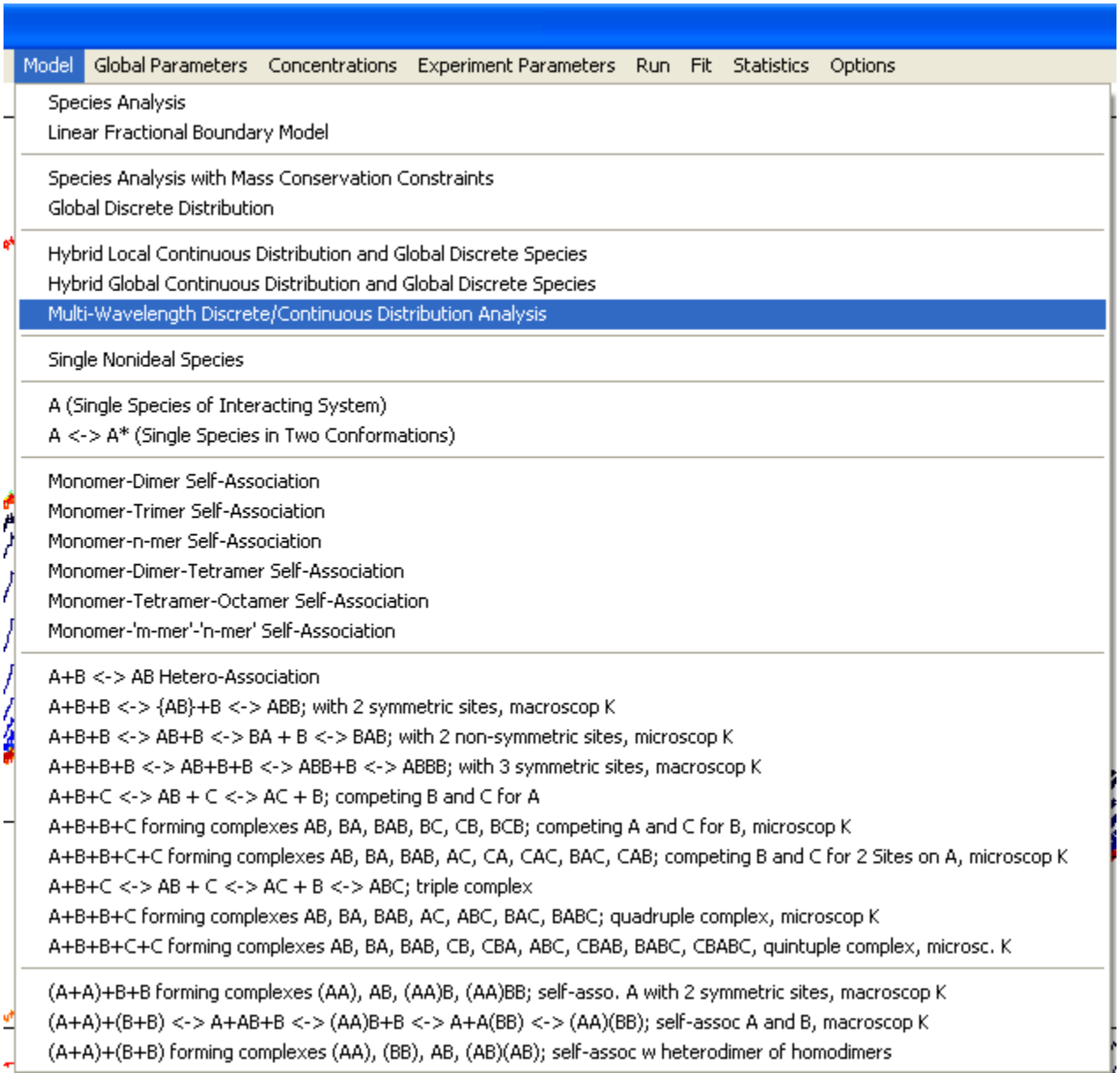
Set Meniscus, Bottom, Fitting Limits for ABS Data



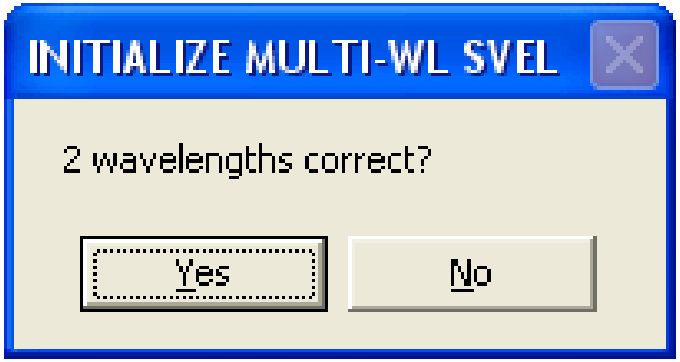
Step 7

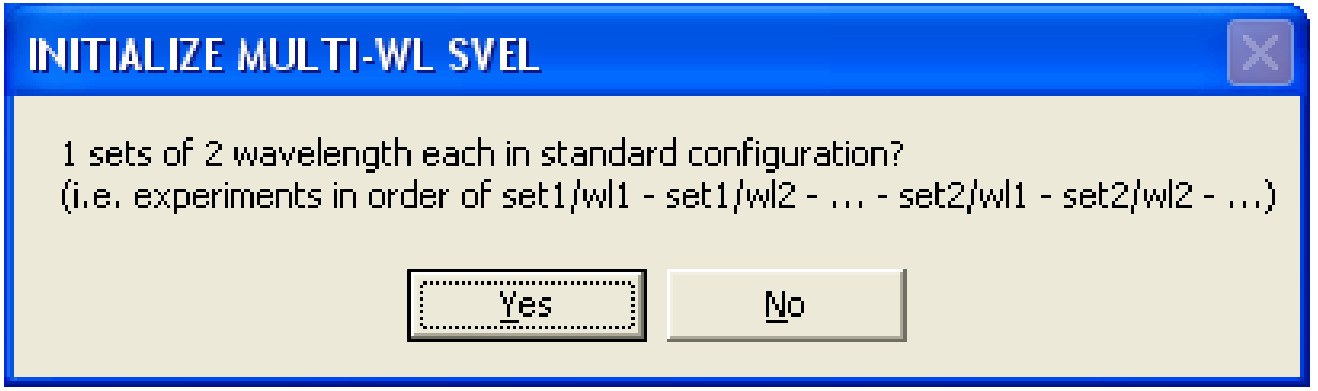
At this point, under the “Model” menu, the model “Multi-Wavelength Discrete/Continuous Distribution Analysis” was chosen. The user was prompted “2 wavelengths correct”; Yes was chosen. The program then asked the user if the data sets were in the “standard configuration,” i.e. if the data sets were loaded in the order wavelengths (N.B.—here, an IF data set is considered a “wavelength”). Yes was chosen. See the next two pages.

Choose Model



Choose Yes in Both Dialogs

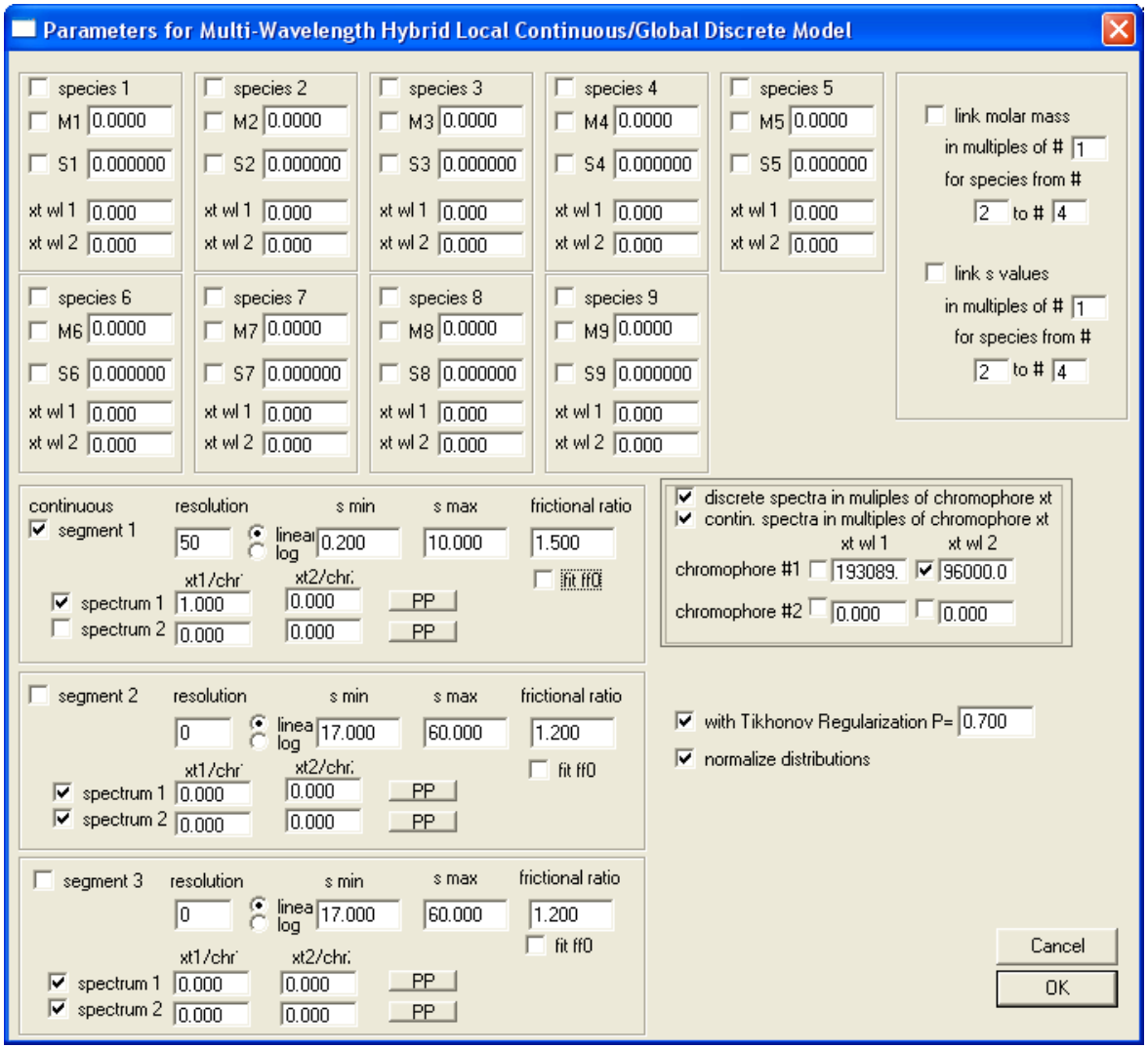




Step 8

The Global Parameters portion of the menu bar was clicked, opening the Global Parameters dialog box. The upper part of the box is for discrete species. This part was not used in this analysis; only continuous species were used. The lower part of the box allows up to three segments of *s*-space to be modeled with continuous distributions. See the next slide

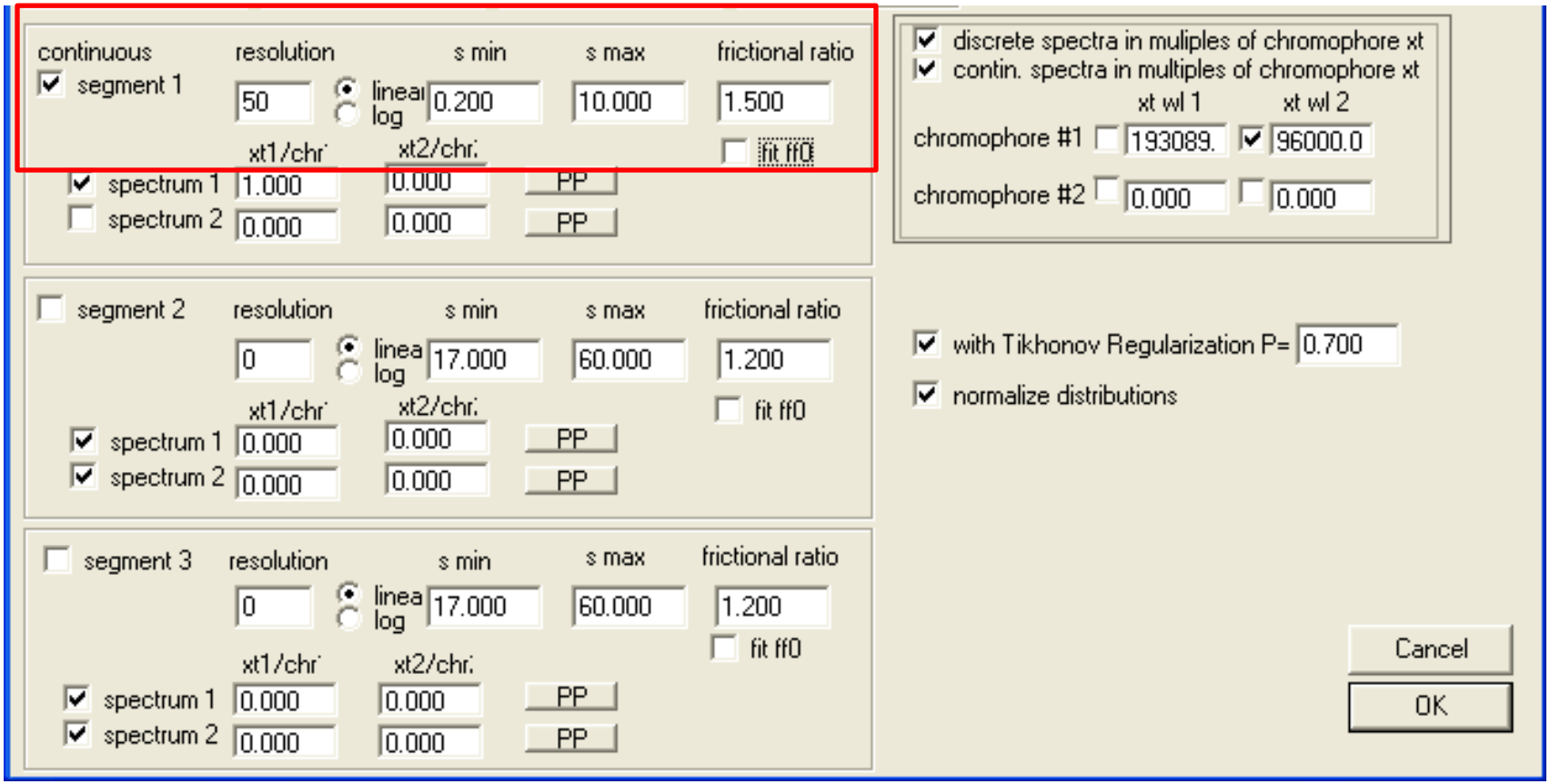
Introducing the MSSV Global Parameter Box.



The upper part of this box will not be used in these analyses; it is omitted in the following.

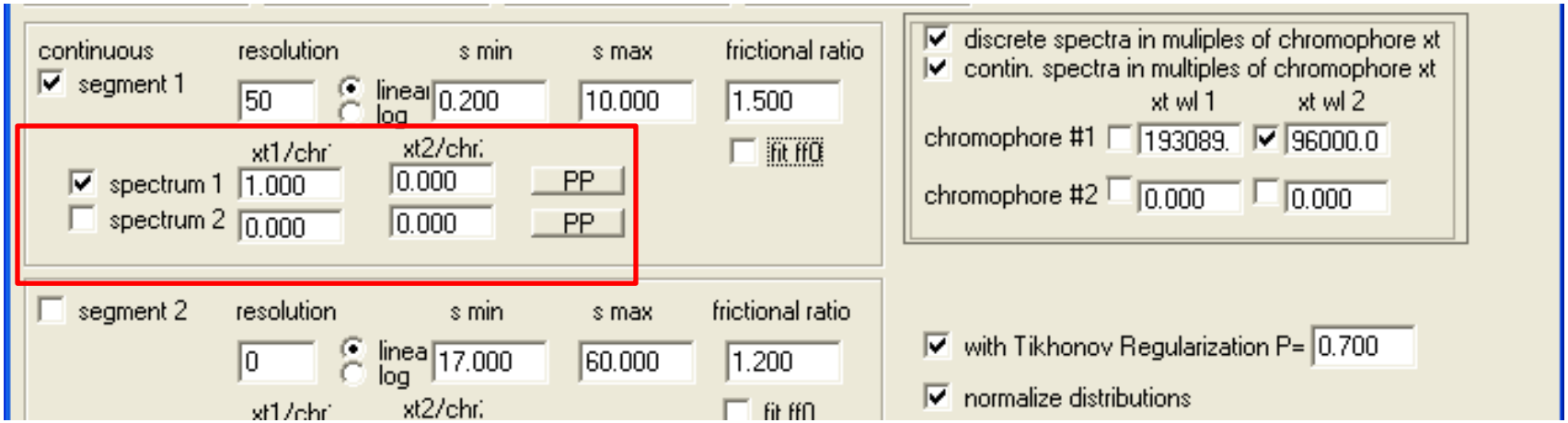
Step 8, continued

The checkbox next to “segment 1” was checked, activating this segment. This segment, which will be evaluated as a $c_{GST\text{-}VCA}(s)$ distribution, was given a resolution of 50, and the “smin” and “smax” values were set to 0.2 and 10, respectively. The frictional ratio was left at the default value of 1.5, and the “fit ff0” checkbox, which would allow the refinement of the frictional ratio, was deactivated.



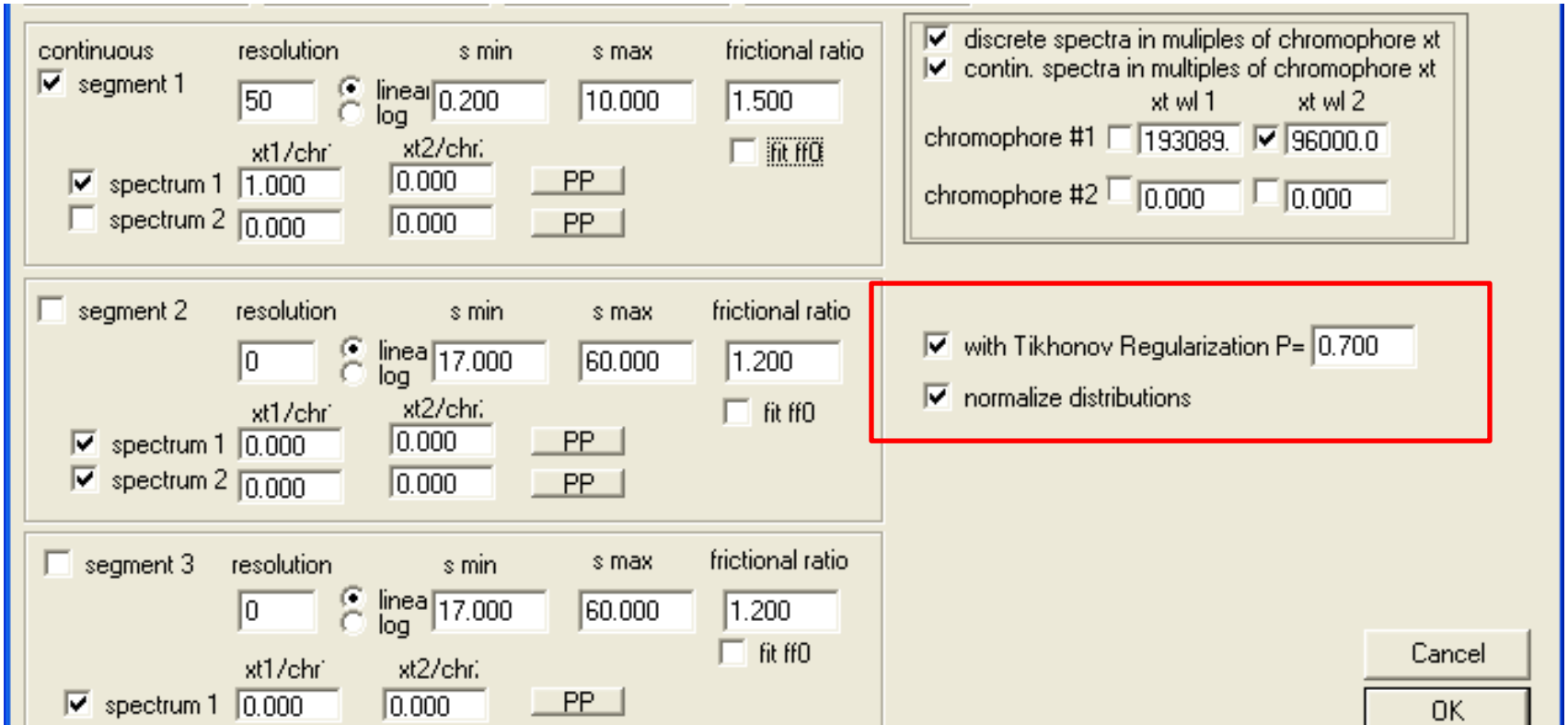
Step 8, continued

Up to two "spectra" may be used in a single segment (for a two-signal experiment). Each spectrum is a single $c_k(s)$ distribution. Because only one component was evaluated here, only a single spectrum is necessary. Therefore, the checkbox to the left of "spectrum 2" was deactivated. The two boxes to the right of "spectrum 1" describe the molar ratio of up to two chromophores (macromolecules) that were to be modeled by the "spectrum 1" $c_k(s)$ distribution. Because only one chromophore is present, a 1:0 molar ratio of chromophore 1 (GST-VCA) to chromophore 2 (not present) was to be modeled. In other words, spectrum 1 will represent a $c_{GST\text{-}VCA}(s)$ distribution. Therefore, "1" and "0" were entered into the left and right boxes, respectively, next to spectrum 1.



Step 8, continued

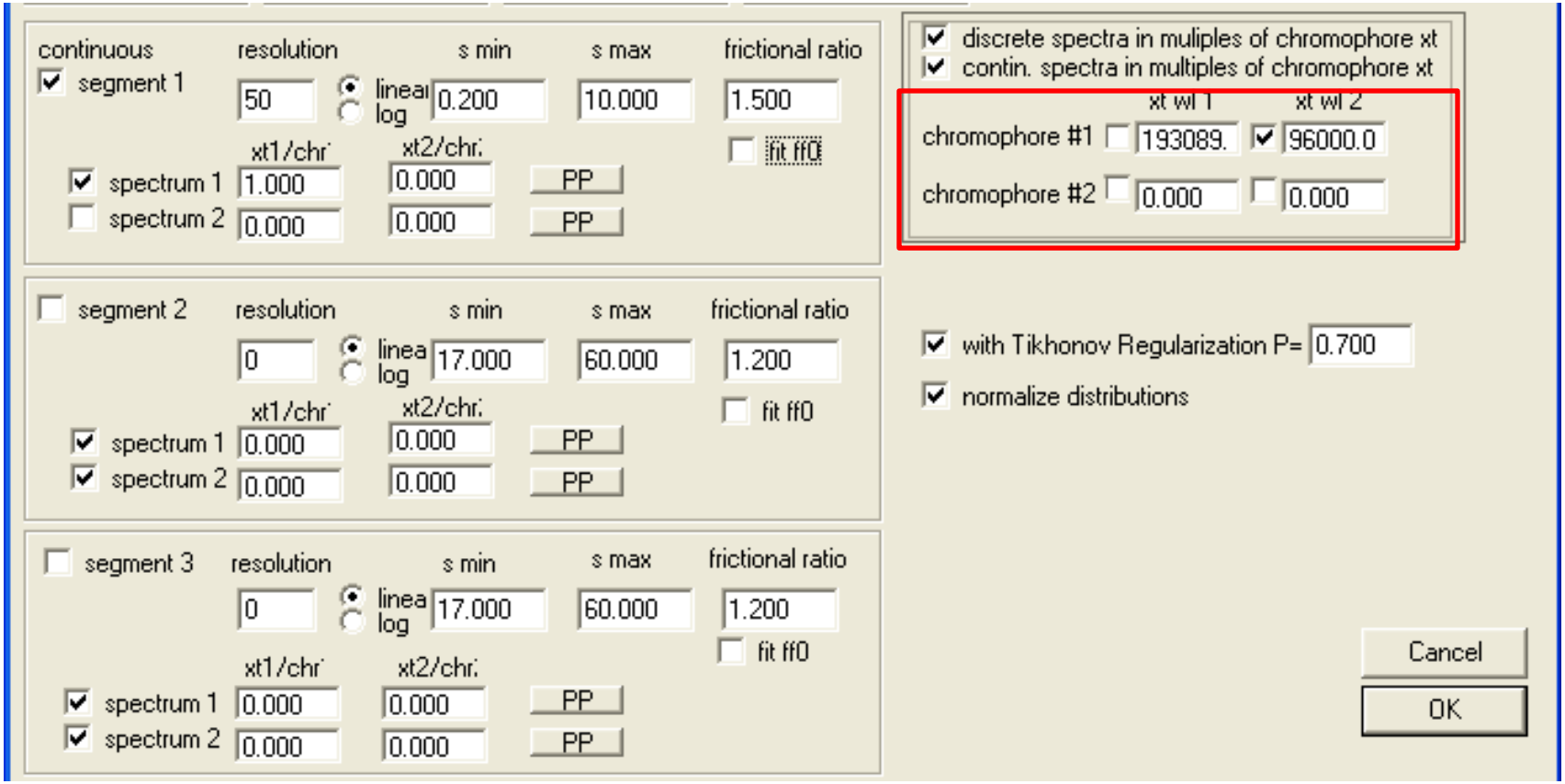
The checkboxes next to “Tikhonov regularization” and “normalize distributions” were activated. The former box regularizes the distribution by suppressing sharp features therein. This step introduces numerical stability in the calculation of *c*(*s*)-style distributions [1]. The latter box normalizes the distribution such that the area beneath a peak is equal to the concentration of the sedimenting material.



Step 8, continued

Finally, the “extinction” information for GST-VCA must be entered in the area of the box above the Tikhonov regularization section. The four boxes in this area represent a 2x2 extinction matrix possible for 2 species with data obtained at 2 wavelengths, which is the matrix E in Equation 7. Because only one “chromophore” is present, only the “chromophore #1” row of the matrix will be used. The left-hand column, labeled “xt wl 1,” is for the signal increments for experiment 1. Here, we use the straightforward and very reliable estimate of the IF signal increment for GST-VCA ($\varepsilon_{\mathrm{IF}}^{\mathrm{GST\text{-}VCA}}$) to be 2.75 multiplied by the molar mass of the protein [2,3]. For GST-VCA, $\varepsilon_{\mathrm{IF}}^{\mathrm{GST\text{-}VCA}}$ = 193,089 fringes·M^{-1}·cm^{-1}. Note that this calculation was based on the dimer molecular mass of GST-VCA (70,214 g/mol, calculated from the sequence of the protein). Thus, all concentrations for this protein in this protocol will be for the dimer, not for the monomer. The $\varepsilon_{\mathrm{IF}}^{\mathrm{GST\text{-}VCA}}$ was fixed in the analysis, and so the checkbox to its left remained unchecked. An estimate for the molar extinction coefficient of GST-VCA at 280 nm ($\varepsilon_{\mathrm{ABS280}}^{\mathrm{GST\text{-}VCA}}$) is 96,000 M^{-1}·cm^{-1}, obtained from SEDNTERP. This value was input into the right-hand box, and the checkbox to its left was activated, allowing this value to refine. See the next page.

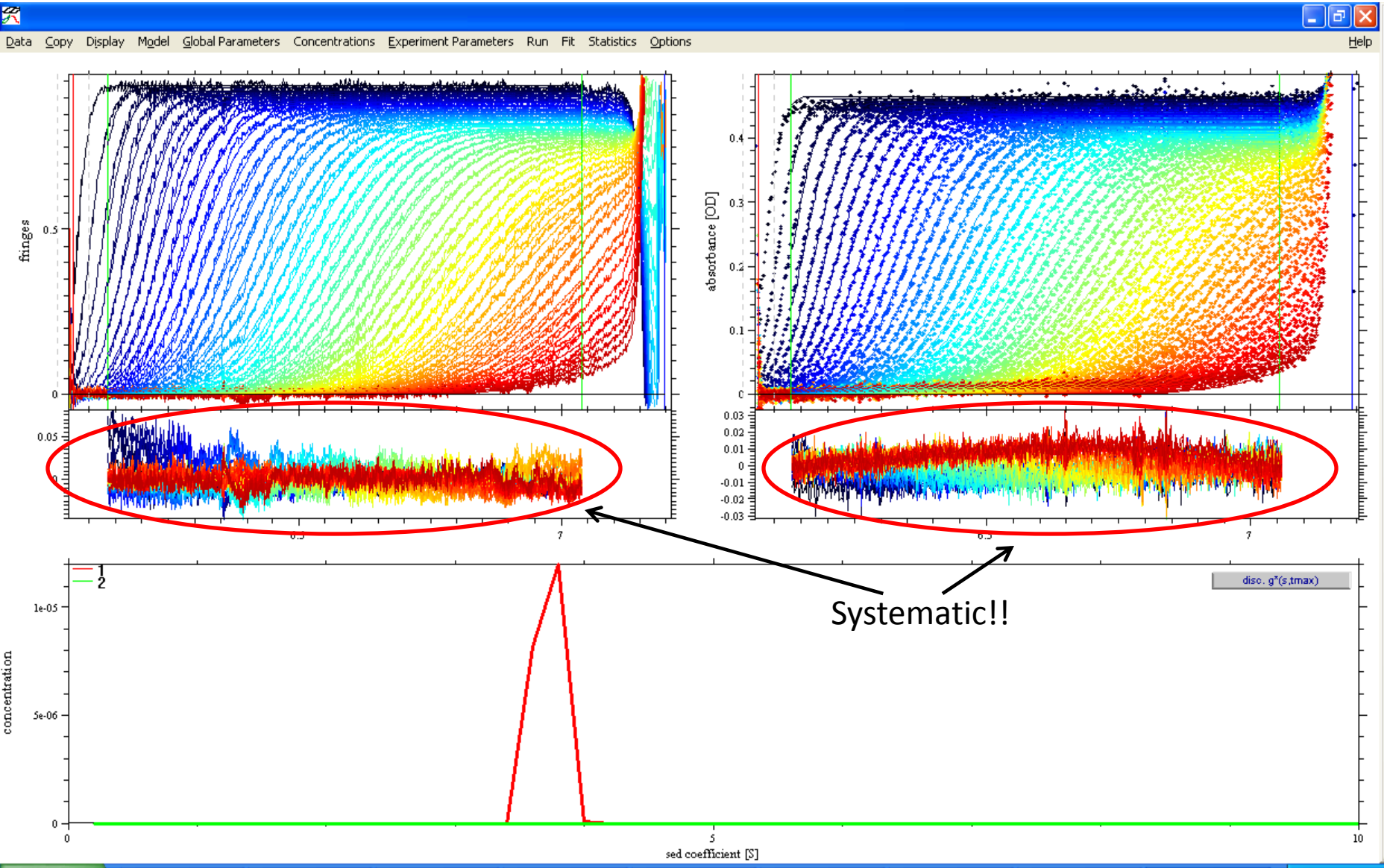
Extinction Info



Step 9

A Global Run was performed by selecting the “Global Run” item in the “Run” dropdown menu. This action optimized the linear parameters, both global (the $c_{GST\text{-}VCA}(s)$ distribution) and local (the noise elements) to each experiment. For the ease of visually inspecting the quality of the fits to the experimental data, the RI and TI noise were subtracted from the data. The quality of the fit (rmsd’s of 0.01 and 0.006 for the IF and A_{280} data, respectively) was deemed to be acceptable to proceed. There is very significant systematicity in the residuals at this point (see circled areas on next page); we hoped that this would be resolved by parameter refinement. The $c_{GST\text{-}VCA}(s)$ distribution shows a single peak at ~3.7 S. See the next page.

After a Global Run



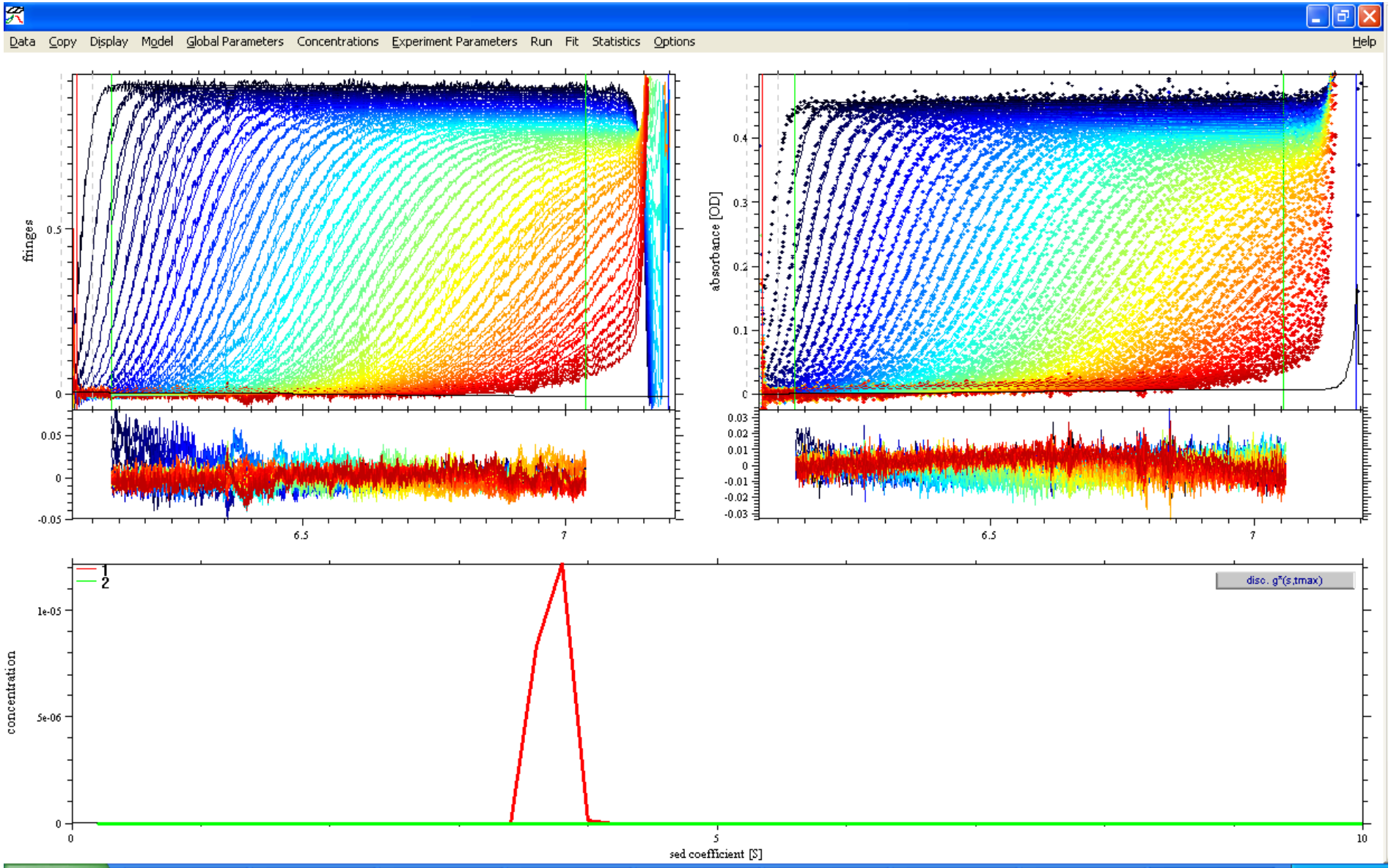
A Note on Noise Subtraction

On the previous page, radially invariant (RI) and time-invariant (TI) noise had been subtracted from both of the experiments. The subtraction can be accomplished in one of two ways: one can go to the “Display” menu item, click on it, then choose “Subtract All Systematic Noise From Raw Data”; or, one can use the simple keyboard shortcut of pressing “Ctrl-N.”

Step 10

A Global Fit was initiated by selecting the menu item Fit and clicking on “Global Fit.” This sequence iteratively optimized the nonlinear parameters. Because the menisci and frictional ratio are fixed, the only nonlinear parameter that was optimized during this analysis was $\varepsilon_{\mathrm{ABS280}}^{\mathrm{GST\text{-}VCA}}$. The fit improved slightly, but the rounded rmsd’s were still 0.01 and 0.006. There is still significant systematicity to the residuals. Examination of the Global Parameters demonstrated that $\varepsilon_{\mathrm{ABS280}}^{\mathrm{GST\text{-}VCA}}$ had refined to 92268.2 $M^{-1} \cdot cm^{-1}$. SEDPHAT can use one of three nonlinear optimization algorithms: Simplex, Marquardt-Levenberg, and Simulated Annealing. Simplex is set by default, and it was this method that was used for this initial optimization. The result is shown on the next page.

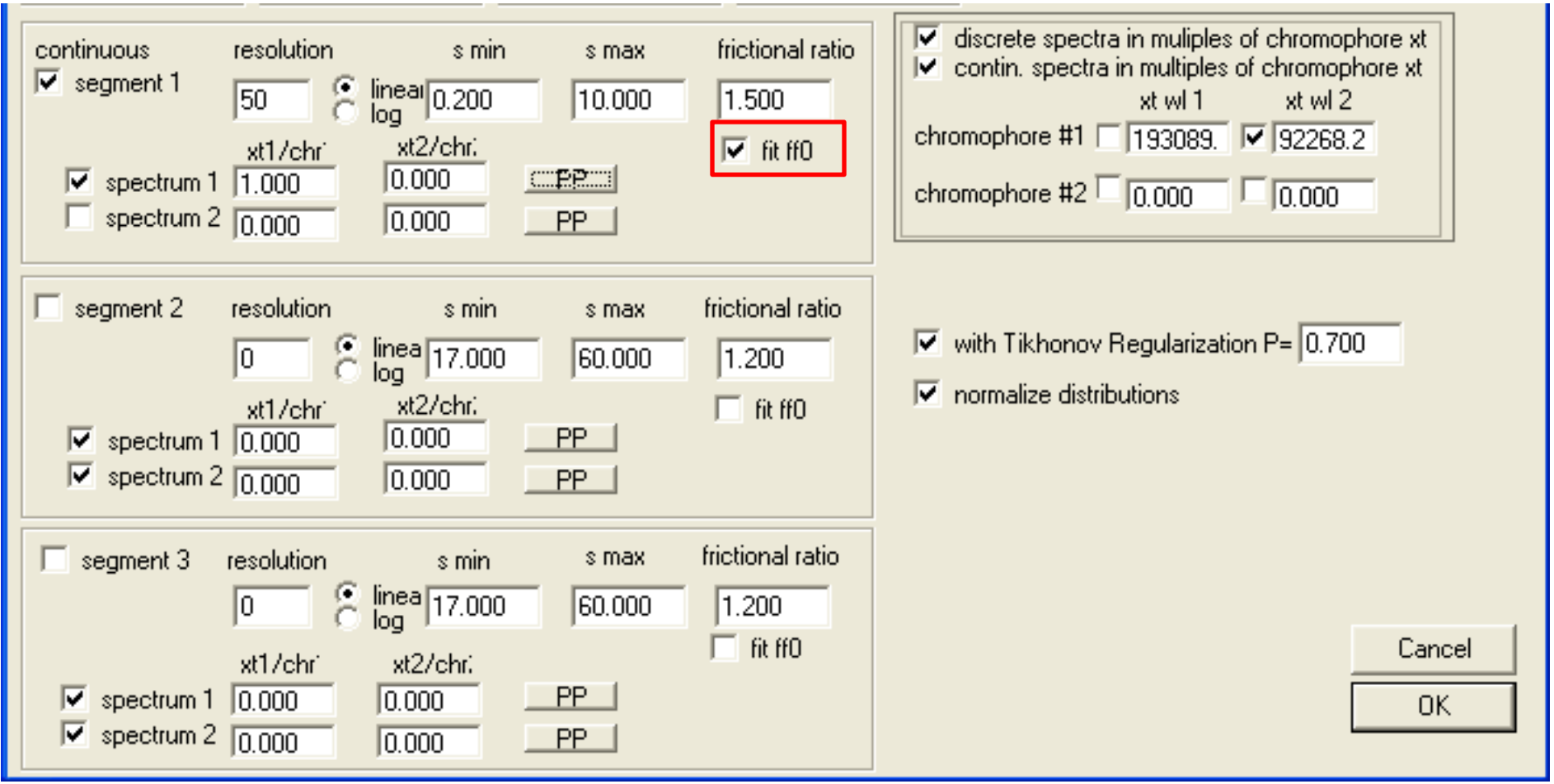
After a Global Fit



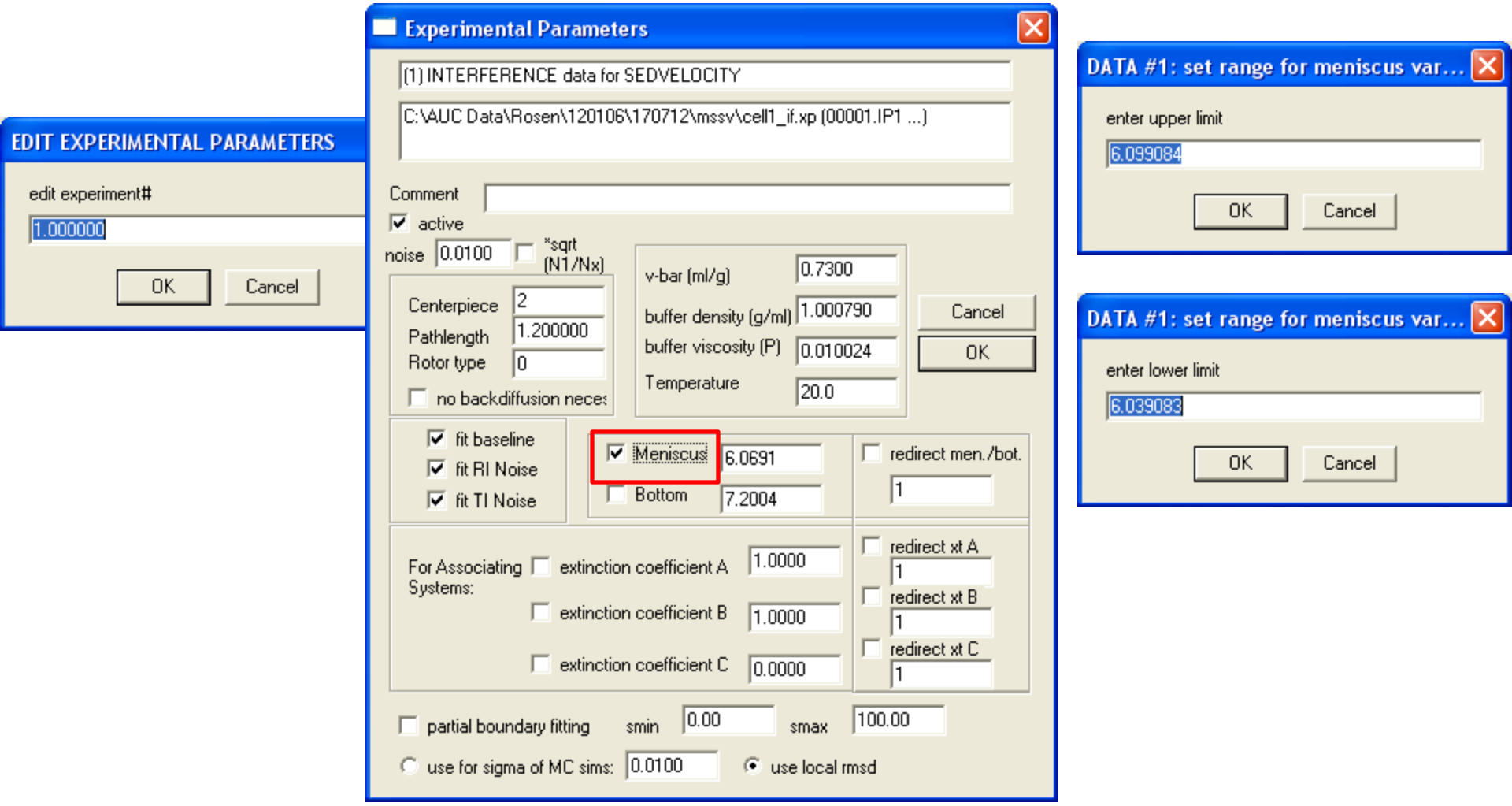
Step 11

The remaining appropriate nonlinear parameters were set to refine by checking the “fit ff0” box for segment 1 in the Global Parameter box, and clicking on the checkbox to the left of “Meniscus” in the Experimental Parameter boxes of both experiments. After accepting each of the Experimental Parameter boxes, the user was prompted for the fitting limits for the respective meniscus. These limits are hard; the program does not allow for the refinement of the meniscus outside these bounds. The default limits were the current meniscus ± 0.03 cm. Because the menisci had been visually chosen to be at acceptable values, the defaults were accepted. Experience has shown that the Marquardt-Levenberg fitting algorithm is the most efficient means to a converged fit in such analyses, so it was selected by clicking on “Options,” then “Fitting Options,” then “Marquardt-Levenberg.” The pictorial view of the substeps of Step 11 are illustrated on the next four pages.

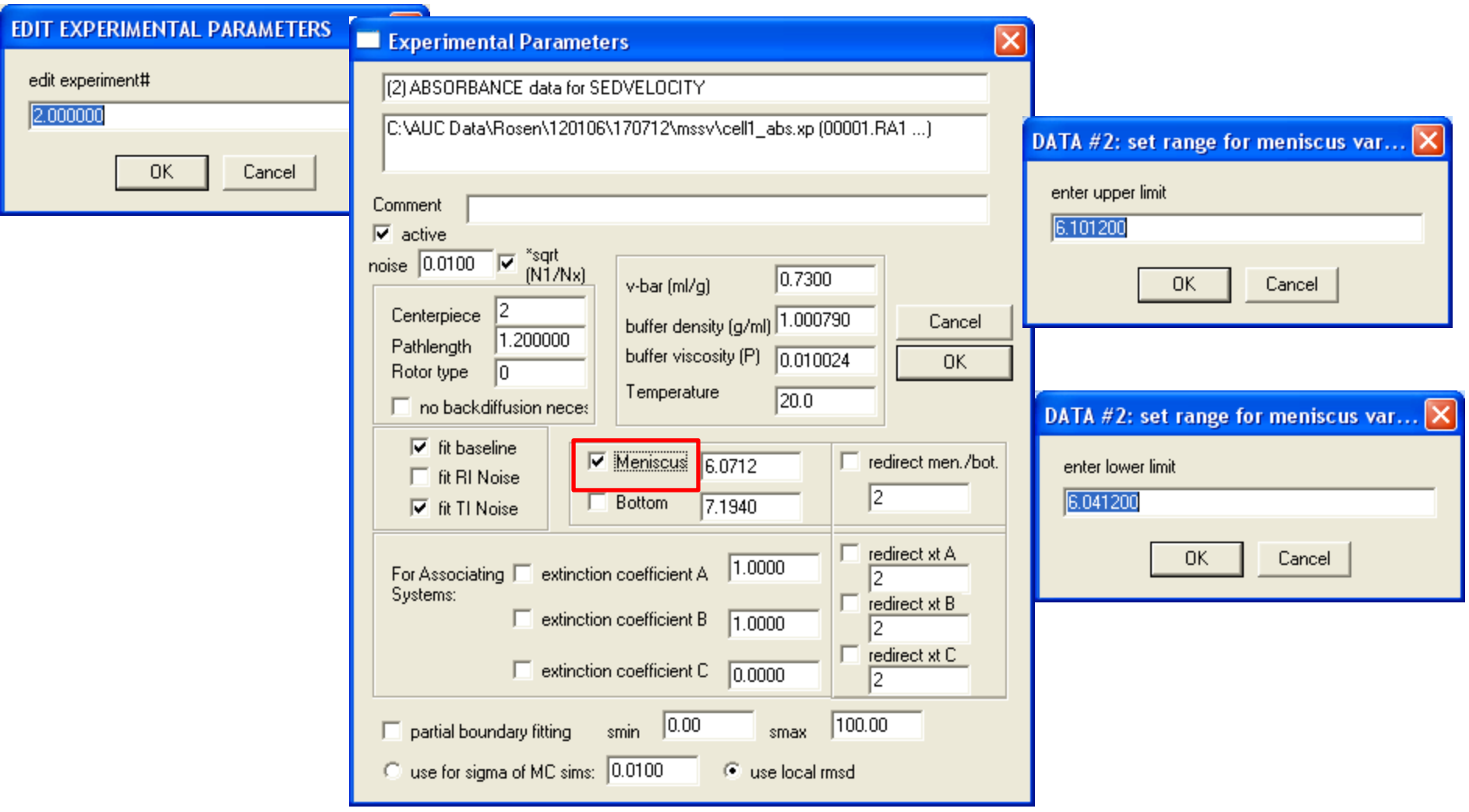
Fit ff0 Now (check its box)



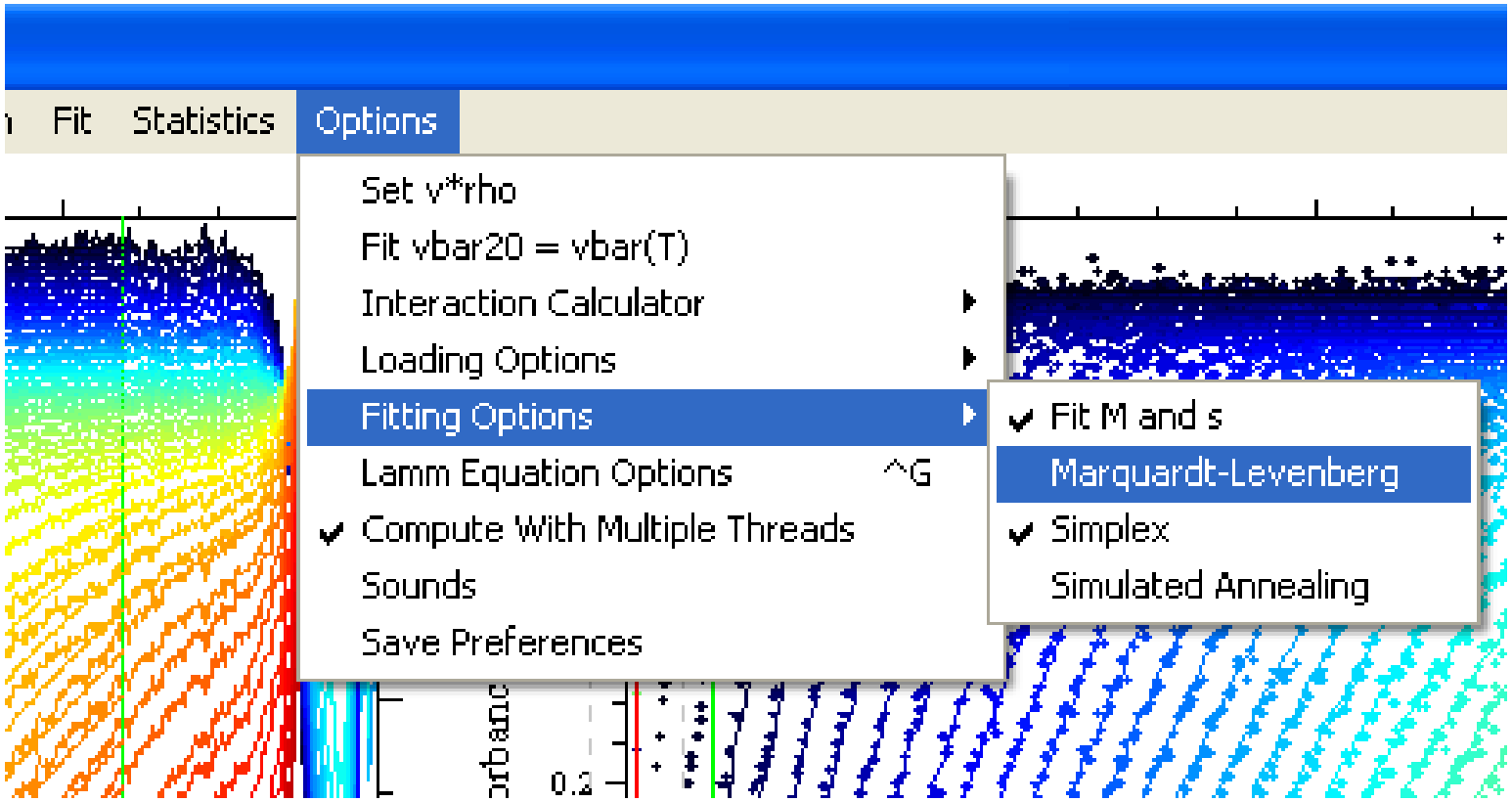
After Clicking on Experimental Parameters, Activate Meniscus Fitting



Activate Meniscus Fitting for Experiment #2



Change the Fitting Algorithm



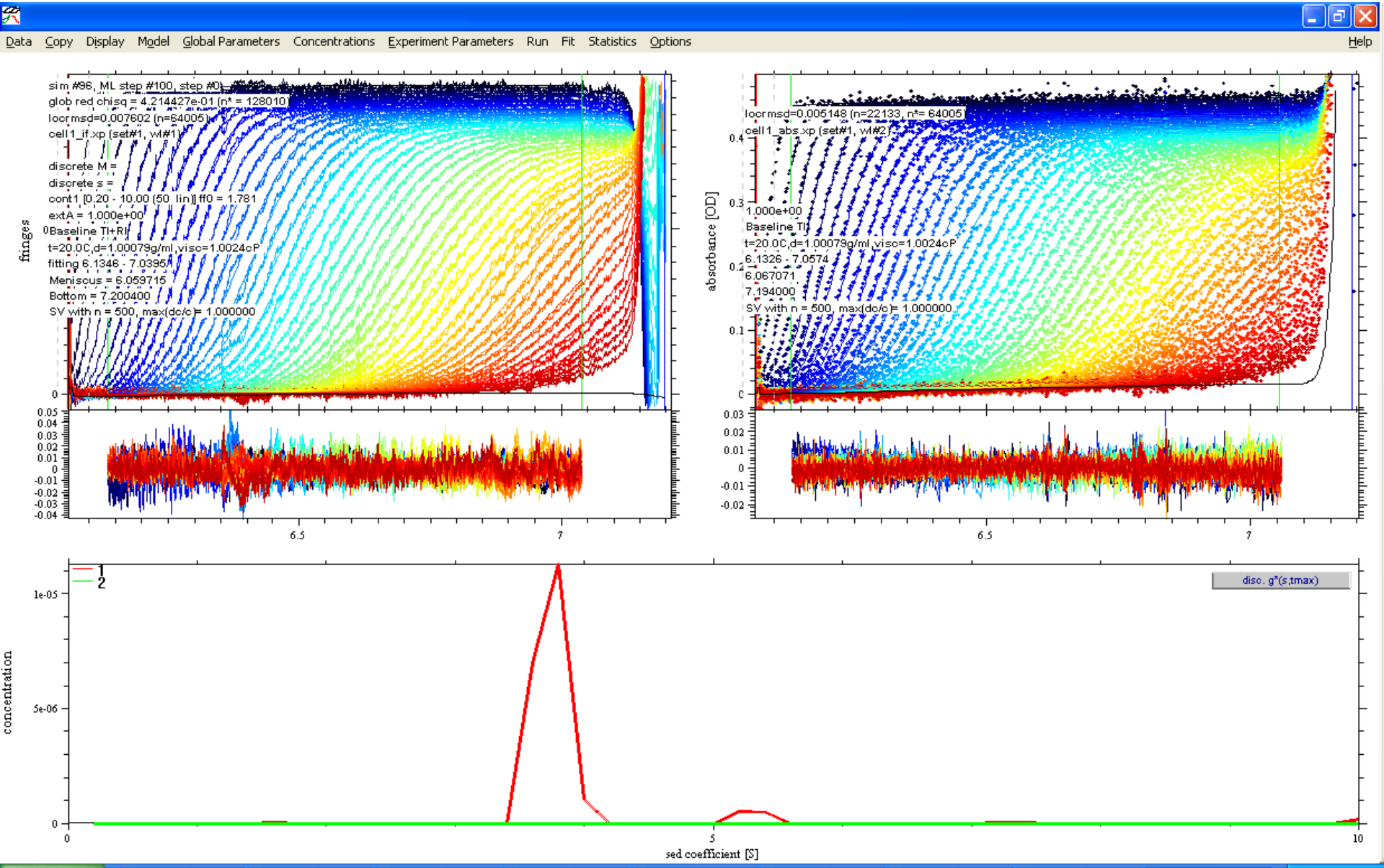
Step 12

Another Global Fit was performed. After about three minutes, the fit converged.

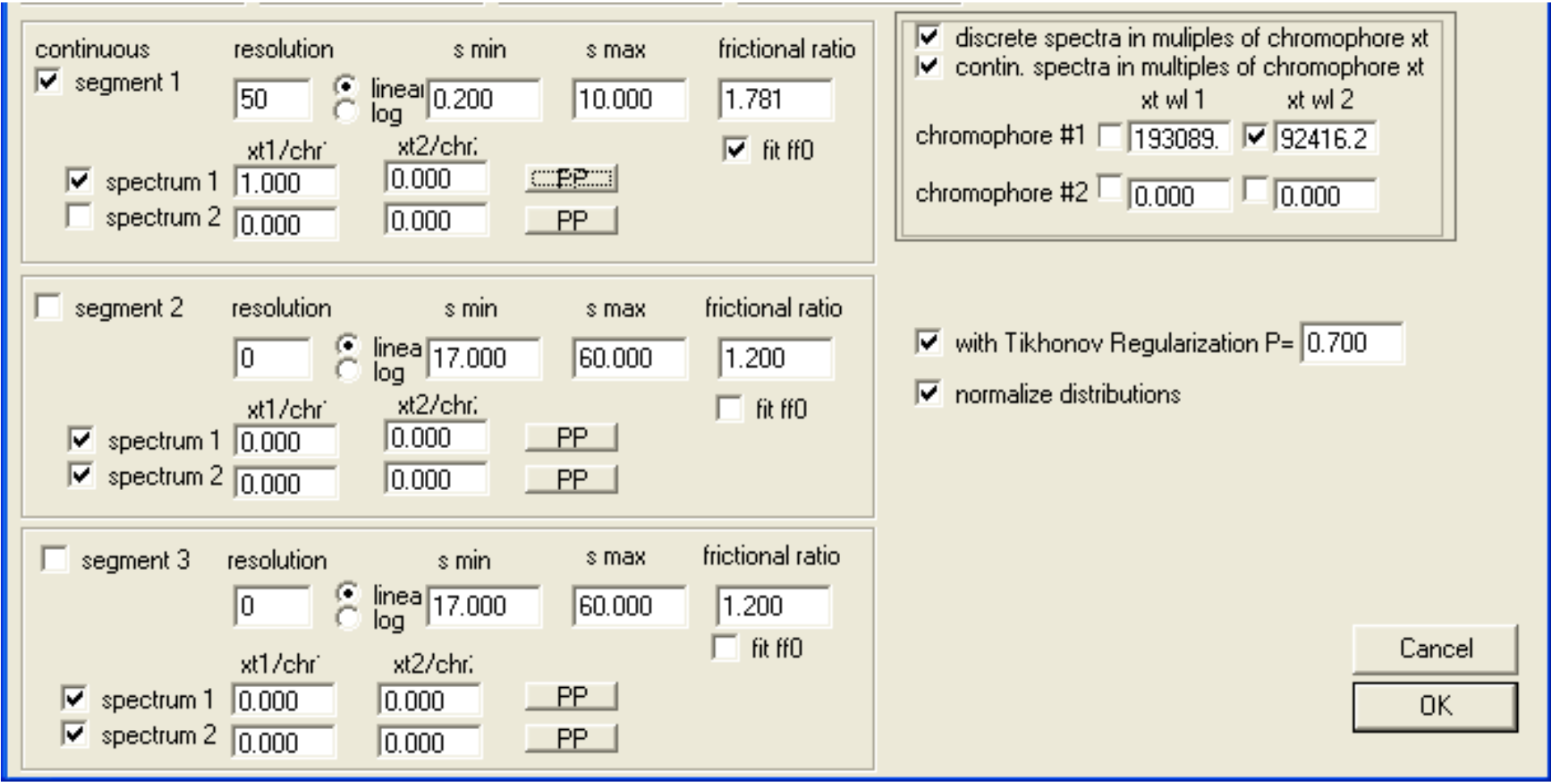
Step 13

The quality of the fit was inspected (see the next two pages). The rmsd's were markedly improved to 0.008 fringes for the IF data and 0.005 OD for the A_{280} data. Additionally, the residuals exhibit much less systematicity. Now, the $c_{\text{GST-VCA}}(s)$ distribution exhibited two peaks: one at 3.7 S, and the other at 5.3 S. Examination of the Global Parameters yielded two significant facts: the $\varepsilon_{\text{ABS280}}^{\text{GST-VCA}}$ had refined to 92,416.2 $M^{-1}\cdot cm^{-1}$, and the f_r had refined to 1.781. The latter value indicated that GST-VCA has a significantly elongated shape in solution. Of course, the refinement of f_r is sensitive to the value of vbar. Thus, the 1.781 value reported here is not the correct value for GST-VCA, because the default value of vbar was accepted (Step 2). However, inputting the correct value of vbar (0.7355cm^3/g) does not significantly change the refined value for f_r (1.743). The next page shows a screenshot of the fitting session after this Global Fit, and the page after shows the resulting parameters in the Global Parameter box.

After another Global Fit



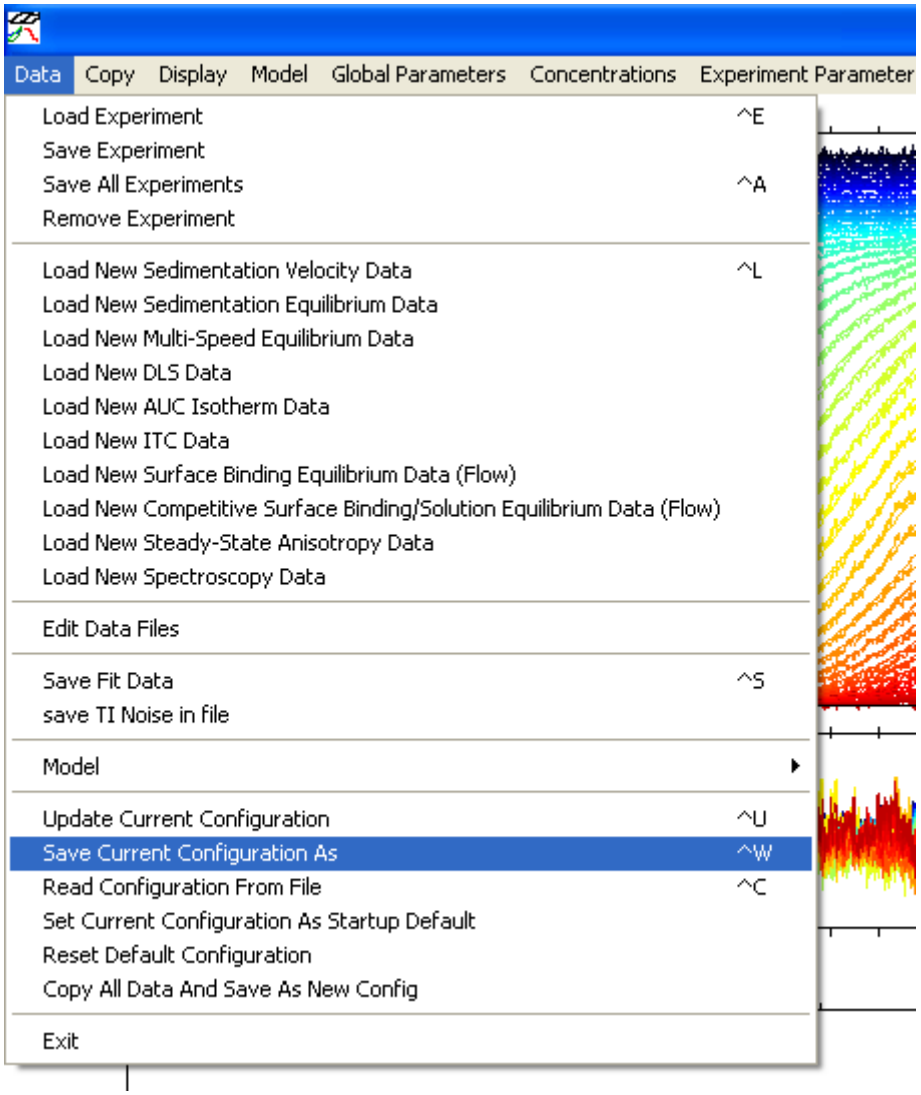
The Resulting Parameters



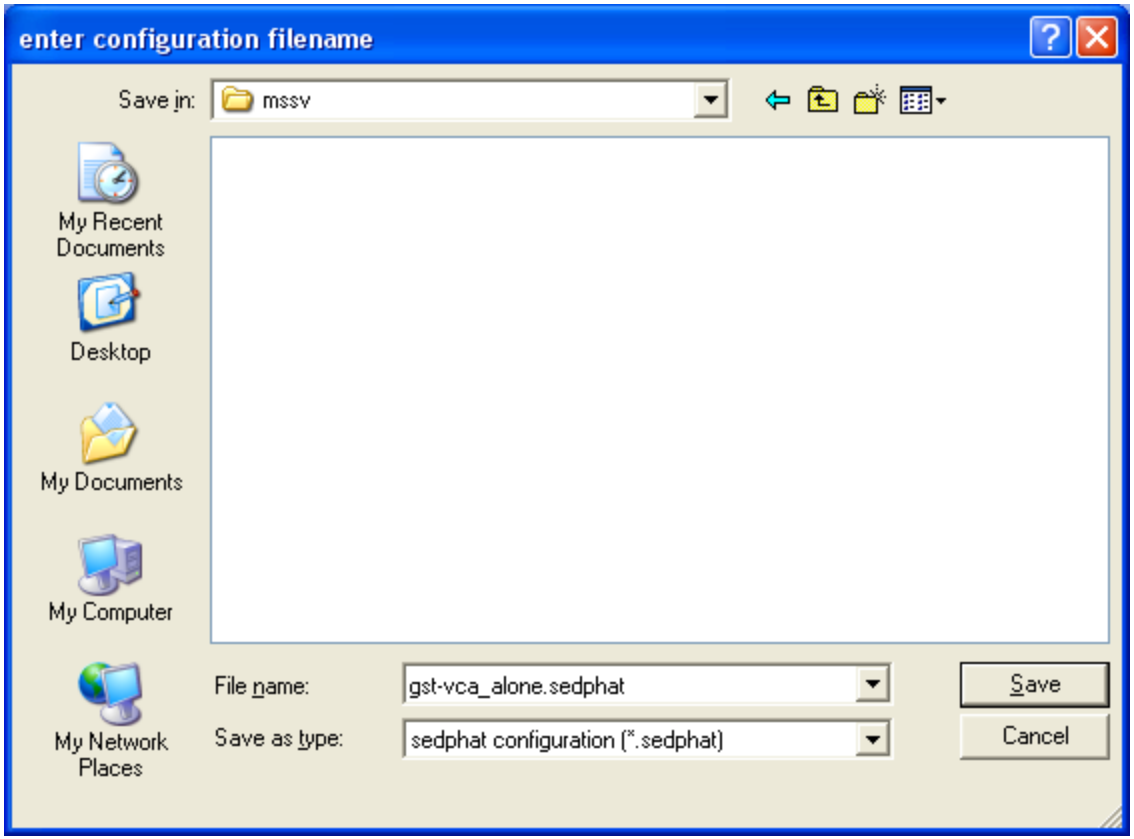
Step 14

The fitted parameters were saved by clicking on the menu item “Data” and selecting “Save Current Configuration As.” The user was prompted to save new copies of the experimental parameter files with updated parameters. This option was accepted. See the next two pages. It is a good idea at this point to note down the refined $\varepsilon_{\mathrm{ABS280}}^{\mathrm{GST\text{-}VCA}}$ so that it can be easily recalled for input later (Step 21). Of course this value can be recalled at a later time by loading the saved configuration into SEDPHAT.

Save the Configuration



Saving…

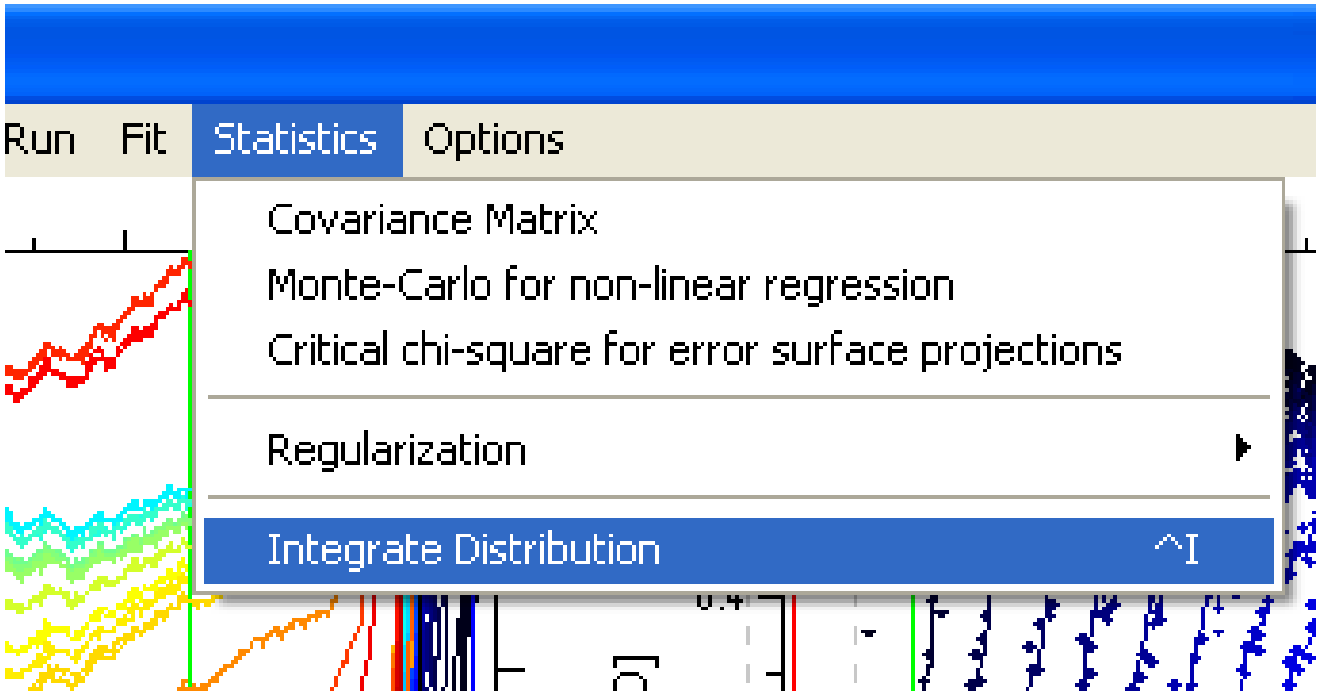

enter configuration filename
Save in:
mssv
My Recent Documents
Desktop
My Documents
My Computer
My Network Places
File name:
gst-vca_alone.sedphat
Save as type:
sedphat configuration (*.sedphat)
Save
Cancel

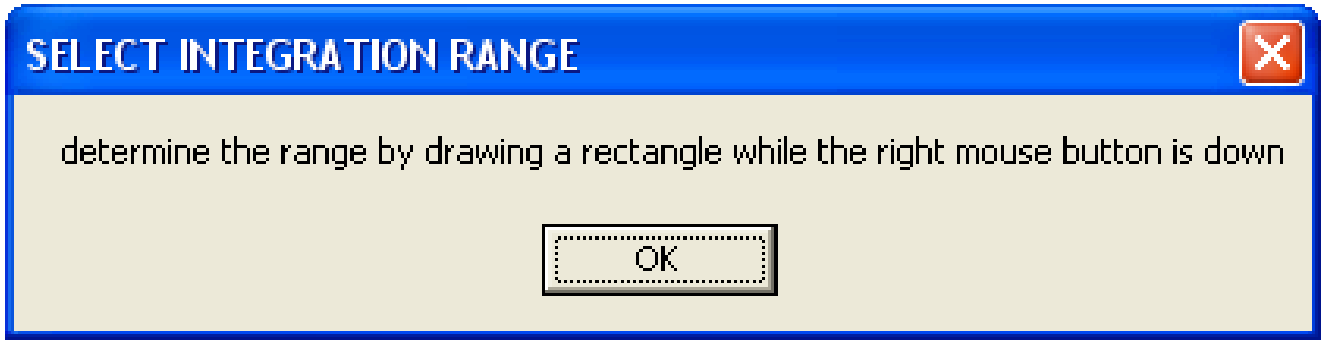

SAVE CONFIGURATION
include new copies of xp-files?
Yes
No

Step 15

Under Statistics, the “Integrate Distribution” option was selected. The user was prompted to “determine the range by drawing a rectangle while the right mouse button is down.” The entire distribution was highlighted by dragging over it with the right mouse button depressed. The resultant [GST-VCA] was 4.09 μM. This process is illustrated in the next 3 pages.

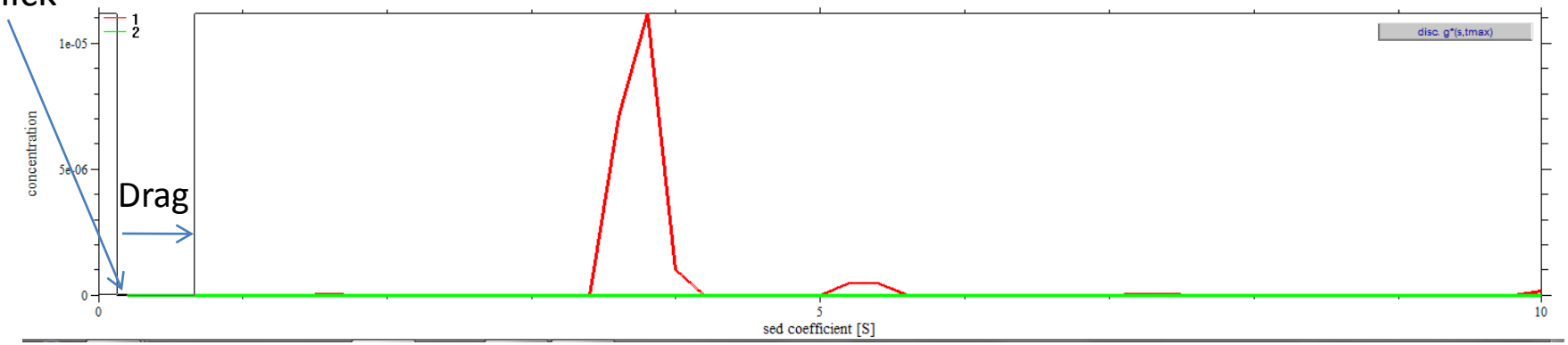
Integrate the Distribution

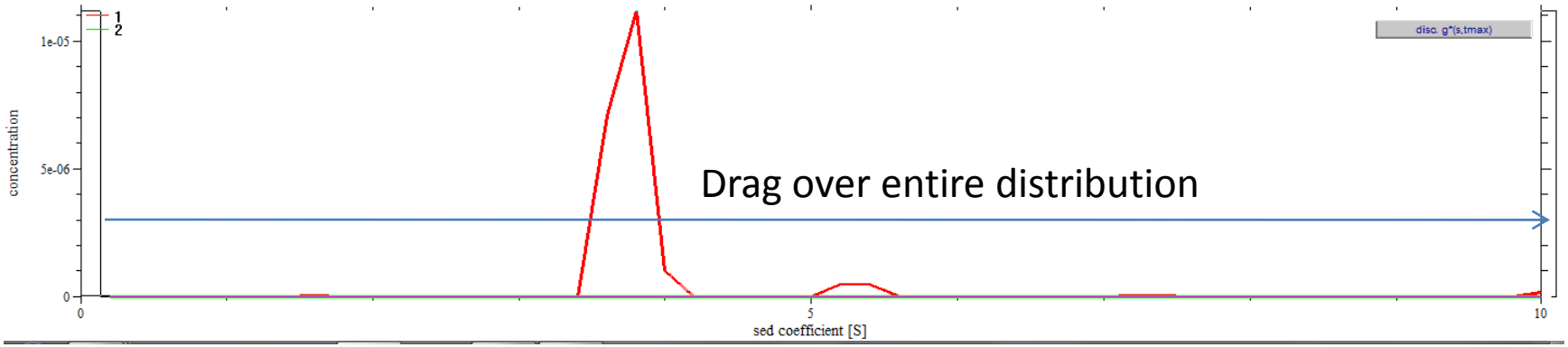




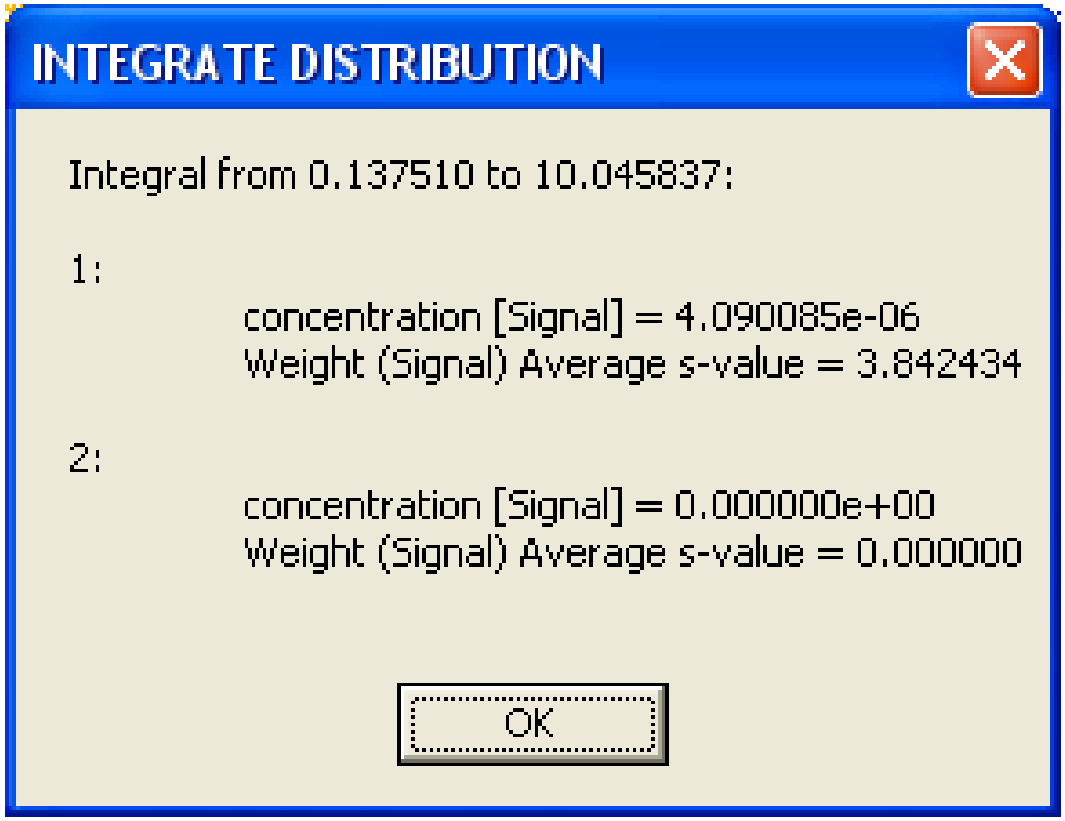
Right Click and Drag

Right click





Integration



Incidentally, SEDPHAT automatically copies these values into the clipboard for you!

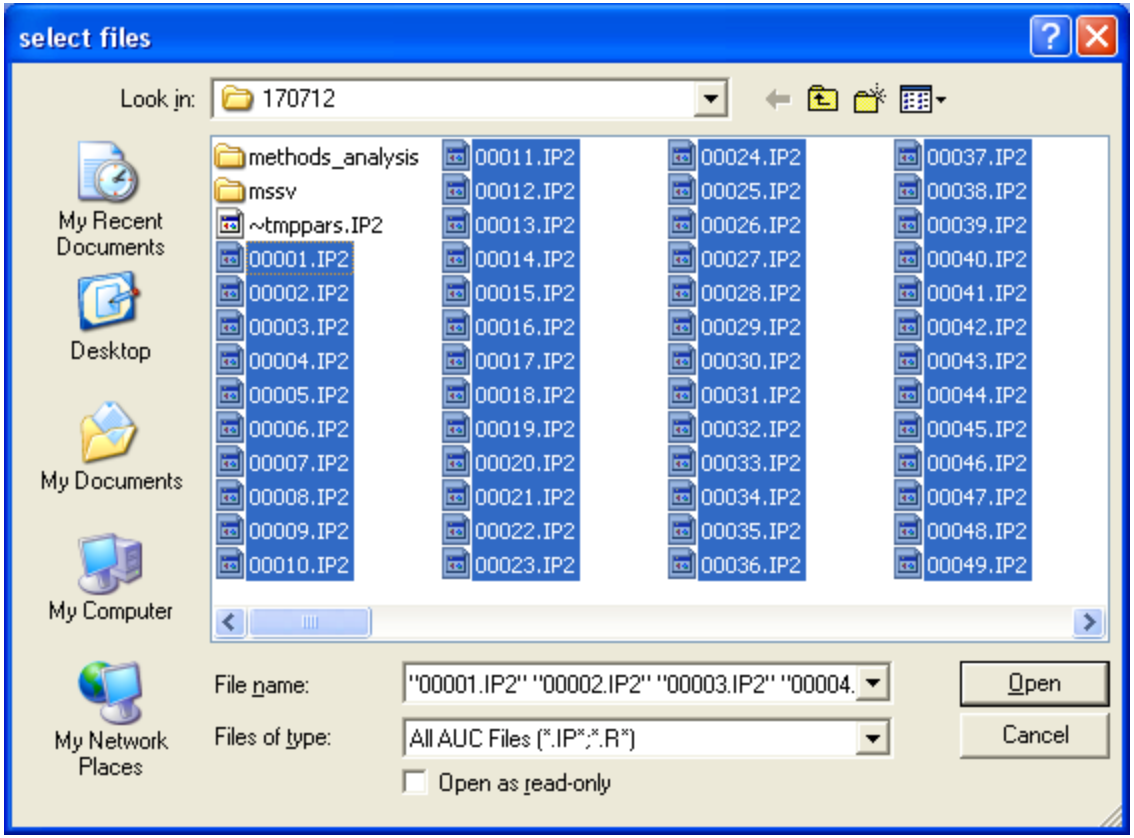
Step 16

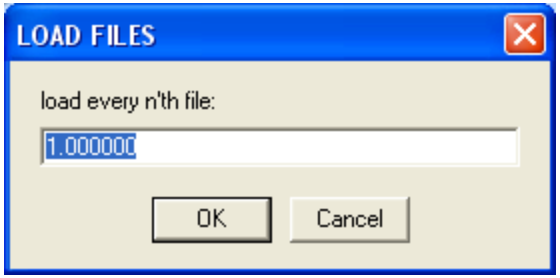
At this point, more fitting could have occurred. For example, the fitting algorithm could have been changed back to Simplex, and the fitting session repeated. However, experience has shown that continued fitting at this point will not significantly change the fitted parameters nor significantly improve the quality of the fit.

Step 17

A new SEDPHAT session was started, and the data from Cell 2 were loaded into the program as in Step 1 above (with the IF data as Experiment 1 and the A_{280} data as Experiment 2; the data range was scans 1-50). These data were obtained from the sample with Arp2/3 alone. Steps 2-7 were performed identically for these data. See Table 4 for the pertinent parameters. This part of the analysis is detailed in the next 8 pages.

A new SEDPHAT session- Load interference data from cell 2 this time

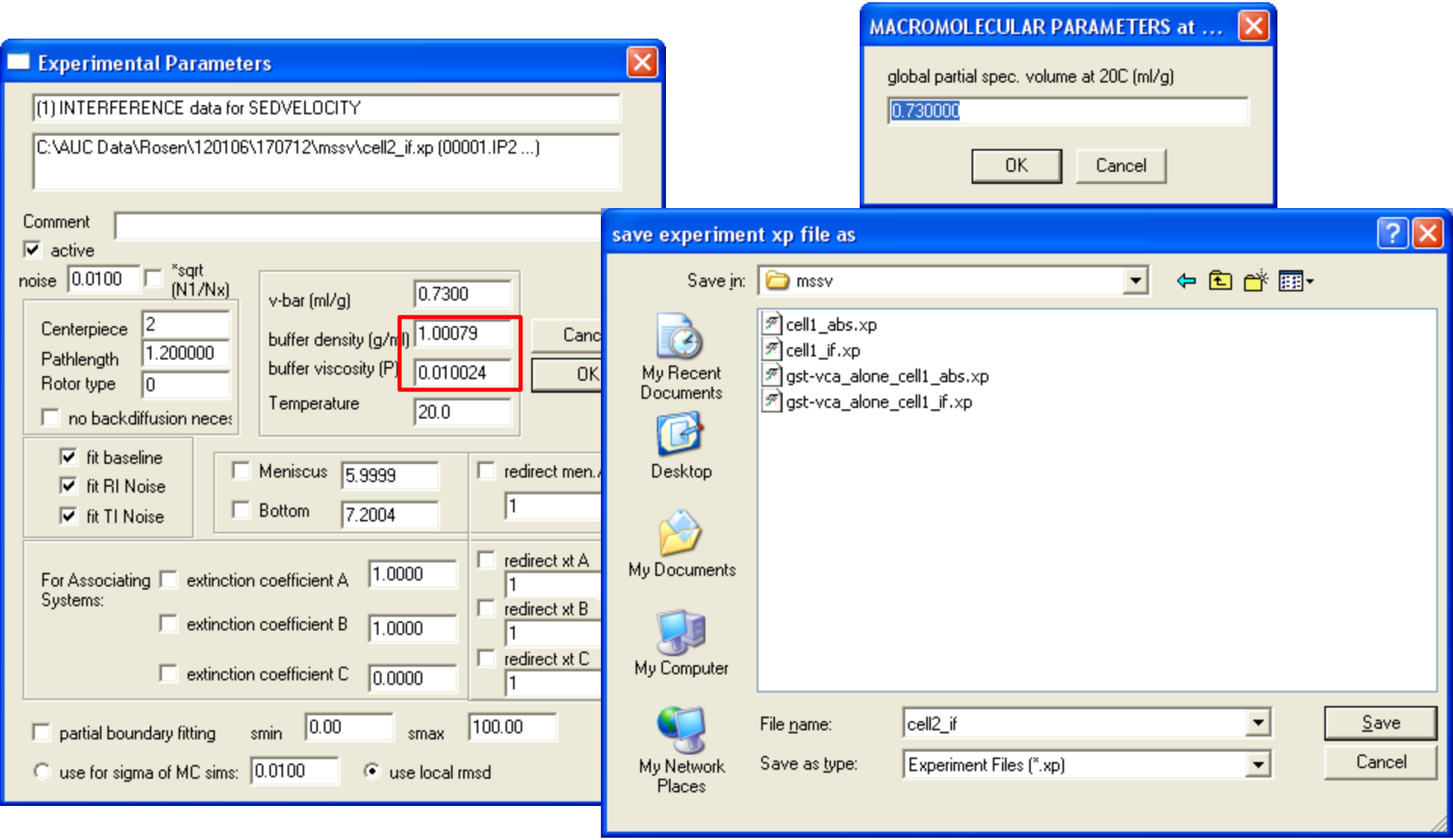




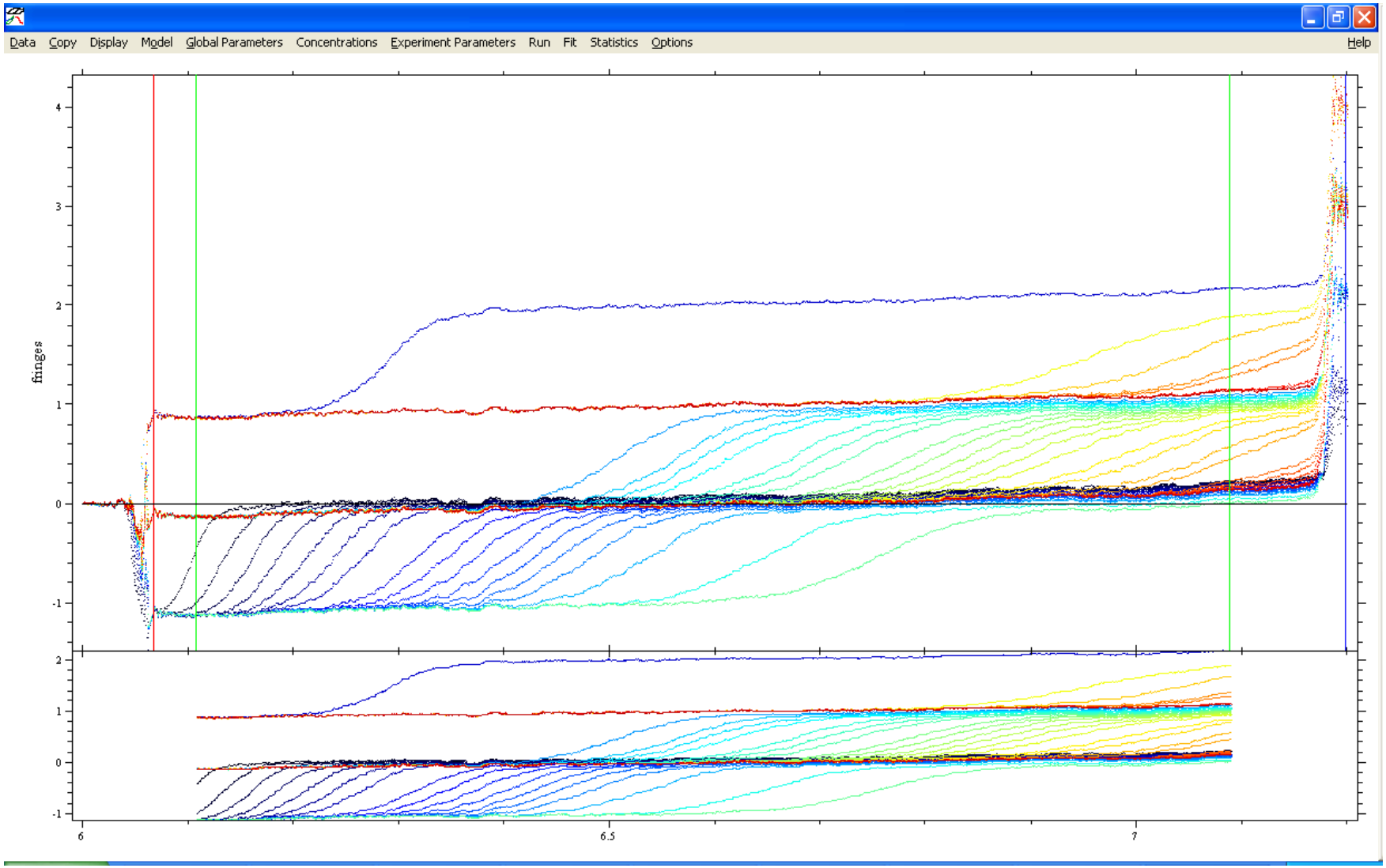
…and every file was loaded.

Scans 1-50 selected…

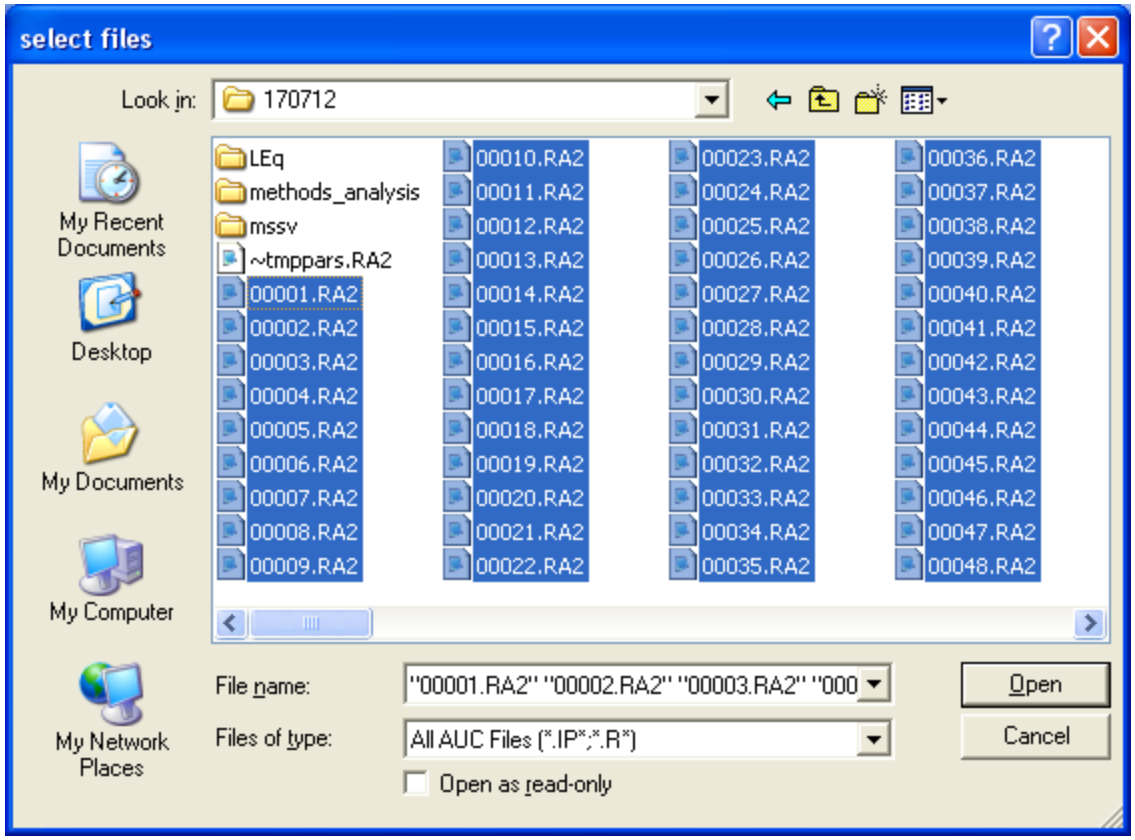
Set Experimental Parameters & Save



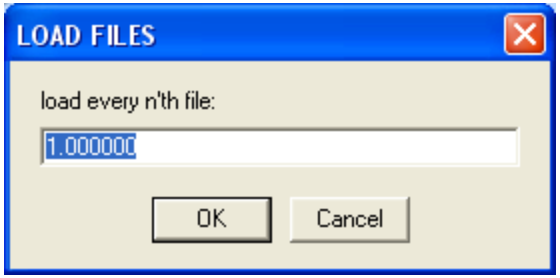
Set the Meniscus, Bottom, Fitting Limits



Load the ABS data

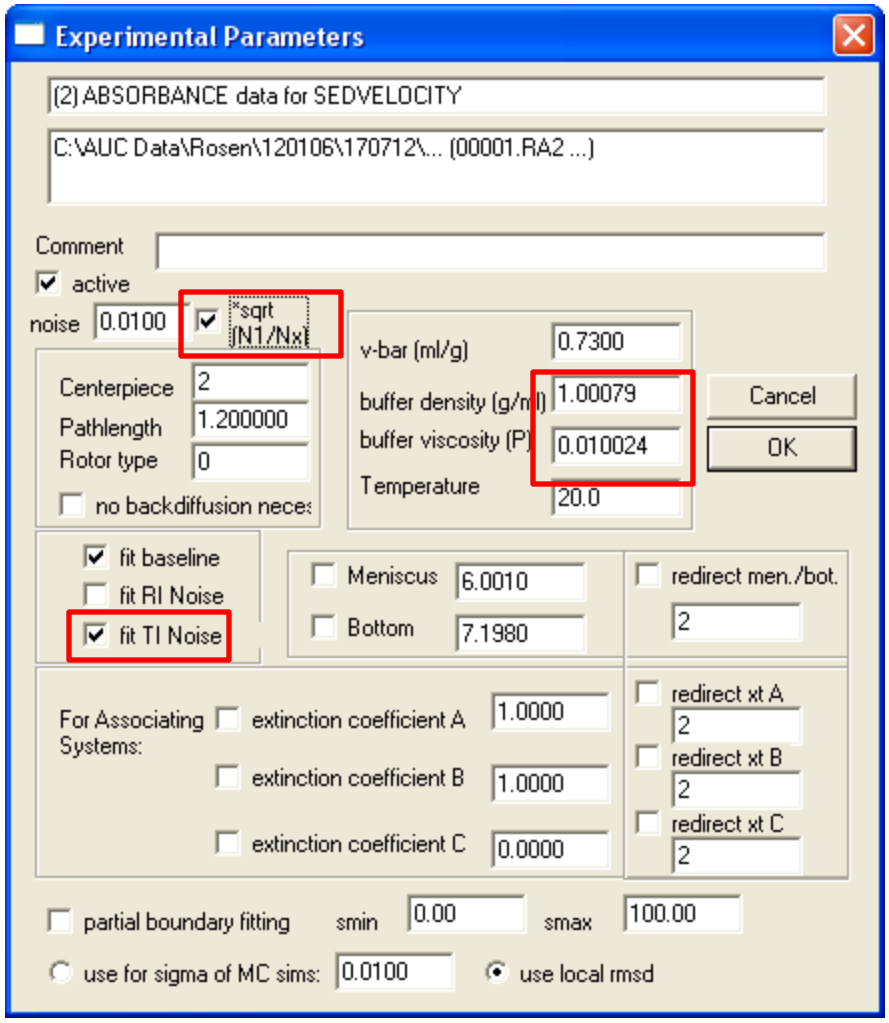


Files 1-50 were selected…

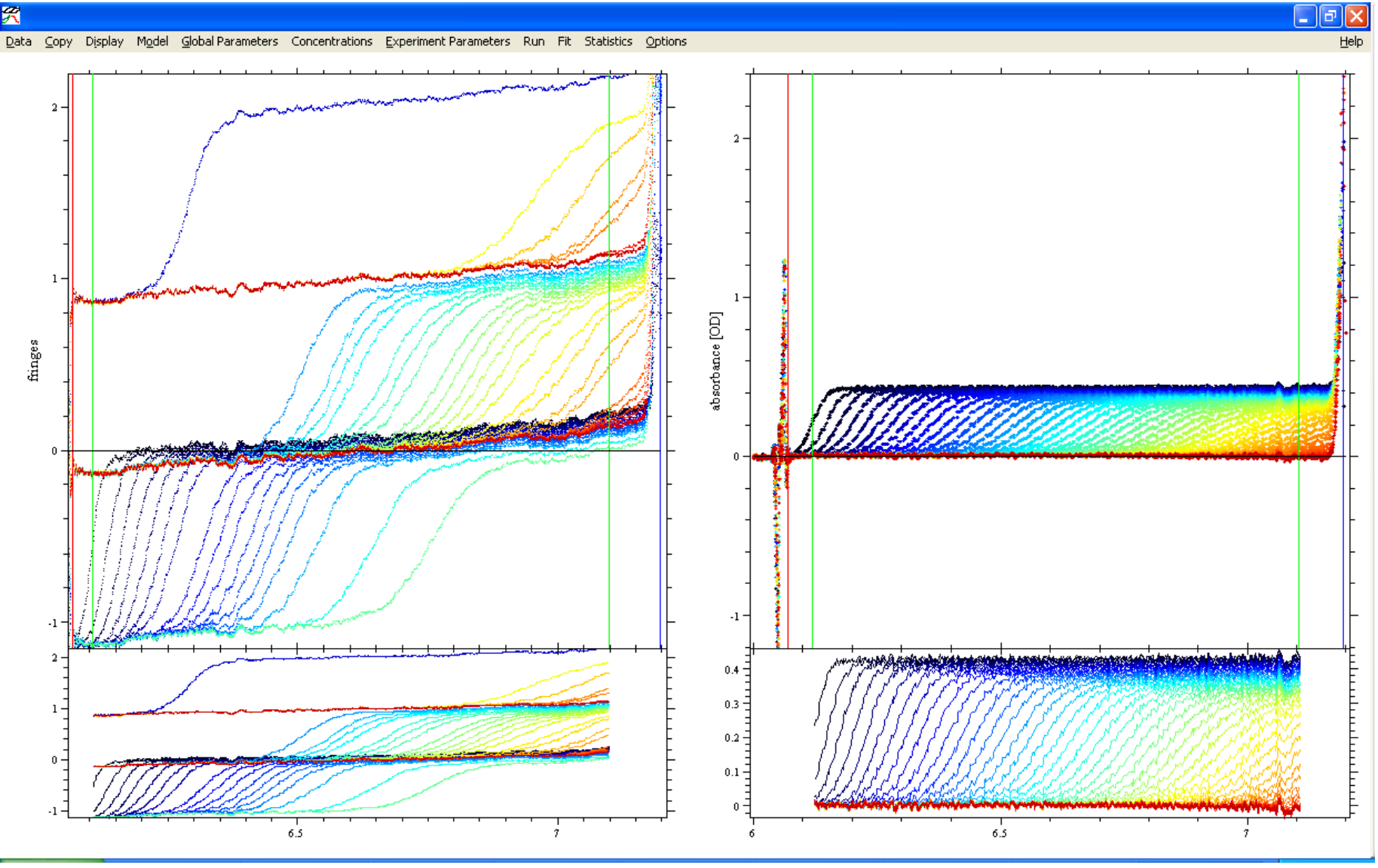


…and every file was loaded.

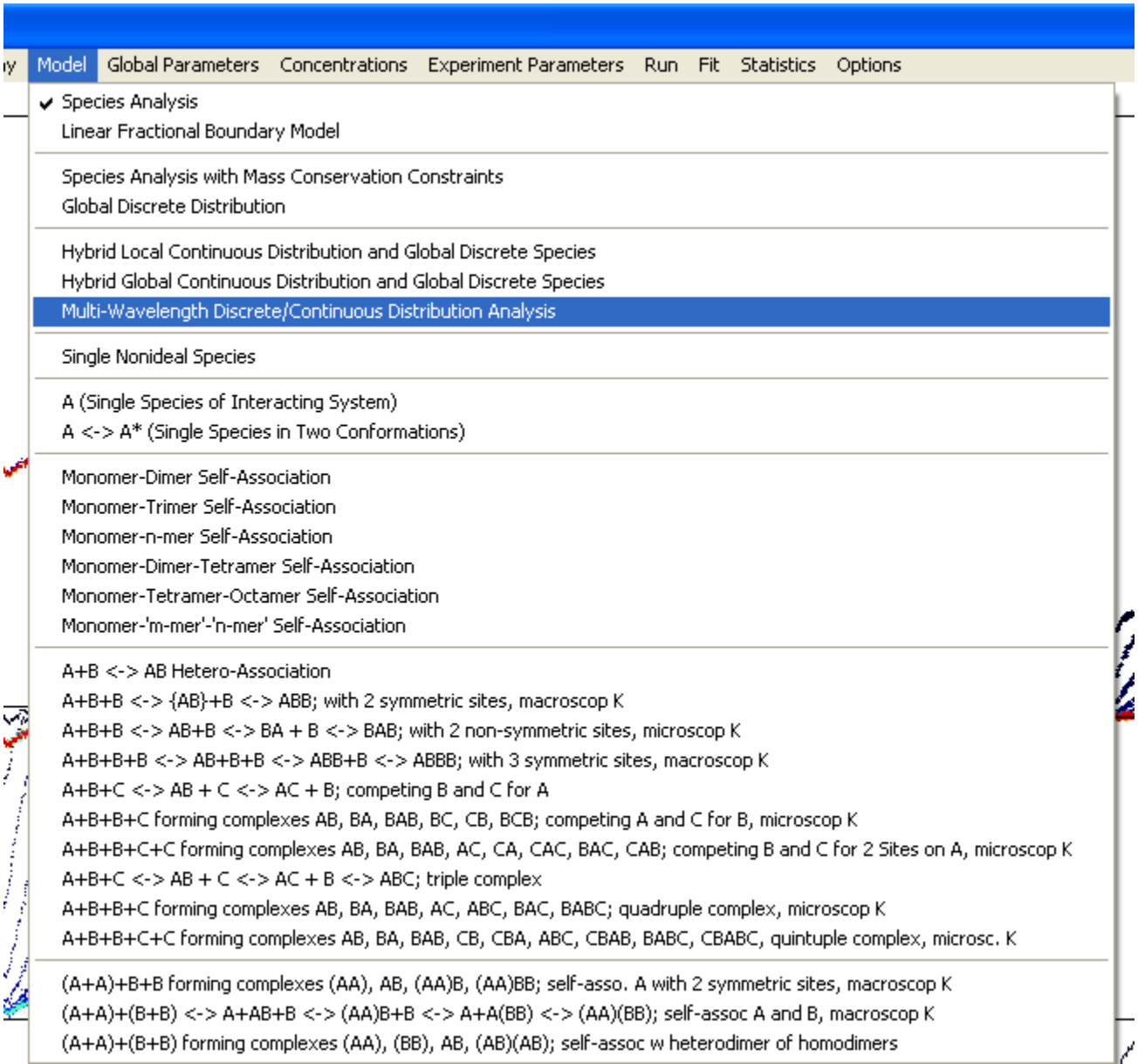
Set the ABS Experimental Parameters



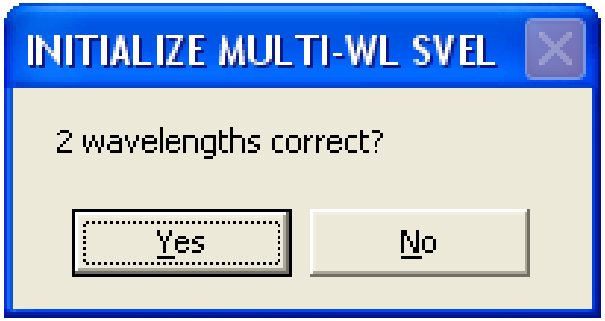
Set the Meniscus, Bottom, Fitting Limits for the ABS Data

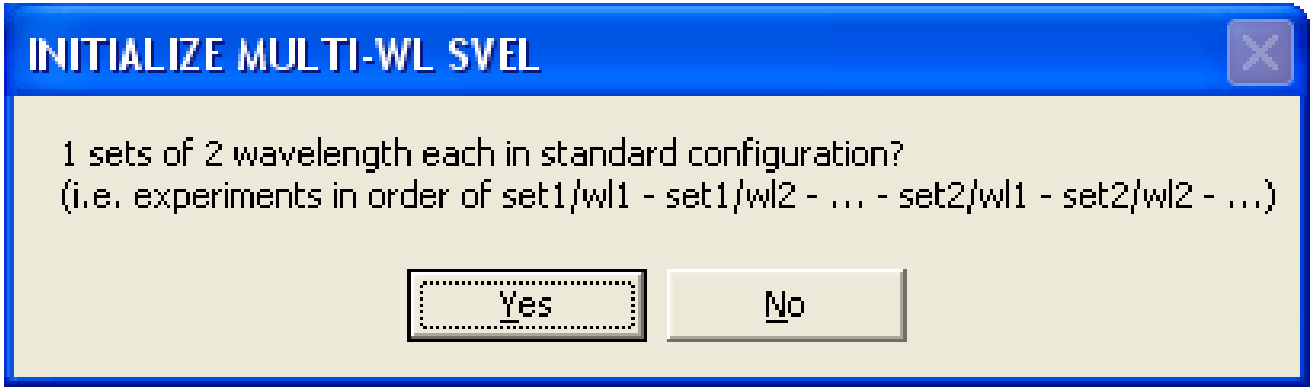


Select the Model



Answer Yes in Both Dialogs

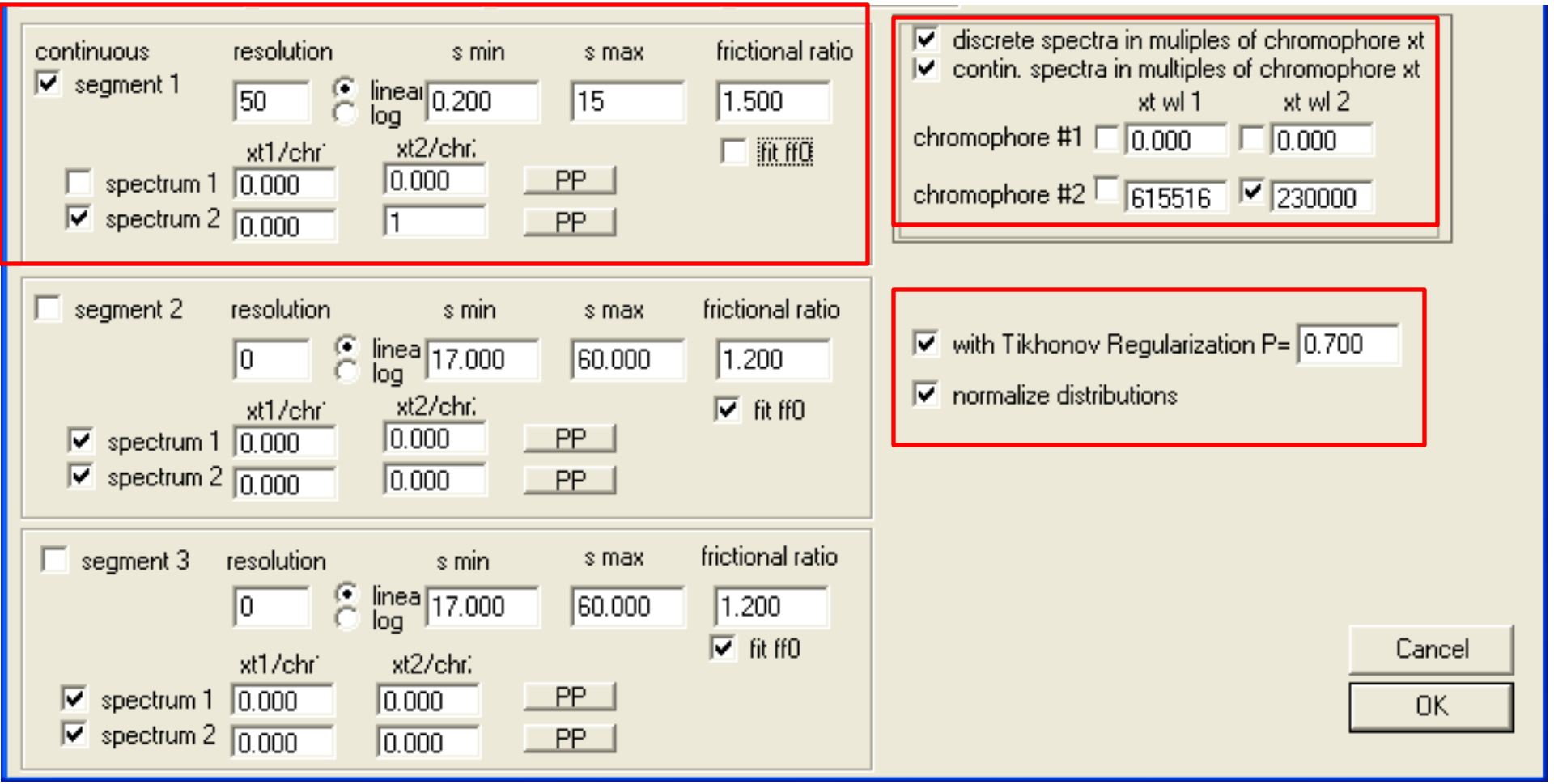




Step 18

See the next page for the values input into the Global Parameter box. Notably, the "chromophore 2" row was used, for consistency with what will follow. The $\varepsilon_{\mathrm{IF}}^{\mathrm{Arp2/3}}$ was fixed at 615,516 fringes·M^{-1}·cm^{-1}, a value based on the estimated molar mass of the protein complex. The $\varepsilon_{\mathrm{ABS280}}^{\mathrm{Arp2/3}}$ was set to 230,000 OD·M^{-1}·cm^{-1}, which was based on a calculation performed by SEDNTERP. This value was allowed to refine in the analysis; indeed, the purpose of this portion of the analysis is to obtain $\varepsilon_{\mathrm{ABS280}}^{\mathrm{Arp2/3}}$.

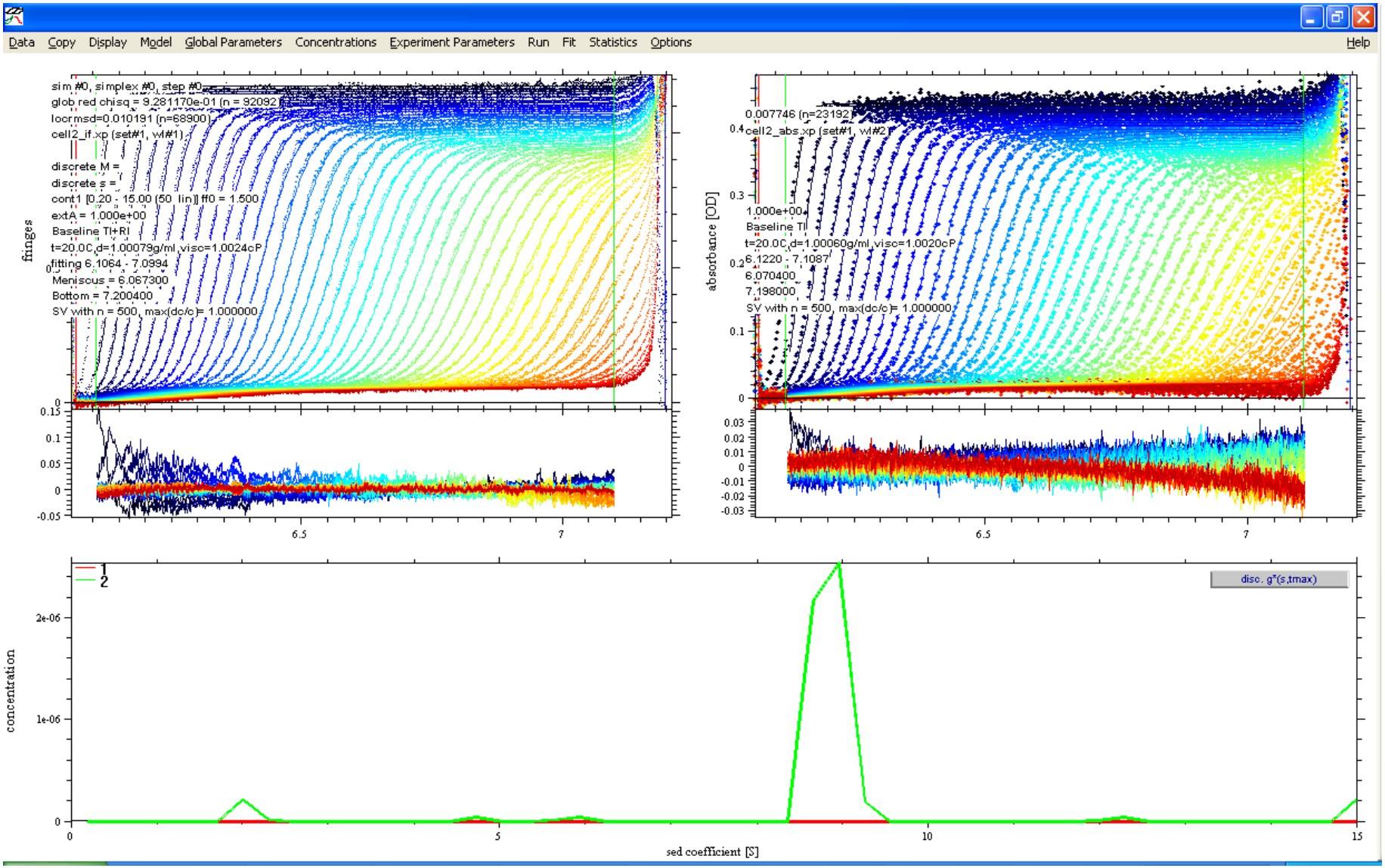
Setup the Global Parameters



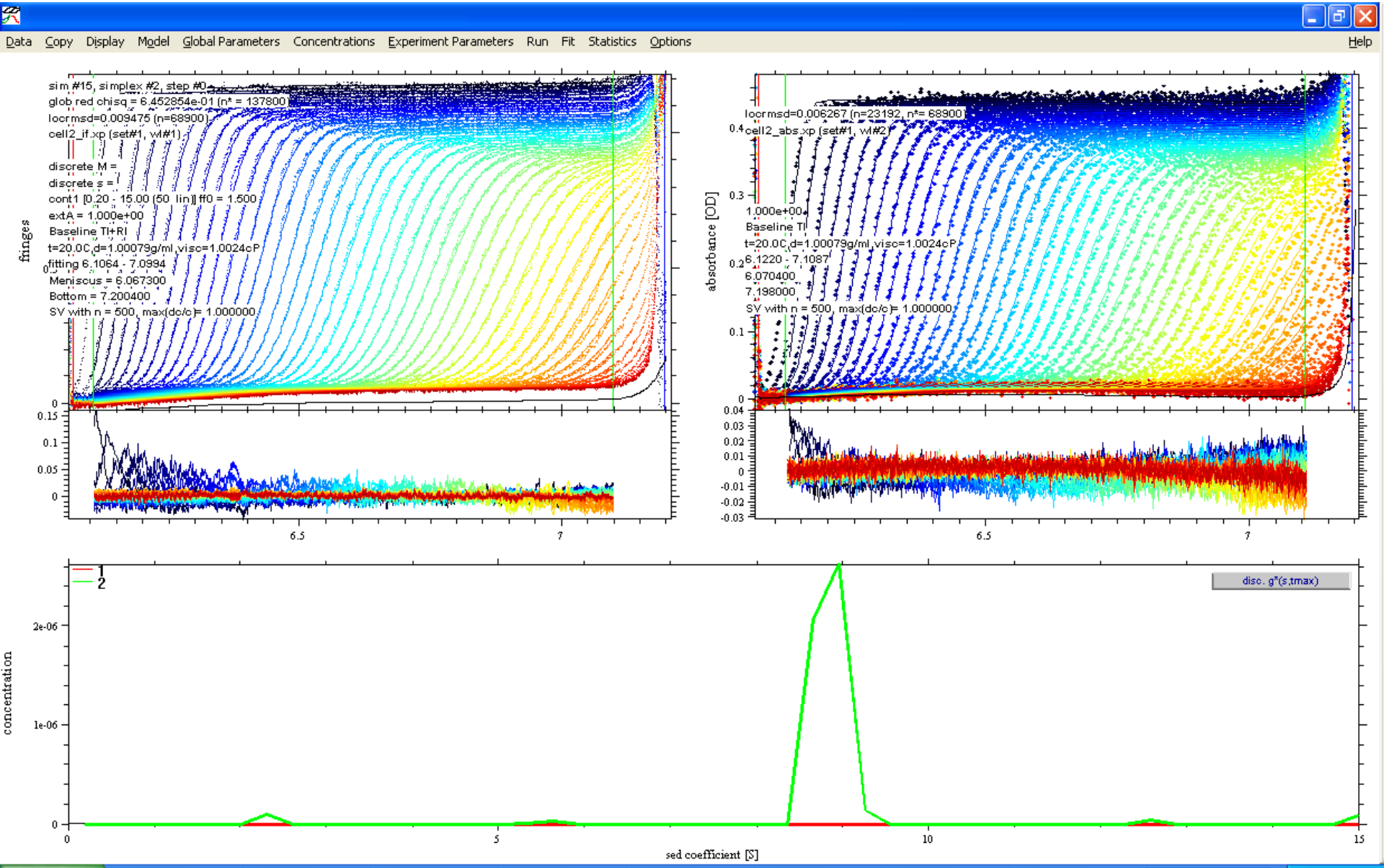
Step 19

A Global Run was performed. The parameters were deemed close enough to initiate a Global Fit, and that was done. After that, other nonlinear parameters were allowed to refine, as in Step 11. The final value of $\varepsilon_{\mathrm{ABS280}}^{\mathrm{Arp2/3}}$ was 244,420 $M^{-1} \cdot cm^{-1}$, and the f_r of Arp2/3 was refined to 1.618. The fit was saved. The [Arp2/3] was determined as in step 15. It was 1.49 μM. These substeps are detailed on the next twelve pages.

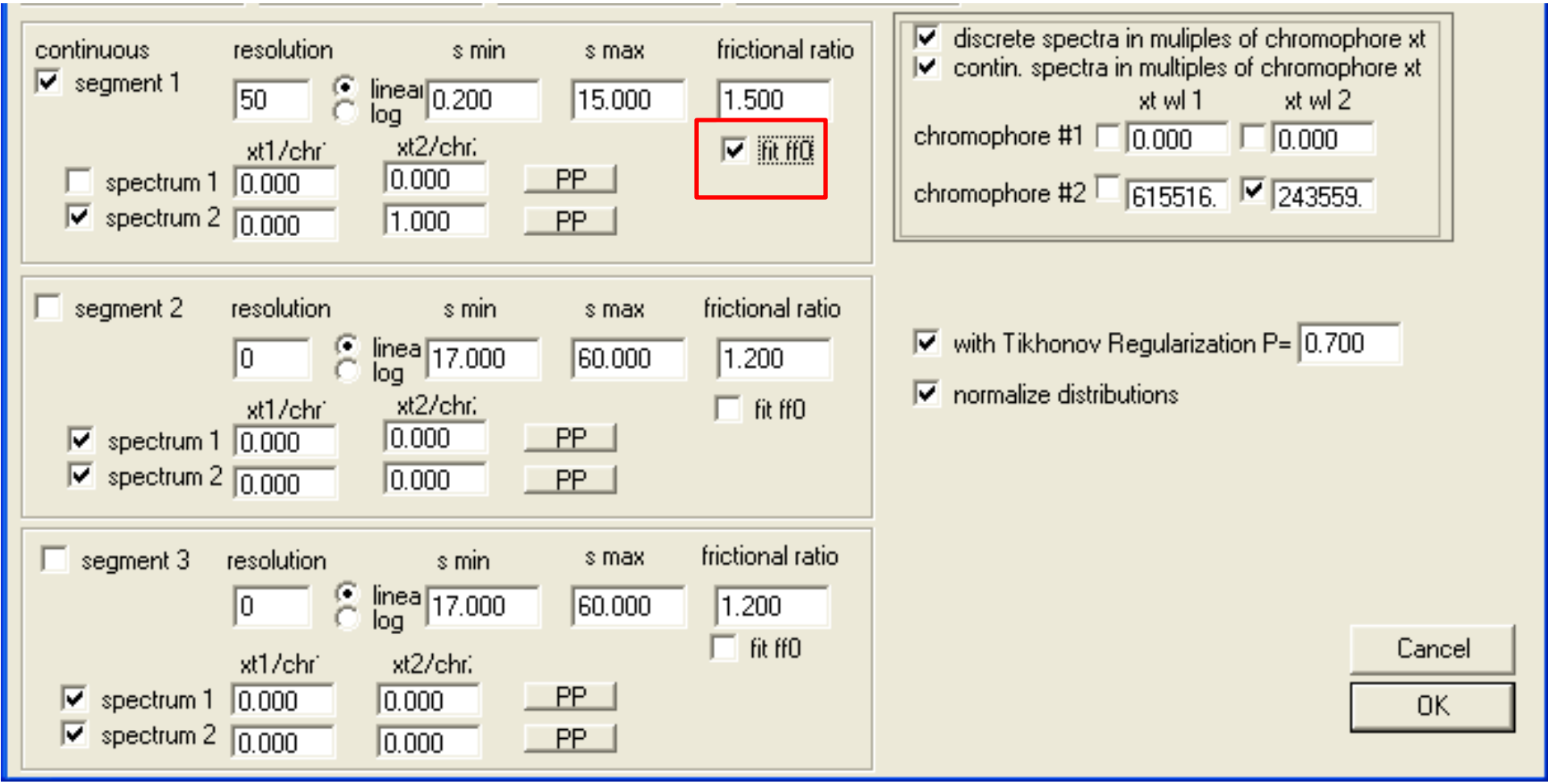
After Global Run



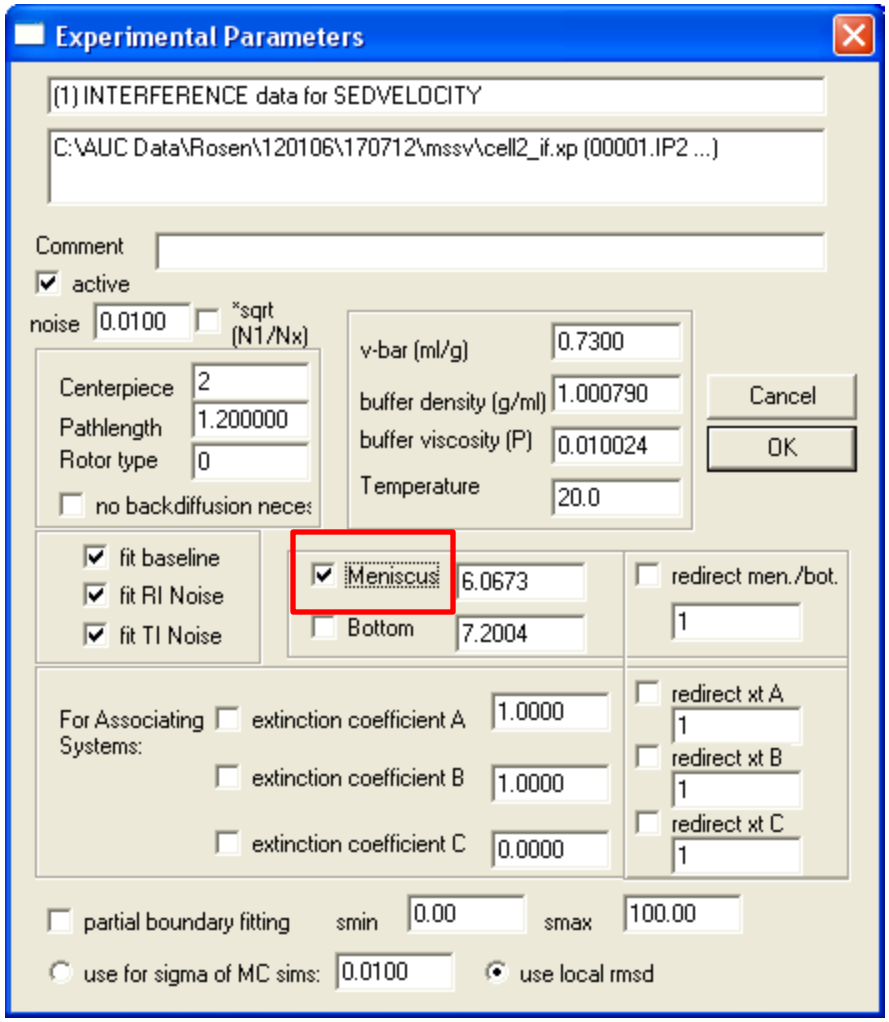
After Global Fit

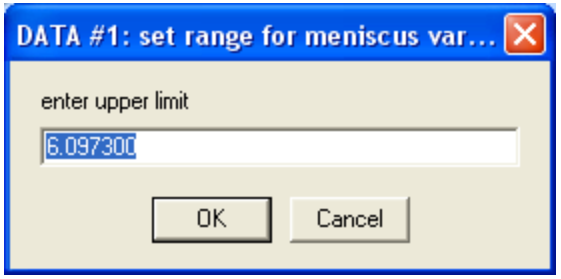


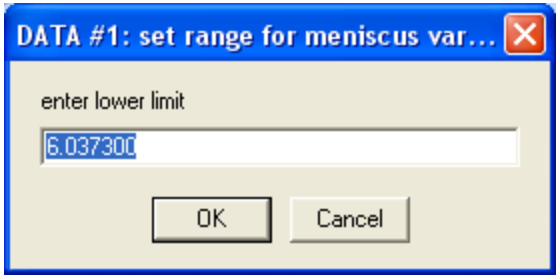
Fit ff0 Now


continuous
segment 1
resolution
50
linear
log
s min
0.200
s max
15.000
frictional ratio
1.500
fit ff0
xt1/chr
xt2/chr.
spectrum 1
0.000
0.000
PP
spectrum 2
0.000
1.000
PP
segment 2
resolution
0
linear
log
s min
17.000
s max
60.000
frictional ratio
1.200
fit ff0
xt1/chr
xt2/chr.
spectrum 1
0.000
0.000
PP
spectrum 2
0.000
0.000
PP
segment 3
resolution
0
linear
log
s min
17.000
s max
60.000
frictional ratio
1.200
fit ff0
xt1/chr
xt2/chr.
spectrum 1
0.000
0.000
PP
spectrum 2
0.000
0.000
PP
discrete spectra in muliples of chromophore xt
contin. spectra in multiples of chromophore xt
xt wl 1
xt wl 2
chromophore #1
0.000
0.000
chromophore #2
615516.
243559.
with Tikhonov Regularization P=
0.700
normalize distributions
Cancel
OK

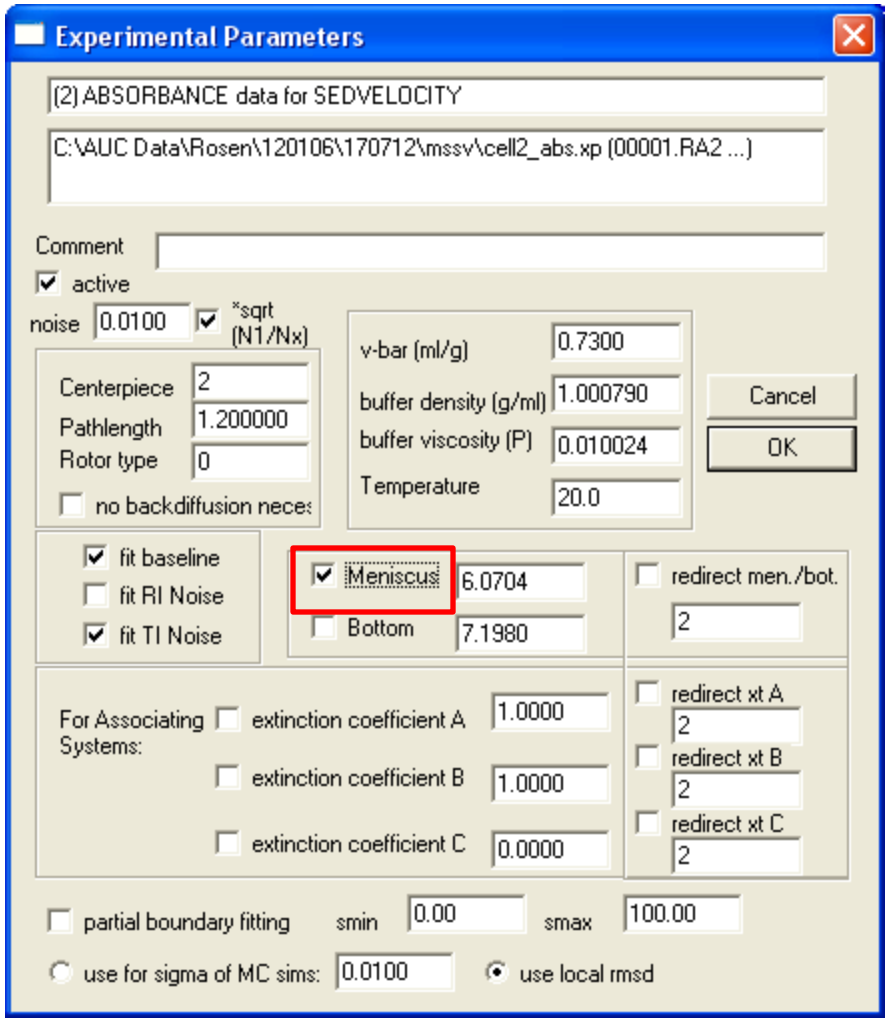
After clicking “Experiment Parameters”, Fit Meniscus #1

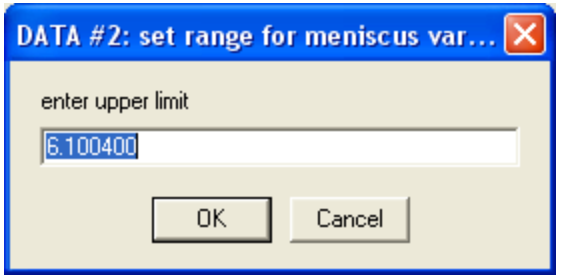


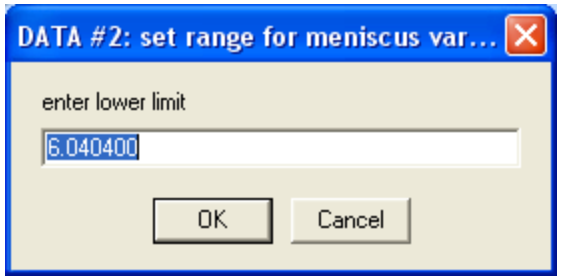




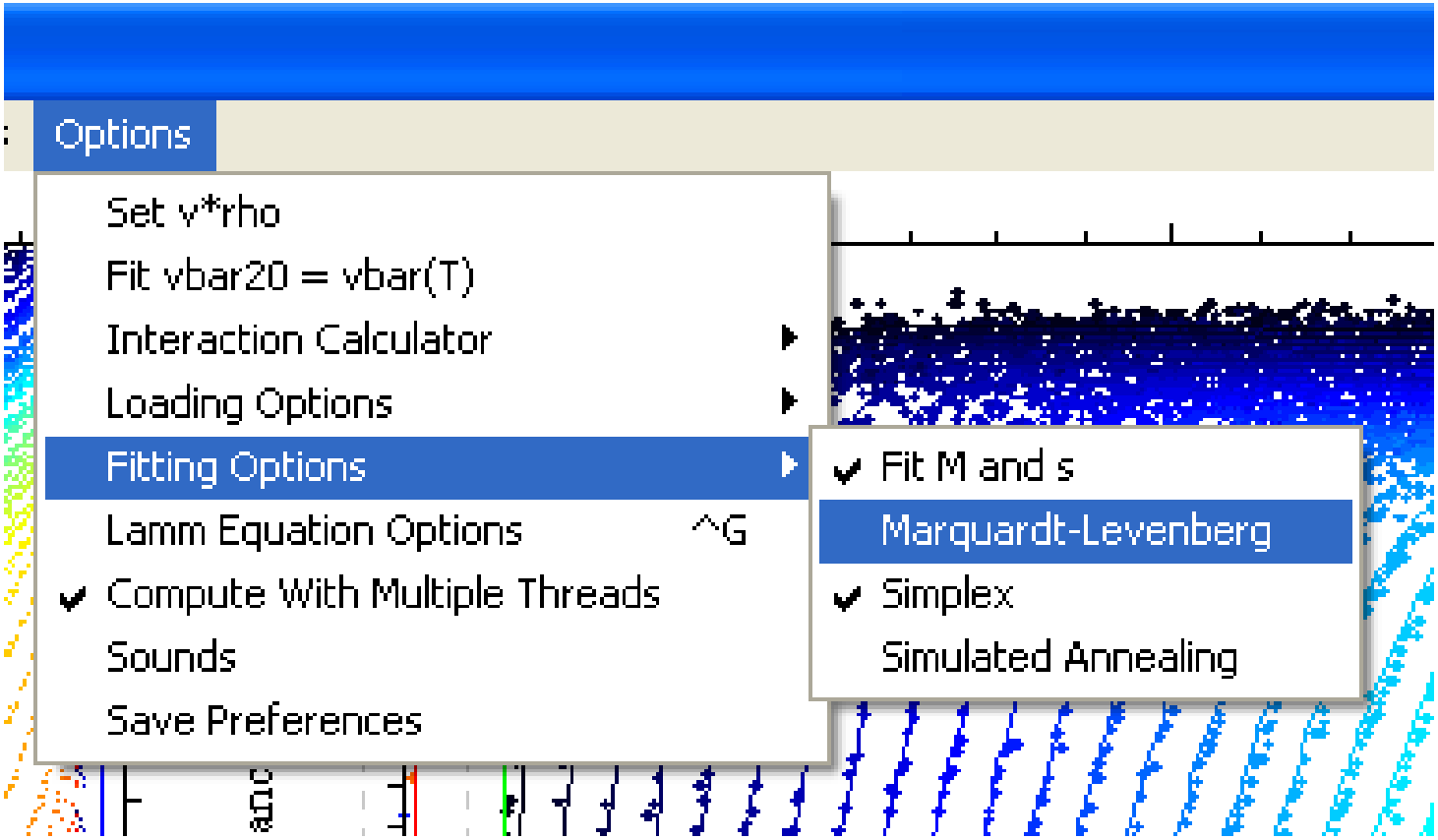
Fit Meniscus #2



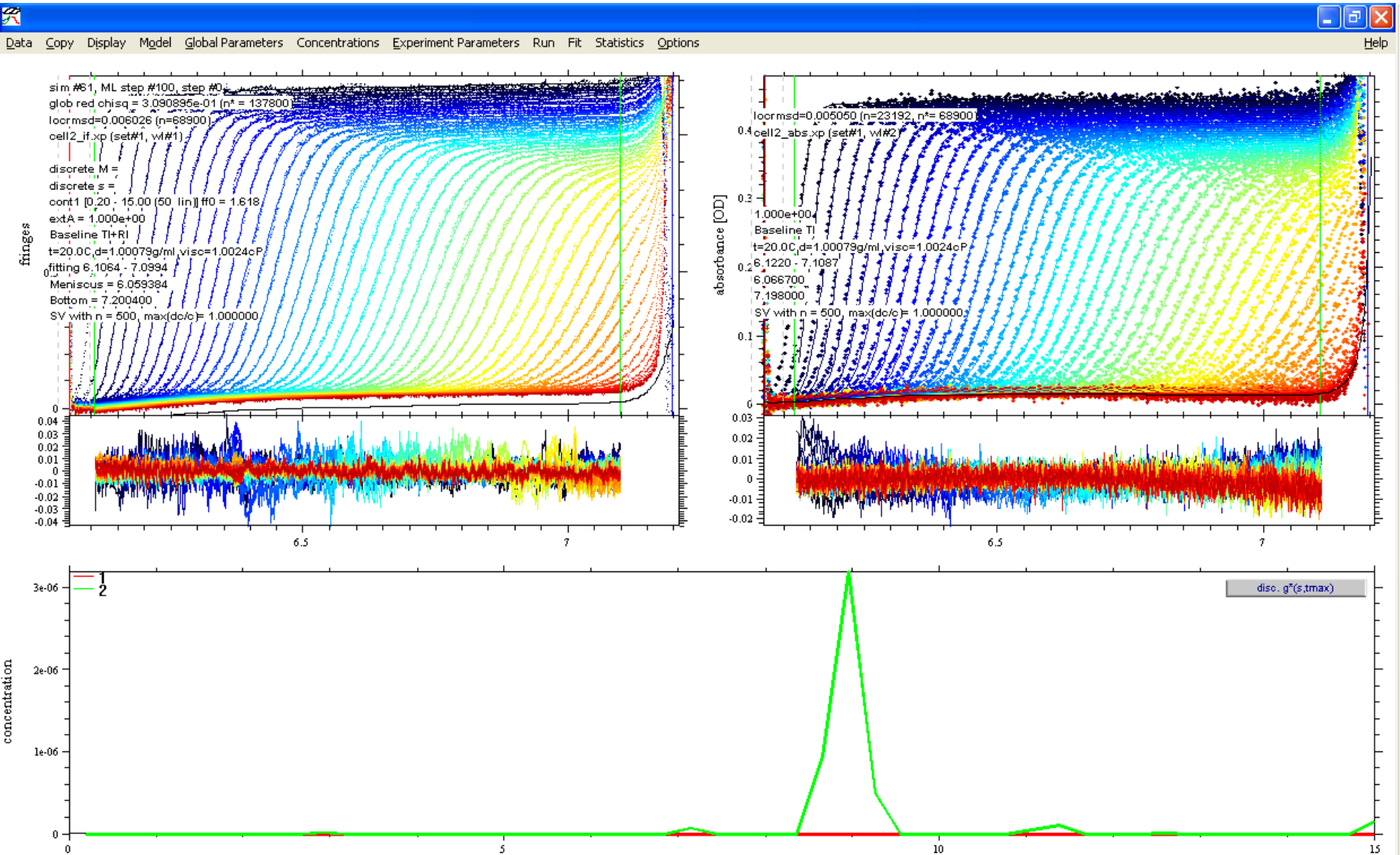




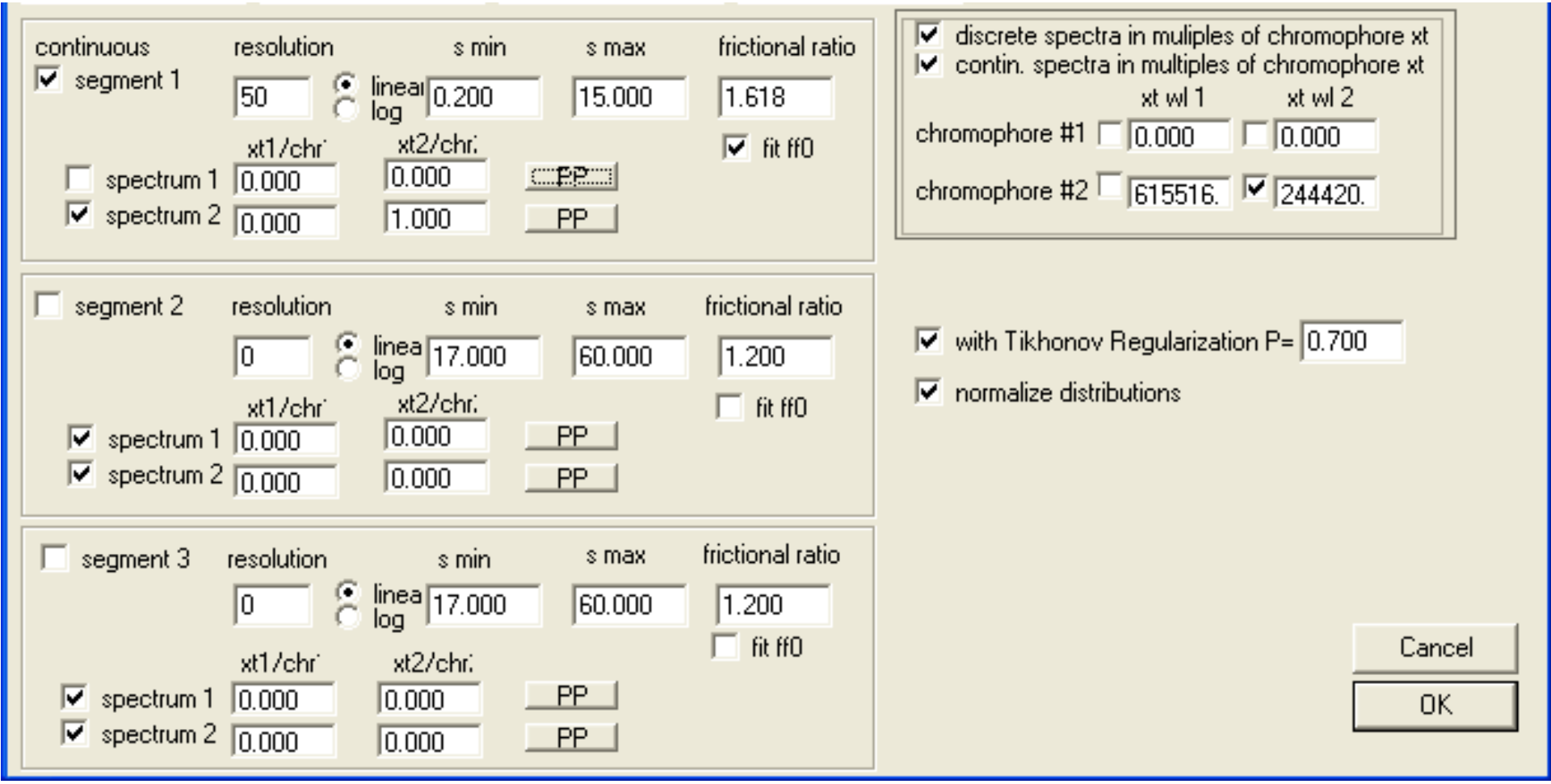
Change Fitting Algorithm to Marquardt Levenberg



After Global Fitting

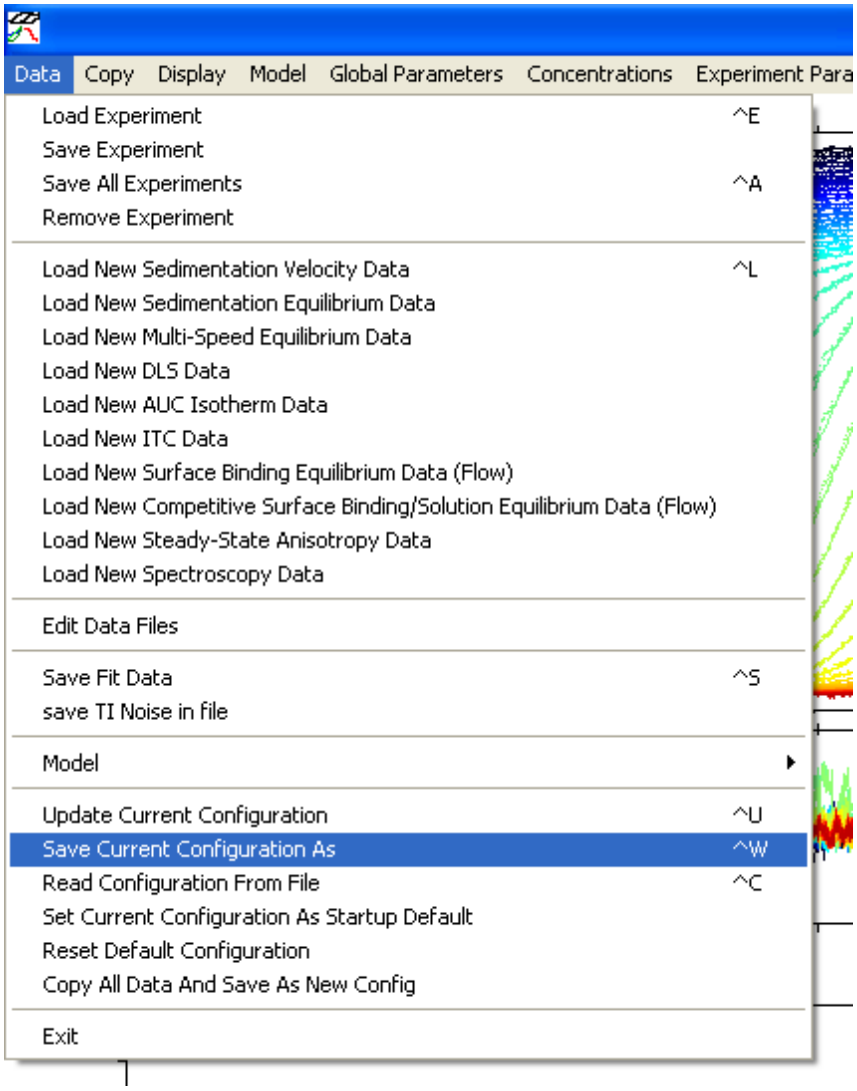


Resultant Parameters

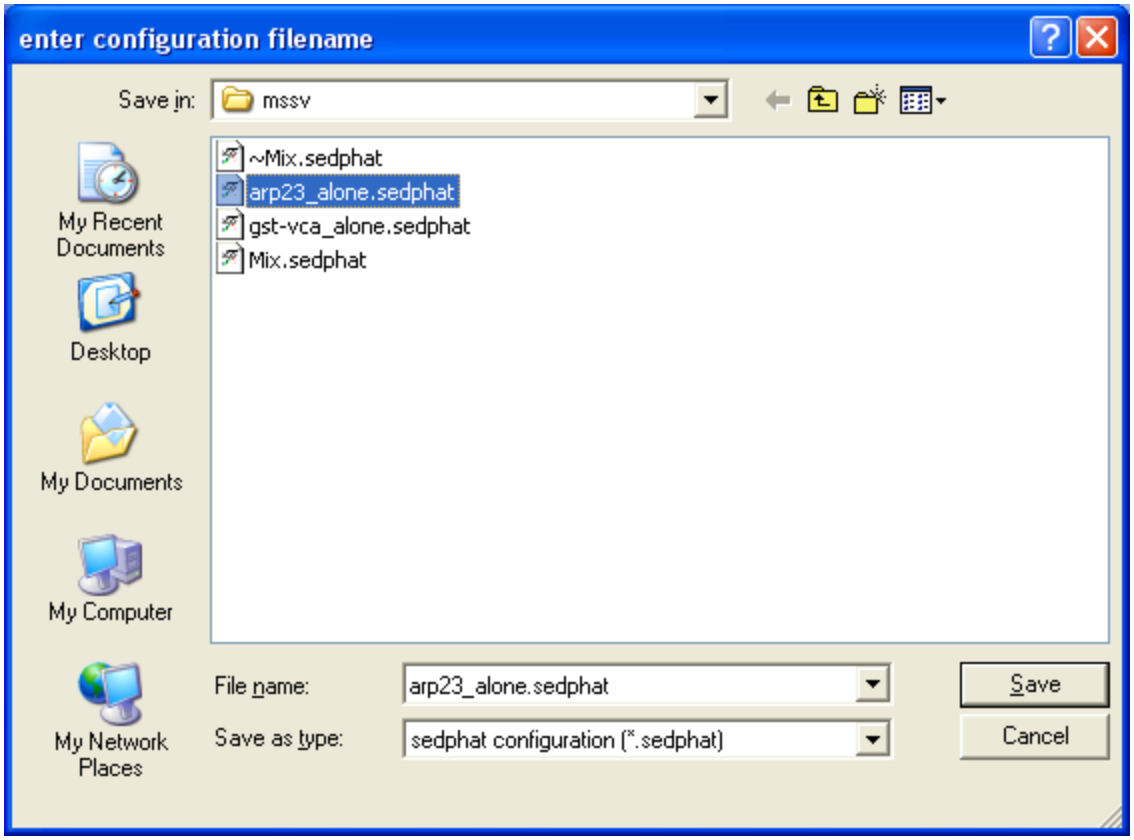


It is a good idea at this point to note down the refined $\varepsilon_{\mathrm{ABS280}}^{\mathrm{Arp2/3}}$ so that it can be easily recalled for input later (Step 21).

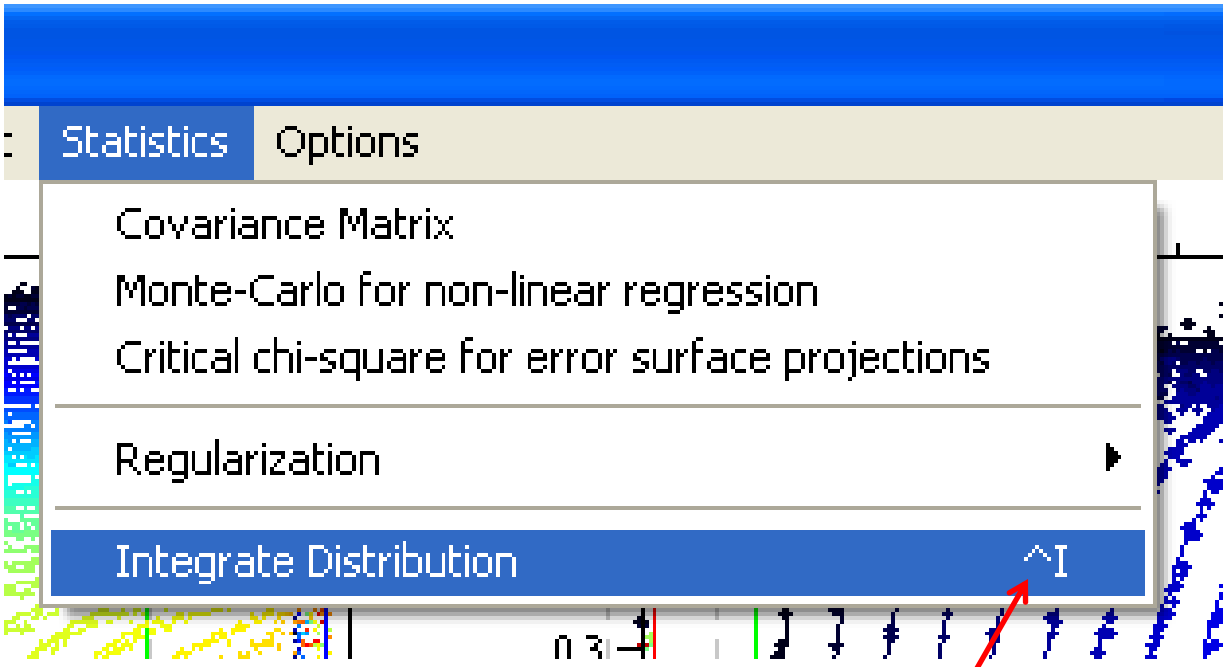
Save the Configuration



Save Configuration Dialogs

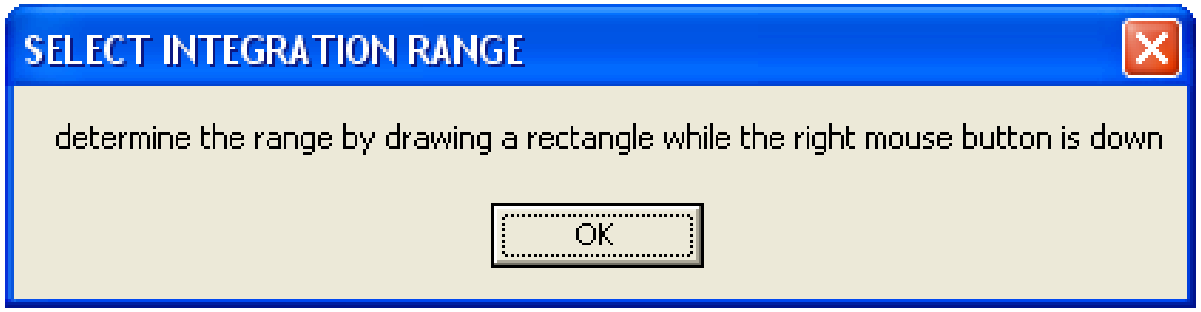


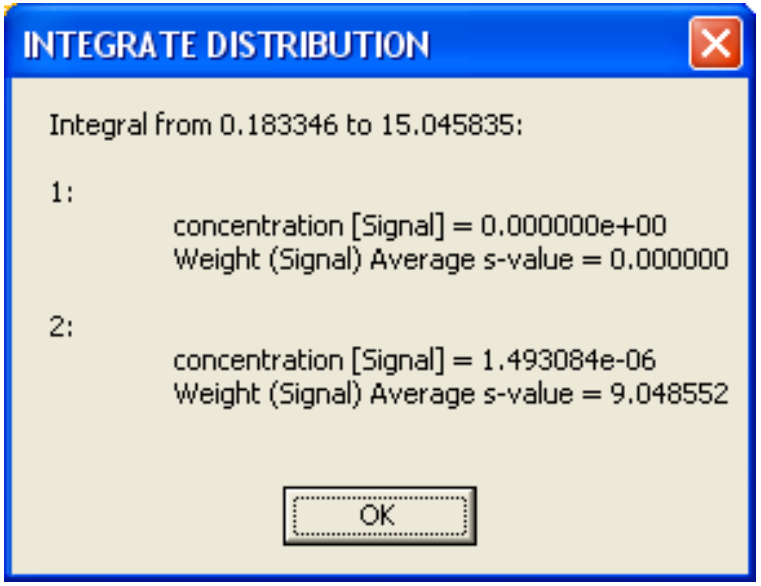
Integrate the distribution



Note that SEDPHAT tells you if a keyboard shortcut is available. Here, instead of choosing the menu item, I could have pressed “Ctrl-I.”

Integration Results

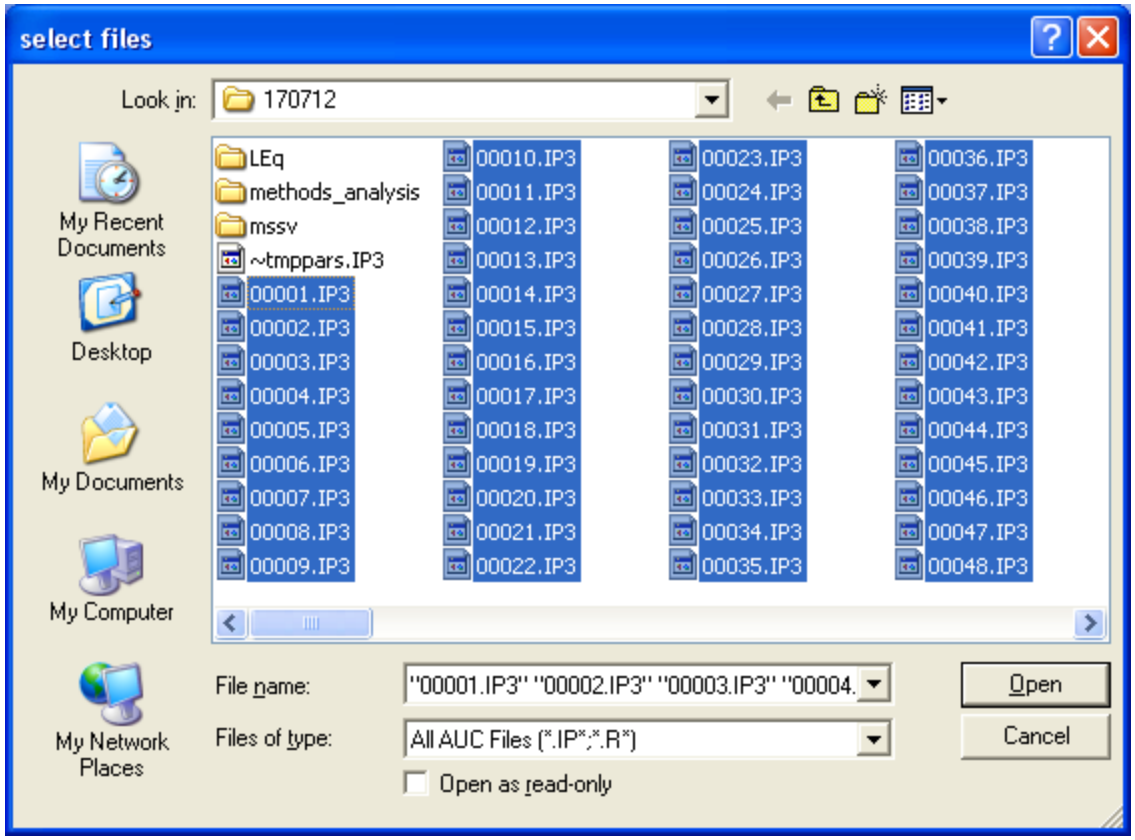


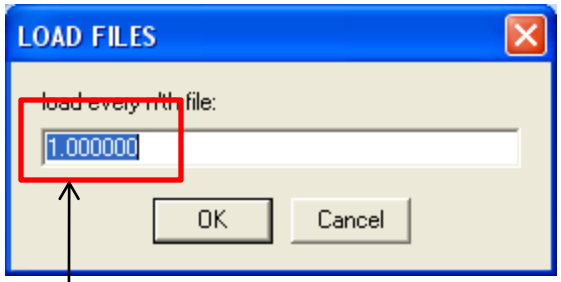


Step 20

The next task was to analyze the mixture. The data from the mixture were acquired from Cell 3. Based on the analyses above and the volumes of the protein stocks pipetted into the mixture, roughly 3.0 μM GST-VCA and 0.4 μM Arp2/3 were expected to be detected. The data were loaded into a new SEDPHAT session as detailed in Steps 1-6 above. These substeps are pictorially described on the next eight pages.

Start a New SEDPHAT Session. Load the Mixture IF data.



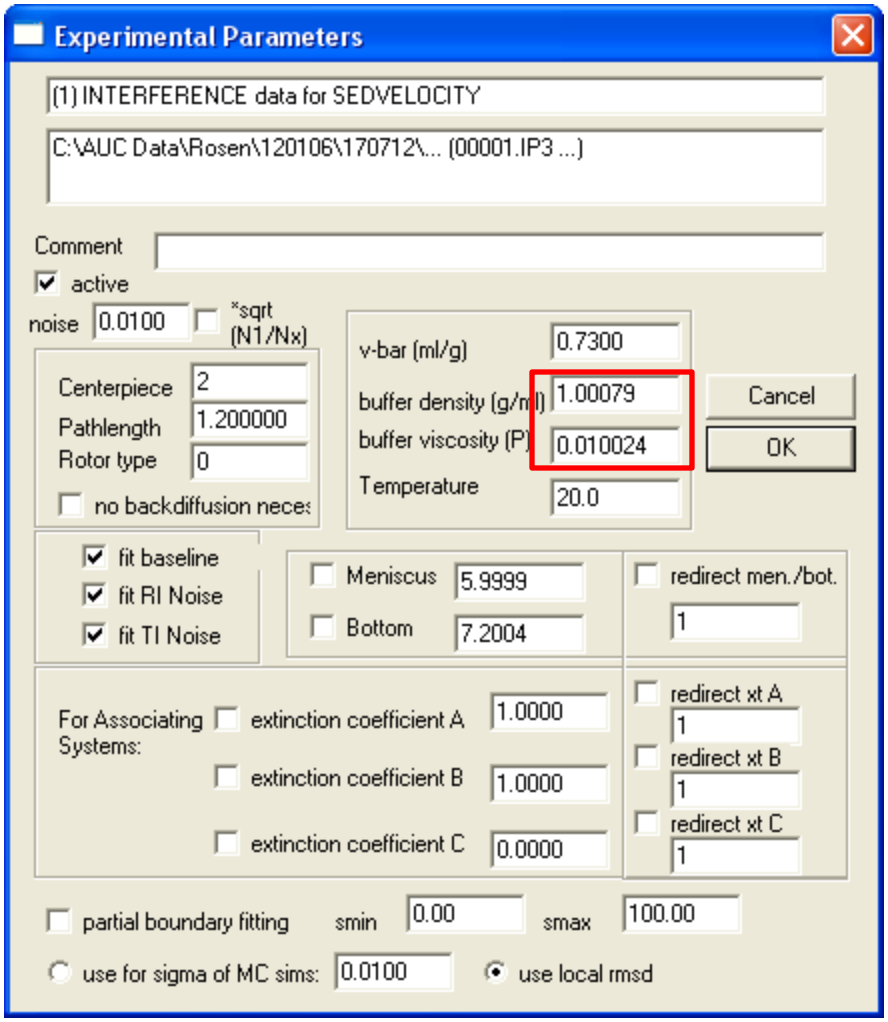


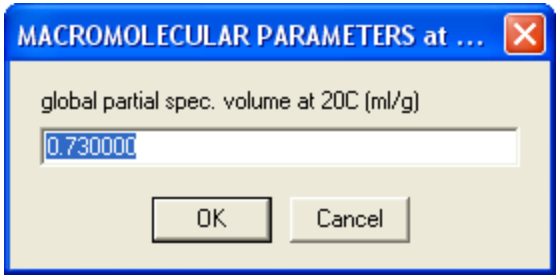
Important!

…and <u>every</u> file is loaded.

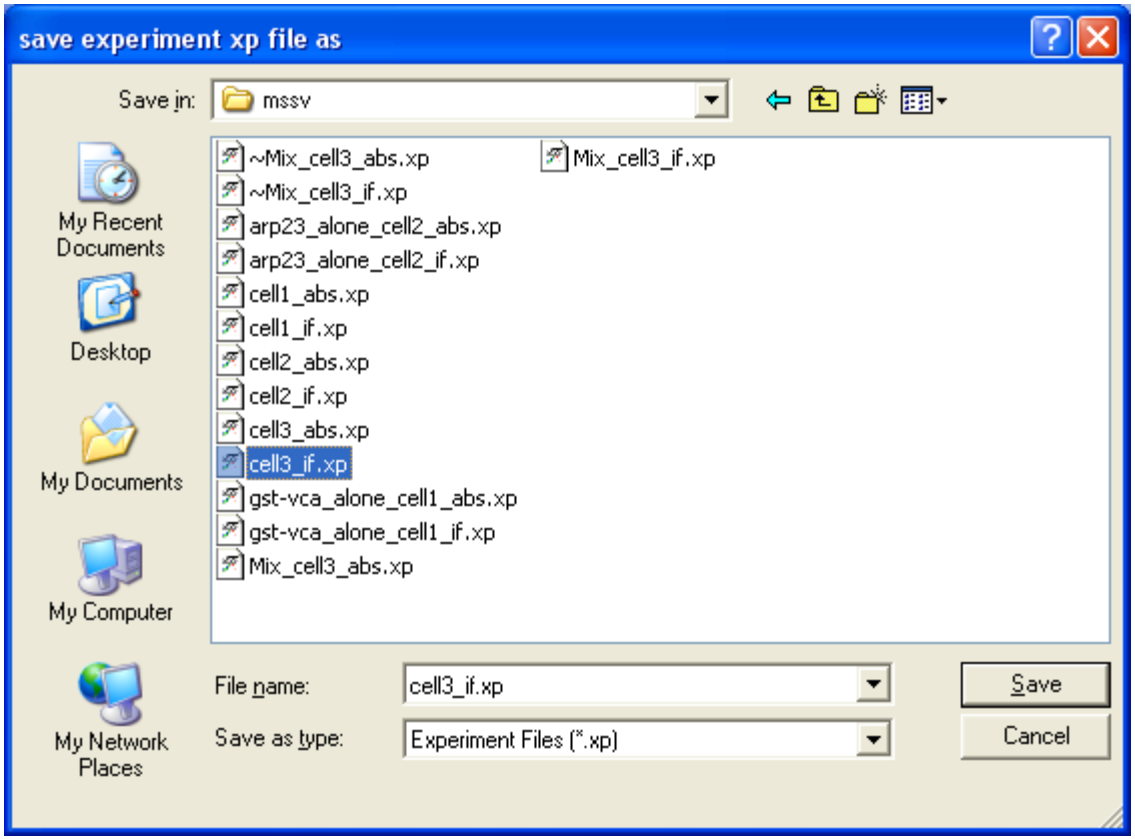
Files 1-101 from Cell 3 are selected…

Input the Experimental Parameters for the Interference Data

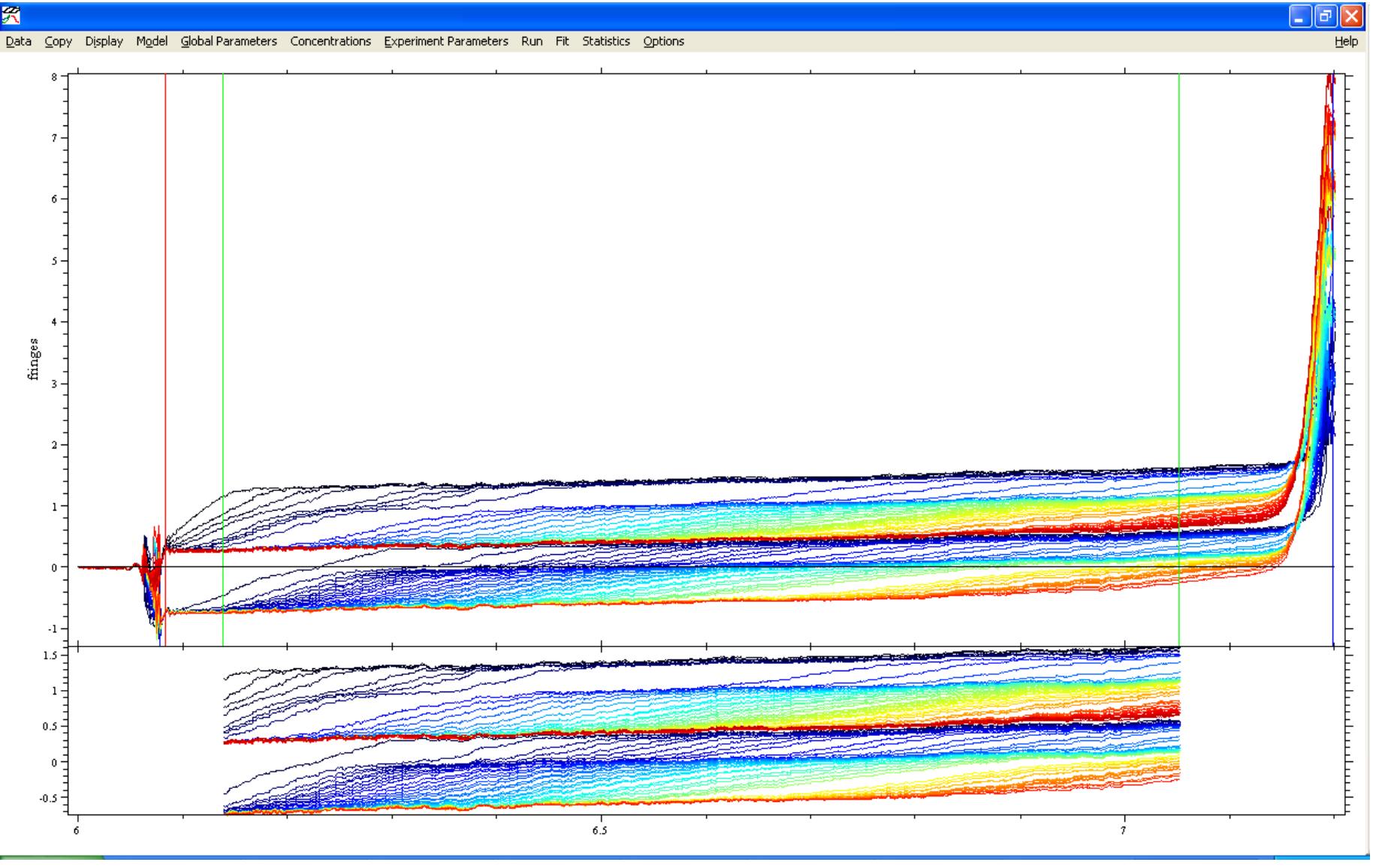




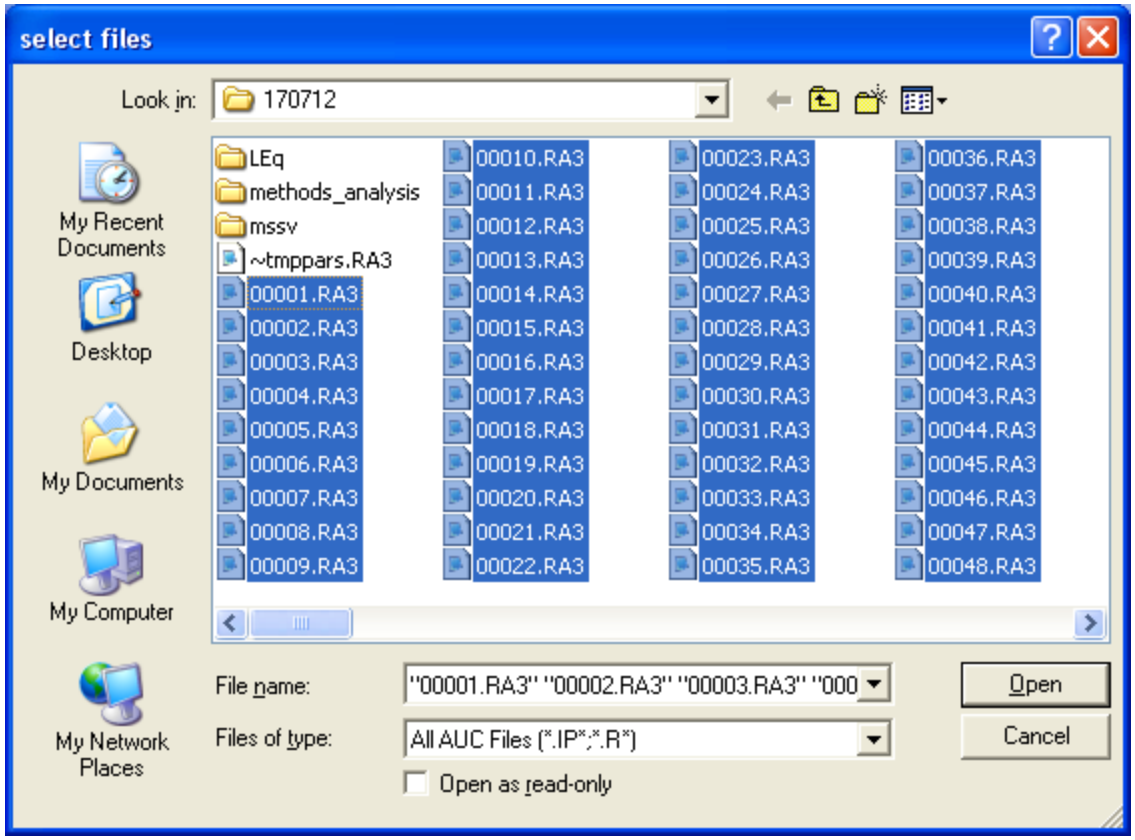
Save the Experiment



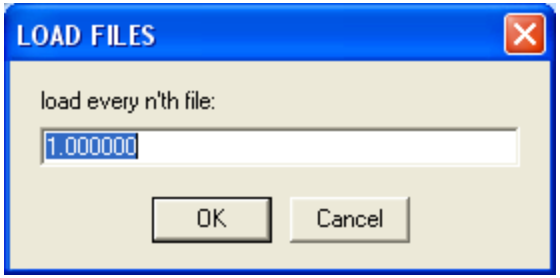
Set the Meniscus, Bottom, Fitting Limits



Load ABS280 for the Mixture

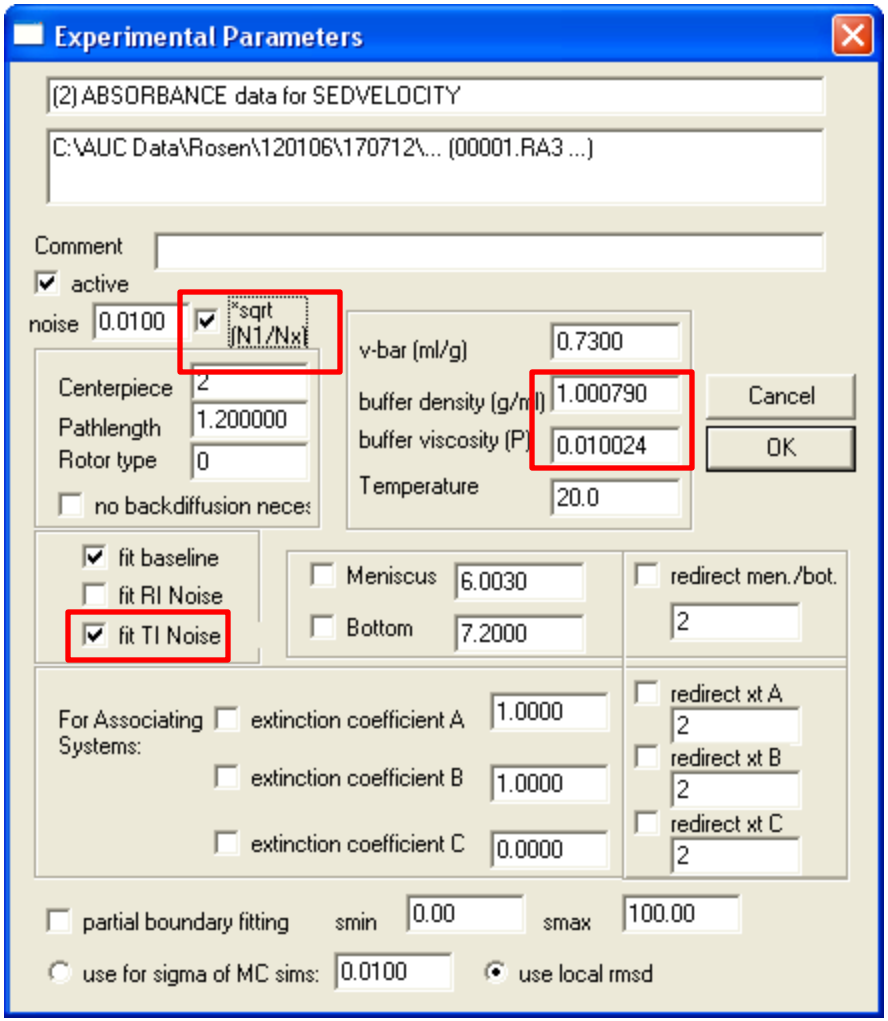


Files 1-101 from Cell 3 are selected...

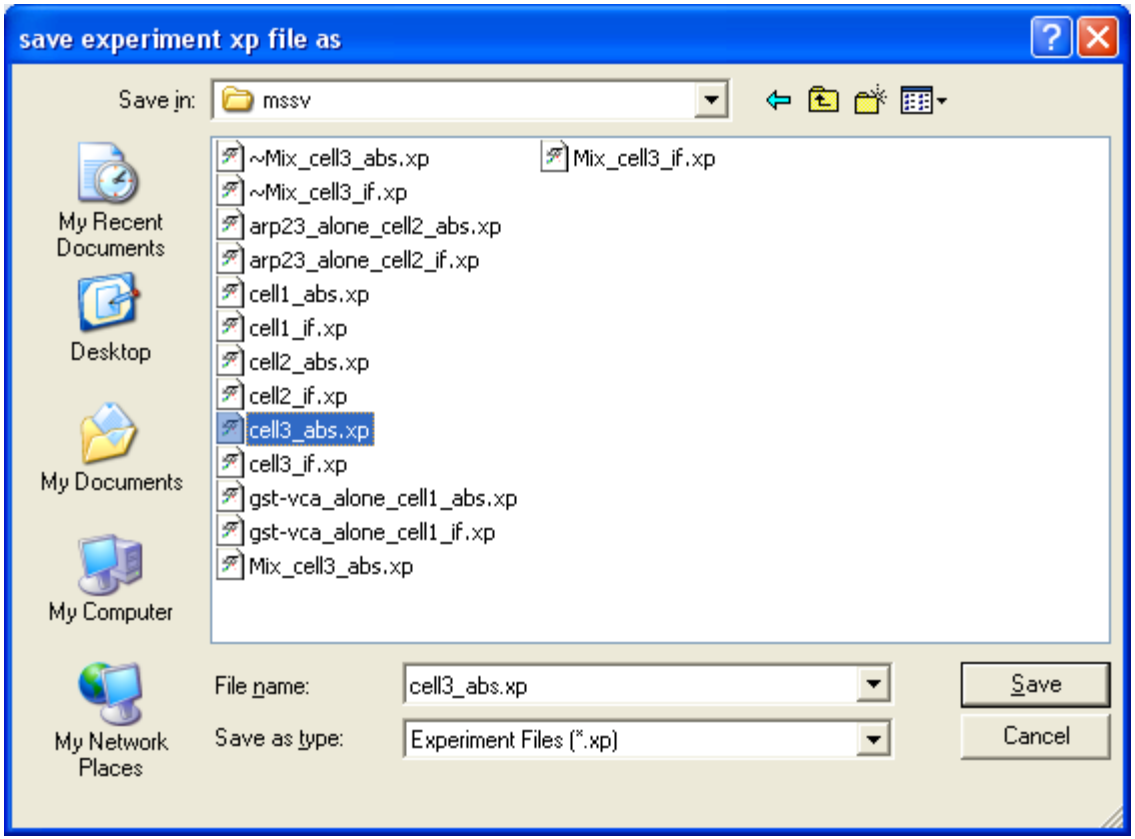


...and <u>every</u> file is loaded.

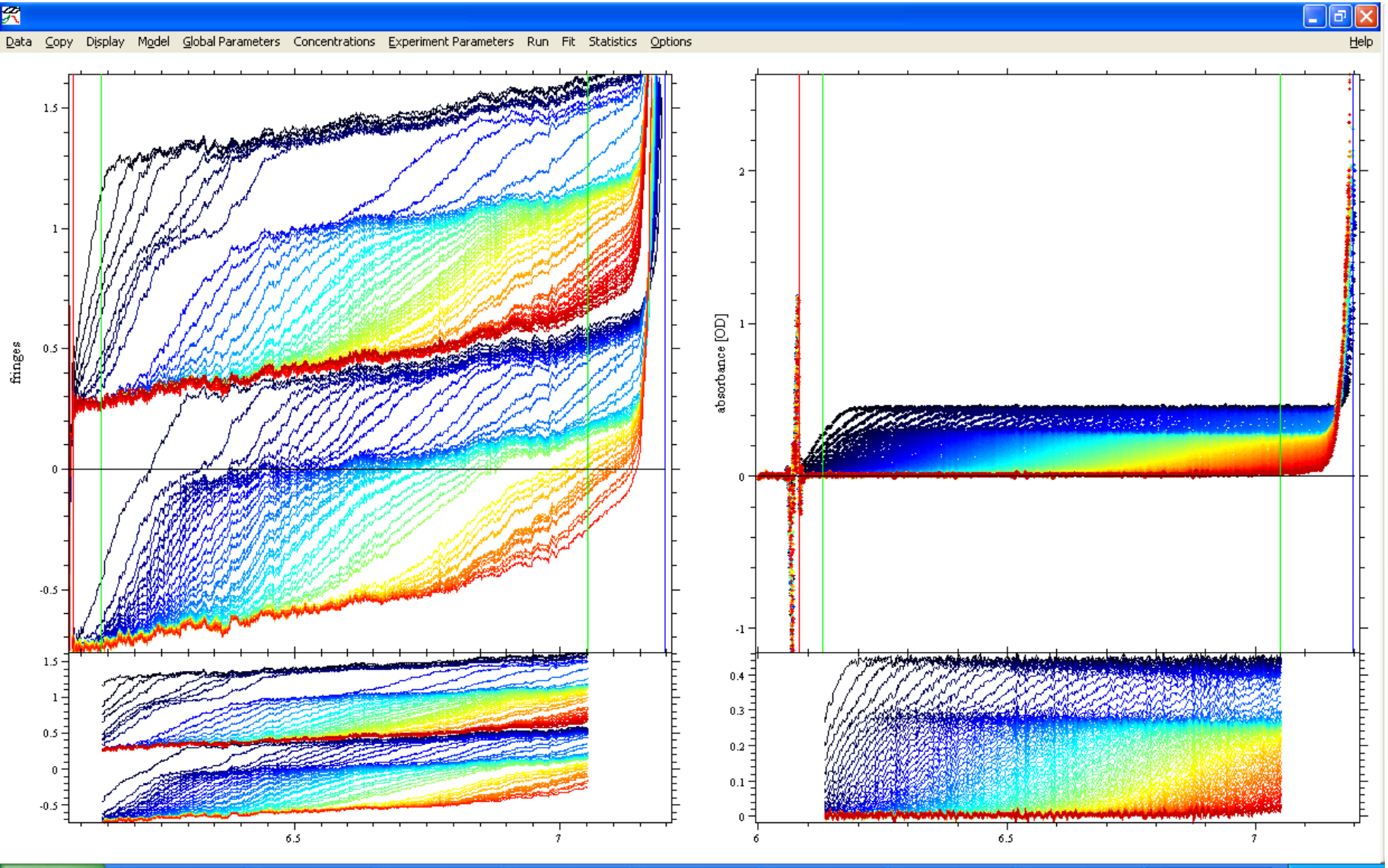
Change the Appropriate Experimental Parameters



Save the Experiment

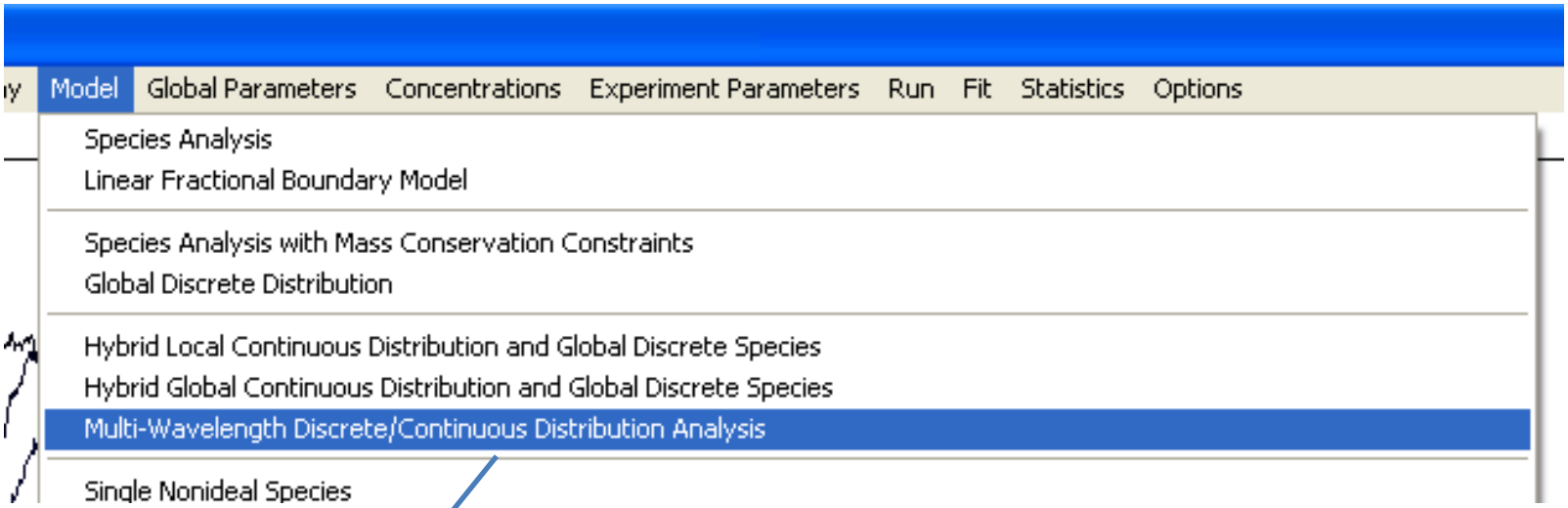


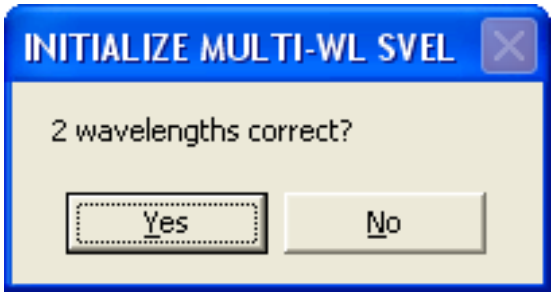
Set the Meniscus, Bottom, Fitting Limits



Step 21

The Multi-Wavelength model was chosen, as before.

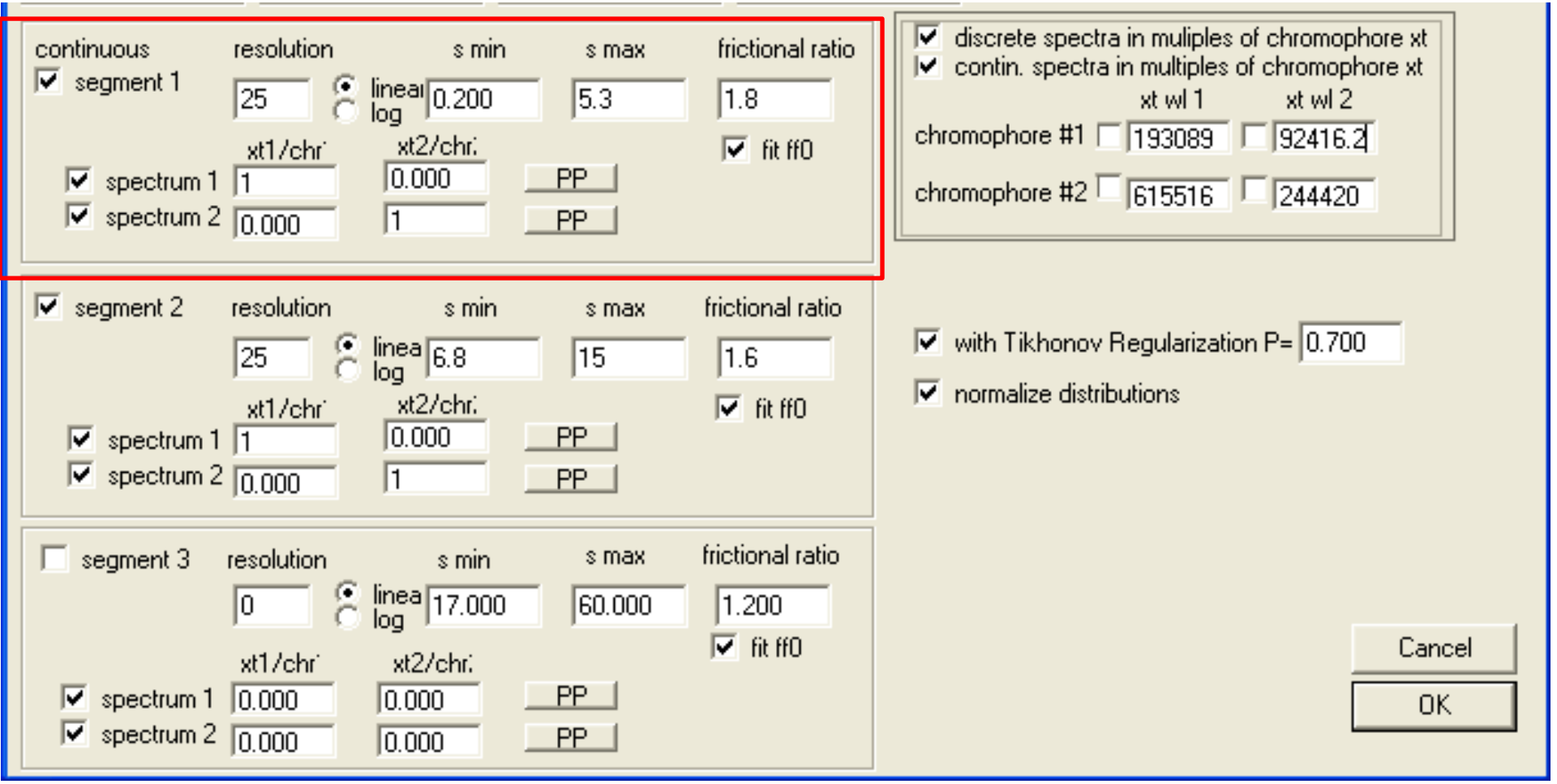




Step 21, continued

The Global Parameters Box was opened. Again, the upper part of the parameter box, which pertains to discrete species, was not used. In this case, two "segments" were activated by checking their respective checkboxes. The reason for this expedient is twofold. First, the f_r of free GST-VCA under these conditions was expected to be substantially different from that of the GST-VCA:Arp2/3 complex. The use of two segments allows two different f_r's to be refined during the analysis. Further, the use of separate segments to describe the two boundaries present in the data allows for stoichiometric constraints to be placed on the data for certain *s*-ranges (see Step 27). Segment 1 was given a resolution of 25, and *s* limits of 0.2 to 5.3 S. The f_r was set to 1.8, and allowed to refine. Both "spectra" were activated: the left-hand and right-hand boxes for spectrum 1 were set to 1 and 0, respectively. Thus spectrum 1 will report on the presence of chromophore 1 (GST-VCA) only; it will be a $c_{\text{GST-VCA}}(s)$ distribution. See the next page.

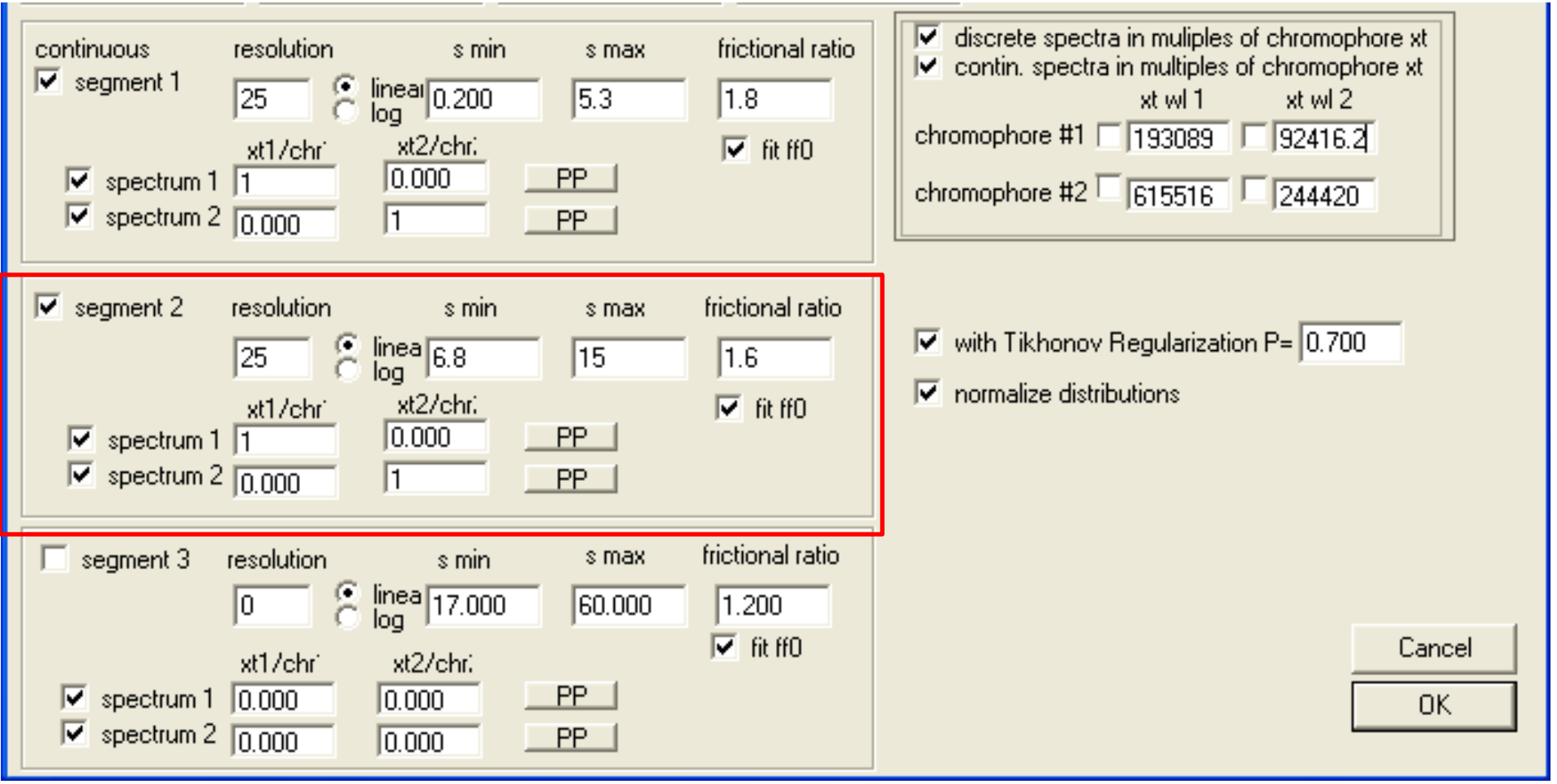
Setup Segment 1



Step 21 (continued)

The boxes for Spectrum 2 were set to 0 and 1, respectively; it will be a $c_{Arp2/3}(s)$ distribution. In Segment 2, a resolution of 25 was entered, and the *s*-range was set to 6.8-15 S. Experience has shown that it is best not to have substantial gaps or overlaps in molar mass- (*M*-) space when calculating segmented distributions. Given the large difference f_r's for the two segments, a 1.5-S gap in *s*-space was needed to ensure the proper coverage of *M*-space. As the analysis below demonstrates, there is no deleterious effect consequent to this gap. The f_r of segment 2 was set to 1.6, and its refinement was enabled. Both spectra were set up exactly as in segment 1. See the next page.

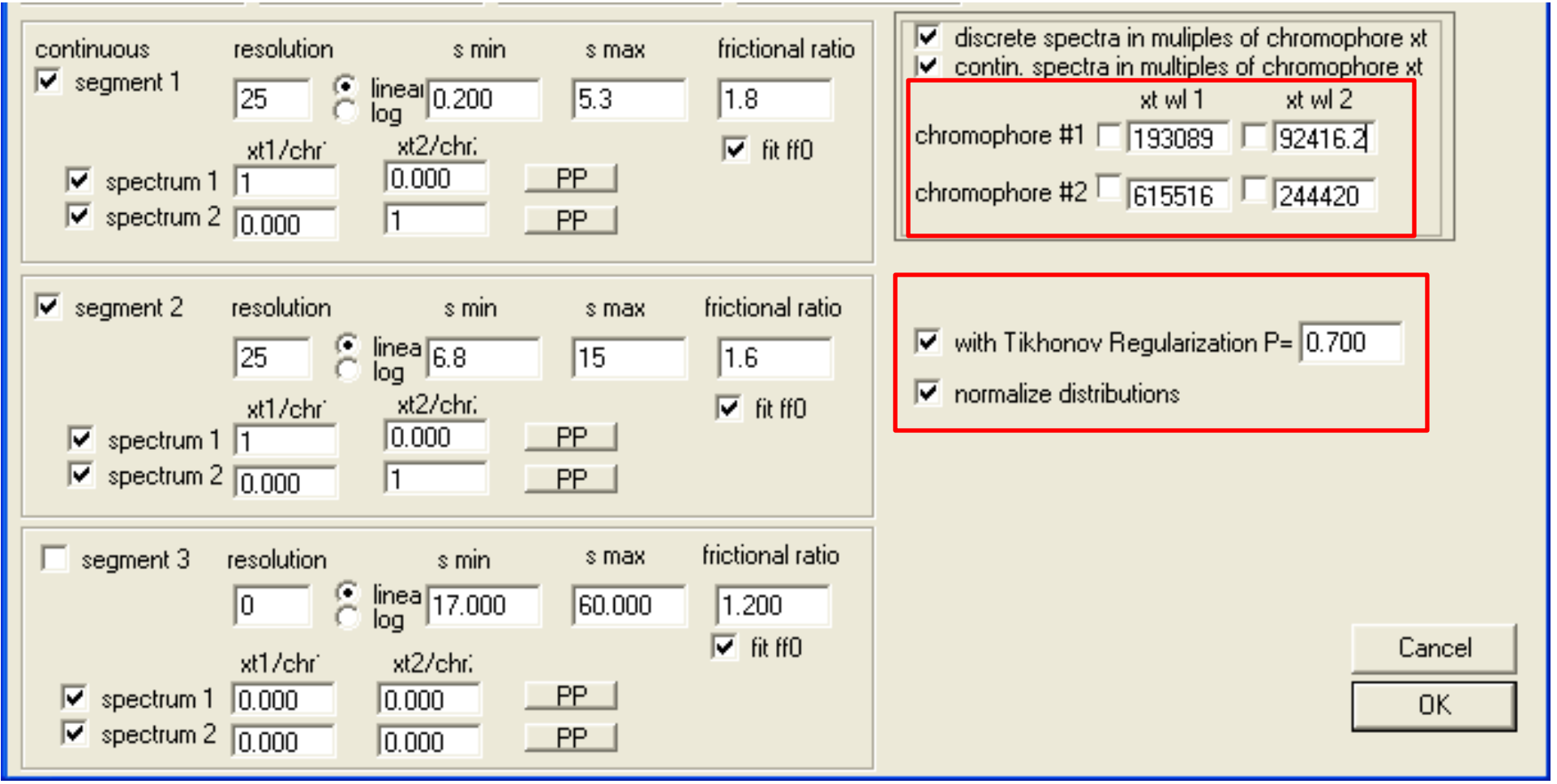
Setup Segment 2



Step 21 continued

In the extinction section, GST-VCA was designated “chromophore #1,” and its extinction properties were entered in this row: 193,089 and 92,420.2 (remembering that Experiment 1 was the IF data and Experiment 2 was the absorbance data). The extinction properties of “chromophore #2,” Arp2/3, were input in the proper entry fields: 615,516 and 244,420. It is very important to disallow the refinement of the extinction parameters in this final phase of the MSSV analysis. That is, none of the checkboxes next to the extinction parameters should be checked. Tikhonov Regularization was activated (P=0.70), as was normalization. See the next page.

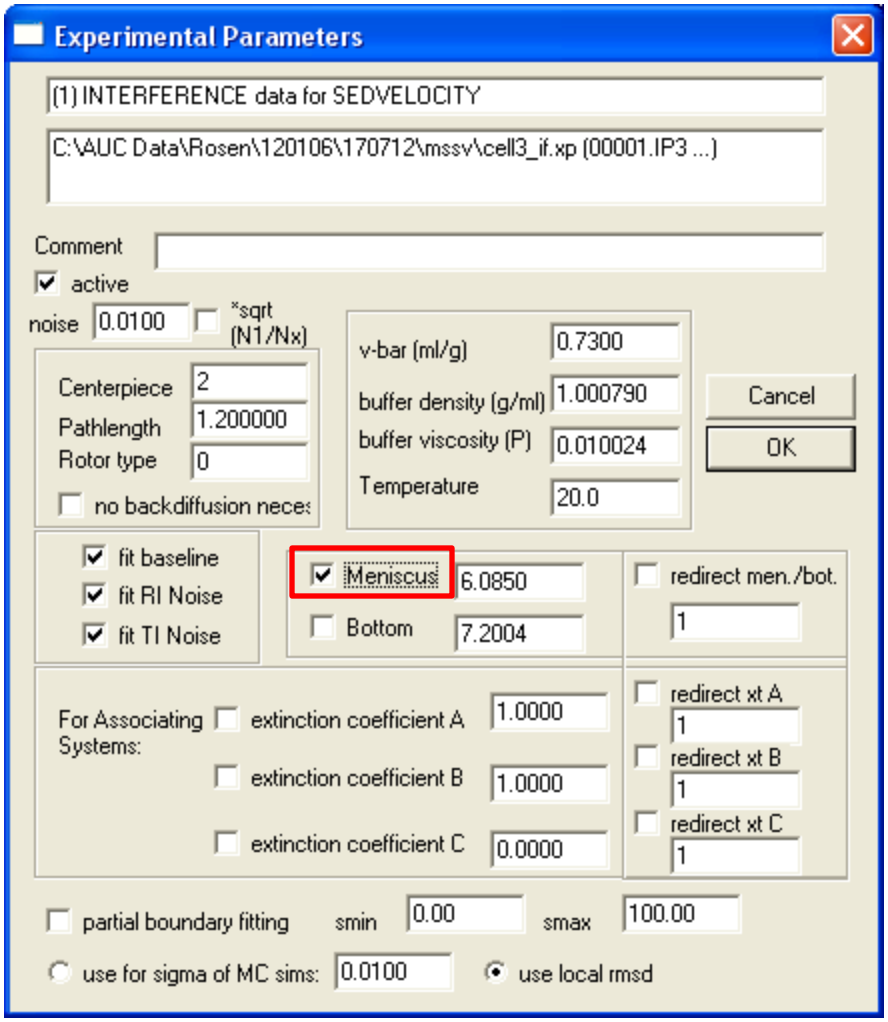
Enter Extinction, Regularization, and Normalization Info

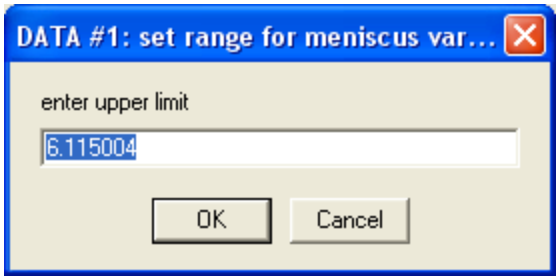


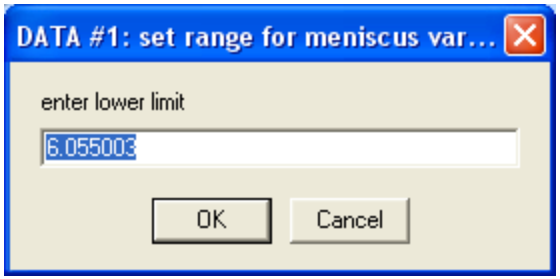
Step 22

The Experimental Parameters boxes were recalled, and the refinements of the respective menisci were enabled. The default values for the refinement limits of the menisci were accepted. See the next two pages.

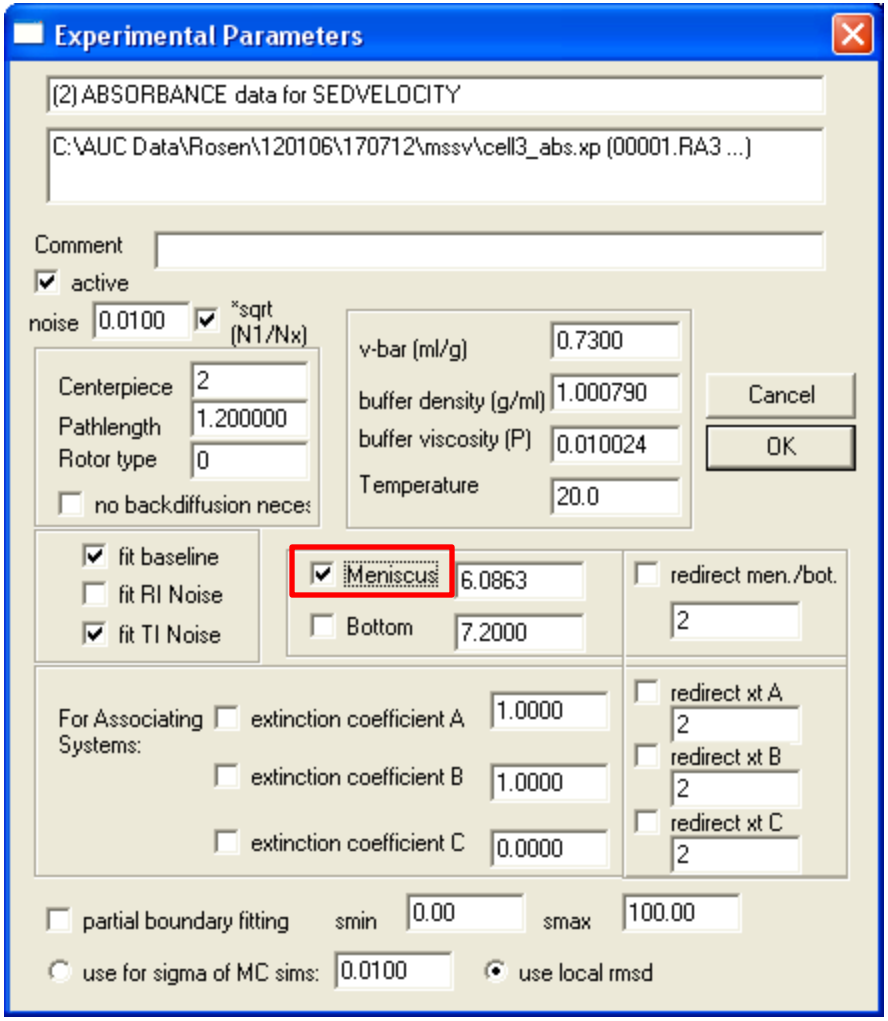
Activate Meniscus Fitting, Experiment #1

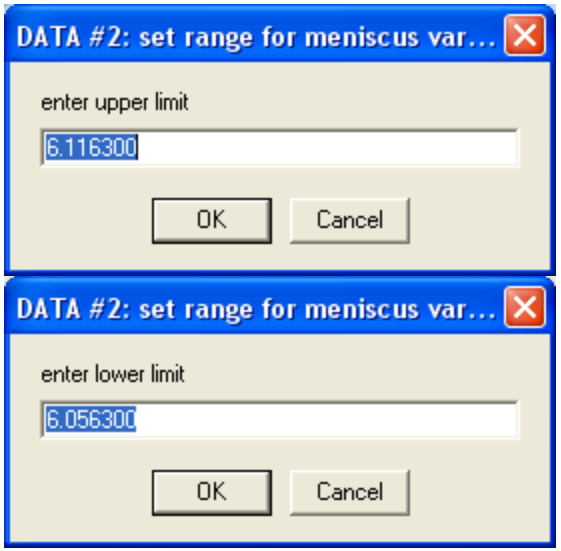






Activate Meniscus Fitting, Experiment #2

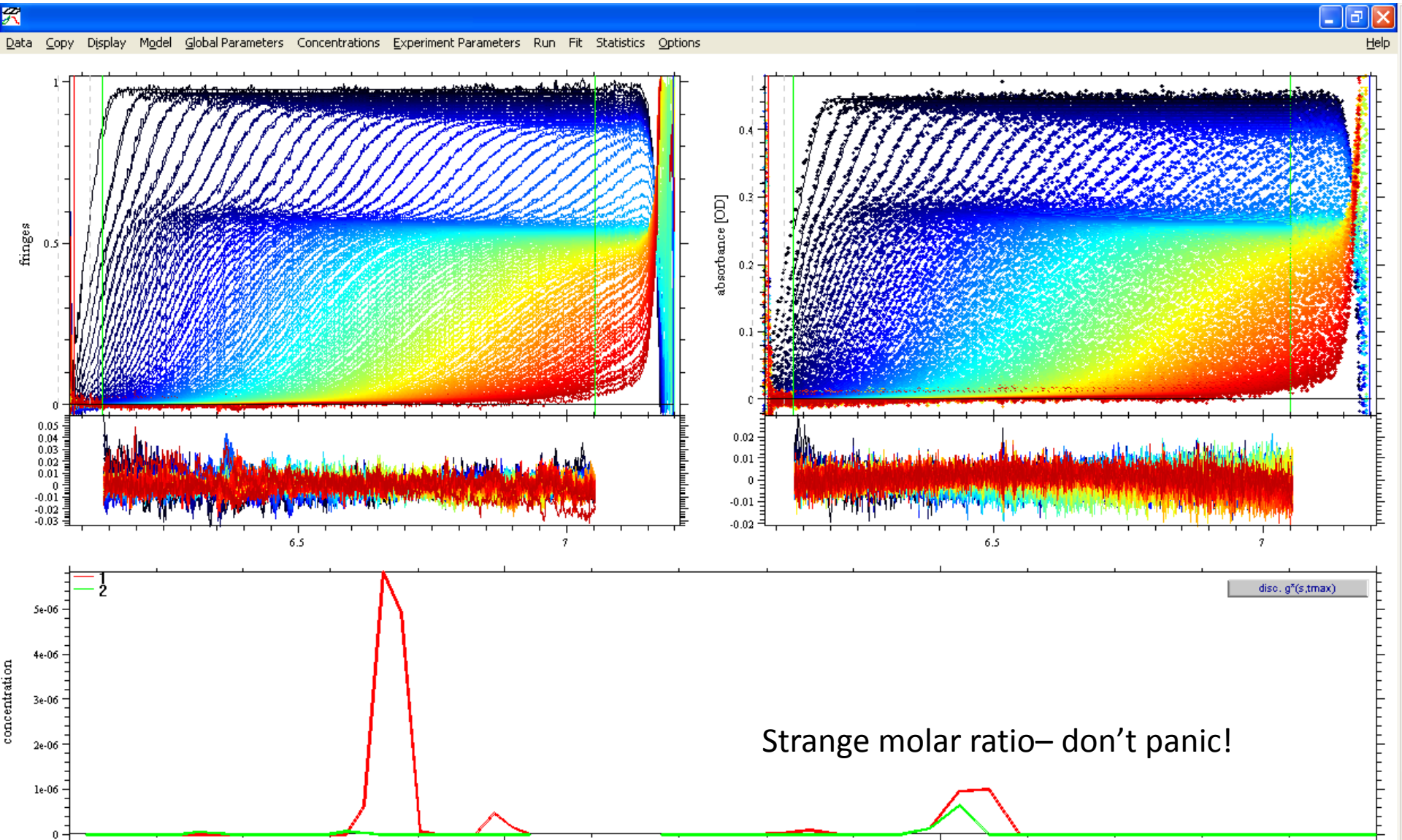




Step 23

A Global Run was initiated, followed by the subtraction of TI and RI noise (where appropriate). The fit, especially in the absorbance data, clearly has systematic features. The distributions ("spectra") in segment 2 show a significant molar excess of GST-VCA complexed to Arp2/3. The reader is cautioned not to attempt an analysis of molar ratio at this point. Experience has demonstrated that, until the fit is optimized, erroneous values for molar ratio can be arrived at. Despite the systematic errors present in the fit residuals, it was deemed close enough to begin an optimization of the non-linear parameters. See the result on the next page.

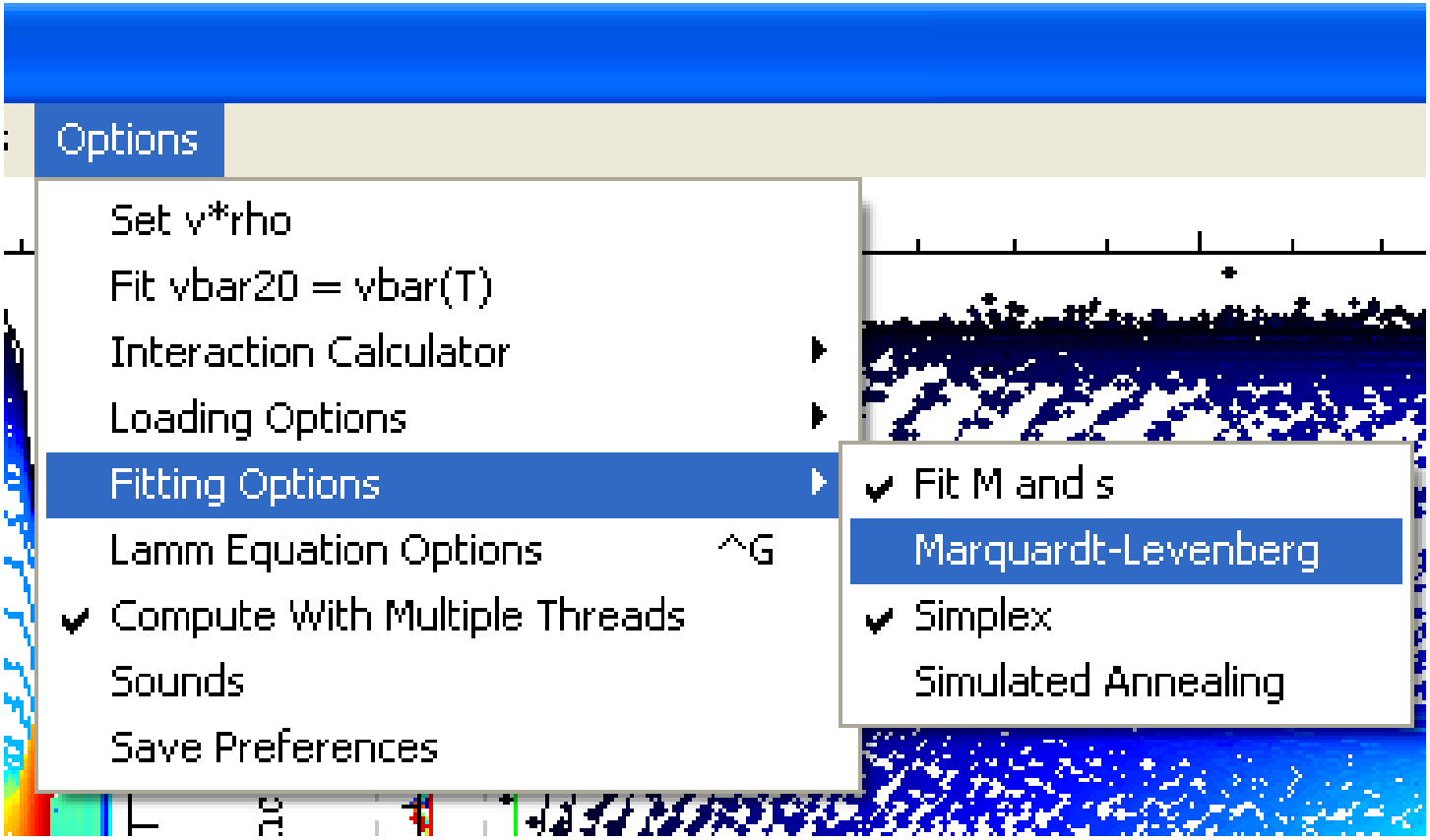
A Global Run


Data Copy Display Model Global Parameters Concentrations Experiment Parameters Run Fit Statistics Options Help
fringes
absorbance [OD]
concentration
disc. g*(s,tmax)
Strange molar ratio– don't panic!

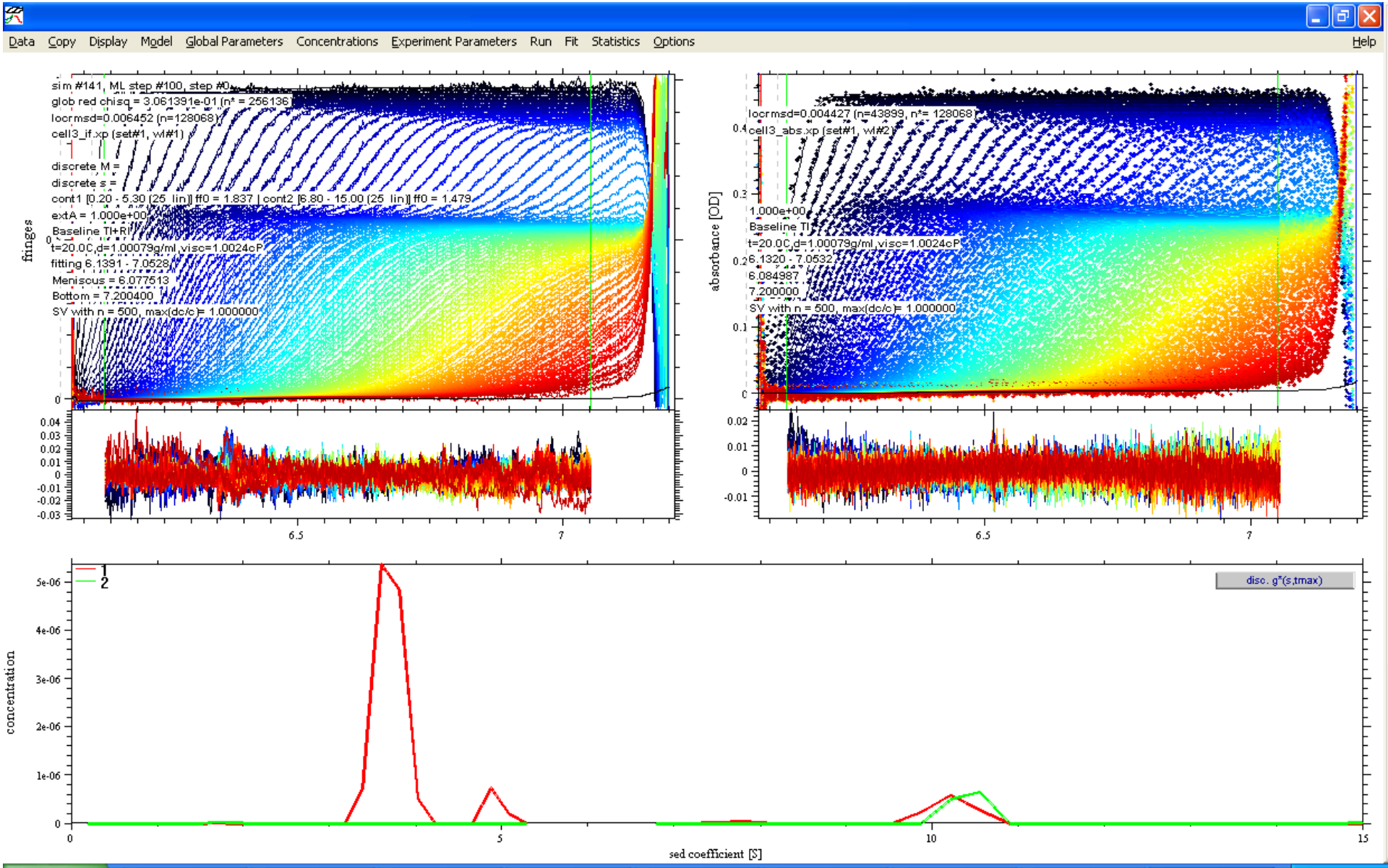
Step 24

A Global Fit was started using the Marquardt-Levenberg minimization algorithm. After about 15 minutes of minimization, convergence was achieved. The configuration was saved. The quality of the fit was judged to be good. The local rmsd for the IF data was 0.006452 and that for the A_{280} data was 0.004427. The residuals were non-systematic. At this point, the fitted χ^2, or χ_b^2, was noted to be 0.3061391. The fit was saved. See the next four pages.

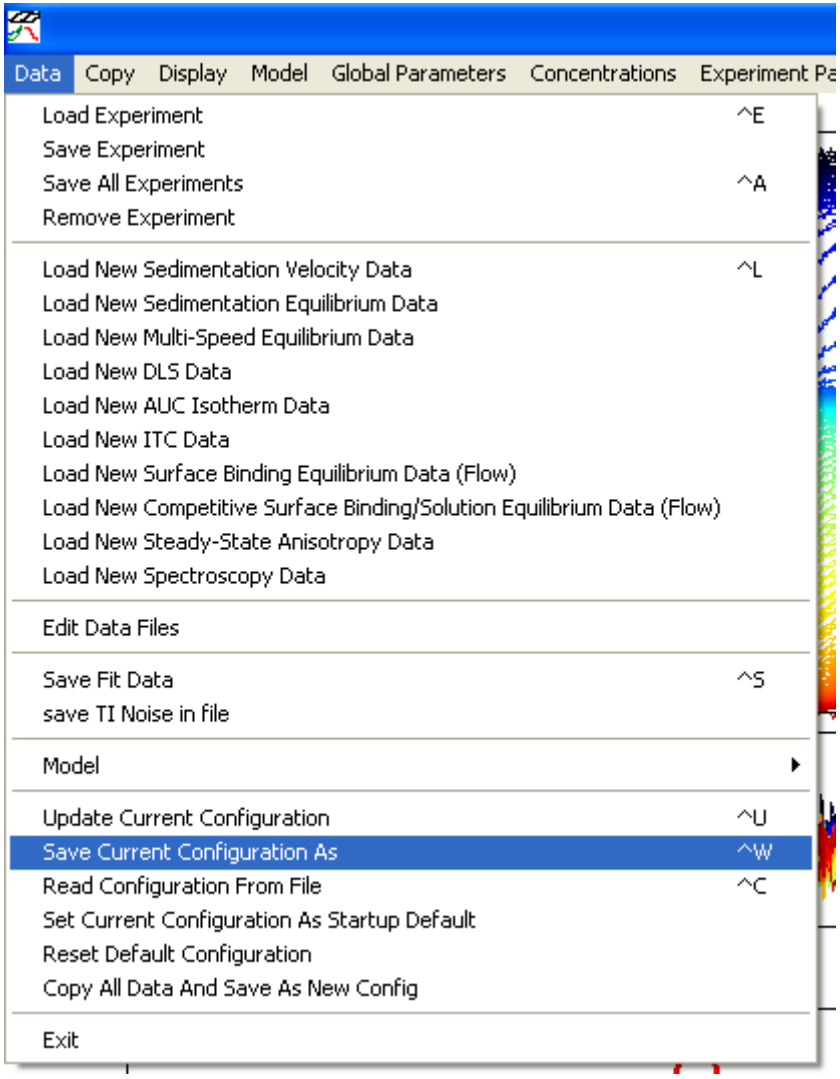
Set the Minimization Algorithm to ML



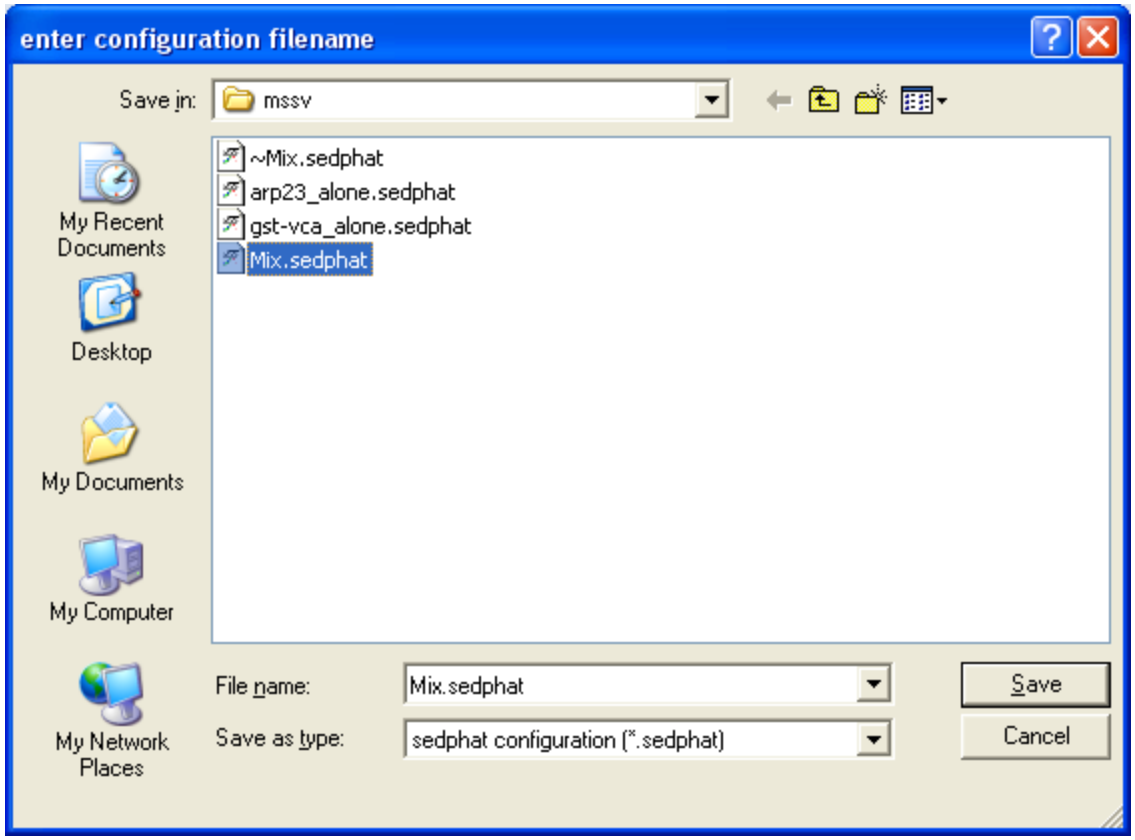
After Global Fit



Save the Configuration



Save Configuration Dialogs



Step 25

At this point, we turned to an initial assessment of the fit's reliability. As mentioned in section 4.1 of the paper, χ_b^2 alone is not a sufficient criterion to assess the outcome. The most important criterion to assess the success of the spectral discrimination is mass conservation; in other words, are the concentrations of components detected in the distributions close to our expectations based on the concentrations obtained from the analyses of the individual components? For example, in the GST-VCA alone experiment, the concentration of the protein was found to be 4.09 μM. Because 87 μL of the stock solution was included in 410 μL of total sample, the stock concentration of GST-VCA was calculated to be 19.1 μM. Because 66 μL of this stock was used to make the 410 μL of the mixture sample, the expected [GST-VCA] is 3.08 μM. By integrating both $c_{\text{GST-VCA}}(s)$ distributions as in Step 15 above (see page 105 of this protocol for this result), it was found that [GST-VCA] = 3.01 μM, in excellent agreement with the expectation. Similarly, the expected [Arp2/3] was calculated to be 0.38 μM, and the actual detected [Arp2/3] = 0.40 μM. Again, the agreement is excellent, and thus the most important criterion for success is easily met.

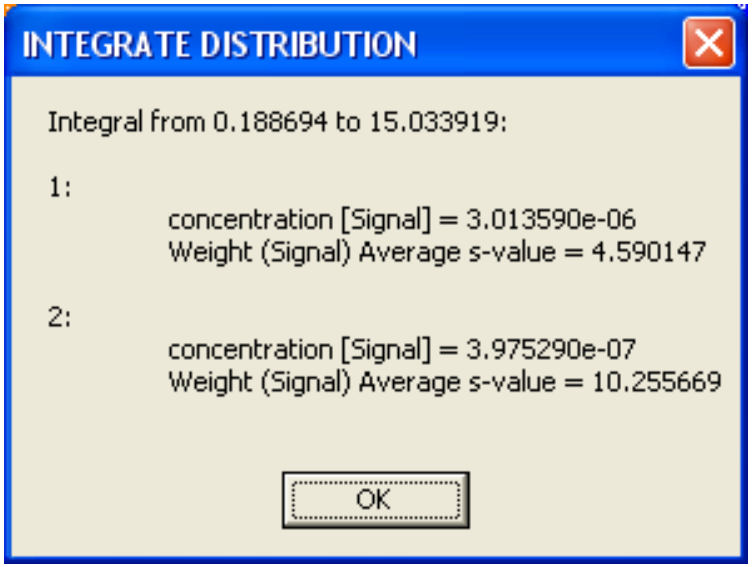
Step 25 continued

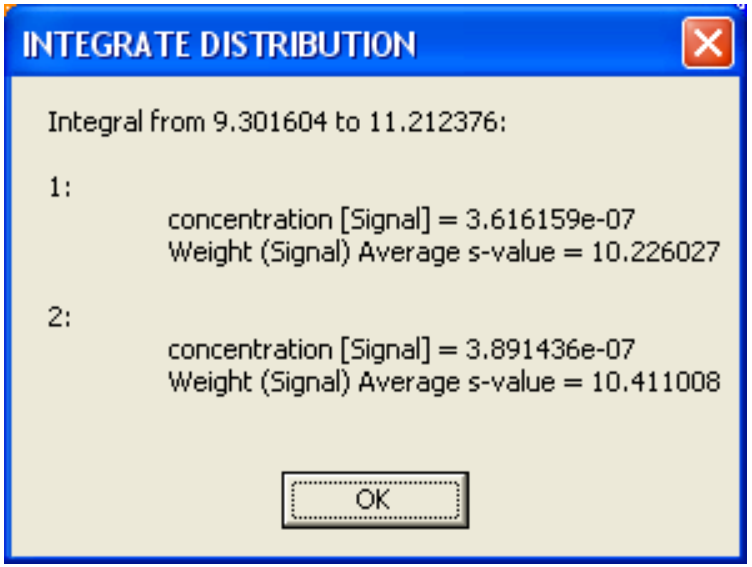
Another important criterion is the absence of compositional contamination. In this case, it is important that little to no Arp2/3, a large protein with a sedimentation coefficient of about 8.9 S, be detected at low *s*-values where the excess GST-VCA was detected. Examination of Page 99 demonstrates that this criterion is also met; there is no signal for Arp2/3 in the peaks that describe the sedimentation of GST-VCA at 3.7 and 4.9 S. Also, as expected, GST-VCA and Arp2/3 appear to be sedimenting in a complex of the two components. Two of the most important criteria for spectral discrimination are therefore met, and the expectation is that this analysis was successful. According to the refined signal increment/extinction information for this system, D_{norm} = 0.068; in retrospect, good spectral discrimination was therefore to be expected for these two proteins, according to the simulations carried out in section 4.1 of the paper.

Step 26

Another integration was performed. This time, only the range 9.3-11.2 S was considered. This was the region of the cosedimenting complex. Here, [GST-VCA] = 0.36 μM and [Arp2/3] = 0.39 μM. Thus, the molar ratio of GST-VCA to Arp2/3 in the complex was calculated to be 0.92 to 1. The molar ratio appears to be close to 1:1.
See the next page.

Some more integration





These conform to expected values.

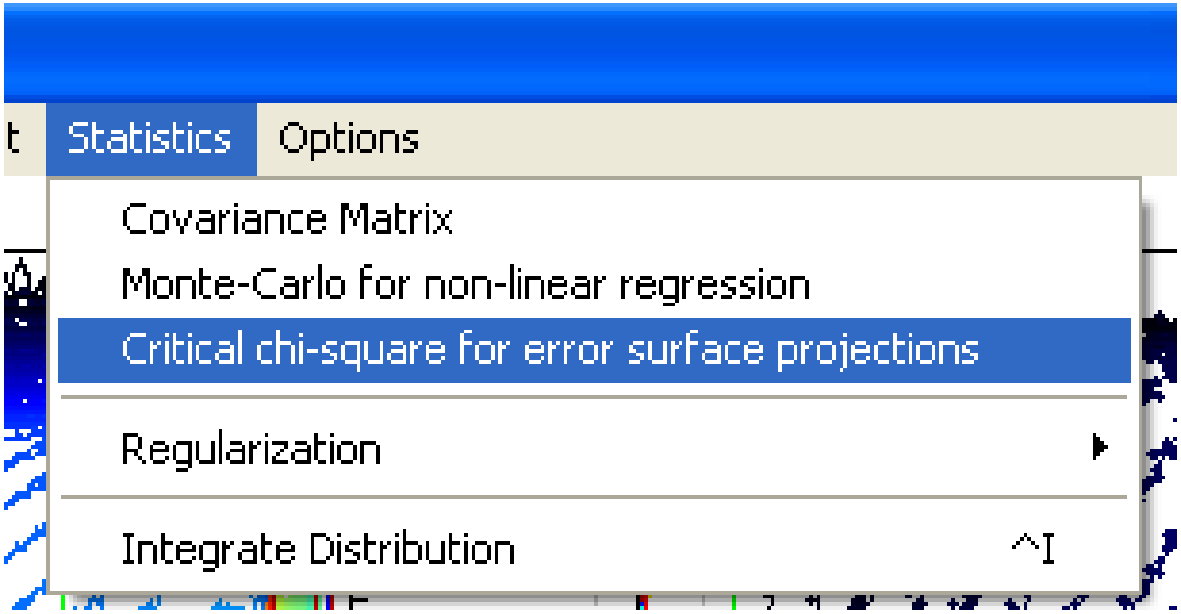
Step 27

Given the data, it seemed reasonable to conclude that all of the material in segment 2 could be represented by 1:1 molar ratios of GST-VCA and Arp2/3. To examine this possibility, a statistical criterion for acceptability must be defined. If the material between 6.8 and 15 S could be modeled using 1:1 complexes with the fit becoming less than 1(σ) worse than the best fit, the 1:1 model could be deemed acceptable. The change in the quality of the fit was to be judged using the χ^2 statistic. The χ^2 of the converged fit (χ_b^2) was 0.3061391. To determine whether any alternative fits to the data are worse by the criterion mentioned above, the χ^2 of a 1(σ) worse fit must be established. To do this, the "Statistics" menu item was clicked, and "Critical chi-square for error surface projections" was selected. A dialog appeared, prompting the user to enter the desired confidence level. Because 1(σ) was desired, the default value of "0.683" was accepted. Another dialog then appeared, telling the user that the critical χ^2 ($\chi_{c,1\sigma}^2$) was 0.306843. Thus, if the χ^2 of any alternative fit (the "test χ^2", or χ_t^2) exceeds this value, the quality of the fit will be deemed statistically worse than that of the best fit.

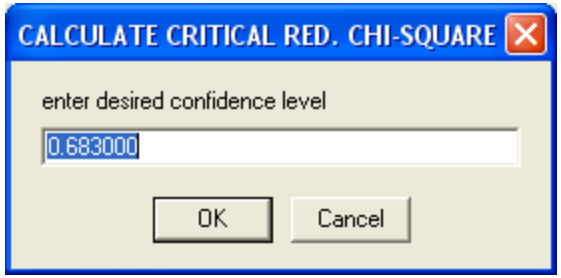
Step 27 continued

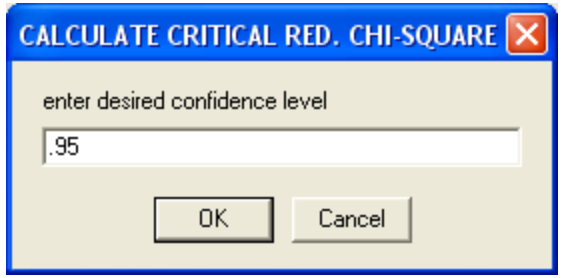
Also, a rejection criterion was established. The same statistical calculation was performed, but the confidence level was set to 0.95. This value was termed "$\chi^2_{c,2\sigma}$" and found to be 0.308574; if χ^2_t for a constrained fit exceeded this value, it could be safely rejected as likely to be incorrect. See the next two pages for a pictorial description of these calculations.

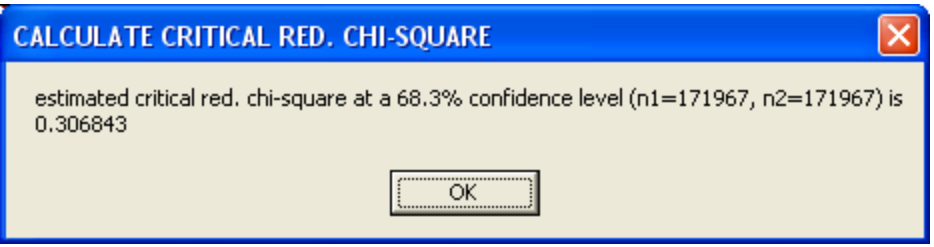
Initiate Statistics Calculations

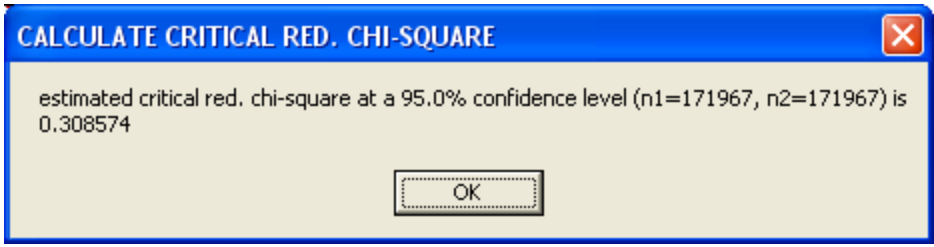


Calculating the critical χ^2's









$$\chi^2_{c,1\sigma}$$

$$\chi^2_{c,2\sigma}$$

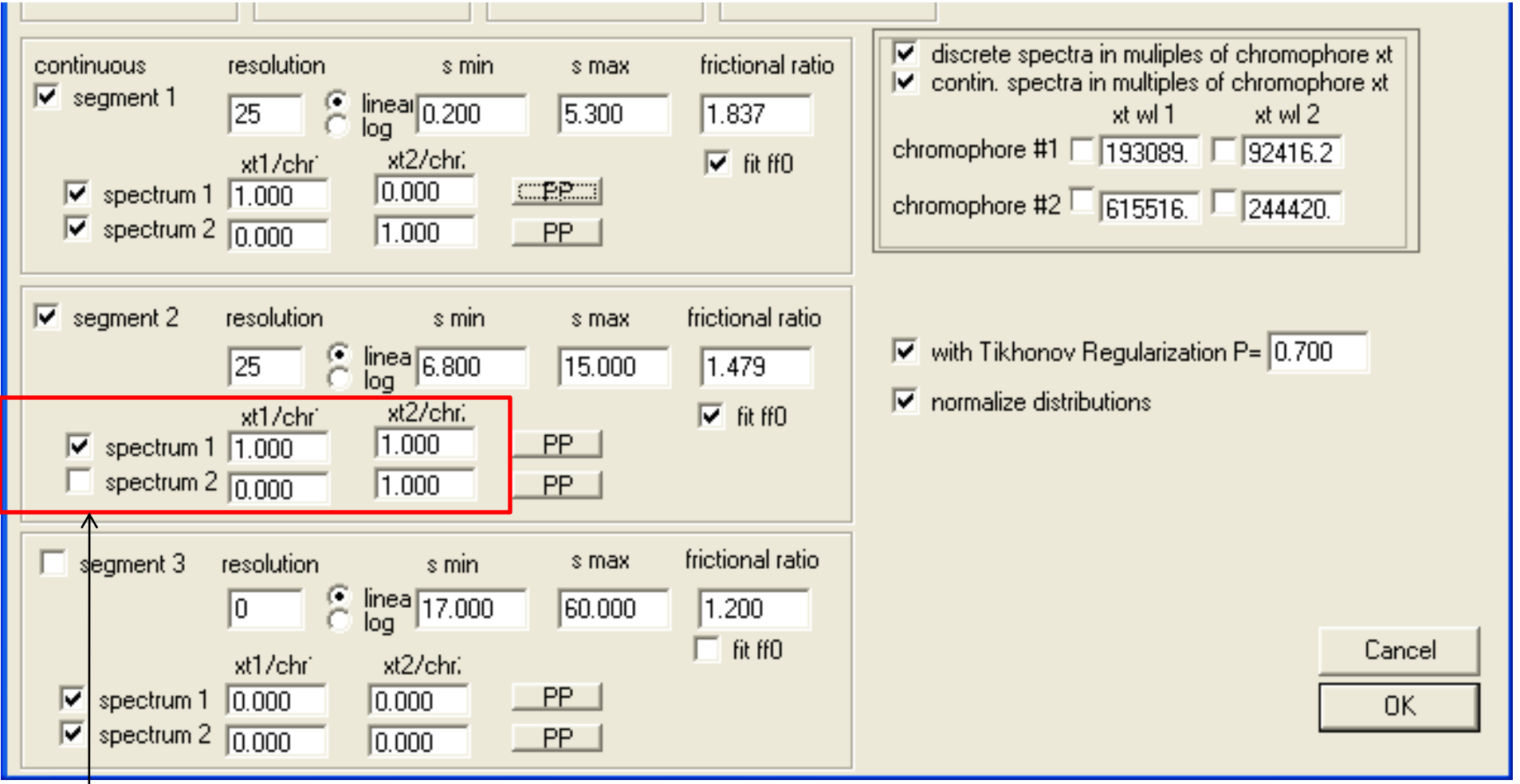
The criterion for "statistical distinguishability"

The criterion for remorseless rejection

Step 28

The Global Parameter Box was recalled. In the "segment 2" section, "spectrum 2" was deactivated, and the boxes for "spectrum 1" were altered such that they both had "1" in them. Thus, the program will attempt to model the material sedimenting between 6.8 and 15 S as 1:1 complexes of GST-VCA and Arp2/3. In other words, a single $c_{\text{GST-VCA:Arp2/3}}(s)$ distribution will be used to model these data in the given *s*-range. The parameter box was dismissed by pressing "OK." See the next page.

First, let's try a 1:1 molar ratio

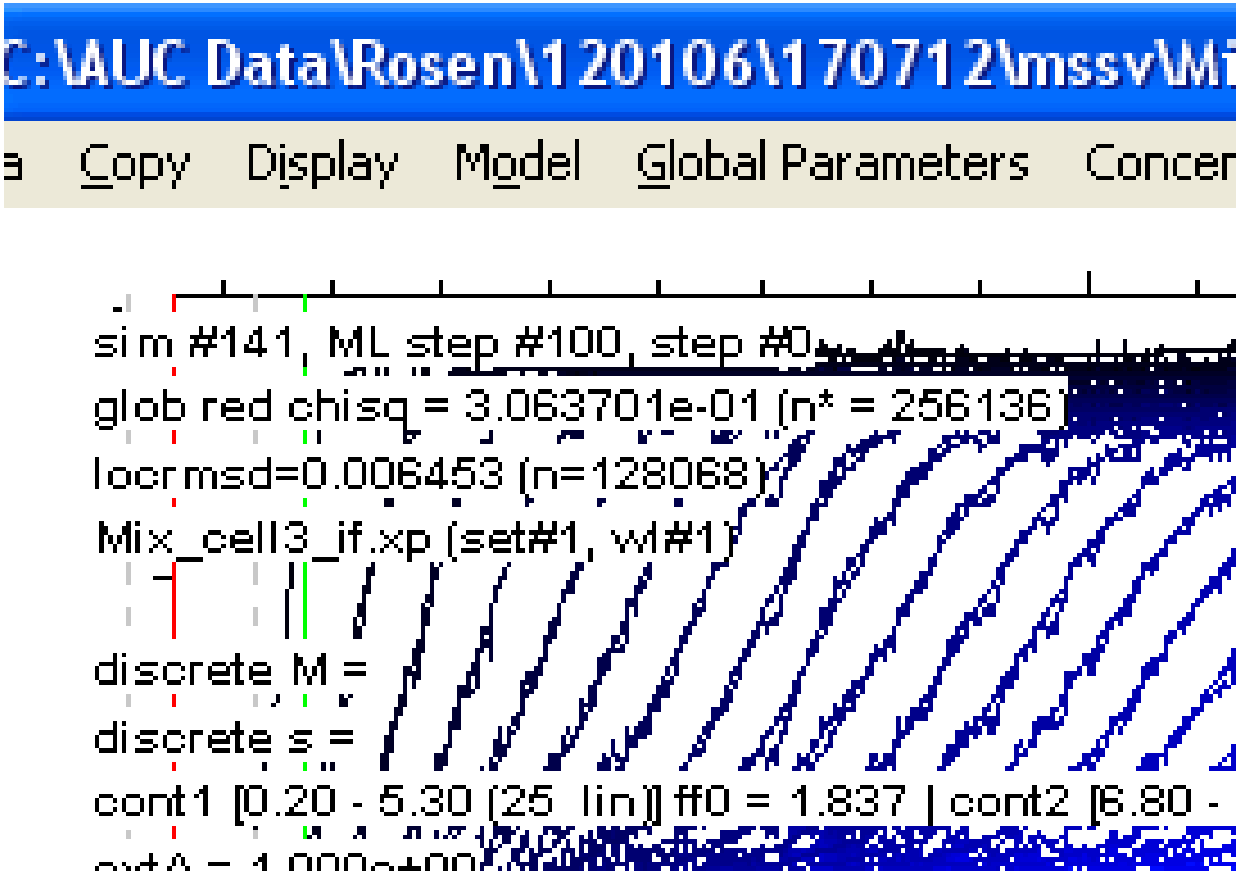


This constrains segment 2 to fit the data with a 1:1 ratio of the two "chromophores"

Step 29

A Global Run was performed. The value of χ_t^2 was 0.3063701. Global fitting at this point would only make the fit better, yet the new fit already meets the established criterion for acceptability. According to the statistical criteria that we have defined, we can accept the 1:1 constraint as consistent with the MSSV data.

Global Run

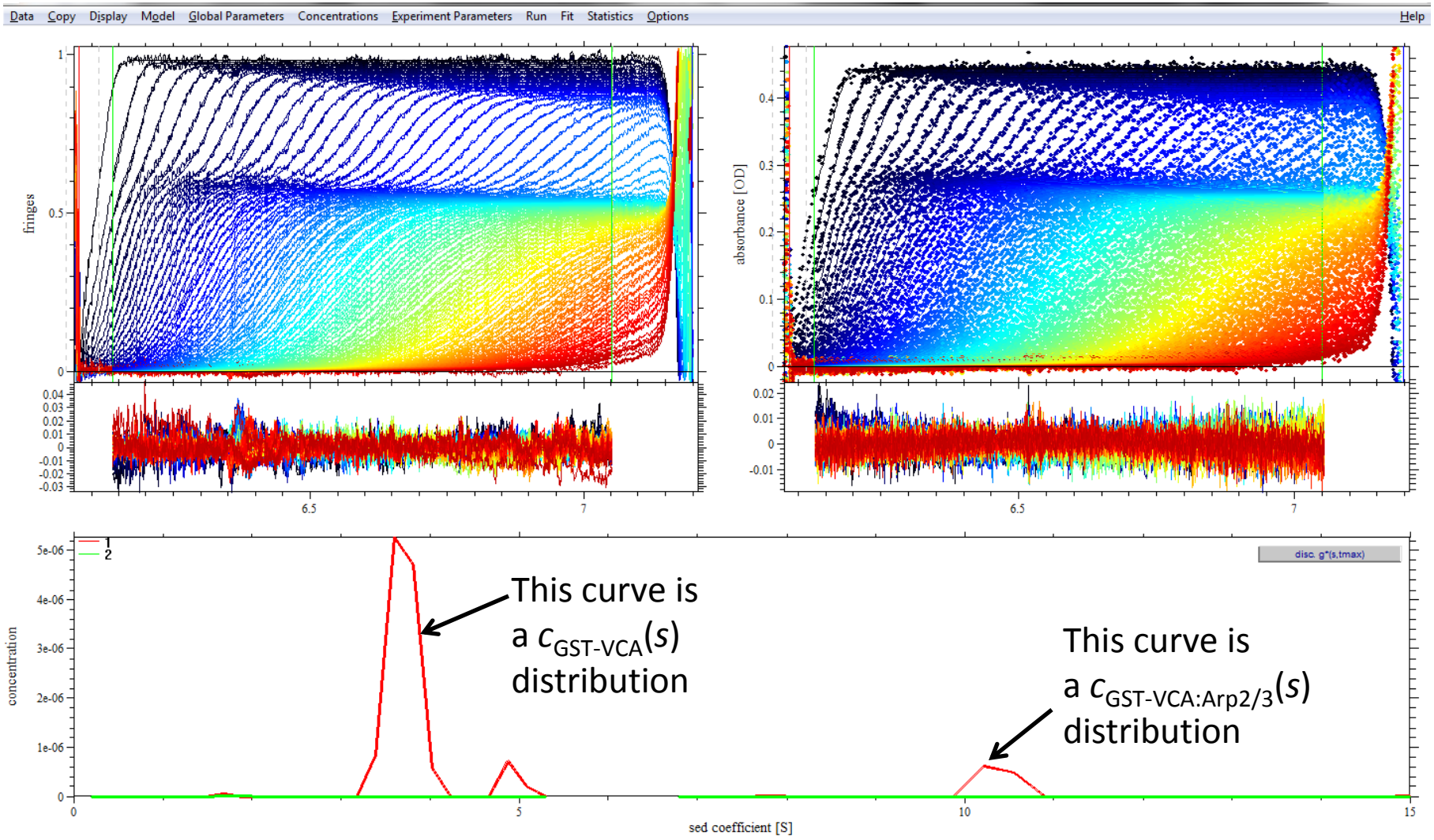


The value isn't above either critical chi-squared, so this molar ratio is consistent with the data.

Step 30

For illustrative purposes, a Global Fit was initiated. After about 10 min., the fit converged. The χ_t^2 was 0.3063701, unchanged from the value obtained from the Global Run in Step 29. As a check on the internal consistency of the result, the 10.3-S peak in the $c_{\text{GST-VCA:Arp2/3}}(s)$ distribution was integrated. Our calculation above indicated that 0.38 μM Arp2/3 was expected, so 0.38 μM of the complex should also be expected. Indeed, the integrated [GST-VCA:Arp2/3] = 0.38 μM. The 1:1 model was therefore accepted as consistent with our MSSV data. See the next two pages.

After a Global Fit



The Integration

Integral from 9.545454 to 11.126814:

1:

concentration [Signal] = 3.818431e-07
Weight (Signal) Average s-value = 10.367032

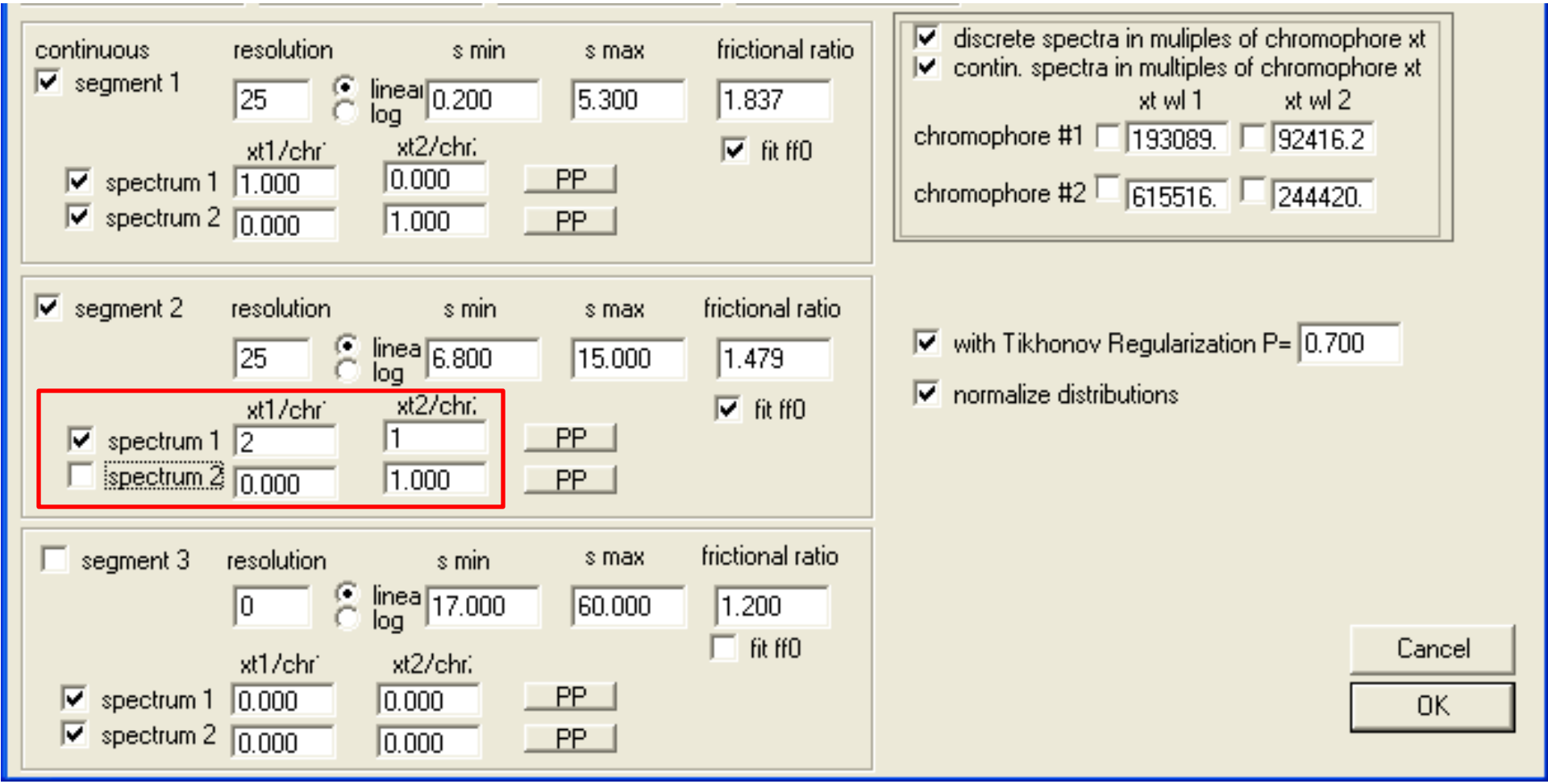
2:

concentration [Signal] = 0.000000e+00
Weight (Signal) Average s-value = 0.000000

Step 31

What about other possible stoichiometries? For the 10.4-S species, the only other stoichiometries that merit consideration are two GST-VCA's to one Arp2/3 complex and two Arp2/3 complexes bound to one GST-VCA dimer (although stoichiometries involving more than one Arp2/3 complex should have greater *s*-values than 10.4 S). To explore these possibilities, the saved, best-fit configuration of unconstrained fit was opened. To do this, the Data menu item was selected, and "Read Configuration from File" was selected. The defaults were accepted, and the Global Parameter box automatically appeared. All of these parameters were correct except for those in the segment 2 portion of the box. First, the possibility of two GST-VCAs binding to one Arp2/3 was examined. As in Step 28, "spectrum 2" was turned off (unchecked). The left and right boxes for spectrum 1 were altered to read "2" and "1", respectively. This fits the 6.8 to 15 S region with a two GST-VCA dimer to one Arp2/3 complex. The Global Parameters box was dismissed by pressing "OK." The next page shows the Global Parameters box for this trial.

Trying 2:1 Now

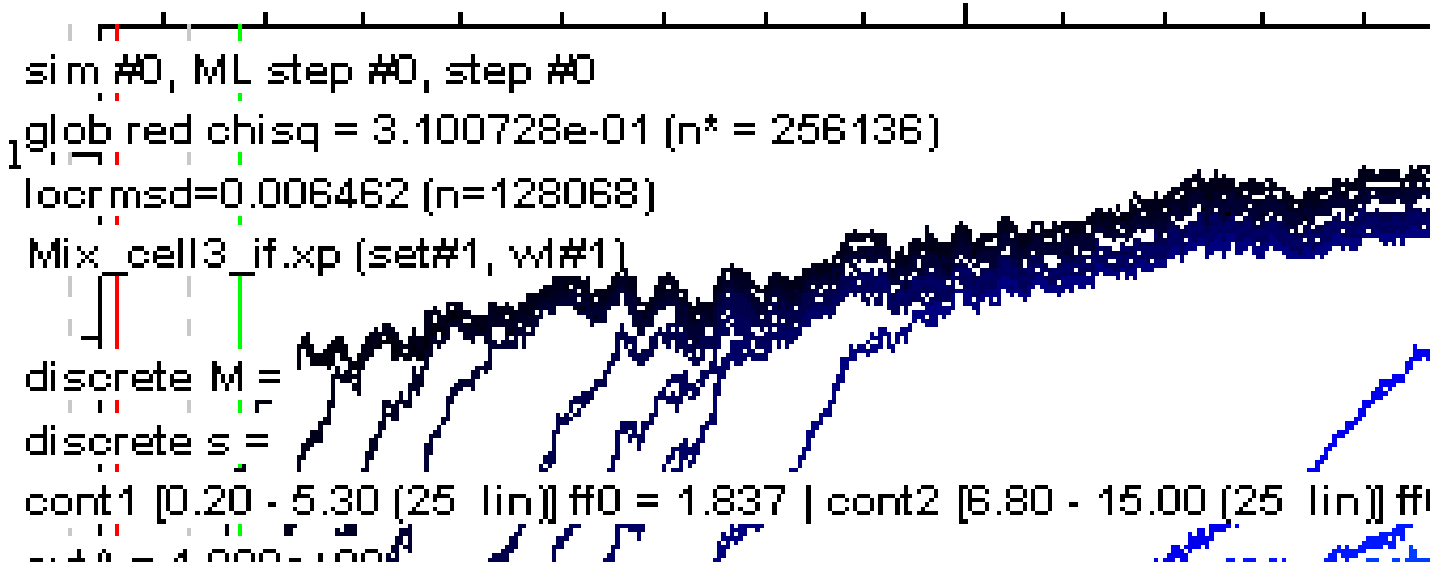


Step 32

A Global Run was initiated. The was 0.3100728, well above both $\chi^2_{c,1\sigma}$ and $\chi^2_{c,2\sigma}$. A Global Fit was then performed. The resulting χ^2_t was 0.3096959, which was greater than $\chi^2_{c,2\sigma}$. Thus, the 2:1 stoichiometric constraint was safely rejected. These statistics are documented on the next two slides.

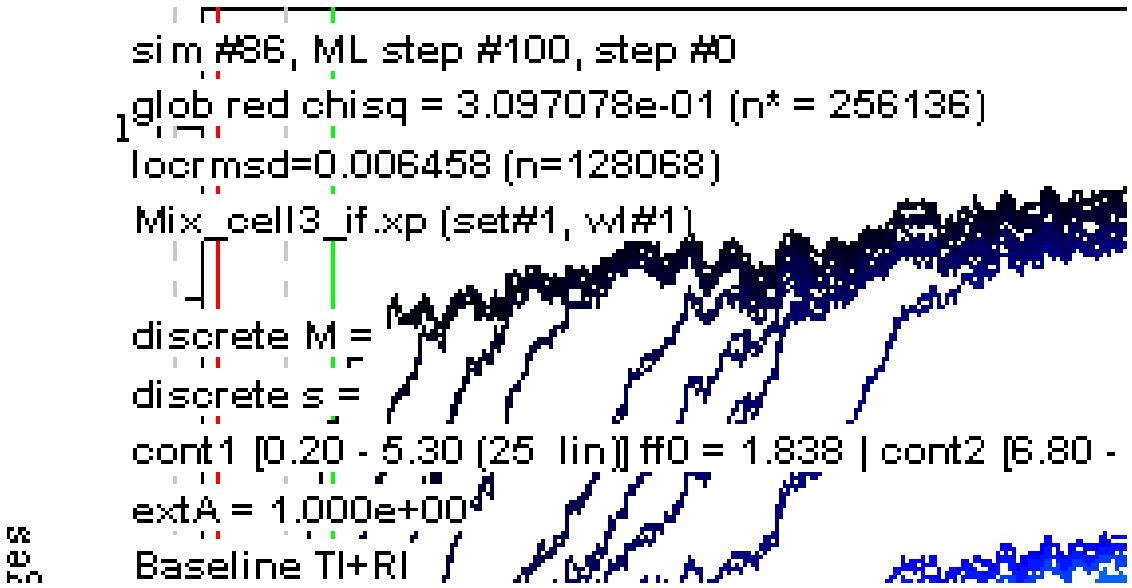
It is worth noting that, in an initial, unshown analysis of these data was carried out with half of the scans shown here. When applying the 2:1 stoichiometric constraint, the situation arose of $\chi^2_{c,1\sigma} < \chi^2_t < \chi^2_{c,2\sigma}$, as in Step 35. By including *all* of the data (above), the 2:1 constraint was easily rejected. This is an example of the fact that including more data results in superior spectral resolution, as mentioned in section 4.1 of the text of the parent paper of this protocol, Padrick & Brautigam, Methods, 2011.

After a Global Run



This χ^2 is above both critical values. We'd better fit to see if the ML algorithm can improve the stat.

After Global Fit



The fit has converged, but the chi-squared is still above both critical values. We may remorselessly reject the 2:1 molar ratio.

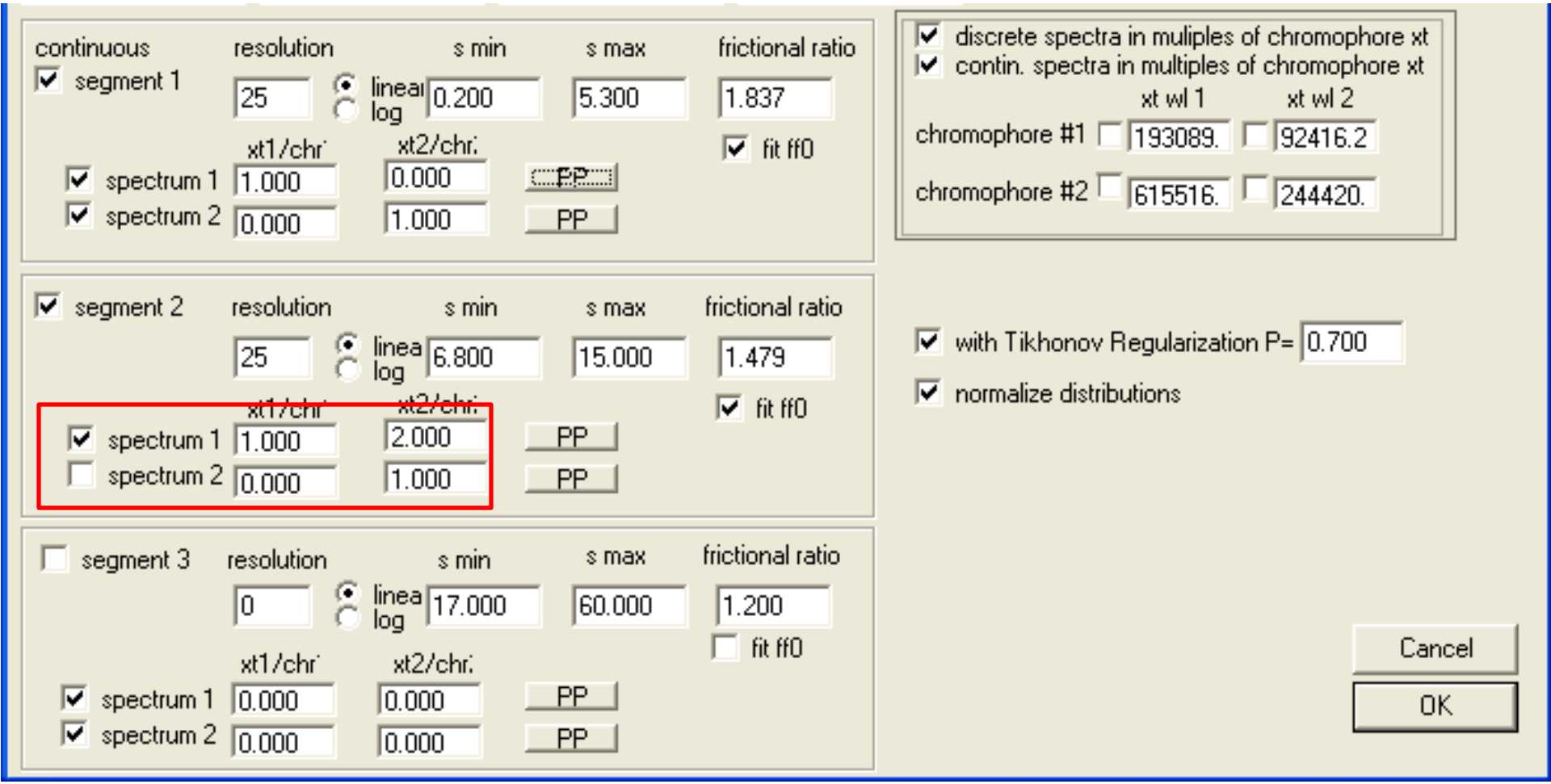
Step 33

To examine the possibility of 1 GST-VCA and 2 Arp2/3's in the complex, a new SEDPHAT session was started, and the best configuration was loaded into the program by selecting "Read Configuration from File" from the Data menu item.

Step 34

The Global Parameter box was summoned, and Step 31 was repeated, but this time, the left and right spectrum 1 boxes were altered to read “1” and “2”, respectively. See the next page.

Trying a 1:2 Molar Ratio

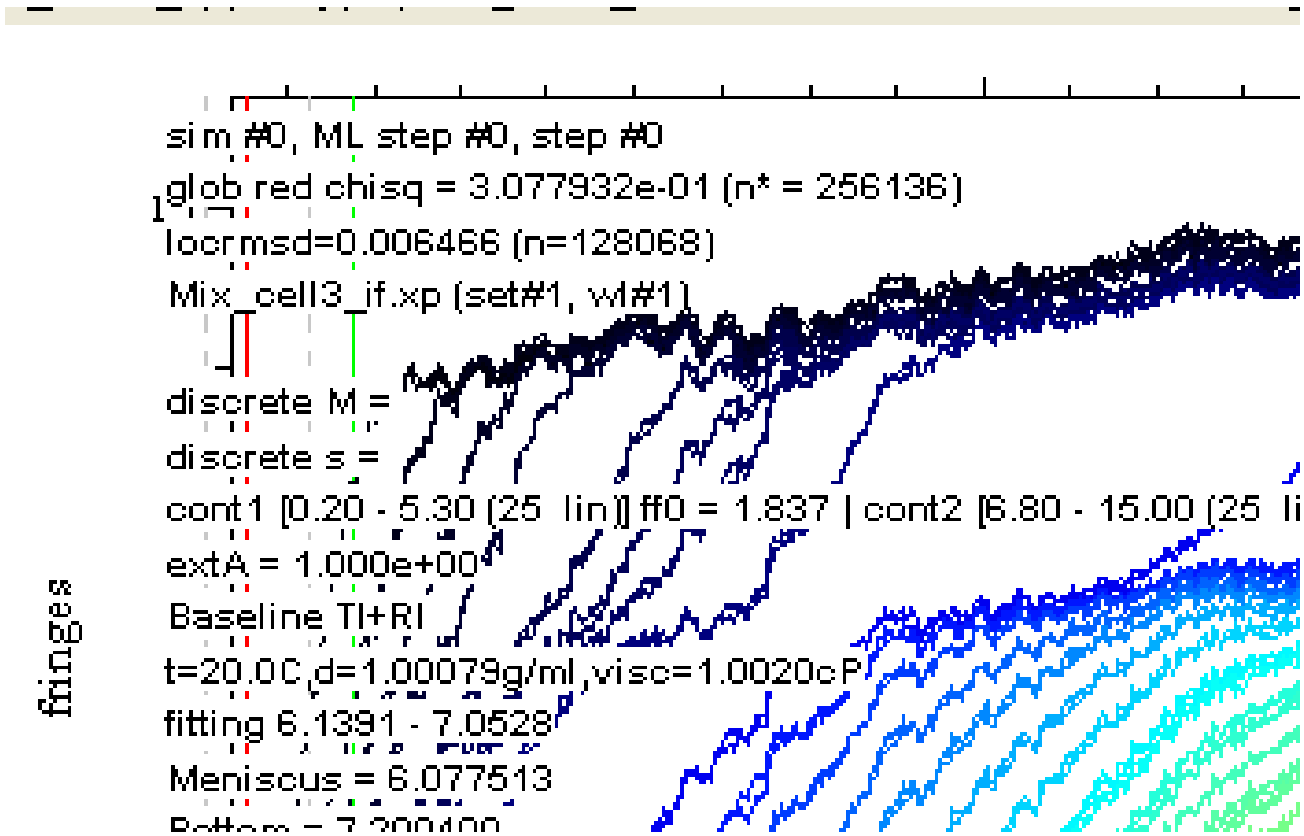


IMPORTANT: This is a new SEDPHAT session– we reloaded the saved fitting parameters, then summoned this box.

Step 35

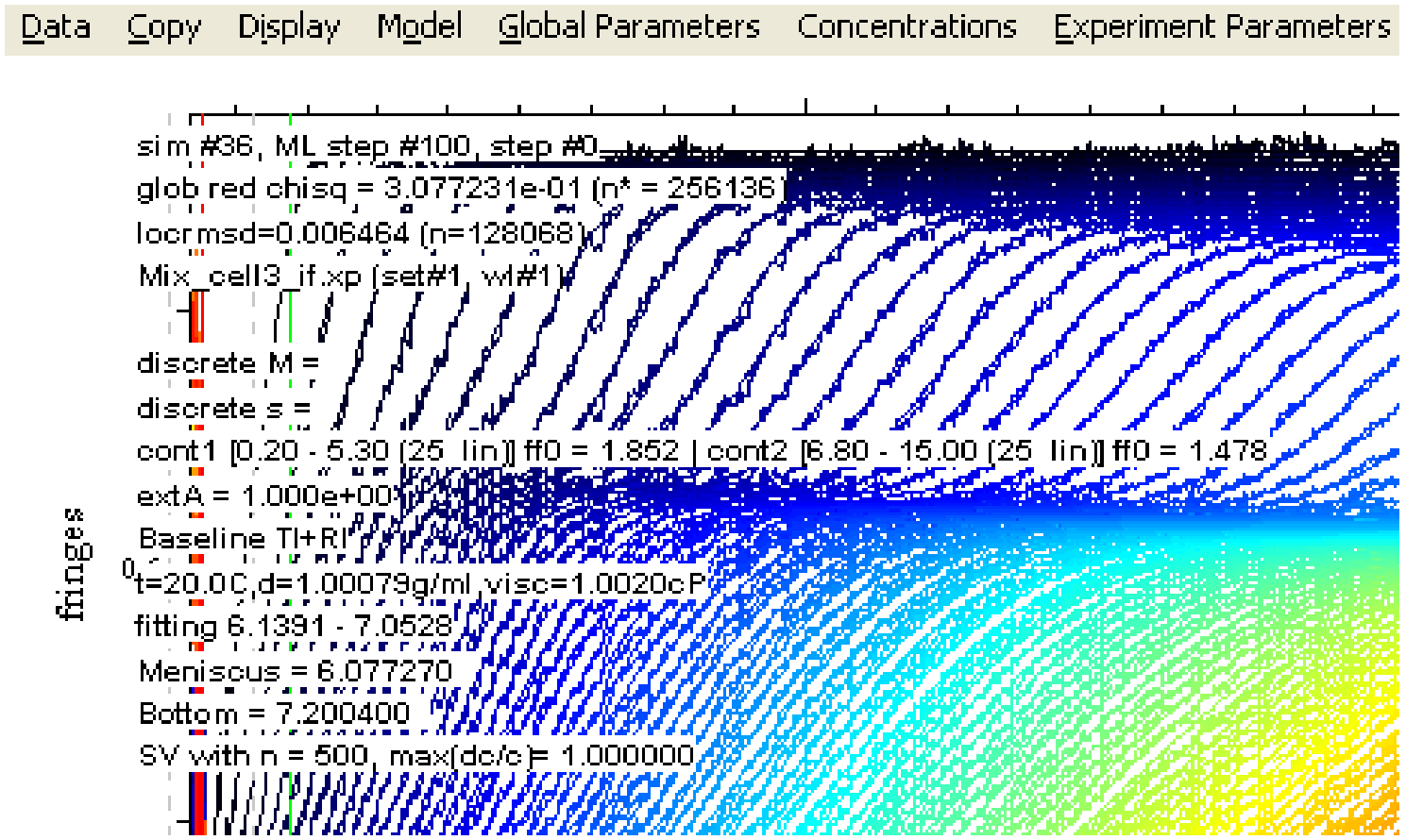
A Global Run was performed. The χ_t^2 value was 0.3077832. A Global Fit was performed using the Marquardt-Levenberg fitting algorithm. The final converged value was 0.3077230. This value is above $\chi_{c,1\sigma}^2$, but below $\chi_{c,2\sigma}^2$. Thus, the stoichiometric constraint has made the fit significantly worse, but the constraint is not rejected by our 2(σ) criterion. However, consideration of conservation of mass casts significant suspicion on this stoichiometry. The Arp2/3 concentration in the complex with the 1:2 stoichiometry was found to be 0.434 μM, 14% above the input concentration of 0.38 μM. Additionally, other factors, such as the hydrodynamic behavior of the complex, make the 1:2 complex of GST-VCA and Arp2/3 very unlikely. Salient points concerning this step are found on the next three pages.

After a Global Run



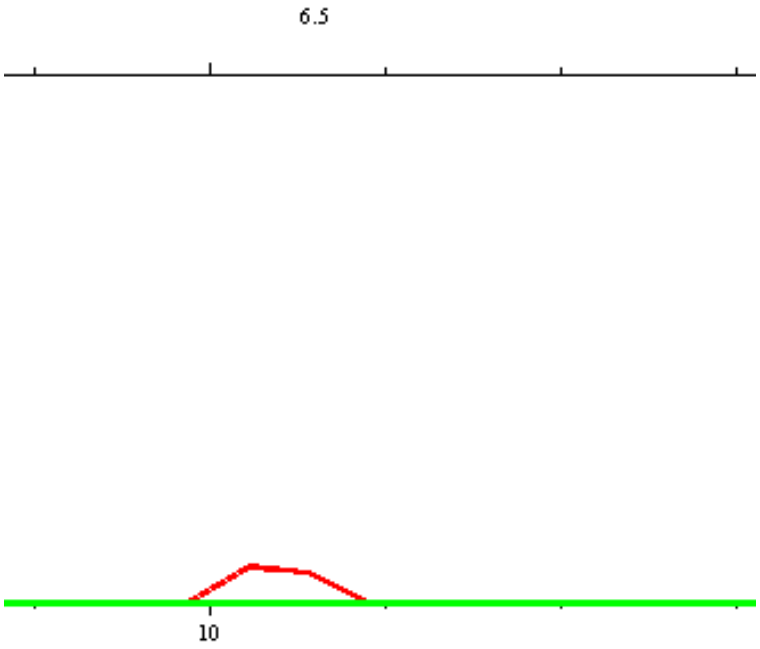
We are already below the 2(σ) cutoff. Can the ML algorithm get us to statistical indistinguishability?

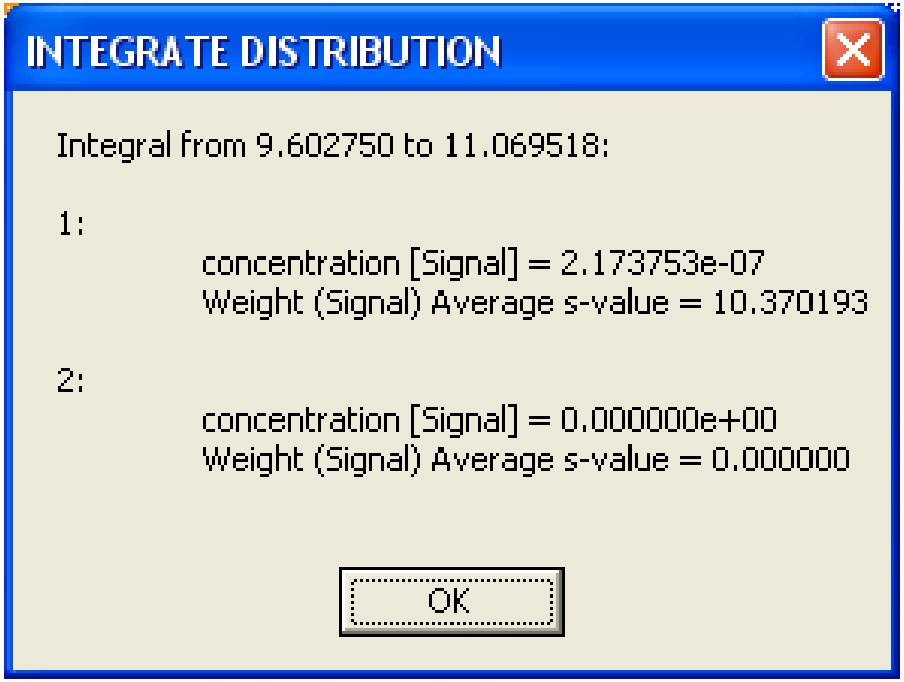
After a Global Fit



No, we didn't get there. This fit is statistically worse than the best, unconstrained fit. But it isn't instantly rejectable, either.

Concentration





But, let's integrate the peak in the second segment. The concentration of the complex is 0.217 μM. That means that [Arp2/3] = 0.434 μM. But we expected 0.38 μM. Our integrated value is thus 14% greater than our expected value.

References

[1] P. Schuck, Size distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophysical J. 78 (2000) 1606-1619.

[2] J.L. Cole, J.W. Lary, T.P. Moody, and T.M. Laue, Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium. in: J.J. Correia, and H.W.I. Detrich, (Eds.), Biophysical Tools for Biologists. Volume One: In Vitro Techniques., Academic Press, 2008, pp. 143-179.

[3] S.B. Padrick, R.K. Deka, J.L. Chuang, R.M. Wynn, D.T. Chuang, M.V. Norgard, M.K. Rosen, and C.A. Brautigam, Determination of protein complex stoichiometry through multisignal sedimentation velocity experiments. Anal. Biochem. 407 (2010) 89-103.

FINIS