Analysing binding stoichiometries in NMR titration experiments using Monte Carlo simulation and resampling techniques

Supporting Information

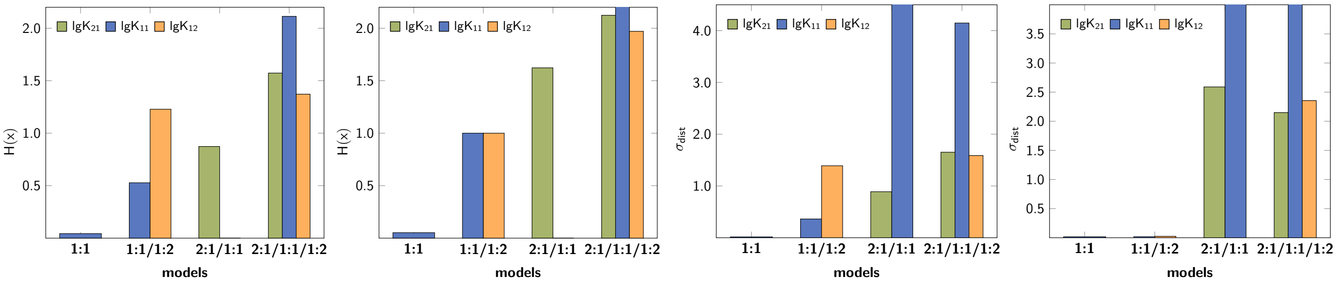
Conrad Hübler

July 28, 2022

Table of Content

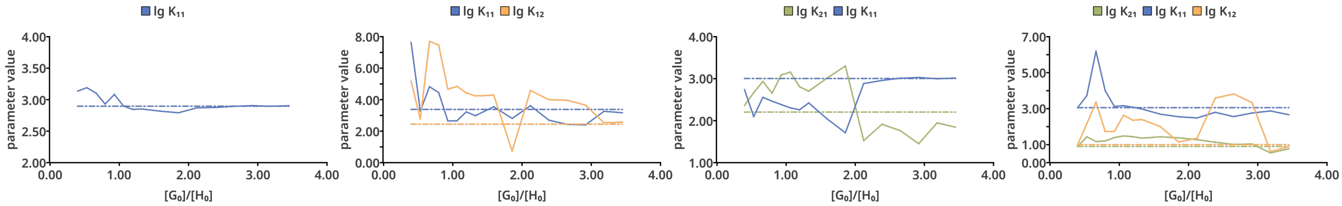
- Collected results of the statistical post-processing for a data set with 1:1 stoichiometry (Figure S1 - Figure S5)
- Collected results of the statistical post-processing for a data set with 1:1/1:2 stoichiometry (Figure S6 - Figure S9)
- Collected results of the statistical post-processing for a data set with 2:1/1:1 stoichiometry (Figure S10 - Figure S14)
- Collected results of the statistical post-processing for a data set with 2:1/1:1/1:2 stoichiometry (Figure S15 - Figure S19)
- Results of the Factor Analysis of the data sets using Sivvu.org (Figure S20)
- Usage of raw data

1 Data set with 1:1 stoichiometry



(a) H(x), S = 2000 (b) H(x), S = 10000 (c) σ, S = 2000 (d) σ, S = 10000

Figure S1: Shannon entropy and standard deviation of the stability constants after Monte Carlo simulation on top of the best-fit models on a simulated data set with 1:1 stoichiometry.



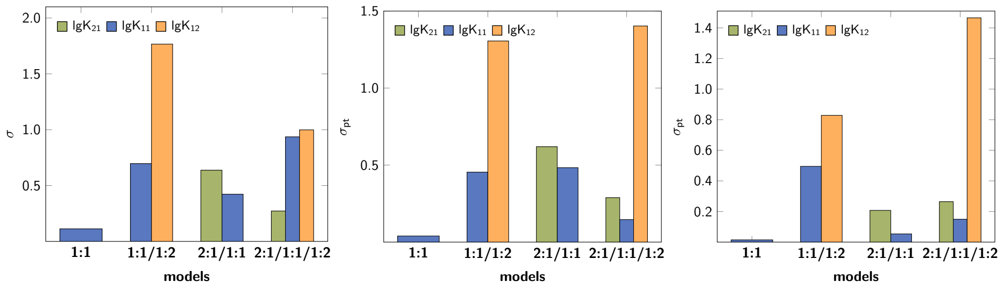
(a) RA of **1:1 model**

(b) RA of **1:1/1:2 model**

(c) RA of **2:1/1:1 model**

(d) RA of **2:1/1:1/1:2 model**

Figure S2: Individual results of the Reduction Analysis for each tested model on a simulated data set with 1:1 stoichiometry.

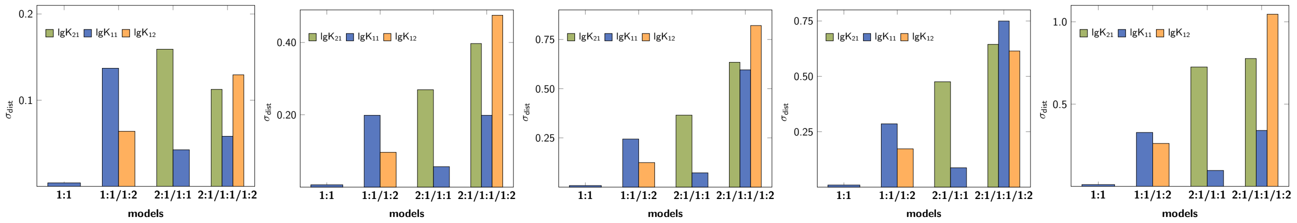


(a) cut-off ratio = 0

(b) cut-off ratio = 1.8

(c) cut-off ratio = 2.0

Figure S3: Calculated σ and σ_{pt} values using different cut-off ratios after Reduction Analyses on top of the best-fit models on a simulated data set with 1:1 stoichiometry.



(a) X = 1

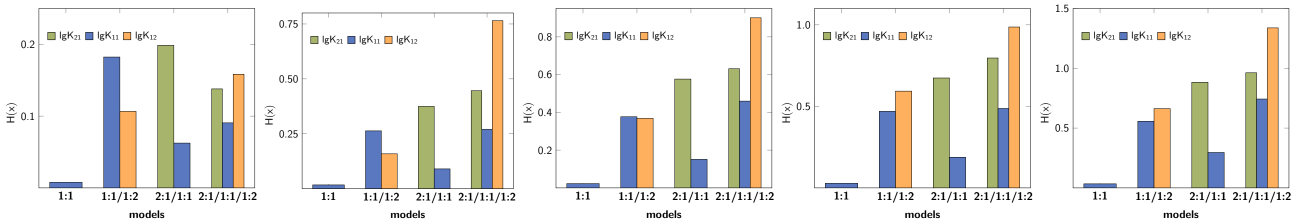
(b) X = 2

(c) X = 3

(d) X = 4

(e) X = 5

Figure S4: Shannon entropy of the histograms obtained for the distribution of individual the stability constants after Cross Validation (CXO) for each model on a simulated data set with 1:1 stoichiometry.



(a) X = 1

(b) X = 2

(c) X = 3

(d) X = 4

(e) X = 5

Figure S5: Standard deviation of the histograms obtained for the distribution of individual the stability constants after Cross Validation (CXO) for each model on a simulated data set with 1:1 stoichiometry.

2 Data set with 1:1/1:2 stoichiometry

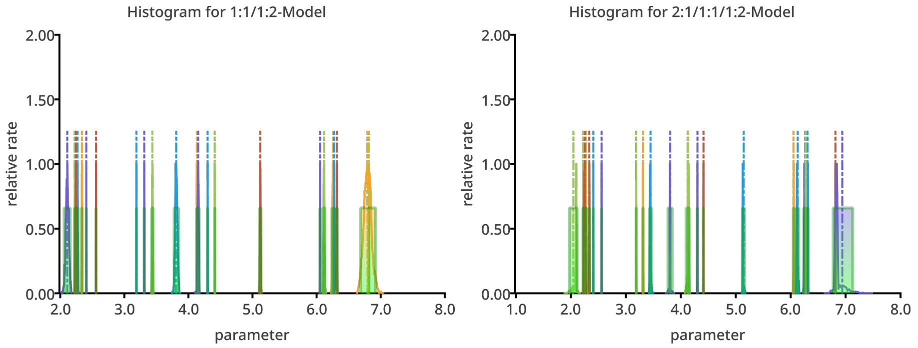
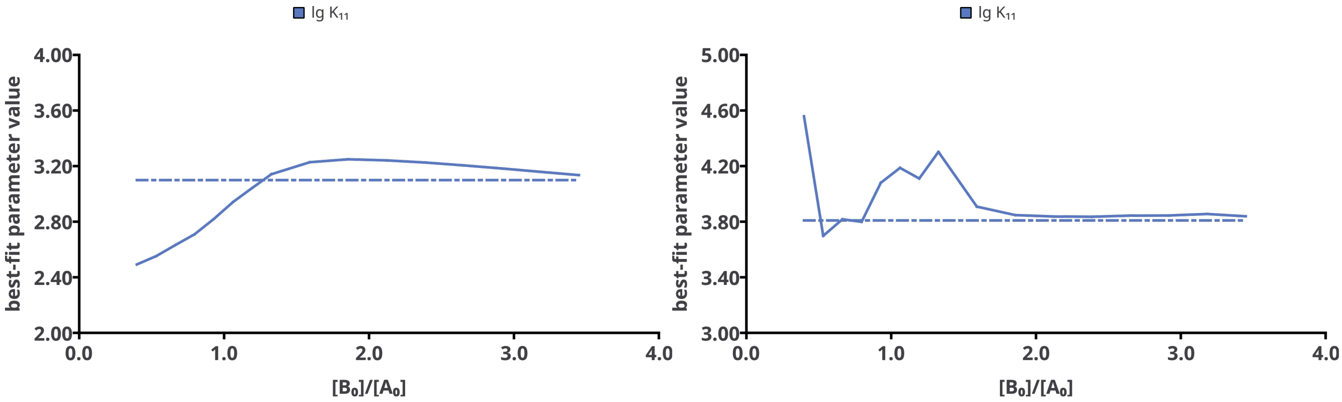


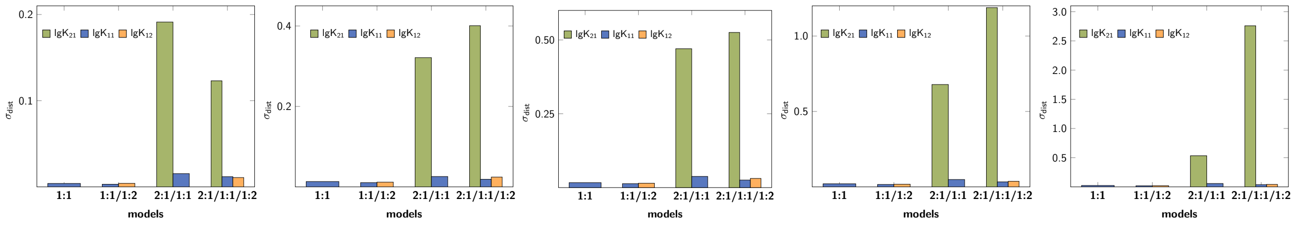
Figure S6: Visual comparison of the shape of the histograms for all parameters involved in 1:1/1:2 model calculation for both tested **1:1/1:2** and **2:1/1:1/1:2 model** show quite narrow distribution. In the current example, correct model parameters in an overall unsuited model may behave differently than the overall incorrect model parameters.



(a) RA of **1:1 model**

(b) RA of **1:1/1:2 model**

Figure S7: Results of the Reduction Analysis for the **1:1** and **1:1/1:2 model** with identical scaling and focusing on the lg K_{11}.



(a) X = 1

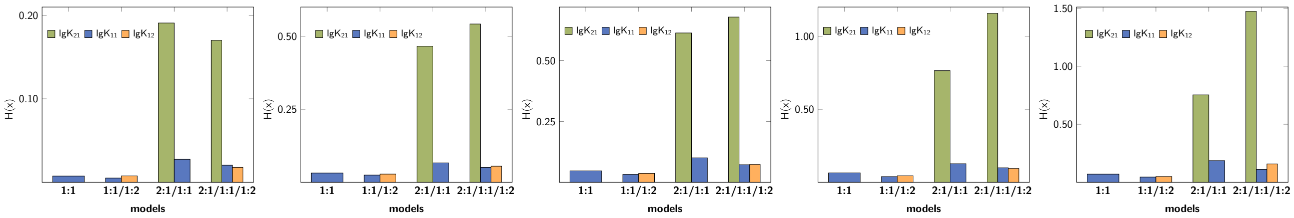
(b) X = 2

(c) X = 3

(d) X = 4

(e) X = 5

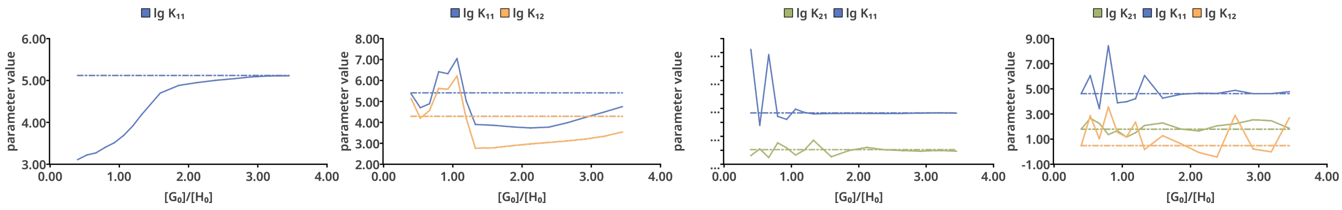
Figure S8: Standard deviation of the histograms obtained for the distribution of individual the stability constants after Cross Validation (CXO) for each model on a simulated data set with 1:1/1:2 stoichiometry.



(a) X = 1 (b) X = 2 (c) X = 3 (d) X = 4 (e) X = 5

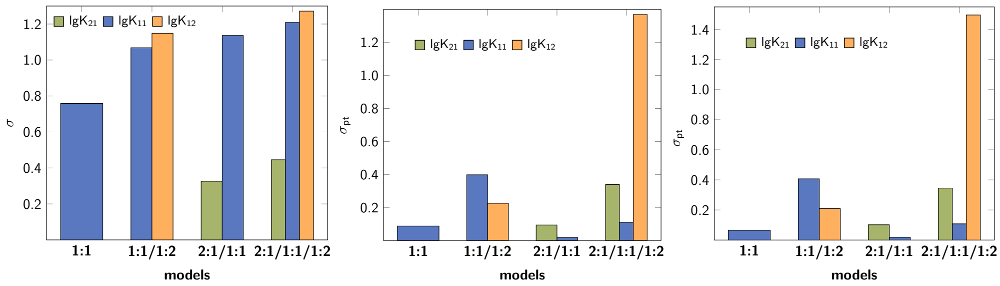
Figure S9: Standard deviation of the histograms obtained for the distribution of individual the stability constants after Cross Validation (CXO) for each model on a simulated data set with 1:1/1:2 stoichiometry.

3 Data set with 2:1/1:1 stoichiometry



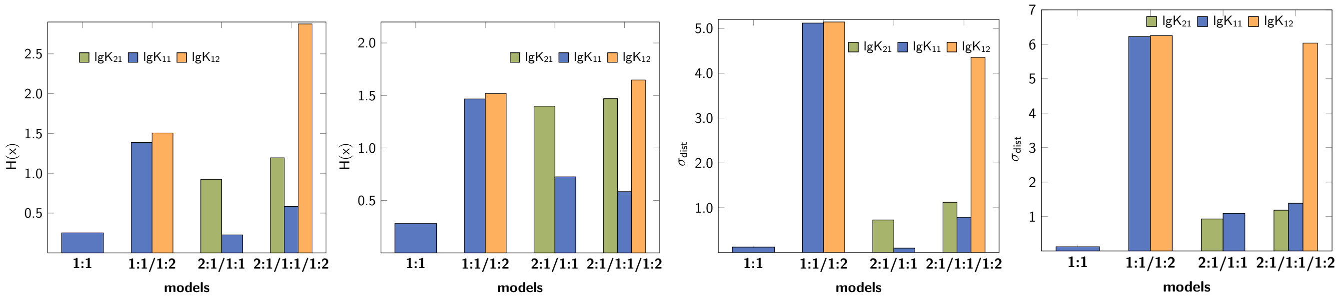
(a) RA of **1:1 model** (b) RA of **1:1/1:2 model** (c) RA of **2:1/1:1 model** (d) RA of **2:1/1:1/1:2 model**

Figure S10: Individual results of the Reduction Analysis for each tested model on a simulated data set with 2:1/1:1 stoichiometry.



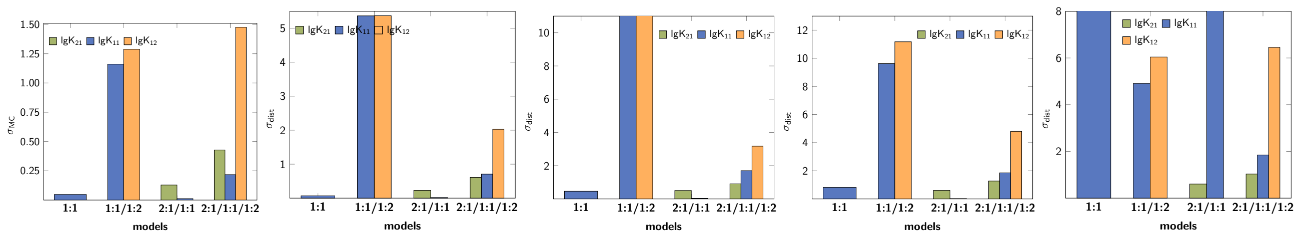
(a) cut-off ratio = 0 (b) cut-off ratio = 1.8 (c) cut-off ratio = 2.0

Figure S11: Calculated σ and σ_{pt} values using different cut-off ratios after Reduction Analyses on top of the best-fit models on a simulated data set with 2:1/1:1 stoichiometry.



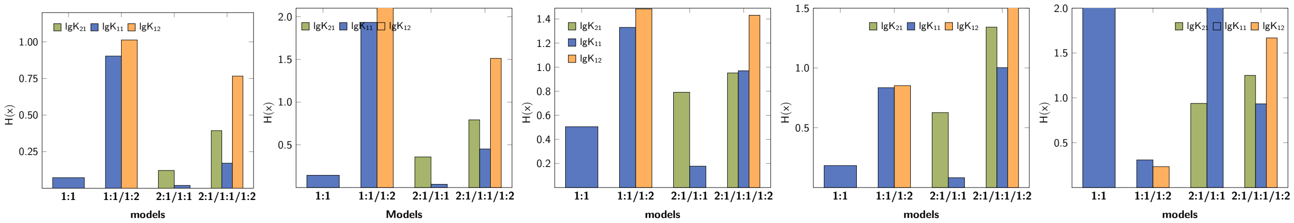
(a) H(x), S = 2000 (b) H(x), S = 10000 (c) σ, S = 2000 (d) σ, S = 10000

Figure S12: Shannon entropy and standard deviation of the stability constants after Monte Carlo simulation on top of the best-fit models on a simulated data set with 2:1/1:1 stoichiometry.



(a) X = 1 (b) X = 2 (c) X = 3 (d) X = 4 (e) X = 5

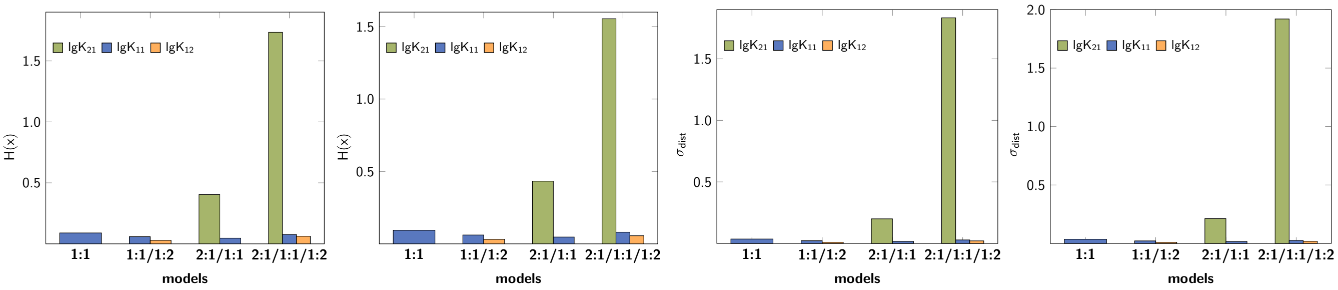
Figure S13: Standard deviation of the histograms obtained for the distribution of individual the stability constants distribution after Cross Validation (CXO) for each model.



(a) X = 1 (b) X = 2 (c) X = 3 (d) X = 4 (e) X = 5

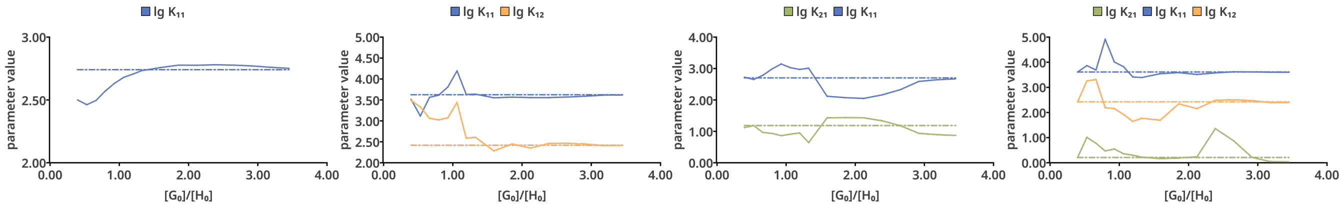
Figure S14: Standard deviation of the histograms obtained for the distribution of individual the stability constants distribution after Cross Validation (CXO) for each model.

4 Data set with 2:1/1:1/1:2 stoichiometry



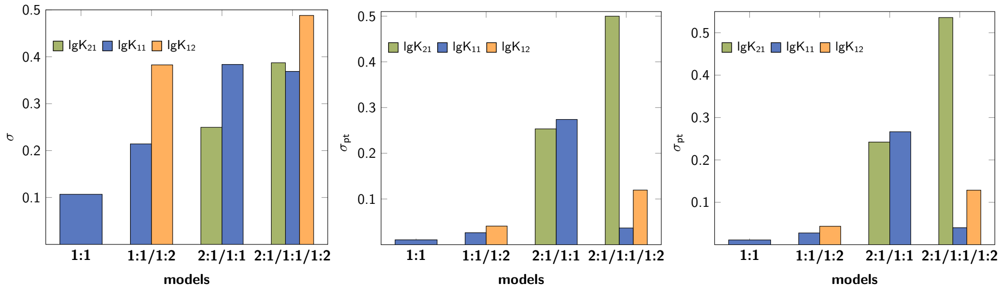
(a) H(x), S = 2000
(b) H(x), S = 10000
(c) σ, S = 2000
(d) σ, S = 10000

Figure S15: Shannon entropy and standard deviation of the stability constants after Monte Carlo simulation on top of the best-fit models on a simulated data set with 2:1/1:1/1:2 stoichiometry.



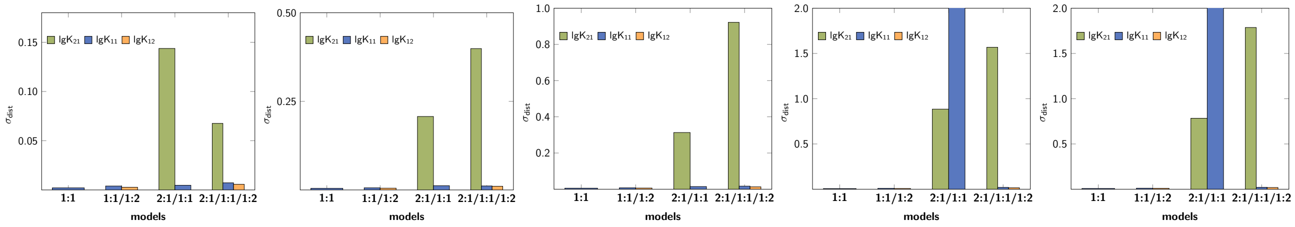
(a) RA of **1:1 model**
(b) RA of **1:1/1:2 model**
(c) RA of **2:1/1:1 model**
(d) RA of **2:1/1:1/1:2 model**

Figure S16: Individual results of the Reduction Analysis for each tested model on a simulated data set with 2:1/1:1/1:2 stoichiometry.



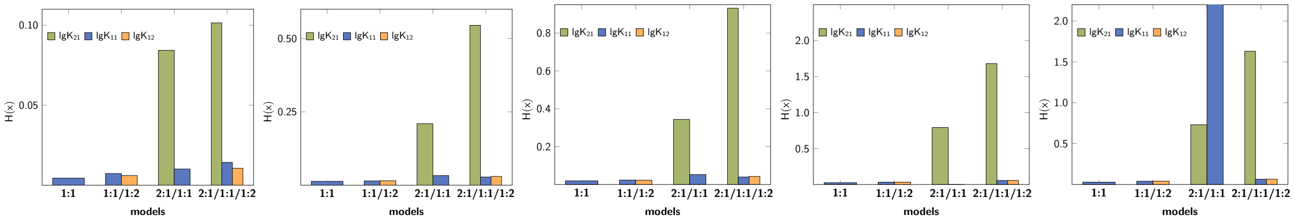
(a) cut-off ratio = 0
(b) cut-off ratio = 1.8
(c) cut-off ratio = 2.0

Figure S17: Calculated σ and σ_{pt} values using different cut-off ratios after Reduction Analyses on top of the best-fit models on a simulated data set with 2:1/1:1/1:2 stoichiometry.



(a) X = 1
(b) X = 2
(c) X = 3
(d) X = 4
(e) X = 5

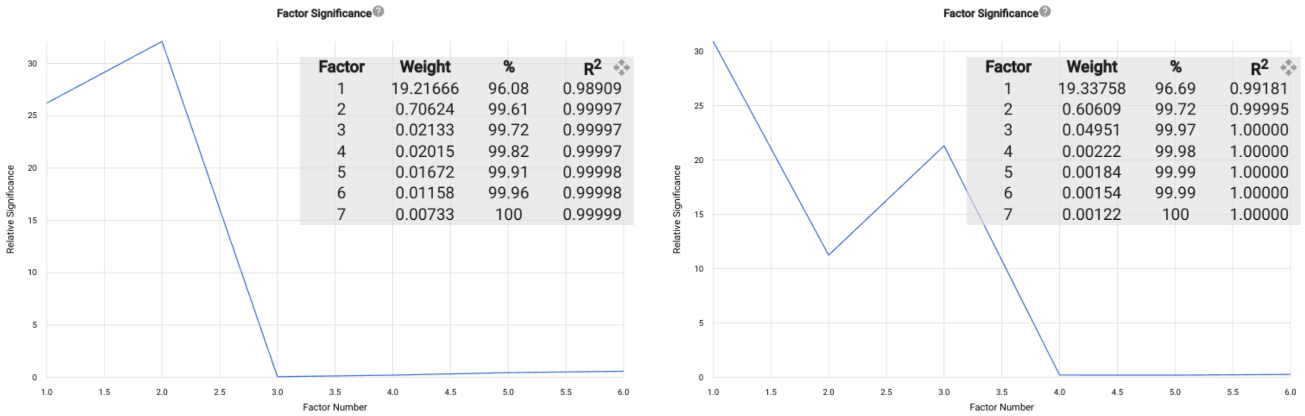
Figure S18: Standard deviation of the histograms obtained for the distribution of individual the stability constants distribution after Cross Validation (CXO) for each model on a simulated data set with 2:1/1:1/1:2 stoichiometry.



(a) X = 1 (b) X = 2 (c) X = 3 (d) X = 4 (e) X = 5

Figure S19: Standard deviation of the histograms obtained for the distribution of individual the stability constants distribution after Cross Validation (CXO) for each model on a simulated data set with 2:1/1:1/1:2 stoichiometry.

5 Factor analysis using Sivvu.org

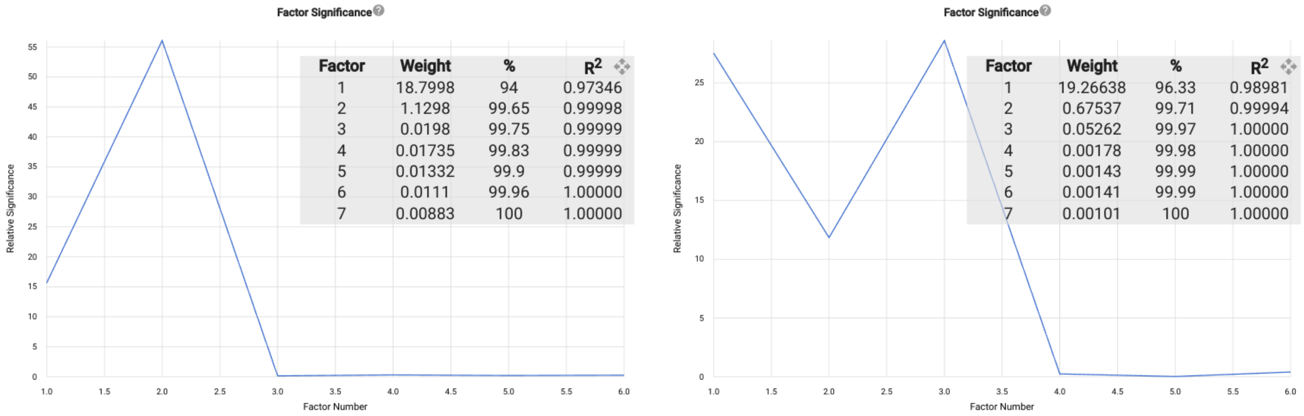


Factor	Weight	%	R^2
1	19.21666	96.08	0.98909
2	0.70624	99.61	0.99997
3	0.02133	99.72	0.99997
4	0.02015	99.82	0.99997
5	0.01672	99.91	0.99998
6	0.01158	99.96	0.99998
7	0.00733	100	0.99999

Factor	Weight	%	R^2
1	19.33758	96.69	0.99181
2	0.60609	99.72	0.99995
3	0.04951	99.97	1.00000
4	0.00222	99.98	1.00000
5	0.00184	99.99	1.00000
6	0.00154	99.99	1.00000
7	0.00122	100	1.00000

(a) 1:1 data set

(b) 1:1/1:2 data set



Factor	Weight	%	R^2
1	18.7998	94	0.97346
2	1.1298	99.65	0.99998
3	0.0198	99.75	0.99999
4	0.01735	99.83	0.99999
5	0.01332	99.9	0.99999
6	0.0111	99.96	1.00000
7	0.00883	100	1.00000

Factor	Weight	%	R^2
1	19.26638	96.33	0.98981
2	0.67537	99.71	0.99994
3	0.05262	99.97	1.00000
4	0.00178	99.98	1.00000
5	0.00143	99.99	1.00000
6	0.00141	99.99	1.00000
7	0.00101	100	1.00000

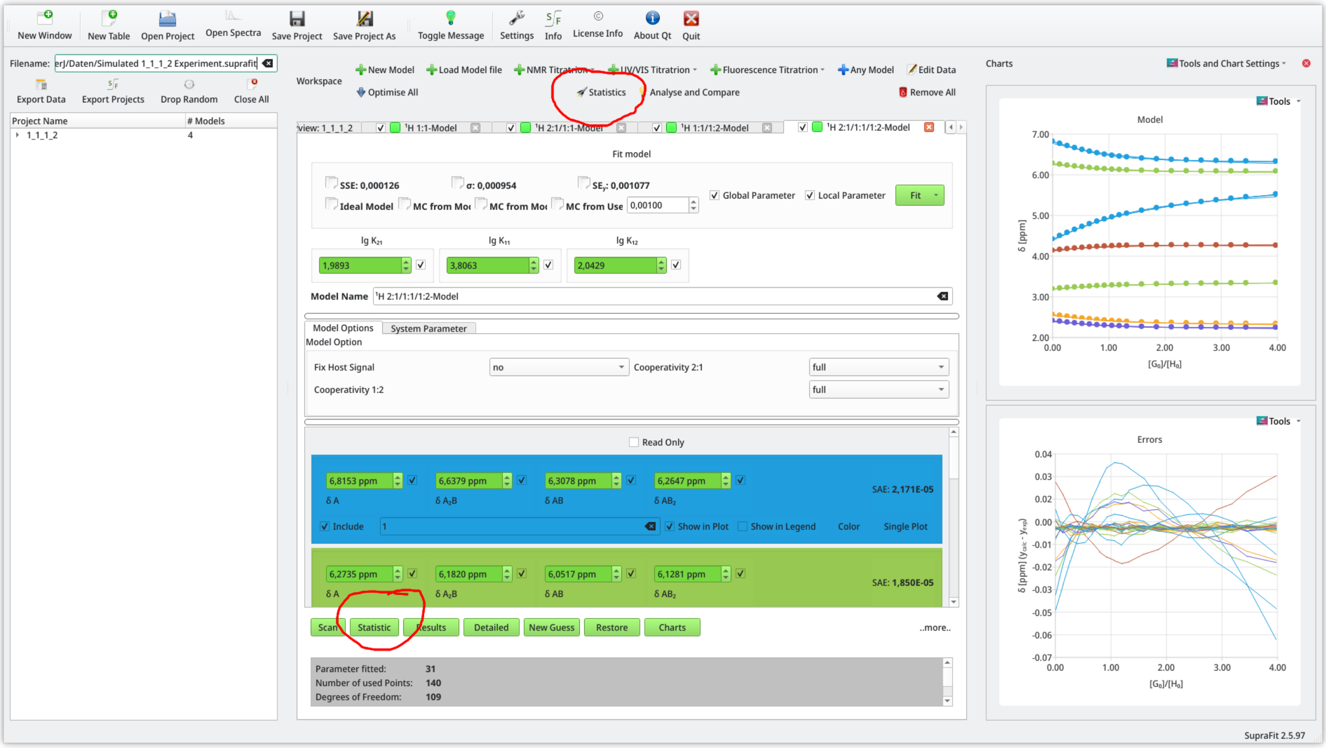
(c) 2:1/1:1 data set

(d) 2:1/1:1/1:2 data set

Figure S20: Results of the factor analysis applied to the data sets according to the implementation using Sivvu.org (D.A. Vander Griend and M. J. DeVries and M. Greeley and Y. Kim and N. Wang and D. Buist and C. Ulry (2021). Sivvu, http://sivvu.org.)

6 Usage of raw data

Raw data and SupraFit binaries for Windows, macOS and Linux are deposited at 10.5281/zenodo.6539577. While the file called *Simulated Experiments.suprafit* can be opened with the stable release of SupraFit (2.0 https://github.com/conradhuebler/SupraFit/releases/tag/2.0.0), the remaining files require a new version, e.g. the one included in the zenodo upload. However, the protocol can be used with SupraFit 2.0 as well. The Monte Carlo simulation results for the 1:1/1:2 model presented in Figure 4 are stored in the file *MC-simulation 1_1_1_2 Experiment.suprafit*.



Filename: erJ/Daten/Simulated 1_1_1_2 Experiment.suprafit

Project Name	# Models
1_1_1_2	4

view: 1_1_1_2 | ☑ ^{1}H 1:1-Model | ☑ ^{1}H 2:1/1:1-Model | ☑ ^{1}H 1:1/1:2-Model | ☑ ^{1}H 2:1/1:1/1:2-Model

☐ SSE: 0,000126 ☐ σ: 0,000954 ☐ SE_y: 0,001077

☐ Ideal Model ☐ MC from Mo ☐ MC from Mo ☐ MC from Use 0,00100 ☑ Global Parameter ☑ Local Parameter

lg K_{21}: 1,9893 ☑ lg K_{11}: 3,8063 ☑ lg K_{12}: 2,0429 ☑

Model Name: ^{1}H 2:1/1:1/1:2-Model

Fix Host Signal: no
Cooperativity 2:1: full
Cooperativity 1:2: full

δ A: 6,8153 ppm ☑ δ A_2B: 6,6379 ppm ☑ δ AB: 6,3078 ppm ☑ δ AB_2: 6,2647 ppm ☑ SAE: 2,171E-05
☑ Include 1 ☑ Show in Plot ☐ Show in Legend

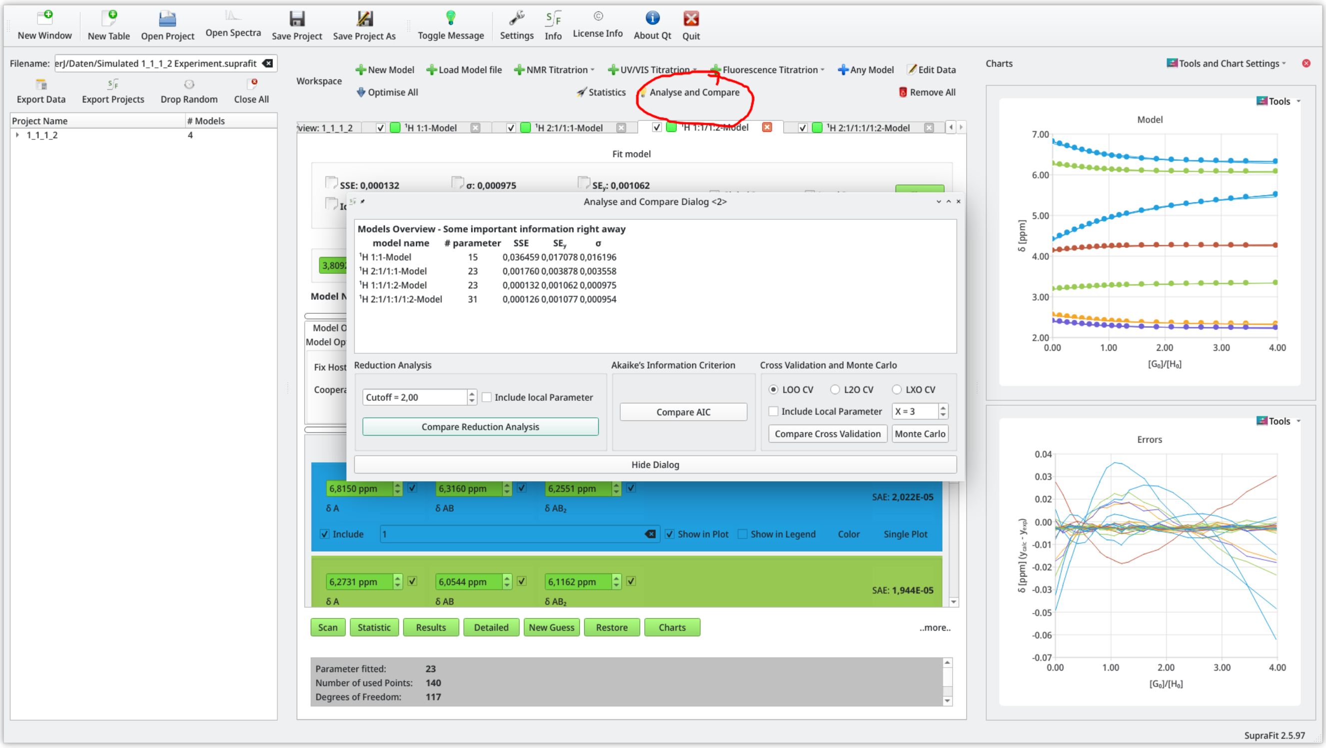
δ A: 6,2735 ppm ☑ δ A_2B: 6,1820 ppm ☑ δ AB: 6,0517 ppm ☑ δ AB_2: 6,1281 ppm ☑ SAE: 1,850E-05

Parameter fitted: 31
Number of used Points: 140
Degrees of Freedom: 109

Figure S21: Statistical post-processing can be started for all open models using *statistics* button (with the rocket) or for the current model only using the green *Statistic* button.

The dialog shown in Figure S21 can be used for Reduction Analysis, Cross Validation and Monte Carlo simulation. SupraFit takes care, that only compatible post-processing analyses results are compared, e.g. only L(X=3)O Cross Validation results. In case of Monte Carlo simulation, this is not always possible. Hence, the number assigned to X represents the index of the Monte Carlo simulation. If these were started using the *Statistics* button (rocket), the indices should be fine (unless selected results were deleted). If however, Monte Carlo simulation were started for each model individually and, for example for the first model a simulation with S=2000 and then the next with S=4000 and for the second model vice versa, the indices should not fit.

Results for a single model can be viewed in detail using the *Results* button (Figure S22 and Figure S23).



Filename: erJ/Daten/Simulated 1_1_1_2 Experiment.suprafit

Project Name	# Models
1_1_1_2	4

SSE: 0,000132 σ: 0,000975 SE_y: 0,001062

Analyse and Compare Dialog <2>

Models Overview - Some important information right away

model name	# parameter	SSE	SE_y	σ
^{1}H 1:1-Model	15	0,036459	0,017078	0,016196
^{1}H 2:1/1:1-Model	23	0,001760	0,003878	0,003558
^{1}H 1:1/1:2-Model	23	0,000132	0,001062	0,000975
^{1}H 2:1/1:1/1:2-Model	31	0,000126	0,001077	0,000954

Reduction Analysis: Cutoff = 2,00 ☐ Include local Parameter — Compare Reduction Analysis

Akaike's Information Criterion: Compare AIC

Cross Validation and Monte Carlo: ● LOO CV ○ L2O CV ○ LXO CV ☐ Include Local Parameter X = 3 — Compare Cross Validation — Monte Carlo

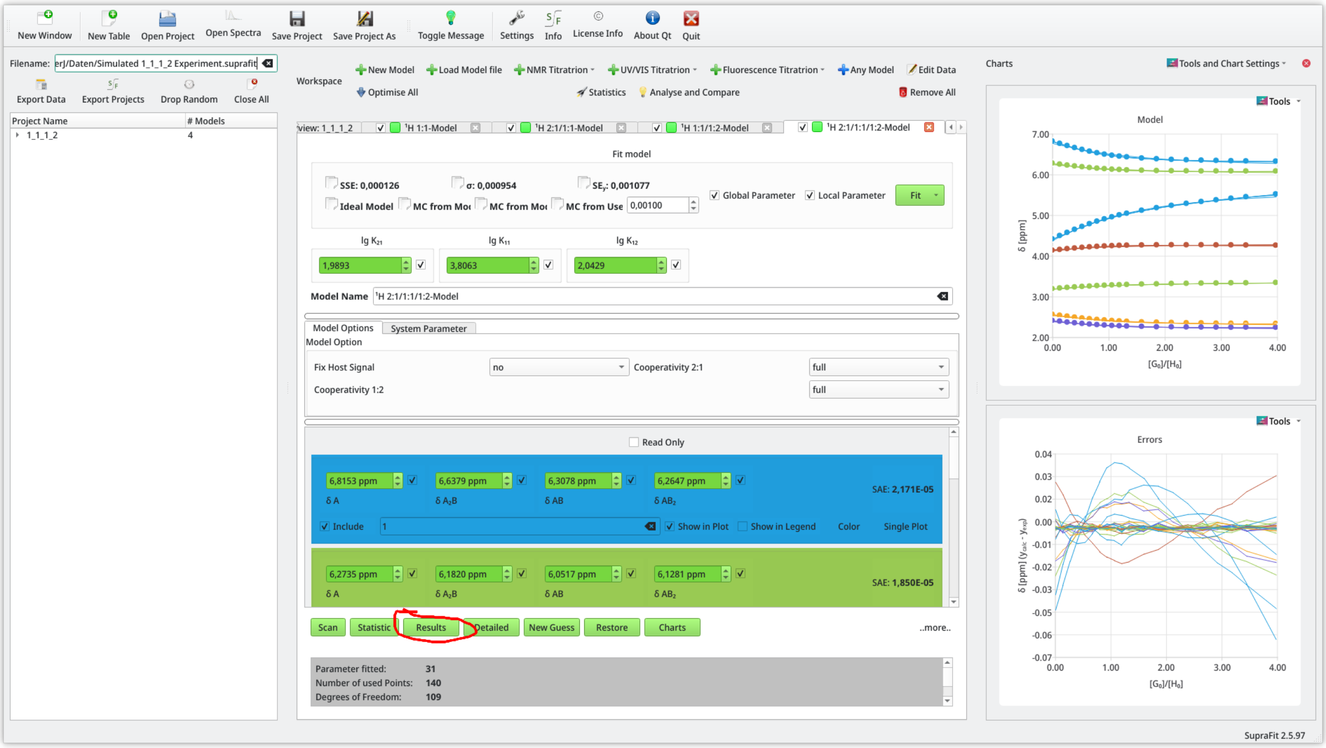
6,8150 ppm δ A; 6,3160 ppm δ AB; 6,2551 ppm δ AB_2 — SAE: 2,022E-05

☑ Include 1 ☑ Show in Plot ☐ Show in Legend Color Single Plot

6,2731 ppm δ A; 6,0544 ppm δ AB; 6,1162 ppm δ AB_2 — SAE: 1,944E-05

Parameter fitted: 23
Number of used Points: 140
Degrees of Freedom: 117

Figure S22: The post-processing dialog of all models can be opened using the *Analyse and Compare* button (the one with the light bulb). Using the dialog, results of Monte Carlo simulation, Reduction Analysis and Cross Validation can be compared.



Filename: erJ/Daten/Simulated 1_1_1_2 Experiment.suprafit

Project Name	# Models
1_1_1_2	4

Fit model

SSE: 0,000126 σ: 0,000954 SE_y: 0,001077

☐ Ideal Model ☐ MC from Mo ☐ MC from Mo ☐ MC from Use 0,00100 ☑ Global Parameter ☑ Local Parameter

lg K_{21}: 1,9893 lg K_{11}: 3,8063 lg K_{12}: 2,0429

Model Name ^{1}H 2:1/1:1/1:2-Model

Model Options | System Parameter

Model Option

Fix Host Signal: no Cooperativity 2:1: full

Cooperativity 1:2: full

☐ Read Only

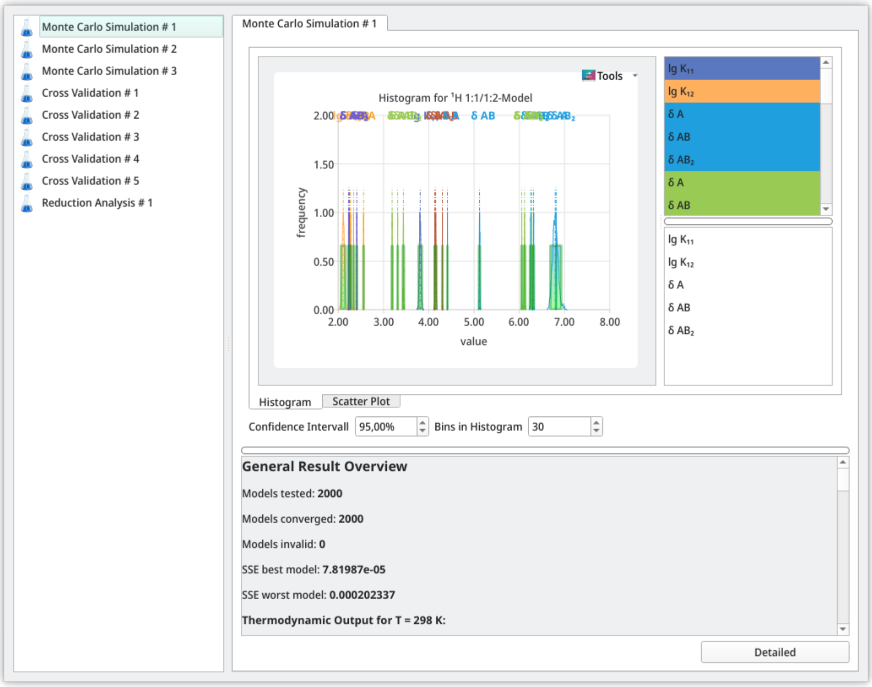
6,8153 ppm δ A; 6,6379 ppm δ A_2B; 6,3078 ppm δ AB; 6,2647 ppm δ AB_2 — SAE: 2,171E-05

☑ Include 1 ☑ Show in Plot ☐ Show in Legend Color Single Plot

6,2735 ppm δ A; 6,1820 ppm δ A_2B; 6,0517 ppm δ AB; 6,1281 ppm δ AB_2 — SAE: 1,850E-05

Parameter fitted: 31
Number of used Points: 140
Degrees of Freedom: 109

Figure S23: Results for individual models can be viewed using the green *Results* button.



Confidence Intervall 95,00% Bins in Histogram 30

General Result Overview

Models tested: **2000**

Models converged: **2000**

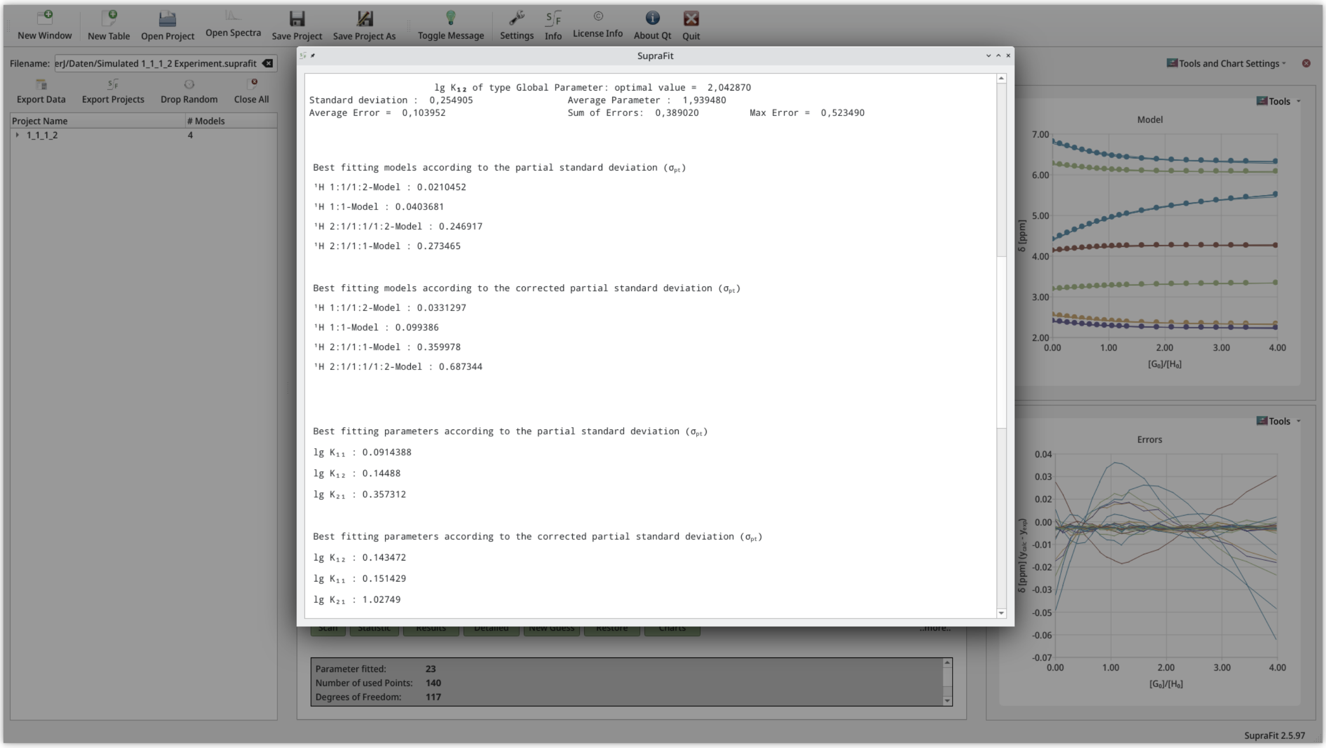
Models invalid: **0**

SSE best model: **7.81987e-05**

SSE worst model: **0.000202337**

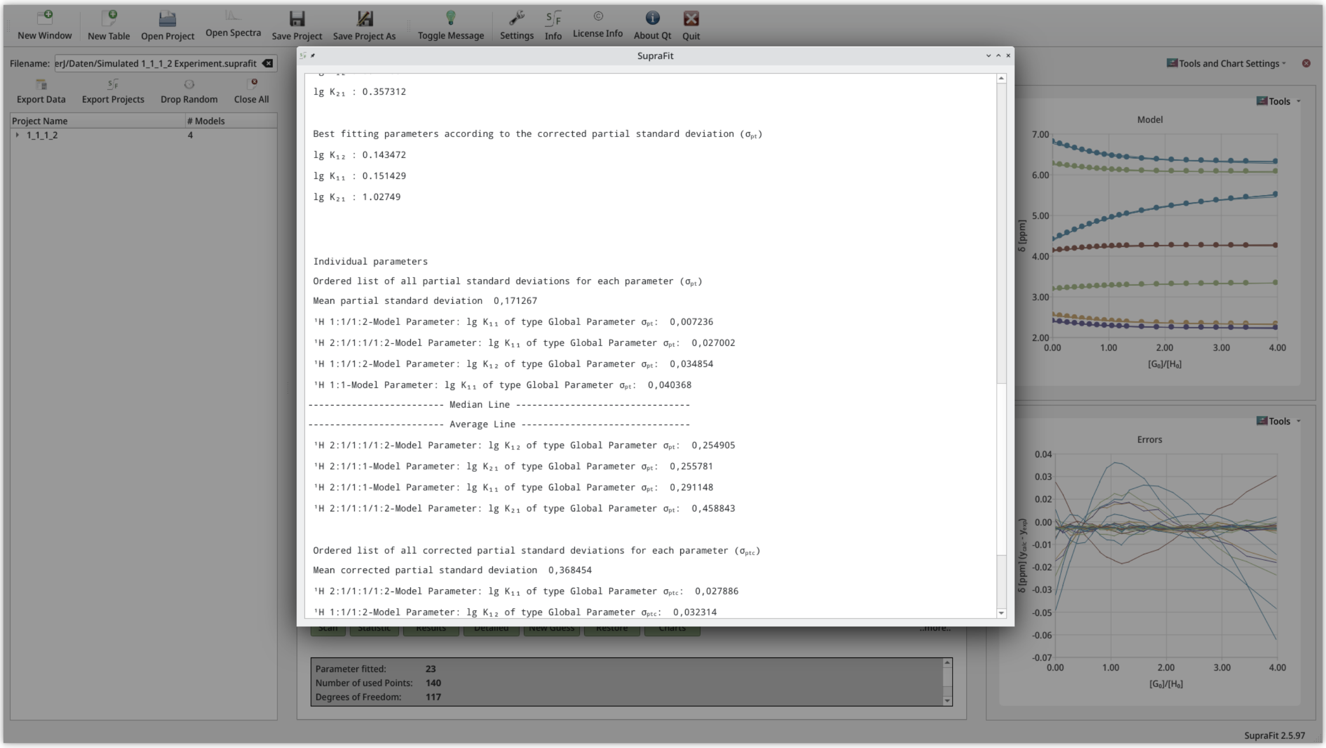
Thermodynamic Output for T = 298 K:

Figure S24: The individual results are presented as charts and text output.



```
                    lg K₁₂ of type Global Parameter: optimal value =  2,042870
Standard deviation :  0,254905                 Average Parameter :  1,939480
Average Error =  0,103952                      Sum of Errors:  0,389020         Max Error =  0,523490



 Best fitting models according to the partial standard deviation (σ_pt)

 ¹H 1:1/1:2-Model : 0.0210452

 ¹H 1:1-Model : 0.0403681

 ¹H 2:1/1:1/1:2-Model : 0.246917

 ¹H 2:1/1:1-Model : 0.273465


 Best fitting models according to the corrected partial standard deviation (σ_pt)

 ¹H 1:1/1:2-Model : 0.0331297

 ¹H 1:1-Model : 0.099386

 ¹H 2:1/1:1-Model : 0.359978

 ¹H 2:1/1:1/1:2-Model : 0.687344



 Best fitting parameters according to the partial standard deviation (σ_pt)

 lg K₁₁ : 0.0914388

 lg K₁₂ : 0.14488

 lg K₂₁ : 0.357312


 Best fitting parameters according to the corrected partial standard deviation (σ_pt)

 lg K₁₂ : 0.143472

 lg K₁₁ : 0.151429

 lg K₂₁ : 1.02749
```

Parameter fitted: 23
Number of used Points: 140
Degrees of Freedom: 117



```
 lg K₂₁ : 0.357312


 Best fitting parameters according to the corrected partial standard deviation (σ_pt)

 lg K₁₂ : 0.143472

 lg K₁₁ : 0.151429

 lg K₂₁ : 1.02749



 Individual parameters

 Ordered list of all partial standard deviations for each parameter (σ_pt)

 Mean partial standard deviation  0,171267

 ¹H 1:1/1:2-Model Parameter: lg K₁₁ of type Global Parameter σ_pt:  0,007236

 ¹H 2:1/1:1/1:2-Model Parameter: lg K₁₁ of type Global Parameter σ_pt:  0,027002

 ¹H 1:1/1:2-Model Parameter: lg K₁₂ of type Global Parameter σ_pt:  0,034854

 ¹H 1:1-Model Parameter: lg K₁₁ of type Global Parameter σ_pt:  0,040368
------------------------ Median Line ------------------------------
------------------------ Average Line -----------------------------
 ¹H 2:1/1:1/1:2-Model Parameter: lg K₁₂ of type Global Parameter σ_pt:  0,254905

 ¹H 2:1/1:1-Model Parameter: lg K₂₁ of type Global Parameter σ_pt:  0,255781

 ¹H 2:1/1:1-Model Parameter: lg K₁₁ of type Global Parameter σ_pt:  0,291148

 ¹H 2:1/1:1/1:2-Model Parameter: lg K₂₁ of type Global Parameter σ_pt:  0,458843


 Ordered list of all corrected partial standard deviations for each parameter (σ_ptc)

 Mean corrected partial standard deviation  0,368454

 ¹H 2:1/1:1/1:2-Model Parameter: lg K₁₁ of type Global Parameter σ_ptc:  0,027886

 ¹H 1:1/1:2-Model Parameter: lg K₁₂ of type Global Parameter σ_ptc:  0,032314
```

Parameter fitted: 23
Number of used Points: 140
Degrees of Freedom: 117

Figure S25: Results of the Reduction Analyses for four models on top of a simulated NMR titration experiment with 1:1/1:2 stoichiometry.