Component Quantification in Solids with the Mixture Analysis Using References Method

Dirk Stueber* and Zachary E. X. Dance*

Department of Analytical Research and Development, Merck Research Laboratories, Merck & Co., Inc., Rahway, New Jersey, 07065, United States

*Corresponding Authors: dirk_stueber@merck.com and zachary_dance@merck.com

Supporting Information

Table of Content

Supporting Material

Section S1. Application of MAR to processed samples S4

Section S2. Limit of detection and quantitation of the MAR method S4

Section S3. MAR comparison with DECRA S5

Section S4. Description of blend preparation S6

Supporting Figures

Figure S1. The three steps of the MAR analysis protocol illustrated for a hypothetical two-component mixture using simulated spectra S7

Figure S2. ^{19}F CPMAS pure component and binary blends spectra S9

Figure S3. ^{19}F relMAR fit for the 31.9 mol%-LevoHH blend from Table 1 S10

Figure S4. Results of the truMAR analysis for all nine Lanso/LevoHH blends from Table 1 as correlation plots S11

Figure S5. ^{13}C and ^{19}F Lanso, LevoHH, and LevoMH pure component, and corresponding ternary blend CPMAS spectra S12

Figure S6. The ^{13}C relMAR and truMAR results of a spectral subsection MAR analysis for all nine binary Lanso/LevoHH blends from Table 1 as correlation plots S13

Figure S7. ^{13}C spectral subsection relMAR fit of the 39.1 mol% LevoHH blend S14

Figure S8. Application of conventional CP-dynamic corrected ssNMR quantification approach to present {^{1}H-^{19}F} CPMAS data S15

Figure S9. Lanso ^{19}F CPMAS spectra as a function of contact time S16

Figure S10. LevoHH ^{19}F CPMAS spectra as a function of contact time S17

Figure S11. LevoMH ^{19}F CPMAS spectra as a function of contact time S18

Figure S12. ^{13}C CPMAS spectra of the pure reference compounds and the corresponding binary blends for the API-FI/API-FII and Ibu/Indo model systems S19

Figure S13. truMAR and relMAR fits of the binary API-FI/API-FII and Ibu/Indo blends S20

Supporting Tables

Table S1. Results of CP-curve Analyses S21

Table S2. Relevant experimental ssNMR parameters used for the Ibu/Indo and API-FI/API-FII model systems S21

Supporting References S21

Section S1. Application of MAR to processed samples.

As noted in the manuscript, a critical requirement for the applicability of MAR is that the pure component species retain their respective line widths and shapes, and T_1 relaxation times, in the mixtures to be analyzed (hereafter called the representative-condition). There are several sample processing techniques used in the pharmaceutical industry, eg, milling and compacting, that could potentially render this assumption invalid. However, in our experience, the assumption that the pure component spectra are still representative for the corresponding components in the mixture holds in the majority of cases. In situations in which the representative-condition does not hold, line width and shape changes are usually observed in conjunction with a change in T_1, and accordingly are easily detectable in the respective MAR fitting residuals. In these cases, it might be possible to prepare more appropriate MAR references and complete the analysis, or MAR might not be applicable and other quantification methods must be applied.

Generally, if there is any indication that the representative-condition is not fulfilled, care must be taken in applying MAR, and additional experimental work can be performed to test the condition. For example, test samples of pure components and mixtures thereof can be prepared by exposing these samples to the relevant processing protocols. Line width and shape and T_1 analyses conducted before and after processing reveals if the representative-condition is valid. Another approach to test the representative-condition is to complete two quantification analyses according to the truMAR methodology (relatively short experiment times) utilizing two different pulse delays (but both pulse delays in the truMAR regime). If the representative-condition is valid, both analyses will predict the same composition. One important criterium for evaluating the representative-condition for MAR, should be the required accuracy of the component quantification for the system under investigation. It is a useful practice to apply the MAR analysis to test blends with known compositions, which, ideally were exposed to the relevant processing conditions and determine if MAR reproduces the compositions with the required accuracy.

Section S2. Limit of detection (LOD) and quantitation (LOQ) of the MAR method.

LOD and LOQ can be determined in a variety of ways.[1, 2] Generally, it is not possible to predict the LOD and LOQ a priori for MAR due to the significant spectral overlap typically observed. One of the strengths of MAR is the capability to quantify components of similar structure, eg, API polymorphs. Naturally, chemical shift differences between such components are small and their corresponding spectra frequently exhibit significant spectral overlap. In favorable cases, in which base-line resolution of relevant peaks is observed, accepted definitions based on the signal-to-noise ratios of the spectra can be used to obtain estimates of LODs and LOQs. However, in order to conduct a rigorous determination of LODs/LOQs with statistical confidence it is necessary to use standard procedures[1, 2] to experimentally build a calibration curve in the range of interest for a given system using the MAR analysis.

Section S3. MAR comparison with DECRA.

Among the chemometric ssNMR tools to quantify mixture components, the direct exponential curve resolution algorithm (DECRA) has been successfully applied to a wide variety of materials, including pharmaceuticals, polymers, and human brain samples. Theoretically, the most attractive benefits of DECRA are that the spectra and T_1 relaxation times of the pure components do not have to be known in advance but, rather, are produced directly by the analysis. The DECRA protocol also provides a validation or rejection of the user-specified number of pure components. Accordingly, DECRA has the potential to reveal unknown phases or impurities in addition to determining the relative amounts of the mixture components, and their respective spectra and T_1s. Principally, the technique is insensitive to spectral overlap. However, mixture component quantification with DECRA also exhibits several significant challenges. The method displays a low tolerance for low signal-to-noise data and experimental artifacts arising from instrumental instabilities which vary between the time increments within a given data set. These artifacts include baseline and phase distortions, as well as departure from exponential behavior, and become increasingly critical for analyzing mixtures containing components with similar T_1 relaxation times. Another intrinsic downside of DECRA is the fact that experiments are performed with pulse delays of five times the T_1 of the slowest relaxing component (5 T_1(long)), which, based on the frequently observed long T_1 relaxation times in solids, may easily result in impractically long experiment times. One of the main considerations for setting up DECRA experiments centers around the appropriate choice of recovery time points based on the T_1 values of the mixture components. For DECRA to yield accurate results, it is required to collect inversion recovery data with equally-spaced time points that (1) extend long enough to define 5 T_1(long), and (2) are simultaneously close enough to resolve the components with the most similar T_1s.[3] These sampling requirements impose significant challenges on the usefulness of the DECRA analysis.

In our experience, the applicability of DECRA to the analysis of our pharmaceutical API systems with ssNMR is limited due to the experimental challenges described above, and the fact that some of its benefits have limited impact. The various API forms naturally exhibit similar and often lengthy T_1 relaxation times, and largely require the detection of ^{13}C at natural abundance, a low-sensitivity nucleus, placing them into the challenging category of compounds to be analyzed with DECRA. Furthermore, for most of our drug systems, the relevant API phases requiring quantification are isolated and characterized as part of the early development process. Accordingly, at the point when form quantification becomes pertinent, the pure component spectra and the respective T_1 relaxation times are naturally available. As a result, obtaining the pure component spectra with respective T_1s from a DECRA analysis has limited value. In our lab, DECRA is utilized in specific cases in which MAR is not applicable, and in which it is necessary to take advantage of the strengths of the DECRA methodology to complete the desired mixture analysis.

Section S4. Description of blend preparation.

The blends for this work were prepared by using quantities approximately 10 times larger than the actual sample size needed for the ssNMR experiments to ensure blend homogeneity (eg, for a 100 mg sample, the respective blend was prepared on 1 g scale). Samples were combined in the respective proportions in glass vials and homogeneously blended in five mixing cycles, alternating between five minutes on a lab-scale vortex blender and five hours on a rotating wheel sample mixer. The powdered samples used in this work presented as free-flowing powders and did not exhibit obvious mixing problems, eg, adhering to the walls of the glass vials used for blending.

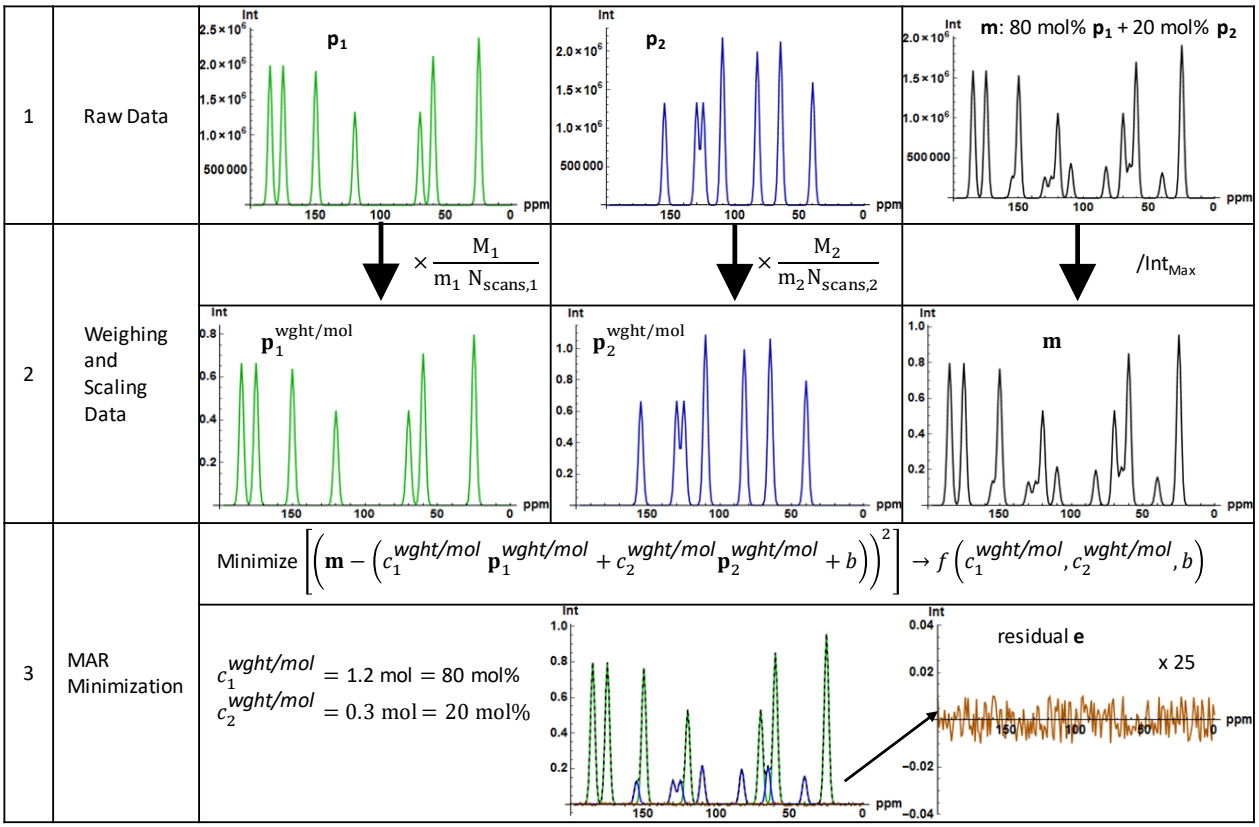


Figure S1. The three steps of the MAR analysis protocol illustrated for a hypothetical two-component mixture using simulated spectra.

1: Step 1 involves the collection of the relevant pure component, $\mathbf{p}_i$, and mixture spectrum, **m**. All spectra are collected and processed using identical NMR parameters (ie, pulse powers, CP conditions, line broadening, etc.). Since the number of scans is part of the weighting (see 2), the number of collected scans for signal averaging, N_{scans}, may vary between the spectra. The relevance of pulse delays for signal averaging is discussed in more detail in the main manuscript.

2: Step 2 consists of weighting the pure component spectra, $\mathbf{P}_i$. Fundamentally, the net magnetization, M^0, which can be detected as an NMR signal observing a spin-1/2 nucleus in a given sample, is proportional to the following variables[4]:

$$M^0 \propto \frac{N_{nuclei}\,\gamma^2\,B_0}{T} \qquad \text{(S1)}$$

N_{nuclei}, γ, B_0, and T, in Equation S1 represent the total number of observed nuclei, the corresponding gyromagnetic ratio of the observed nucleus, the external magnetic field, and the absolute temperature, respectively. When data for a series of pure components and mixture spectra under investigation are collected at the same field and temperature, B_0 and T are rendered constants for all spectra in the series. All γ-values for the spectra in the series are naturally identical, reducing γ in Equation S1 to a constant as well. By combining γ, B_0, and T, and all contributions to signal scaling from instrumental and electronic factors (eg, receiver gain, filters, etc.) into a constant of proportionality, S_0, and introducing the number of scans collected, N_{scans}, the experimentally observed NMR signal, S_{exp}, in arbitrary area units (A) for a given sample can be expressed as:

$$S_{exp} = N_{nuclei}\,N_{scans}\,S_0 \qquad \text{(S2)}$$

S_0 in Equation S2 includes all parameters that are constants for the system under investigation and exhibits the unit of A nucleus^{-1} scans^{-1}. Equation S2 can be expanded to:

$$S_{exp} = N_{nuclei/molecule}\, n_{molecule}\, N_A\, N_{scans}\, S_0 \qquad \text{(S3)}$$

where $N_{nuclei/molecule}$, $n_{molecule}$, and N_A, are the number of nuclei per molecule, the number of moles of the given molecule, and Avogadros's Number, respectively. From Equation S3, S_{exp} can be weighted to represent the area per mole of molecules and per scan ($A\, n_{molecule}{}^{-1}\, scan^{-1}$), S_{exp}^{mol}:

$$S_{exp}^{mol} = \frac{S_{exp}}{N_{scans}\, n_{molecules}} = N_{nuclei/molecule}\ N_A\ S_0 \qquad \text{(S4)}$$

In the MAR quantification analysis, the pure component reference spectra, $\mathbf{p}_i$, are weighted based on Equation S4, yielding:

$$\mathbf{p}_i^{wght/mol} = \frac{\mathbf{p}_i}{n_{sample,i}\, N_{scans}} = \frac{\mathbf{p}_i\, M_i}{m_{sample,i}\, N_{scans}} \qquad \text{(S5)}$$

In Equation S5, $m_{sample,i}$ and M_i are the sample mass and molecular mass of component *i*, respectively. In practice, each pure component vector, $\mathbf{p}_i$, is multiplied by the molecular mass of the respective component and divided by the product of sample mass of component *i* and the number of scans collected for the spectrum of component *i.* When weighting the pure component spectra according to Equation S5, the MAR minimization according to Equation 5 in the main manuscript yields the mixture composition as mole fractions of the components, ie, the c_i exhibit the unit of mole. For practical purposes, the mixture spectrum may be scaled to adjust its spectral area close to the spectral areas of the corresponding pure component spectra, for example dividing it by its maximum intensity. Scaling the mixture spectrum in such a way permits the mathematical minimization procedure, Equation 5 in the main manuscript, to reach a minimum faster and more robustly.

3: In step3, the two coefficients $c_1^{wght/mol}$ and $c_2^{wght/mol}$, and the baseline offset, *b*, are varied to minimize the difference in Equation 5, resulting in the MAR fit (rows 4 and 5). The residual of the fit is the difference between the mixture spectrum and the linear combination spectrum, shown enlarged on the right in row 5. For a perfect fit, the residual is expected to represent the corresponding NMR noise. In fact, deviations of MAR residuals from NMR noise can be used as a diagnostic tool to evaluate the validity and quality of the analysis.

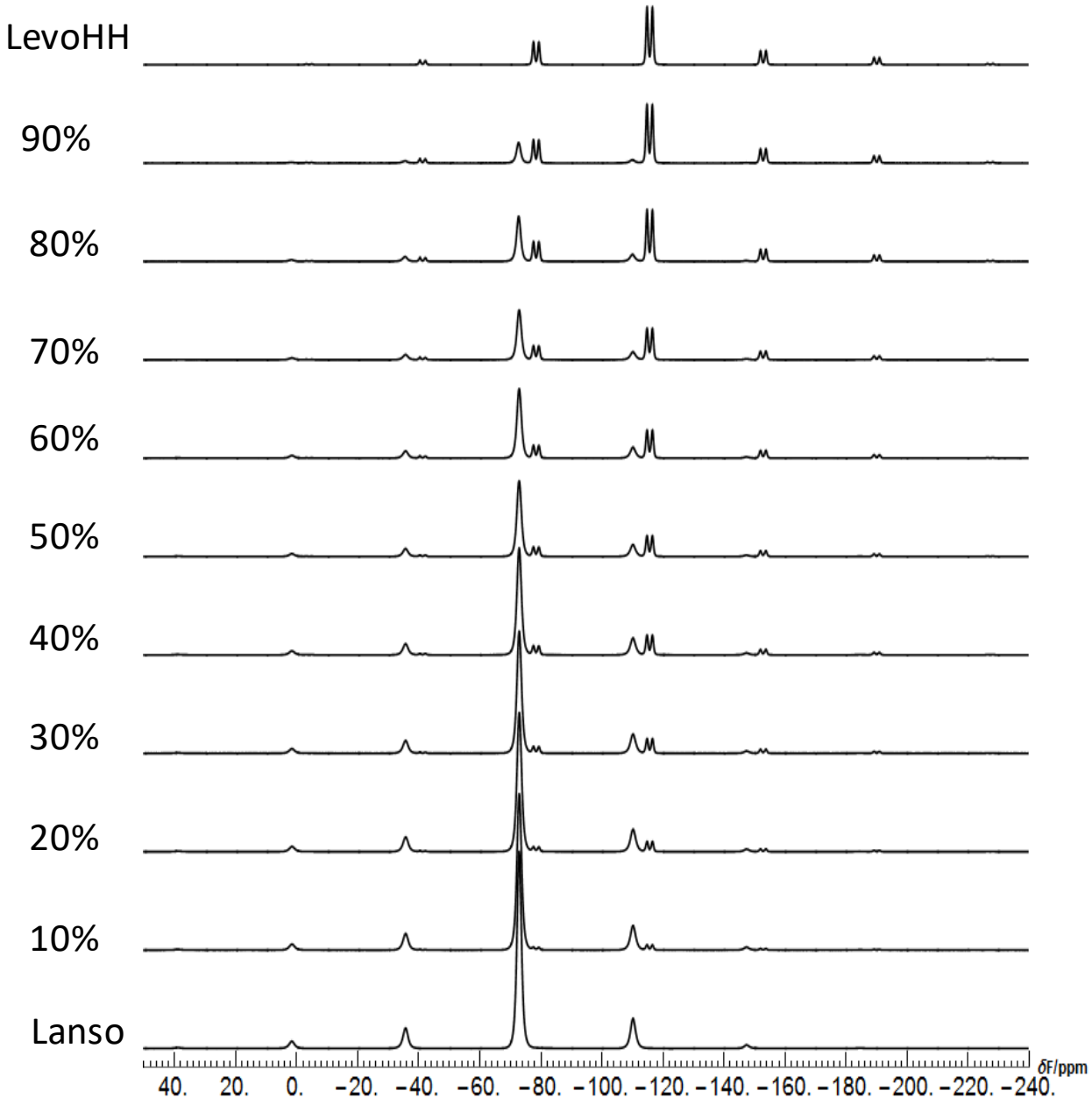


Figure S2. ^{19}F CPMAS pure component spectra for the model compounds, LevoHH (top) and Lanso (bottom), and of the corresponding nine binary blends with decreasing LevoHH content of about 90 mol% to about 10 mol% (exact compositions are given in Table 1).

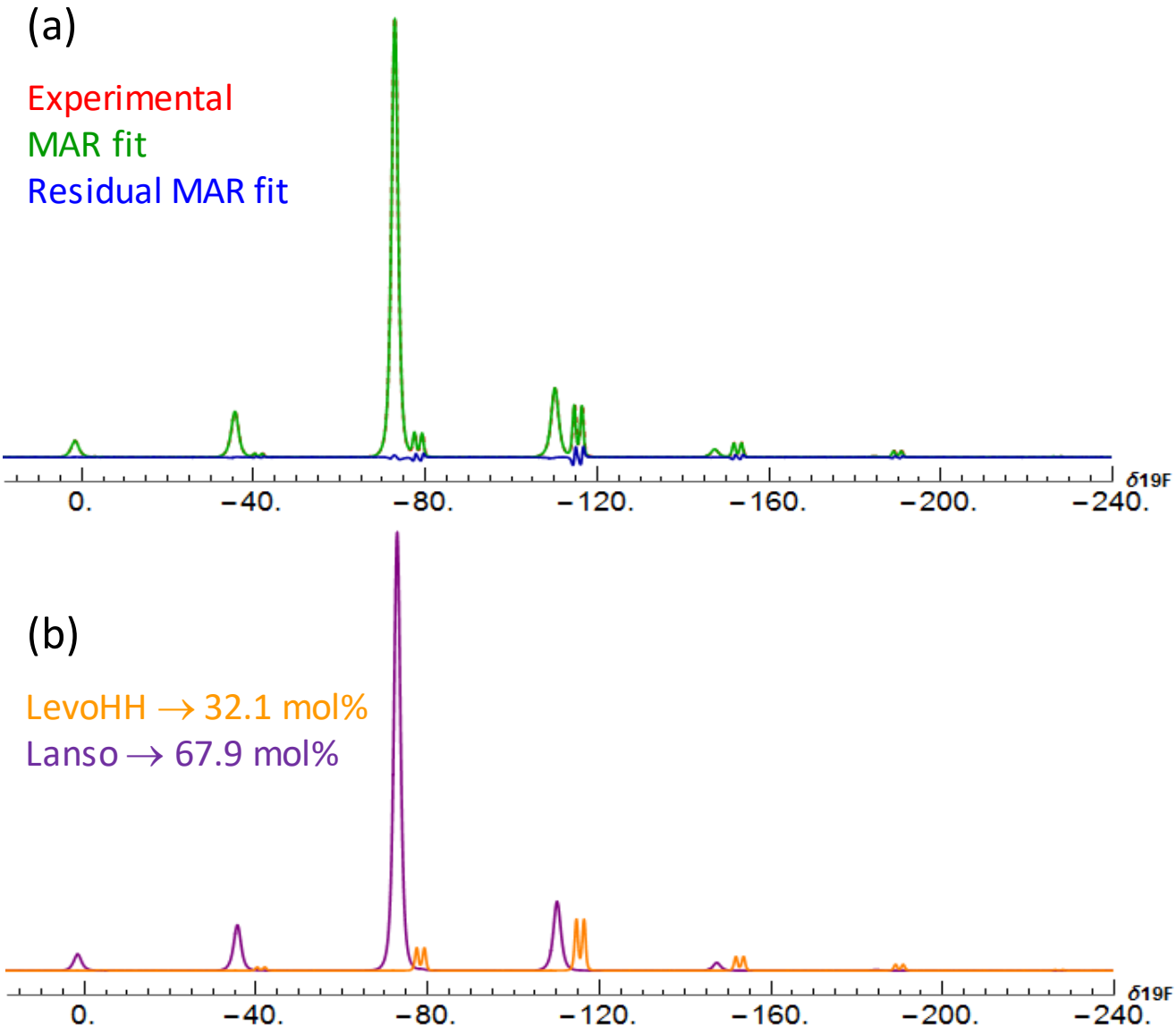


Figure S3. ^{19}F relMAR fit for the 31.9 mol%-LevoHH blend (Table 1). Panels (a) and (b) display the overall, and the LevoHH and Lanso contributions to the relMAR fit. The corresponding predicted composition is given in the inset in panels (b).

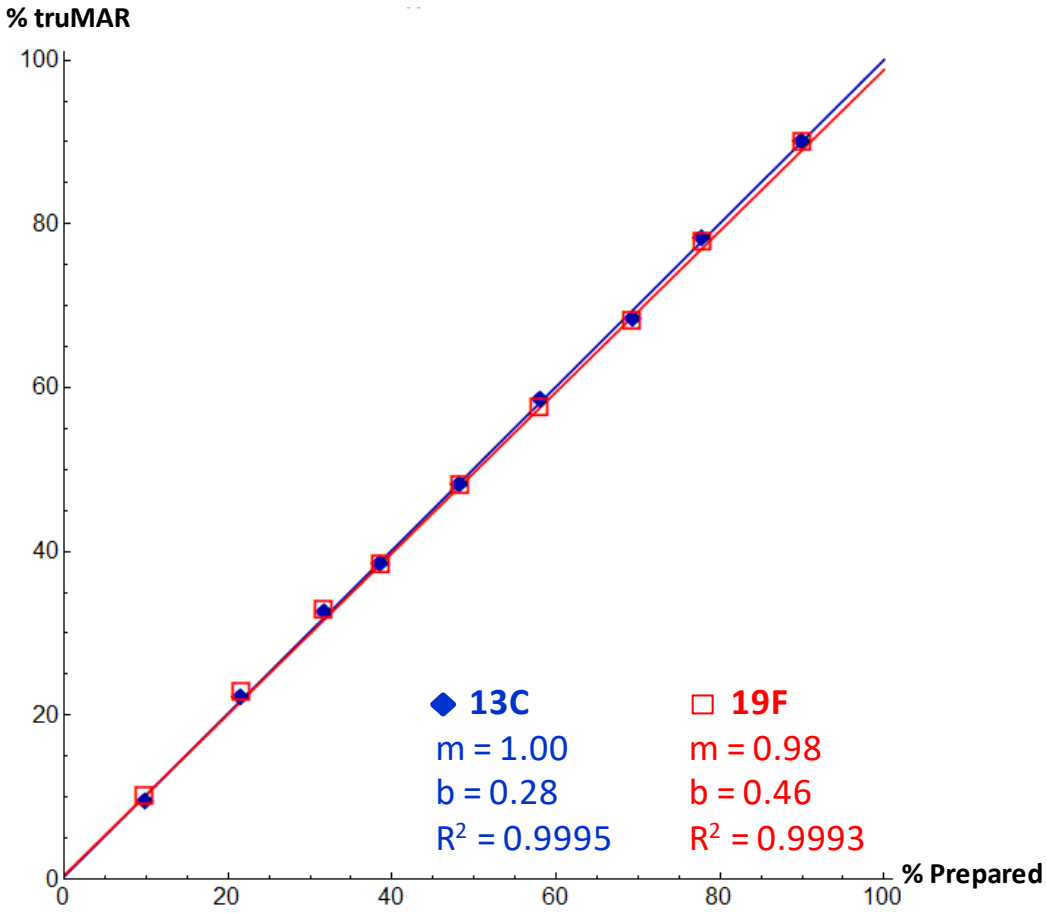


Figure S4. Results of the truMAR analysis of ^{13}C (blue, solid diamonds) and ^{19}F (red, open squares) data for all nine Lanso/LevoHH blends from Table 1 as correlation plots. The truMAR-predicted LevoHH mol% values are plotted versus the corresponding prepared composition values. The slopes (m), intercepts (b), and correlation coefficients (R^2) from linear regressions are presented.

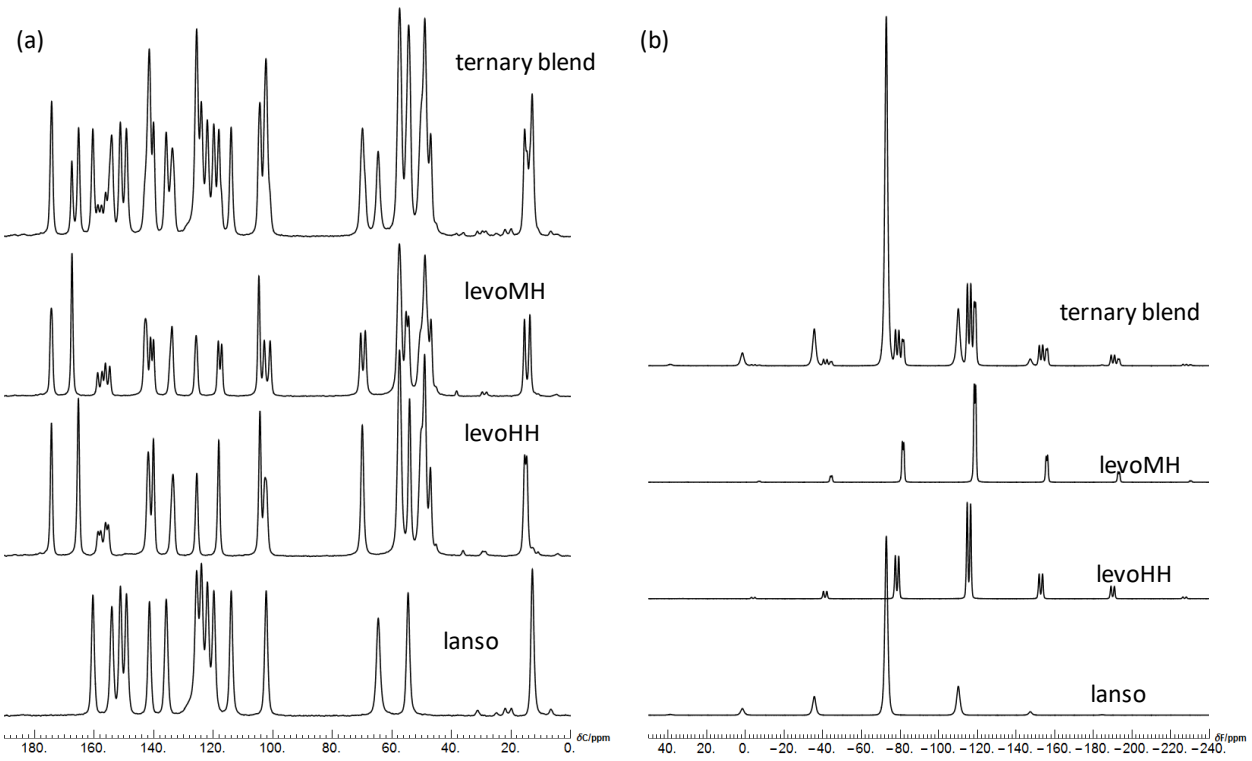


Figure S5. ^{13}C (a) and ^{19}F (b) CPMAS pure component spectra of the Lanso, LevoHH, and LevoMH, model compounds and of the corresponding ternary blend. The blend was prepared with a composition of 52.94 mol% Lanso, 29.13 mol% LevoHH, and 17.93 mol% LevoMH.

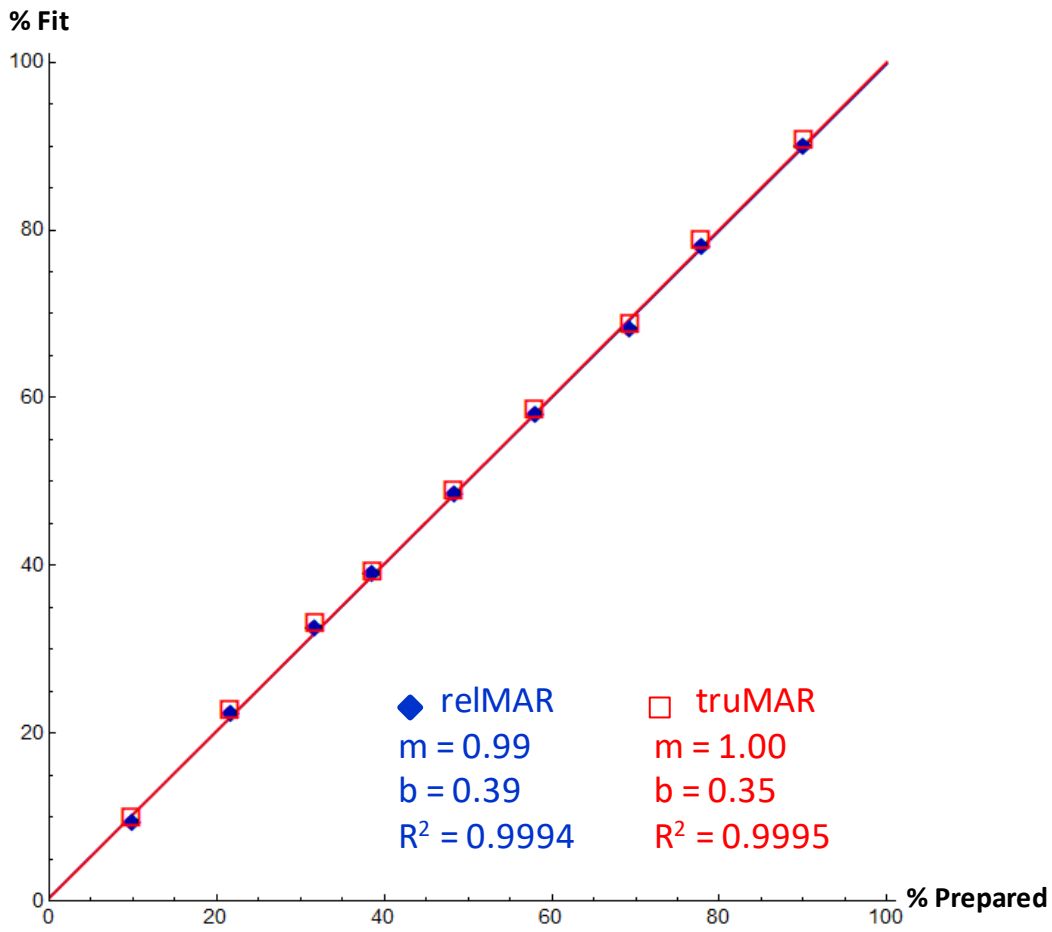


Figure S6. The relMAR (blue, solid diamonds) and truMAR (red, open squares) results of the ^{13}C spectral subsection (91 ppm to 145 ppm) MAR analysis for all nine binary Lanso/LevoHH blends from Table 1 as correlation plots. The MAR-predicted LevoHH mol-percentages are plotted versus the corresponding prepared values. The slopes (m), intercepts (b), and correlation coefficients (R^2) from linear regressions are displayed.

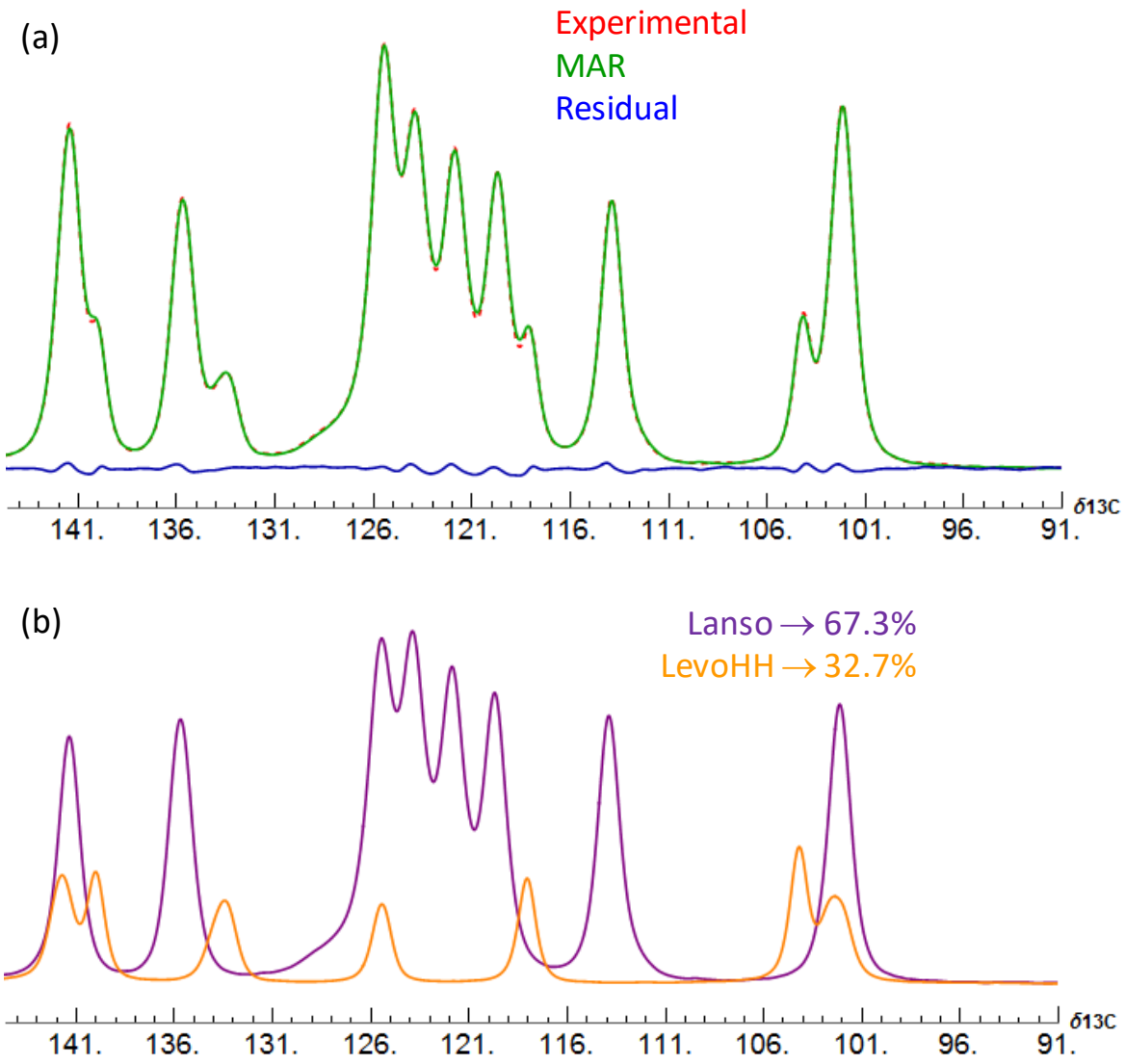


Figure S7. ^{13}C spectral subsection relMAR fit of the 39.1 mol% LevoHH blend. The overall fit is shown in (a), whereas panel (b) shows the contributions from each component to the fit. In addition, panel (b) displays the relMAR-predicted composition in the inset.

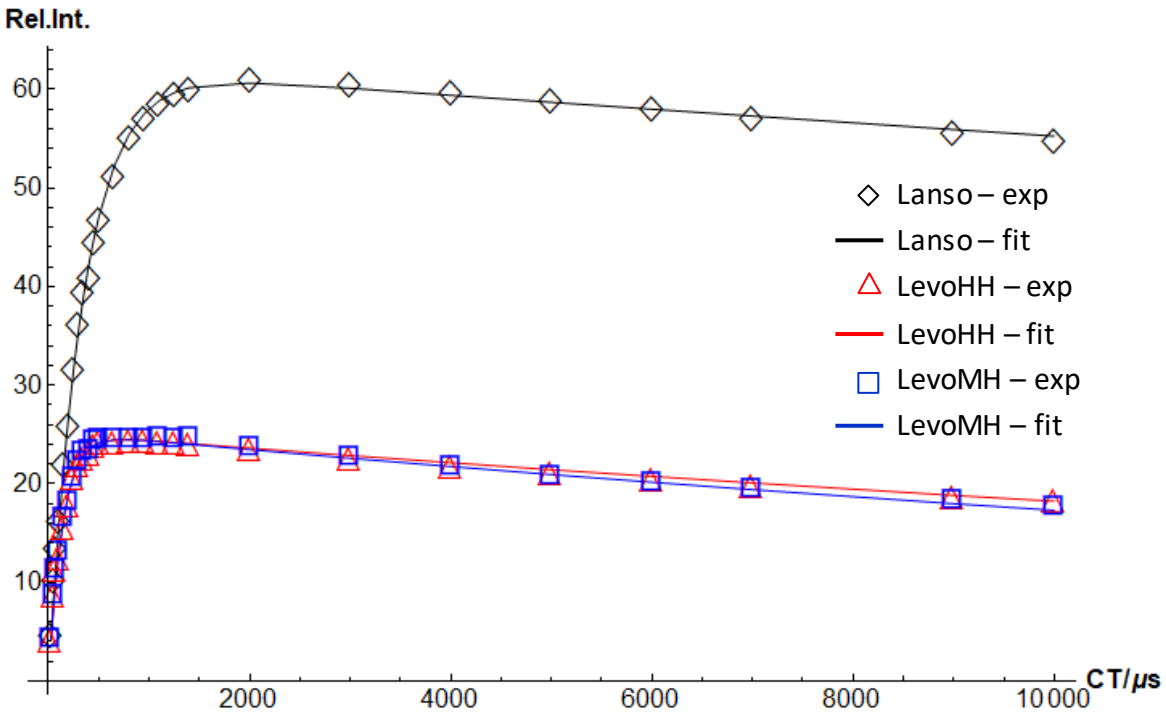


Figure S8. Application of the CP-dynamic corrected quantification approach (corrCP) discribed by Gao et al.[5] to our {^{1}H-^{19}F} CPMAS data for all binary Lanso/LevoHH blends in Table 1 and the ternary Lanso/LevoHH/LevoMH blend in Table 3. The figure shows the CP-curves, ie, signal area versus contact time, for Lanso, LevoHH, and LevoMH. Each of the three curves was built from 23 ^{19}F CPMAS spectra with contact times ranging from 25 to 10000 µs. The open symbols show the experimental total peak areas and the solid lines the corresponding fits to the theoretical model outlined by Gao et al.:

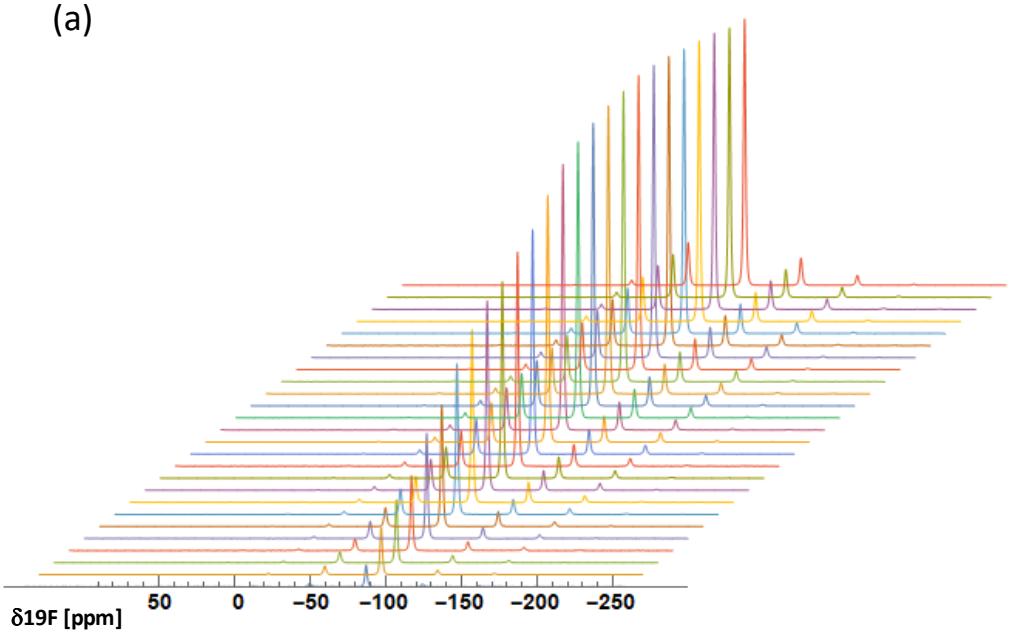
$$I(\tau) = \frac{M^0\left(\frac{\gamma_H}{\gamma_F}\right)\left[Exp\left(-\frac{\tau}{T_{1\rho H}}\right)-Exp\left(-\frac{\tau}{T_{FH}}\right)\right]}{\left(1-\frac{T_{FH}}{T_{1\rho H}}\right)} \qquad \text{(S6)}$$

In Equation S6, M^0, T_{FH}, $T_{1\rho H}$, γ_H/γ_F, are the X-nucleus equilibrium net magnetization value, the CP transfer rate, the proton spin-relaxation time in the rotating frame, and the relevant gyromagnetic ratios, respectively. Fits of the curves according to Equation 12 yield M^0, T_{FH}, and $T_{1\rho H}$ for the three model compounds as listed in Table S1. The experimental cp- curve data and the corresponding extracted peak areas are provided in Figures S9 to S11 for Lanso, LevoHH, and LevoMH, respectively. Once the CP-dynamic parameters were established, the relative peak areas for the components as measured from the spectra can be adjusted pair-wise utilizing a correction factor, $F_{A/B}$, that accounts for differences in CP dynamics according to equation S7:[5]

$$\frac{M_A^0}{M_B^0} = \frac{\left(1-\frac{T_{FH}^A}{T_{1\rho H}^A}\right)\left[Exp\left(-\frac{\tau}{T_{1\rho H}^B}\right)-Exp\left(-\frac{\tau}{T_{FH}^B}\right)\right]}{\left(1-\frac{T_{FH}^B}{T_{1\rho H}^B}\right)\left[Exp\left(-\frac{\tau}{T_{1\rho H}^A}\right)-Exp\left(-\frac{\tau}{T_{FH}^A}\right)\right]}\frac{I_{A(\tau)}}{I_{B(\tau)}} = F_{A/B}\left(\frac{I_A(\tau)}{I_B(\tau)}\right) \qquad \text{(S7)}$$

The two pair-wise correction factors necessary for our three-component system were determined as $F_{\text{Lanso}/\text{LevoHH}} = 1.04$ and $F_{\text{Lanso}/\text{LevoMH}} = 1.03$.

(a)

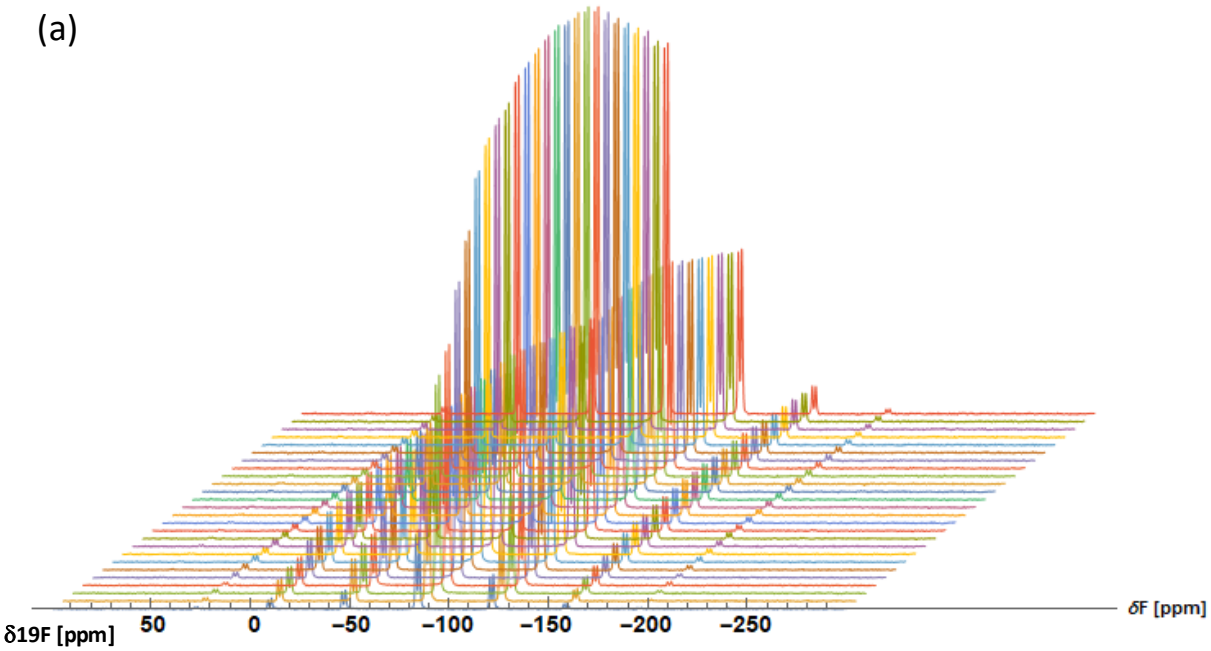


(b)

Contact time [μs]	Relative Intensity
25	4.83
50	10.32
75	13.68
100	16.37
150	22.17
200	25.96
250	31.69
300	36.30
350	39.63
400	41.08
450	44.72
500	46.97
650	51.37
800	55.21
950	57.28
1100	58.64
1250	59.71
1400	60.22
2000	61.13
3000	60.69
4000	59.88
5000	59.09
6000	58.13
7000	57.22
9000	55.72
10000	55.01

Figure S9: (a) 23 Lanso ^{19}F CPMAS spectra as a function of increasing contact time (front to back). The table in (b) displays the actual contact times as well as the corresponding extracted total peak areas.

(a)

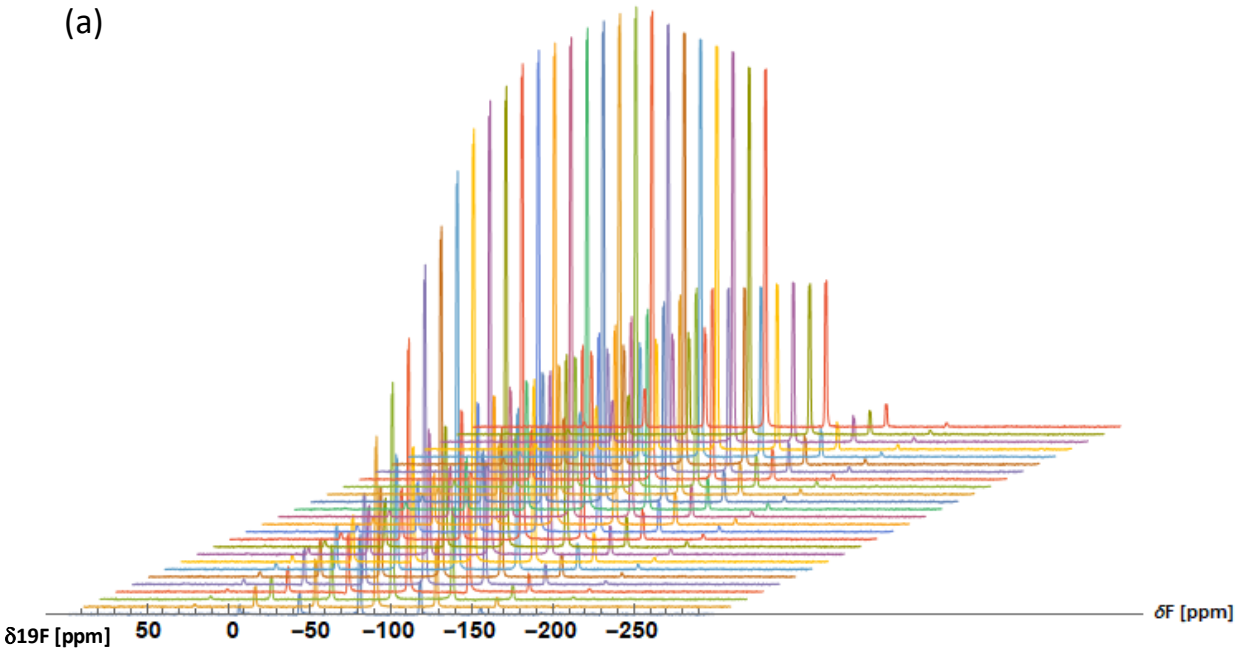


(b)

Contact time [µs]	Relative Intensity
25	4.30
50	8.86
75	11.44
100	12.67
150	15.72
200	17.96
250	20.79
300	22.13
350	22.75
400	23.27
450	24.03
500	24.33
650	24.40
800	24.56
950	24.55
1100	24.44
1250	24.34
1400	24.20
2000	23.67
3000	22.80
4000	21.95
5000	21.28
6000	20.58
7000	19.99
9000	18.88
10000	18.45

Figure S10: (a) 23 LevoHH ^{19}F CPMAS spectra as a function of increasing contact time (front to back). The table in (b) displays the actual contact times as well as the corresponding extracted total peak areas.

(a)



(b)

Contact time [µs]	Relative Intensity
25	4.22
50	8.62
75	11.18
100	13.04
150	16.32
200	18.01
250	20.49
300	22.18
350	23.06
400	23.31
450	24.24
500	24.41
650	24.46
800	24.43
950	24.38
1100	24.55
1250	24.45
1400	24.49
2000	23.50
3000	22.61
4000	21.62
5000	20.69
6000	19.96
7000	19.38
9000	18.16
10000	17.57

Figure S11: (a) 23 LevoMH ^{19}F CPMAS spectra as a function of increasing contact time (front to back). The table in (b) displays the actual contact times as well as the corresponding extracted total peak areas.

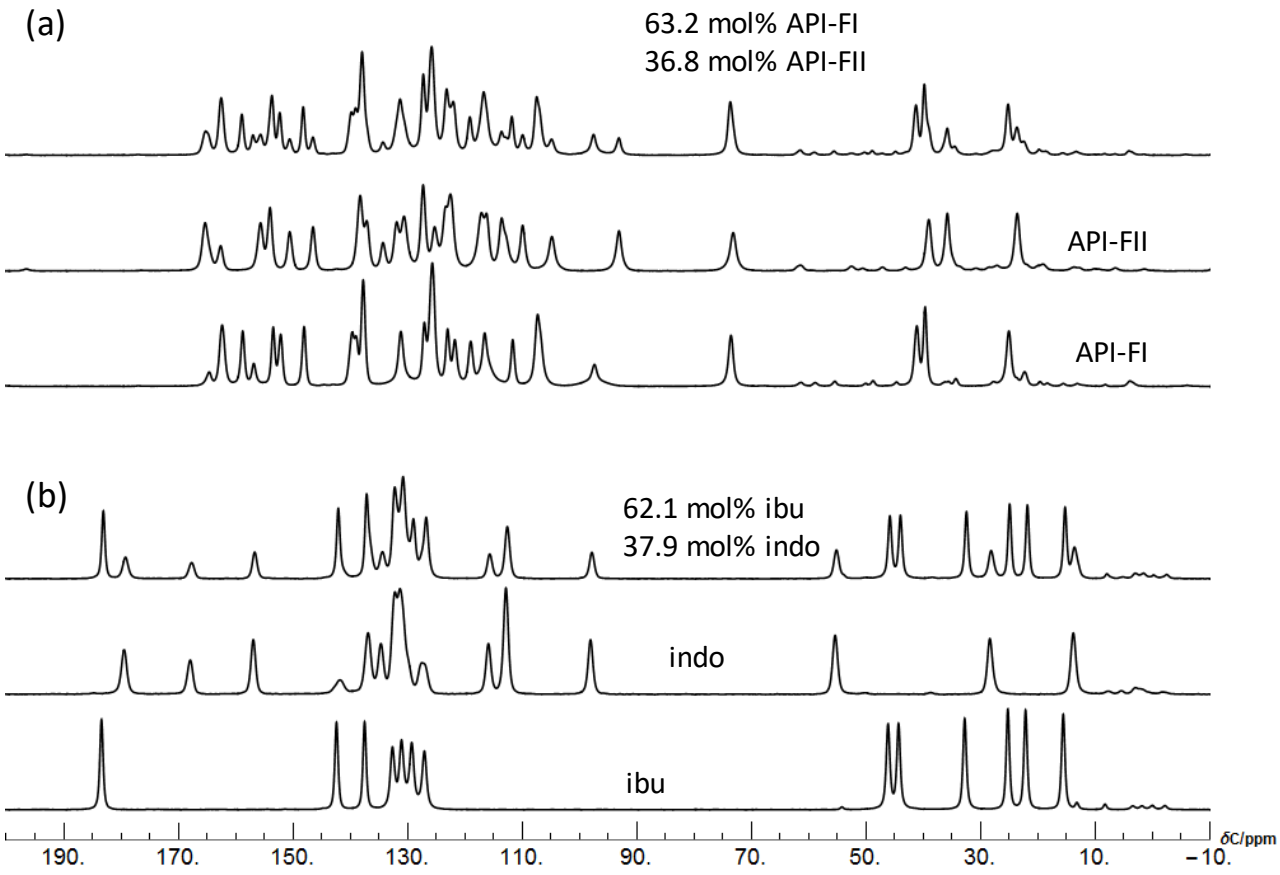


Figure S12. ^{13}C CPMAS spectra of the pure reference compounds and the corresponding binary blends for the API-FI/API-FII (a) and Ibu/Indo (b) model systems. The prepared compositions of each binary blend are given in the respective insets.

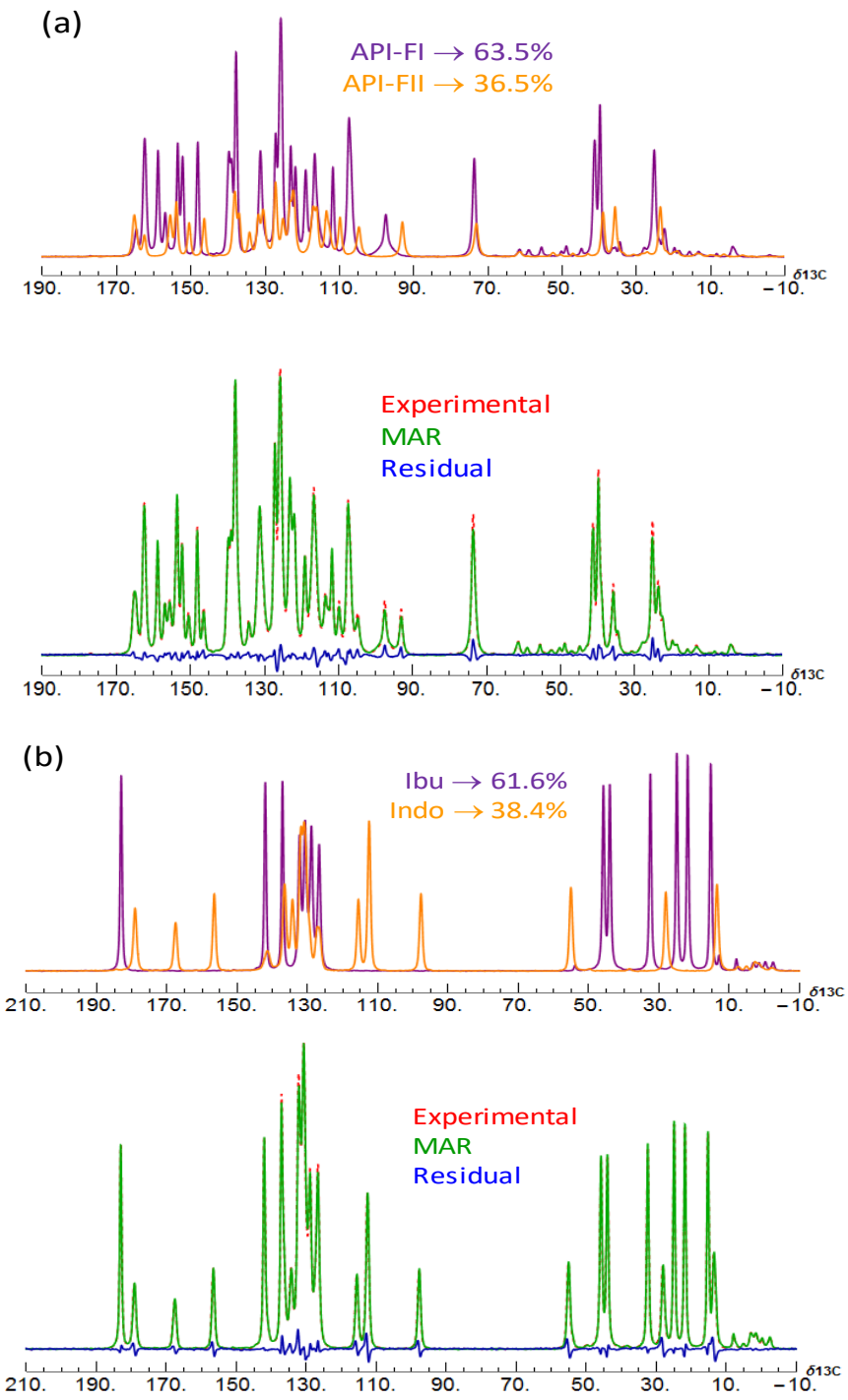


Figure S13. truMAR (a) and relMAR (b) fits of the binary API-FI/API-FII and Ibu/Indo blends, prepared as 63.2 mol% API-FI/36.8 mol% API-FII and 62.1 mol% Ibu/37.9 mol% Indo, respectively. The bottom panels show the corresponding overall fits, the top panels show the contributions from each component.

Table S1. Results of the CP-curve analyses described in Figure S8

Component	T_{FH} [ms]	$T_{1\rho H}$ [ms]	M^0	rms	$F_{lanso/xx}$
Lanso	349	83100	58.4	0.71	---
LevoHH	139	30900	23.6	0.44	1.04
LevoMH	140	26300	23.7	0.39	1.03

Table S2. Relevant experimental ssNMR parameters used for the Ibu/Indo and API-FI/API-FII model systems

Sample	d_{pulse} [s]	N_{scans}
Ibu	6.2 (5 T_1 Ibu)	1000
Indo	40.0 (5 T_1 long)	1000
Ibu/Indo blend	40.0 (5 T_1 long)	1600
API-FI	5.52 (1.2 T_1 long)	9000
API-FII	5.52 (1.2 T_1 long)	9000
API-FI/API-FII blend	5.52 (1.2 T_1 long)	18000

References

1. Little, T. A., Method Validation Essentials, Limit of Blank, Limit of Detection, and Limit of Quantitation. *BioPharm International* **2015,** *28* (4), 48-51.
2. Shrivastava, A.; Gupta, V. B., Methods for the determination of limit of detection and limit of quantitation of the analytical methods. *Chron Young Sci* **2011,** *2* (1), 21-25.
3. Zumbulyadis, N.; Antalek, B.; Windig, W.; Scaringe, R. P.; Lanzafame, A. M.; Blanton, T.; Helber, M., Elucidation of polymorph mixtures using solid-state C-13 CP/MAS NMR spectroscopy and direct exponential curve resolution algorithm. *Journal of the American Chemical Society* **1999,** *121* (49), 11554-11557.
4. Apperley, D. C.; Harris, R. K.; Hodgkinson, P., *Solid-state NMR : basic principles & practice*. Momentum Press: [New York, N.Y.] (222 East 46th Street, New York, NY 10017), 2012; p 36-38.
5. Gao, P., Determination of the Composition of Delavirdine Mesylate Polymorph and Pseudopolymorph Mixtures Using 13C CP/MAS NMR. *Pharmaceutical Research* **1996,** *13* (7), 1095-1104.