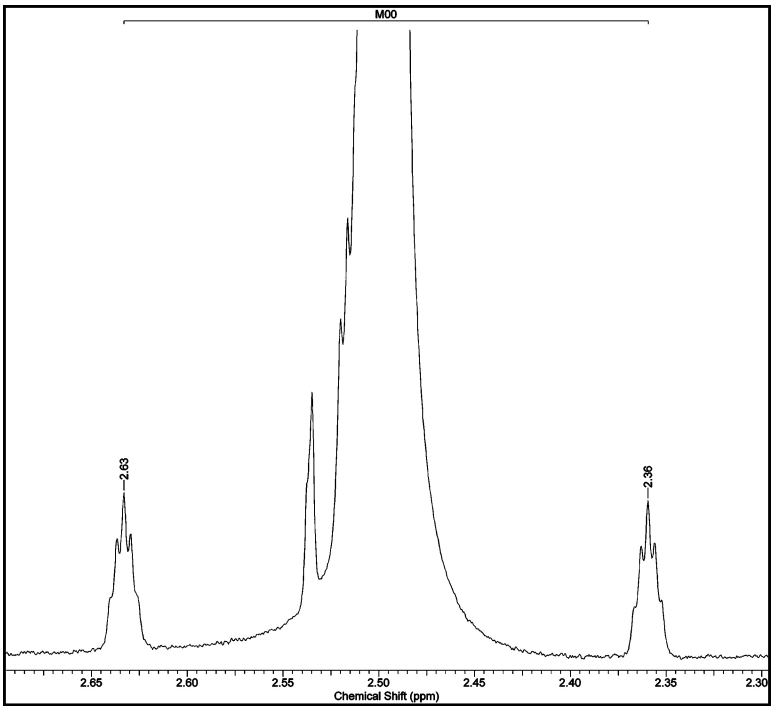
JOURNAL OF combinatorial CHEMISTRY

Article

Quality Control in Combinatorial Chemistry: Determinations of Amounts and Comparison of the "Purity" of LC–MS-Purified Samples by NMR, LC–UV and CLND

Elodie Letot, Guido Koch, Rocco Falchetto, Gnter Bovermann, Lukas Oberer, and Hans-Jrg Roth

J. Comb. Chem., **2005**, 7 (3), 364-371• DOI: 10.1021/cc049850r • Publication Date (Web): 16 March 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 7 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

Journal of Combinatorial Chemistry is published by the American Chemical Society.
1155 Sixteenth Street N.W., Washington, DC 20036

Articles

Quality Control in Combinatorial Chemistry: Determinations of Amounts and Comparison of the "Purity" of LC−MS-Purified Samples by NMR, LC−UV and CLND

Elodie Letot, Guido Koch, Rocco Falchetto, Günter Bovermann, Lukas Oberer, and Hans-Jörg Roth*

Novartis Institutes for BioMedical Research, Lead Synthesis & Chemogenetics and Analytical & Imaging Sciences, CH-4002 Basel, Switzerland

Received September 23, 2004

The absolute purities of 20 purified samples from a combinatorial library have been determined by a new method that uses the DMSO sidebands [$^1J\{^{13}C-^1H\}$] as an internal standard for quantification. The obtained absolute amounts are compared with the amounts of sample obtained by weighing, with the calculated weights obtained by chemiluminescent nitrogen detection (CLND) chromatography and with the relative purities obtained by LC−UV chromatography.

Introduction

In the last couple of years, most of the major pharmaceutical companies significantly invested in building up large compound archives in order to benefit from the increased screening capacities that have been developed during the past decade. Today, a typical highly automated compound archive consists of 0.5−2 million compounds. The time required to screen such a compound collection is usually less than 1 week. Larger numbers of new compounds which cannot be delivered by traditional synthesis methods are therefore required.

Combinatorial chemistry is becoming more and more a significant source of compounds for building up these large and diverse compound archives. It is understood that the value of the screening data is highly dependent on the quality of the compound archive. Therefore, quality assurance of combinatorial libraries became an important issue in the last years.[1,2] The discussion of the quality of compounds from combinatorial or high-throughput chemistry can be divided into two fields: (1) drug likeliness (physical properties as well as diversity) and (2) chemical quality (identity, purity, and quantity). Here, we would like to address the latter by reporting the results of a study on the identity, purity, and quantity of 20 randomly chosen compounds from a single library example.

In 1998, we occasionally observed a 10−20-fold weight reduction of crude combinatorial chemistry samples cleaved from solid phase after preparative HPLC purification. This observation was true for samples for which the $LC-UV_{214nm}$ purity before and after purification was determined to be >90%. Therefore, it became clear to us that crude samples from solid-phase synthesis in general contain large amounts of impurities not detected by LC−UV analysis.[1] In selected cases, if isolated pure samples were available, we confirmed this observation by UV calibration curves (obtained from LC−UV analysis of the crude samples), which allowed the determination of the absolute amount of the desired compound in the crude sample. On the basis of these experiences, we decided that all library compounds requested HPLC purification, regardless of the relative LC−UV purity of the crude samples. Our standard process, therefore, consists of an MS-triggered HPLC purification of all library compounds followed by an LC−UV quality analysis and weighing of the dried samples. On the basis of the relative LC−UV purity and the obtained weights, we performed a select-and-compress process. Samples with a relative LC−UV purity higher than 85% are submitted to the archive as DMSO solutions with a concentration of 10 mM. The majority (∼70%) of isolated amounts of compound of a typical library are in the range of 3−5 mg. This corresponds, depending on the molecular weight of the compounds, to 6−10 μmol. The scale for the last reaction step of a library production is typically 40 μmol.

It was the first and original aim of the study reported here to investigate how relevant the weighing of such samples is. How much of the measured weight belongs to the expected compound and how much belongs to solvent inclusions, water, and salts? Furthermore, we asked the question on the absolute purity of small library samples in

* To whom correspondence should be addressed. Phone: +41 61 324 30 96. Fax: +41 324 67 26. E-mail: hansjoerg.roth@pharma.novartis.com.

10.1021/cc049850r CCC: $30.25
Published on Web 03/16/2005

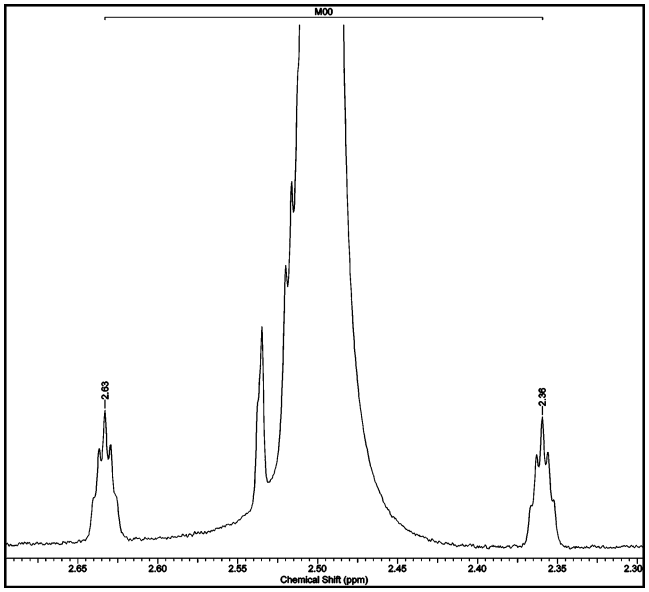


Figure 1. ^{1}H NMR of DMSO-d_6 (500 MHz,); $\delta = 2.5$ ppm.

Table 1. Calibration of the ^{13}C Side Bands with Acetanilide as an External Standard

concn acetanilide, mM	0.21	0.42	1.05	10.50
^{13}C satellite integrals[a]	0.269	0.136	0.0532	0.005 44
calcd ^{13}C satellite "concn"[b], mM	0.169	0.172	0.168	0.171

[a] The ^{13}C satellite integral values were obtained after calibrating the integral of the acetanilide methyl group signal as 1.000. [b] The concentration of the ^{13}C sidebands is calculated by $c = c_{\text{acetanilide}} \times I^{13}\text{C}/0.3333$

comparison with compounds prepared by traditional methods and in quantities around 100 mg. To answer these questions, we applied an in-house-developed method that allowed the NMR-based quantification of compounds in the absence of external standards. In addition to this absolute quantification, we performed a visual inspection of the NMR data. This gave us an impression of the comparability with typical pure medicinal chemistry samples. Finally, we completed the data set with the corresponding LC−MS, LC−UV, and LC−CLND (chemiluminescent nitrogen detection) analysis in order to validate our standard high-throughput quality control process.

Material and Methods

^{1}H NMR Quantification of Compound Samples by ^{13}C Satellites. Deuterated dimethyl sulfoxid (DMSO-d_6) is one of the most commonly used solvents in nuclear magnetic resonance (NMR) spectroscopy and is often regarded as the "universal solvent" because of its ability to dissolve most organic compounds. It stabilizes exchangeable protons of functional groups, such as OH and NH, and makes them easily detectable.

Due to 1.1% naturally abundant ^{13}C, the signal of DMSO-d_6 is accompanied by two ^{13}C satellites, each with an intensity of 0.55% of the parent signal. The distance between the two sidebands reflects the $^1J\{^{13}\text{C}-^1\text{H}\}$ coupling constant of 136.9 Hz. (Figure 1).

Since the integrals of the ^{13}C sidebands of DMSO-d_6 are constant within the same batch of DMSO-d_6, they can be used as an internal standard after their calibration. The absolute concentration of any sample can then be determined by the comparison of the integral of the sideband with that of a compound signal, as long as one of the sidebands is visible in the spectrum and its intensity is sufficient for a reliable integral.

Calibration of the ^{13}C Sidebands with an Internal Standard. Acetanilide was used as an internal standard for the calibration of the ^{13}C satellites. The observed linear correlation between the concentration of acetanilide and the integral of the ^{13}C satellite signal allows the determination of a sample concentration that corresponds to the ^{13}C satellite integral of a given DMSO-d_6 batch[3]. In the reported sample,

Table 2. Amounts Obtained by Weighing and by NMR and CLND Quantification

entry	sample	amt by weighing, mg[a]	amt by NMR quantification, mg[b]	ratio TFA/cmpd	amt incl. TFA by NMR quantification, mg	purity by NMR, %[c]	purity by CLND, %
1[d]	**621-083**	**7.9**	**1.5**	**0.6**	**1.7**	**22**	**26**
2	621-084	10.5	4.2	0.7	5.2	50	32
3	621-086	5.3	3.0	0.5	3.5	66	43
4	**621-092**	**11.1**	**6.5**	**1.3**	**8.9**	**80**	**72**
5	621-095	11.2	5.6	1.6	8.3	74	67
6	621-104	4.7	1.7	1.1	2.3	48	44
7	621-106	8.7	3.6	1.2	5.0	58	67
8	621-109	9.7	4.3	0.8	5.3	55	59
9	**621-124**	**7.0**	**5.1**	**1.0**	**6.1**	**87**	**79**
10	621-129	7.0	4.1	0.7	4.9	71	71
11	621-131	10.3	4.2	2.2	7.2	70	68
12	621-145	5.8	2.9	0.9	3.5	60	36
13	621-151	9.8	7.3	0.9	8.5	87	100
14	621-154	10.1	4.3	0.7	4.8	48	
15	621-173	3.3	0.7	1.4	0.9	28	18
16	621-196	9.4	1.7	0.9	2.1	22	19
17	**621-213**	**6.5**	**3.0**	**0.6**	**3.4**	**51**	**41**
18	621-236	8.0	6.9	0.6	7.8	86	65
19	621-260	8.6	1.9	1.0	2.5	30	33
20	621-281	4.1	1.6	1.5	2.1	50	52

[a] After purification. [b] Calculated (based on the NMR results) amount of the sample that belongs to the expected compound. [c] Percent of expected compound obtained by the NMR measurement of the obtained weights after purification. [d] Rows in bold are discussed in detail in the text.

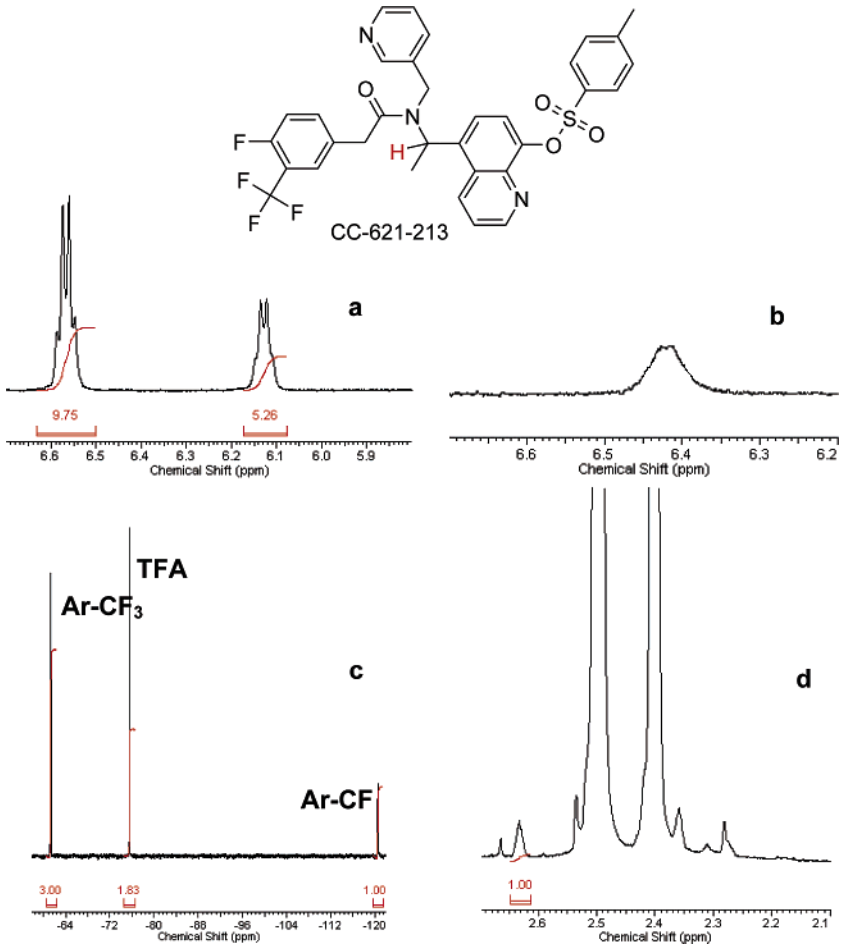


Figure 2. (a) ^{1}H NMR region of the α-amido proton (red) at room temperature, showing the signals of the two rotamers; (b) ^{1}H NMR region of the same protons at 393 K; (c) ^{19}F NMR spectrum at room temperature; and (d) integration of the ^{13}C sidebands.

the ^{13}C satellite integral of the DMSO-d_6 batch corresponded to a "signal concentration" of 0.17 mM for one proton (Table 1).

From this result, the absolute amount of compound in a sample or the concentration of the sample is determined on the basis of the following formula:

$$m = (C_{\text{side band}} IVDM)/N \quad (1)$$

$$C = (C_{\text{side band}} I)/N \quad (2)$$

where $C_{\text{sideband}} = 0.170$ mM, I is the integral of the peak when one satellite is integrated as 1, V is the volume of the solution in the NMR tube (mL); D is a dilution factor, M is the molecular weight of the compound, and N is the number of responsible protons for a given peak.

^{19}F NMR Quantification of TFA Amounts. Since all selected compounds of interest contain one or more fluorine atoms, a semiquantitative measurement of the amount of TFA was possible by simply comparing the ^{19}F integral of TFA with that of the ^{19}F signal(s) of the compound.

Sample Preparation. Stock solutions from the purified samples were prepared, using DMSO-d_6 of an identical batch, for which the calibration with actetanilide was done previously. To obtain a concentration in a range of 10−30 mM for the stock solutions, either 0.5 or 1 mL DMSO-d_6 was taken. For the final NMR solutions, 150 μL of each of these stock solutions was pipetted into the NMR tubes, and 400 μL of the same DMSO-d_6 batch was finally added to each, resulting in a total volume (V) (eq 1 and 2) of 550 μL and a dilution factor (D) of 3.33 or 6.66, respectively.

NMR Equipment and Parameters

nucleus	^{1}H	^{1}H	^{19}F
temperature, K	300	393	300
spectral width, kHz	10	8.278	41.408
acquisition time, s	3.5	2.6	0.4
relaxation time, s	1.5	1	1
no. of scans	64	64	32
RF pulse, °	30	30	30
frequency, MHz	500.13	400.13	376.59
spectrometer	BRUKER DRX500	BRUKER DPX400	BRUKER DPX400
probe	TXI-CryoProbe	TXI	QNP

In all experiments, the receiver gain was not automatically kept constant from one measurement to the other, since a high precision in the results was not the goal of the present study.

RP-HPLC/UV/CLND/MS Analysis. CLND-based quantification and UV/MS identification were carried out according to Taylor et al. with minor adaptations.[4] RP-HPLC was performed using a ReproSil-pur C18 ODS-3 column, 3-μm particle size, 120-Å pore size, 1.0 × 60 mm (Dr. Maisch HPLC-Gmbh, Ammerbuch, Germany) on a Shi-

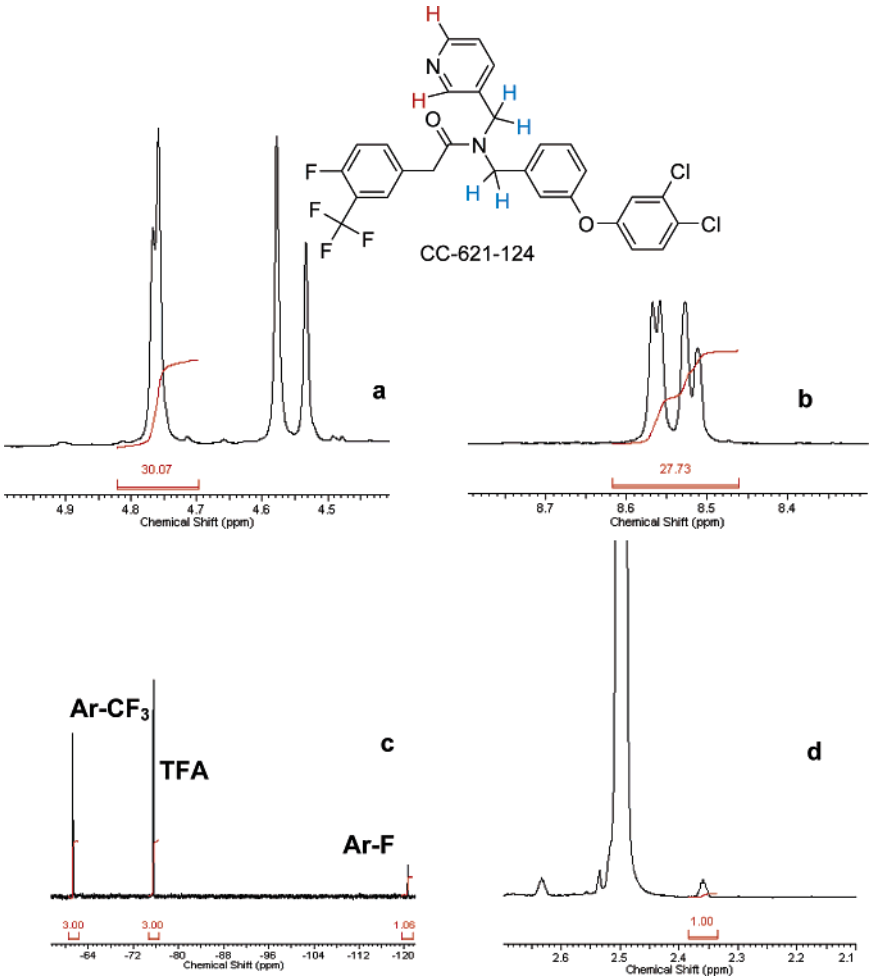


Figure 3. (a) ^{1}H NMR signal region of the rotamers of two benzylic groups (blue); (b) ^{1}H signal region of two pyrridyl protons (red); (c) ^{19}F spectrum, showing the 1:1 ratio between expected compound and TFA; and (d) ^{1}H NMR region of the ^{13}C sideband, integral set to 1 as a reference.

madzu LC-10AD VP pump equipped with an SCL-10A VP controller and a SIL-10AD VP auto injector. The column was heated at 45 °C in a Cera 150 column heater (Cera Inc., Baldwin Park, CA) and eluted at 150 μL/min with a linear gradient from 20 to 95% buffer B in 10 min. Buffer A was 0.1% (v/v) TFA in water, and buffer B was 0.08% (v/v) TFA in 75% (v/v) MeOH, 25% (v/v) IPA (isopropylic alcohol). The HPLC flow was directed into a Linear UVIS 204 monitor (Linear Instruments Corp., Reno, NV) set at 214 nm. Following UV detection, the eluate was divided in two through a splitting tee. About 100 μL/min was directed into an Antek 8060 CLND nitrogen detector (Antek Instruments, Inc., Houston, TX) equipped with a Meinhard TL-HEN-1220-AA glass nebulizer (Meinhard Glass Products, Golden, CO); 50 μL/min was directed into a Finnigan LCQ ion trap mass spectrometer equipped with an electrospray ion source. The CLND was calibrated using diphenhydramine (Sigma, St. Louis, MO) as the standard nitrogen-containing compound.

Purification of Library Compounds. Separations were carried out by linear gradient elution of 10 min from 5 to 95% buffer B. Buffer A contained 0.1% TFA in water, and buffer B contained 0.1% TFA in ACN. Samples were eluted on a 19 × 100 mm Waters Xterra 5-μm column using a flow rate of 20 mL/min. The target compounds were identified by electrospray ionization and collected by mass-directed triggering of the fraction collector. Fractions were collected using a Gilson 215 fraction collector. The expected product from each sample present in the input rack was collected in one fraction (max. 8 mL), on the basis of mass detection, and placed at the same position in the output rack (1:1 mapping).

Results

Quantification. To quantify the absolute amount of an expected compound in the purified samples, each spectrum needs at least one isolated and assignable NMR signal that belongs clearly to the compound of interest, and at least one of the ^{13}C satellites of DMSO-d_6 also has to be isolated from other signals. The integral(s) of this signal(s) is/are then compared with that of one of the ^{13}C satellites of DMSO-d_6. From the integration ratio of the ^{13}C satellite versus compound signals, the absolute concentration in the NMR sample can be determined. This value allows a back calculation of the concentration of the stock solution, from which the absolute amount of compound can be determined. In the case of compound CC-621-213 (Table 2, entry 17), the ^{1}H signals of the α-amido proton are well-isolated and observable at room temperature as signals of two rotamers (Figure 2a). At 393 K, the two signals become one broad

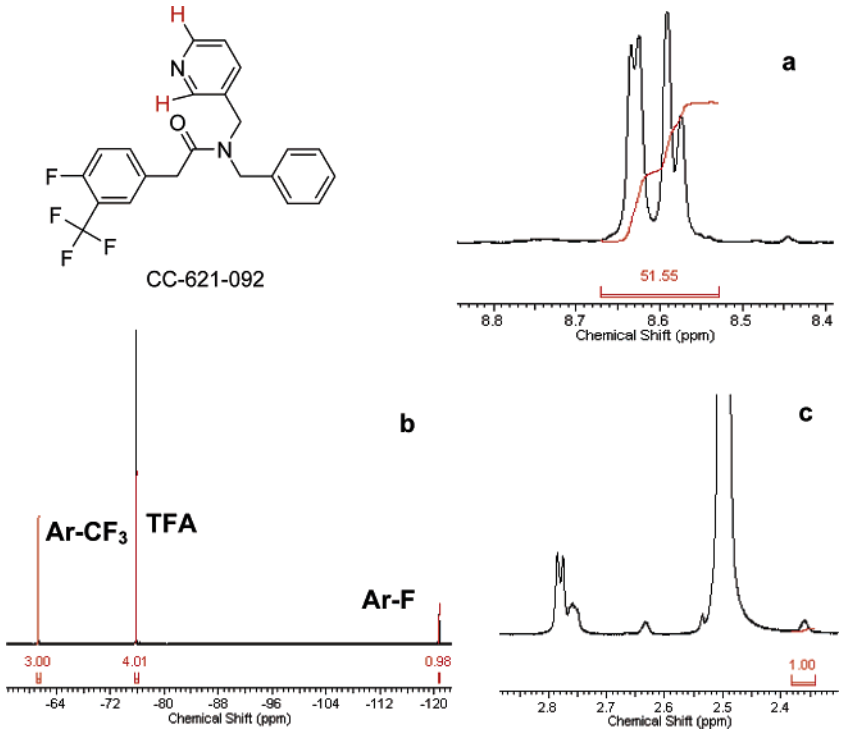


Figure 4. (a) ^{1}H NMR signal of the two pyrridyl protons (red) next to the ring nitrogen; (b) ^{19}F spectrum, showing the 1:1.3 ratio between the expected compound and TFA; and (c) ^{1}H NMR region of the ^{13}C sideband, integral set to 1 as a reference.

signal (Figure 2b). One of the ^{13}C satellites is nicely isolated from the other signals. Its integral is set to 1 as a reference (Figure 2d). The calculation of the absolute amount of expected compound was done according to eq 1 on the basis of the integrals of the spectrum at room temperature: $C_{\text{satellite}}$ = 0.17 mM; sample volume (V), 0.55 mL; integral (I) of assigned proton, 9.75 + 5.26 = 15.01; dilution factor (D), 3.33; no. of integrated protons (N) = 1; molecular weight (M), 637.66. Therefore, $m = 0.00017 \times 15.01 \times 0.55 \times 3.33 \times 637.66/1 = 2.98$ mg of expected compound in sample 621−213.

The intregration ratio of the ^{19}F signals of the trifluoromethyl group of CC-621−213 at −62 ppm and the signal of TFA at −76 ppm (1.83:3 = 0.61) showed that sample 621-213 contains 0.6 equiv of TFA (Figure 2c). The observed amount of 3.0 mg of expected product without TFA corresponds to 4.7 μmol. The observed 0.6 equiv of TFA (2.8 μmol) corresponds to 0.3 mg. The total amount of expected product (neutral species or as TFA salt) is, therefore, 3.3 mg. The sample amount determined by the automated weighing was 6.5 mg (Table 2, entry 17). The isolated sample has, therefore, an absolute purity of 51% (3.3 of 6.5 mg). The quantification by CLND gave an absolute purity of 41%. Because the ^{1}H NMR spectrum (Figure 6c), the high-throughput LC−UV analysis (100% purity, not shown), and the analytical LC−MS (Figure 7c) did not show any significant organic impurities, it is presumed that the difference in the weight between the results of the weighing and the NMR determination is due to water inclusions, invisible inorganic salts, a weighing error, or a combination thereof.

The ^{1}H NMR spectrum of sample 621-124 (Table 2, entry 9) contains two regions of signals, allowing a proper integration necessary for the quantification. The region between 4.5 and 4.8 ppm shows the signals (of the rotamers) of two benzylic groups (Figure 3a); the region between 8.5 and 8.6 ppm, the signals of two pyrridylic protons (Figure 3b). The ^{13}C sideband signal was integrated and set to 1 as a reference (Figure 3d). The corresponding calculation (0.00017. 57.80. 0.55. 6.66. 563.83/4) gave 5.1 mg (9 μmol). The ^{19}F spectrum showed 1 equiv of TFA in the sample (1.0 mg). The total amount of expected product (neural species and TFA salt) is, therefore, 6.1 mg, which is 87% of the amount obtained by weighing. The amount obtained by CLND corresponds to an absolute purity of 79%. In good agreement with the good quality of the ^{1}H NMR spectrum (Figure 6b), the high-throughput (not shown) and the analytical LC−UV chromatogram (Figure 7b) do not show any significant organic impurities.

The quantification of sample CC-621-092 (Table 2, entry 4) is also based on the signals of the two ortho pyrridyl protons (Figure 4a). The corresponding calculation (0.00017 × 51.55 × 0.55 × 3.33 × 402.39/4) delivered an amount of 6.5 mg (16 μmol) of the expected compound. The ratio of the corresponding signals in the ^{19}F spectrum (Figure 4b) delivers a value of 1.3 equiv of TFA (2.4 mg). The total calculated amount (neutral species and TFA salt) of 8.9 mg corresponds to 80% of the amount obtained by weighing (11.1 mg). The amount obtained by CLND corresponds to an absolute purity of 72%. In good agreement with the good quality of the ^{1}H NMR spectrum (Figure 6a), the high-throughput (not shown) and the analytical LC−UV chromatogram (Figure 7a) do not show any significant organic impurities.

All three examples discussed above show in accordance with their ^{1}H NMR spectra and their HPLC chromatograms high-purity samples (Figure 6a−c, Figure 7a−c). The differences in the weights between the NMR or CLND

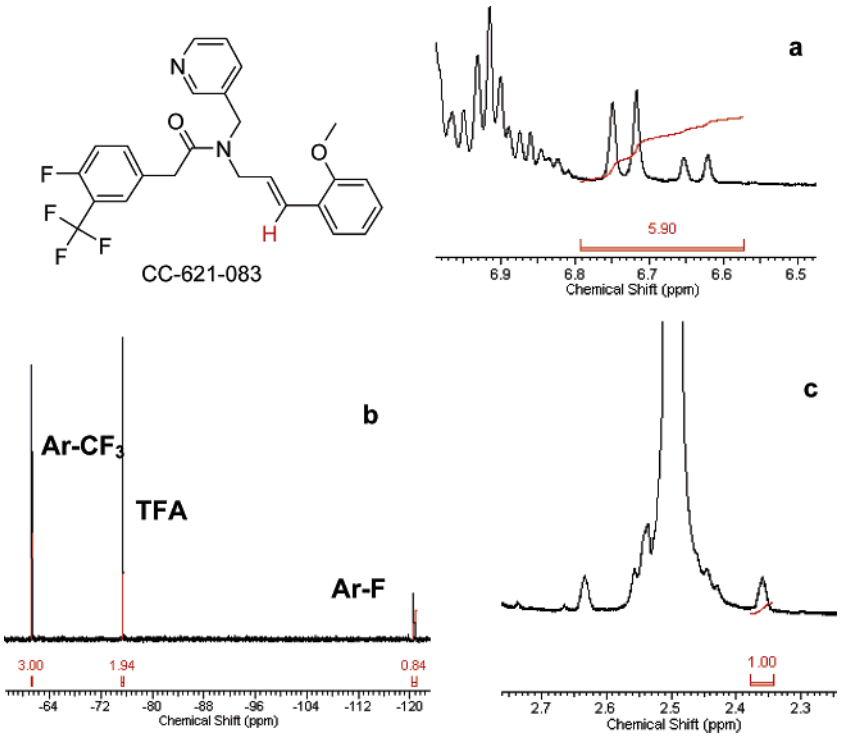


Figure 5. (a) Signals of the two rotamers of the vinylic protons (red); (b) ^{19}F spectrum; and (c) ^{13}C satellite signal, integral set to 1 as a reference.

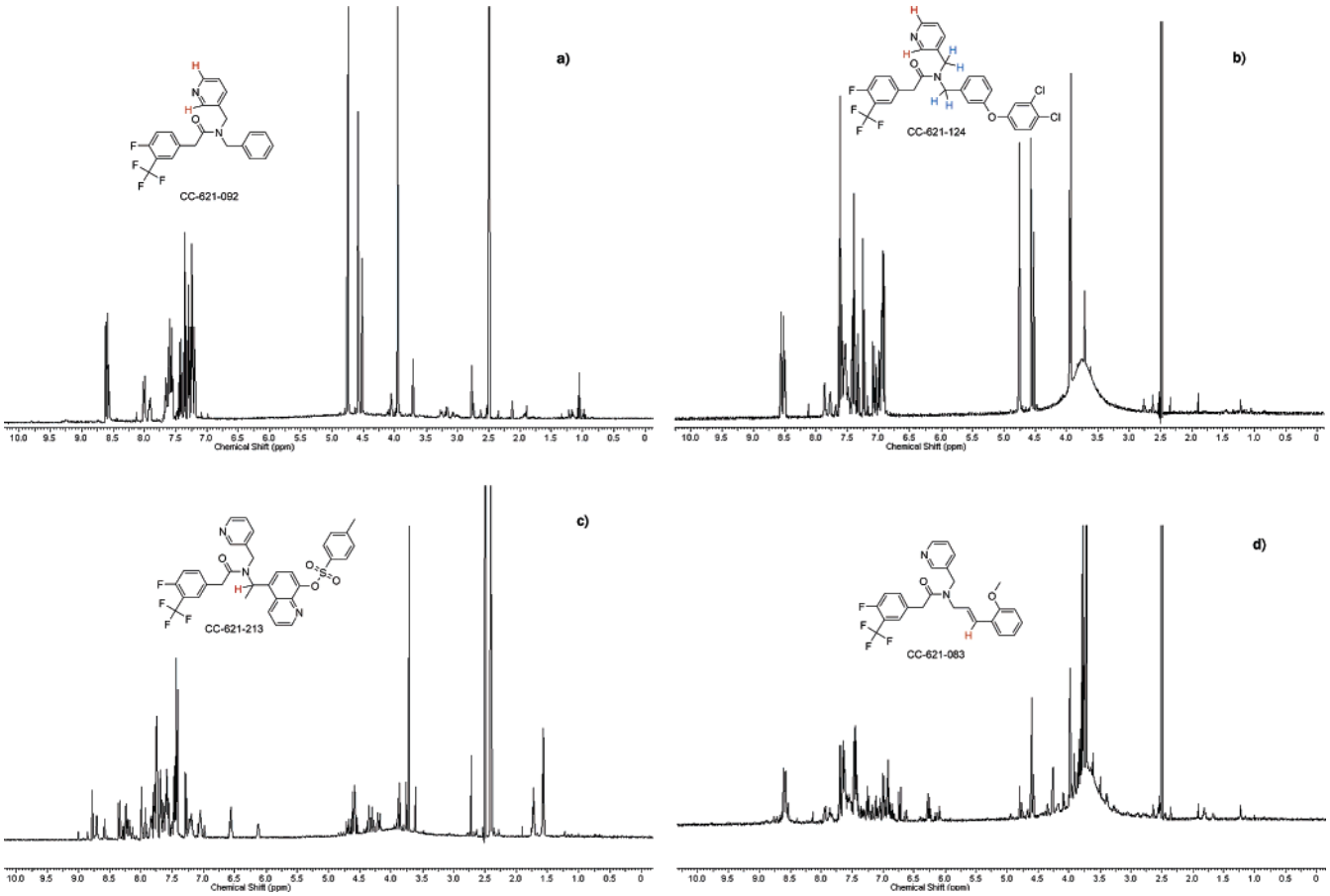


Figure 6. (a−c) ^1H spectra (0−10 ppm) of the three samples with absolute amounts ≥52%, which represents the average quality of the investigated batch. (d) ^1H spectrum (0−10 ppm) of the sample with an absolute amount of 22%, which is worse than the average.

quantification and the weighing are, therefore, explained by water inclusions, invisible salts, and weighing errors. Sample CC-621-083 (Table 2, entry 1) is significantly different. Already, the visual inspection of the ^1H NMR spectrum

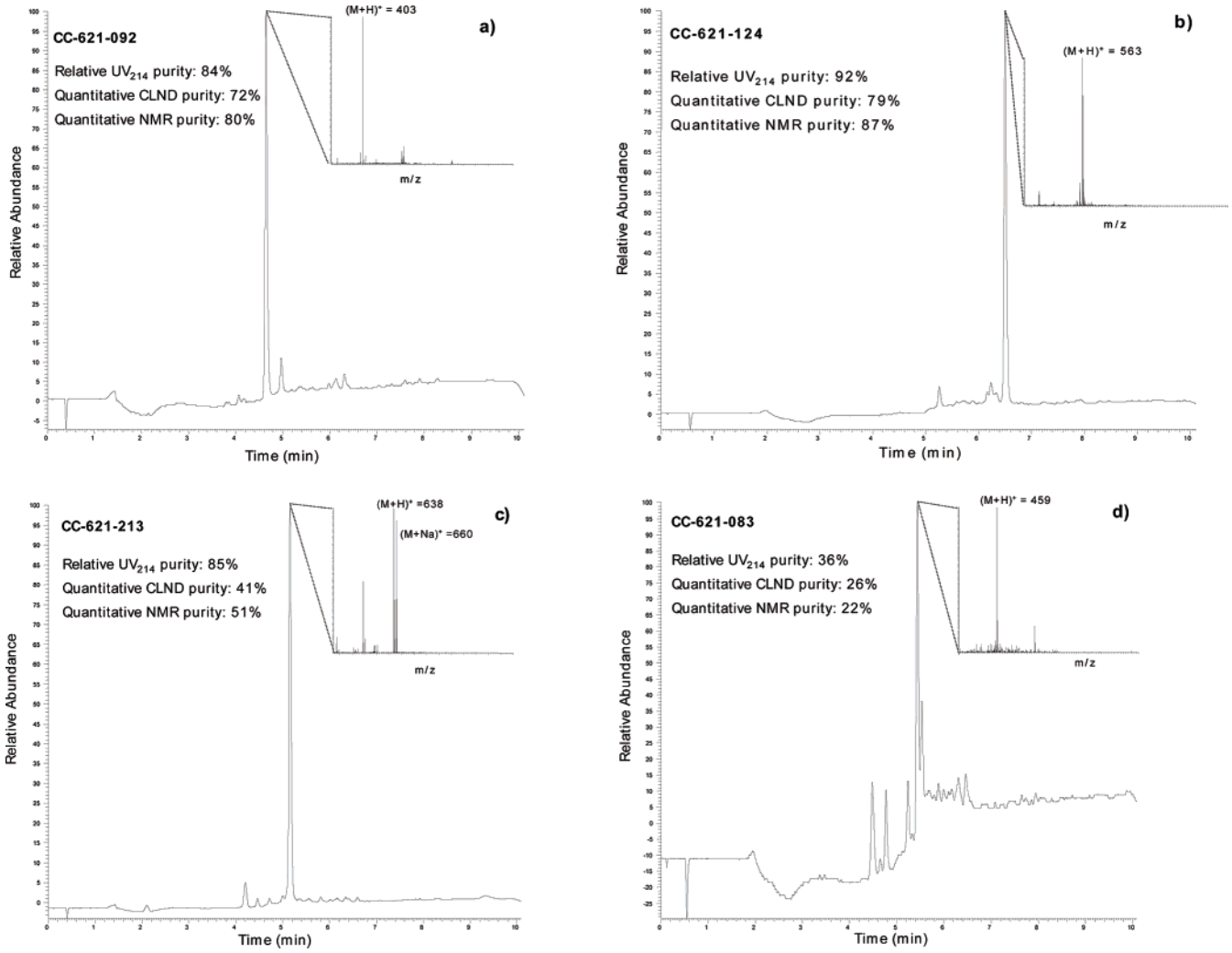


Figure 7. (a−c) Analytical LC−UV chromatograms at 214 nm of the three samples with absolute amounts (NMR) ≥51%, which represents the average quality of the investigated batch. (d) Analytical LC−UV chromatogram at 214 nm of the sample with an absolute amount (NMR) of 22%, which is worse than the average. The absolute amounts obtained by CLND are in good agreement with the NMR purities. The relative LC−UV purities correspond to the absolute purities in the cases in which the absolute purity is high. The insets display the corresponding MS spectra. The LC−UV traces were obtained from the RP-HPLC−UV−CLND−MS^n analysis.

(Figure 6d) shows several organic impurities. Nevertheless, the vinylic protons of the rotamers between 6.6 and 6.8 ppm (Figure 5a) could be used for quantification ($0.00017 \times 5.90 \times 0.55 \times 3.33 \times 458.45/1$) and delivered 1.3 mg (3.4 μmol). The ^{19}F spectra showed 0.6 equiv of TFA in the sample (0.2 mg), which results in an absolute purity of 22% of sample CC-621-083. In agreement with this result, the CLND quantification resulted in an absolute purity of 26%. In this example, the purification was not effective enough to separate organic impurities. This is confirmed by the analytical LC−MS which contains several additional peaks (Figure 7d).

As discussed in detail for the four samples, 621-083, 621-092, 621-124, and 621-213, the determination of the absolute purity was done for all 20 samples (Table 2). The purities obtained by ^{1}H NMR quantification varied from 22 to 87%, and the purities obtained by CLND quantification varied from 18 to 100%.

Discussion

Comparison between the ^{1}H NMR and the CLND Quantifications. In general, the purities obtained by NMR quantification can be considered to be more accurate than those obtained by CLND quantification. Indeed, in the NMR method, both reference signals (^{13}C-satellite and compound signals) are on the same spectrum, which means that the absolute purity of a sample is determined with a reference under the same conditions. The CLND quantification relies on a calibration and is, therefore, not obtained under absolutely identical conditions, such as the analyzed sample.

However, the differences in the purity values are not always of the same order for the two methods. For low concentrations, NMR may lose a part of its accuracy: the lower the concentration, the smaller the NMR signal of the compound, and the lower the signal-to-noise ratio. If this ratio is too small, it becomes difficult to integrate the signal properly due to the higher baseline fluctuations. This explains why bigger differences between the results of the two methods are often observed at lower sample concentrations (entries 3, 7, 12, 15, 17). The presence of impurity signals close to the peak of interest may also lead to unsound results in the NMR quantification method. The CLND method detects only compounds with nitrogen(s); if impurities do

not contain nitrogen, they are not detected, and the obtained absolute purity from CLND is higher than that from NMR (entries 7, 8, 13, 19, and 20).

^{1}H NMR Quantification. The visual inspection of all the ^{1}H NMR spectra at 393 K immediately demonstrates that there are only two principal classes of qualities observed: (1) good purity without any significant organic impurities in the sample; demonstrated for 621-092, 621-124, and 621-213, entries 4,9,17 (Figure 6a-c), but also true for entries 2, 3, 5−8, 10−14, 18, and 20; and (2) medium purity; signals of the expected compound can still be recognized, but a couple of minor organic impurities are observable; demonstrated for 621-083, entry 1 (Figure 6d), but also true for entries 15, 16, and 19.

A judgment of the value of the weighing cannot be made for class 2, because in these cases, the difference between weighing and NMR quantification cannot be explained only by water inclusion, invisible salts, and the weighing error, but also by other organic impurities which have not been quantified by NMR. However, the absolute purities in class (a) are always ≥48%. This means that the absolute error with respect to sample amounts in cases in which the purification could separate the expected product from other organic impurities and the isolated amounts of the expected product are ≥2.3 mg (entry 6) is, in general, less than a factor of 2. Of course, this factor might vary slightly from library to library; however, it was the goal of this study to get an impression of the order of magnitude of the weighing error. In our opinion, the observed dimension of the error factor, in comparison with the accumulated errors in reported high-throughput screening results, is not relevant and does not impact any screening results or any early attempts for structure−activity relationship information.

With these results we have reported, to our knowledge for the first time, data that show the absolute purity of isolated combichem samples and the relevance of weighing these samples at a scale of a few milligrams.

References and Notes

(1) Yan, B.; Fang, L.; Irving, M.; Zhang, S.; Boldi, A. M.; Woolard, F.; Johnson, C. R.; Kshirsagar, T.; Figliozzi, G. M.; Krueger, C. A.; Collins, N. *J. Comb. Chem*. **2003**, *5*, 547.

(2) Felder, E. R.; Scarpella, S.; Tatò, M.; *CHIMIA* **2003**, *57*, 229.

(3) Indeed, the method is so sensitive that differences in the concentration of the satellites from one batch to the other were detected.

(4) Taylor, E. W.; Qian, M. G.; Dollinger, G. D. *Anal. Chem.* **1998**, *70*, 3339.

CC049850R