Vidaud *Journal of Analytical Science and Technology* (2018) 9:13
https://doi.org/10.1186/s40543-018-0145-1

Journal of Analytical Science and Technology

RESEARCH ARTICLE Open Access

Mat-CHN software for drug substance titrations in discovery and development

Lionel Vidaud

Abstract

Background: Elemental Analysis (microanalysis) is a reference method widely used to make assessments of thequality of chemical or biochemical substances. Using very small quantities of substance and no methoddevelopment, this analytical technique delivers absolute assays. Compared to other methods needing larger samplequantities and most of the time reference standards, the Elemental Analysis is the preferred technique of scientists.

Methods: One of the difficulties with this method is the definition of the theoretical elemental composition when several complementary assays values, generally delivered in w/w percentages (e.g., water content by Karl Fischer method), instead of mole numbers (as this is the case for NMR determination), must be taken into account. The situation is manageable with only one of this type of additional assay; however, this exercise becomes more complicated when the scientists must integrate two or more assay values. Mat-CHN software was developed specifically to resolve these difficulties, by delivering the individual mole numberof each constituent assayed in the compound analyzed, leading to an accurate comparison of the theoretical and experimental results of centesimal compositions.

Results: Apart its large capabilities for simulations, Mat-CHN increases the interest of the Elemental Analysis technique for scientists working in discovery fields or when new molecular entities are promoted to pre-development and development stages, either for absolute titration of their reference standards or for direct titrations of the drug substance itself. The software conducts to simple titration determinations, based on Nitrogen "Kjeldhal" assays and/or molecular weight reports, compared to data of the pure substance.

Conclusions: Based on experimental analytical data, Mat-CHN software delivers accurate elemental content determinations and titration of the substance analyzed. These results should be considered also experimental ones, because Mat-CHN resolves the equations without any convergence or external parameters.

Background

Information: *raw formulas presented as examples in this document may correspond to some molecules described in other papers. If this situation should occur, it shall be considered as an absolute involuntary error from my part. But realistic raw formulas were needed to make the demonstrations of the practical interest and of the efficiency of Mat-CHN software.*

Decimal values reported in the different screenshots of the user interface are presented with the European decimal format (e.g., coma instead of dot).

Correspondence: lionel.vidaud@evotec.com
EVOTEC France SAS, 195 Route d'Espagne, 31036 Toulouse, France

Microanalysis capabilities–experimental and theoretical centesimal composition

Elemental Analysis equipment are able today to provide the assays of carbon (%C), hydrogen (%H), nitrogen (%N), sulfur (%S), and oxygen (%O), with a high degree of precision for known and, most important, for unknown products (Rouessac and Rouessac 2007). It must be emphasized that these results of centesimal composition are obtained with a very limited quantity of sample, generally around 1 mg and can be delivered with a medium throughput.

According to the different suppliers of equipment, a reliability of ± 0.3% for the three first elements (C, H, N) and of ± 0.6% for the sulfur made a consensus and seems sustained by the different users. Some laboratories offer more precise confidence intervals, according to the percentage

Springer Open

© The Author(s). 2018 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

levels of the element determined. For the three first elements (C, H, N), the more the content will be elevated, the more the enlargement of the confidence interval will be, generally ± 0.3%, but never more than ± 0.5% for the highest contents and never less than ± 0.01% for the lowest contents (not less than 0.10%) (Krotz et al. 2015).

Each assay of these elements corresponds to a direct absolute determination, just as it is provided by a potentiometric titration. With a quantity of approximately 100 fold less consumed per sample compared to the titration, the microanalysis technique takes a major place in discovery activities (Paul 2015; Fadeeva et al. 2008), where synthesized compounds are delivered in a range of 1 to 100 mg per batch.

If the technical aspect of this methodology has apparently no major drawback, except the cautions to apply for the weight of a very small amount, we should not forget that each element content must be compared to the theoretical one, to evaluate the degree of purity of the sample analyzed. If the objective of the scientist is to obtain a pure substance, and if theoretical contents are equal to the experimental ones, comparisons will be immediate and assessment of 100% of purity is easily achieved.

Most of the time, these two sets of data are not perfectly superimposable and scientists try to identify the supplementary components which are responsible of these differences. Consequently, the initial and expected formula of the pure compound will be re-adjusted with proportions of other constituents, until the theoretical elemental composition of the redefined formula will match with the one experimentally determined.

In addition, the scientists cannot make such proposals of additional mole numbers, if these suggested numbers are not supported by complementary analytical determinations. For instance, the number of moles of water in the main substance can be determined by a specific assay, based on the Karl Fisher method. But in this case and for the major part of the analytical determinations (except, the NMR analysis), results are delivered in weight/weight percentages. Consequently, the conversion of these results into mole numbers is mandatory to re-calculate the theoretical centesimal composition of the redefined raw formula of the compound. If this mole number is easily accessible for only one component included in the substance analyzed, the difficulty to calculate the mole number of each component increases with the number of components to be considered around the main compound.

Mat-CHN Software principles

Mat-CHN software was specifically developed, to simplify microanalysis interpretations, based on a unique recalculated composition, including all the individual assayed components and/or impurities in the analyzed substance. In a few seconds, Mat-CHN generates the corresponding mole number and composition of the whole redefined raw formula, leading to the possible comparison of the experimental microanalysis results and finally to the titration of the analyzed substance.

The principles of calculations of Mat-CHN do not include any approximation, recursive processes with target options (e.g., percent of convergence), and need nothing more than individual data experimentally determined. This means also that Mat-CHN's recalculated centesimal compositions must be considered as experimental results.

Mat-CHN software– calculation

Interface

The first operation of this software consists in the capture of experimental results and formula of the free molecule.

Figure 1 shows the user interface for entries of the needed data in the software and how the final results are provided.

Different boxes have been defined to give a view of the functionalities in this global table:

- {**Box 1**}: the user enters in the second column the raw formula of the pure compound and will obtain the redefined raw formula, after the calculation, in the fourth column (in red or gray). The fifth column will deliver the elemental composition re-calculated, which will be used for the comparisons with the experimental elemental analysis results (see {Box 6}).
- {**Box 2**}: four sets of supplementary entries are proposed to add raw formula and moles number of solvents (or impurities) quantified by NMR. As for box 1 only, the second column in each space is accessible by the user. This space is dedicated only to results directly delivered in mole numbers.
- {**Box 3**}: this box is very specific for entries of contents (in %*w*/*w*) of components assayed by dedicated techniques (except for water content, see {Box 5}). This box is supported by an application (button "organic and mineral assays residues"), to manage a personal library of pre-described and new-described components (including raw formula, molecular weight, charge, number of hydrogen atoms to add or to subtract). This box 3 can integrate up to 36 (18 per set of columns) contents of components. After the calculation process, mole numbers are reported (column "Nb moles," results in red or gray).
- {**Box 4**}: corresponds to the entry of the experimental results of the Elemental Analysis for the five elements (C, H, N, S, and O). Two sets of experimental data are recordable.

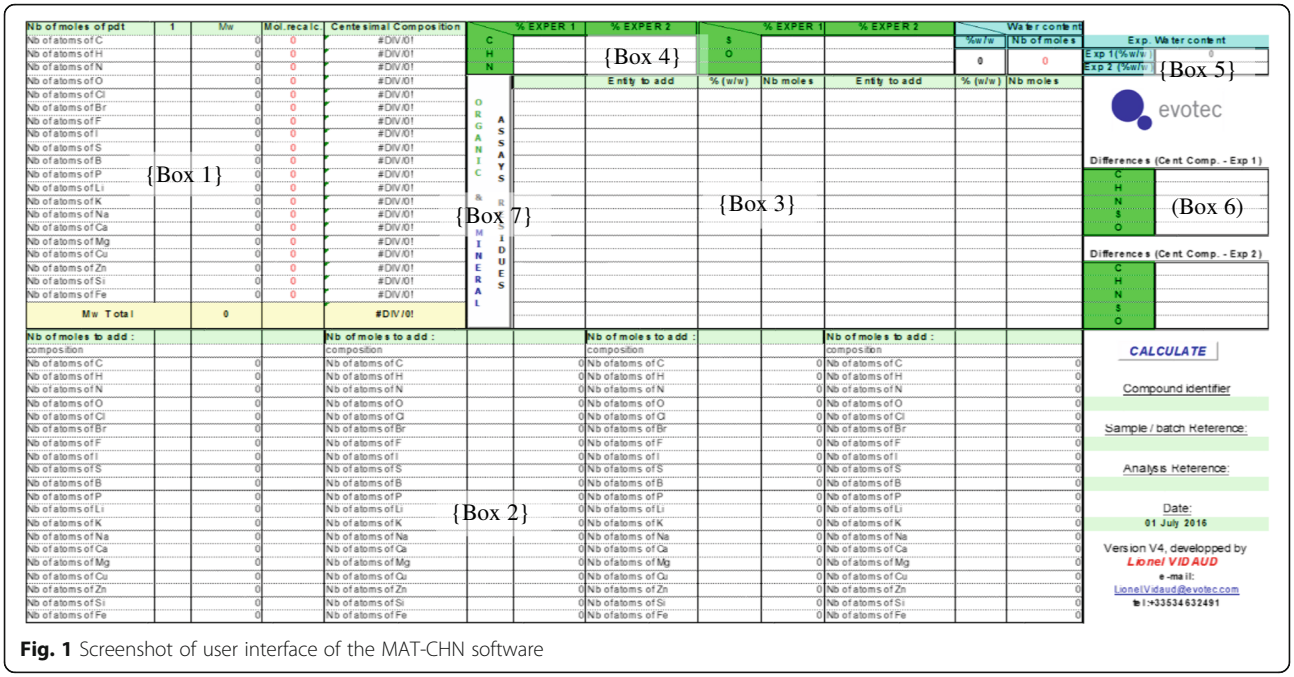

Fig. 1 Screenshot of user interface of the MAT-CHN software

- {**Box 5**}: is the specific place of the water content determination, two experimental results are recordable. On the left, the average content calculation is provided.
- {**Box 6**}: corresponds to the comparisons of the experimental and the recalculated elemental compositions, when the "CALCULATE" button is activated. A color coding is established to highlight results with acceptable criteria (≤± 0.3%) in green, critical results in yellow (between ±0.3% and ± 0.4%) and out of criteria (≥± 0.4%) in red.
- {**Box 7**}: corresponds to the button "Organics and mineral assays residues," opening the personal library of the users containing pre-defined substances, or ions which, when selected, will appear in the list of the {Box 3}. The user can add or suppress these defined compounds. The added substances are recorded, of course, with their respective raw formulae and charges. This functionality is needed to use Mat-CHN, but is not described here, because not directly belonging to the calculation principles.

Atomic elements proposed in this table are the most frequently observed in our chemical structures. Of course, the software can be modified according to the needs of the users and number of lines related to these elements can be increased in needed cases or decreased for a better clarity.

The molecular weights of each element used in this software are in accordance with IUPAC 2015 (IUPAC Commission on Isotopic Abundances and Atomic Weights 2015).

Calculations

Principles

The principal objective of the exploitation of the Elemental Analysis results is to determine the titration value of the sample analyzed. This value corresponds to the relationship between the experimental and theoretical centesimal compositions of the compound tested. Most of the time, scientists are used to calculate this relationship based on the carbon or the nitrogen content percentages.

As explained above, no difficulty is expected when mole number of each component of the compound is defined. On a simple and not consolidated approach, a scientist can make different suppositions of components and contents of components, to ensure the experimental elemental composition values are matching with the theoretical composition.

These suppositions are consolidated by complementary assays and the different delivered results (in *w*/*w*%) related to the substance tested must be converted into mole numbers.

The base of the calculation is summarized in the generic equation Eq. (1):

$$\frac{A}{100} = \frac{(nb.of\ mol\ x_A) \times m_A}{Mt} \quad (1)$$

m_A : molecular weight of the element or the component A

Mt : total molecular weight of the compound analyzed

Considering A as an element (C, H, N,...) or as a component (H_2O, impurity, Na^+ or Cl^- ions,...), the content,

resulting from the dedicated analysis, expressed in *w*/*w*% will follow this equation Eq. (1). The nb. of mol x_A is unknown, but can be easily extracted from this equation Eq. (1) if Mt is known.

The Mt value, corresponding to the total molecular weight of the analyzed compound, including components, should be determined according to this second equation:

$$\mathrm{Mt} = \mathrm{M(free)} + \sum C_x \times m_x - (\mathrm{e} \times C_x \times m_h) + \sum A_y \times m_y + (\mathrm{e} \times A_y \times m_h) + W_z \times m_z + \sum I_t \times m_t \quad (2)$$

in which:

- M(free) corresponds to the molecular weight of the pure substance
- $\Sigma C_x \times m_x - (\mathrm{e} \times C_x \times m_h)$, corresponds to the "molecular weight" of the moles number of cations, excluding the same mole number of hydrogen in the analyzed substance, as a result of the neutralization of acidic functions (in the case of alkaline salts). C_x is the mole number of the cation x and m_x its molecular weight, e is the charge number of x and m_h the molecular weight of the hydrogen atom.
- $\Sigma A_y \times m_y + (\mathrm{e} \times A_y \times m_h)$, corresponds to the "molecular weight" of the mole number of the anions, including the same mole number of hydrogen in the analyzed substance as a result of neutralization of the basic functions (in the case of acid salts). A_y represents the mole number of the anion y and m_y its molecular weight, e is the charge number of y and m_h the molecular weight of the hydrogen atom.

Remark: according to these second and third terms of the equation Eq. (2), in case of mineral components assays (e.g., sodium chloride, NaCl) by anionic and cationic titrations, the added hydrogen moles (A_y) coming from the anion content and the ones subtracted coming from the cation content (C_x) will produce of course the expected null sum ($e \times A_y \times m_h + e \times C_x \times m_h = 0$).

- $W_z \times m_z$, corresponds to the "molecular weight" of the mole number of water. W_z is the mole number of water and m_z its molecular weight.
- $\Sigma I_t \times m_t$, corresponds to the "molecular weight" of mole number of impurities quantified through dedicated techniques (HPLC, GC, etc..). I_t is the mole number of the impurity t and m_t its molecular weight.

In this equation Eq. (2), all the C_x, A_y, W_z, and I_t variables corresponding to mole numbers are unknown, making the direct calculation of Mt not accessible. As this can be anticipated, the resolution of the system of Equations (1) and (2) will produce the value of each unknown variable included in the expression of Mt, finally leading to the Mt value.

Of course, the two expressions of %A and of Mt can be completed with other specific analytical determinations, not reported here. However, it appeared that the two equations proposed covered the majority of the cases of our compounds synthesized in our context of the discovery and the development of new molecular entities (NMEs) for therapeutic applications.

Development of a mathematical software

The following approach is based on the two equations proposed (Eqs. (1) and (2)), which should calculate the number of moles of each component included in 1 mol of the substance analyzed, based on their individual analytical results. At the end of the process, the centesimal composition of the redefined and complete raw formula of the compound analyzed will be accessible and then compared to the experimental Elemental Analysis results.

As a simple example, if a substance contains only a few moles of water (represented by the %**W** content, determined for instance by Karl Fisher technique), combination of equations Eqs. (1) and (2), will produce to the simplified expression:

$$\frac{W_z \times m_z}{\mathrm{M(free)} + W_z \times m_z} = \frac{\%\mathbf{W}}{100}$$

in which Wz (nb of moles of water) can be easily extracted:

$W_z \times m_z \times (100 - \%\mathbf{W}) = \%\mathbf{W} \times \mathrm{M(free)}$ leading to $W_z = \frac{\%\mathbf{W} \times \mathrm{M(free)}}{m_z \times (100 - \%\mathbf{W})}$.

Based on this example, the raw formula of the compound is redefined, by adding $2 \times W_z$ moles of hydrogen atom and W_z moles of oxygen atom to the formula of the free compound and the theoretical centesimal composition recalculated accordingly.

Consequently, we show that based on an individual assay performed on the tested substance, a mole number of the assayed entity can be extracted by the resolution of the combined equations Eqs. (1) and (2). If multiple supplementary assays are performed, each individual mole number corresponding to each individual assay must be calculated by the resolution of a system of equations in which the Mt term (equation Eq. (2)) will be extended with the number of individual assay.

Today, it seems that each laboratory has its own approach to resolve these equations, even if some principles are identical for everyone. Some developed software proposes multiple combinations of contents of components included in the substance, offering the same centesimal composition results. These approaches are interesting to initiate a final choice of the most

appropriate combination by the scientist, but do not deliver a unique or the right solution of this composition.

The mathematical software development presented here was focused on this particular aspect, leading to a unique solution of contents of components to reach only one redefined raw formula of the analyzed compound, which is a critical point in our activities for research and development of NMEs.

For reasons of costs and also for having a simple and direct access to centesimal composition calculations, the software proposed here was designed to be supported by a generalized "Office" interface implemented on each computer in the company. The program is written in Visual Basic for Excel® and does not need any knowledge in this language for the user.

The objective of this program is not to replace what the other software can provide, but to give a simple and quick access to centesimal compositions, to everyone, through accurate calculations based on the resolution of a system of equations.

Resolution of one equation

Let us now consider the example of the content determination for one cation, C_1. The dedicated analytical method delivered the result %Ct (%*w*/*w*), assayed in the sampled substance.

Then, inserting this result in the equation Eq. (1) and replacing the Mt value by the equation Eq. (2), the expression of %Ct corresponds to the equation Eq. (3):

$$\frac{\boldsymbol{C_1}\times m_1}{\mathrm{M(free)}+\sum[\boldsymbol{C_x}\times m_x-(\mathrm{e}\times\boldsymbol{C_x}\times m_h)]+\sum[A_y\times m_y+(\mathrm{e}\times A_y\times m_h)]+W_z\times m_z+\sum I_t\times m_t}=\frac{\%\mathrm{Ct}}{100} \tag{3}$$

Consequently, C_1 can be extracted from the expression of this equation, according to the successive equations (Eqs. (4), (5), and (6):

$$\%\mathrm{Ct}\times\left\{\mathrm{M(free)}+\sum_1^{n1}[\boldsymbol{C_x}\times m_x-(\mathrm{e}\times\boldsymbol{C_x}\times m_h)]+\sum_1^{n2}\left[A_y\times m_y+\left(\mathrm{e}\times A_y\times m_h\right)\right]+W_z\times m_z+\sum_1^{n3}I_t\times m_t\right\}=\boldsymbol{C_1}\times m_1\times 100 \tag{4}$$

$$\boldsymbol{C_1}\times m_1\times 100-\%\mathrm{Ct}\times[\boldsymbol{C_1}\times m_1-(\mathrm{e}\times\boldsymbol{C_1}\times m_h)]-\%\mathrm{Ct}\times\left\{\sum_2^{n1}[C_x\times m_x-(\mathrm{e}\times C_x\times m_h)]+\sum_1^{n2}[A_y\times m_y+(\mathrm{e}\times A_y\times m_h)]+W_z\times m_z+\sum_1^{n3}I_t\times m_t\right\}=\%\mathrm{Ct}\times\mathrm{M(free)} \tag{5}$$

$$\boldsymbol{C_1}\times[m_1\times(100-\%\mathrm{Ct})-\%\mathrm{Ct}\times\mathrm{e}\times m_h]-\%\mathrm{Ct}\times\left\{\sum_2^{n1}[C_x\times m_x-(\mathrm{e}\times C_x\times m_h)]+\sum_1^{n2}[A_y\times m_y+(\mathrm{e}\times A_y\times m_h)]+W_z\times m_z+\sum_1^{n3}I_t\times m_t\right\}=\%\mathrm{Ct}\times\mathrm{M(free)} \tag{6}$$

The different parameters, %C, m_x, m_y, m_h, e, m_z, m_t, have known experimental or defined values. The other variables C_1, C_x, A_y, W_z, and I_t should be determined.

Trying to simplify the different terms of the equation Eq. (6), $P_{r,\,x}$ is included and represents the known and summarized parameter:

$$\boldsymbol{C_1}\times[m_1\times(100-\%\mathrm{Ct})-\%\mathrm{Ct}\times\mathrm{e}\times m_h]=\boldsymbol{C_1}\times P_{1,1} \tag{6a}$$

$$\%\mathrm{Ct}\times\left\{\sum_2^{n1}[C_x\times m_x-(\mathrm{e}\times C_x\times m_h)]\right\}=\sum_2^{n1}C_x\times[\%\mathrm{Ct}\times m_x-(\mathrm{e}\times\%\mathrm{Ct}\times m_h)]=\sum_2^{n1}C_x\times P_{2,x} \tag{6b}$$

$$\%\mathrm{Ct}\times\sum_1^{n2}\left[A_y\times m_y+\left(\mathrm{e}\times A_y\times m_h\right)\right]=\sum_1^{n2}\left[A_y\times\left[\%\mathrm{Ct}\times m_y+(\mathrm{e}\times\%\mathrm{Ct}\times m_h)\right]\right]=\sum_1^{n2}A_y\times P_{3,y} \tag{6c}$$

$$\%\mathrm{Ct}\times[W_z\times m_z]=W_z\times[\%\mathrm{Ct}\times m_z]=W_z\times P_{4,z} \tag{6d}$$

$$\%\mathrm{Ct} \times \left[\sum_{1}^{n3} I_t \times m_t\right] = \sum_{1}^{n3} I_t \times [\%\mathrm{Ct} \times m_t] = \sum_{1}^{n3} I_t \times P_{5,t} \quad (6e)$$

$$\%\mathrm{Ct} \times \mathrm{M(free)} = Y_1 \quad (6f)$$

According to these intermediate simplifications, the equation Eq. (6) can be clarified in equation Eq. (7), leading to the result of C_1, in equation Eq. (8):

$$P_{1,1} \times \boldsymbol{C_1} + \sum_{2}^{n1} C_x \times P_{2,x} + \sum_{1}^{n2} A_y \times P_{3,y} + W_z \times P_{4,z} + \sum_{1}^{n3} I_t \times P_{5,t} = Y_1 \quad (7)$$

$$\boldsymbol{C_1} = \frac{Y_1 - \left(\sum_{2}^{n1} C_x \times P_{2,x} + \sum_{1}^{n2} A_y \times P_{3,y} + W_z \times P_{4,z} + \sum_{1}^{n3} I_t \times P_{5,t}\right)}{P_{1,1}} \quad (8)$$

Generalization of the resolution of the system of equations

Equation (7) can conduct to a generalized system of equations, leading to the determinations of C_x, A_y, W_z, and I_t summarized in the following matrix equation:

$$P \times C = Y \quad (9)$$

in which:

-P is the square matrix containing the known parameters and their calculated combinations as described above (see equations Eq. (6)a–e.

-C corresponds to the matrix of mole numbers that must be calculated

-Y corresponds to the column of results containing each experimental assay result multiplied by the molecular weight of the free compound (according to Eq. (6f)).

The resolution of this system of equations, according to a matrix calculation, is proposed in equation Eq. (10) as following:

$$P^{-1} \times P \times C = P^{-1} \times Y$$

$$I \times C = P^{-1} \times Y \ (I : \text{identity matrix})$$

$$C = P^{-1} \times Y \quad (10)$$

The resolution conducts to the C matrix, containing each individual mole number of each component assayed.

This solution is valid if the P matrix is invertible, meaning that the determinant of this square matrix P is different from 0.

If, instead of the mean value, we want to integrate each individual result delivered for one assay of one component, the matrix P will never be a square matrix, and consequently, the resolution of the system of equations must be proposed differently.

Starting with the Equation (9), now embedding a non-square matrix P, the solution must pass by the transformation of this non-square matrix to a square matrix. For this, we must multiply this P matrix by its transposed (P^t) form:

$$P^t \times P \times C = P^t \times Y$$

Then, reproducing the same method conducting the equation (10):

$$(P^t \times P)^{-1} \times P^t \times P \times C = (P^t \times P)^{-1} \times P^t \times Y$$

$$I \times C = (P^t \times P)^{-1} \times P^t \times Y \ (I : \text{identity matrix})$$

$$\boldsymbol{C} = (\boldsymbol{P^t} \times \boldsymbol{P})^{-1} \times \boldsymbol{P^t} \times \boldsymbol{Y} \quad (11)$$

Remark: in the particular case of integrations of all individual experimental results collected for one assay and to preserve the generic Equation (11), each considered value must be previously divided by the total number of the related experiments performed.

Having defined the mathematical expression of the C Matrix of results, the algorithm of the software was established accordingly. In general terms, the calculation of the matrix C was based on the Gauss Seidel method, meaning that all the diagonal values of the matrix P must be different from 0. Then, a previous preparation of the matrix P was also included in the algorithm, based on the individual assay results. A management of a library of pre-defined components was also taken into account in this algorithm to simplify the entries of experimental results.

However, the objective of this work is not to detail the full software step by step, but to make the demonstration of the interest of this application, for initiated and non-initiated users.

USES of Mat-CHN in the discovery context

Performances of the software

As explained in the previous paragraph, the core of Mat-CHN software is a solver of systems of equations based on matrix calculations.

A second and non-negligible functionality of this software (not developed here) is the by-user management of its personal library of molecules or ions, for integration of pre-defined and new-defined components in the calculations. Using this library, the user will be unlimited in the choice of components. He will have also the possibilities to add or to delete components in this library every time.

It may be important to notice here that no recursive process is used in Mat-CHN for these mathematical resolutions, or convergence calculations, but only one direct calculation. This is also a critical aspect to obtain a quick and unique result.

To validate these performances, the best way is to illustrate some of them through a particular example, in this case quite complex, a situation which will be probably never encountered in reality.

Let us start with a crude compound (raw formula $C_{25}H_{25}N_8O_4Cl$, HCl) isolated before purification steps. This compound contains residual water, residual solvents, mineral charges, synthesis impurities, and traces of the precursor salt. We will make the assumption that each of these components have been assayed by dedicated techniques (e.g., water by Karl Fisher method, solvents by GC/head space, mineral charges by ion chromatography, and atomic absorption and impurities by liquid chromatography). Each of the 20 results in this example has been reported in *w*/*w* percentages. No quantified result was provided by NMR.

Calculation duration

The time spent to capture each experimental result and the pure raw formula not taken into account, the time needed to make the resolution of the system of 20 equations with 20 unknown variables (mole numbers), and the recalculation of the elemental composition were less than 3 s on a typical laptop (2.6 Ghz processor) that are currently available on the market today.

Results of these recalculations are illustrated in Figs. 2, 3, and 4 and as expected with a unique set of data. Results in Fig. 4 validates the correlation of the two sets of experimental elemental compositions with those recalculated for the three elements.

Range of values supported by the software:

As shown in this example, the software has resolved the system of equations whatever the range of data considered. To illustrate this performance, the content of chloride (5.6%), which represents a "major" content for the hydrochloride salt definition of this molecule, was recalculated into 1 mol per mole of compound.

On the other hand, the content of the residual benzene (8E−5%, or 0.8 ppm), delivered 6.5E−6 mol of this solvent per mole of compound. Similar ppm levels of residual contents have been also managed by Mat-CHN to deliver the corresponding low numbers of moles.

Interpretations of separated ion contents

As shown in Fig. 3 and because this was not directly apparent with results provided only in *w*/*w* percentages, we can observe:

-Ammonium content (0.73%) and acetate content (2.4%) correspond to ~ 0.26 mol of residual ammonium acetate salt, per mole of compound

-Calcium content (0.4%) and carbonate content (0.6%) correspond to ~ 0.06 mol of residual calcium carbonate salt, per mole of compound.

-Copper content (2E−5%) and iodide content (8E−5%) correspond to ~2e−6 mol of residual copper iodide salt, per mole of compound.

-Finally, para-toluene sulfonate content (0.35%) and potassium content (0.08%), correspond to ~ 0.13 mol of residual potassium para-toluene sulfonate salt, per mole of compound.

These interpretations would have never been accessible, if the corresponding mole numbers have not been calculated. This is another aspect of the performance of this software, useful for chemists and analysts.

Reliability of the software

Total number of hydrogen atoms in the redefined raw formula

As shown in Fig. 2, the initial hydrogen atom number was entered to 25 for the pure compound.

At the end of the calculation, this number increased to 29.93. Consequently, 4.93 hydrogen atoms are coming from the different components included in the final structure and this number is interpreted as presented in the Table 1:

The total of the figures reported in Table 1 represents the expected 4.93 of supplementary hydrogen atoms in the redefined formula. If we compare the exact number of hydrogen atoms added in the full structure (4.9279304. see Fig. 2) and the one recalculated here (4.926299912. see Table 1), the very low difference is interpreted by the non-absolute stoichiometries (linked to experimental assay errors) of mineral salts, bearing protons or not.

Nb of moles of pdt	1	Mw	Mol.recalc.	Centesimal Composition
Nb of atoms of C	25	312,5384921	26,0214984	49,369734
Nb of atoms of H	25	30,16576765	29,9279304	4,765096
Nb of atoms of N	8	117,2171689	8,36862184	18,516057
Nb of atoms of O	4	93,89103529	5,86839877	14,831374
Nb of atoms of Cl	1	71,45861096	2,01559908	11,287866
Nb of atoms of Br		0	0	0,000000
Nb of atoms of F		0	0	0,000000
Nb of atoms of I		0,000506445	3,9908E-06	0,000080
Nb of atoms of S		2,214306424	0,06905336	0,349780
Nb of atoms of B		0	0	0,000000
Nb of atoms of P		0	0	0,000000
Nb of atoms of Li		0	0	0,000000
Nb of atoms of K		0,506445497	0,01295313	0,080000
Nb of atoms of Na		2,459043049	0,10696249	0,388440
Nb of atoms of Ca		2,532227483	0,06318185	0,400000
Nb of atoms of Mg		0,063913618	0,00262964	0,010096
Nb of atoms of Cu		0,000126611	1,9924E-06	0,000020
Nb of atoms of Zn		0,009226664	0,00014112	0,001457
Nb of atoms of Si		0	0	0,000000
Nb of atoms of Fe		0	0	0,000000
Mw Total		**633,056871**		**100,000**

Fig. 2 User interface screenshot––initial (2nd column), recalculated (3rd, 4th columns) raw formulae of the compound and final recalculated centesimal composition (5th column)

Reconcile experimental and calculated contents of the components, based on their mole numbers and the molecular weight of the redefined raw formula

For this verification, we can consider the calculated mole number of the residual benzene and chloride.

In the case of the benzene, (molecular weight 78.112362 g.mol^{-1} proposed by the software), the calculated mole number is 6.48355E−06 (see Fig. 3) and the molecular weight of the complete structure is 633.054436 g.mol^{-1}, (see Fig. 2). The content is then calculated to 8.00E−5%*w*/*w*, in perfect accordance with the value experimentally obtained to 8E−5% *w*/*w*.

For the case of the chloride content (molecular weight 35.45279 g.mol^{-1}, proposed by the software) calculated to 1.000 mol, the same process produces to 5.60%*w*/*w*, exactly the same as the experimental value.

This simple exercise extended for the other components of Fig. 3 validates the perfect recovery between the

	% EXPER 1	% EXPER 2		% EXPER 1	% EXPER 2	Water content	
C	49,54	49,27	S			%w/w	Nb of moles
H	4,55	4,68	O			2,4	0,843357849
N	18,62	18,38					

		Entity to add	% (w/w)	Nb moles	Entity to add	% (w/w)	Nb moles
ORGANIC & MINERAL	ASSAYS RESIDUES	Benzene (C6H6)	0,00008	6,48355E-06	Chloride (Cl -)	5,6	0,999954722
		Sodium Sulfate (Na2SO4)	1,2	0,053481245			
		des-Cl Impurity (C25H26N8O4)	0,5	0,006298725			
		Ethanol (C2H6O)	0,0002	2,74832E-05			
		Ammonium (NH4 +)	0,73	0,256191268			
		Acetate (CH3COO -)	2,4	0,25732162			
		Calcium (Ca 2+)	0,4	0,063181851			
		Carbonate (CO3 2-)	0,6	0,063296119			
		Pentane (C5H12)	0,000065	5,70327E-06			
		Di-Cl Impurity (C25H24N8O4Cl2)	0,7	0,007755097			
		Cupper II (Cu 2+)	0,00002	1,99243E-06			
		Iodide (I -)	0,00008	3,99076E-06			
		Para Toluene Sulfonate (C7H7SO3 -)	0,35	0,012942471			
		Potassium (K +)	0,08	0,012953133			
		Ethyl Acetate (C4H8O2)	0,003	0,000215556			
		Magnesium Sulfate (MgSO4)	0,05	0,002629642			
		Methylene Chloride (CH2Cl2)	0,0009	6,7083E-05			
		Zinc fumarate (ZnC4H2O4)	0,004	0,000141119			

Fig. 3 User interface screenshot––experimental Elemental Analysis results (CHNSO)––recalculated nb of moles of water (8th column)––recalculated nb of moles (5th and 8th columns) of each of the 20 components

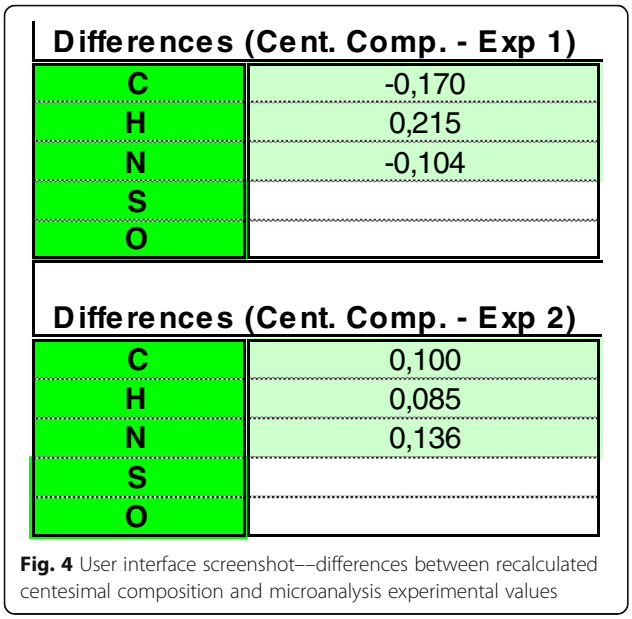

Fig. 4 User interface screenshot––differences between recalculated centesimal composition and microanalysis experimental values

calculated and the experimental results, as shown in Table 2.

Advantages provided by Mat-CHN in discovery field of fine chemistry

As this was developed above, the Elemental Analysis technique is today a robust methodology delivering absolute contents of the principal elements included in organic substances, mainly %C. %H, and %N. The reliability of this technique is principally dependent of the precision of the micro-balances used to weight the very low quantities of sample to be analyzed. However, this is a serious advantage in the discovery activities field, where compounds are synthesized around and frequently less than the 1 g scale.

If the experimental aspects have no drawbacks, the difficulty consists in the interpretation of the results provided by this technique to make accurate correlation with the molecular structure analyzed and refined through several complementary assays.

Non-absolute selectivity of the steps in chemical synthesis or hemisynthesis or substance extractions are most of the time followed by isolation and purification steps leading to expected purified compounds. All of these successive operations are generating organic and mineral impurities and residual traces of solvents, frequently included at low levels in the final products. We should not exclude also the impurities arising from the quality of the different raw materials (structure analogs) involved in the synthesis processes. Impurities resulting from non-fully selective chemical reactions must be considered also.

Aware of these principal aspects of the chemistry, chemical companies and providers of analytical equipment have enlarged their analytical efforts to characterize these by-products (covering a large panel of dedicated techniques able to quantify from percent to ppb content levels), leading to supplementary results. But few efforts were made to conciliate these complementary results with those provided by the Elemental Analysis experiments.

Our experience with Mat-CHN in the context of the discovery activities for medicinal chemistry demonstrates our capabilities to understand deeply how the synthesis and purification steps have been conducted, leading to improvements in the quality of the expected active substances. The analytical support, in which Elemental

Table 1 Protons added in the complete molecular structure

Entities (raw formula)––{Box3} and {Box5}	Nb of H per molecular structure	Nb of moles calculated	Nb of H atoms re-calculated
Hydrogen chloride	1	0.999954722	0.999954722
Water	2	0.843357849	1.686715698
Benzene (C6H6)	6	6.48355E−06	3.89013E−05
Des-Cl impurity (C25H26N8O4)	26	0.006298725	0.163766841
Ethanol (C2H6O)	6	2.74832E−05	0.000164899
Ammonium (NH4 +)	4	0.256191268	1.024765072
Acetate (CH3COO −)	3	0.25732162	0.77196486
Pentane (C5H12)	12	5.70327E−06	6.84393E−05
Di-Cl impurity (C25H24N8O4Cl2)	24	0.007755097	0.186122325
Para-toluene sulfonate (C7H7O3S −)	7	0.012942471	0.090597299
Ethyl acetate (C4H8O2)	8	0.000215556	0.00172445
Methylene chloride (CH2Cl2)	2	6.7083E−05	0.000134166
Zinc fumarate (ZnC4H2O4)	2	0.000141119	0.000282239
Total			4.926299912

Table 2 Comparison of experimental and Mat-CHN recalculated %*w/w* of each component

MAT-CHN - Molecular weight of the complete structure: 633.052084 g.mol^{-1}					
Components	Mat-CHN Mw (g mol^{-1})	Mat-CHN Nb of moles calculated	%*w/w* re-calculated	Experimental %*w/w*	Δ (% recalculated −% experimental)
Chloride	35.452790	0.999954722	5.600042344	5.6	4.23435E−05
Water	18.015324	0.843357849	2.400018147	2.4	1.81472E−05
Benzene (C6H6)	78.112362	6.48355E−06	8.00006E−05	0.00008	6.04907E−10
Sodium sulfate (Na2SO4)	142.043860	0.053481245	1.200009074	1.2	9.07361E−06
Des-Cl impurity (C25H26N8O4)	502.527818	0.006298725	0.500003781	0.5	3.78067E−06
Ethanol (C2H6O)	46.068672	2.74832E−05	0.000200002	0.0002	1.51227E−09
Ammonium (NH4 +)	18.038535	0.256191268	0.73000552	0.73	5.51978E−06
Acetate (CH3COO −)	59.044261	0.25732162	2.400018147	2.4	1.81472E−05
Calcium (Ca 2+)	40.078400	0.063181851	0.400003025	0.4	3.02454E−06
Carbonate (CO3 2−)	60.009070	0.063296119	0.600004537	0.6	4.5368E−06
Pentane (C5H12)	72.149264	5.70327E−06	6.50005E−05	0.000065	4.91487E−10
Di-Cl impurity (C25H24N8O4Cl2)	571.417504	0.007755097	0.700005293	0.7	5.29294E−06
Cupper II (Cu 2+)	63.546300	1.99243E−06	2.00002E−05	0.00002	1.51227E−10
Iodide (I −)	126.904473	3.99076E−06	8.00006E−05	0.00008	6.04907E−10
Para-toluene sulfonate (C7H8O3S −)	171.195979	0.012942471	0.350002646	0.35	2.64647E−06
Potassium (K +)	39.098300	0.012953133	0.080000605	0.08	6.04907E−07
Ethyl acetate (C4H8O2)	88.105556	0.000215556	0.003000023	0.003	2.2684E−08
Magnesium sulfate (MgSO4)	120.369380	0.002629642	0.050000378	0.05	3.78067E−07
Methylene chloride (CH2Cl2)	84.932254	6.7083E−05	0.000900007	0.0009	6.80521E−09
Zinc fumarate (ZnC4H2O4)	179.438734	0.000141119	0.00400003	0.004	3.02454E−08

Analysis is the unique analytical technique used to make absolute assays at this stage of the value chain, is now consolidated when combined with Mat-CHN. This software is probably one of the first able to reconcile all the individual complementary analytical determinations leading to the redefinition of molecular structure of compounds, conducting to their titration values. Having in hand such precious result, decision regarding the compound analyzed can be to reprocess the batch to eliminate too elevated contents of undesirable by-products, or to use the ratio provided between the active substance and its redefined structure, to anticipate its accurate amount to be administered for in vitro or in vivo tests.

Uses of Mat-CHN software for pre-development and development purposes

Another and non-negligible aspect of the Elemental Analysis technique is the one related to the pre-development and development fields of activities.

Currently, titration determinations for batch controls of drug substances are provided against reference standards, for which definition of the strength is based on specific chemistry processes and accuracy of analytical methods.

This means that the absolute assay value of the primary standard (governing the assay value of the reference standards batches) should be determined accurately and verified on a regular basis.

Of course, the greatest the purity of this particular batch will be, the less the difficulty to define the assay value around 100% will be. This is frequently the first strategy applied by the companies, which make the statement on the absolute purity of the primary standard batch, when the last recycled purification did not lead to a lower impurity profile than the one obtained during the same penultimate operation.

In the case of small molecules and due to larger quantities of sample available, absolute methods, like potentiometric titrations in non-aqueous media are also used, with respect to the repeatability of the determinations which must be in the expected range. This exercise needs some effort to develop the appropriate method. Most of the time is spent to find the adapted non-aqueous medium to get the best potentiometric amplitude signal to increase the accuracy of the inflection point of the titration curve.

However, quite frequently, these titrations are not as accurate as expected and sometimes not applicable due to the low reactivity of the substance against reactants. More significant are the cases of peptides, proteins, and antibody drug conjugates (ADCs).

In these situations, other validated analytical techniques must be put in place, and due to the impact on the regulatory aspects, they must be recognized by the health authorities (Crowther 2001).

Applicable to all the considered substances, the particular case of the nitrogen content assay is probably the second method of choice after unsuccessful trials with the potentiometric titrations. The assay recognized by the authorities is the nitrogen assay by the Kjeldhal method (referenced 2.5.9 in the European Pharmacopeia (EP) and <461> method II in the United States Pharmacopeia (USP)). This method assay substances bearing nitrogen atoms, including at the same time some impurities, as this is the case with the potentiometric titration.

Since the time when the Elemental Analysis technique was developed and improved, a number of comparisons with the Kjeldhal method were conducted (Gervasio Pereira et al. 2006; Oftedal et al. 2014; dos Santos et al. 2009), and globally, it was accepted that the two methods lead to close results.

In an initiative of the USP Monograph Modernization (Ouderkirk and Seo 2010), nitrogen assay test (USP <461>), was pointed out as non-specific in the monographs of the Povidone, Crospovidone, and Copovidone. Following this initiative, in 2012, the expert panel made recommendations to add some complementary and specific tests (Block 2013). Elemental Analysis which was not included until that time was finally suggested as an identity method combined with the total sulfate ash content test, for adulteration controls (e.g., melamine) showing the increasing interest of this analytical method in the area of controls for drug substances.

In the case of peptides, proteins, and antibody drug conjugates, the Elemental Analysis method was recommended as an additional test by the USP (e.g., <461> method 7) and EP (e.g., 2.5.33 method 7), for the assay of these bio-molecules.

According to these preliminary information, in the field of the pre-development and development of new molecular entities (NMEs), Elemental Analysis cannot be excluded from the standard methodologies to control these active substances and their reference standards (Crowther 2001) even as a main method and/or complementary method.

As discussed previously, this method can be suggested for assay determinations, based on the nitrogen content, equivalent to the Kjeldhal procedure.

In the light of the developed example in the paragraph of the discovery field of activities, the paradox of the Elemental Analysis technique, delivering experimental reliable results, comes from the evaluation or recalculations of the centesimal composition of the compound after integration of its associated mineral and organic component contents. This is also what underlined indirectly the expert panel of the US Pharmacopeia by making the association of the Elemental Analysis results with the sulfated ashes in the Povidone case discussed above.

Again, Mat-CHN was developed to make these missing reconciliations of these independent experimental results, leading to a perfect resolution of the initiated system of equations, to get a fine-tuned centesimal composition of the analyzed compound.

Case study of a drug substance analyzed for development purposes

Example of analytical results:

To illustrate the use of Mat-CHN in the scope of the development activities, another example is given here. The same substance previously studied, (raw formula $C_{25}H_{25}N_8O_4Cl$, HCl), is now considered purified, for development purposes. Only the impurity profile is remaining, meaning des-chlorinated and di-chlorinated impurities, with residual water and residual butanol-1 contents.

The hydrochloride salt content is determined based on three individual trials.

Analytical results are provided in the Table 3.

A first approach of the elemental composition

A first Mat-CHN use, not taking into account the contents of the two impurities (des- and di-chlorinated impurities) and residual butanol-1, is tested according to Fig. 5.

After less than 1 s of calculation, Mat-CHN returned the results reported in Fig. 6.

Different observations can be issued from Fig. 6:

1) Experimental Elemental Analysis results (part c) are not in perfect accordance with those provided by Mat-CHN, due to the 0.350% difference for the nitrogen content, beyond the acceptance criteria. Consequently, the set of chloride assays alone (part b), integrated into these calculations was not sufficient to make an acceptable correlation of the recalculated composition with the Elemental Analysis results.
2) The hydrochloride content corresponds approximately to 1 mol per mole of substance, more precisely 0.98 mol [according to the total chloride contents (1.98–1, 1 coming from the substance itself, see Fig. 6 part a), or according to the sum of the individual chloride assays, (*each individual result divided by the total number of*

Table 3 Experimental results for the compound $C_{25}H_{25}N_8O_4Cl$, HCl

Analytical determination	Method	Experimental results (%*w*/*w*)
Water content	Karl Fischer	0.80
Water content	Karl Fischer	0.85
Butanol-1	CPG-head space	0.35 (3500 ppm)
Chloride ion content	Ionic chromatography	6.21[(1)]
Chloride ion content	Ionic chromatography	5.82[(1)]
Chloride ion content	Ionic chromatography	6.03[(1)]
Des-chlorinated Impurity	HPLC	0.3
Di-chlorinated Impurity	HPLC	0.6
%C	Elemental analysis	51.720
%H	Elemental analysis	4.708
%N	Elemental analysis	19.054

[1]As a remainder, each of these individual results will be divided by 3 (three values for the same series of results) for Mat-CHN exploitation

assays, three in this series), 0.337 + 0.316 + 0.327 = 0.980, see Fig. 6 part b)].

3) The water content corresponds to 0.27 mol per mole of substance and justifies the increase of the total number of hydrogen atoms in the redefined compound (2 × 0.265 coming from water, 0.98 for the hydrochloride and the initial 25 of the substance, leading to the total of 26.51 hydrogen atoms).

Optimization of the elemental composition

To improve this first step of calculation, the set of chloride assay values was completed with the contents of residual butanol-1 (0.35%) and of each impurity, respectively, 0.3% for the des-chlorinated impurity and 0.6% for the di-chlorinated impurity. This new set of data was re-submitted to Mat-CHN and results are reported in Fig. 7.

From this second exploitation by Mat-CHN, we can conclude:

1) All the individual assay results included in the calculation are now consistent with the experimental Elemental Analysis values because differences observed for each %C, %H, and %N element are less than ± 0.3%, within the range of the acceptance criteria.
2) The hydrochloride content is close to 1.00 mol per mole of compound (0.994 mol corresponding to the sum of the three individual number of moles re-calculated, see second table of Fig. 7). Finally, the chloride mole number (2.006) is decomposed in 1 mol belonging to the substance, 0.0124 mol (2 × 0.0062) for the di-chlorinated impurity and 0.994 mol for the hydrochloride. But the two impurities are also hydrochlorinated. Then, the mole number of hydrochloric acid per mole of active

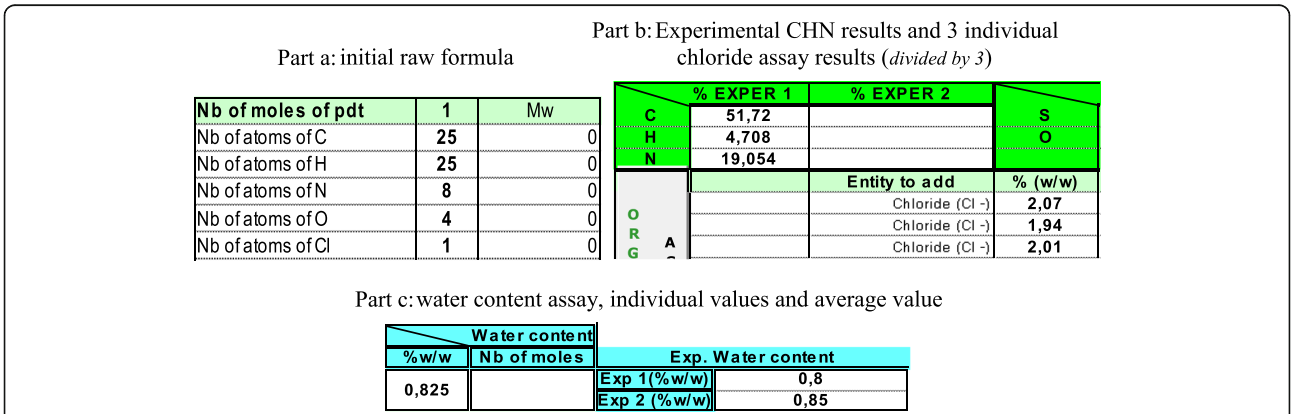

Fig. 5 User interface screenshots (parts a, b, and c)––first set of experimental results, except those related to the impurity profile, inserted in Mat-CHN for the compound $C_{25}H_{25}N_8O_4Cl$, HCl

Part a: recalculated numbers of atoms of the redefined formula

Nb of moles of pdt	1	Mw	Mol.recalc.	Centesimal Composition
Nb of atoms of C	25	300,2695	25	51,995596
Nb of atoms of H	25	26,72018429	26,5095132	4,626950
Nb of atoms of N	8	112,053976	8	19,403613
Nb of atoms of O	4	68,22889556	4,26445789	11,814727
Nb of atoms of Cl	1	70,21770362	1,9805974	12,159115

Part b: recalculated moles numbers of water and individual chloride ions (*divided by 3*)

	% EXPER 1	% EXPER 2		% EXPER 1	% EXPER 2	Water content	
C	51,72		S			%w/w	Nb of moles
H	4,708		O			0,825	0,264457894
N	19,054						

	Entity to add	% (w/w)	Nb moles	Entity to add	% (w/w)	Nb moles
ORG A	Chloride (Cl -)	2,07	0,337182162			
	Chloride (Cl -)	1,94	0,31600647			
	Chloride (Cl -)	2,01	0,327408766			

Part c: Differences of recalculated centesimal composition and microanalysis experimental values

Differences (Cent. Comp. - Exp 1)	
C	0,276
H	-0,081
N	0,350

Fig. 6 User interface screenshots (parts a, b, c)—Mat-CHN exploitation tacking into account the first set of results for the compound $C_{25}H_{25}N_8O_4Cl$, HCl

substance is 0.984 (0.994–(0.0062 + 0.0035)). This result is very close to the one observed in the first exploitation (0.980).

3) The water content was found, after calculations to 0.27 mol per mole of compound, in perfect accordance with the previous result (0.27 mol).
4) Consequently, the redefined raw formula is the following:
$C_{25}H_{25}N_8O_4Cl$, 0.994HCl, 0.0035 $C_{25}H_{26}N_8O_4$, 0.0062 $C_{25}H_{24}N_8O_4Cl_2$, 0.028 $C_4H_{10}O$, 0.27H_2O.

The titration of the drug substance in this sample can be now calculated based on the ratio of the molecular weight of the non-salified substance to the weight of the refined raw formula described above:

$\frac{536.973\times100}{585.359} = 91.73\%$ (expressed on the dried and non-salified substance).

or, based on the nitrogen assay, in regard with a Kjeldahl assay analogy, (see nitrogen result in the first table of the Fig. 7), the calculation leads to:

Part a: recalculated numbers of atoms of the redefined formula

Nb of moles of pdt	1	Mw	Mol.recalc.	Centesimal Composition
Nb of atoms of C	25	304,4922908	25,3515834	52,018010
Nb of atoms of H	25	27,25978138	27,0448559	4,656931
Nb of atoms of N	8	113,1342751	8,07712705	19,327320
Nb of atoms of O	4	69,34577526	4,33426536	11,846701
Nb of atoms of Cl	1	71,12723696	2,00625217	12,151038

Part b: recalculated moles numbers of water and individual impurities

	% EXPER 1	% EXPER 2		% EXPER 1	% EXPER 2	Water content	
C	51,72		S			%w/w	Nb of moles
H	4,708		O			0,825	0,268061497
N	19,054						

	Entity to add	% (w/w)	Nb moles	Entity to add	% (w/w)	Nb moles
ORGANIC ASSAY	Chloride (Cl -)	2,07	0,341776733			
	Chloride (Cl -)	1,94	0,320312494			
	Chloride (Cl -)	2,01	0,331870161			
	Butanol-1 (C4H10O)	0,35	0,027640339			
	De-Chlorinated Impurity (C25H26N8O4)	0,3	0,003494489			
	Di-Chlorinated Impurity (C25H24N8O4Cl2)	0,6	0,006146392			

Part c: Differences of recalculated centesimal composition and microanalysis experimental values

Differences (Cent. Comp. - Exp 1)	
C	0,298
H	-0,051
N	0,273

Fig. 7 User interface screenshots (parts a, b, c)––second exploitation of experimental results, including those related to the impurity profile, by Mat-CHN for the compound $C_{25}H_{25}N_8O_4Cl$, HCl

$$\frac{19.327 \times \left[\left(\frac{8}{8}\right) - 0.0035 \times \left(\frac{8}{8}\right) - 0.0062 \times \left(\frac{8}{8}\right)\right] \times 536.973}{8 \times 14.006747} = 91.72\%$$

(expressed on the dried and non-salified substance) according to the percentage of nitrogen of the pure substance (20.868%).

As shown in the two results expressed above, only data provided by Mat-CHN are considered (molecular weight of the redefined substance for the first and nitrogen content and mole numbers of nitrogen impurities for the second). The titration value, 91.72%, must be compared to a quick calculation based on the sum of the impurities quantified subtracted from 100%, leading to 91.91%, in the range of 0.3% of reliability for the microanalysis technique:

$$100 - [2.07 + 1.94 + 2.01 + 0.35 + 0.3 + 0.6 + 0.825] = 91.91\%$$

Reliability of the drug substance titration

Uncertainty of the titration value can be easily evaluated because this is dependent on the uncertainty of the measurements of each experimental value.

Examples of confidence intervals of the individual assays have been used to illustrate the impacts on the nitrogen titration, as this is summarized in Table 4.

As shown in Table 4, the impact of the different uncertainties of measurements of each individual assay, on the titration calculation, generates approximately 0.038% of relative uncertainty for the final result 91.72% (see last row of columns 4 and 5), corresponding to ~ 0.04% of absolute uncertainty. Consequently, the reliability of the titration, including the reliability of the microanalysis determination for the nitrogen content is 0.34% (0.3 + 0.04%).

Contribution of the impurity profile in the refined raw formula

As discussed above, the contributions of the impurity profile and of the residual solvent cannot be dissociated from this analysis. This aspect can be deeply interpreted, considering the influence of the different combinations of the corresponding contents. This study is summarized in Table 5.

As shown in Table 5, the combination "Nb 1" corresponds to the case developed above where residual solvent and impurity profile were not included in the calculations. Except the combination Nb 6, in which the des-chlorinated impurity content is ignored, only the combination Nb 8 verifies the experimental Elemental Analysis results, in which the content differences for each of the three elements (C, H, and N) are in the acceptance criteria (± 0.3%).

Supported by the quick calculations issued from Mat-CHN, the eight combinations of contents of each of the three components tested demonstrate that they cannot be dissociated to ensure the correlation with the experimental Elemental Analysis results.

Impurity profile in the refined raw formula––cases of related substances and of enantiomeric impurity

Mat-CHN is also able to integrate the contents of the related substances and the ones of the enantiomers (or stereoisomers).

In these particular cases, the user is invited to record the raw formula of the active substance in the personal library and to associate the related contents (see example in Fig. 8).

For the related substances, the user can detail each content of each impurity (as this was applied for the chloride contents in the previous example) or provide the sum of each content.

After calculations, these contents will be converted into mole numbers, which will impact the titration value of the analyzed substance. Of course, no change is expected regarding the elemental composition because the raw formula of these impurities is the same than the one of the pure compound.

For instance, 0.5% of the enantiomer of the analyzed substance discussed in this paragraph was added in the last updated impurity profile. As expected, the titration value became 91.23%, instead of 91.73%, based on the "Kjeldahl" calculation. Of course and more easy is the direct subtraction of the 0.5% of the enantiomer content from the 91.73% titration value. In fact, this second way of calculation validates the 0.0055 mol of the enantiomer calculated by Mat-CHN, leading to the titration value obtained by the Kjeldahl approach.

Also anticipated, the differences between the experimental elemental composition and the recalculated one stayed unchanged, with or without the insertion of this enantiomeric content.

These features are very helpful because the user can enter all the analytical results obtained during the control of the drug substance, paying attention to indicate the raw formula of each considered impurity.

Advantages provided by Mat-CHN in the field of pre-development and development activities

The pre-cited advantages, identified for the discovery field of activities, can also be applied in the case of the development and pre-development fields.

A particular point of interest is to help the users (chemists and analysts), with the mineral by-product

Table 4 Confidence interval of the titration value based on the nitrogen content, according to examples of uncertainty of measurements of each individual assay––Mat-CHN recalculations

Experimental determination	Value (%*w*/*w*)	Uncertainty of measurement (examples)	Lower value (%*w*/*w*)	Higher value (%*w*/*w*)	Lower mole number recalculated	Higher mole number recalculated
Water	0.80	5% (relative)	0.76	0.84	0.2540	0.2820
Water	0.85	5% (relative)	0.81	0.89		
Butanol-1	0.35	5% (relative)	0.3325	0.3675	0.0262	0.0292
Des-Chlorinated impurity	0.3	30% (relative)	0.21	0.39	0.0024	0.0046
Di-chlorinated impurity	0.6	20% (relative)	0.48	0.72	0,0049	0.0074
Chloride	6.21	2% (relative)	6.09	6.33	0.3338	0.3500
Chloride	5.82	2% (relative)	5.70	5.94	0.3124	0.3283
Chloride	6.03	2% (relative)	5.91	6.15	0.3239	0.3399
%N recalculated	19.327 (e)	$\frac{\lvert r1-e\rvert+\lvert r2-e\rvert}{2} = \pm 0.038\%$ (relative)	19.365** (*r1*)	19.289*** (*r2*)		
Titration* based on Nitrogen content	91.72	0.038% (relative)	91.68	91.75		

*Titration calculation (see previous paragraph)
**%N recalculated with each value of the different contents of the same column
***%N recalculated with each value of the different contents of the same column

identifications (e.g., mole number for cation is equivalent, according to the valence, to the mole number of anion, see Table 2). Without calculated mole numbers, there is no evidence to recognize which cation is combined with which anion.

Another particular point is the capability of the software to resolve its calculations, by managing low contents (ppm and ppb scale) mixed with higher contents (percent scale), which is a frequent situation in the cases of control monographs, where ppm of residual solvents results are presented with percent contents of water or other impurities.

The exercise presented above shows clearly the positive impact of Mat-CHN used beside other analytical results, including of course those of Elemental Analysis.

Proof is provided here, that Mat-CHN can resolve calculations integrating all the available results (included those of related substances and isomers impurities) expressed in %*w*/*w* and in mole numbers (e.g., NMR), to produce finally an accurate titration of the substance analyzed.

No choice is presented to the user for the redefined structure of the molecule analyzed. Only one structure is possible. This point constitutes a serious advantage, in cases where the user should justify the best selected structure among different possibilities (as this is proposed by some software), based on criteria that should be clarified in the context of the development of NMEs.

Table 5 Contributions of the different combinations of residual solvent and impurities contents on the experimental Elemental Analysis results

Impurities and residual solvent contents combinations inserted in Mat-CHN for centesimal composition calculations				Experimental Elemental Analysis results		
				%C: 51.720	%H: 4.708	%N: 19.054
Nb	Butanol-1 (%)	Des-chlorinated impurity (%)	Di-chlorinated impurity (%)	CHN content differences between combinations and experimental results		
				Δ *%C*	Δ *%H*	Δ *%N*
1				0.276	− 0.081	*0.350*
2	0.35			*0.307*	− 0.050	0.277
3		0.3		0.287	− 0.079	*0.354*
4			0.6	0.255	− 0.084	*0.342*
5	0.35	0.3		*0.318*	− 0.048	0.281
6	0.35		0.6	0.287	− 0.053	0.269
7		0.3	0.6	0.267	− 0.082	*0.346*
8	0.35	0.3	0.6	0.298	− 0.051	0.273

Italic figures are out of the acceptance criteria

Part a: recalculated numbers of atoms of the redefined formula

Nb of moles of pdt	1	Mw	Mol.recalc.	Centesimal Composition
Nb of atoms of C	25	306,1610371	25,4905208	52,018010
Nb of atoms of H	25	27,40917649	27,1930731	4,656931
Nb of atoms of N	8	113,7542987	8,12139312	19,327320
Nb of atoms of O	4	69,72581939	4,35801897	11,846701
Nb of atoms of Cl	1	71,51704424	2,01724728	12,151038

Part b: recalculated moles numbers of water and individual impurities

	% EXPER 1	% EXPER 2		% EXPER 1	% EXPER 2	Water content	
C	51,72		S			%w/w	Nb of moles
H	4,708		O			0,825	0,269530587
N	19,054						

	Entity to add	% (w/w)	Nb moles	Entity to add	% (w/w)	Nb moles
ORGANIC ASSAYS	Chloride (Cl -)	2,07	0,34364981			
	Chloride (Cl -)	1,94	0,32206794			
	Chloride (Cl -)	2,01	0,33368895			
	Butanol-1 (C4H10O)	0,35	0,02779182			
	-Chlorinated Impurity (C25H26N8O4)	0,3	0,00351364			
	lorinated Impurity (C25H24N8O4Cl2)	0,6	0,00618008			
	Enantiomer (C25H25N8O4Cl)	0,5	0,00548042			

Part c: Differences of recalculated centesimal composition and microanalysis experimental values

Differences (Cent. Comp. - Exp 1)	
C	0,298
H	-0,051
N	0,273
S	
O	

Fig. 8 User interface screenshots (parts a, b, c)––example of enantiomeric impurity insertion in the impurity profile, for the compound $C_{25}H_{25}N_8O_4Cl$, HCl

Another interesting point is the capability to integrate impurity profile contents (with known structures), contents of residual solvents, and others, which are most of the time included in control monographs. Because this aspect was not presented in these examples, related substance contents (meaning contents of substances assimilated to the structure of the substance analyzed) can also be added in the calculations with Mat-CHN. In this last case, the impurity will have the same raw formula than the one of the substance analyzed. This is a real new performance delivered by this software, and it could be considered as a part of the answer to the expert panel confronted with the case of the Povidone monograph.

In a choice between use of Elemental Analysis combined with Mat-CHN and the assay against a reference standard or the absolute assay by a potentiometric method (frequently difficult to develop), the Elemental Analysis associated to Mat-CHN method has substantial benefits in the reduction sample consumption and the time spent.

The last interesting point is related to the results provided by Mat-CHN. As this was demonstrated along this work, Mat-CHN uses only experimental results (except the raw formula of the pure substance) and makes calculations with these experimental data. Because no recursive process and no other value are inserted by the user (to manage these calculations), the final results delivered by Mat-CHN should be considered also as an experimental one. Consequently, the titration value issued from Mat-CHN must be considered at the same level than the one obtained by a direct potentiometric titration.

Conclusions

The Mat-CHN principles of calculations have been presented in this work in the theoretical aspect and through practical examples in the contexts of the discovery and the development activities for new molecular entities.

This software has the advantage of delivering a unique and reliable solution to the system of equations, resulting in individual tests dedicated to determine mineral and organic contents (expressed in percent and in mole numbers) for substances associated to the main compound analyzed.

Mat-CHN software can support analytical determinations obtained in the range of ppm and ppb units, mixed with other results delivered at the percent scale.

Without recursive process or convergence test, in a very quick time (around 1 s) and implemented on a classical computer, the scientist is directly oriented to the unique redefined structure, to make comparisons of experimental and recalculated Elemental Analysis results.

Titration of the substance analyzed can be easily delivered, based on the ratio of the molecular weight of the pure compound to the one of the refined molecule in the context of discovery and based on the nitrogen content ratio in the case of development activities, as recommended by the regulations. In these two cases, Mat-CHN software plays a central role to deliver the right figures to make the appropriate calculations. And because no recursive process and no other data added than the ones experimentally obtained, the titration calculations delivered by Mat-CHN should be considered also, as experimental values.

If the interest in Elemental Analysis was sometimes limited by the missing capabilities to complete the definition of the structure of the analyzed compound, Mat-CHN is probably one of the first software to support and to reinforce this analytical technique. It has the main advantage of delivering an absolute assay of each of the principal elements (carbon, hydrogen, nitrogen) or to allow the titration of the substance analyzed, by using very low quantities of substance, compared to current methods based on comparative results and management of reference standards.

In the case of the development of NMEs, the Elemental Analysis associated with Mat-CHN presents a real advantage for scientists, allowing routine absolute titration of the compound analyzed, without any development of method, also taking into account impurity profile, residual solvents, mineral and metal traces, and other contents of defined substances. In the case of preferred titrations against reference standards, for the same reasons, the absolute titrations of these reference substances will be relevant of the Elemental Analysis associated to Mat-CHN.

Not developed in this work but easy to understand, it is also important to underline here that Mat-CHN constitutes a simulator tool to anticipate contents of by-products that can appear during synthesis studies or drug product elaborations, very helpful for the chemists, pharmacists, and analysts in discovery and development steps.

Abbreviations

%C, %H, %N, %O, %S: Percentages of carbon, hydrogen, nitrogen, oxygen and sulfur elements; ADCs: Antibody-drug conjugates; C, H, N, O, S: Carbon, hydrogen, nitrogen, oxygen and sulfur elements; CHN: Carbon, hydrogen, nitrogen. Also used for microanalysis determinations; EA: Elemental Analysis; EP: European Pharmacopeia; Exper.: Experimental; $g.mol^{-1}$: Gram per mole; GC: Gaz chromatography analysis; Ghz: Giga Hertz; HPLC: High-performance liquid chromatography analysis; IUPAC: International Union of Pure and Applied Chemistry; Mat-CHN: Name of the Software, based on mathematical (Matrix) microanalysis calculations; Mol, mol: Mole; Mw: Molecular weight; Nb: Number; NME: New molecular entity; NMR: Nuclear magnetic resonance; ppb: Parts per billion; ppm: Parts per million; Reclac.: Recalculated; USP: United States Pharmacopeia; *w/w*, %*w/w*: Weight to weight report, weight to weight percentage; **Δ** (Delta): Delta: difference between values

Acknowledgements

This work was initiated for editing certificates of analysis in our R&D environment. Consequently, without help of chemists and analysts involved respectively in batch syntheses and analytical results, this work would have never been put in place. I thank all these teams for all the different and complicated cases, they submitted to me, to validate the robustness of this software.

I will never forget the involvement of the analytical team of Nathalie Hasel, in charge of the analytical support to the medicinal chemistry, for structural analysis purposes with Armèle de Guinée, Stéphanie Lessertois, and Valérie Bègue also in charge of the elemental analysis activities and for purity determinations with Christine Arnal, Daisy Leguevaques, and Fabien Millac also in charge of ion analysis activities.

I want also thank particularly Anthony Padfield, who joined recently our team and accepted the role of external reviewer of this article.

Availability of data and materials

Data and software are stored in the computer of the company of the author (see "Print Name" section).

Author's contributions

LV is the only author of the manuscript and inventor of the software. Consequently, he contributed alone to this work. The author read and approved the final manuscript

Competing interests

The author declares that there are no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 1 March 2018 Accepted: 5 June 2018
Published online: 21 June 2018

References

Block LH. The USP Excipients Stakeholder Forum, Excipient Monograph Modernization, June 7 2013. 2013. https://www.usp.org/sites/default/files/usp/document/get-involved/stakeholder-forums/5-excipient-modernization-2013-06-07.pdf. Accessed 28 Jun 2016.

Crowther JB. Validation of pharmaceutical test methods. Sep Sci Technol. 2001;3:415–43.

dos Santos ZM, Croni ALPF, Pereira MR, da Silva DR, Fonseca JLC. Determination of deacetylation degree of chitosan: a comparison between conductometric titration and CHN elemental analysis. Carbohydr Res. 2009;344:2591–5.

Fadeeva VP, Tikhova VD, Nikulicheva ON. Elemental analysis of organic compounds with the use of automated CHNS analyzers. J Anal Chem. 2008; 63(11):1094–106.

Gervasio Pereira M, Espindula A Jr, Souza Valladares G, dos Anjos LHC, de Melo Benites V, Schultz N. Comparison of total nitrogen methods applied to histosols and soil horizons with high organic matter content. Commun Soil Sci Plant Anal. 2006;37:939–43.

Hemming PE. 2015. https://www.labmate-online.com/article/laboratory-products/3/paul_e.hemming/implications_of_instrument_design_on_chn_microanalysis_data_quality_/1925. Accessed 28 Jun 2016.

IUPAC Commission on Isotopic Abundances and Atomic Weights. Atomic weights of the elements. 2015. http://www.sbcs.qmul.ac.uk/iupac/AtWt/. Accessed 28 Jun 2016.

Krotz L, Francesco L, Giazzi G. Performance of the Thermo Scientific FLASH 2000 elemental analyzer at low, medium and high concentration. In: Thermo fisher scientific technical note, N°42234; 2015. p. 1–4.

Oftedal OT, Eisert R, Barrell GK. Comparison of analytical and predictive methods for water, protein, fat, sugar and gross energy in marine mammal milk. J Dairy Sci. 2014;97:4713–32.

Ouderkirk L, Seo P. Monograph modernization history––2010-11-16 letter from FDA task group to USP. 2010. http://www.usp.org/sites/default/files/usp/document/get-involved/monograph-modernization/2010-11-16-letter-from-fda-task-group.pdf. Accessed 28 Jun 2016.

Rouessac F, Rouessac A. Chemical Analysis: second edition, J. Wiley & sons, Ltd, 2007, chapter 18 - Elemental analysis, p441-451, [ISBN:978-0-470-85902-5 (HB)].

Submit your manuscript to a SpringerOpen journal and benefit from:

- ► Convenient online submission
- ► Rigorous peer review
- ► Open access: articles freely available online
- ► High visibility within the field
- ► Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com