

Identification	CGRPFSTR150020	Ind.	A
Nom du FDR :	HAVL Argile		
Diffusion (*):	Communicable		

Émetteur externe : AMPHOS 21
Noms des signataires : M. Grivé, I. Campos, E. Colàs, J. Perrone
Référence : Project TC1 Date d'émission: July 2015
N° de commande ou de l'accord de coopération : 20064522

THERMOCHIMIE-TDB GUIDELINES

1. Data selection and uncertainties
2. Data integration and consistency
3. Validation and verification
4. Data traceability
5. Bibliographic references recording process
6. Data quality

This work has been performed by Amphos 21, and co-funded by Andra and NDA/RWM within their partnership for the ThermoChimie-TDB development

***CE DOCUMENT EST LA PROPRIÉTÉ DE L'ANDRA ET NE PEUT ÊTRE REPRODUIT OU COMMUNIQUÉ QUE SUIVANT LA MENTION INDIQUÉE CI-DESSUS EN DIFFUSION**

Communicable : document pouvant être diffusé à tout public
Limitée : document pouvant être diffusé à tout le personnel Andra ainsi qu'au public averti
Andra : document pouvant être diffusé au seul personnel Andra
Confidentielle : document dont la diffusion est interdite à d'autres destinataires que ceux indiqués sur le document

Document répondant au cahier des charges mais dont le contenu traduit le point de vue du prestataire et n'engage que ce dernier.

Observation éventuelle du prescripteur :

Nom/visa Andra pour diffusion

Eric GIFFAUT

Date : 15/10/2015

Introduction to ThermoChimie guidelines

Project ANDRA TC1

**Mireia Grivé, Elisenda Colàs,
Isaac Campos, Jane Perrone**

July 2015

Written by:	Revised by:	Validated by:
 Elisenda Colàs Isaac Campos	 Jane Perrone	 Mireia Grivé

Index

1. THE THERMOCHIMIE DATABASE	2
1.1 PRINCIPLES OF DEVELOPMENT	2
1.2 CONTENTS OF THE DATABASE	4
1.2.1 <i>Thermodynamic data</i>	5
1.2.2 <i>Kinetic data</i>	6
1.2.3 <i>Database formats</i>	6
2. THE THERMOCHIMIE GUIDELINES	7
3. BIBLIOGRAPHY	8

1. The ThermoChimie database

ThermoChimie is a thermodynamic database created and developed by Andra, the French National Radioactive Waste Management Agency, since 1996. From September 2014 onwards, Radioactive Waste Management Limited (RWM, UK) shares the financing support of the development of ThermoChimie database together with Andra.

ThermoChimie is addressed to geochemical modelling and performance assessment tasks. It was born from the need of using a reliable and comprehensive chemical thermodynamic database to feed conceptual and numerical models used in Performance Assessments (PA) of Nuclear Waste Repositories. These models are mainly used to assess the geochemical evolution of a repository both in terms of performance of the engineering barrier system and of the migration/retention behaviour of radionuclides. Data acquisition is supported by an experimental programme, in addition to active literature survey and the development of predictive models. Several expert groups and scientific organizations contribute to this continuous update process.

The ThermoChimie database is available on <http://www.thermochimie-tdb.com/>. A general description of the database can be found in Giffaut et al. (2014). Specific information for radionuclides, minerals and kinetic data can be found in Grivé et al. (2015), Blanc et al. (2015) and Marty et al. (2015) respectively.

1.1 Principles of development

The ThermoChimie database has been developed to meet the six following objectives, in compliance with Performance Assessment requirements:

1. Determination of radioelement and chemotoxic element aqueous speciation and solubility.

-
2. Study of geochemical evolution of both the near- and far-fields of the repository, covering the stability of clay minerals, bentonite clays, and aluminosilicate systems.
 3. Assessment of the process of cement degradation (considering a broad composition range with respect to formula) to account for the stability of cementitious phases.
 4. Assessment of the evolution of metallic/steel components present in the containers, liners, reinforced waste packages, etc. with special emphasis on iron/steel corrosion processes and secondary phases formed.
 5. Assessment of the role of simple organic ligands and their impact on the mobilization of radionuclides, including: a) organic ligands present in wastes; b) organic ligands formed by degradation of organic materials present among the wastes or in the engineered barriers; c) those organic ligands derived from the degradation of Natural Organic Matter (NOM) present in the Callovo-Oxfordian clay stone.
 6. Tools and applications of the thermochemical database under different disposal scenarios.

ThermoChimie is built according to the following main requirements:

- **Consistency:** The thermodynamic functions included in ThermoChimie must be consistent. Furthermore, the different thermodynamic functions for the whole chemical system are also consistent among them. Further details about consistency can be found in Guideline 2: Data integration and consistency.
- **Exhaustivity:** The database must contain all data needed to represent the expected behaviour of the systems studied, under the ranges of conditions likely to occur in the different disposal situations. This exhaustivity requirement is only applied to the systems and disposal situations for which the database has been developed, as described above. The contents of the database are described in section 1.2 of the present document; details about data selection procedures are provided in Guideline 1: Data selection.
- **Traceability:** Selected data must be traceable to the original source. The calculations used to obtain the data must also be traceable. Further details can be found in Guideline 4: Data traceability.

1.2.1 Thermodynamic data

For each element, a basic component is defined in ThermoChimie. The basic component is usually a free cation (e.g., Sr^{2+}), an oxycation (e.g., UO_2^{2+}) or an oxyanion (e.g., SeO_4^{2-}) of the element of interest.

Data included in ThermoChimie for **basic components** are:

- $\Delta_f \mathbf{G}_m^0$ ($\text{KJ}\cdot\text{mol}^{-1}$) at 25 °C: Standard molar Gibbs energy of formation.
- $\Delta_f \mathbf{H}_m^0$ ($\text{KJ}\cdot\text{mol}^{-1}$) at 25 °C: Standard molar enthalpy of formation
- \mathbf{S}_m^0 ($\text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$) at 25 °C: Standard molar entropy.
- **Ion interaction coefficients $\epsilon(j, k)$** ($\text{Kg}\cdot\text{mol}^{-1}$) for the interaction of the species with Na^+ , Cl^- , ClO_4^- and/or NO_3^- , when available, at 25 °C.

Basic components (together with e^- and H^+) constitute the building blocks for the construction of the formation reactions of the remaining species and solid phases in the database.

In ThermoChimie, aqueous species are defined by chemical reactions of the basic components. The thermodynamic data included in the database for **aqueous species** are:

- **$\text{Log}_{10}\mathbf{K}^0$** at 25 °C: Equilibrium constant of the reaction.
- $\Delta_r \mathbf{G}_m^0$ ($\text{KJ}\cdot\text{mol}^{-1}$) at 25 °C: Molar Gibbs energy of reaction.
- $\Delta_r \mathbf{H}_m^0$ ($\text{KJ}\cdot\text{mol}^{-1}$) at 25 °C: Molar enthalpy of reaction.
- $\Delta_r \mathbf{S}_m^0$ ($\text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$) at 25 °C: Molar entropy of reaction.
- $\Delta_f \mathbf{G}_m^0$ ($\text{KJ}\cdot\text{mol}^{-1}$) at 25 °C: Standard molar Gibbs energy of formation.
- $\Delta_f \mathbf{H}_m^0$ ($\text{KJ}\cdot\text{mol}^{-1}$) at 25 °C: Enthalpy of formation.
- \mathbf{S}_m^0 ($\text{J}\cdot\text{K}^{-1}\cdot\text{mol}^{-1}$) at 25 °C: Standard molar entropy.
- **Ion interaction coefficients $\epsilon(j, k)$** ($\text{Kg}\cdot\text{mol}^{-1}$) for the interaction of the species with Na^+ , Cl^- , ClO_4^- and/or NO_3^- , when available, at 25 °C.

For **solid compounds**, ThermoChimie includes values for the same variables than for aqueous species. Other parameters are also provided, when possible, for some of them (especially solid clays and cement mineral phases):

-
- $C_{p,m}^0$ ($J \cdot K^{-1} \cdot mol^{-1}$): Standard molar heat capacity at constant pressure.
 - V_m^0 ($cm^3 \cdot mol^{-1}$): Molar volume.

All the values included in ThermoChimie are accompanied by a bibliographic reference of the source from where the datum has been adapted and, when possible, given with their associated uncertainties.

It is foreseen that in the future, the values will also include qualifiers; those qualifiers are specific text labels that will provide information about the quality and overall reliability of each datum (see Guideline 6: Data quality).

1.2.2 Kinetic data

A compilation of kinetic parameters in the form of individual sheets (one for each mineral phase) is proposed in ThermoChimie. The selection is currently limited to some of the mineral phases present in clay rich rocks and cements. The data selected are written in agreement with the formalism implemented in the geochemical code Toughreact (Xu et al., 2011) considering the Transition State Theory (TST), but can be easily applied to other codes such as PhreeqC (Parkhurst and Appelo, 2013) or Crunch (Steeffel, 2009). The exact variables compiled in those data sheets depend on the mineral, but in general the parameters k_n (the kinetic constant), θ and η (parameters describing the dependence between the reaction velocity and the saturation index) are provided.

1.2.3 Database formats

The individual thermodynamic and kinetic data for each species included in ThermoChimie (as described in sections above) can be reached in the web page <http://www.thermochimie-tdb.com/>.

Furthermore, in order to fulfil its usability requirements, the ThermoChimie database has been extracted into data formats compatible with the input needs of several numerical codes commonly used for geochemical applications: PhreeqC (Parkhurst and Appelo, 2013), Crunch (Steeffel, 2009), Toughreact (Xu et al., 2011), Chess (Van der Lee and De Windt, 2002).

These extraction capabilities are being extended to other codes as, for example, the Geochemist's Workbench code (Bethke and Yeakel, 2015).

2. The ThermoChimie guidelines

The main requirements of ThermoChimie (consistency, exhaustivity, traceability and usability) have led to the development of a significantly large database, which is being continuously updated. For this purpose, several groups of experts from different scientific disciplines participate in the development of ThermoChimie.

In order to guarantee that these requirements, which constitute the quality insurance of the database, are fulfilled, a number of procedures that must be followed by the experts have to be clearly defined. Therefore, a series of guidelines, which contain the necessary information, have been developed:

- **Guideline 1: Data selection and uncertainties.** This document describes how the data selection process must be carried out. This guideline guarantees the consistency in the selection process for the different groups of experts involved in the development of ThermoChimie, while taking into account the different particularities of each dataset.
- **Guideline 2: Data integration and consistency.** This guideline describes the different processes that must be fulfilled before the inclusion, elimination or modification of any value in ThermoChimie. Those processes guarantee the consistency of the database.
- **Guideline 3: Validation exercises.** The document defines the validation tests to be applied to the data included in ThermoChimie. These tests must take into account the diversity of systems in geochemical modelling for which the database is applied.
- **Guideline 4: Data traceability.** This guideline describes how to produce the documents that keep track of the different changes, updates and releases of the database.
- **Guideline 5: Bibliographic reference recording process:** This document defines how the bibliographic references must be cited (style) and which

reference manager system is used in order to guarantee consistency in the bibliographic citation for all ThermoChimie developers.

- **Guideline 6: Data quality:** This document defines how different qualifiers will be assigned to the data included in future ThermoChimie versions.

These guidelines are available for all ThermoChimie developers and also for all ThermoChimie users, in order to ensure that the ThermoChimie requirements are fulfilled and that the database development and usage is carried out in the best possible way.

3. Bibliography

Bethke, C. M., Yeakel, S. 2015. The Geochemist's Workbench: GWB Essentials Guide. Aqueous Solutions, LLC

Blanc, P., Vieillard, P., Gailhanou, H., Gaboreau, S., Marty, N., Claret, F., Madé, B., & Giffaut, E. 2015. ThermoChimie database developments in the framework of cement/clay interactions. *Applied Geochemistry* 55, 95-107.

Giffaut, E., Grivé, M., Blanc, P., Vieillard, P., Colàs, E., Gailhanou, H., Gaboreau, S., Marty, N., Madé, B., & Duro, L. 2014. Andra thermodynamic database for performance assessment: ThermoChimie. *Applied Geochemistry*, 49, 225-236.

Grivé, M., Duro, L., Colàs, E., & Giffaut, E. 2015. Thermodynamic data selection applied to radionuclides and chemotoxic elements: An overview of the ThermoChimie-TDB. *Applied Geochemistry* 55, 85-94.

Marty, N. C., Claret, F., Lassin, A., Tremosa, J., Blanc, P., Madé, B., Giffaut, E., Cochevin, B. & Tournassat, C. 2015. A database of dissolution and precipitation rates for clay-rocks minerals. *Applied Geochemistry* 55, 108–118.

Parkhurst, D.L., Appelo, C.A.J. 2013. *Description of input and Examples for PHREEQC Version 3 — A Computer Program for Speciation, Batch-reaction, Onedimensional Transport, and Inverse Geochemical Calculations*. U.S. Geological Survey Techniques and Methods, book 6, Chap. A43.

Steefel C.I., 2009. Software for Modeling Multicomponent Reactive Flow and Transport. Rapport technique Lawrence Berkeley National Laboratory.

Van der Lee, J., De Windt, L. 2002. CHESS Tutorial and Cookbook. Updated for Version 3.0. Users Manual Nr. LHM/RD/02/13, Ecole des Mines de Paris, Fontainebleau, France.

Xu, T., Spycher, N., Sonnenthal, E., Zhang, G., Zheng, L., Pruess, K. 2011. TOUGHREACT Version 2.0: a simulator for subsurface reactive transport under non-isothermal multiphase flow conditions. *Comput. Geosci.* 37 (6), 763–774.

ThermoChimie guidelines 1. Data selection and uncertainties

Project ANDRA TC1

**Mireia Grivé, Elisenda Colàs,
Isaac Campos, Jane Perrone**

July 2015

Written by:	Revised by:	Validated by:
 Elisenda Colàs Isaac Campos	 Jane Perrone	 Mireia Grivé

Index

1. INTRODUCTION	2
2. SELECTION PROCESS	2
2.1 PRIMARY MASTER SPECIES AND REFERENCE STATES OF THE ELEMENTS	3
2.2 RADIOELEMENTS AND CHEMOTOXIC ELEMENTS	3
2.3 LOW MOLECULAR WEIGHT ORGANIC COMPLEXES	8
2.4 CLAY MINERALS	10
2.5 CEMENT MINERALS	13
2.6 ZEOLITES	15
2.7 KINETIC DATA	16
3. UNCERTAINTIES	18
3.1 UNCERTAINTIES IN DATA FROM DEDICATED EXPERIMENTAL PROGRAMS	19
3.2 UNCERTAINTIES IN DATA FROM OPEN LITERATURE	21
3.3 UNCERTAINTIES IN ESTIMATED DATA	21
3.4 OTHER CASES	22
4. BIBLIOGRAPHY	22

1. Introduction

The ThermoChimie database contains different datasets including thermodynamic data for major elements, radioactive and toxic elements, organic ligands, cement hydrates, clay minerals and zeolites, and kinetic data for some minerals present in clay rich rocks and cements. Each of these groups has its own chemical characteristics, and thus the data selection process for each group has its own particularities and requirements. Nevertheless, all these data sets are interdependent when applied to Performance Assessment exercises. Thus, all those data sets have to be built in a consistent way, in order to be applied in the different systems for which ThermoChimie has been designed¹.

The purpose of the present guideline is to provide the indications necessary for data selection processes, taking into account the particularities of each data group but also the need to satisfy the requirements of ThermoChimie, including consistency and applicability to Performance Assessment exercises.

2. Selection process

The data selection process includes an exhaustive work of literature research, comparison and evaluation of different data sources:

- Open scientific literature
- Specific experimental programmes
- Previous thermodynamic data compilations
- Estimations

During the selection process, data are analysed and tested in front of independent values. The selection process takes also into account the thermodynamic data reported in previous compilations. The iterative behaviour of the selection process is always present because of the continuously appearing data in the open scientific

¹ See the Introduction to ThermoChimie guidelines for an exhaustive description of the systems for which ThermoChimie has been designed for.

literature, which implies the revision and final update of previously selected data or the inclusion of new thermodynamic data.

2.1 Primary master species and reference states of the elements

When possible, ThermoChimie relies on **CODATA recommendations** (Cox et al. 1989) for the thermodynamic data of the **primary master species**² and the **reference states of the elements** (Giffaut et al. 2014). However, in some specific cases (Al and Si) the values selected for the basic species are not exactly the same as the ones recommended by CODATA (see Blanc et al. 2006). Those changes require a solid discussion and supplementary justifications, as well as an exhaustive analysis of the practical consequences that derive from this decision, i.e., exhaustive consistency checking of the whole database (see Guideline 2: Data integration and consistency).

2.2 Radioelements and chemotoxic elements

Each individual datum for the radionuclides and chemotoxic elements selected in ThermoChimie is the result of an exhaustive work of literature research, data intercomparison, estimation of thermodynamic data and, in some specific cases, development of a focused experimental programme to fill in important data gaps (Grivé et al. 2015).

The results of the **NEA-TDB project** are particularly considered for the selection of thermodynamic data of some radionuclides, given their high quality. Nevertheless, in some cases the standard of the selection published in the Chemical Thermodynamics series excludes data that are needed to explain a system of interest. One of the guiding principles of ThermoChimie is that it must be sufficiently complete to allow for calculation of the behaviour of the complex systems of interest for radioactive waste management calculations. This means that it must include all those species known to exist under the conditions of interest; then, in some cases, data not selected in other reviews (e.g. the NEA-TDB project) are included in ThermoChimie to ensure the

completeness of the assessment. Furthermore, the NEA data does not cover all the elements and conditions of interest of ThermoChimie application; then, the selection has to go beyond the NEA-TDB project limits.

The data included in ThermoChimie are selected according to the procedure described below.

1. When possible, **aqueous stability constants and solubility equilibria ($\log_{10}K^0$)** are selected as main data. Values of $\log_{10}K^0$ and $\Delta_rG_m^0$ are thus the first parameters selected. The selection process for $\log_{10}K^0$ is summarized in Figure 1.

² Primary master species are the minimal set of aqueous species required to formulate product species; thus, they are the basic building blocks in order to formulate formation reactions.

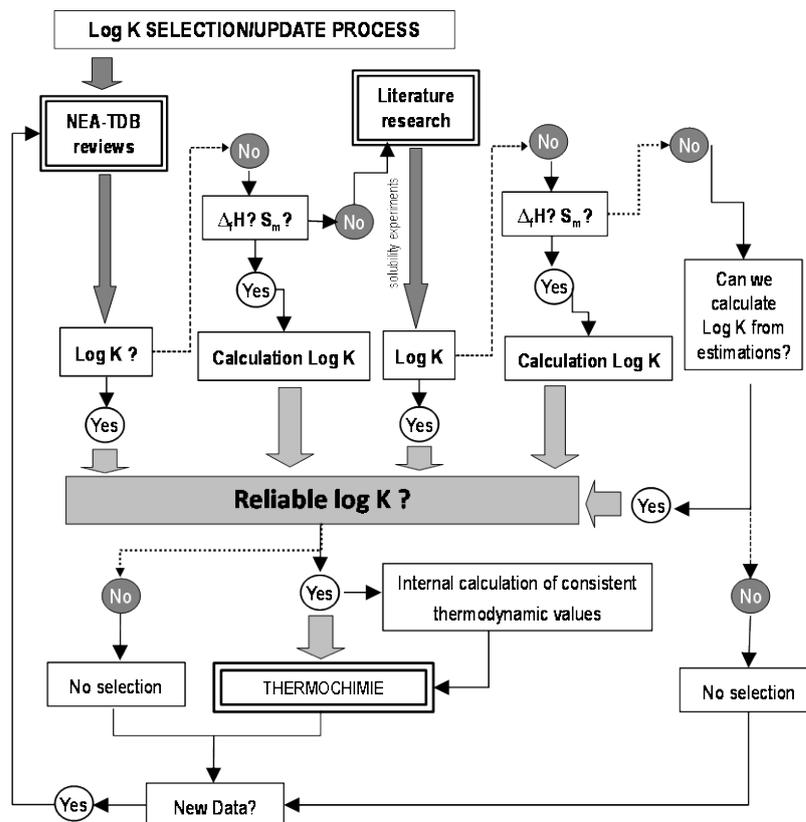


Figure 1. Summary of the selection and update process for log K values of radionuclides and chemotoxic elements in ThermoChimie (Grivé et al. 2015).

These data allow the calculation of the Gibbs energy of formation ($\Delta_f G_m^0$) of the species as far as the Gibbs energy of formation of the basic components is available.

Solubility equilibrium constants can be obtained by two different ways.

- Calculation of the solubility product ($\log K_s^0$) from solubility experiments: In those cases, the aqueous concentrations of solutes in equilibrium with the solid phase are used to obtain the equilibrium constant. This implies the assumption of an aqueous speciation scheme and associated stabilities of the aqueous complexes of the element of interest.
- Calculation of $\log K_s^0$ from formation magnitudes obtained by calorimetric measurements, from which values of enthalpy and/or entropy can be measured.

Since ThermoChimie is used in solubility and performance assessment of deep underground repositories, preference is given in many cases to the value obtained from solubility experiments. During the dissolution process, the mineral surface can suffer changes such as hydration, precipitation of amorphous minerals, ion exchange and adsorption, that may affect the value of the equilibrium constants. The solubility constants obtained from classical dissolution experiments includes all these effects, whereas $\Delta_r G_m^0$ values derived from calorimetric data refer to a perfectly crystalline surface. The surface of a crystalline solid may differ from the surface of the most amorphous solid phases, which are the ones in contact with the groundwater and, therefore, controlling the concentrations of the dissolved species in the solution.

Redox reactions may be an exception to this general rule. In this case, different methodologies (e.g., potentiometry) could be considered as appropriate in order to provide accurate thermodynamic data.

2. Secondly, **enthalpies or entropies** are selected. If available, experimental data are preferred, however, when reliable data are not available, estimations are performed (see point 5).

3. When values have been selected for two of these three variables ($\Delta_r G_m^0$, $\Delta_r H_m^0$ and S_m^0), the rest of the data are internally calculated using the following equations (Eq. 1, Eq. 2):

$$\Delta_r G_m^0 = \Delta_r H_m^0 - T \left(S_m^0 + \left(\frac{Z}{2} \right) x S_{m,H2(g)}^0 - \sum_i \frac{\nu_i}{\nu_{i,ref}} S_{m,i}^0 \right) \quad \text{Eq. 1}$$

$$\Delta_r G_m^0 = \Delta_r H_m^0 - T \Delta_r S_m^0 \quad \text{Eq. 2}$$

4. Finally, ThermoChimie primes *SIT (Specific ion Interaction Theory)* for **activity corrections** to the derivation of the stability constants to the standard state.

The SIT approach takes into account short-range non-electrostatic interactions by adding terms to the Debye-Hückel expression (Eq. 3):

$$\log(\gamma_i) = -z_i^2 \left(\frac{A\sqrt{I_m}}{1 + Ba_i\sqrt{I_m}} \right) + \sum_k \varepsilon(i,k,I_m)m_k \quad \text{Eq. 3}$$

If individual, reliable and consistent interaction coefficients are available, it is possible to calculate $\log_{10}K^0$ values accordingly. If the individual interaction coefficients are not reported, it is possible to calculate $\log_{10}K^0$ values using stability constants at different ionic strengths, by interpolating $\log_{10}K^0$ and $\Delta\varepsilon$ from a linear regression.

RELATIONSHIP OF LogK° AND LogK FOLLOWING SIT

$$\log K^0 = \log K - \Delta z^2 \times D + \Delta\varepsilon \times I$$

$$D = \frac{-0.5\sqrt{I}}{1 + 1.5\sqrt{I}}$$

$$\Delta z^2 = \sum_{\text{products}} z_i^2 - \sum_{\text{reactants}} z_i^2$$

$$\Delta\varepsilon = \sum_{\text{products}} \varepsilon_i - \sum_{\text{reactants}} \varepsilon_i$$

z: ionic charge; I: ionic strength; ε: interaction coefficients between ions

Figure 2. Relationship among SIT equation, $\log K^0$ and $\Delta\varepsilon$ values.

5. In some cases, due to the lack of data or to the inaccuracy of the existing data, **estimations** are needed.

Three main types of gaps are commonly identified in thermodynamic databases:

- gaps related to stability constants;
- gaps related to enthalpy and/or entropy values;
- and gaps related to ionic strength correction interaction coefficients (SIT interaction coefficients).

Different estimation approaches can be used provide values for these different kinds of thermodynamic data.

The choice of the estimation procedure depends on many factors (type of data, availability of the original values, studied element,...) and can be of different forms (analogies with other elements and complexes, correlations considering the charge z

and the ionic radii r of different elements, empirical algorithms,...). The validity and accuracy of the estimation must be individually evaluated, and consistency between estimated data with the rest of the database must be maintained (see Guideline 2: Data integration and consistency).

6. Verification is carried out by comparison of selected values with additional independent data (when available), or with calculations including correlations or analogies, as LFER (see Guideline 2: Data integration and consistency).

2.3 Low molecular weight organic complexes

The selection for organic ligands is focused on those low molecular weight organic species likely to play a role under disposal conditions. This includes:

- bitumen degradation products (e.g., oxalate, acetate);
- ligands that may be used as a surrogate of the effect of cement additives or their degradation products (e.g., gluconate);
- chemicals used for decontamination and cleaning processes (e.g., EDTA, NTA);
- degradation and leaching products from plastics, filters or resins (e.g., adipate, phthalate);
- cellulose degradation products (e.g., isosaccharinate);
- low molecular weight organic species identified in natural clay systems in the presence of kerogen (e.g., malonic and succinic acids).

The methodology followed for the selection of the relevant thermodynamic data is described below.

1. Identification of the available thermodynamic data for the organic ligands of interest, including open literature data and previous compilations.

The results of the NEA-TDB project are also taken into account in the selection. However, the NEA review for organic ligands (Hummel et al. 2005) contains only a limited number of data for citrate, oxalate, EDTA and isosaccharinate with a selected number of major ions and radionuclides, and thus provides a limited amount of

information. Data selection has to go beyond the NEA-TDB scope in order to cover all the fields of application of ThermoChimie.

2. Selection of reliable **experimental** data (*stability constants, $\log_{10}K$*) available for each system and extrapolation to $I = 0$ using the specific-ion interaction theory (SIT). Data acquired in the context of the experimental programme developed within the ThermoChimie project are specially taken into account.

In order to maintain the consistency of the dataset, the selection is carried out in three successive steps:

- 2.a) selection of hydrolysis data of the organic complexes;
- 2.b) selection of organic complexation data with the alkaline-earth elements Ca and Mg, consistently with data in step 2.a;
- 2.c) selection of organic complexation data with the radionuclides of interest included in ThermoChimie, consistently with data in steps 2.a and 2.b.

3. Available experimental data for organic ligands of interest are limited. When possible, relevant data gaps for stability constants are covered with **estimations** based on analogies or correlations.

4. Finally, **entropy or enthalpy** data are selected in order to evaluate the influence of temperature. However, it should be noticed that the information on those parameters is quasi inexistent among the open literature data, and thus the amount of information included in ThermoChimie on this subject is limited. However, this lack of knowledge is not considered as relevant considering the focus on non exothermic organic wastes.

5. Verification is carried out:

- by comparison of selected values with additional independent data available in the literature, e.g., estimations or values reported in other thermodynamic data compilations.
- with calculations including correlations or analogies.

These verification exercises allow to increase the confidence in the selection performed and to confirm that a reasonable uncertainty has been assigned to the selected thermodynamic data.

2.4 Clay minerals

The clay mineral thermodynamic data are obtained using three different approaches (Blanc et al. 2015):

- from a critical selection of the literature data;
- from dedicated experimental measurements (e.g. Blanc et al. 2014);
- using estimated values produced by predictive models. Those predictive models play a key role in thermodynamic data selection for clays, in order to estimate missing values, verify the consistency of data selection and provide a guidance for the critical data selection and for dedicated experiments.

1. When possible, $\log_{10}K(T)$ functions are calculated by using **calorimetric data**, either obtained from literature or obtained in specific experiments dedicated to fill in data gaps. Calorimetry is especially well suited for this purpose because the results do not depend on the attainment of equilibrium, which is difficult to prove, especially in low-temperature solution experiments (Blanc et al. 2013). Fitting $\log_{10}K(T)$ functions, as well as averaging equilibrium constants, is avoided as far as possible, in order to increase data traceability.

2. However, literature data are limited to micas and chlorite groups. On the other hand, trying to acquire the thermodynamic datasets for all possible compositions of clay minerals is not realistic. To solve this problem, **predictive models**, able to predict the thermodynamic properties of clay minerals for the most common compositions, are developed.

The properties of formation for a hydrated phase are in many cases deduced from the properties of formation of the anhydrous phase, adding afterwards the properties of hydration. This implies avoiding to consider directly the properties of formation of the

hydrated phase and, rather, to split the studies into the anhydrous phase and the properties of hydration

2.a) Predictive models for the properties of formation of the anhydrous clay minerals. In the case of anhydrous clay minerals, formation enthalpy is calculated in most cases following the approach developed by Vieillard, (1994a and b). The principle of calculation for this method relies on the difference in electronegativity between two cations around a common oxygen atom. The thermodynamic functions are completed by using the polyhedral decomposition model. The combination of both models allows calculating an estimated Gibbs energy for anhydrous phyllosilicates.

2.b) Predictive models of the thermodynamic properties of the hydration of clay minerals. The predictive models for hydration properties are based on an improvement of the original model by Ransom and Helgeson (1994), as fully described in Vieillard et al. (2011). The thermodynamic model was initially developed considering an asymmetric subregular binary solid solution between a fully hydrated and an anhydrous smectite, by calculating the effective amount of hydration water as the difference between the so-called gravimetric water and the surface covering water.

The thermodynamic properties of the hydrated clay mineral end members are calculated from the properties of the anhydrous end members. Then, the hydration model from Vieillard et al. (2011) is used to calculate the number of H₂O moles, the enthalpies and entropies of hydration, for P/P₀ = 1 (water saturation). For the moment, the hydration model does not take into account the charge and its location in the TOT sheet.

3. The calculations are then processed in the following way:
- first, formation enthalpies, entropies, Cp(T) functions and volumes are calculated for different compositions of clay minerals;
 - then, formation enthalpies and entropies are combined in order to provide the Gibbs energy;

-
- finally, the equilibrium constant $\log_{10}K(T)$ is calculated as a function of temperature, using primary basic species included into the ThermoChimie database.

4. Verification is carried out by drawing activity diagrams involving the minerals of interest in the chemical subsystems of concern and comparing phase relationships are compared with the observations gathered from the literature. The selection of thermodynamic constants allows realising speciation/predominance diagrams in different chemical sub systems. Those diagrams allow verifying the predicted phase relations and the consistency of the data set. Furthermore, they may be helpful to identify data gaps and weak points in data selection. Thus, this is a key step in the process, in order to guarantee the consistency of the different approaches used in the development of the system.

When possible, $\log_{10}K$ values calculated from calorimetric data and the results of predictive models are also compared with results of solution equilibria, when available, for verification. Experiments using data from a literature selection or experimental data specifically obtained for this purpose are used to fulfil this objective (Blanc et al. 2015). In any case, and taking into account the specific pitfalls of these types of experiments (compositional variability, presence of impurities, difficulties in the equilibrium achievement...) the application of this verification procedure is limited.

A scheme summarizing the process of data selection for clay minerals is shown in Figure 3.

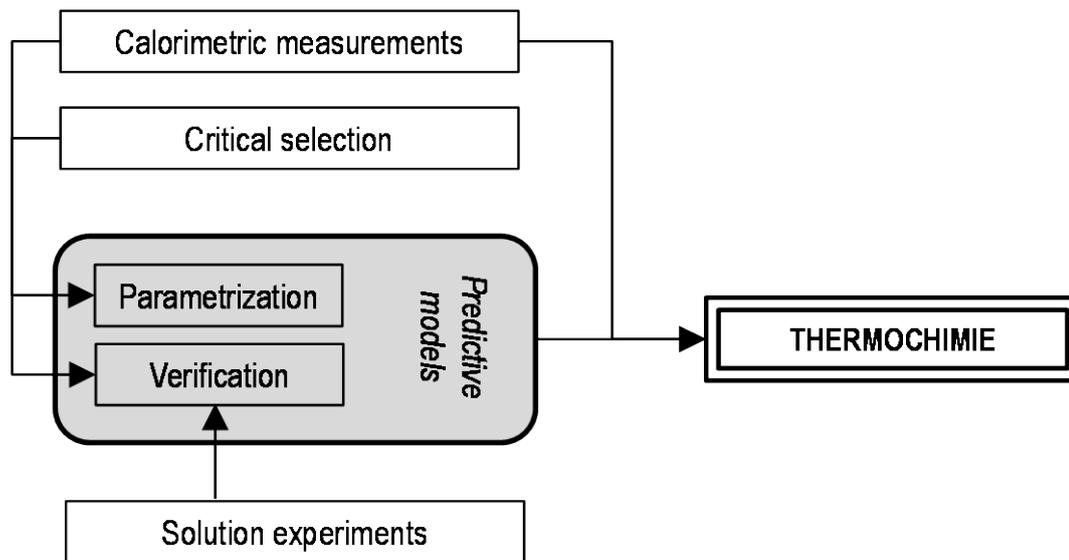


Figure 3. Scheme illustrating the process for the selection of thermodynamic datasets of clay minerals (from Blanc et al. 2015).

2.5 Cement minerals

Process of data selection for cement minerals is different for C-S-H phases and for the other cementitious phases, as described in Blanc et al. (2015).

1. For **C-S-H phases**, the selection of thermodynamic properties is organized the following way:
 - The experimental data available for the solubility or calorimetry of C–S–H phases at 25 °C and at $T > 25$ °C are compiled and analysed. When possible, properties are taken from the available experimental works.
 - Equilibrium constants for nanocrystalline C–S–H and their corresponding compositions are extracted from a set of experimental data.
 - For crystalline C–S–H, the properties datasets are completed with the results of the phase diagram refinement.
 - The lacking properties (formation enthalpy and heat capacity) are obtained by using a polyhedral decomposition method applied to the system $\text{SiO}_2\text{--CaO--H}_2\text{O}$ (Blanc et al. 2010a). This method is derived from the oxide summation method.

2. For **phases other than C-S-H**, the selection process is organized in the following way (Blanc et al. 2010b):

- For each phase, the equilibrium constant at 25 °C is selected from a single experimental work. For the equilibration experiments in aqueous media, only the aqueous solutions displaying a low ionic charge imbalance (smaller than 5% when possible) are retained. The duration and reversibility of the experiments, in addition to the electroneutrality of the solutions, are also considered in the selection and analysis of the experimental works.
- The formation enthalpy and the $C_p(T)$ function are taken from the literature or estimated. The estimation technique of Helgeson et al. (1978) is usually applied. This technique consists in considering a fictive solid–solid transformation reaction; in this reaction, the thermodynamic properties of all the solids involved (except one) are known. The unknown properties of this solid can then be estimated using as a basis the hypothesis that $\Delta_r C_p(298) = 0$ for the fictive solid-solid transformation reaction.
- Finally the $\text{Log}K(T)$ function is calculated by combination of the $\log_{10}K$ at 25 °C, enthalpy data and heat capacity data.

3. **Verification** is initially carried out by comparing the $\text{Log}_{10}K(T)$ function with the reaction quotient $\text{Log}_{10}Q$ from the literature, at different temperatures. An important step in the verification procedure is to draw speciation/predominance diagrams corresponding to defined chemical subsystems. Phase relationships are compared with the observations gathered from the literature. Those diagrams allow verifying the predicted phase relations and the consistency of the data set. Furthermore, they may be helpful to identify data gaps and weak points in data selection. Thus, as in the case of clays, this is a key step in the process, in order to guarantee the consistency of the different approaches used in the development of the system. Those verification tests, involving the C-S-H phases and the other cementitious phases, are done before the data are integrated into the database.

A summary of the processes are shown in Figure 4.

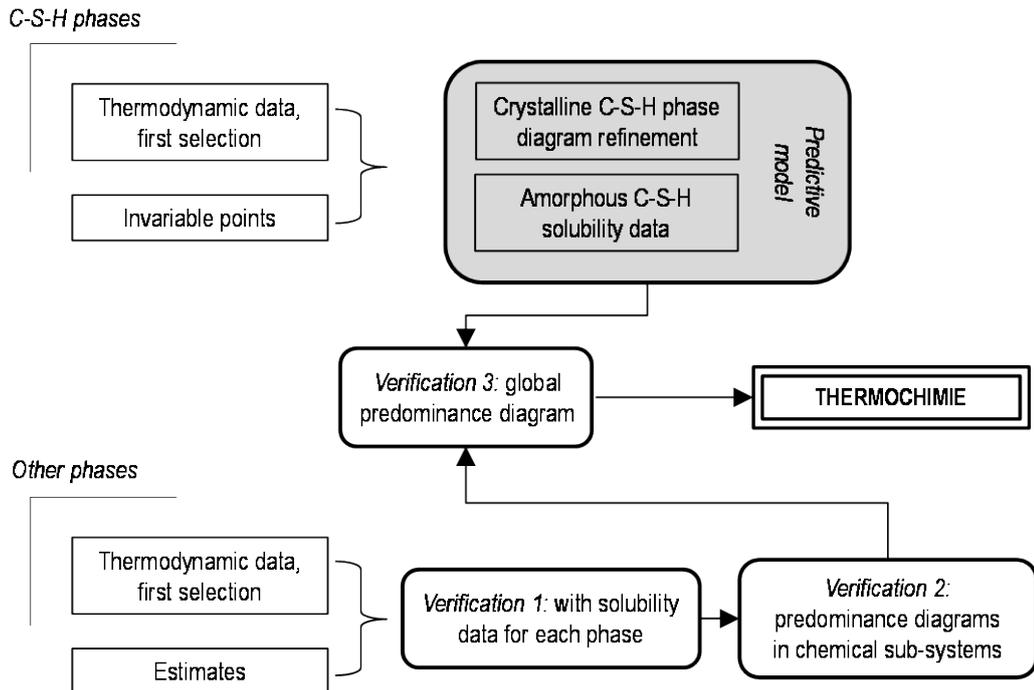


Figure 4. Scheme illustrating the process for the selection of thermodynamic datasets of cement minerals (from Blanc et al. 2015).

2.6 Zeolites

The selection for zeolites is limited to the phases likely to play a role during cement/clay interaction in deep disposal and formed at temperatures below 90 °C (i.e. within the range of application of ThermoChimie. Details on the selection can be found in Blanc et al. (2015).

1. When possible, the thermodynamic properties of zeolites are obtained from a **critical selection of literature data**. The selection of thermodynamic properties is focused on the collection of direct and traceable measurements. Fitting $\log_{10}K(T)$ functions, as well as averaging equilibrium constants, is avoided as far as possible, in order to increase data traceability. When possible, calorimetric data are preferred over solution equilibria; solution equilibria data are used mainly for comparison and verification purposes. If it is necessary to use solution equilibria results, those are accurately selected according to the electroneutrality condition, the duration of the

experiment, the analysis of the final solid products and the experimental protocol used in the original work.

2. The literature data selection is sometimes limited when trying to obtain a complete thermodynamic dataset for a unique zeolite; for example, enthalpy and entropy data reported in the literature could be measured on minerals displaying slightly different compositions, which limits the selection. In order to complete the thermodynamic datasets, when necessary, missing heat capacities are estimated using the **predictive model** developed by Vieillard (2010), based on a multiple linear regression method. The main interest of the estimation method resides in the use of zeolitic minerals as a base for the calibration of the model, as the method exclusively targets zeolites.

3. Finally, data selection is **verified** by the following approaches:

- Additional experimental or field studies for individual phase properties are used for verification purposes, when available. The main limitation for the selection of appropriate works for verification is the availability of data, as mineralogical determinations based on XRD alone do not provide precise enough information for chemical composition and degree of hydration. Errors are also possible when trying to derive thermodynamic properties from experiments performed in supersaturated conditions, if equilibrium is not reached in the experiments.
- Verification is also carried out by drawing predominance diagrams and investigating phase relations with respect to literature.
- An alternative verification procedure is carried out by comparing the data selection (based on literature experimental data) with model predictions.

2.7 Kinetic data

A compilation of kinetic parameters in the form of individual sheets (one for each mineral phase) is proposed in ThermoChimie (see Marty et al. 2015). The selection of minerals is currently limited to those present in clay rich rocks and cements.

The proposed selection is done (i) to ensure consistency of data selection criteria and data treatment and (ii) to ease the use of common kinetic parameters that are

independent of the chosen geochemical modelling code. Thus, the purpose of these sheets is not to provide a realistic description of mechanisms involved in dissolution/precipitation processes, but to provide data with direct application in widely used geochemical codes. For those two reasons, the kinetic formalism of the transition state theory (TST), as described by Lasaga (1981) (Eq. 4), is chosen in the kinetic description of data.

$$r_n = \pm k_n S_n |1 - \Omega_n^\theta|^n \quad \text{Eq. 4}$$

The selection process is summarized below (see Marty et al. 2015).

1. The **criteria for data selection** comply as well as possible with the pressure and temperature ranges considered in ThermoChimie. The selection depends directly on the parameters to be estimated (i.e. kinetic dissolution, precipitation rate constant, activation energy, etc.). Three types of selection are considered, depending on the process to be modelled:

- dependence on pH, with the selected data covering a pH range as large as possible;
- dependence on temperature, preferentially selecting data in temperatures that incorporate the effect of pH;
- dependence on the saturation ratio.

The selection is done according to the acquisition method, the stoichiometry of the reaction, and the dispersion of data that can be extracted from the literature. Stirred flow-through reactor experiments allow a good control of reaction thermochemical conditions; thus, rates resulting from these kinds of experiments are preferentially selected.

2. **Consistency** is always taken into account in the calculations. Reaction rates are defined with respect to given reactions and since the equilibrium constant term is involved in the equation of the rate, the thermodynamic data used for the calculations are always consistent with the ThermoChimie database. Kinetic rate values also depend on the structural formula considered, which must be the same in both the database and the reaction under consideration.

3. Although some data for the **reactive surface area** of each mineral is provided in the individual sheets, no final data selection for this parameter is recommended. The reasoning behind this decision is the fact that reactive surface area is specific to each studied system and it must be selected by the modeller when defining the geochemical/reactive transport system.

4. The data selected are written in accordance with the **formalism** implemented in the geochemical code Toughreact (Xu et al., 2011), considering the Transition State Theory. This form is chosen because it can be easily transformed into the equations needed by other geochemical codes such as PhreeqC (Parkhurst and Appelo, 2013) or Crunch (Steeffel, 2009).

5. The models developed are finally **compared** with the data available in the literature. This approach allows to assess the uncertainties of the selected rates.

3. Uncertainties

Chemical thermodynamic data may have uncertainties due to both random and systematic errors. Random errors are inherently unpredictable and are scattered about the true value. All data are prone to random error and they cannot be avoided; however, those errors can be estimated and quantified. On the contrary, systematic errors are related to biases in data; those type of errors are reduced and avoided if an accurate selection process as described above is followed.

All the magnitudes included in ThermoChimie must have their own associated uncertainties, when possible. The data selection process includes evaluation of different data sources (dedicated experimental programs, open scientific literature, thermodynamic data compilations, estimations...). The procedure for uncertainty assignment in ThermoChimie is different depending on the data source. Different cases are considered, as described below:

- Data from dedicated experimental programmes
- Data from open literature (original experimental sources or thermodynamic data compilations)

-
- Data from estimations
 - Other cases

In all cases, the consistency and traceability requirements of ThermoChimie (see Guideline 2: Data integration and consistency and guideline 4: Data traceability) apply also to the uncertainty data. A discussion of the possible sources of uncertainty and the appropriate quantification procedure must be reported together with the discussion on the thermodynamic data itself.

3.1 Uncertainties in data from dedicated experimental programs

Many different aspects are involved in the determination of the uncertainty associated to the thermodynamic data acquired from experiments (Figure 5). Its quantification requires looking closely at all the possible sources of uncertainty; however, a detailed study of this kind may require a considerable effort. It is then important to identify the most significant sources of uncertainty; a good estimate of uncertainty can be made by concentrating effort on the identification and quantification of the largest contributions (EURACHEM, 2012). For example: in solubility experiments, the uncertainty related to the calibration of the balance used to weight the solid could be of minor importance in front of the uncertainty related to the solid homogenisation procedure.

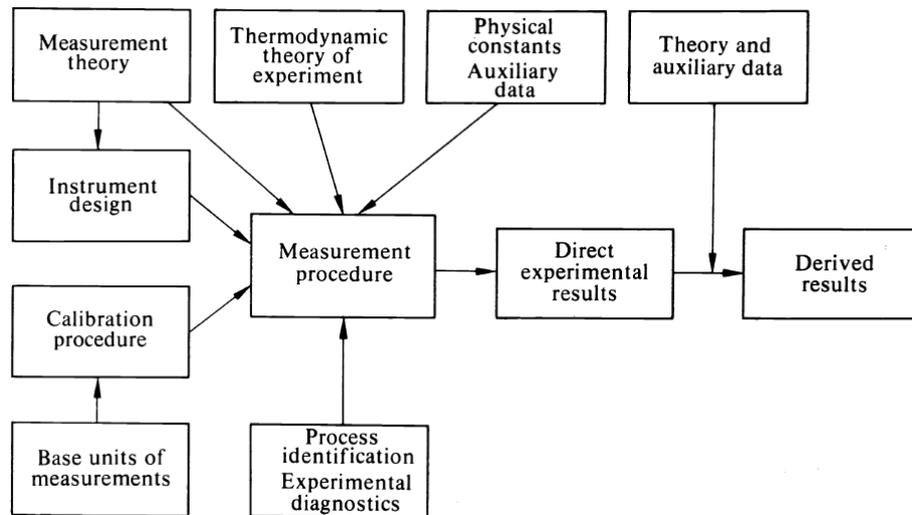


Figure 5. Scheme of different aspects involved in the determination of thermodynamic data (from Olofsson et al. 1981).

When data are obtained from original experimental sources in dedicated experimental programs, the uncertainty assignment should be done by a person familiar with the experimental method applied (EURACHEM, 2012). In this sense, the uncertainty will be representative of the reliability and reproducibility of the experimental value.

The uncertainty assignment will of course depend on the specific system to be studied and on the technique applied, and it is not possible to provide a detailed description of the uncertainty assignment for each case.

However, replicate measurements under unchanged conditions are always advisable (when possible) as they provide indications on the variability associated to the experimental procedure (Olofsson et al. 1981), as heterogeneity of the solid, random errors in analytical determinations, uncertainty associated to volumetric and weight procedures...

Comparison between results obtained from complementary/alternative techniques (e.g. solubility experiments and spectroscopic measurements; solubility from undersaturation and oversaturation directions; comparison between calorimetry and

solubility data, etc...) can also be helpful to provide an indication on the uncertainty associated to a particular experiment.

As discussed above, the possible sources of uncertainty and their appropriate quantification must be reported together with the discussion on the thermodynamic data itself, in order to guarantee the consistency and traceability of the database.

3.2 Uncertainties in data from open literature

The criteria to assign uncertainty to data obtained from the open literature is provided below:

- When the data are obtained from different original experimental sources from the literature, the uncertainty range is selected to include all the data range in the literature (unless very clear indications of the quality of one single publication over the others exist).
- When data are selected from one of the NEA TDB compilations, the uncertainty provided in the NEA review is given.
- When data are derived from compilations other than the NEA TDB, a review of the selection process and also of the original sources is conducted. This review provides information to assign the appropriate uncertainty to the selected value.

3.3 Uncertainties in estimated data

When data are obtained from estimations, the uncertainty is assigned by comparing estimated data with available experimental data and calculating observed deviations. In order to make this, the methodology used in the estimations is used to estimate both missing and available data with similar characteristics. The calculated results are then compared to the available data and the comparison provides indications on the uncertainty range related to that particular estimation procedure.

3.4 Other cases

If additional calculations are needed or in case of internally calculated values, error propagation is applied (Eq. 5). The uncertainty is assessed by using the NEA recommended procedures for the propagation of errors specified in the NEA TDB guidelines (Wanner and Östhols, 1999).

$$\sigma_X^2 = \sum_{i=1}^N \left(\frac{\partial X}{\partial Y_i} \sigma_{Y_i} \right)^2 \quad \text{Eq. 5}$$

4. Bibliography

Blanc, P., Bourbon, X., Lassin, A., & Gaucher, E. C. 2010b. Chemical model for cement-based materials: Thermodynamic data assessment for phases other than C–S–H. *Cement and Concrete Research*, 40(9), 1360-1374.

Blanc, P., Bourbon, X., Lassin, A., Gaucher, E.C. 2010a. Chemical model for cement-based materials: Temperature dependence of thermodynamic functions for nanocrystalline and crystalline C-S-H phases. *Cem. Concr. Res.* 40, 851-866.

Blanc, P., Gailhanou, H., Rogez, J., Mikaelian, G., Kawaji, H., Warmont, F, Gaboreau, S., Grangeon, S., Grenèche, J.M., Vieillard, P., Fialips, C. I., Giffaut, E., Gaucher, E.C. & Claret, F. (2014). Thermodynamic properties of chlorite and berthierine derived from calorimetric measurements. *Physics and Chemistry of Minerals*, 41(8), 603-615.

Blanc, P., Piantone, P., Lassin, A. and Burnol, A. 2006. *ThermoChimie: Sélection de constantes thermodynamiques pour les éléments majeurs, le plomb et le cadmium*. BRGM, Rapport Final (BRGM/RP-54902-FR).

Blanc, P., Vieillard, P., Gailhanou, H., Gaboreau, S. 2013. Chapter 6 - Thermodynamics of Clay Minerals, in: *Developments in Clay Science*. Bergaya, F. and Lagaly, G. (Ed.), Elsevier, pp. 173-210.

Blanc, P., Vieillard, P., Gailhanou, H., Gaboreau, S., Marty, N., Claret, F., Madé, B., & Giffaut, E. 2015. ThermoChimie database developments in the framework of cement/clay interactions. *Applied Geochemistry* 55, 95-107.

Colàs, E., Grivé, M., Rojo, I., & Duro, L. 2011. Solubility of ThO₂·H₂O(am) in the presence of gluconate. *Radiochimica Acta* 99(5), 269-273.

Cox, J.D., Wagman, D.D. and Medvedev, M.A. 1989. CODATA Key Values for Thermodynamics, New York Hemisphere Publishing Corporation.

EURACHEM. 2012. Quantifying Uncertainty in Chemical Measurement. Third Edition. Ellison, S.L.R., Williams, A. (Ed.) QUAM:2012.P1. EURACHEM.

Giffaut, E., Grivé, M., Blanc, P., Vieillard, P., Colàs, E., Gailhanou, H., Gaboreau, S., Marty, N., Madé, B., & Duro, L. 2014. Andra thermodynamic database for performance assessment: ThermoChimie. *Applied Geochemistry*, 49, 225-236.

Grivé, M., Duro, L., Colàs, E., & Giffaut, E. 2015. Thermodynamic data selection applied to radionuclides and chemotoxic elements: An overview of the ThermoChimie-TDB. *Applied Geochemistry* 55, 85-94.

Helgeson, H.C., Delany, J.M., Nesbitt, H.W., Bird, D.K. 1978. Summary and critique of the thermodynamic properties of rock-forming minerals. *Am. J. Sci.* 278-A

Hummel, W., Anderegg, G., Puigdoménech, I., Rao, L., and Tochiyama, O. 2005. Chemical Thermodynamics Series Volume 9: Chemical thermodynamics of compounds and complexes of U, Np, Pu, Am, Tc, Se, Ni and Zr with selected organic ligands. OECD NEA.

Marty, N. C., Claret, F., Lassin, A., Tremosa, J., Blanc, P., Madé, B., Giffaut, E., Cochetin, B. & Tournassat, C. 2015. A database of dissolution and precipitation rates for clay-rocks minerals. *Applied Geochemistry* 55, 108–118.

Olofsson, G., Angus, S., Armstrong, G. T., & Kornilov, A. N. 1981. A report of IUPAC Commission 1.2 on Thermodynamics. Assignment and presentation of uncertainties

of the numerical results of thermodynamic measurements. *J. Chem. Thermodyn*, 13, 603-622.

Parkhurst, D.L., Appelo, C.A.J. 2013. Description of input and Examples for PHREEQC Version 3 — A Computer Program for Speciation, Batch-reaction, Onedimensional Transport, and Inverse Geochemical Calculations: U.S. Geological Survey Techniques and Methods, book 6, Chap. A43.

Ransom B., Helgeson, H.C. 1994. A Chemical and Thermodynamic Model of Aluminous Dioctahedral-2/1 Layer Clay Minerals in Diagenetic Processes - Regular Solution Representation of Interlayer Dehydration in Smectite. *American Journal of Science* 294, 449-484.

Steeffel C.I., 2009. Software for Modeling Multicomponent Reactive Flow and Transport. Rapport technique Lawrence Berkeley National Laboratory.

Vieillard, P. 1994a. Prediction of enthalpy of formation based on refined crystal structures of multisite compounds. 1. Theories and examples. *Geochimica Cosmochimica Acta* 58, 4049-4063.

Vieillard, P. 1994b. Prediction of enthalpy of formation based on refined crystal structures of multisite compounds. 2. Application to minerals belonging to the system $\text{Li}_2\text{O}-\text{Na}_2\text{O}-\text{K}_2\text{O}-\text{BeO}-\text{MgO}-\text{CaO}-\text{MnO}-\text{FeO}-\text{Fe}_2\text{O}_3-\text{Al}_2\text{O}_3-\text{SiO}_2-\text{H}_2\text{O}$. Results and discussion. *Geochimica et Cosmochimica Acta*, 58(19), 4065-4107.

Vieillard, P. 2010. A predictive model for the entropies and heat capacities of zeolites. *Eur. J. Mineral.* 22, 823-836.

Vieillard, P., Blanc, P., Fialips, C.I., Gailhanou, H., Gaboreau, S. 2011. Hydration thermodynamics of the SWy-1 montmorillonite saturated with alkali and alkaline-earth cations: a predictive model. *Geochimica et Cosmochimica Acta* 75, 5664-5685.

Wanner, H., Östholms, E. 1999 TDB-3. Guidelines for the assignment of uncertainties. Version of 1st June 1999. NEA OECD.

Xu, T., Spycher, N., Sonnenthal, E., Zhang, G., Zheng, L., Pruess, K. 2011. TOUGHREACT Version 2.0: a simulator for subsurface reactive transport under non-isothermal multiphase flow conditions. *Comput. Geosci.* 37 (6), 763–774.

ThermoChimie guidelines 2. Data integration and consistency

Project ANDRA TC1

**Mireia Grivé, Elisenda Colàs,
Isaac Campos, Jane Perrone**

July 2015

Written by:	Revised by:	Validated by:
 Isaac Campos Elisenda Colàs	 Jane Perrone	 Mireia Grivé

Index

1. INTRODUCTION	2
2. CONSISTENCY	2
2.1 CONSISTENCY WITH THE FUNDAMENTAL LAWS OF THERMODYNAMICS	3
2.2 CONSISTENCY WITHIN A CHEMICAL MODEL.....	4
2.3 CONSISTENCY BETWEEN ESTIMATED DATA AND EXPERIMENTAL DATA.	4
2.4 CONSISTENCY WITH AUXILIARY DATA.	5
2.5 CONSISTENCY IN THE DATA CORRECTION PROCEDURES.	6
2.5.1 <i>Data correction to zero ionic strength.....</i>	<i>6</i>
2.5.2 <i>Data correction with temperature</i>	<i>7</i>
3. DATA INTEGRATION PROCEDURE	8
3.1 THERMOCHIMIE INFORMATICS DEVELOPMENT	10
3.2 THE EXCEL DATA SHEET.....	12
3.2.1 <i>The information in the excel data sheet.....</i>	<i>12</i>
3.2.2 <i>Internal calculations.....</i>	<i>14</i>
3.2.3 <i>Procedures for data handling.....</i>	<i>16</i>
4. BIBLIOGRAPHY	16
ANNEX A: EXCEL SHEET TEMPLATE	18

1. Introduction

One of the main requirements of ThermoChimie database is consistency. Consistency of a thermodynamic database involves different aspects, and requires the knowledge of a considerably amount of primary information (Wanner, 1991). The purpose of the present guideline is to describe the different processes that must be fulfilled before the inclusion, elimination or modification of any value in ThermoChimie in order to guarantee the consistency of the database.

With this objective, the present document has been divided in two different parts:

- A summary of the different aspects that are involved in the consistency of a thermodynamic database, from the scientific point of view, is provided in section 2.
- When a value is judged appropriate for its inclusion in ThermoChimie, its introduction in the database follows a standard procedure described in section 3. This procedure takes into account both the consistency of the database and the informatics development of ThermoChimie.

2. Consistency

Consistency of a thermodynamic database involves different aspects:

- a) Consistency with the fundamental laws of thermodynamics.
- b) Consistency within a chemical model.
- c) Consistency between estimated data and experimental data.
- d) Consistency with auxiliary data.
- e) Consistency in data correction procedures.

Maintaining the consistency of a database requires a detailed amount of primary information. Furthermore, the very different data sets included in ThermoChimie have different particularities and characteristics that have to be taken into account.

In a general way, consistency checking in ThermoChimie is carried out using as a basis the excel data sheet (see section 3.2). The excel data sheet is build up in order

to facilitate consistency checking with the fundamental laws of thermodynamics and with auxiliary data (points a) and d) in the list above). Consistency within a chemical model, between estimated and experimental data and consistency in data correction procedures (points b), c) and e) in the list above) are maintained if data selection procedures (see Guideline 1: Data selection and uncertainties) are taken into account in the selection process. In order to facilitate the update and improvement of the ThermoChimie database in a consistent way, track-change documents are available (see Guideline 4: Track-changes and track-error document). Those track-change documents describe the evolution of the database and allow understanding the differences between ThermoChimie versions; thus, they can help database developers to identify which systems have to be taken into account in order to maintain consistency. Finally, validation and verification processes (see Guideline 3: Validation and verification) are of the utmost important to ensure that the selected thermodynamic data are accurate and consistent.

The different aspects related with consistency are described below in a more detailed way.

2.1 Consistency with the fundamental laws of thermodynamics

The thermodynamic quantities Gibbs energy, enthalpy and entropy are closely related. When a value has been selected for two of these three variables, the rest of the data are internally calculated using the following equations (Eq. 1, Eq. 2):

$$\Delta_f G_m^0 = \Delta_f H_m^0 - T \left(S_m^0 + \left(\frac{Z}{2} \right) x S_{m,H2(g)}^0 - \sum_i \frac{\nu_i}{\nu_{i,ref}} S_{m,i}^0 \right) \quad \text{Eq. 1}$$

$$\Delta_r G_m^0 = \Delta_r H_m^0 - T \Delta_r S_m^0 \quad \text{Eq. 2}$$

In ThermoChimie, this consistency is ensured by making the appropriate calculations in traceable excel sheets, as fully explained in section 3.2.

2.2 Consistency within a chemical model

The interpretation of data obtained for a chemical system implies a chemical model. In order to maintain the consistency of the database, it is necessary to use the chemical model included in ThermoChimie in order to interpret any new data.

For example, the selection of organic complexation data with radionuclides needs to take into account the protonation constants of ligands in order to evaluate the chemical species in the system. Those protonation constants used in the calculations must be consistent with the selection of hydrolysis data of the organic complexes included in ThermoChimie, as described in Guideline 1: Data selection and uncertainties).

Another example is the selection for cement phases. Before any modification of the database, predominance diagrams corresponding to defined chemical subsystems are drawn, and phase relationships are compared with the observations gathered from the literature. Those tests are a key step in database development and allow maintaining consistency within the chemical model for cement phases.

Consistency must be taken into account in all the developments of the database, including kinetics. Reaction rates are defined with respect to given reactions and since the equilibrium constant term is involved in the equation of the rate, the thermodynamic data used for the calculations must be consistent with the data included in ThermoChimie.

The data selection process for ThermoChimie (see Guideline 1: Data selection and uncertainties) does take into account the aspects related with consistency within a chemical model.

2.3 Consistency between estimated data and experimental data.

Data estimations, analogies and correlations depend on original values (see Guideline 1: Data selection and uncertainties). In order to maintain the consistency of the database, the original values used in the calculations should be the ones included in ThermoChimie database. If the original values are modified, the estimated data should be modified accordingly.

In this sense, Linear Free-Energy Relationships (LFER) are one of the main procedures to take into account. LFER allows establishing the linear dependence of the stability constants with different parameters as the charge density (z^2/r_i), the reciprocal of the ionic radii ($1/r_i$) or the Gibbs energy of formation ($\Delta_f G^\circ$) of the aqueous metal cation.

LFER procedures can be used in different ways:

- LFER are often used to estimate missing stability constants along a chemical series, filling in the thermodynamic gaps caused by the lack of proper experimental data. In those calculations, known stability constants must be used in order to establish the linear relationship. Those known stability constants should be the ones included in ThermoChimie database in order to maintain the consistency of the database. If the known stability data are changed, a new LFER has to be established and the estimated data have to be recalculated accordingly.
- When different stability constants from a common chemical series are included in ThermoChimie, LFER correlations provide a good indication of the quality of the selection and the consistency among the different values in the database. They also provide a good mechanism in order to validate the database contents along the particular chemical series studied.

For example, LFER calculations have been used at different stages of database development in order to fill in data gaps or to check the consistency of the data selection, as in the case of Fe(III)-carbonate system (Grivé, 2005) or in the system An(IV)-gluconate (Gaona et al. 2008, Colàs 2014).

2.4 Consistency with auxiliary data.

The interpretation of experimental measurements and the evaluation of chemical data require the use of auxiliary data (for example data related with primary master species of elements). In order to maintain the consistency of the database, the auxiliary data used in calculations must be consistent with the auxiliary data included in ThermoChimie.

As described in Guideline 1: Data selection and uncertainties, ThermoChimie relies on CODATA recommendations (Cox et al. 1989) for the primary master species and the reference states of the elements, when possible. Any update or modification of the CODATA values or of the auxiliary data published in the NEA TDB project is taken into account. In any case, those data are also reviewed and verified during the selection process, following the data selection criteria applied to any other chemical species in the database.

In some specific cases (Al and Si) thermodynamic data selected for basic species are not exactly the same as the values recommended by CODATA (see Blanc et al. 2006). Those changes require a solid discussion and supplementary justifications, as well as an exhaustive analysis of the practical consequences that derivate from this decision, i.e., exhaustive consistency checking for the whole database. Further updates and verifications on this particular issue are envisaged during ThermoChimie development.

2.5 Consistency in the data correction procedures.

ThermoChimie contains values that refer to standard conditions (zero ionic strength) and to a reference temperature (25°C). However, the experimental determination of equilibrium constants and other thermodynamic data is done at non-standard conditions. The data used to perform those corrections should be consistent with the data included in the database.

2.5.1 Data correction to zero ionic strength

In the case of data correction to zero ionic strength, ThermoChimie primes the Specific Interaction Theory (SIT), which is the selected ionic strength correction approach used by ThermoChimie, as reference for further developments of the data base so that this approach should then be used, when possible, in order to obtain the corrected values at $I=0$.

SIT is not frequently used in data base compilations (with the exception of the NEA TDB project) or in other thermodynamic databases. Thus, special attention has to be drawn during the data update and modification processes, in order to maintain database consistency.

Data selection procedure for ion interaction coefficients is described in Guideline 1: Data selection and uncertainties. The SIT equation is shown in Eq. 3.

$$\log(\gamma_i) = -z_i^2 \left(\frac{A\sqrt{I_m}}{1 + Ba_i\sqrt{I_m}} \right) + \sum_k \varepsilon(i, k, I_m) m_k \quad \text{Eq. 3}$$

If individual, reliable and consistent interaction coefficients are available, it is possible to calculate $\log_{10}K^0$ values accordingly. If the individual interaction coefficients are not reported, it is possible to calculate $\log_{10}K^0$ values using stability constants at different ionic strengths, by interpolating $\log_{10}K^0$ and $\Delta\varepsilon$ from a linear regression. In other cases, it is possible to estimate the missing interaction coefficients, by using analogies or correlations.

The SIT approach takes into account short-range non-electrostatic interactions by adding terms to the Debye-Hückel expression (Eq. 3). Thus, in the case that it is not possible to use SIT in the correction (due to limited amount of data at different ionic strengths and the lack of appropriate ion interaction coefficients) the Debye-Hückel approach may then be employed.

2.5.2 Data correction with temperature

ThermoChimie contains very different data sets including thermodynamic data for major elements, radioactive and toxic elements, organic ligands, cement hydrates, clay minerals and zeolites, and kinetic data for some minerals. Each of these groups has its own characteristics, and handling of data at temperatures different than 25°C depends on the particularities of the type of data, as described in Guideline 1: Data selection and uncertainties.

In any case, the extrapolation of equilibrium constants with the temperature is usually done by using various modifications of the so-called second- and third-law methods, and involves the use of some or all of the following parameters: enthalpy, entropy and/or heat capacity data.

The temperature dependence of the equilibrium constant can be written as a function of the standard enthalpy and the standard heat capacity (Eq. 4):

$$\log_{10} K^\circ(T) = \log_{10} K^\circ(T_0) - \frac{\Delta_r H_m^\circ(T_0)}{R \ln(10)} \left(\frac{1}{T} - \frac{1}{T_0} \right) - \frac{1}{R \ln(10)} \int_{T_0}^T \Delta_r C_{p,m}^\circ dT + \frac{1}{R \ln(10)} \int_{T_0}^T \frac{\Delta_r C_{p,m}^\circ}{T} dT, \quad \text{Eq. 4}$$

A possible approximation is to assume that the heat capacity of the reaction does not vary with temperature, which leads to the constant enthalpy of reaction equation (Eq. 5).

$$\log_{10} K^\circ(T) = \log_{10} K^\circ(T_0) + \frac{\Delta_r H_m^\circ(T_0)}{R \ln(10)} \left(\frac{1}{T_0} - \frac{1}{T} \right) + \frac{\Delta_r C_{p,m}^\circ}{R \ln(10)} [(T_0/T) - 1 + \ln(T/T_0)] \quad \text{Eq. 5}$$

Unfortunately, experimentally derived heat capacity data have not been measured for many relevant aqueous species and solid phases. Then, the simplest assumption to be made is that the heat capacity change of reaction is zero at all temperatures (i.e., the standard molar enthalpy of reaction does not vary with temperature) In that case, the Van't Hoff expression is obtained (Eq. 6).

$$\log_{10} K^\circ(T) = \log_{10} K^\circ(T_0) + \frac{\Delta_r H_m^\circ(T_0)}{R \ln(10)} \left(\frac{1}{T_0} - \frac{1}{T} \right) \quad \text{Eq. 6}$$

As far as enthalpies of reaction (and in some cases Cp values) are included in ThermoChimie, temperature correction calculations can be performed in a consistent way. At the present stage of development of the database, temperature corrections are mainly done by using the Van't Hoff approach. Further updates on this particular issue are envisaged during the ThermoChimie development.

3. Data integration procedure

The different processes summarized in Figure 1 must be fulfilled previous to the inclusion, the elimination or the modification of any value in ThermoChimie database. This standard procedure is defined to maintain the quality and consistency of the

database, and takes into account both the consistency requirements described above and the informatics development of ThermoChimie explained below.

After data selection, the excel data sheet described in section 3.2 is filled and provided to database administrators. Then, the appropriate consistency checking is carried out and the data is introduced in the informatics system of ThermoChimie (Access and Oracle). At the same time, the appropriate supporting documents are generated (see Guideline 4: Data traceability and Guideline 5: Bibliographic reference recording process).

The data in the Oracle format are then extracted into the different formats appropriate for relevant geochemical codes. The adequate validation and/or verification procedures (see Guideline 3: Validation and verification) are carried out before a new ThermoChimie release is done in the web site.

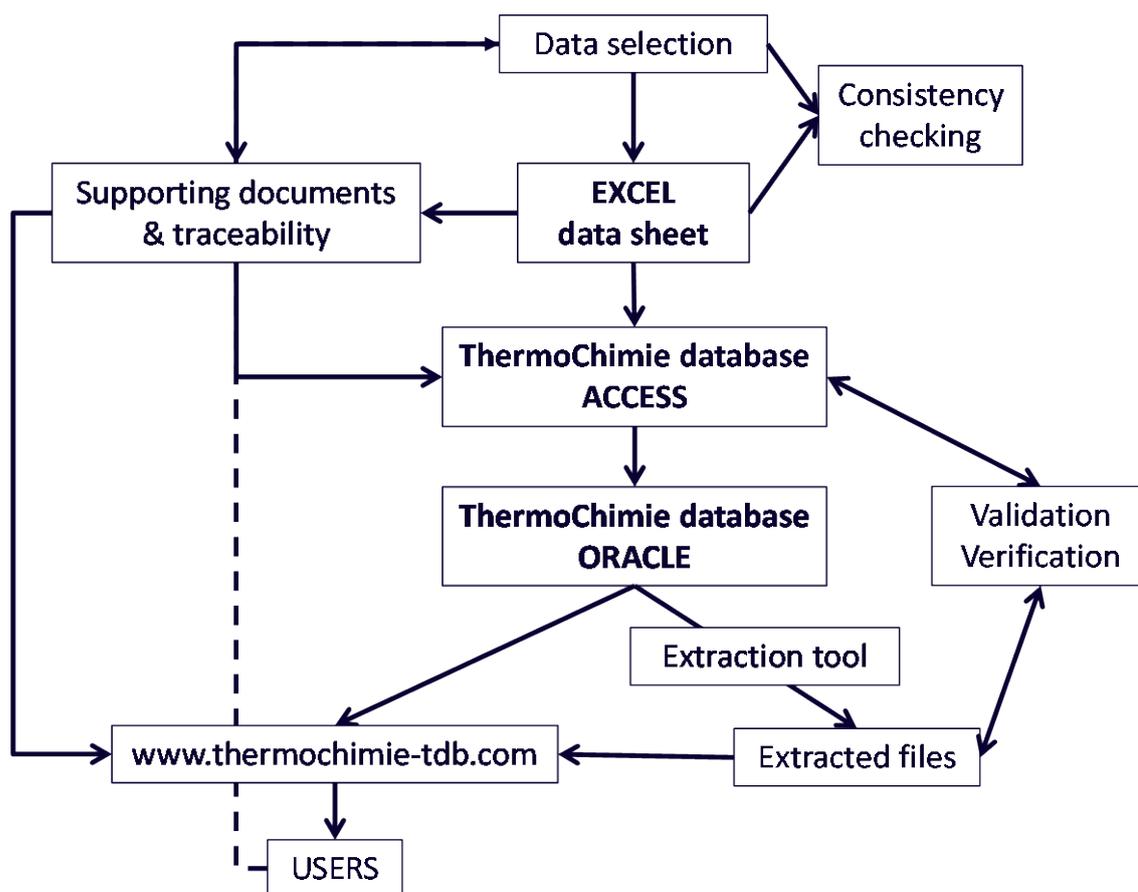


Figure 1. Summary of relevant processes in ThermoChimie database development.

The next sections intend to explain in a more detailed way the informatics development of ThermoChimie (Access, Oracle, and extraction tool) and to explain in a detailed way how the excel sheet must be filled.

3.1 ThermoChimie informatics development

The informatics structure of ThermoChimie (Figure 2) is based on an ACCESS environment. The consultation interface of ThermoChimie supported by this ACCESS application allows database administrators to perform requests and consultation of species and values, but also to include, modify or eliminate data in order to update new database versions. Then, once a value is judged appropriate for its inclusion in

ThermoChimie, it is initially introduced in ThermoChimie database in the ACCESS format.

As well as the ACCESS environment, ThermoChimie database is also restructured following an Oracle basis (Figure 2). This restructuration allows the database to be compatible to an extraction tool application and with the website development.

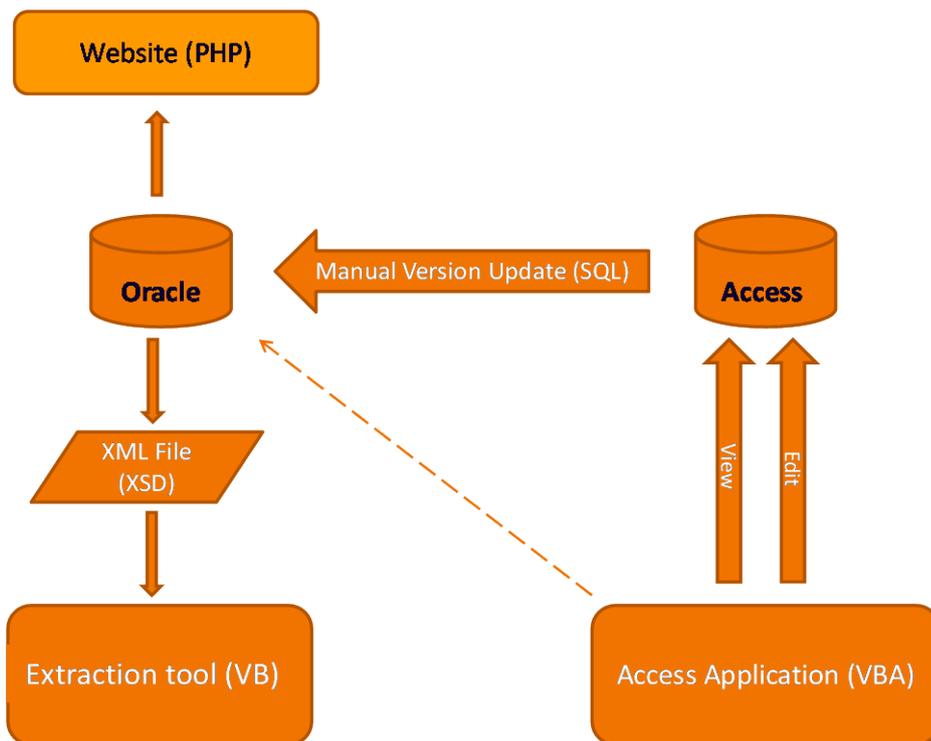


Figure 2. ThermoChimie informatics system overview.

ThermoChimie has an additional functionality, a Visual Basic extraction tool. The main objective of this tool is to extract the information contained in ThermoChimie into the different text format files that the geochemical codes (e.g. Phreeqc, Crunch, ToughReact, EQ3, CHESS, Geochemist Workbench) need to understand thermodynamic data. This application allows a wide diffusion of ThermoChimie to the scientific community, in order to fulfil the requirement of applicability (usability) of the database.

Finally, the restructuration of ThermoChimie following an Oracle basis allows the compatibility of the database and a website (<https://www.thermochimie-tdb.com/>). This website provides the user the opportunity to:

-
- consult the raw thermodynamic data contained in the different ThermoChimie versions
 - download the ThermoChimie database in the different formats compatible with different geochemical codes
 - consult ThermoChimie supporting information and guidelines
 - provide feedback

3.2 The excel data sheet

The database includes information on elements, aqueous species (comprising primary master species, secondary master species, and product species), minerals and other solid and gases. Prior to the introduction, modification or elimination of any of those data in the ThermoChimie database in the ACCESS format, an EXCEL data sheet (Figure 1) must be filled and provided to database administrators. A template of the Excel data sheet will be provided to any database contributors who will fill in by following the indications in this guideline.

An example of an excel data sheet is provided in Annex A; an explanation on how to fill in this excel data sheet is provided below.

3.2.1 The information in the excel data sheet

The data contained in the excel sheet is divided in two different parts, one for master species and another one for the selected data for product species.

3.2.1.1 Part A: Master species

The **master species** are the building blocks of the formation reactions for product species, either aqueous species, solids or gases. There must be at least one master species for each element to be included in the database. Two types of master species are defined: **primary master species and secondary master species**.

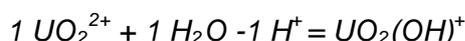
Example (see Annex A): In the case of the element Uranium (U), UO_2^{2+} is the primary master species in ThermoChimie. Secondary master species defined in the database are U^{3+} , U^{4+} , and UO^{2+} .

Thermodynamic properties for the master species ($\Delta_f G_m^0$, $\Delta_f H_m^0$ and S_m^0 , as well as their corresponding uncertainties) are shown in the top of the Excel sheet (see Annex A). These data are used to perform the calculations of thermodynamic data of selected species. The inclusion of these data in the excel sheet allows database administrations to verify that the consistency with auxiliary data is fulfilled in the calculations.

3.2.1.2 Part B: Product species

Product species are defined by reaction among master species, with the appropriate reaction coefficients.

Example (see Annex A): The reaction for the formation of the product species $UO_2(OH)^+$ is written in terms of the master species UO_2^{2+} , H_2O and H^+ .



Data for product species are shown in the bottom part of the excel sheet. Data included in the excel sheet are described below (see also Guideline 1: Data selection and uncertainties):

- **Log₁₀K⁰** at 25 °C: Equilibrium constant of the reaction.
- **$\Delta_r G_m^0$** (KJ·mol⁻¹) at 25 °C: Molar Gibbs energy of reaction.
- **$\Delta_r H_m^0$** (KJ·mol⁻¹) at 25 °C: Molar enthalpy of reaction.
- **$\Delta_r S_m^0$** (J·K⁻¹·mol⁻¹) at 25 °C: Molar entropy of reaction.
- **$\Delta_f G_m^0$** (KJ·mol⁻¹) at 25 °C: Standard molar Gibbs energy of formation.
- **$\Delta_f H_m^0$** (KJ·mol⁻¹) at 25 °C: Enthalpy of formation.
- **S_m^0** (J·K⁻¹·mol⁻¹) at 25 °C: Standard molar entropy.
- **Ion interaction coefficients $\epsilon(j, k)$** (Kg·mol⁻¹) for the interaction of the species with Na⁺, Cl⁻, ClO₄⁻ and/or NO₃⁻, when available, at 25 °C (for aqueous species only).

-
- $C_{p,m}^{\circ}$ ($J \cdot K^{-1} \cdot mol^{-1}$): Standard molar heat capacity at constant pressure (for solid compounds only).
 - V_m° ($cm^3 \cdot mol^{-1}$): Molar volume (for solid compounds only).

All the values included in ThermoChimie are accompanied by:

- a bibliographic reference of the source from where the datum has been adapted
- the associated uncertainty.
- It is foreseen that in the future, the values will also include qualifiers; those qualifiers are specific labels that will provide information about the quality and overall reliability of each datum (see Guideline 6: Data quality).

The following notation is used in order to fill in the data in the excel sheet:

- Values in bold and blue are the original values selected.
- Values in white correspond to internally calculated values (see section 3.2.2).
- Ion interaction coefficients $\epsilon(j, k)$ are shown in black and regular source.

Example (see Annex A): Selected thermodynamic data for the product species $UO_2(OH)^+$ are the equilibrium constant of the reaction and the standard molar entropy, which are the parameters in bold and blue in the table.

3.2.2 Internal calculations

When values have been selected for two of the main thermodynamic variables (Gibbs energy, enthalpy or entropy), the rest of the data must be internally calculated in order to fulfil consistency requirements with the fundamental laws of thermodynamics (see section 2.1). The use of the excel sheet allows database administrations to verify that this requirement is fulfilled before the inclusion of the data in ThermoChimie.

Descriptions of the thermodynamic equations that have to be applied in the internal calculations are detailed below. Notice that in the excel sheet, enthalpy and entropy units are $KJ \cdot mol^{-1}$ and entropy units are $J \cdot K^{-1} \cdot mol^{-1}$. This must be taken into account when calculations are performed.

Eq. 7 and Eq. 8 describe the fundamental laws to be fulfilled:

$$\Delta_f G_m^0 = \Delta_f H_m^0 - T \left(S_m^0 + \left(\frac{Z}{2} \right) x S_{m,H2(g)}^0 - \sum_i \frac{\nu_i}{\nu_{i,ref}} S_{m,i}^0 \right) \quad \text{Eq. 7}$$

$$\Delta_r G_m^0 = \Delta_r H_m^0 - T \Delta_r S_m^0 \quad \text{Eq. 8}$$

If standard molar Gibbs energy of formation of all reactants and products are available, the Gibbs energy of reaction can also be calculated using Eq. 9. Results obtained with Eq. 9 and with Eq. 8 should be equivalent. A similar approach can be applied in the case of the enthalpy and the entropy of reaction (Eq. 10 and Eq. 11).

$$\Delta_r G_m^0 = \sum_i \nu_i \Delta_f G_{m,products}^0 - \sum_i \nu_i \Delta_f G_{m,reactants}^0 \quad \text{Eq. 9}$$

$$\Delta_r H_m^0 = \sum_i \nu_i \Delta_f H_{m,products}^0 - \sum_i \nu_i \Delta_f H_{m,reactants}^0 \quad \text{Eq. 10}$$

$$\Delta_r S_m^0 = \sum_i \nu_i \Delta_f S_{m,products}^0 - \sum_i \nu_i \Delta_f S_{m,reactants}^0 \quad \text{Eq. 11}$$

The relationship between the equilibrium constant $\log_{10} K^0$ and $\Delta_r G_m^0$ is shown in Eq. 12:

$$\log K^0 = \frac{-\Delta_r G_m^0}{RT \ln(10)} \quad \text{Eq. 12}$$

For the uncertainty of internally calculated values, error propagation is applied as shown in Eq. 13 (see also Guideline 1: Data selection and uncertainties).

$$\sigma_X^2 = \sum_{i=1}^N \left(\frac{\partial X}{\partial Y_i} \sigma_{Y_i} \right)^2 \quad \text{Eq. 13}$$

Example (see Annex A): In the case of the $\text{UO}_2(\text{OH})^+$ species, the equilibrium constant of the reaction is the selected value. $\Delta_r G_m^0$ is calculated using Eq. 12 and $\Delta_r G_m^0$ using Eq. 9.

3.2.3 Procedures for data handling

*It is important to notice that **all calculations and operations must be carried out in full numbers**, and only the final results are rounded, in order to avoid loss of information.*

Independently of the magnitude of the uncertainty, in ThermoChimie database $\Delta_r G_m^0$, $\Delta_r H_m^0$, $\Delta_r G_m^0$ and $\Delta_r H_m^0$ and $\varepsilon(j, k)$ are provided with **three decimal places**; $\log_{10} K^0$, $C_{p,m}^0$ and V_m^0 data are provided with **two decimal places**; and, finally, $\Delta_r S_m^0$ and S_m^0 with **one decimal place**. The same criteria are applied for uncertainties.

Rules to be used for rounding are the same that recommended by NEA TDB guidelines (Wanner and Östhols, 1999):

- i) When the digit following the last digit to be rounded is less than 5, last digit keeps unchanged.
- ii) When the digit following the last digit to be rounded is greater than 5, last digit retained is increased by 1.

4. Bibliography

Blanc, P., Piantone, P., Lassin, A. and Burnol, A. 2006. ThermoChimie: Sélection de constantes thermodynamiques pour les éléments majeurs, le plomb et le cadmium. BRGM, Rapport Final (BRGM/RP-54902-FR).

Colàs, E. 2014. Complexation of Th(IV) and U(VI) by polyhydroxy and polyamino carboxylic acids. PhD Thesis, Universitat Politècnica de Catalunya, Barcelona.

Cox, J.D., Wagman, D.D. and Medvedev, M.A. 1989. CODATA Key Values for Thermodynamics, New York Hemisphere Publishing Corporation.

Gaona, X., Montoya, V., Colàs, E., Grivé, M., & Duro, L. 2008. Review of the complexation of tetravalent actinides by ISA and gluconate under alkaline to hyperalkaline conditions. *Journal of Contaminant Hydrology*, 102(3), 217-227.

Grivé, M. (2005). The linkage between uranium, iron and carbon cycling processes at interfaces: evidences from combined solution chemical and spectroscopic studies. PhD Thesis, Universitat Politècnica de Catalunya, Barcelona.

Wanner, H. 1991. On the problem of consistency of chemical thermodynamic data bases. *Mat. Res. Soc. Symp. Proc. Vol. 212*, 815-822.

Annex A: Excel Sheet template

A template for the excel sheet described in section 3.2 is shown below and will be provided to any database contributors who will fill in, by following the indications in this guideline.

PART A: Basic components:

Specie	UO ₂ (OH) ⁺
Species	UO ₂ (OH) ⁺
coef UO ₂ ⁺²	1
coef e-	
coef H ₂ O	1
coef H ⁺	-1
Elt. 1	UO ₂ ⁺²
Coef. 1	1
dGf	-952.551
idGf	1.747
dHf	-1019.000
idHf	1.500
Sf	-98.2
iSf	3.0
Elt. 2	e-
Coef. 2	0
dGf	0.000
idGf	0.000
dHf	0.000
idHf	0.000
Sf	65.34
iSf	0.001
Elt. 3	H ₂ O
Coef. 3	1
dGf	-237.140
idGf	0.041
dHf	-285.830
idHf	0.040
Sf	70.0
iSf	0.0
Elt. 4	H ⁺
Coef. 4	-1
dGf	0.000
idGf	0.000
dHf	0.000
idHf	0.000
Sf	0.0
iSf	0.0
Elt. 5	
Coef. 5	

Specie	Calcite
Species	CaCO ₃
coef Ca ⁺²	1
coef CO ₃ ⁻²	1
coef H ₂ O	
coef H ⁺	
Elt. 1	Ca ⁺²
Coef. 1	1
dGf	-552.806
idGf	1.044
dHf	-543.000
idHf	1.000
Sf	-56.2
iSf	1.0
Elt. 2	CO ₃ ⁻²
Coef. 2	1
dGf	-527.900
idGf	0.390
dHf	-675.230
idHf	0.250
Sf	-50.0
iSf	1.0
Elt. 3	H ₂ O
Coef. 3	0
dGf	-237.140
idGf	0.041
dHf	-285.830
idHf	0.040
Sf	70.0
iSf	0.0
Elt. 4	H ⁺
Coef. 4	0
dGf	0.000
idGf	0.000
dHf	0.000
idHf	0.000
Sf	0.0
iSf	0.0
Elt. 5	
Coef. 5	

PART B: Selected species:

Specie	UO ₂ (OH) ⁺
Species	UO ₂ (OH) ⁺
Log K ^o	-5.25
± Log K ^o	0.24
Δ _r G ^o	29.967
± Δ _r G ^o	1.370
Δ _r H ^o	43.458
± Δ _r H ^o	14.997
Δ _r S ^o	45.3
± Δ _r S ^o	50.1
Δ _f G ^o	-1159.724
± Δ _f G ^o	2.220
Δ _f H ^o	-1261.372
± Δ _f H ^o	15.072
S ^o	17.00
± S ^o	50.00
C _p	
± C _p	
V _m	
± V _m	
ε M+/Cl-	-0.003
i.e M+/Cl-	0.231
ε M+/ClO ₄ ⁻	-0.060
i.e M+/ClO ₄ ⁻	0.400
ε M+/NO ₃ ⁻	0.510
i.e M+/NO ₃ ⁻	1.400
comment	
Ref. Log K ^o / Δ _r G ^o / Δ _f G ^o	03GUI/FAN
Ref. Δ _r H ^o / Δ _f H ^o	
Ref. Δ _r S ^o / S ^o	92GRE/FUG
Ref. C _p	
Ref. V _m	
Ref. ε Cl-	a)
Ref. ε ClO ₄ ⁻	92GRE/FUG
Ref. ε NO ₃ ⁻	92GRE/FUG
Quality label Log K ^o	B1
Quality label Δ _r G ^o	D1
Quality label Δ _r H ^o	D1
Quality label Δ _r S ^o	D1
Quality label Δ _f G ^o	D1
Quality label Δ _f H ^o	D1
Quality label S ^o	B 1
Quality label C _p	
Quality label V _m	
Quality label ε Cl-	B1
Quality label ε ClO ₄ ⁻	C1
Quality label ε NO ₃ ⁻	B1
Comments	
a) Calculated according to the value with ClO ₄ ⁻ recommended by 95SIL/BID	

Specie	Calcite
Species	CaCO ₃
Log K ^o	8.48
± Log K ^o	0.02
Δ _r G ^o	-48.404
± Δ _r G ^o	0.114
Δ _r H ^o	10.620
± Δ _r H ^o	
Δ _r S ^o	198.0
± Δ _r S ^o	
Δ _f G ^o	-1129.110
± Δ _f G ^o	1.120
Δ _f H ^o	-1207.610
± Δ _f H ^o	
S ^o	91.77
± S ^o	
C _p	83.45
± C _p	
V _m	36.93
± V _m	
ε M+/Cl-	
i.e M+/Cl-	
ε M+/ClO ₄ ⁻	
i.e M+/ClO ₄ ⁻	
ε M+/NO ₃ ⁻	
i.e M+/NO ₃ ⁻	
comment	
Ref. Log K ^o / Δ _r G ^o / Δ _f G ^o	82PLU/BUS
Ref. Δ _r H ^o / Δ _f H ^o	82PLU/BUS
Ref. Δ _r S ^o / S ^o	
Ref. C _p	95ROB/HEM
Ref. V _m	10BLA/BOU
Ref. ε Cl-	
Ref. ε ClO ₄ ⁻	
Ref. ε NO ₃ ⁻	
Quality label Log K ^o	B1
Quality label Δ _r G ^o	D1
Quality label Δ _r H ^o	D1
Quality label Δ _r S ^o	D1
Quality label Δ _f G ^o	D1
Quality label Δ _f H ^o	D1
Quality label S ^o	D1
Quality label C _p	B1
Quality label V _m	B1
Quality label ε Cl-	
Quality label ε ClO ₄ ⁻	
Quality label ε NO ₃ ⁻	
Comments	

ThermoChimie guidelines 3. Validation and verification

Project ANDRA TC1

**Mireia Grivé, Elisenda Colàs,
Isaac Campos, Jane Perrone**

July 2015

Written by:	Revised by:	Validated by:
 Isaac Campos Elisenda Colàs	 Jane Perrone	 Mireia Grivé

AMPHOS²¹



Index

1. INTRODUCTION	2
2. VERIFICATION	4
3. VALIDATION	6
4. BIBLIOGRAPHY	10

1. Introduction

The reliability of a thermodynamic database is often judged by its capability to reproduce independent experimental data gathered from either laboratory experiments or field observations. In order to guarantee this requirement, a strict **validation of the selected thermodynamic data** is essential. Validation and verification procedures are important to ensure that all the aqueous species and solid phases relevant under the conditions of the applicability of the database are included. These procedures also provide information on the consistency and accuracy of the database and allow to identify whether the adequate uncertainty is assigned to less accurate data.

Figure 1 shows a general scheme of the validation and verification process in the context of ThermoChimie database development. In this context, the present guideline intends to explain the different verification and validation exercises carried out.

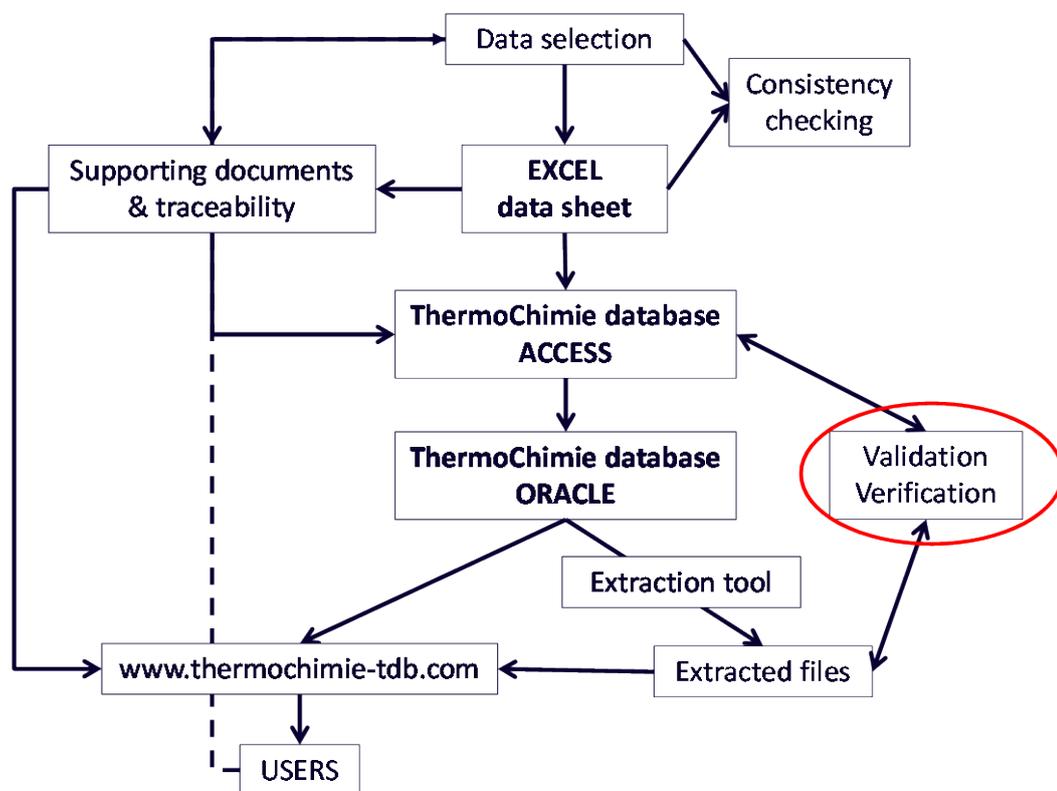


Figure 1. Validation and verification processes in the context of ThermoChimie database development.

The terms “validation” and “verification” are sometimes used as synonyms; however, their meaning is slightly different.

Verification refers to the process of determining that the implementation of the model accurately represents the developer’s description of this model. In the specific case of ThermoChimie, it will include all the processes developed to check that the data are correctly stored in and extracted from the database¹.

On the other hand, **validation** refers to the process of determining the degree to which a model is an accurate representation of the real world (within the limits for

¹ Further details on data integration procedures and extraction can be found in Guideline 2: Data integration and consistency.

which the model has been developed for). In the case of ThermoChimie, validation will include all the process developed to ensure that the database is accurate and exhaustive enough to achieve geochemical modelling and performance assessment tasks for which it has been designed for².

2. Verification

ThermoChimie database verification is mainly focused on checking that 1) the data are correctly stored in the database, and b) that no data are lost, altered or modified during the application of the extraction routines.

Different types of verification exercises are used:

1. **Direct comparison exercises** among different versions of the ThermoChimie database. When a new database version is ready to be released, database administrators perform a direct comparison between the data included in the new version and data included in former versions, using an excel macro specifically designed for this task. The main objective of the exercise is to ensure that data are not lost or altered when introduced in the Access version of ThermoChimie. Those comparisons increase the traceability of the database and allow confirming that all the relevant information is included in the track-changes and track-errors document (see Guideline 4: Data traceability).
2. **Benchmarking exercises.** In this type of exercises, a geochemical model is run several times with different geochemical codes (PhreeqC, Crunch, ToughReact, Chess...)³, using ThermoChimie database. Different geochemical models are considered in this type of exercises, in order to cover all the different systems that are covered by the database. The calculations are framed under conditions relevant for database application. Some examples are:
 - Radionuclide solubility and speciation under clay or cement conditions.

² Further details on ThermoChimie database objectives and principles of development are detailed in the Introduction to ThermoChimie guidelines.

-
- Radionuclide solubility and speciation at different temperatures.
 - Organic speciation under clay or cement conditions.
 - Speciation of major ions as a function of temperature.
 - Radionuclide and major ions speciation under different ionic strength conditions, using different ionic strength correction approaches (SIT, extended Debye-Hückel, Davies...)
 - Definition of cement phase assemblages representative of a cement at a specific degradation stage
 - Calculation of the porewater composition in equilibrium with Callovo-Oxfordian formation at different temperatures.
 - Simplified 1-D transport calculations

It should be noticed that not all the geochemical codes are capable of handling all calculations; thus, only the appropriate codes are used in each exercise.

These type of exercises are run when a major ThermoChimie version is released, or when significant updates and improvements are performed in the extraction tool. The analysis of the results obtained help to ensure that the extraction of ThermoChimie into the different formats needed by different geochemical codes (see Guideline 2: Data integration and consistency) is correctly done. In addition, these exercises help to adapt the extracted files to the needs of the codes.

3. Modelling of experimental data. In this type of exercises, experimental data used in the development of ThermoChimie are modelled using ThermoChimie and an appropriate geochemical code.

- Radionuclide solubility calculation and comparison with the original experimental data used in database update.
- Organic speciation calculation and comparison with the original experimental data used in database update.

³ See the Introduction to ThermoChimie guidelines for additional information on available formats of ThermoChimie.

-
- Radionuclide and major ions speciation under different ionic strength conditions calculation and comparison with the original experimental data used in database update.
 - Definition of cement phases predominance diagrams corresponding to defined chemical subsystems.
 - Definition of cement phases predominance diagrams corresponding to defined chemical subsystem.
 - Drawing activity diagrams involving clay minerals in a defined chemical subsystem.

The performance of the database in the interpretation of the data allows ensuring that the data are correctly stored in the database.

3. Validation

Validation exercises based on geochemical calculations using ThermoChimie database are continuously undertaken. Those exercises cover the whole range of aqueous species and solid phases relevant for the different scenarios of interest.

Validation of thermodynamic data of ThermoChimie should be accomplished by conducting tests or studies that show that the geochemical model, implemented by the combination of the user's input parameters, the thermodynamic database, and the geochemical code, produces an outcome that is an acceptable representation of reality.

Because of the high number of solid phases and aqueous species included in the database, it is not possible to test all of them simultaneously; therefore, different validation exercises are proposed to cover the whole range of aqueous species and solid phases appearing in the different scenarios of interest.

One of the exercises described below are done for each data selection, when possible. The choice of one or another validation test is performed depending on the availability of the appropriate data.

Three different types of validation exercises are proposed:

1. Direct comparison exercises among ThermoChimie and another thermodynamic database. These comparisons are focused on checking both the species and minerals included in each database, and the numerical values of the stability constants included in each database. The investigation of the reasons behind the similitudes and discrepancies helps to identify additional information sources and to check the completeness of the database; this is, that the database does contain all data needed to account for the expected behaviour of the studied systems.

Databases used in the comparison exercises have to fulfill some requisites:

- The field of application for which the database has been designed for must have common points with the ThermoChimie application fields;
- Databases have to be public, available through public web sites or through public printed documents;
- Its development must be independent of that of ThermoChimie.

Some examples of databases that can be included in the comparison exercises are: the **LLNL database** of the Lawrence Livermore National Laboratory (<https://www.llnl.gov/>); **JNC database** of the Japan Atomic Energy Agency (JAEA) (<https://www.jaea.go.jp/english/index.html>); the **Nagra-PSI TDB** of the Paul Scherrer Institute (PSI) (<http://www.psi.ch/les/database>); or the **Thereda database**, a cooperative project of several institutions on the field of radioactive waste disposal in Germany (<https://www.thereda.de/>), among others.

2. Benchmarking exercises. In this type of exercises, literature is reviewed to get an example of a geochemical model within the application fields of ThermoChimie. Then, this geochemical model is run several times with the same code but with different thermodynamic databases, one of them being ThermoChimie. Results obtained are compared and conclusions extracted.

The detailed analysis of the results obtained allows identifying new possibilities of improvement for ThermoChimie:

-
- The exercises may point out data gaps: solids or aqueous species not included in ThermoChimie that could be relevant for the accurate representation of the geochemical model.
 - The exercises may point out data discrepancies: significant differences in thermodynamic data for solids or aqueous species included in the different databases that lead to different results in the representation of the geochemical model.

In those cases, a detailed study on the different compounds identified during the benchmark exercises is done, following the ThermoChimie criteria for data selection (see Guideline 1: Data selection and uncertainties). In some cases, a devoted experimental program could be defined in order to cover the identified data gaps. If adequate, the new or updated data will be included in ThermoChimie following the standard procedure, and a new version of the database will be released (see Guideline 2: Data integration and consistency).

A particular example of application of these validation exercises is to check the performance of the SIT approach. ThermoChimie primes the Specific Interaction Theory (SIT) for ionic strength corrections. Other databases use different ionic strength correction approaches (e.g. Pitzer or extended Debye-Huckel equations). In this specific case, validation exercises are carried out by modelling solubility data for different minerals under high ionic strength conditions, using different databases with different ionic strength approaches. The detailed analysis of the results obtained allows identifying new possibilities of improvement for ThermoChimie:

- The exercises provide indications of the validity of the database in terms of the ionic strength of the system. Specifically, it helps to demonstrate the suitability of the SIT approach in front of Pitzer. Although the Pitzer approach should be in principle more adequate for modelling systems at high ionic strength conditions, those tests have demonstrated

that the lack of the appropriate Pitzer parameters can be a serious handicap.

- The exercises may point out significant data gaps in terms of missing or inaccurate ion interaction coefficients. This helps to focus the efforts for database development in the most relevant points.
- Exercises also help to adequate the database to the requirements of different geochemical codes.

3. Modelling of experimental data. A validated thermodynamic database is characterized by a good agreement between calculated and available experimental results. In this type of exercises, literature is reviewed to get experimental data that have not been previously used in the data selection process, when available. Then, these new sets of data are modelled using ThermoChimie and an appropriate geochemical code. The performance of the database in the interpretation of the new data is tested, results are interpreted and conclusions extracted.

- The open literature is continuously reviewed in order to identify suitable experimental data. These data should be independent (not used in the selection and update process of the database), reliable (the experiments must be performed with a sound scientific basis, without systematic errors), and detailed (extensive analysis of the aqueous phase, solid characterization, etc. should be provided in the paper). This is a limiting step in the process, because most of the relevant experimental studies available in the literature have already been used during database development.
- During the update procedure (and before any major ThermoChimie release), the available data are modelled using ThermoChimie. Taking into account the high number of compounds included in the database, the validation exercises are focused on the systems that are being updated in that particular database version (see Guideline 4: Data traceability). For example: if the cement system has been significantly updated, efforts will be focused on validation exercises on cement systems.

-
- The results of the validation exercise will point out the database strengths and the possibilities of improvement for the system studied. If the results are not satisfactory, ThermoChimie will be improved before the new release is done.

In addition to the three validation procedures defined above, **modelling of natural analogue data** has also been used for ThermoChimie validation, although in a less systematic way. Modelling natural analogues goes beyond the purposes of database validation, as it includes testing the conceptual geochemical models and the required numerical codes as well as the associated thermodynamic databases.

An example of modelling of natural analogues is showed in the work of Bruno et al. (2002) where radionuclide solubilities calculated with ThermoChimie were systematically compared with the data gathered from natural analogues. Poços de Cladas (Brazil), Cigar Lake (Canada), Maqurin (Jordan), El Berrocal (Spain), Oklo (Gabon) or Palmottu (Finland) are some of the natural analogues that can be considered.

4. Bibliography

Bruno, J., Duro, L. and Grive, M. 2002. The applicability and limitations of Thermodynamic geochemical models to simulate trace element behaviour in natural waters. Lessons learned from natural analogue studies. *Chemical Geology*, 190, 371-393.

Grivé, M., Duro, L., Colàs, E. and Giffaut, E. 2015. Thermodynamic data selection applied to radionuclides and chemotoxic elements: An overview of the ThermoChimie-TDB. *Applied Geochemistry* 55, 85-94.

ThermoChimie guidelines 4. Data traceability

**Mireia Grivé, Isaac Campos,
Elisenda Colàs, Jane Perrone**

July 2015

Written by:	Revised by:	Validated by:
 Isaac Campos Elisenda Colàs	 Jane Perrone	 Mireia Grivé

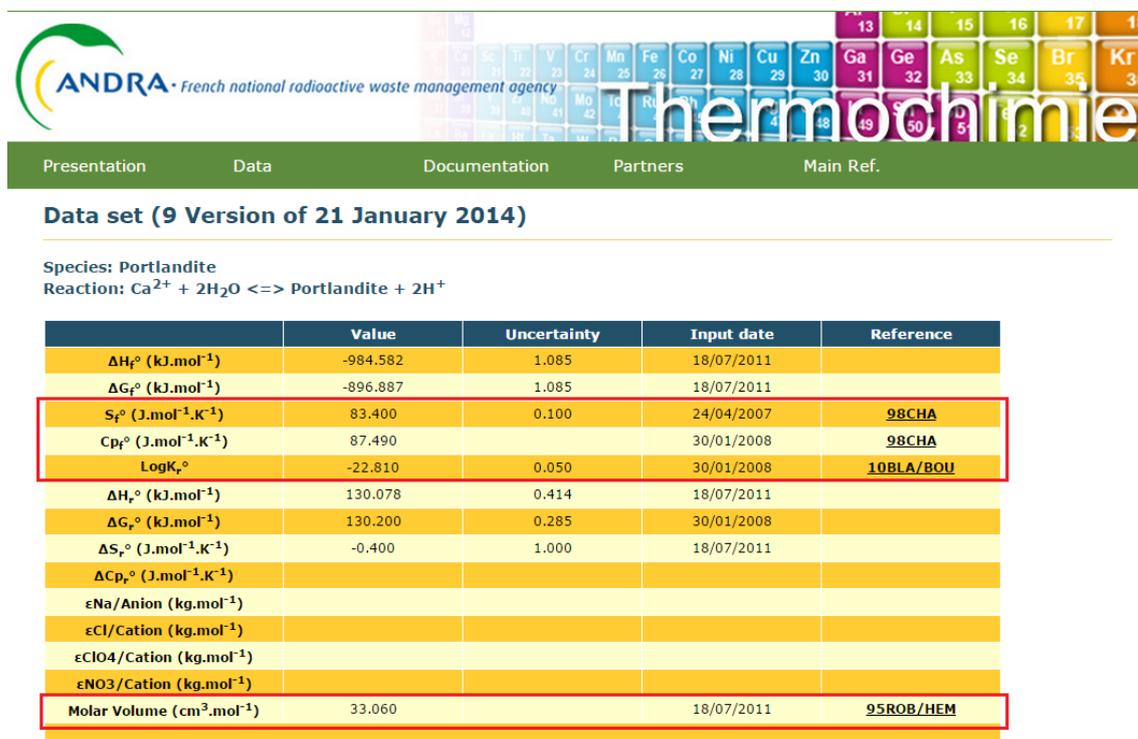
Index

1. INTRODUCTION	2
2. TRACK-CHANGES DOCUMENT	3
3. TRACK-ERRORS DOCUMENT	4

1. Introduction

Traceability is one of the main requirements in the ThermoChimie (TC) database. A database is traceable whether it is possible to verify the history of the values included in it, keeping an appropriate track of the changes made.

Not only selected data must be traceable to the original source (identifying each thermodynamic parameter with its original source, see Figure 1), but also the calculations used to obtain the data must also be traceable (making the appropriate calculations in traceable excel sheet, included in Guideline 2: Data integration and consistency).



ANDRA · French national radioactive waste management agency

ThermoChimie

Presentation Data Documentation Partners Main Ref.

Data set (9 Version of 21 January 2014)

Species: Portlandite
Reaction: $\text{Ca}^{2+} + 2\text{H}_2\text{O} \rightleftharpoons \text{Portlandite} + 2\text{H}^+$

	Value	Uncertainty	Input date	Reference
ΔH_f° (kJ.mol ⁻¹)	-984.582	1.085	18/07/2011	
ΔG_f° (kJ.mol ⁻¹)	-896.887	1.085	18/07/2011	
S_f° (J.mol ⁻¹ .K ⁻¹)	83.400	0.100	24/04/2007	98CHA
$C_{p_f}^\circ$ (J.mol ⁻¹ .K ⁻¹)	87.490		30/01/2008	98CHA
LogK _p ^o	-22.810	0.050	30/01/2008	10BLA/BOU
ΔH_r° (kJ.mol ⁻¹)	130.078	0.414	18/07/2011	
ΔG_r° (kJ.mol ⁻¹)	130.200	0.285	30/01/2008	
ΔS_r° (J.mol ⁻¹ .K ⁻¹)	-0.400	1.000	18/07/2011	
$\Delta C_{p_r}^\circ$ (J.mol ⁻¹ .K ⁻¹)				
$\epsilon_{\text{Na/Anion}}$ (kg.mol ⁻¹)				
$\epsilon_{\text{Cl/Cation}}$ (kg.mol ⁻¹)				
$\epsilon_{\text{ClO}_4/\text{Cation}}$ (kg.mol ⁻¹)				
$\epsilon_{\text{NO}_3/\text{Cation}}$ (kg.mol ⁻¹)				
Molar Volume (cm ³ .mol ⁻¹)	33.060		18/07/2011	95ROB/HEM

Figure 1. Traceability of selected data in ThermoChimie database with its corresponding original source (in red square)

Likewise, every single change, update and/or correction made during the development of the ThermoChimie database must be registered to ensure the

traceability and thus, allow the identification of the modifications among the different versions of the database.

The purpose of the present guideline is to provide the indications necessary to elaborate the document describing the evolution of the database (***Track-changes document***) and the document describing the corrections carried out between different versions of ThermoChimie (***Track-error document***). Notice that, when applicable, both documents can be integrated in a single document.

The elaboration of these documents will be responsibility of the ThermoChimie data base management who, after each update of the database (and after collecting and evaluating the feedback from the ThermoChimie project contributors), will be in charge of their development and distribution. The responsibility of the different project contributors is to provide all the necessary related information to ThermoChimie database management.

2. Track-changes document

The update and use of the ThermoChimie database needs the consultation, evaluation and application of studies continuously produced within the scientific community. This information is continuously checked in order evaluate it and incorporate it into the database if judged necessary.

Thus, the ThermoChimie database is continuously modified/updated in order to ensure that it is up to date. All the changes done in the database must be registered in a document (the so called ***track-changes document***) that summarizes the differences between the different versions of the database. The document must allow the ThermoChimie end users to understand the differences found in results when different versions of ThermoChimie are used in geochemical calculations.

Information related with the following issues must be summarized in a ***brief and practical way*** in the track-changes document (when applicable):

- **Added systems:** New systems (e.g., a new element, a new organic ligand) added in the database must be documented.

-
- **Updated systems:** Updated thermodynamic data for families or groups of elements (e.g. cementitious system, organics, lanthanides...) already included in the database, must be documented.
 - **Other changes:** Other changes in the database not defined in the previous categories as: elimination of species, changes in the name of an existing compound, new complexes introduced in an existent system, inclusion of SIT coefficients,.... must be documented.

The Track-changes document must clearly and accurately define the versions that are being compared, and the date of the modifications. In this case, the reference of the new version is quoted in a second level number (e.g. version 9.b.3, 'b' being corresponding to changes reported in the track-change document).

3. Track-errors document

Once the release in the web site of a new version of ThermoChimie is done, a continuous revision of the database is performed in order to check for possible errors. Feedback and comments from final users and from the different groups involved in the development of ThermoChimie are then taken into account. Feedback and comments are registered, evaluated and, when appropriate, accepted and included in a new version of ThermoChimie. In this case, the reference of the new version is quoted in a third level number (e.g. version 9.b.3, '3' being corresponding to error correction and related track-error document).

The identified errors and the solutions applied in the further version are registered in a **track-error document**. Once those errors have been corrected following the track-error document, a new release of the database is ready. Examples of typical errors that must be documented can be:

- Incorrect values of thermodynamic data.
- Incorrect notation of aqueous and solid species.
- Incorrect equations.

ThermoChimie guidelines 5. Bibliographic references recording process

Project ANDRA TC1

**Mireia Grivé, Isaac Campos,
Elisenda Colàs, Jane Perrone**

July 2015

Written by:	Revised by:	Validated by:
 Isaac Campos Elisenda Colàs	 Jane Perrone	 Mireia Grivé

Index

1.	INTRODUCTION	1
2.	THE ENDNOTE MANAGER SYSTEM.....	2
3.	BIBLIOGRAPHIC DATABASE	3
4.	CITED LITERATURE REFERENCES IN REPORTS.....	3
4.1	REFERENCE LIST.....	4
4.2	IN-TEXT CITATIONS	6

1. Introduction

The update and use of the ThermoChimie database (<http://www.thermochimie-tdb.com/>) requires the consultation, evaluation and application of studies continuously produced within the scientific community. This process implies at the same time, the consultation of scientific publications where the results of those studies are described and/or evaluated and the production of new technical reports.

The management of this information generated within the ThermoChimie project is hence important and needs an organization and systematization of the whole bibliographic information related with the database, in order to keep the traceability of the process.

In order to handle all this bibliographic documentation and help with the storage of bibliographic references in a useful and comprehensible way, a **bibliography reference manager** system is defined and implemented.

The overall objective of this guideline is to define the way of providing a database that contains a complete list of bibliographic references (books, articles in journals, proceedings, technical reports, etc.) reviewed for the development of ThermoChimie. The document will also provide instructions on how the different types of publications should be referenced in the project reports.

Bibliographic database will include:

- References available in the open literature (for example articles in peer-reviewed journals, relevant technical documents available to the public on the web...)
- References not available in the open literature. This is specifically the case of results coming from the acquisition program supported by ThermoChimie, and those documents will be quoted as technical reports not available in open literature. It should be noticed that the general policy of ThermoChimie development aims to provide this information in the form of peer-reviewed

journals; but this cannot be done in a systematic way due to time constraints and delays in publication procedures.

The construction of a bibliographic database will increase the traceability of ThermoChimie.

2. The EndNote manager system

The **EndNote** software (<http://endnote.com/>) has been defined as reference manager system for the ThermoChimie project. The main function of EndNote is to store, manage and search references in a personal library. In addition, it allows exporting the reference library in various file formats.

The use of EndNote to manage the bibliographic references in ThermoChimie has several advantages:

- The information is easy to manage in order to keep traceability of the overall documents produced and referenced within the process of updating the ThermoChimie database.
- Bibliographic references can be inserted while writing a document.
- It allows online searching to retrieve references from online databases.
- It offers the possibility to read and annotate attached PDFs.
- Local storage is unlimited.

Other references manager systems can be used if they allow exporting files to EndNote. For example:

- **JabRef**: a free open source bibliography reference manager (<http://jabref.sourceforge.net/>).
- **Mendeley**: a free reference manager and PDF organizer (<https://www.mendeley.com/>)

3. Bibliographic database

Each contractor participating in the ThermoChimie project has to provide to the database administrators a bibliographic database containing the overall literature references cited in their work. These files will be managed by the database administrators, in order to create a bibliographic list that will increase ThermoChimie traceability.

Depending on the reference manager system used by the contractor, the bibliographic files can be provided in different formats:

- a) **EndNote:** A full EndNote library (*.enl file*) can be provided if EndNote is used by the contractor.
- b) **Other reference manager system:** It is possible to generate list of references compatible with EndNote using other bibliographic systems such as JabRef, Mendeley or Microsoft Word.
 - Jabref can generate import Text files (*.txt files*) compatible with EndNote
 - Mendeley can generate Reference Manager files (*.ris files*) compatible with EndNote.
 - The bibliographical sources manager of Microsoft word can generate references files (*.XML files*) compatible with EndNote.
- c) **No reference manager system:** If the contractor does not use any reference manager compatible with EndNote, a full list of bibliographic references in Harvard citation style (as defined in the next section) can be provided in a Word document.

4. Cited literature references in reports

Bibliographic references given in the reports must contain sufficient information to identify and trace the original sources of the thermodynamic data in ThermoChimie,

and it is advisable that this information is presented in a consistent and standardized manner throughout the project reports.

A reference file under EndNote compatible format (as described above), including all the reference cited in the text, must be provided to database administrators with each new technical report generated by contractors.

4.1 Reference list

The references cited in the text of a research paper have to be systematically presented at the end of the document in a reference list or bibliography.

It is required to follow the **Harvard citation style**. Other citation styles can be used if enough information to identify and trace the original sources are provided.

Authors should be reported with the last name first, followed a comma and by first names initials.

The references should be arranged in alphabetical order by authors' last names. For multiple articles by the same author, or authors listed in the same order, the references should be listed in chronological order, from the earliest to the most recent.

Author names should be separated by commas, while the last author name should be preceded by the symbol "&" or the word "and" when listing multiple authors of a single work.

Examples of formatting of the references are presented below for different types of publications.

- **Articles in journals, magazines, and newspapers**

A complete reference to an article from a journal/periodical should include the following information: author(s), date of publication, article title, journal title, volume number, issue number (if applicable), and page numbers, as shown in the example below.

Giffaut, E., Grivé, M., Blanc, P., Viellard, P., Colàs, E., Gailhanou, H., Gaboreau, S., Marty, N., Madé, B. & Duro, L. 2014. Andra Thermodynamic database for performance assessment: ThermoChimie. *Applied Geochemistry*, 49, 225-236.

- **Books and InBooks**

Reference to an entire book or to a book chapter must include the following elements: author(s) or editor(s), date of publication, title, and the name of the publisher, as shown in the example below.

Grenthe, I., Puigdomenech, I. & Allard, B. 1997. *Modelling In Aquatic Chemistry*, Nuclear Energy Agency (NEA).

Blanc, P., Vieillard, P., Gailhanou, H., Gaboreau, S. 2013. Chapter 6 - Thermodynamics of Clay Minerals, in: *Developments in Clay Science*. Bergaya, F. and Lagaly, G. (Ed.), Elsevier, pp. 173-210.

- **Technical reports**

Reference to a report should include the following elements: author(s), date of publication, title and name of publisher, as shown in the example below. If the issuing organization assigned a number (e.g., report number, contract number, or monograph number) to the report, it should be given in parentheses immediately after the title.

ANDRA. 2009. Inventaire national des matières et déchets radioactifs: Rapport de synthèse. The National Radioactive Waste Management Agency (ANDRA).

- **PhD thesis**

The information needed to reference a thesis includes the following elements: author of the thesis, date of publication, title of the thesis, type of thesis (eg PhD, MSc) and university where the thesis was completed, as shown in the example below.

Colàs, E. 2014. Complexation of Th(IV) and U(VI) by polyhydroxy and polyamino carboxylic acids. PhD Thesis, Universitat Politècnica de Catalunya, Barcelona.

- **Entire Web site**

When citing an entire Web site, no reference list entry is required if the address for the site is cited in the text of the document. When possible, the date of the consultation should be included.

<https://www.thermochimie-tdb.com/> (24 June 2015)

4.2 In-Text Citations

Citations of sources are placed within the text of the paper to briefly identify the sources and enable the reader to locate the source of the cited information in the reference list.

In general, the following rules should be followed:

- In a work with two authors, both authors should be cited in the signal phrase or in the parentheses, followed by the year of publication. The symbol “&” or the word “and” should be used between the authors:

Parkhurst and Appelo (2013) or (Parkhurst and Appelo, 2013)

- In a work with more than two authors, only the use of the first author’s last name followed by “et.al “ and the year of publication, should be used in the signal phrase or in parentheses:

Guiffaut et al. (2014) explained.... or (Guiffaut et al., 2014).

- In the case of two cites with the same author and year, they must be differenced by adding correlative alphabetic letters.

• (Guiffaut et al. 2014a).... or (Guiffaut et al., 2014b).

- If the author is an organization or a government agency, mention the organization in the signal phrase or in the parentheses:

According to the Nuclear Energy Agency (2013)..... or (Nuclear Energy Agency, 2013)

ThermoChimie guidelines. 6. Data quality

Project ANDRA TC1

**Mireia Grivé, Isaac Campos,
Elisenda Colàs, Jane Perrone**

July 2015

Written by:	Revised by:	Validated by:
 Isaac Campos Elisenda Colàs	 Jane Perrone	 Mireia Grivé

AMPHOS²¹



Index

1. INTRODUCTION	2
2. QUALITY CRITERIA	2
2.1 DATA SOURCE.....	3
2.2 DATA QUALITY LEVEL	4
3. EXAMPLE	5
4. BIBLIOGRAPHY	5
ANNEX A: EXCEL SHEET TEMPLATE	6

1. Introduction

The reliability of the results of safety assessments of radioactive waste disposal systems depends on the quality of the thermodynamic database used in calculations. Data quality is thus a critical issue that must be handled as an inseparable part of the database management.

This guideline describes the methodology followed for the quality assessment of the data contained in the ThermoChimie database. This methodology has not been applied in the existing versions of the database (versions 7c, 8 and 9 of ThermoChimie available in the web site <https://www.thermochimie-tdb.com/>), but it will be applied to future ThermoChimie releases. The data already present in the database will be revised and quality labels will be progressively assigned during the continuous development of ThermoChimie.

At the present moment, special efforts are being done in order to fill remaining gaps in the uncertainties of some stability constants and reaction enthalpies available in the database. The assignment of quality labels to ThermoChimie data is strongly related to the uncertainty assignment. Then, quality labels will only be included after this improvement task is complete.

2. Quality criteria

Each individual datum selected in ThermoChimie is the result of an exhaustive work of literature research, data intercomparison, estimation of thermodynamic data and, in specific cases, development of a focused experimental programme to fill in important data gaps, in accordance with the indications provided in ThermoChimie Guideline 1: Data selection and uncertainties.

All the data included in the ThermoChimie database will be classified in different levels as is described below, and a series of **quality labels** will be assigned to each datum in order to define classes, source and quality of different datum.

2.1 Data source

Data are classified taking into account the scientific and technical origin of the data source. The corresponding quality label is defined by a **capital letter** (A, B, C, D, E).

A. By definition: These are convention fixed values. CODATA key values (Cox et al. 1989) are included in this category. ThermoChimie relies on CODATA recommendations for the primary master species and the reference states of the elements (For more information see Guideline 1: Data selection and uncertainties).

E.g.: $\Delta_f G_m^0$ is equal to 0 for species or compounds such as H^+ or $Ca(s)$.

B. Experimental data: Concerns data obtained from experimental sources. This category includes:

- the original experimental data from the literature that are directly selected;
- experimental data from literature that are merged or averaged in order to obtain a final value;
- data that have to be treated or corrected before being included in the database.

E.g: stability constants available in the literature are normally obtained at $I \neq 0$. Stability constants from different experimental sources at different I values are treated in order to obtain a single corrected value at $I=0$. The corrected stability constant will be labelled as "B: experimental data".

C. Estimated values: Concerns values obtained using procedures with a sound scientific basis for making estimates of otherwise unavailable data.

E.g: data obtained using analogies between elements (e.g. Am and Cm) or log K values obtained using ionic-radii correlations will be labelled as "C: estimated values".

E.g: clay data obtained using predictive models (see Guideline 1: Data selection and uncertainties) will be labelled as "C: estimated values".

D. Internally calculated data: When a value has been selected for two of these three variables ($\Delta_r G_m^0$, $\Delta_r H_m^0$ and S_m^0), the rest of the data are internally calculated using the following equations (Eq. 1, Eq. 2) (For more details see Guideline 2: Data integration and consistency). Data obtained using this internal calculation procedures will be labelled as “D: internally calculated data”.

$$\Delta_f G_m^0 = \Delta_f H_m^0 - T \left(S_m^0 + \left(\frac{Z}{2} \right) x S_{m,H2(g)}^0 - \sum_i \frac{\nu_i}{\nu_{i,ref}} S_{m,i}^0 \right) \quad \text{Eq. 1}$$

$$\Delta_r G_m^0 = \Delta_r H_m^0 - T \Delta_r S_m^0 \quad \text{Eq. 2}$$

E. Data class not defined. Data not included in the categories described above. In this case, one explanation of the origin of the data should be provided in the “comments” section of the excel sheet (see the excel sheet template in Annex A).

2.2 Data quality level

Data are also classified on the basis of the assessment of the confidence (reliability) of the values selected. The corresponding quality label is defined by an **ordinal number** (1, 2, 3, 4, 5).

1. Reliable data: Well-established data. No doubts exist on the existence of the species and there is an agreement within different sources for its thermodynamic properties and formation conditions.

2. Reliable data with a wide uncertainty: Reliable but inaccurate data, that is, data with a large degree of associated uncertainty.

E.g., data for which a very limited number of experimental values are available will be labelled as “2: reliable data with a wide uncertainty”.

3. Questionable data: Data that are not reliable enough, but that are necessary for the appropriate description of the system and that must be included in the database in

order to fulfill the completeness and applicability criteria of ThermoChimie. Further efforts in order to improve the quality of these data should be done.

4. Missing but needed data: This quality label is useful in order to identify where efforts for ThermoChimie development should be focused on. It is applied to data that are not available in the database but are necessary for the appropriate description of the system.

Missing but needed data will be identified in the context of validation and verification exercises (see Guideline 3: Validation and verification). *E.g., enthalpy of reaction data missing in ThermoChimie database that are relevant to perform adequate temperature corrections will be labelled as “4: missing but needed data”.*

5. Data quality not defined: Data not included in the categories described above. In this case, an explanation of the reason to assign this category to the data should be provided in the “comments” section of the excel sheet (see the excel sheet template in Annex A).

3. Example

ANNEXE A shows an example to explain how the quality labels are included in the Excel sheet template. Further details on the excel sheet template are provided in Guideline 2: Data integration and consistency.

4. Bibliography

Cox, J.D., Wagman, D.D. and Medvedev, M.A. 1989. CODATA Key Values for Thermodynamics. New York Hemisphere Publishing Corporation.

Annex A: Excel Sheet template

PART A: Basic components:

Specie	UO ₂ (OH) ⁺
Species	UO ₂ (OH) ⁺
coef UO ₂ ⁺²	1
coef e-	
coef H ₂ O	1
coef H ⁺	-1
Elt. 1	UO ₂ ⁺²
Coef. 1	1
dGf	-952.551
idGf	1.747
dHf	-1019.000
idHf	1.500
Sf	-98.2
iSf	3.0
Elt. 2	e-
Coef. 2	0
dGf	0.000
idGf	0.000
dHf	0.000
idHf	0.000
Sf	65.34
iSf	0.001
Elt. 3	H ₂ O
Coef. 3	1
dGf	-237.140
idGf	0.041
dHf	-285.830
idHf	0.040
Sf	70.0
iSf	0.0
Elt. 4	H ⁺
Coef. 4	-1
dGf	0.000
idGf	0.000
dHf	0.000
idHf	0.000
Sf	0.0
iSf	0.0
Elt. 5	
Coef. 5	

Specie	Calcite
Species	CaCO ₃
coef Ca ⁺²	1
coef CO ₃ ⁻²	1
coef H ₂ O	
coef H ⁺	
Elt. 1	Ca ⁺²
Coef. 1	1
dGf	-552.806
idGf	1.044
dHf	-543.000
idHf	1.000
Sf	-56.2
iSf	1.0
Elt. 2	CO ₃ ⁻²
Coef. 2	1
dGf	-527.900
idGf	0.390
dHf	-675.230
idHf	0.250
Sf	-50.0
iSf	1.0
Elt. 3	H ₂ O
Coef. 3	0
dGf	-237.140
idGf	0.041
dHf	-285.830
idHf	0.040
Sf	70.0
iSf	0.0
Elt. 4	H ⁺
Coef. 4	0
dGf	0.000
idGf	0.000
dHf	0.000
idHf	0.000
Sf	0.0
iSf	0.0
Elt. 5	
Coef. 5	

PART B: Selected species:

Specie	UO ₂ (OH) ⁺
Species	UO ₂ (OH) ⁺
Log K ^o	-5.25
± Log K ^o	0.24
Δ _r G ^o	29.967
± Δ _r G ^o	1.370
Δ _r H ^o	43.458
± Δ _r H ^o	14.997
Δ _r S ^o	45.3
± Δ _r S ^o	50.1
Δ _r G ^o	-1159.724
± Δ _r G ^o	2.220
Δ _r H ^o	-1261.372
± Δ _r H ^o	15.072
S ^o	17.00
± S ^o	50.00
C _p	
± C _p	
V _m	
± V _m	
ε _{M+/O-}	-0.003
iε _{M+/O-}	0.231
ε _{M+/OO4-}	-0.060
iε _{M+/OO4-}	0.400
ε _{M+/NO3-}	0.510
iε _{M+/NO3-}	1.400
comment	
Ref. Log K ^o / Δ _r G ^o / Δ _r H ^o	03GUI/FAN
Ref. Δ _r H ^o / Δ _r H ^o	
Ref. Δ _r S ^o / S ^o	92GRE/FUG
Ref. C _p	
Ref. V _m	
Ref. ε Cl ⁻	a)
Ref. ε ClO ₄ ⁻	92GRE/FUG
Ref. ε NO ₃ ⁻	92GRE/FUG
Quality label Log K ^o	B1
Quality label Δ _r G ^o	D1
Quality label Δ _r H ^o	D1
Quality label Δ _r S ^o	D1
Quality label Δ _r G ^o	D1
Quality label Δ _r H ^o	D1
Quality label S ^o	B 1
Quality label C _p	
Quality label V _m	
Quality label ε Cl ⁻	B1
Quality label ε ClO ₄ ⁻	C1
Quality label ε NO ₃ ⁻	B1
Comments	
a) Calculated according to the value with ClO ₄ ⁻ recommended by 95SL/BID	

Specie	Calcite
Species	CaCO ₃
Log K ^o	8.48
± Log K ^o	0.02
Δ _r G ^o	-48.404
± Δ _r G ^o	0.114
Δ _r H ^o	10.620
± Δ _r H ^o	
Δ _r S ^o	198.0
± Δ _r S ^o	
Δ _r G ^o	-1129.110
± Δ _r G ^o	1.120
Δ _r H ^o	-1207.610
± Δ _r H ^o	
S ^o	91.77
± S ^o	
C _p	83.45
± C _p	
V _m	36.93
± V _m	
ε _{M+/O-}	
iε _{M+/O-}	
ε _{M+/OO4-}	
iε _{M+/OO4-}	
ε _{M+/NO3-}	
iε _{M+/NO3-}	
comment	
Ref. Log K ^o / Δ _r G ^o / Δ _r H ^o	82PLU/BUS
Ref. Δ _r H ^o / Δ _r H ^o	82PLU/BUS
Ref. Δ _r S ^o / S ^o	
Ref. C _p	95ROB/HEM
Ref. V _m	10BLA/BOU
Ref. ε Cl ⁻	
Ref. ε ClO ₄ ⁻	
Ref. ε NO ₃ ⁻	
Quality label Log K ^o	B1
Quality label Δ _r G ^o	D1
Quality label Δ _r H ^o	D1
Quality label Δ _r S ^o	D1
Quality label Δ _r G ^o	D1
Quality label Δ _r H ^o	D1
Quality label S ^o	D1
Quality label C _p	B1
Quality label V _m	B1
Quality label ε Cl ⁻	
Quality label ε ClO ₄ ⁻	
Quality label ε NO ₃ ⁻	
Comments	