# Thermodynamic Database and System Settings

## Description

Run `data(thermo)` to initialize or reset all of the data used in CHNOSZ. This includes the computational settings, thermodynamic database, and system settings (chemical species).

The system settings are changed using `basis` and `species`. To restore the default system settings (no species loaded), run `basis("")`.

The thermodynamic database is changed using `add.obigt` and `mod.obigt`. To restore the default database, run `data(OBIGT)`.

The computational settings are changed using `water`, `P.units`, `T.units`, `E.units`, and some other commands (e.g. `mod.buffer`).

Some settings can only be changed by direct manipulation of `thermo$opt`. In an interactive session, this should be done using the super-assignment operator (e.g. `thermo$opt$Berman <<- TRUE`) so that the `thermo` object is not copied to the global environment. (Doing so would cause problems, as many functions are designed to access the `thermo` object in the `CHNOSZ` environment.)

To restore the default computational settings, thermodynamic database, and system settings, run `data(thermo)`.

The data files provided with CHNOSZ are in the `data` and `extdata/OBIGT` directories of the package. The `*.csv` files in these directories are used to build the `thermo` data object in an environment named `CHNOSZ`. The structure of the `thermo` object is described below.

## Usage

```
data(thermo)
data(OBIGT)
```

## Format

- `thermo$opt` List of computational settings. Square brackets indicate default values.

| | | |
|---|---|---|
| `cutoff` | numeric | Cutoff below which values are taken to be zero [1e-10] (see `makeup`) |
| `E.units` | character | The user's units of energy ([`cal`] or `J`) (see `subcrt`) |
| `T.units` | character | The user's units of temperature ([`C`] or `K`) |
| `P.units` | character | The user's units of pressure ([`bar`] or `MPa`) |
| `state` | character | The default physical state for searching species [`aq`] (see `info`) |

| | | |
|---|---|---|
| water | character | Computational option for properties of water ([SUPCRT] or IAPWS; see water) |
| G.tol | numeric | Difference in G above which checkGHS produces a message (cal $mol^{-1}$) [100] |
| Cp.tol | numeric | Difference in Cp above which checkEOS produces a message (cal $K^{-1}$ $mol^{-1}$) [1] |
| V.tol | numeric | Difference in V above which checkEOS produces a message ($cm^3$ $mol^{-1}$) [1] |
| varP | logical | Use variable-pressure standard state for gases? [FALSE] (see subcrt) |
| IAPWS.sat | character | State of water for saturation properties [liquid] (see util.water) |
| paramin | integer | Minimum number of calculations to launch parallel processes [1000] (see palply) |
| ideal.H | logical | Should nonideal ignore the proton? [TRUE] |
| ideal.e | logical | Should nonideal ignore the electron? [TRUE] |
| nonideal | character | Method for nonideal [Helgeson] |
| Berman | logical | Should info preferentially return Berman minerals? [FALSE] |

- thermo$element Dataframe containing the thermodynamic properties of elements taken from Cox et al., 1989 and Wagman et al., 1982. The standard molal entropy (*S*(Z)) at 25 °C and 1 bar for the "element" of charge (Z) was calculated from *S*(H2,g) + 2*S*(Z) = 2*S*(H+), where the standard molal entropies of H2,g and H+ were taken from Cox et al., 1989. The mass of Z is taken to be zero. Accessing this data frame using mass or entropy will select the first entry found for a given element; i.e., values from Wagman et al., 1982 will only be retrieved if the properties of the element are not found from Cox et al., 1989.

| | | |
|---|---|---|
| element | character | Symbol of element |
| state | character | Stable state of element at 25 °C and 1 bar |
| source | character | Source of data |
| mass | numeric | Mass of element (in natural isotopic distribution; referenced to a mass of 12 for *12*C) |
| s | numeric | Entropy of the compound of the element in its stable state at 25 °C and 1 bar (cal K^-*1* mol^-*1*) |
| n | numeric | Number of atoms of the element in its stable compound at 25 °C and 1 bar |

- thermo$obigt

  This dataframe is a thermodynamic database of standard molal thermodynamic properties and equations of state parameters of species. Note the following database conventions:

  - The combination of name and state defines a species in thermo$obigt. A

species can not be duplicated (this is checked when running `data(thermo)`).

- English names of gases are used only for the gas state. The dissolved species is named with the chemical formula. Therefore, `info("oxygen")` refers to the gas, and `info("O2")` refers to the aqueous species.

`OrganoBioGeoTherm` is the name of a GUI program to use SUPCRT in Windows, produced in Harold C. Helgeson's Laboratory of Theoretical Geochemistry and Biogeochemistry at the University of California, Berkeley. The OBIGT database was originally developed for that program, and has been ported to CHNOSZ, with additional modifications. There may be an additional meaning for the acronym: “One BIG Table” of thermodynamic data.

Each entry is referenced to one or two literature sources listed in `thermo$refs`. Use `thermo.refs` to look up the citation information for the references. OBIGT was initially built from the SUPCRT92 (Johnson et al., 1992) and slop98.dat data files (Shock et al., 1998), and the references in those files are included here. Some data in slop98.dat were corrected or modified as noted in that file; these modifications are indicated in OBIGT by having `SLOP98` as one of the sources of data. Other additions or modifications used in CHNOSZ are indicated by having `CHNOSZ` as one of the sources of data. See the vignette *Thermodynamic data in CHNOSZ* for a complete description of the sources of data.

In order to represent thermodynamic data for minerals with phase transitions, the higher-temperature phases of these minerals are represented as phase species that have states denoted by `cr2`, `cr3`, etc. The standard molar thermodynamic properties at 25 °C and 1 bar (*Pr* and *Pr*) of the `cr2` phase species of minerals were generated by first calculating those of the `cr` (lowest-T) phase species at the transition temperature (*Ttr*) and 1 bar then taking account of the volume and entropy of transition (the latter can be retrieved by combining the former with the Clausius-Clapeyron equation and values of *(dP/dT)* of transitions taken from the SUPCRT92 data file) to calculate the standard molar entropy of the `cr2` phase species at *Ttr*, and taking account of the enthalpy of transition (*DeltaH0*, taken from the SUPCRT92 data file) to calculate the standard molar enthalpy of the `cr2` phase species at *Ttr*. The standard molar properties of the `cr2` phase species at *Ttr* and 1 bar calculated in this manner were combined with the equations-of-state parameters of the species to generate values of the standard molar properties at 25 °C and 1 bar. This process was repeated as necessary to generate the standard molar properties of phase species represented by `cr3` and `cr4`, referencing at each iteration the previously calculated values of the standard molar properties of the lower-temperature phase species (i.e., `cr2` and `cr3`). A consequence of tabulating the standard molar thermodynamic properties of the phase species is that the values of *(dP/dT)* and *DeltaH0* of phase transitions can be calculated using the equations of state and therefore do not need to be stored in the thermodynamic database. However, the transition temperatures (*Ttr*) generally can not be assessed by comparing the Gibbs energies of phase species and are tabulated in the database.

The identification of species and their standard molal thermodynamic properties at 25 °C and 1 bar are located in the first 12 columns of `thermo$obigt`:

| | |
|---|---|
| `name` | character Species name |
| `abbrv` | character Species abbreviation |

| | | |
|---|---|---|
| `formula` | character | Species formula |
| `state` | character | Physical state |
| `ref1` | character | Primary source |
| `ref2` | character | Secondary source |
| `date` | character | Date of data entry (formatted as in SUPCRT92) |
| `G` | numeric | Standard molal Gibbs energy of formation from the elements (cal mol^-1) |
| `H` | numeric | Standard molal enthalpy of formation from the elements (cal mol^-1) |
| `S` | numeric | Standard molal entropy (cal mol^-1 K^-1) |
| `Cp` | numeric | Standard molal isobaric heat capacity (cal mol^-1 K^-1) |
| `V` | numeric | Standard molal volume (cm^3 mol^-1) |

The meanings of the remaining columns depend on the physical state of a particular species. If it is aqueous, the values in these columns represent parameters in the revised HKF equations of state (see `hkf`), otherwise they denote parameters in a general equations for crystalline, gas and liquid species (see `cgl`). The names of these columns are compounded from those of the parameters in each of the equations of state (for example, column 13 is named `a1.a`). Scaling of the values by orders of magnitude is adopted for some of the parameters, following common usage in the literature.

Columns 13-20 for aqueous species (parameters in the revised HKF equations of state):

| | | |
|---|---|---|
| `a1` | numeric | *a1 * 10* (cal mol^-1 bar^-1) |
| `a2` | numeric | *a2 * 10^{-2}* (cal mol^-1) |
| `a3` | numeric | *a3* (cal K mol^-1 bar^-1) |
| `a4` | numeric | *a4 * 10^-4* (cal mol^-1 K) |
| `c1` | numeric | *c1* (cal mol^-1 K^-1) |
| `c2` | numeric | *c2 * 10^-4* (cal mol^-1 K) |
| `omega` | numeric | *omega * 10^-5* (cal mol^-1) |
| `Z` | numeric | Charge |

Columns 13-20 for crystalline, gas and liquid species (*Cp = a + bT + cT^-2 + dT^-0.5 + eT^2 + fT^lambda*).

| | | |
|---|---|---|
| `a` | numeric | *a* (cal K^-1 mol^-1) |
| `b` | numeric | *b * 10^3* (cal K^-2 mol^-1) |
| `c` | numeric | *c * 10^-5* (cal K mol^-1) |
| `d` | numeric | *d* (cal K^-0.5 mol^-1) |
| `e` | numeric | *e * 10^5* (cal K^-3 mol^-1) |
| `f` | numeric | *f* (cal K-*lambda*-1 mol^-1) |
| `lambda` | numeric | *lambda* (exponent on the *f* term) |
| `T` | numeric | Temperature of phase transition or upper |

temperature limit of validity of extrapolation (K)

- `thermo$refs` Dataframe of references to sources of thermodynamic data.

| | | |
|---|---|---|
| `key` | character | Source key |
| `author` | character | Author(s) |
| `year` | character | Year |
| `citation` | character | Citation (journal title, volume, and article number or pages; or book or report title) |
| `note` | character | Short description of the compounds or species in this data source |
| `URL` | character | URL |

- `thermo$buffers`

  Dataframe which contains definitions of buffers of chemical activity. Each named buffer can be composed of one or more species, which may include any species in the thermodynamic database and/or any protein. The calculations provided by `buffer` do not take into account phase transitions of minerals, so individual phase species of such minerals must be specified in the buffers.

| | | |
|---|---|---|
| `name` | character | Name of buffer |
| `species` | character | Name of species |
| `state` | character | Physical state of species |
| `logact` | numeric | Logarithm of activity (fugacity for gases) |

- `thermo$protein` Data frame of amino acid compositions of selected proteins. Most of the compositions were taken from the SWISS-PROT/UniProt online database (Boeckmann et al., 2003) and the protein and organism names usually follow the conventions adopted there. In some cases different isoforms of proteins are identified using modifications of the protein names; for example, `MOD5.M` and `MOD5.N` proteins of `YEAST` denote the mitochondrial and nuclear isoforms of this protein. See `pinfo` to search this data frame by protein name, and other functions to work with the amino acid compositions.

| | | |
|---|---|---|
| `protein` | character | Identification of protein |
| `organism` | character | Identification of organism |
| `ref` | character | Reference key for source of compositional data |
| `abbrv` | character | Abbreviation or other ID for protein |
| `chains` | numeric | Number of polypeptide chains in the protein |
| `Ala...Tyr` | numeric | Number of each amino acid in the protein |

- `thermo$groups` This is a dataframe with 22 columns for the amino acid sidechain, backbone and protein backbone groups ([Ala]..[Tyr],[AABB],[UPBB]) whose rows correspond to the elements C, H, N, O, S. It is used to quickly calculate the chemical formulas of proteins that are selected using the `iprotein` argument in `affinity`.

- `thermo$basis` Initially `NULL`, reserved for a dataframe written by `basis` upon definition of

the basis species. The number of rows of this dataframe is equal to the number of columns in "..." (one for each element).

| | | |
|---|---|---|
| `...` | numeric | One or more columns of stoichiometric coefficients of elements in the basis species |
| `ispecies` | numeric | Rownumber of basis species in `thermo$obigt` |
| `logact` | numeric | Logarithm of activity or fugacity of basis species |
| `state` | character | Physical state of basis species |

- `thermo$species` Initially `NULL`, reserved for a dataframe generated by `species` to define the species of interest. The number of columns in "..." is equal to the number of basis species (i.e., rows of `thermo$basis`).

| | | |
|---|---|---|
| `...` | numeric | One or more columns of stoichiometric coefficients of basis species in the species of interest |
| `ispecies` | numeric | Rownumber of species in `thermo$obigt` |
| `logact` | numeric | Logarithm of activity or fugacity of species |
| `state` | character | Physical state of species |
| `name` | character | Name of species |

## References


Cox, J. D., Wagman, D. D. and Medvedev, V. A., eds. (1989) *CODATA Key Values for Thermodynamics*. Hemisphere Publishing Corporation, New York, 271 p. http://www.worldcat.org/oclc/18559968

Johnson, J. W., Oelkers, E. H. and Helgeson, H. C. (1992) SUPCRT92: A software package for calculating the standard molal thermodynamic properties of minerals, gases, aqueous species, and reactions from 1 to 5000 bar and 0 to 1000°C. *Comp. Geosci.* **18**, 899–947. https://doi.org/10.1016/0098-3004(92)90029-Q

Shock, E. L. et al. 1998 *SLOP98.dat* (computer data file). http://geopig.asu.edu/supcrt92_data/slop98.dat, accessed on 2005-11-05. Current location: http://geopig.asu.edu/?q=tools.

Wagman, D. D., Evans, W. H., Parker, V. B., Schumm, R. H., Halow, I., Bailey, S. M., Churney, K. L. and Nuttall, R. L. (1982) The NBS tables of chemical thermodynamic properties. Selected values for inorganic and C*1* and C*2* organic substances in SI units. *J. Phys. Chem. Ref. Data* **11** (supp. 2), 1–392. https://srd.nist.gov/JPCRD/jpcrdS2Vol11.pdf


## See Also

Other data files, including those supporting the examples and vignettes, are documented separately at `extdata`.

## Examples

```
## where are the data files in CHNOSZ?
system.file("data", package="CHNOSZ")
system.file("extdata", package="CHNOSZ")

## exploring thermo$obigt
# what physical states there are
unique(thermo$obigt$state)
# formulas of ten random species
n <- nrow(thermo$obigt)
thermo$obigt$formula[runif(10)*n]
```

---