



CH 153K

Finding Physical and Chemical Property Data

Fall 2009

Science cannot happen without data.¹ Data generated by experimentation reinforce, confirm, illustrate, modify or even refute hypotheses and theories and enable the understanding of natural systems. At all stages of the experimental process, the scientist and engineer must be able to locate needed data, assess their accuracy and quality, and use them appropriately.

This is not as easy as it sounds. Searching for data can be time-consuming, frustrating, and sometimes fruitless. Fortunately, the data required for most undergraduate experiments are usually not difficult to find. It's just a question of knowing what sources to consult, and how to use them properly.

When choosing a source of data, **reliability** is a crucial factor. This point will be made again and again below.

1. Published Sources of Chemical Data

1a. Databases

Databases may be electronic versions of printed books (e.g. the CRC Handbook), or they may be digital-native. Some are freely available on the Web (such as the NIST Chemistry WebBook), but most require a subscription to view the content.

1b. Handbooks

“Handbook” is a loose term for a collection of useful information drawn from other sources. A handbook might be a handy one-volume reference tool such as the *CRC Handbook of Chemistry and Physics*. It might be a database, such as DIPPR. It might be a massive multivolume work such as the *Beilstein Handbook of Organic Chemistry* or its online equivalent. It might cover a topic broadly, such as *Lange's Handbook of Chemistry*. Or it might focus on a very narrow topic: *Handbook of Solid-Liquid Equilibria in Systems of Anhydrous Inorganic Salts*. Handbooks come in every size and scope. As secondary literature, handbooks do not report original research. But they save you time by gathering pertinent data from the primary literature and organizing it in a useful format.

1. **Data** is usually a plural noun, the singular form of which is **datum**. (It is a form of the Latin verb *dare*, to give.) When used in a sentence it takes the plural form of verbs and adjectives: “The data presented here suggest that...”; “These data are clearly in error...” When the word refers to a collection of data to be regarded as a whole unit, it can be used in the singular: “The data was distributed to the class.”

1c. Journals

Journals are the primary literature of science. New research is first formally reported in journals. Journals are **peer-reviewed**, meaning that a submitted article is read and critiqued by anonymous reviewers who determine whether it is worthy of publication and what revisions must be made first. Reviewers are looking for internal consistency, valid conclusions, and obvious errors in data analysis, statistical methods, and experimental procedures, but they do not reproduce experiments or analyze the reported data.

Since most journals emphasize brevity, raw data are rarely given. Summary tables and graphs present experimental data in a more digestible form. A few journals are devoted to the reporting of extensive and reliable physical data: the *Journal of Physical and Chemical Reference Data* and the *Journal of Chemical and Engineering Data* are two important examples. Articles in these journals can be long and consist mostly of tables.

1d. Monographs

“Monograph” is just a fancy word for book. In the sciences, books are often collections of chapters written by individual experts and compiled by an editor. Books are by definition secondary literature – they do not report new research, but summarize what is already known. As such, books can provide a good overview of a topic, and include tables of otherwise hard-to-find scientific data drawn from the journal literature. However, books are not peer-reviewed and are only as reliable as their authors make them.

The trouble is that data content in books is difficult to locate unless you already know that the book contains that data; otherwise you’ll probably run across it only by accident. The normal discovery tools for books – the library catalog, WorldCat, Amazon.com, etc. – aren’t of much help in identifying books that might contain specific data. You have to get the book and look inside. The Google Books project (<http://books.google.com/>), which is scanning millions of books from major library collections, offers a new tool to virtually look inside a book without actually holding it in your hand.

1e. Textbooks

This is often the first place a student will look for needed data. Textbooks often include brief tables of commonly used data within chapters or in appendices. But once your need goes beyond the most basic data and compounds, textbooks will not suffice.

2. Where to Start?

Let’s face it, you probably will not have thought of any of these concepts until you’re faced with needing a piece of data right away. Your first reaction will likely be to thumb through your textbook hoping to find it right there. When that fails, you’ll have to confront the disorderly and mysterious universe of scientific information described above, and hope for the best.

What you should do first is ask yourself some questions. The answers will start you off.

- What kind of data?

This will usually be spelled out in your lab or procedure. But beyond the specific property in question, what kind of data is this? Is it thermodynamic, i.e. temperature-dependent? Is it a physical property, a kinetic property, or a chemical property? What units will be used?

- What kind of compound?

This is more important than you might think. Chemical data are usually organized by type of compound or system. Small organic molecules ($C \leq 8$ or so) and simple gases are often well characterized and data for them are relatively easy to locate. Beyond that, all bets are off. Larger more complex molecules, biomolecules, polymers, drugs, commercial compounds and mixtures may have had little if any physical data reported for them. Even common inorganic compounds such as metal salts can be elusive. The same issues apply to multicomponent systems. Binary and ternary systems of well known organic substances aren't too difficult. Anything else is likely to be a challenge.

- What conditions apply?

Most temperature- or pressure-dependent data fall within fairly narrow ranges of conditions. If you need data for unusual conditions, such as very high or low pressures or temperatures, your chance of success drops sharply.

2a. Starting Points

UT Libraries web pages serve as gateways for identifying and linking to the best sources.

- Finding Thermodynamic and Physical Property Data

<http://www.lib.utexas.edu/chem/info/thermo.html>

This guide lists some of the most important reference sources of property data, both on the Web and in print. Critically evaluated sources (the most reliable) are marked with a check-mark icon. The "Where to Look First" section lists the resources where almost every data search should begin.

- Chemistry Library Home Page

<http://www.lib.utexas.edu/chem/>

This is your main gateway to chemical information at UT Austin. The home page contains direct links to some of the most important online tools.

- ThermoDex

<http://www.lib.utexas.edu/thermodex/>

A finding aid for over 300 data compilations, handbooks, and reference sources that contain physical property data needed by chemists. If you need to identify a more specialized handbook when the First Stop tools don't have what you need, use this database.

So what *are* the best sources to start with?

Much of what you'll need as an undergraduate (including all the answers to your exercise questions) can be found in one or more of these four tools. All are available electronically. Links to them can be found on the Chemistry Library's home page.

1	CRC Handbook of Chemistry & Physics	<p>A mix of PDF full text and interactive web tables. You can browse the table-of-contents in the left bar, but it's usually easier to do a free-text search in the search box using property and compound keywords to locate specific tables. Example: solubility benzoic acid</p> <p>Multiple editions are located on the Handbook Table in the Chemistry Library. The print version is published annually. The content changes slightly each year, so when you cite the CRC be sure to indicate the edition and year. Consult the index in the back under the desired property name.</p>
2	DIPPR	<p>Evaluated and checked thermochemical and temperature-dependent data for about 1900 pure organic chemicals. Search by chemical name, Registry number, formula, or by property value.</p>
3	NIST Chemistry WebBook	<p>Highly reliable collection of critically evaluated physical, kinetic, thermodynamic and spectral data for thousands of chemicals, drawn from NIST/NBS publications and databases, and the technical literature. Search by name, formula, Registry number, structure, or selected physical property values.</p>
4	NIST/TRC Web Thermo Tables	<p>Evaluated data on 7800+ organic compounds.</p>

2b. What's in a Name?

When starting a search for property data, you obviously need to be very clear about what compound you're after. Typically you're starting with a name of a compound as given in your lab book, and maybe nothing else. But chemical names can be troublesome. Common names for well known compounds (e.g. benzene, naphthalene, toluene, cyclohexane, benzoic acid, etc.) are not so bad because everyone knows what they mean. But as molecules get more complex, so do the names. Names can be ambiguous, and a given structure might have many different "correct" names: trivial or common names, systematic IUPAC-style names, trade names, abbreviations, and so on. You may need to verify the compound by checking its structure, molecular formula, or CAS Registry Number.

2c. Watch your Units

Pay attention to the units used in a given resource. Be sure to record them along with the data, and convert them when necessary. Record the conditions such as temperature and pressure, as well.

3. Surfing the Web

“Spending an hour on the web can save you five minutes in the library.”

You may be tempted to search for property data out on the open Web, using search engines like Google or online tools like Wikipedia. You should resist this urge. Why?

- The Web is an unstable, amorphous, disordered, and unvetted ocean of information, with a very low signal-to-noise ratio. Using unreliable data in a laboratory setting is irresponsible and potentially dangerous. Using it in a lab report or paper can get you in trouble too.
- Search engines can turn into a huge waste of your time: you might spend an hour sifting through irrelevant Google results, getting sidetracked and clicking random links – hunting for something that you could find in 30 seconds using the *CRC Handbook*. And even then you wouldn't be able to tell if it was reliable or not.
- The “deep web,” where proprietary information lives behind firewalls and subscriptions, is not well indexed by search engines. What you're really looking for is likely invisible to Google.
- You'll have to do a legitimate data search eventually anyway, so you may as well save time and do it right from the beginning.

4. Doing a Literature Search

Secondary reference sources (handbooks, databases, etc.) draw data *selectively* from the primary literature (journals). But only a small fraction of scientific data reported in journals ever makes it into the secondary reference tools. In some cases you'll have to search the full literature to find data or become relatively confident that what you're looking for hasn't been reported. But don't do this until you've exhausted the easier options.

4a. Finding Other Data Compilations

There are hundreds of chemical data compilations that have been published over the years. They contain a wealth of data gathered painstakingly from the primary literature or from direct experimental work. For the most part, they exist only in printed format. But how do you find out about them? They are scattered around the library, and the Library Catalog often is not detailed enough to reveal what's included in them. We have created a Web tool to help: ThermoDex (<http://www.lib.utexas.edu/thermodex/>). This database indexes about 300 compilations according to the types of property data and compounds included.

4b. Finding Articles: SciFinder and Beilstein

Finding physical data in the journal literature is not as easy as it sounds, but thanks to modern indexing database tools it's much faster than it used to be. Searching the literature is beyond the scope of this introduction, but further information about the tools can be found here:

<http://www.lib.utexas.edu/chem/info/thermo-lit.html>

5. When to Ask for Help

Consider the Law of Diminishing Returns, or the Twenty-Minute Rule: If you haven't found what you're looking for after twenty minutes or so of looking in the *best sources*, stop and ask for help. You might have been looking in the wrong places, or making some incorrect assumptions. Further hunting in that direction could well be a waste of your time. Students have two options for help:

- **Your Instructors.** These are the experts in the science behind your labs. They can tell you when you're barking up the wrong tree, and they can help reframe your problem and suggest practical solutions. You may not even need what you think you need – or it might be easily derived from what you already know.
- **Your Librarians.** These are the experts with the literature and published data sources. They know what's out there, where it is, whether it's online or not. After you've exhausted the basics on your own, they can guide you to more specialized tools that you might not know about. When reference sources fail, they can show you how to do a full literature search.

6. Evaluating Information

So you've found a piece of data you're looking for. Now what? Do you just plug it in and go? Is it really the "literature value" you've been told to find? Not necessarily. All data sources are not created equal.

A literature value, as the name implies, is a data point that has been published in the scientific literature. Its accuracy is therefore implied, but should not be assumed absolutely. Some uncertainty is inevitable in any measurement. Variations in data values are usually due to differences in experimental procedures, precision of instruments, impurities in compounds, and human error. Data that fall outside the range of acceptable uncertainty should be used with caution. Some large data sources collect many such values from the literature, compare and assess them, and recommend some and discard others. These are known as "**critically evaluated**" sources. DIPPR is an excellent example. But most sources are not critically evaluated.

The skills you learn as a student scientist will help you examine all resources critically, and accept some pieces of information and discard others. You should get in the habit of asking yourself questions like:

- Where does this information come from?
- Who did it and when?
- Is it consistent with other sources you've found?
- Does it cite its sources?
- Is it safe?

Every piece of information you retrieve should be evaluated in this way. Naturally you're on somewhat safer ground when dealing with journal articles, books and reputable databases, as opposed to web pages. That's not to say that journals and books don't sometimes contain erroneous information. Errors do occur, both in the lab and in the publishing process. When you take data from gray-area web resources, use caution. Sources like MSDS might look authoritative and official, but they can be very unreliable.

7. Citing Your Sources

Whenever you use data or information that you did not generate yourself from your own experimental work, and that is not considered "common knowledge in the field," you must cite your source clearly. If you do not cite a source for a statement or fact in your lab reports or research papers, you are implying that you are the creator of that information. So always give credit where credit is due.

Legitimate sources to cite include:

- Journal articles
- Books and book chapters
- Conference papers
- Patents
- Handbooks
- Encyclopedias and textbooks
- Reputable factual databases (such as NIST WebBook, DIPPR, etc.)

7a. Reference Style

A journal article is routinely cited by its author(s), source, and bibliographic data (volume, year, pages). The article title is usually not given in a reference, to save space. It does not matter whether you obtained the document online or in print. It is not standard practice to provide an article's URL in a reference. URLs are not stable over time.

The goal of the reference is to provide just enough information to allow a reader anywhere to locate the document via whatever paths s/he has at his or her disposal. It is your responsibility to cite sources accurately.

Follow the citation formats to journals and other types of publications as provided in Chapter 14 of the *ACS Style Guide* (3rd edition), which is available on reserve in the Chemistry Library.

Some examples based on ACS style are available at

<http://chemistry.library.wisc.edu/writing/acs-style-guidelines.html>

The actual reference format can vary slightly by journal and publisher, but in chemistry it usually follows this form:

Smith, J.A.; Jones, P.T.; Wang, T. *J. Phys. Chem. Ref. Data* **2004**, 31, 698-710.
 [Authors in published order, last name first, first and middle initials, separated by semicolons. *Journal Title abbreviation** **Year**, *Volume*, pagination.]

* Journal titles are now fully spelled out in most index databases, but are still routinely abbreviated in reference lists. Abbreviations follow certain conventions, but can sometimes be non-intuitive. If you have difficulty creating or deciphering an abbreviation, consult CASSI, which is Chemical Abstracts' cumulative listing of sources indexed since it began in 1907. CASSI is located on the Chemistry Library's circulation desk. A list of 1000 common chemistry journal abbreviations can be found in the *ACS Style Guide*.

When citing other kinds of sources, such as books, conferences, encyclopedias, databases, etc., consult the Style Guide for examples.

If you wish to cite an "unpublished" web page, you must of course include the its URL, but you must be very judicious when doing this. Many editors (and professors) will reject a citation that is not to a "legitimate" published source. You can't guarantee that future readers will be able to find a web page that may disappear or move in the interim.

7b. Copyright

Beyond the requirement of citing your sources properly, you must respect copyright when you wish to quote or reproduce content created by someone else. It is not legal to reproduce tables, graphics, or images from a published or unpublished source without first obtaining permission from the copyright holder. While it is rare in scientific writing to quote text verbatim, you may do so with proper attribution, as long as you don't quote too much.

Further Reading

Perrot, Pierre. *A to Z of thermodynamics*. (Oxford Univ. Press, 1998).
 QC 310.3 P47 1998 Chem Reference

Poling, Bruce E. et al. *The Properties of Gases and Liquids*. 5th ed. (McGraw-Hill, 2001).
 TP 242 P62 2001 Chem Reserves

Smith, J.M., Van Ness, H.C., Abbott, M.M. *Introduction to Chemical Engineering Thermodynamics*. 7th ed. (McGraw-Hill, 2007).
TP 155.2 T45 S58 2007 Chem Reserves

More Questions? Need Help?

Visit the Chemistry Library and ask for assistance: reference help is usually available weekdays before 5pm. Or, you can send email to the Ask a Librarian page at <http://www.lib.utexas.edu/services/reference/> and we will try to respond within 24 hours.

Appendix. Worked Sample Questions

The “first choice” sources used to answer these questions are illustrative and aren’t necessarily the only places where such data could be found. Consider trying several sources sequentially and comparing the data you find. Remember to record units and conditions such as temperature and pressure.

- a. Find the solubility of magnesium iodide (MgI_2) in water at 30° C.
 - First choice: CRC Handbook (print)
 - Index: Look under “Solubility”: note table “Aqueous solubility of inorganic compounds at various temperatures”.
 - Result: 60.8
 - Units: mass percent of solute (from table introduction)

- b. Find the speed of sound in neon gas.
 - First choice: CRC Handbook (Web)
 - Search: “Speed of sound neon”
 - Table: “Speed of Sound in Various Media”
 - Value: 435 m/s
 - Temp: 0° C
 - Pressure: 1 atm

- c. Find the value for the heat of formation (ideal gas) of 1,1,1-trifluoroethane.
 - First choice: DIPPR database
 - Search: Name=trifluoroethane
 - Value: -7.36400×10^8
 - Units: J/kmol
 - Reliability: < 3%

- d. Graph values of the heat of vaporization of toluene by temperature.
 - First choice: DIPPR
 - Search: name=toluene
 - Click “Temperature Dependent” button
 - Select “Heat of vaporization”

- Click graph link
 - Units: J/kmol
- e. Locate Antoine coefficients for calculating the vapor pressure of cyclohexane.
- First choice: NIST WebBook
 - Search: name=cyclohexane
 - Select “Phase change data”
 - Note that you can view a plot of all reported data.
- f. Find values of constant pressure heat capacity C_p for liquid 1-pentanol.
- First choice: NIST WebBook
 - Search: name=pentanol
 - Select “Condensed phase thermochemistry data”
- g. Find solubility data for tantalum pentachloride ($TaCl_5$) in carbon tetrachloride (CCl_4).
- First choices: not found
 - Second choice: ThermoDex
 - Search: “solubility” as property and “inorganic” as compound type
 - Select from results: *Solubilities of inorganic and metal organic compounds*, by Linke and Seidell. Note location in Chemistry Library: QD 543 S4 1958 Chem Reference.
 - Consult printed handbook: Index under $TaCl_5$: refers to vol.2, p.1525 for organic solvents.

© 2009 by The University of Texas at Austin. This material may be quoted or reproduced for educational purposes without prior permission, provided appropriate credit is given. Any commercial use of this material is prohibited without prior permission from The UT Libraries.