Active Thermochemical Tables: thermochemistry for the 21st century

Branko Ruscic,[a] **Reinhardt E. Pinzon,**[a] **Gregor von Laszewski,**[b]
Deepti Kodeboyina,[b] **Alexander Burcat,**[a,c] **David Leahy,**[d]
David Montoya,[e] **and Albert F. Wagner**[a]

[a] Chemistry Division, Argonne National Laboratory, Argonne, IL 60439
[b] Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL 60439
[c] on sabbatical leave from Faculty of Aerospace Engineering, Technion – Israel Institute of Technology, Haifa 32000, Israel
[d] Sandia National Laboratories, Livermore, CA 94551
[e] Los Alamos National Laboratory, Los Alamos, NM 87545

E-mail: ruscic@anl.gov

Abstract. Active Thermochemical Tables (ATcT) are a good example of a significant breakthrough in chemical science that is directly enabled by the U.S. DOE SciDAC initiative. ATcT is a new paradigm of how to obtain accurate, reliable, and internally consistent thermochemistry and overcome the limitations that are intrinsic to the traditional sequential approach to thermochemistry. The availability of high-quality consistent thermochemical values is critical in many areas of chemistry, including the development of realistic predictive models of complex chemical environments such as combustion or the atmosphere, or development and improvement of sophisticated high-fidelity electronic structure computational treatments. As opposed to the traditional sequential evolution of thermochemical values for the chemical species of interest, ATcT utilizes the Thermochemical Network (TN) approach. This approach explicitly exposes the maze of inherent interdependencies normally ignored by the conventional treatment, and allows, *inter alia*, a statistical analysis of the individual measurements that define the TN. The end result is the extraction of the best possible thermochemistry, based on optimal use of all the currently available knowledge, hence making conventional tabulations of thermochemical values obsolete. Moreover, ATcT offer a number of additional features that are neither present nor possible in the traditional approach. With ATcT, new knowledge can be painlessly propagated through all affected thermochemical values. ATcT also allows hypothesis testing and evaluation, as well as discovery of weak links in the TN. The latter provides pointers to new experimental or theoretical determinations that can most efficiently improve the underlying thermochemical body of knowledge.

1. Introduction

Active Thermochemical Tables (ATcT) [1] are a novel approach to thermochemistry, centered on a distinctively different paradigm of how to derive accurate, reliable, and internally consistent thermochemical values, and are rapidly becoming the archetypal approach to thermochemistry for the 21st century [2]. As opposed to the traditional sequential approach to developing thermochemical tables, ATcT uses the Thermochemical Network (TN) approach.

Knowledge of thermochemical stability for a broad range of chemical entities (species), encompassing both stable substances and ephemeral chemical moieties (such as free radicals and reaction intermediates), is central to many aspects of chemistry and critical in a number of industries. While the accuracy requirements of thermochemical values for general chemistry deliberations may often be somewhat relaxed (values need to be reasonable, but not necessarily of ultimate accuracy), the availability of accurate, reliable, and internally consistent thermochemistry is a necessary condition in many fields of physical chemistry, ranging from kinetics, through construction of chemical reaction mechanisms, to formulation of realistic models of complex chemical environments, such as flames or the atmosphere. The availability of highly accurate thermochemical benchmark values is also important for the development and testing of sophisticated electronic structure methods.

Another important but currently underplayed aspect relates to properly quantified uncertainties ("error limits") that accompany the thermochemical values. The uncertainties found in traditional tabulations should approximately correspond to the diagonal elements of the covariance matrix scaled by the appropriate coverage factor to obtain 95% confidence intervals. However, the off-diagonal elements that quantify the cross-correlation of the values for different species are absent. Even the available uncertainties are mostly ignored or used only as a qualitative guide to the level of confidence that can be placed in the listed thermochemical values. However, as the sophistication and fidelity levels of complex computer models are rapidly increasing, proper uncertainty quantification is becoming a pressing issue. As the modeled behavior comes closer to the empirical reality, modelers need to be able to quantitatively separate the degree of uncertainty attributable to the underlying uncertainties in the input parameters (such as thermochemical values) from residual imperfections in the model itself. Given the rapid pace of progress in modeling capabilities, it would be rather surprising not to see in the near future a large wave of systematic implementation of full uncertainty propagation from input parameters to the computed output in many models.

2. Traditional thermochemical tables and the sequential approach to thermochemistry

Traditional thermochemical tables are tabulations of thermochemical properties conveniently sorted by chemical species. Their quality ranges from fully-documented critical data evaluations [3-5] or evaluated compilations containing references but not documenting the reasons for selecting a particular value [6], through lists providing fully referenced multiple values without attempting an evaluation [7], to tabulations that simply list values taken from other compilations [8] (sometimes without clear referencing of the individual listed values). The type of listed thermochemical varies considerably: the enthalpy of formation is virtually always provided, frequently accompanied by the Gibbs energy of formation, heat capacity, entropy, enthalpy increment, etc. The temperatures for which these properties are listed typically include room temperature (298.15 K), and sometimes a selection of other temperatures. Some compilations present the thermochemistry in form of polynomials rather than tables [9, 10]. Critical data evaluations are the most valuable kind of tabulations, and, indeed, the source of recommended new (or least added-value) thermochemical data that is acceptable for scientific use. The fully-documented variants are particularly useful, because, in addition to providing high-quality data, they allow the user to re-examine, if needed, the original process that led to the recommended thermochemical value. Critical data evaluations are slow and labour-intensive undertakings, and, particularly in the case of large compilations (several hundred species or more), they are a substantial collaborative effort of a group of experts that typically extends over a number of years or even decades, and consists of initial data accumulation and evaluation, followed by a manual sequential evolution of the pivotal thermochemical quantity for each species.

2.1. The traditional sequential approach to thermochemistry

The central thermochemical property of any chemical species is not directly measurable. Rather, it is derived from more basic determinations that, for the purpose of this discussion, generally fall into two categories: species-specific and species-interrelating. Some properties, such as heat capacity, $C^o_p(T)$, entropy, $S^o(T)$, enthalpy increment, H_T-H_0 or H_T-H_{298} (a.k.a. integrated heat capacity) can be derived directly from species-specific information, such as spectroscopic determinations (or, alternatively, high-level ab initio computations) of the electronic states and their rovibronic levels for gas-phase

species, direct measurements of heat capacity for condensed-phase species, etc. However, the central thermochemical property for any chemical species, the enthalpy of formation, $\Delta_f H°(T)$, or the related Gibbs energy of formation, $\Delta_f G°(T)$, are obtained from determinations that express these quantities relative to other chemical species. Examples of species-interrelating determinations are enthalpies of chemical reactions, $\Delta_r H°(T)$, bond dissociation energies, D_0 or BDE(T), kinetic equilibria, $K_{eq}(T)$, electrode potentials, $V°(T)$, solubility data, $K_{sol}(T)$, atomization enthalpies, $\Delta_{atom} H°(T)$ (usually computed for polyatomic moieties rather than directly measured), etc. These determinations provide the change in the enthalpy/Gibbs energy for the overall chemical reaction, involving by definition at least two, and in most cases several different chemical entities.

The fact that information leading to enthalpy/Gibbs energy is accessible only through species-interrelating data leads to considerable complications. These are traditionally resolved by using a sequential approach that allows the compilers of thermochemical tables to manually work their way through the available data. The process consists of a series of steps, each focusing on one and only one target chemical species. The available species-interrelating determinations relating the target species to those (and *only* those) species that have been "solved" during previous steps is critically examined. The determination that is considered to be the "best" in the eyes of the evaluator (or some average of the "best" few) is selected, and used to derive the enthalpy of formation of the target species at some temperature. Of course, this can be accomplished only if the enthalpies of the other species involved in the selected "best" determination are known from previous steps, hence the fundamental restriction on the eligible pool of species-interrelating determinations. Once the enthalpy of formation of the target species is obtained, the available species-specific can be used to determine the remaining complement of thermochemical quantities for that species and their temperature dependence. The thermochemistry of the target species is then hibernated, and the sequence shifts to the next step where the procedure is reiterated for another chemical entity. As pioneered by the NBS compilation and its predecessors [11], this sequence usually follows the "standard order of elements" (oxygen, hydrogen, Noble gases, halogens, chalcogens, pnicogens, the carbon group, etc), adopted from the ICT [12] sorting order.

2.2. The deficiencies of the traditional sequential approach

Apart from the fact that any manual procedure is bound to be slow, a forced linear sequential procedure to obtain solutions from a set of intrinsically cross-related information entails considerable simplifications, and the resulting thermochemical tabulations have a number of serious difficulties. The biggest problem is the maze of hidden progenitor-progeny dependencies, which frustrates any attempt to update the tabulation with new knowledge. In practice, all one can do is to use new species-interrelating determinations to update the properties of one species, which is tantamount to revising one of the steps in the middle of the original sequence. While this may improve things locally for one target species, it immediately introduces new inconsistencies across the tabulation. Namely, there will be other species in the tabulation that are pegged directly or indirectly to the old value of the revised species and would also need to be updated - which those may be, is far from clear. Not only that, but the sequential procedure lacks corrective feedback to the thermochemistry of species that have been determined in previous steps and frozen, leading to cumulative errors. The traditional process also produces uncertainties that are not necessarily properly quantified and do not reflect the information content that is being used in other (later) parts of the tabulation. It can be argued that, even under the best of circumstances, the available information is used only partially during the sequential approach.

3. Thermochemical Networks: the approach of ATcT

The intricate cross-relational nature of thermochemical determinations means that the dependencies do not form a linear sequence but they span an elaborate network (Thermochemical Network, TN). The realizations that this is the case, together with pointers to the need to incorporate some form of statistical analysis and pre-treatment of the network, have been expressed quite some time ago [13,14]. However, limited computational capabilities available at the time prevented any serious (large-scale) developments. After a few early uses (late seventies/early eighties) on limited small blocks of species [11, 3, 4, 15] that were described via petite localized TNs and involved manual adjustments in lieu of a

statistical pre-treatment (and in most cases no pre-treatment at all), the idea went largely dormant, until a revival of TN usage several years ago [17, 18]. Other early attempts to computerize some aspects of the process of solving TNs, such as those of Pedley *et al.* [16], while producing valuable results, are arguably more akin to a computer-assisted sequential process than to a real TN treatment.

3.1. The TN approach of ATcT

Figure 1 provides a graphical representation of a very small section of our current Core (Argonne) Thermochemical Network. The depicted section focuses on chemical species related to the recent discovery that the generally accepted value of the bond dissociation energy in water is in error by ~2 kJ/mol [19]. In general, the TN corresponds to a mathematical graph with two types of vertices. The primary vertices (denoted by squares in Fig. 1) represent the enthalpies of formation of the chemical species (which need to be determined), while the secondary vertices (ovals) represent the chemical reactions that span the graph. The graph edges (arrows) are bo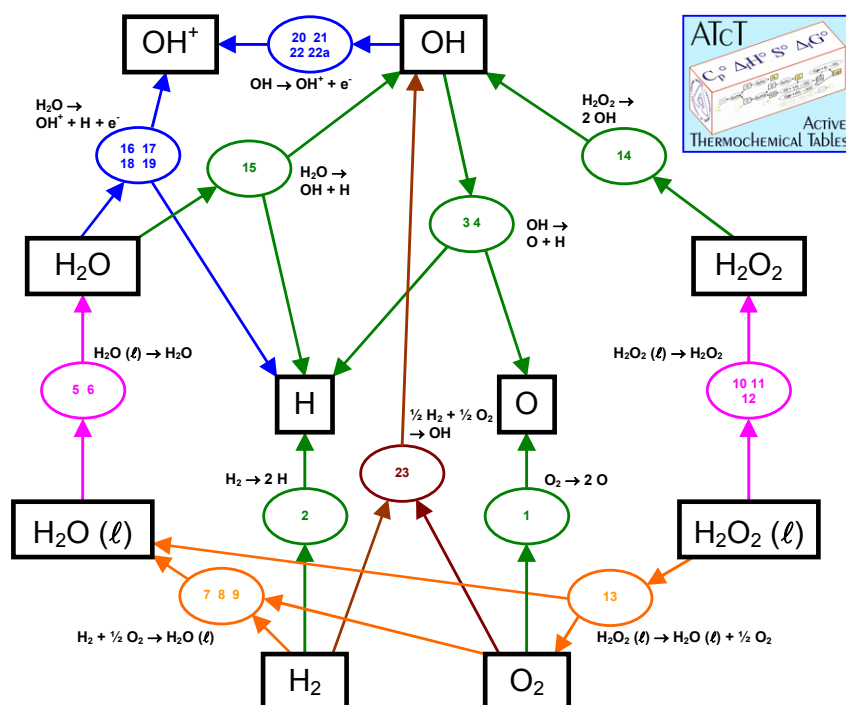th directed and weighted and map the participation of chemical species in the chemical reactions. The TN graph conforms to a number of curious rules. First neighbors, and indeed all $(2k+1)$-neighbor vertices are always of a different kind, and second neighbors (and all other $2k$-neighbors) are both either primary or secondary vertices. The details of the topology of the TN graph are driven by the chemical reactions it describes, which, through stoichiometric rules, determines the weights of the edges and rigorously limits the set of vertices that can ever become first, second, and third neighbors. Secondary vertices may have multiple degeneracies (reflecting competing species-interrelating

Figure 1. A small subsection of the current Core (Argonne) Thermochemical Network. The full network currently contains >600 primary vertices and >3200 secondary vertices. See text for further details

determinations of the same chemical reaction; each of the numbers contained in the ovals in Fig. 1 references one thermochemically-relevant species-interrelating determination at a particular temperature that is being considered).

The Core (Argonne) Thermochemical Network, whose subsection is represented in Figure 1, corresponds to a *global* TN (a.k.a. *ab ovo* TN). In such a TN, all primary vertices, except for reference elements in standard states (in the current example H_2 and O_2 at the bottom of the graph) are treated as unknowns, and nothing is held artificially fixed at a pre-selected value. Global networks start with the elements in their standard states, and need to include "key" chemical species (CODATA term [3], but species correspond to major hubs in the TN, such as H_2O, CO_2, NH_3, gas phase atoms, etc.), before progressing toward larger species. Also, stable parent species typically need to be included before the related free radicals can be inserted. The "ultimate" (or "complete") global TN would be extremely large and would include all known thermochemically-relevant determinations that were ever made.

A different and less complex variant is a local TN, which can be obtained by modifying some of the primary vertices and setting them to pre-selected fixed values. Such vertices then become terminators, and, using elementary mathematical manipulations, can be formally removed from the network, hence severing some of the links and simplifying the TN. Formally, any local TN can be imagined as being derived from a larger global TN by removing the primary vertices for which thermochemistry is considered to be firmly known from prior considerations and not sensitive to the thermochemistry encompassed within the local TN. While the latter is never completely true, in certain cases it can be a very good approximation. Using some perspicacity, local TNs that fairly reasonably define the thermochemistry of a small block of related chemical species can be generated. However, treatment of a global TN should be the preferred approach.

A visual examination of the TN depicted in Fig. 1 immediately reveals that there are many allowed paths (thermochemical cycles) through the graph between any two arbitrarily selected primary vertices. The traditional sequential approach uses only a minute subset of possible paths. Starting at a reference element in a standard state (which can be removed from the graph, since by basic thermochemical definition the reference states define the origin of the enthalpy/free energy scale), each of the sequential steps corresponds to selecting one trivial subgraph that consists of one secondary vertex and one adjunct primary vertex. The trivial graph, which corresponds to a linear equation with one unknown, is solved producing the value for the primary vertex. The primary vertex then becomes fixed at that value, and can be removed from the network, together with any secondary vertices that have become disjoint during this procedure.

However, from the statistical viewpoint, the proper solution for the TN is obtained not by favouring arbitrarily one path and ignoring the rest, but by considering all possible paths. This can be accomplished by finding the simultaneous solution of the whole system via minimization of a suitable statistical measure. While other possibilities can be also contemplated, it can be shown that the conventional χ^2 is indeed a reasonable statistical measure that can be minimized in order to obtain the proper solution [1]. However, this will work only if the adjunct uncertainties are an honest representation of the underlying confidence in the species-interrelating determination present in the TN. If the latter is not fulfilled, "optimistic" uncertainties correspond to outliers that will skew the results. Hence, the simultaneous solution in weighted space needs to be preceded by a statistical analysis that will detect and correct possible "optimistic" uncertainties. Among various statistical approaches that are possible, the "worst offender" analysis, explained in detail elsewhere [1], has proven to be very successful. The isolation of "offenders" is made possible by the redundancy of information present in the graph, created both by competing determinations and the existence of alternative paths in the graph. Once the TN is adjusted to self-consistency, the final (simultaneous) solutions for all primary vertices can be computed. These correspond to the enthalpies of formation of all involved chemical species, which can be then coupled to existing species-specific information to develop the full complement of thermochemical information for each species.

3.2. The advantages of the TN approach

The pivotal advantage of the TN approach is that the thermochemical values that are obtained by treating simultaneously the whole TN are significantly superior to those obtained by a sequential approach. The solutions and their adjunct uncertainties (including the complete covariance matrix, rather than just variances) are by definition guaranteed to be consistent both internally and with all the knowledge stored in the TN. The advantage of deriving thermochemical values by optimally using all the available knowledge (as opposed to using it only partially, as done during the sequential approach) is self-evident, and produces values that are typically significantly (and occasionally spectacularly) more accurate than those derived by a sequential approach from the same pool of original determinations. Internal consistency of the final values is actually also extremely important, for a very simple reason: Namely, chemists almost never need the individual enthalpies/Gibbs energies of formation of chemical species *per se*. Rather, the typical goal is to compute the enthalpy/Gibbs energy of a chemical reaction, which corresponds to sums and differences of the individual values of the species involved. A necessary (though not sufficient) condition to obtain the correct answer is that the used values are internally consistent.

Parenthetically, both the aspect of internal consistency and the virtual absence of the off-diagonal elements of the covariance matrix are a serious issue in current practices. Namely, in a typical situation, the thermochemical values for the individual species that are used to compute the overall thermochemistry of a reaction are – based on their availability - collected from a mixture of sources, including various thermochemical tabulations, as well as individual research papers. While the intent is to collect the best (and sometimes the only available) set of needed parameters, the collected values are not likely to be internally consistent. An illustrative case of both types of challenges would be a situation where, for example, a researcher has just obtained a spectroscopic measurement of some bond dissociation energy BDE(AB-C) = 456.78 ± 0.12 kJ/mol. He finds in a standard thermochemical tabulation the enthalpies of formation $\Delta_f H^o(ABC)$ = -234.5 ± 6.0 kJ/mol and $\Delta_f H^o(C)$ = 123.4 ± 2.0 kJ/mol, and uses them to sequentially obtain the (so far) unknown enthalpy of formation of the radical AB, $\Delta_f H^o(AB)$ = 98.9 ± 6.3 kJ/mol, where the quoted uncertainty is obtained by normal error propagation rules. However, without the full covariance matrix, the bond dissociation energy that is recoverable from $\Delta_f H^o(AB)$, even when identical values of $\Delta_f H^o(ABC)$ and $\Delta_f H^o(C)$ are used, has a seriously inflated uncertainty: BDE(AB-C) = 456.8 ± 8.9 kJ/mol. Worse yet, if the derived value of $\Delta_f H^o(AB)$ is coupled to $\Delta_f H^o(ABC)$ and $\Delta_f H^o(C)$ that are slightly different because there are extracted from a different tabulation, even the recovered value of BDE(AB-C) becomes incorrect.

Besides the pivotal advantages regarding the quality of the results discussed above, the TN approach has also other important advantages. A very important feature is that the resulting thermochemistry is easily updated with new knowledge: new information is simply added to the TN, and the system is solved again, instantly propagating the consequences of the new determination throughout the resulting values. Also, the TN can accept and happily commix experimental and theoretical determinations – providing that one can at least roughly estimate a realistic uncertainty for the computation, which in many cases is quite a challenge. Some help can be obtained from the fact that many standardized electronic structure computations have a generic measure of fidelity usually expressed as the "mean absolute deviation", MAD. However, the use of MAD as *the* uncertainty for the computed values produces utterly optimistic error limits. MAD needs to be multiplied by a factor of at least 2, and typically 3 or more, to obtain an estimate of the 95% confidence limit, which is customary in thermochemistry. (The use of 95% confidence limits in thermochemistry has been introduced by Rossini [20] more than seven decades ago, and subsequently embraced by virtually all thermochemical tabulations. Before meaningfully comparing values, e. g. from theory and experiment, care needs to be taken to have the uncertainties on the same footing.) Another type of help to assess the appropriate uncertainties is obtained from the "worst offender" analysis of the TN, which indicates which determinations (experimental or theoretical) have an "optimistic" uncertainty.

Another exciting feature of the TN approach is the possibility of running "what if" scenarios (hypothesis testing): new or tentative data is easily tested for consistency or lack thereof against the existing body of knowledge. If the tested hypothesis is inconsistent with existing knowledge, the statistical analysis can identify the conflicting information, allowing its re-analysis. Similarly, a statistical analysis of the TN is capable of isolating the "weakest links" in the existing TN, thus identifying new experiments or computations will best enhance the resulting thermochemistry. This alone is a new paradigm of how to use efficiently the limited scientific laboratory and computational resources.

4. Active Thermochemical Tables within the Collaboratory for Multi-Scale Chemical Science

Since today's commodity computers exceed the capabilities of yesterday's supercomputers, the original justification for practicing the sequential thermochemistry approach instead of the computationally more demanding TN approach has vanished. Active Thermochemical Tables are a software suite based on the TN approach [1,2]. ATcT have been designed with a broad purpose in mind, and can be exposed to the world through the web, but also run as a self-standing application on a PC. ATcT are currently exposed through a framework of web and grid services using servers at Argonne National Laboratory [21,22]. These services are an integral part of the Collaboratory for Multi-Scale Chemical Science (CMCS) [23], and are accessible to users via the CMCS Portal.

4.1. Collaboratory for Multi-Scale Chemical Science

CMCS [23] is one of the SciDAC (Scientific Discovery through Advanced Computing) National Collaboratory projects funded by the Office of Science of the U.S. Department of Energy (DOE). CMCS brings together researchers and developers across multiple DOE laboratories, other government laboratories, and academic institutions to develop an open knowledge grid for synthesizing multi-scale information to support a system-based research approach. While these developments are general in nature, CMCS is currently geared to support combustion research. The core activity is the development of a multi-scale informatics toolkit using advanced collaboration and metadata-based data management technologies to address a number of issues core to the emerging concept of knowledge grids, including provenance tracking and lightweight federation of both data and application resources into cross-scale information flows. The CMCS portal is currently in use by a number of high-profile pilot groups and is playing a significant role in enabling their efforts to improve and extend community maintained chemical reference information.

4.2. ATcT and the Core (Argonne) Thermochemical Network

ATcT is the central application on the thermochemical scale of CMCS. ATcT consists of several parts: the software kernel, the user interface with the framework of services, and the underlying data bases. Both the ATcT software and the databases are under continuous development at Argonne National Laboratory. The software suite allows queries of existing data, as well as visualization and manipulation of Thermochemical Networks. The users can set a variety of options that direct the mode in which the databases are queried, results presented, or handling of networks is parameterized.

The ATcT databases are organized as a series of Libraries. The Main Library contains the Core (Argonne) Thermochemical Network that is currently being constructed and is growing on a daily basis, and which, as of version 1.051, encompasses >3200 experimental and theoretical determinations spanning >600 thermochemically-distinct species. The Auxiliary Libraries aim to reproduce historical data from popular thermochemical compilations for reference purposes. Users have an opportunity to establish their own mini-libraries (Notes), in which they can store data that supplements or modifies the information contained in the Main and Auxiliary Libraries.

4.2.1. Early scientific results from Active Thermochemical Tables

The advent of ATcT has indeed ushered a new era for thermochemistry. Though we have barely scratched the surface on exploring and using the new capabilities, ATcT has already produced a number of significant scientific results. We are currently in the process of preparing several formal scientific papers that will expose and discuss the new thermochemical results. A few illustrative examples of early (mostly unpublished) results are given below.

The significant improvement in the accuracy of the thermochemistry that is obtained by using the ATcT TN approach has become immediately obvious even as we have started solving the first few initial versions (and quite smaller at the time) of the Core (Argonne) Thermochemical Network. The network includes so far chemical species containing H, O, C, N, and halogens. As mentioned earlier on, since this is a global TN, it needs to be built from the base up, and hence, by necessity, incorporates a significant number of "key" thermochemical species, as defined by the CODATA compilation [3]. Our aim was not to try to improve on these "key" values, but to explore the thermochemistry of species that are "higher up" in hierarchy. It is generally accepted that the thermochemistry of "key" species is by now well-defined, based both on the fact that the CODATA Task Group has made a meticulous and dedicated effort to evaluate them, and by the fact that the thermochemistry of these important species has been subsequently implicitly validated through (literally) daily use. In fact, virtually all thermochemical compilations [4-10] that appeared since the CODATA report [3] simply adopt the CODATA values without questioning them, and concentrate on evaluating other species. However, to our complete surprise, even the earliest solutions have demonstrated that there is apparently room for improvement in the thermochemistry of the "key" species. As the network was growing to include more and more data on these species, the output thermochemistry was further improving. The ATcT improvement is based on better (more complete, and, in fact, synergistic) use of the available information through the TN approach, as well as on the

ability to easily incorporate new data that appeared since the CODATA report. A subset of the new ATcT values for "key" species has been recently reported as an illustration of the power of the ATcT paradigm [1]. In many cases the uncertainties of the "key" values are lowered by a factor of 2 or more, and in some cases closer to an order of magnitude or even more. Some examples of such improvements are the enthalpy of formation of CO_2 (an order of magnitude better), NH_3 (a factor of 12), HCl (a factor of 15), and O atom (a rather spectacular factor of 50). Clearly, the reported amount of improvement critically depends on the soundness of the ATcT uncertainties. These are indeed proper 95% confidence limits – a point that has been separately verified by running on several occasions a separate Monte Carlo (both regular and Latin hypercube flavor) analysis of the TN.

Another recent success was the exploitation of the hypotheses testing ability of ATcT. Namely, we have recently revised the enthalpy of formation of OH radical (or, equivalently, the bond dissociation energy of water) [19]. This is a very important species playing a prominent role both in combustion and in atmospheric chemistry. The revision was achieved by combining new multiple measurements, a critical analysis of older spectroscopic data, and state-of-the-art electronic structure computations within a framework of conventional sequential thermochemistry arguments. The new value has been since then independently confirmed [24], and is now generally accepted as the correct one. However, the sequential thermochemistry approach, though providing the scaffolding to support a transparent and convincing argumentation, prevented us both from using efficiently all the available data and from removing a small (from the thermochemical viewpoint) residual discrepancy of ~25 cm^{-1} between the best available measurements, hence limiting the reported accuracy to ±0.3 kJ/mol. The ATcT approach has, of course, no problem in digesting additional data, and, in fact, provides further verification that our revision was correct. Moreover, the statistical analysis that is used to check the TN for self-consistency easily pinpoints the source of the discrepancy, which can be further verified beyond any reasonable doubt by carrying hypothesis tests. These help shift the argument around and fully explore alternative possibilities. The end result is that the uncertainty in $\Delta_f H°(OH)$ has been now further reduced by a factor of 7.5, to ±0.04 kJ/mol, thus definitely removing this species from the list of potential sources of uncertainty in interpreting kinetic data and combustion/atmospheric modeling.

The process of constructing the Core (Argonne) Thermochemical Network has allowed us to discover a number of "weak links" in the TN. "Weak links" are measurements that, if improved on, may be pivotal in improving the thermochemistry for some important species (usually a hub which in turn may improve the thermochemistry of many other species). These are now being pursued in collaborative efforts to obtain additional experimental and theoretical determinations. One important weak link was $BDE(N_2)$, limiting the accuracy of $\Delta_f H°(N)$ to ±0.40 kJ/mol. As a result, new measurements have now been performed [25], and they already lower the uncertainty by a factor of 4. The same new data, when used through ATcT with the other available measurements, results in an additionally enhanced accuracy, increasing the overall improvement factor to approximately one order of magnitude [26]. A similar activity is pursued to lower the uncertainty in $\Delta_f H°(C)$ [27]. Gas-phase atoms are important thermochemical pegs, directly affecting the results of all electronic structure calculations that derive enthalpies of formation from atomization energies.

The IUPAC Task Group on Thermochemistry of Selected Free Radicals, IUPAC project 2003-024-1-100, has used an evaluation procedure based on the ATcT TN approach in the final step of the evaluation process which produced IUPAC recommended values for about a dozen important radicals [28]. The Task Group is now one of the pilot users of CMCS, taking advantage of the available rich collaborative environment to ease their task. Also, the data accumulated by the Task Group is currently being imported into the Core (Argonne) Thermochemical Network. In addition, another unrelated IUPAC Task Group, Standard Potentials of Radicals, IUPAC project 2001-015-1-100 [29], has been using ATcT to optimize the thermochemistry of radicals in aqueous medium.

We are currently engaged in further expansion of the Core (Argonne) Thermochemical Network, and implementation of additional features in the ATcT software, as well as working on developing automated links between the largest existing polynomialized thermochemical database [10] and ATcT.

5. Acknowledgments

References

[1] Ruscic B, Pinzon R E, Morton M L, von Laszewski G, Bittner S J, Nijsure S G, Amin K A, Minkoff M and Wagner A F 2004 *J. Phys. Chem. A* **108** 9979

[2] Ruscic B 2004 "Active Thermochemical Tables", in: *2005 Yearbook of Science and Technology* (a yearly update to *McGraw-Hill Encyclopedia of Science and Technology*), McGraw-Hill, New York, p. 3

[3] Cox J D, Wagman D D and Medvedev V A 1989 "CODATA Key Values for Thermodynamics", Hemisphere, New York

[4] Chase M W, Davies C A, Downey J R Jr., Frurip D J, McDonald R A, and Syverud A N 1985 "JANAF Thermochemical Tables", 3rd ed.; *J. Phys. Chem. Ref. Data* **14**, Suppl. 1

Chase M W Jr. (Editor) 1998, "NIST-JANAF Thermochemical Tables", 4th ed., *J. Phys. Chem. Ref. Data*, Monograph No. 9

[5] Gurvich L V, Veyts I V and Alcock C B 1989-1996 "Thermodynamic Properties of Individual Substances", Vols. 1-3, Hemisphere, New York and Begell House, New York

[6] Kerr A J 2004, "Strengths of Chemical Bonds", in: CRC Handbook of Chemistry and Physics, Lide D (Editor), 85th ed., CRC Press, Boca Raton, Sec. 9 p. 52

[7] Linstrom P J and Mallard W G (Editors) 2005 "NIST Chemistry WebBook", NIST Standard Reference Database Number 69, National Institute of Standards and Technology, Gaithersburg MD, 2003 (http://webbook.nist.gov)

[8] "Standard Thermodynamic Properties of Chemical Substances", in: Lide D (Editor) 2004 CRC Handbook of Chemistry and Physics, 85th ed., CRC Press, Boca Raton, Sec. 5 p. 4

[9] McBride B J, Zehe M J and Gordon S 2002 "NASA Glenn Coefficients for Calculating Thermodynamic Properties of Individual Species", NASA Technical Paper 2002-211556, NASA Glenn Research Center, Cleveland, OH

McBride B J, Zehe M J and Gordon S "Properties From Coefficients(PFC)" at http://cea.grc.nasa.gov/index.html

[10] Burcat A and Ruscic B 2005 "Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion with updates from Active Thermochemical Tables, Technion and Argonne National Laboratory", (in press),   pre-release available at ftp://ftp.technion.ac.il/pub/supported/aetdd/thermodynamics (mirrored at http://garfield.elte.chem/Burcat/burcat.html)

Burcat A 2001 "Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion", Technion Aerospace Engineering (TAE) Report # 867, Technion, Haifa

[11] Rossini F D, Wagman D D, Evans W H, Levine S and Jaffe I 1952 "Selected Values of Chemical Thermodynamic Properties", NBS Circular 500, National Bureau of Standards, Washington DC

Wagman D D, Evans W H, Parker V B, et al. 1965–1981 "Selected Values of Chemical Thermodynamic Properties", Parts 1–8, NBS Technical Note 270-1 through 270-8, National Bureau of Standards, Washington DC

Wagman D D, Evans W H, Parker V B, Schumm R H, Halow I, Bailey S M, Churney K L and R. L. Nuttall, "The NBS Tables of Chemical Thermodynamic Properties", J. Phys. Chem. Ref. Data, 11, Suppl. 2 (1982)

[12] Washburn E W (Editor) 1926-1930 "International Critical Tables of Numerical Data, Physics, Chemistry and Technology" Vols I-VII, McGraw-Hill, New York

[13] Guest M F, Pedley J B and Horn H J 1969 *Thermodynamics* **1**, 345

[14] Garvin D, Parker V B, Wagman D D, and Evans W H 1977 *Proc. Fifth Biennal International CODATA Conf.*, Dreyfus B (Editor), Pergamon, Oxford, p. 567

[15] Garvin D, Parker V B and White H J Jr 1987 "CODATA Thermodynamic Tables: Selections for for Some Compunds of Calcium and Related Mixtures", Hemisphere, Washington

[16] Pedley J B, Naylor R D, Kirby S P 1986 "Thermochemical Data of Organic Compunds", 2nd Ed, Capman and Hall, London

[17] Ruscic B, Michael J V, Redfern P C and Curtiss L A 1998 *J. Phys. Chem. A* **102**, 10889

[18] Ruscic B, Litorja M, Asher R L 1999 *J. Phys. Chem. A* **103**, 8625

[19] Ruscic B, Wagner A F, Harding L B, Asher R L, Feller D, Dixon D A, Peterson K A, Song Y, Qian X, Ng C Y, Liu J, Chen W and Schwenke D W 2002 *J. Phys. Chem. A* **106**, 2727

Ruscic B, Feller D, Dixon D A, Peterson K A, Harding L B, Asher R L and Wagner A F 2001 *J. Phys. Chem. A* **105**, 1

[20] Rossini F D 1931 *J. Res. NBS* **6**, 1

[21] von Laszewski G, Ruscic B, Wagstrom P, Krishnan S, Amin K, Nijsure S, Bittner S, Pinzon R, Hewson J C, Morton M L, Minkoff M and Wagner A F 2003 "A Grid Service-Based Active Thermochemical Table Framework", Lecture Notes in Comp. Sci. **2536**, Springer, Berlin

[22] von Laszewski G, Ruscic B, Amin K, Wagstrom P, Krishnan S and Nijsure S 2003 *Int. J. High Perform. Comp. Applicat.* **17**, 431

[23] Myers J D, Allison T C, Bittner S, Didier B, Frenklach M, Green W H, Jr., Ho Y L, Hewson J, Koegler W, Lansing C, Leahy D, Lee M, McCoy R, Minkoff M, Nijsure S, von Laszewski G, Montoya D, Oluwole L, Pancerella C, Pinzon R, Pitz W, Rahn L A, Ruscic B, Schuchardt K, Stephan K, Wagner A L, Windus T and Yang C 2005 Cluster Computing 8, (in press)

See also http://cmcs.org

[24] Herbon J T, Hanson R K, Golden D M and Bowman C T 2002 *Proc. Combust. Inst.* **29**, 1201

[25] Tang X, Hou Y, Ng C Y and Ruscic B 2005 *J. Chem. Phys.* (in press)

[26] Ruscic B, Pinzon R E, Tang X, Hou Y and Ng C Y, manuscript in preparation

[27] Ruscic B, Pinzon R E, Csaszar A G, Stanton J F and Kallay M , manuscript in preparation

[28] Ruscic B, Boggs J E, Burcat A, Csaszar A G, Demaison J, Janoschek R, Martin J M L, Morton M L, Rossi M J, Stanton J F, Szalay P G, Westmoreland P R, Zabel F, and Berces T 2005 *J. Phys. Chem. Ref. Data* **34**, 573

See also http://www.iupac.org/projects/2003/2003-024-1-100.html and the predecessor IUPAC project 2000-013-1-100 at http://www.iupac.org/projects/2000/2000-013-1-100.html

[29] IUPAC Task Group on Standard Potentials of Radicals, Chair: Stanbury D M, see http://www.iupac.org/projects/2001/2001-015-1-100.html