



## Summary Report:

# Symposium on Computational Thermochemistry

**Sponsored by  
the Computers in Chemistry Division,  
American Chemical Society  
held at the 212th ACS National Meeting,  
Orlando, Florida, August 25 - 29, 1996**

**Karl K. Irikura**

National Institute of Standards and Technology

**David J. Frurip**

Dow Chemical Company

U.S. DEPARTMENT OF COMMERCE  
Technology Administration  
National Institute of Standards  
and Technology  
Gaithersburg, MD 20899-0001

QC  
100  
.U56  
NO.5973  
1997

**NIST**



## Summary Report:

# Symposium on Computational Thermochemistry

Sponsored by  
the Computers in Chemistry Division,  
American Chemical Society  
held at the 212th ACS National Meeting,  
Orlando, Florida, August 25 - 29, 1996

**Karl K. Irikura**

National Institute of Standards and Technology

**David J. Frurip**

Dow Chemical Company

U.S. DEPARTMENT OF COMMERCE  
Technology Administration  
National Institute of Standards  
and Technology  
Gaithersburg, MD 20899-0001

June 1997



U.S. DEPARTMENT OF COMMERCE  
William M. Daley, Secretary

TECHNOLOGY ADMINISTRATION  
Gary Bachula, Acting Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS  
AND TECHNOLOGY  
Robert E. Hebner, Acting Director



**Summary Report**  
**Symposium on Computational Thermochemistry**  
Computers in Chemistry Division  
American Chemical Society  
212th National Meeting, Orlando, Florida, August 25-29, 1996

*Karl K. Irikura (National Institute of Standards and Technology) and  
David J. Frurip (Dow Chemical Co.), Organizers*

**SPONSORS**

National Institute of Standards and Technology  
(Physical and Chemical Properties Division)

International Business Machines  
(Scientific & Technical Systems and Solutions Group)

Petroleum Research Fund

Gaussian, Inc.

Cray Research

**CONTENTS**

Abstract .....	2
Symposium Papers and Schedule .....	3
Purpose of Meeting and Panel Discussion .....	5
Executive Summary of Panel Discussion .....	6
Synopsis of Panel Discussion .....	7
Synopses of Technical Presentations .....	13
Poster Abstracts .....	30

## ABSTRACT

This was the first symposium or conference ever held on the topic of *Computational Thermochemistry*. It was organized to address problems in molecular thermochemistry using all available computational methods, ranging from purely empirical schemes to state-of-the-art, *ab initio* quantum mechanical methods. Thirty-three papers were presented (18 invited) and there were five members on the discussion panel. Most of the participants were from the US, but there were two from Canada, two from Portugal, and one each from Spain, South Korea, and Sweden.

We strove to encourage the different scientific and engineering communities to mix by not placing similar talks in adjacent slots. A two-hour panel discussion was designed to identify specific needs for thermochemical data and to encourage communication between the users and suppliers of thermochemical data. Attendance ranged between 50 and 100 with good audience participation. Although some of the talks covered unfamiliar topics quite different from the rest, even these talks were well-attended and generated questions and discussions. We feel that the symposium papers and panel discussion achieved their respective goals very successfully.

**Symposium Papers and Schedule.** Invited talks are indicated by asterisks (\*).

Time	Speaker, co-author(s)	Title
<i>Monday, Aug. 26 (Omni Rosen hotel, Grand Ballroom D)</i>		
<i>Karl K. Irikura, Chair</i>		
1:30	[Karl K. Irikura, NIST]	[Introductory remarks]
1:40*	<b>Sidney W. Benson, USC;</b> Norman Cohen	Current Status of Group Additivity
2:10*	<b>Russell S. Drago, Univ. Fla.</b>	The ECT Model for Determining Trends in Chemistry
2:40*	<b>Joseph T. Golab, Amoco;</b> Michael R. Green	Practical Quantum Chemistry Applied to Process Design Studies
3:10*	<b>David A. Dixon, PNNL</b>	Computational Chemistry and the Prediction of Thermochemical and Thermophysical Properties
3:40*	<b>Tsan H. Lay, NJIT;</b> Lev N. Krasnoperov; Joseph W. Bozzelli	Calculations of Thermodynamic Properties of Alkyl Hydroperoxide Compounds with Chlorine or Fluorine Substituents Using Molecular Orbital Methods
4:10	<b>Andrzej Anderko, OLI Systems Inc.;</b> M. M. Lencka	Applications of Thermochemical and Thermodynamic Software to Corrosion Simulation and Materials Processing
4:40	<b>Roberta G. Susnow, Exxon Res. &amp; Eng. Co.;</b> W. H. Green; A. M. Dean; P. K. Peczak; L. J. Broadbelt	Computer-generated Chemical Kinetic Models
<i>Tuesday, Aug. 27 (Omni Rosen hotel, Salon 15)</i>		
<i>David J. Frurip, Chair</i>		
9:00*	<b>Larry A. Curtiss, Argonne NL</b>	Computational Methods for Calculation of Accurate Bond Energies, Electron Affinities, and Ionization Energies
9:30*	<b>Donald W. Rogers, Long Island Univ.</b>	Enthalpies of Hydrogenation
10:00*	<b>Panel Discussion</b>	Dr. David Frurip (Dow Chemical), <i>Leader</i> Dr. Joseph Golab (Amoco Corp.) Prof. Klavs Jensen (MIT) Maj. Walter Lauderdale, Ph.D. (Wright-Patterson AFB) Dr. George Thomson (DIPPR)
<i>Tuesday, Aug. 27 (Omni Rosen hotel, Salon 15)</i>		
<i>William F. Schneider, Chair</i>		
1:30	<b>Chul Soo Lee, Korea Univ.;</b> Ki-Poong Yoo	A New Group Contribution Method for Configurational Properties of Pure Fluids and Mixtures
2:00*	<b>Axel D. Becke, Queen's Univ. (Ontario)</b>	Density-functional Thermochemistry
2:30*	Candee C. Chambers; David J. Giesen; Zhen Gu; Gregory D. Hawkins; <b>Christopher J. Cramer, Univ. Minn.;</b> Donald G. Truhlar	Methods for Calculating Free Energies in Solution and Free Energies of Transfer
3:00	S. H. Hilal; <b>Lionel A. Carreira, Univ. Ga.;</b> S. W. Karickhoff	Vapor Pressure, Boiling Point and Activity Coefficient Calculations by SPARC
3:30*	<b>Carl F. Melius, Sandia NL</b>	Application of the BAC-MP4 Method and Its Variations in Determining Thermochemical Properties of Molecules
4:00	<b>Sherif A. Kafafi, Johns Hopkins Sch. of Med.</b>	Computation of Accurate Atomization Energies of Molecules from Density Functional Methods
4:30	<b>Jeffry D. Madura, Univ. So. Ala.;</b> B. Keith Harrison; C.-L. Huang; J. Dolfing	Evaluation of Estimation Methods and Application for Predicting Dehalogenation Pathways

Time	Speaker, co-author(s)	Title
<i>Wednesday, Aug. 28 (Omni Rosen hotel, Salon 11)</i>		
<i>Timothy J. Lee, Chair</i>		
8:30*	<b>Tom Ziegler, Univ. Calgary</b>	A Study of Periodic Trends in the Bond Energies of Transition Metal Complexes by Density Functional Theory
9:00*	<b>B. Keith Harrison, Univ. So. Ala.</b>	CHETAH, a Program for Convenient Thermochemistry Applications
9:30*	<b>George A. Petersson, Wesleyan Univ.</b>	Complete Basis Set Thermochemistry
10:00	<b>William F. Schneider, Ford Motor Co.</b>	Applications of Computational Thermochemistry to Halocarbon Atmospheric Chemistry: Successes and Failures
10:30*	Michael L. Mavrouniotis; <b>Michele M. Foisy, Northwestern Univ.</b>	A Conjugation-based Version of the UNIFAC Method
11:00*	<b>Peter Politzer, Univ. New Orleans</b>	Density Functional Calculations of Enthalpies of Formation and Reaction Energetics
<i>Wednesday, Aug. 28 (Omni Rosen hotel, Salon 11)</i>		
<i>Jeffry D. Madura, Chair</i>		
1:30	<b>Hussein Y. Afeefy, Univ. Md. Balt. Co.;</b> Joel F. Liebman; Suzanne W. Slayden	Thermochemical Data: There is More Than You Think But Less Than You Need
1:50	<b>Manuel E. Minas da Piedade, Inst. Sup. Técn. (Lisbon);</b> José A. Martinho Simões	Estimation of Entropy Changes in Organometallic Reactions Using Data for Organic and Inorganic Model Reactions
2:10*	<b>Timothy J. Lee, NASA Ames Res. Ctr.</b>	Determination of Highly Accurate Heats of Formation
2:40	<b>Rajiv J. Berry, Wright-Patterson AFB;</b> Martin Schwartz; Paul Marshall	Ab Initio Investigation of Halocarbon Thermochemistry and Kinetics
3:10*	<b>Per E. M. Siegbahn, Stockholm Univ.</b>	Calculation of Bond Strengths for Transition Metal Complexes
3:40*	<b>Andrew J. Holder, Univ. of Mo. K. C. and Semichem</b>	SAM1 Semi-empirical Parameters for Transition Metals
4:10	<b>David J. Frurip, Dow;</b> Nelson Rondan; Joey Storer	Implementation and Application of Computational Thermochemistry to Industrial Process Design at the Dow Chemical Company
4:40	<b>Klavs F. Jensen, MIT;</b> Harsono Simka	Thermochemistry of Organometallic Precursors for Chemical Vapor Deposition
<i>Poster Sessions</i>		
<i>Monday, Aug. 26, 8:00 pm (Convention Center, Valencia Room)</i>		
<i>Tuesday, Aug. 27, 7:00 pm (Omni Rosen hotel, Grand Ballroom E)</i>		
<i>George R. Famini, Chair</i>		
	<b>James S. Chickos, Univ. Mo. St. Louis;</b> William E. Acree Jr.	Advances in the Estimation of Fusion Entropies and Enthalpies by Group Additivity
	<b>Albert Davydov, Univ. Fla.;</b> C. H. Chang; T. J. Anderson	Assessment of Thermochemical and Phase Diagram Data for Selected Compound Semiconductors
	<b>Maricel Torrent, Univ. Girona (Spain);</b> Miquel Duran; Miquel Solà	The Decarbonylation Step of the Dötz Reaction: Density Functional Predictions



## PURPOSE OF MEETING AND PANEL DISCUSSION

Thermophysical and thermochemical prediction techniques have received increased attention from industry in the last few years. One major reason is increased global competition, which is shrinking the timelines for industrial process scale-up (from laboratory to production). This has resulted in an increased reliance on, and confidence in, computerized prediction techniques, modeling and simulations to replace many (expensive) experiments. Concurrently, there has been a rapid increase in computational "horsepower," allowing the development of new predictive techniques that are practical in an industrial environment.

The focus of this symposium was "Computational Thermochemistry." The organizers (Karl K. Irikura of NIST and David J. Frurip of Dow Chemical) took a broad approach to include any computer-based means for predicting the thermochemical properties of a substance. This encompassed everything from empirical group methods to high-level *ab initio* calculations. Topic areas included:

- \* Advances in Group Contribution Techniques
- \* New Software to Predict Thermochemistry
- \* New Computational Techniques
- \* Validation of Predictive Methods
- \* Applications of Techniques to Practical Problems

In addition to the technical presentations, there was a Panel Discussion of the needs for and applications of accurate thermochemical data, with an emphasis on prediction. The panel represented the users of thermochemical data in areas such as chemical processing, microelectronics, academia, and the military. The discussion offered a unique opportunity for both the users and those who are developing the techniques to discuss, in an open forum, common problems and how to work toward common goals which are mutually beneficial. The organizers anticipate that the results could help to guide developments in this field. The members of the panel are listed on the following page.

The topic areas for the panel included, among others, the following areas:

- \* Thermochemical data needs in industry, academia, military, etc.
- \* Bottlenecks in generating data
- \* Experimental verification needs
- \* Suggestions for techniques development (e.g., user-friendliness)
- \* Condensed phase data (e.g., vaporization heats, sublimation heats)
- \* Common, consensus test data set for model verification

## EXECUTIVE SUMMARY OF PANEL DISCUSSION

The significant conclusions resulting from this panel are given below.

1. A need exists to develop and make available a critically evaluated database of thermochemical data. This database can then be used by those developing the techniques as a standard test for comparing the accuracy of different models *in a fair and unbiased way*.
2. Technique development efforts need to continue to be broadened to allow the accurate prediction of a wider spectrum of engineering properties such as kinetics, vapor-liquid equilibrium parameters, critical data, and the properties of molecules on surfaces.
3. An effort needs to be made toward improved dissemination of these techniques to the engineering and industrial community in general. Key to this is a realization by the developers of these techniques that they must be straightforward (easy) to use.
4. Corporate "cultures" should be changed to allow for the dissemination of private but non-proprietary (industrial) thermochemical data to the general public, or at least to those who maintain resources such as databases and estimation programs.

### Panel Members

Dr. Joseph Golab  
Amoco Corp.  
Naperville, IL

Prof. Klavs Jensen  
Dept. of Chemical Engineering  
Massachusetts Institute of Technology  
Cambridge, MA

Maj. Walter Lauderdale, Ph.D.  
Wright-Patterson AFB, OH

Dr. George Thomson  
Design Institute for Physical Property Research (DIPPR)  
Bartlesville, OK

Dr. David Frurip (*Discussion Leader*)  
The Dow Chemical Co.  
Midland, MI

## SYNOPSIS OF PANEL DISCUSSION

The panel discussion began with a brief presentation by each panel member (5 min. with an optional transparency). Discussion was then opened to the audience.

**George Thomson (DIPPR)** opened the discussion by citing four of the earlier technical papers in the symposium that showed impressively useful results from predictive modeling. As a potential user in his engineering community, he doesn't want to use quantum mechanics because it's expensive and requires major computing resources. In contrast, Benson's method (and similar methods) are easy to use (although he'd prefer not to calculate symmetry numbers) and are reasonably accurate. The quantum methods have many drawbacks and thermochemical properties are only a minor part of the interests of DIPPR members, so most of this symposium is irrelevant to DIPPR's projects. Properties such as critical temperature and viscosity are much more useful.

**Joe Golab (Amoco)** emphasized that computational methods can have impact only if they are easy to use and well-integrated with standard engineering modeling codes such as Aspen. Chemists must talk to chemical engineers to learn what will be useful and not merely possible. There is seldom time for methods validation, and calculations are most often invoked during brainstorming sessions. Calculations are not trusted as much as experiments. Thermodynamics is not enough; kinetics is also needed. Many databases exist but they're not intelligent enough. They may be poorly evaluated and they don't indicate how values were derived, such as by experiment or by estimation. Databases should also be able to check themselves for internal consistency and to suggest how to generate needed but missing data. Transition-metal surfaces are one area in which more data are needed. Other areas are data in solvents, under various conditions of temperature and pressure, multi-phase systems, systems containing impurities, and polymers. Finally, standardization is needed to allow different databases to communicate.

**Klavs Jensen (MIT)** pointed out that semiconductor processing is very capital-intensive, at about \$10<sup>9</sup> per plant, and that time is always urgent. The industry is moving beyond silicon to compound semiconductors. From a cultural point of view, the industry is more willing than most to use computers for simulations. Fluid mechanics and the shapes of vias [etched channels] are often modeled successfully, but chemistry is often merely guessed. This is becoming inadequate as the technology evolves. Most of the researchers in this area are trained in solid-state physics. The data needs are very broad, including reactive processing, chemical vapor deposition (CVD), plasma etching, and surface cleaning. Most of the periodic table is important, and most of the interesting compounds are organometallic. Rate constant data are needed, not just thermochemistry, and predictive methods must be of known accuracy and sufficiently fast to be used for screening. Surfaces are critical, and thermochemistry on surfaces is a major unfulfilled need.

**Maj. Walt Lauderdale (Wright-Patterson AFB)** presented the perspective from the fuels and lubrication division, and disclaimed representing the entire Air Force. Jet fuel is the principal heat sink on-board jet aircraft. As it is thermally stressed it will decompose into various gums, varnishes, and other undesirables. Inhibitors are needed to negate these decomposition reactions and hence yield increased heat sink. Other increased heat sink strategies include 1) operating the fuel under supercritical conditions and 2) endothermic reactions. Lubrication systems must endure high temperatures (> 260 °C), so exotic alternatives must be considered such as vapor-phase

delivery. Combustion modeling is another important area. Unfortunately, computational fluid dynamics (CFD) with chemistry is very hard to do for real fuels. Reduced mechanisms/reaction sets (global chemistry) are needed to make the problems tractable. More of both theory and experiment are needed on the effects of supercritical solvents on the rates and mechanisms of chemical reactions. Better data on surface-molecule interactions are needed to model physisorption and chemisorption on different surfaces. This information will help determine the factors that limit bearing lifetime. There are also other applications of this information which extend beyond lubrication. One must be able to model the effects of fuel contaminants and heteroatomic systems, since they are known to affect the formation of gums, varnishes, and coke in/on fuel-wetted components. Potential combustion intermediates must be considered in order to learn how to reduce pollutant emissions and to develop fire extinguishants to replace halons. Dr. Lauderdale summarized by pointing out that these are very complex systems, such as modeling the wear of bearings. At this point one just needs to know what the most important effects are; there's no time to wait for the perfect methods to be developed.

**Dave Frurip (Dow)** presented Dow's major perceived need for computational thermochemistry as providing enthalpies of reaction. Typically an accuracy of  $\pm 12$  kJ/mol will satisfy an engineer who's buying a heat exchanger, since she or he will over-specify anyway. The principal method they use currently is group estimation, mostly through the CHETAH program (Dr. Frurip is one of the current developers). Some specific needs are (a) missing groups, (b) condensed-phase data, as by the group methods of Domalski (NIST), (c) data should be computerized for easy dissemination, (d) energy-release properties such as enthalpies of detonation. He also made some suggestions to make quantum-chemistry more useful: (1) report standard enthalpies of formation instead of atomization energies, (2) agree on a standard set of test data, (3) actively seek and report limitations of methods, (4) make methods easy to use ("black box"), and (5) avoid unnecessary computational expense by making the packages more robust, with better error-checking. For item (2), he suggested that a group or committee might be helpful for identifying the classes of molecule that should be represented in the database and for collecting and scrutinizing the data. He mentioned the data sets of Melius and Curtiss and suggested that NIST might be appropriate for the project.

Dr. Frurip then opened the floor for general discussion.

**Joel Liebman (Univ. Md. Baltimore Co.):** As will be discussed in tomorrow's talk, a new database is under development at NIST involving Hussein Afeefy and me; please contribute your "secret" data that are not in the open literature.

**George Petersson (Wesleyan Univ.):** It's essential that the database that Dave Frurip suggested be critically evaluated. The uncertainty estimates should also be accurate.

**Peter Politzer (Univ. of New Orleans):** For explosives applications, both gas- and solid-phase energetics are needed.

**George Parks (Phillips Petroleum):** I have participated in a meeting of representatives of 15 companies who met to discuss the influence and uses of computational thermochemistry. We concluded that group additivity is powerful and useful but that more sophisticated quantum methods have negligible impact. Enthalpies of formation are not particularly interesting. Phase equilibria and

activity coefficients are much more important, and computations appear promising for them. The world runs on modeling codes such as Aspen. One question we focus on is that of urgency vs. method. How would you obtain a free-energy of reaction value if you had 1 hour vs. 24 hours vs. 1 week?

**Fruip:** First seek a suitable analog for which data were available. Failing that, try Benson's group methods. If the group is missing then turn to quantum calculations. A G2 calculation can be done overnight under good conditions.

**Golab:** At Amoco, if an engineer calls he's probably already tried his own estimations and failed, and he's probably willing to wait longer than the times just cited.

**Fruip:** Quantum mechanics is not used more because the methods have only recently been made reliable to "chemical accuracy."

**Fred Breitbeil (organic chemist in academia):** Many of my colleagues don't trust quantum chemical results and are biased against them (even within the physical chemistry community). Is it merely a generation gap or a more problematic barrier?

**Fruip:** The methods can only be trusted in the light of systematic validations, such as those done by Curtiss.

**Dave Dixon (Pacific Northwest NL):** An effort to use quantum chemistry as an aid in designing chemical processes and plants was begun more than 10 years ago at DuPont. G1 and G2 are not the only way to attack the computational thermochemistry problem. For many molecules the standard tricks, such as isodesmic reactions, are very good. Quantum chemical methods are needed because group additivity schemes sometimes have problems, for example, in treating partially fluorinated systems. The real job for computational thermochemistry is to solve problems that can't be solved in the lab because of time or dollar constraints. However, the computational results have to be carefully validated. For actual processes, we really need high accuracy. For example, the choice of a new catalyst for a commercial process may be dependent on its performance being a factor of 2 to 4 better than an alternative process. In terms of a binding energy or an activation energy, this corresponds to only 0.4 to 0.8 kJ/mol. A similar feature is noted for the design of separation systems. We are clearly not at this stage of accuracy yet but we should not stop at 4 to 8 kJ/mol of accuracy and think that as computational thermodynamicists that we have finished the job. New thermophysical property prediction methods are being developed for gas phase and solution phase species. Reliable condensed-phase predictions will require accurate intermolecular potentials which will probably become increasingly available over the next five years. "Green" process design and environmental-cleanup applications will require high quality thermochemical and kinetic data.

**Don Rogers (Long Island Univ.):** Progress is rapid but we are still at an early stage. Academic colleagues must remember that they are training the next generation.

**Carl Melius (Sandia NL):** Transition states must be calculated accurately to get reaction mechanisms and processes right. Computer power is increasing and costs are decreasing, so computational methods will become increasingly important. There is a hierarchy of tools currently

available. When using a method you must also know when to believe its predictions. Unlike experimental measurements, the errors in G2 predictions are not random; systematic errors can be corrected. Reference data are needed, and unfortunately the available data are not well-matched to current needs. We need some new, accurate experimental data. A group that makes a list must consider this in making recommendations.

**Petersson:** Accurate gas-phase data are needed to develop methods for predicting solvation energies.

**Thomson:** There are only two experimental labs in the US that do accurate combustion calorimetry: Bill Steele (NIPER; oxygen) and Pat O'Hare (NIST; fluorine). DIPPR is also trying to get new Benson groups, but the bottleneck is experimental measurements.

**Fruerip:** Maybe any committee should also recommend a set of experimental measurements.

**José Artur Martinho Simões (Univ. of Lisbon):** I'm developing group methods for organometallics and find that the experimental database is very thin and unreliable. I'm currently working with NIST to develop a database of organometallic thermochemistry.

**Manuel Minas da Piedade (Inst. Sup. Técn., Lisbon):** In my experimental combustion calorimetry work I find that compounds are often available only in small amounts. Only two other groups do microcalorimetry.

**Dave Golden (SRI):** I'm on a NASA data evaluation panel. In some applications, the greatest need is for qualitative information about reaction pathways, not for more precise thermochemical data. In one project studying the combustion of natural gas, we found that we had to change the well-established rate constant for  $\text{CH}_3 + \text{H}$  by 40% to get the model to agree with experiment!

**Lauderdale:** Many experiments are very difficult, such as those involving surfaces. Often the need is for qualitative or semi-quantitative information to guide experimental research and resource allocation. For such screening purposes, don't wait for the experimental data to catch up--relative comparisons are often what's needed.

**Bill Schneider (Ford):** Accuracies must be quantified, but the future is in topics such as solution and heterogeneous chemistry, for which it is difficult to establish the accuracy of predictions, and for which accuracies of  $\pm 8$  kJ/mol are not required to make useful predictions.

**Rogers:** Are the complainers about the lack of data sitting on their own?

**Fruerip:** Yes, of course. Usually there isn't enough push or incentive to disseminate data.

**Liebman:** Often in examining data, quantity is more useful than quality.

**Dixon:** If you want to work with experimental people you must establish credibility by always having reliable results. Surfaces and solids are clearly an important area for computational science

but today it is hard to achieve accuracy in such predictions or high performance computational performance. Improved methods for surfaces and solids are being developed today.

**Sid Benson (Univ. of So. Calif.):** In condensed-phase there can be complex donor-acceptor interactions such as hydrogen-bonding in alcohols, which results in tetramers, not individual molecules. In physical chemistry there is still no theory of viscosity for the liquid state, as there is for gases.

**Butch Carreira (Univ. of Georgia):** Later talks in the symposium will address some of the other topics and properties that have been mentioned.

**Joe Bozzelli (New Jersey Inst. Tech.):** Reactive intermediates such as adducts from combination, insertion and addition reactions are important for estimating and understanding kinetics and reaction pathways. It is important when reporting thermodynamic property results to include the entropy as well as the more common enthalpy of formation, because entropy is an important component for determining the pre-exponential factor. Reasonable estimates, within an hour or so, of thermodynamic parameters,  $\Delta_f H$ ,  $S$ ,  $C_p(T)$  can be obtained from group additivity where unknown groups can be estimated from an isodesmic reaction calculation performed at the semi-empirical level such as MOPAC PM3. One can also search literature databases, such as that of Melius, for more accurate enthalpy data on molecules which have a needed group for GA.

**Benson:** Bond additivity does much better than DFT in many cases.

**Fruip:** Is a committee needed to create and design a reference database?

**Thomson:** Sharing or merging the databases already mentioned might be a good starting point.

**Benson:** Money could be a problem; nobody wants to pay for thermochemical data compilation.

**Dixon:** Don't restrict the list to stable molecules; include ions and radicals as well. Radicals are the foundations for isodesmic bond-strength calculations. For entropy, it's still hard to solve systems of coupled rotors exactly. Is this fundable as "basic new theory"?

**Allan Smith (Drexel Univ.):** Condensed-phase data will be needed in the database.

**Fruip:** The database must be useful for developing and testing quantum-mechanical and other techniques.

**Melius:** Different communities have different needs, so reaching a consensus on the data needs may be difficult. NIST is probably the appropriate organization for such a project.

**Liebman:** Please include in the report of this discussion the e-mail addresses of the participants.

**Lauderdale:** We need to think of these data activities in terms of customers and suppliers. What resources are needed to solve the most important problems?

The following significant conclusions and plans serve to summarize the discussion.

1. **A need exists to develop and make available a critically evaluated database of thermochemical data for technique developers to use to test the accuracy of their models.** This seemed to be the one clear cut item which most thought would help the community of technique developers and users. It seems that most “developers” (both quantum mechanical and group methods) have their own databases. Most data are from commonly available public sources. Larry Curtiss has offered for public use his new, larger database compiled for testing of the G2 (and variants) method and said that it would be available on the Internet soon. Others who offered to share their databases are Joe Bozzelli, Carl Melius, and Dave Dixon. We discussed that perhaps the best institution for compiling these data is NIST.
2. **Efforts need to be focused on the accurate prediction of other properties of engineering importance and extension of current techniques to a wider variety of molecular types.** It was pointed out many times during the discussion that there are efforts in this direction, and in fact, several papers were presented at the symposium which discussed this (e.g., Politzer, Carreira). Whereas thermochemical properties are industrially important, there are many more properties (e.g., kinetics, vapor-liquid equilibrium data, critical properties) which are key to making a process an economic success. Extension to organometallics and to the properties of molecules on surfaces is also important.
3. **Dissemination of these techniques to the engineering community.** It appeared clear that although a few companies (Dow, DuPont, Amoco, among others) have strong and active molecular modeling departments, many don't. How do we get the word out to the user community that these techniques exist and are beneficial? This symposium is one method but other efforts are needed as well. One clear message is that the techniques must be “easy” to use to gain widespread acceptance and use.
4. **Make private but non-proprietary (industrial) data public.** If a mechanism were in place for this to occur more readily, we'd all be better off. Unfortunately, most corporate cultures do not allow this to happen easily, due to complicated and cumbersome clearance processes. Keeping these data private sometimes provides a competitive advantage, but often this is probably not the case. The first step might be for Joel Liebman to compile and distribute a list of needs, i.e. where he sees the “holes” in the data tables.



## SYNOPSIS OF TECHNICAL PRESENTATIONS

**Sid Benson** (Univ. of So. Calif.) gave the opening talk of the symposium. He identified the historical motivation for the development of empirical group additivity (commonly called the Benson group method) as the need to predict chemical reactivity in organic systems. Earlier bond additivity schemes were refined into the group additivity (GA) methods used today to predict many properties. For thermochemical quantities ( $\Delta_f H^\circ_{298}$ ,  $C_p^\circ(T)$ ,  $S^\circ_{298}$ ), hydrocarbon values are now reproduced to within the experimental uncertainties, and singly- and doubly-functionalized organics give good results. Some discrepancies remain for molecules that are highly branched or contain multiply-substituted polar groups. Prof. Benson pointed out that the experimental generation of thermochemical data has nearly ceased (aside from phase diagrams and mixtures), which makes it difficult to develop values for new groups. He continued by presenting methods for estimating enthalpies of vaporization, with a typical accuracy of better than 2 kJ/mol. "Homothermal pairs" are groups similar in size and electronegativity and for which metathesis reactions are nearly thermoneutral. Several pairs have been discovered and they increase the range of applicability of GA. A modified electronegativity scale has been developed that also extends the range of GA, and methods have been developed for  $\Delta_f H^\circ_{298}$  of solid salts (good to 4-8 kJ/mol).

**Russell Drago** (Univ. of Florida) described his "ECT" model, which was designed with the dual goals of (1) predicting chemical properties such as enthalpies, rate constants, nmr shifts, and quadrupole coupling constants, and (2) testing models of chemical bonding. He stressed that such a model must contain two parameters, for electrostatic and covalent interactions, and that single-parameter models such as electronegativity can only succeed for limited data sets or with heavy constraints. The experimental enthalpies for 500 reactions, both organic and inorganic, were fitted to  $\pm 1$  kJ/mol using 188 ECT parameters. Systems with important steric effects or  $\pi$ -backbonding are reproduced less accurately. Gas-phase ion-molecule reactions are problematic for the simple ECT method, which fails to predict, for example, the 40 kJ/mol bond in the  $\text{LiKr}^+$  molecule. However, another term can be added to the ECT equation to yield good results. This also allows gas-phase/solution differences to be decomposed into donor-acceptor and pure solvation effects. Similar decomposition of basic ECT results into covalent and ionic contributions leads to the concepts of "catimer" and "animer" parts of a molecule (incipient cation and anion, respectively).

**Joe Golab** (Amoco Res. Ctr.) examined computational thermochemistry as one part of the larger area of chemical modeling. He then examined chemical modeling (theory and simulations) as an "enabling technology" that can provide data for use in making technical business decisions. Technical applications at Amoco include thermochemistry, phase changes, chemical kinetics, reaction mechanisms, and several kinds of spectroscopy: IR, UV-vis, NMR ( $^{13}\text{C}$ ,  $^{29}\text{Si}$ ), diffraction (electron, neutral, crystal, powder, fiber), EXAFS. Even inexpensive quantum methods often yield acceptably accurate results, and they have the big advantage of being applicable to any compound. Dr. Golab presented a number of real applications in which calculations have had impact. Catalytic conversion of methane to methanol by 5d transition metals was studied by a wide variety of methods for all the metals from Hf through Pt. The original candidate metal identified experimentally was tantalum. The calculations accurately predicted the performance of other promising candidates and explained the differences in reactivity among them. Modeling at the engineering level (e.g., using ASPEN) is still very difficult, but is promising and being pursued. He also discussed styrene process design (alkylation and trans-alkylation), including CHEMKIN modeling of gas-phase kinetics. Supporting thermodynamic data were from databases such as NIST's or from AM1 semi-empirical

calculations using the GAMESS program. The results were very successful and not unique; several other successes were listed, including light naphtha desulfurization. Computational technologies are useful for process assessment and improvement, for discovering reaction mechanisms, for product development, and even for addressing marketing questions. In summary, Amoco uses computational chemistry for a competitive advantage in bringing ideas to market rapidly. As computing resources become cheaper, modeling is expected to be important for increasingly complex problems, such as corrosion.

**David Dixon** (Pacific Northwest Natl. Lab.) is concerned with environmental problems related to the Dept. of Energy. *Ab initio* thermochemistry is used either directly or to produce new groups for Benson's group additivity (GA) methods. The specific problem, and the number of molecules for which data are needed, dictate the strategy followed. Molecular orbital (MO) and density functional theory (DFT) methods are the *ab initio* alternatives. DFT is cheaper but less accurate and more recent, so its reliability is more unknown. Of course, GA is used when the necessary group values are available, but they are often lacking. The rates of catalytic reactions are often of interest. To illustrate the accuracy that may be required, note that at 25 °C a factor of 10 in the rate corresponds to a precision of about 5.7 kJ/mol in the activation energy. Likewise, the difference between a product selectivity of 50:50 vs. 99:1, or between an equilibrium constant of 1 vs. 100, is only 11 kJ/mol. For *ab initio* calculations, the one-particle basis set problem is largely solved by the correlation-consistent basis sets developed by Dunning and co-workers at PNNL. For the many-electron problem, multi-configuration approaches are problematic to implement because of the important but somewhat arbitrary choices that the user must make. Configuration interaction methods are not size-consistent. This leaves Møller-Plesset and coupled-cluster theories, and possibly DFT. Transition-metals remain a very difficult special case. For practical work, it's important to use isodesmic reactions and auxiliary experimental data, but at PNNL they want to avoid putting empirical corrections or parameters into the underlying theoretical method. They would like to generate enough data to fill a database, so methods must be tailored to molecule size. For geometries and vibrational frequencies, Hartree-Fock (HF) theory is useful up to about 500 basis functions (20-25 heavy atoms with pVDZ basis), and DFT to twice this size. For single-point energies, MP2/cc-pVTZ is useful to about 1000 basis functions (20-25 heavy atoms), CCSD(T)/cc-pVQZ to about 300-500 basis functions (6-10 heavy atoms), and DFT/cc-pVDZ to about 2500 basis functions (100-150 heavy atoms). At PNNL they have done benchmark thermochemical calculations for small molecules using val-CCSD(T)/cc-pVnZ calculations (through  $n=5$  or 6). They find  $n=4$  good for single bonds but not for double or triple bonds or for polar molecules. Core-valence correlation is usually about 4 kJ/mol. Dr. Dixon pointed out that sometimes the electronic energy calculation is fine but the other parts of the calculation (nuclear motions) are very hard, as for floppy molecules. In these cases, the vibrational zero-point energy (ZPE) and entropy depend strongly upon the calculated vibrational frequencies and anharmonicity constants. The temperature dependencies of the thermodynamic properties are also difficult to calculate in these cases. Some of their very careful benchmark calculations have been for the proton affinities of H<sub>2</sub>O and of NH<sub>3</sub>, and for the binding energy of the water dimer, (H<sub>2</sub>O)<sub>2</sub>. For small basis sets, counterpoise calculations for the basis set superposition error (BSSE) can over-correct. At the opposite extreme, energetics for isodesmic reactions can often be calculated quite accurately using only MP2 theory. This was used (in his prior job at DuPont) to explain the decomposition mechanism of Teflon.

**Tsan Lay**, from **Joe Bozzelli's** group (New Jersey Inst. of Tech.), presented a number of topics in addition to the one in the title. Their work is directed toward combustion intermediates, for which few good experimental data are available. As a result, their database makes use of canonical bond strengths and of entropies and heat capacities calculated semi-empirically using the MOPAC program. For hindered rotors, they can't use Pitzer's tables because of the torsional asymmetry in molecules such as  $\text{CH}_3\text{-CXY-OOH}$ . Instead, they express the torsional potential in a Fourier expansion and diagonalize in a free-rotor basis. Some data for radicals were estimated using "hydrogen atom bond increments," a method that they published (*JPC* 1995, 99, 14514). Data also had to be estimated for many oxacyclic compounds. Formation enthalpies from semi-empirical PM3 calculations were found to correlate linearly with experimental values, so they used the regression line to correct PM3 results. They revised some GA ring corrections and also challenge some  $\Delta_f H^\circ$  estimates made by Dorofeeva, based upon their MO calculations. They are currently trying to develop Benson-type groups that are applicable to transition-states.

**Andrzej Anderko** (OLI Systems Inc., Morris Plains, NJ) presented impressively accurate results of modeling multi-component, heterogeneous systems. [This topic is beyond the expertise of the symposium organizers, so this summary may be deficient.] Wherever their models were compared with experimental observations, the agreement was extremely good. They use a thermodynamic model that includes many non-ideal phases (aqueous/vapor/solid), using a Helgeson-Kirkham-Flowers equation of state. Solvation was treated by a Born model with empirical specific interactions, for a total of seven parameters to describe species including simple ions and hydrolysis products, and properties including partial molar volume and heat capacity. A Bromley-Zemaitis-Pitzer model was used for activity coefficients and was expected to be accurate for dense fluids in the range from -50 to 300 °C, < 30 mol/kg, and < 500 bar. The reference database included thermochemical, thermophysical, and interaction data. Dr. Anderko's first, "simple" example was corrosion in LiBr/ZnBr concentrated solutions used in refrigeration. Corrosion is problematic because  $\text{H}_2$  is evolved, which raises the system pressure and hurts efficiency; they want to find inhibitors. The model system includes simple and complex ions, vapor, aqueous neutral solutes, hydrated and anhydrous solids, and many iron compounds deriving from corrosion. Equilibria are modeled under various conditions (e.g., of pH) to produce real-solution analogs of Pourbaix diagrams. These plots of oxidation potential vs. pH show the origin of the problems and suggest that chromate might inhibit the corrosion. From the modeling,  $\text{Li}_2\text{CrO}_4$  looks good at 300 °C but not at lower temperatures. It's impractical just to raise the pH, as shown by modeling addition of LiOH. Arsenate ( $\text{AsO}_4^{3-}$ ) looks good at low temperature but not above 280 °C. Since only partial solutions have been found, at either low or high temperature, this work is still underway. A more complex application is the hydrothermal synthesis of piezoelectric ceramic powders. They use simulations to investigate various precursors (mostly titanates and zirconates) and reaction conditions (such as  $\text{CO}_2$  contamination). They produce stability/yield diagrams that agree well with experiments and indicate that high pH is needed to reach 99% yield. Switching to more alkaline precursors such as  $\text{ZrO}_2$  or  $\text{Sr}(\text{OH})_2$  avoids the need for added base. They have also modeled even more complex piezoelectrics such as  $\text{PbZr}_x\text{Ti}_{1-x}\text{O}_3$ . [Editorial comment: another use of the term "computational thermochemistry," different from that of this symposium, is in the area of phase equilibria; more information is available from the Centre for Research in Computational Thermochemistry at <http://www.crct.polymtl.ca/>.]

**Roberta Susnow** (Exxon Res. & Eng. Co.) presented a paper somewhat off the topic of the symposium but that generated a fair amount of discussion. She described a program called NETGEN

that is under development to predict an entire chemical kinetic model based upon a small set of elementary reaction types. It's based upon earlier work by Klein at the Univ. of Delaware. The program uses quantitative data from NIST's database and from semi-empirical calculations using MOPAC. Reaction rates are estimated based upon linear free-energy relations as needed. After a discussion of technical details such as convergence criteria and decoupling differential equations, Dr. Susnow showed the results of a simulation of the pyrolysis of ethane. Program behavior is stable to input convergence parameters but better rate constants are needed for quantitative results. Further testing will include non-linear, autocatalytic systems, and a future goal is to include mass transport in the simulations.

**Larry Curtiss** (Argonne Natl. Lab.) described progress made in reply to various criticisms of the G1- and G2-type theories. The reference, experimental data set has been expanded to include much larger molecules, up to six heavy atoms, and to include some atoms from the third row (Ga-Kr). The new set includes 148 molecules. The test reactions include atomization energies, ionization energies, electron affinities, and proton affinities. There are 40 test reactions involving third-row elements. Note that for some atoms (Al, Be, Si) the experimental gas-phase uncertainty in  $\Delta_f H^\circ_{298}$  is  $\geq 4.2$  kJ/mol. G2(MP2) is less expensive than G2 theory, and only slightly less accurate. G2(MP2, SVP) uses only the 6-31G\* basis set for the QCISD(T) calculation and is less expensive than G2(MP2), although of equal accuracy. All  $\Delta_f H^\circ$  calculations are now done at 298 K instead of 0 K but are still based upon atomization energies. Experimental enthalpy differences are used for atoms. Experimental spin-orbit corrections are now included for  $^2\text{P}$  and  $^3\text{P}$  atoms and for  $^2\Pi$  molecules. This correction can be large; for Cl atom it is 3.5 kJ/mol. Seven DFT methods were also evaluated, using the 6-311+G(3df,2p) basis with HF ZPEs and MP2 geometries for a fair comparison of electronic energies. The mean absolute error is higher with the new test set than with the old; it is now 10.5 kJ/mol for the non-hydrides (the worst case). Adjusting the "high-level correction" (HLC) was found to make no improvement. The spin-orbit correction improves the results for chlorides, but for fluorides results are improved for first-row compounds and degraded for second-row compounds. For cyclic molecules the mean absolute error is about 8 kJ/mol. Unsaturated rings such as furan are the most problematic. All acyclic hydrocarbons are well-predicted. Mean absolute errors in kJ/mol were as follow: G2=6.61, G2(MP2)=8.54, G2(MP2, SVP)=8.08, B3LYP=13.1, B3PW91=14.7, B3P86=75.2, BLYP=29.7, BPW91=32.9, BP86=84.5. For DFT calculations, no HLC was applied and the B3 parameters were not re-optimized. Using the smaller 6-31G\* basis with B3LYP gives results about twice as bad. The non-hydrides are still the most problematic for B3LYP. Dr. Curtiss presented error histograms for all the methods investigated. Among the DFT methods, even B3LYP has long tails in its histogram; only about 50% of the data are within 8.4 kJ/mol of the experimental values. For the G2 method the tails are not as bad and 70% of the data lie within 8.4 kJ/mol of the experimental values. [*Editorial comment:* Although the distributions didn't really look Gaussian, we can interpret the results crudely as standard deviations ( $1\sigma$ ) of about 12.4 kJ/mol for B3LYP and 8.1 kJ/mol for G2.]

**Don Rogers** (Long Island Univ.) discussed the reliability and utility of calculating enthalpies of hydrogenation of hydrocarbons. Most such hydrogenations are exothermic; about 500 have been measured with a typical precision of 0.4-0.8 kJ/mol. For example, the stepwise hydrogenation of benzene to cyclohexane has  $\Delta_f H^\circ$  of +23, -112, and -124 kJ/mol. He uses the G2, G2(MP2), and G2(MP2,SVP) calculational methods. Isodesmic reactions typically lead to accurate calculated enthalpies but may have large stoichiometric coefficients that multiply small errors. Single hydrogenation reactions don't have this problem. They do, however, lead to changes in

hybridization, which is not a problem for isodesmic reactions. Networks (“cat’s cradles”) of hypothetical reactions, including isomerizations, can be constructed and used for consistency checks. It turns out that simple hydrogenation enthalpies are calculated accurately but isomerization enthalpies less well. There appears to be a small systematic error for ring-opening reactions. The high quality of the calculated hydrogenation enthalpies is preserved as the molecules become larger (he has gone as large as C<sub>4</sub>, including rings).

**Chul Soo Lee** (Korea Univ.) pointed out that the group methods in his work are not of the Benson type. [This topic is beyond the expertise of the symposium organizers, so this summary may be deficient.] He began with a brief review of the prior work in equations of state, with emphasis on lattice-based theories. Earlier rigid lattice theories (excess Gibbs function models) are by Flory-Huggins (1941) and incorporated in the UNIQUAC program (1976). Later lattice-hole theories are due to Sanchez-Lacombe (1976), Okada-Nose (1981), Panayioutu-Vera (1982), and Kumar et al. (1987). The subsequent development by Victorov-Smirnova (1989) introduced the use of group methods. Prof. Lee’s work develops these methods further (several recent publications). The problem is divided into random and non-random contributions and then solutions are determined in zeroth order (athermal) and first-order (residue). The quasi-chemical solution used is exact only for two-components, considered as a single component and random hole mixture. To achieve practical numerical solutions, the consistency and computational efficiency had to be improved. Surprisingly, the common “two-liquid approximation” works better than the rigorous solution. Pure-component parameters were taken from pVT data, vapor pressure and other data for liquids. Empirical binary parameters were also included. The group contribution method was used to infer parameter values. “Groups” here refer to the number of branches in the molecule (CH<sub>3</sub>, CH<sub>2</sub>, etc.). Prof. Lee has applied the method to many thermophysical properties (e.g., vapor pressure, critical pressure), with good agreement with experiment for the critical properties of hydrocarbons. The theory works well even for critical properties of heavy alkanes, which are usually not well-predicted by lattice theories. Alcohols and ketones were also tested with good results. Predictions for binary mixtures agree with available experimental data and can be used for systems such as C<sub>40</sub>/C<sub>48</sub> mixtures for which no data are available. Parameters are still under development to extend the range of applicability, for example to alkene-alkane systems. Many such systems look good so far, such as heptane-polyethylene densities (liquid and vapor). Bimodal curves also agree well with experiment and are far cheaper than molecular dynamics simulations. Excess enthalpy predictions are sensitive to the temperature-dependence of the parameters, yet results are good for ethanol/MEK mixtures. Enthalpy is not yet incorporated into the lattice model.

**Axel Becke** (Queen’s Univ.) began with a review of Kohn-Sham theory and a discussion of the classes of correlation-exchange functionals. In KS theory, the reference system is of non-interacting electrons in an effective potential  $V_{eff}$  (the KS potential). This is the potential that gives the same total density as in the real, fully-interacting problem. The expression for the energy is thus  $E = T_0 + \int V_{nuc} \rho + \frac{1}{2} \iint \rho(1)\rho(2)/r_{12} + E_{xc}$ . Because of the difference between the reference and real problems, the term  $E_{xc}$  actually includes some kinetic as well as potential energy; this point is often overlooked. If  $h_{xc}(1,2)$  is the exchange-correlation hole as in conventional molecular orbital (MO) theory, then  $E_{xc} = \frac{1}{2} \iint \rho(1)h_{xc}(1,2)/r_{12}$ . The connection between the reference and real systems is made using a coupling-strength parameter  $\lambda$ ,  $h_{xc}(1,2) = \int_0^1 h_{xc}^\lambda(1,2) d\lambda$ . This is known as the “adiabatic connection formula” and gives KS theory its rigor. Since the total density is required to be independent of  $\lambda$ , the one-body potential  $V_\lambda$  must depend upon  $\lambda$ . This integration puts kinetic energy into  $E_{xc}$ . At this point, this is still an *ab initio* theory with no approximations. The

approximations begin when one designs functionals  $E_{xc}[\rho]$ ; the “right”, universally-applicable functional is merely known to exist and has not been identified. We can only test our functionals for universality by trying them on many different chemical systems. The local-density approximation (LDA) uses the exact hole functional from the uniform electron gas, which is a better approximation than it sounds. Generalized gradient approximations (GGAs) depend upon  $\nabla\rho$  as well as  $\rho$  (they are often called “non-local” but that’s a misnomer). Fancier functionals that include dependencies such as  $\nabla^2\rho$  are not much better and generally not used. “Hybrid” functionals combine LDA and GGA and exact exchange (nearly the same as Hartree-Fock exchange) to get the right results in the limit  $\lambda=0$ . A novel idea was to include only a fraction of the exact exchange and led to parametrized methods he described in *JCP* **1993**, *98*, 5648. The best one is often referred to as “ACM” (adiabatic connection model) or “B3P” (Becke-3-parameter plus Perdew). All these methods have been tested on the G2 data set. LDA usually overbinds severely. Mean (and maximum) absolute errors for comparable methods, in kJ/mol, are HF=326 (732), LDA=151 (352), GGA=25 (84) and hybrid=8 (34). Another recent paper, describing a new hybrid functional, is *JCP* **1996**, *104*, 1040. But he doesn’t recommend it, because of a pathology he just discovered for weakly-bound systems (van der Waals molecules)! In his current work, Prof. Becke is screening many hybrid functionals for their performance on properties including activation energies and binding energies of van der Waals complexes; dispersion remains a challenge because the densities of the fragments must overlap. The current functionals include new gradient-correction terms for both exchange and correlation, and a small number of empirical parameters. Unlike the LYP correlation functional, the new functionals are correct in the limit of the uniform electron gas (theoretically satisfying but of unknown relevance to chemical problems). In response to a question, Prof. Becke pointed out that his computer program does not use basis sets, so that the parameters might benefit from re-optimization if they are to be used with small basis sets (as is commonly done). He parametrizes his new functionals on the G2 experimental data set and on the total energies of the atoms H-Ne. Compared with the B3P functional, the new B96 functionals reduce the mean absolute error in atomization energies from 10.0 to 7.9 kJ/mol. Even electron affinities are well-reproduced. So far, reaction barrier heights (activation energies) appear low compared with MP2 and CCSD calculations, but he would like to have an *experimental* data set for activation energies! As shown in *CPL* **1994**, *221*, 100 for the reaction  $H + H_2$ , DFT has the most trouble with light atoms. The new functionals do improve this, and even do well for the dimer  $(HF)_2$ . In summary, Prof. Becke feels that DFT will probably never be as accurate as high-level MO methods, but that it has the big advantage of universality (transition metals, multi-determinant systems, etc.).

**Chris Cramer** (Univ. of Minnesota) described joint research, with his Minnesota colleague Don Truhlar, on modeling solvation energetics quantitatively. Solvation is important in thermodynamics, chemical kinetics, in phenomena such as solvatochromism, and in intermolecular interactions. For example, thiophene is fairly non-polar in the gas phase but quite polar in water ( $\epsilon=78$ ). The alternative treatments are (1) empirical (UNIFAC, LFER, neural networks), (2) explicit inclusion of solvent molecules, and (3) implicit models such as mean-field, bulk dielectric (“reaction field”) models. Their work falls in the third, intermediate category. Most of their underlying calculations are at the semi-empirical AM1 level. Free energy effects are broadly partitioned into electronic and nuclear polarization (ENP) and the rest: cavitation, dispersion and structural effects (CDS),  $\Delta G^\circ_{solv} = \Delta G^\circ_{ENP} + \Delta G^\circ_{CDS}$ . Polarization is treated in the Born model, modified for non-point charges, and is included at the Fock matrix level. Results depend strongly upon the partial charges adopted. The CDS terms are important, but relatively less so for highly-

charged systems. Exact treatment is very hard, so they assume proportionality to the solvent-accessible surface area of the solute molecule and parametrize the microscopic surface tension. The latest model (called "SM5", *JPC* in press) considers entire functional groups instead of just atoms. OSM5 ("organic solvation model no. 5") includes the parameters  $\epsilon$  (dielectric constant),  $n$  (index of refraction),  $\gamma$  (microscopic surface tension), and the Abraham hydrogen-bond acidity and basicity parameters  $\alpha$  and  $\beta$ . For 1784 solute-solvent pairs, the mean absolute error is 2 kJ/mol. They obtained good agreement with big molecular dynamics (MD) simulations for  $\Delta G_{solv}$  of methylated nucleic acids. Good results were also obtained for conformational equilibria and for kinetic solvent effects.

**Butch Carreira** (Univ. of Georgia) described the performance of his computer program SPARC. It uses empirical structure-property relationships based upon a set of molecular descriptors that are tied to observables. Molecular volume is very important in the model, and may be calculated by the program or input by the user. SPARC is able to predict a wide variety of properties with excellent agreement with experiment, including molecular polarizability, bond dipoles, and Taft-Abraham  $\alpha$  and  $\beta$  hydrogen-bonding parameters (including sterics). The  $\alpha/\beta$  calculations are similar to  $pK_a$  calculations in SPARC. Refractive indices are typically correct to about 0.004. Volumes are calculated very accurately (rms error of 1.78) and as a function of temperature. Vapor pressure (log p) looks good, rms error = 0.076 at 25 °C. Activity coefficients depend upon dispersion, induction and dipole interactions (Flory-Huggins). In about 30 solvents, as compared with Eckert's well-screened data set, predicted activity coefficients have an rms error of 0.064. Solubilities (actually quite different from activity coefficients) are not as good for solids as for liquids because it's hard to get the crystal energies right. Solubilities in mixed solvents are done by iterating the activity coefficient to self-consistency. Boiling points at various pressures are good, rms error = 5.7 °C. Prof. Carreira also provided a few details of the underlying (empirical) equations used by the program. For example, a hydrogen-bonding cavity energy is needed to differentiate methanol and decanol. [*Editorial note*: information about the SPARC program is available at <http://www.als.com/nalp/appls/sparc/sparc.html>.]

**Carl Melius** (Sandia Natl. Lab.) presented an overview of the BAC-MP4 and related methods, a selection of some interesting applications, and an indication of problems and future directions. The BAC-MP4 method (bond-additivity corrections applied to fourth-order perturbation theory) is now about 12 years old. It's based upon the idea that systematic errors can be corrected. HF theory is used to calculate geometries and vibrational frequencies and the electronic energy of the molecule is calculated at the MP4 level. It was found that the errors in the bond energies depend exponentially upon the bond length, with smaller contributions from neighboring atoms and from spin-contamination. Hindered-rotor corrections are applied for non-zero temperatures. An error-estimation procedure is built into the program to flag problem molecules. Large databases of thermochemical results (for a few thousand molecules) are available at <http://herzberg.ca.sandia.gov/~melius/>. Transition-state calculations are done to predict rate constants; note that it's the free energy that's important, not the enthalpy. Important results included those for the five-centered decomposition of nitroethane,  $\text{CH}_3\text{CH}_2\text{NO}_2 \rightarrow \text{C}_2\text{H}_4 + \text{HONO}$ , and of trioxane,  $\text{C}_3\text{H}_6\text{O}_3 \rightarrow 3 \text{CH}_2\text{O}$ . Chemical reaction pathways (i.e., reaction mechanisms) can also be predicted, as for aromatic hydrocarbons from the reaction  $\text{C}_3\text{H}_3 + \text{C}_3\text{H}_3$ . Unfortunately, TEOS (tetraethoxy silane or tetraethyl orthosilicate) is too large a molecule (7 heavy atoms is about the limit of the method). [*Editorial comment*: This limit sounds too low; NIST users of the BAC-MP4 method cite a limit of about 12 heavy atoms.] However, the accuracy required depends upon the

question being asked; less-expensive methods are often satisfactory. BAC-MP2 agrees pretty well with BAC-MP4 but can go up to five-ring aromatics. At some point, disagreeing *experimental* values must be questioned. For reactions in large systems such as  $2 \text{C}_5\text{H}_5$  (cyclopentadienyl radical) +  $\text{C}_{10}\text{H}_8$  (naphthalene) *many* calculations are necessary, so each calculation must be fast. In most cases, high accuracy is not as important as getting the chemistry (i.e., reaction mechanism) right. For a Si/B system experimental data were unavailable, so the BAC parameters had to be derived from thermochemical results from G2 calculations (an example of coupling theory with theory, in contrast to coupling theory with experiment). In the decomposition of  $(\text{EtO})_2(\text{MeO})\text{PO}$  to lose ethylene or ethanol, solvent effects and entropic effects are important and must be considered. New BAC methods are currently under development to describe the following: larger molecules, positive and negative ions, extreme oxidation states, heavier elements (even transition metals?), better geometries and frequencies, better vibrational entropies, and better load balancing to avoid computational bottlenecks. Candidates include (MP2 or B3LYP)/6-31++G(2df,2pd)//B3LYP/6-31G\*. For chloromethanes  $\text{CH}_n\text{Cl}_m$ , BAC-MP2 looks a bit better than DFT. Since the transition state of a reaction should be the  $\Delta G$  bottleneck (not  $\Delta H$ ), sometimes HF theory actually gives better results. One problem with the BAC method is that “bonds” must be located. This can cause discontinuities and means that potential energy surfaces can’t be mapped using BAC methods.

**Sherif Kafafi** (Johns Hopkins School of Medicine) described recent progress in parametrized DFT calculations of accurate thermochemistry for about 250 molecules of the first- and second-row elements. The B3LYP method uses a mixture of about 20% HF exchange and 80% Slater exchange energy. Using B3LYP/6-311G\*//6-31G\* atomization energies, he found large errors in  $\Delta_f H^\circ$  (many tens of kJ/mol) that are ascribed to shortcomings in the basis set, not the DFT method. Empirically correcting the atom energies leads to a simple method in excellent agreement with experiment. (For example, the energy for the carbon atom is raised by about 8 kJ/mol.) Good results can be obtained even with a smaller basis set: 3-21+G\* on heavy atoms, 3-21G on H, 3-21+G on H bonded to N, O, F, or Cl. Some molecules, such as sulfur fluorides, give poor results even with the atomic corrections. A big improvement is realized by adjusting the HF/Slater exchange mixture to only 15% HF. Then most of the deviations from experiment are  $\leq 8$  kJ/mol, even for sulfur fluorides and for large molecules such as  $\text{C}_{10}\text{H}_8$ . But oxygen still appears problematic:  $\text{PF}_3\text{O}$ ,  $\text{SO}_2$ , and  $\text{SO}_3$  are off by 25, 25, and 42 kJ/mol respectively. The basis set is apparently the problem here, and the method is under further development.

**Jeffrey Madura** (Univ. of So. Alabama) described applications of computational thermochemistry to bioremediation technology, such as the destruction of dioxins. Anaerobic bacteria degrade such compounds by reductive dechlorination, and he sought to predict the dechlorination pathways using thermochemical data. In particular,  $\Delta_f H^\circ$ ,  $S^\circ$ , vapor pressure, aqueous solubilities, and ionization constants are needed to get the aqueous  $\Delta_f G^\circ$ . Such experimental data are not generally available, so he wanted to estimate at least ideal-gas data and test its predictive value. Since the molecules are large, the methods used were the semi-empirical MNDO, AM1, PM3, and MNDO/d methods and the empirical group additivity schemes in the programs from NIST, CHETAH, Bozzelli, and Shaub. CHETAH seemed to work best where contact with experiments could be made. The other methods were therefore compared with CHETAH for 75 compounds of interest, with mostly constant (systematic) discrepancies. Finally, Prof. Madura found that the dechlorination sequence for a given molecule *can* generally be predicted based upon the calculated  $\Delta G$  (or reduction potential). He speculated that this may be through microbial adaptation to extract the maximum energy content from these “food” molecules.



**Tom Ziegler** (Univ. of Calgary) discussed the application of DFT to understand trends in the bonding of transition metal complexes. He uses the ADF program, which can incorporate relativistic effects at a fairly high level. For bond energies  $MH_3-X$  ( $M=C, Si, Ge, Sn$ ), BP86 gets the trend correct and MP4 does no better; G2 does better but is far more expensive. His explanation of periodic trends is based upon arguments by Kutzelnigg (*Angew. Chem.* 1984) about intra-atomic Pauli repulsion. In the series  $CH_3-X$ , the overlap increases from  $X=F$  to  $I$ , but  $EA(X)$  decreases, so ionic bonding decreases. In the series  $MH_3-Cl$ , the overlap is constant, but  $IP(M)$  drops after  $M=C$ . The  $H_3M-MH_3$  bond energy decreases from  $M=C$  to  $Ge$  because of interatomic core repulsions for the heavier elements. For  $M-H$  and  $M_2$  ( $M=Cu, Ag, Au$ ),  $Ag$  has the weakest bonds (experimentally), and relativistic effects are needed to get the trend correct. The major relativistic effect is the core contraction due to the high-velocity mass increase. For organometallics such as  $(CO)_5Mn-H$ , reasonable agreement with experiment is obtained (see *JACS* 1993), and he expects BP86 results to be within 5 kcal/mol of the correct result (e.g.,  $(CO)_5Mn-Mn(CO)_5$ ). Reliable predictions of trends across the d-block are also expected. The contributions to the bonding can be determined using a decomposition method based upon 4-electron/2-orbital interactions,  $2e/2o$  interactions, and reorganization energies. For the double-bond energies  $H_2M=MH_2$  ( $M=$ group 14:  $C, Si, Ge, Sn, Pb$ ) the trend obtained agrees with prior work. But analysis shows that the ratio  $\Delta E_\sigma/\Delta E_\pi$  is 2.8, 2.4, 2.1, 2.1, and 1.4 across this series, indicating that the  $\pi$ -bond is not solely responsible for the bond-weakening, as is typically argued. The  $\sigma$ -bond also weakens, because of interatomic core repulsion for  $M>C$ , as before. Similar analysis of  $(CO)_4Cr=MH_2$  ( $M=$ group 14) indicates strong  $\pi$ -backbonding to  $C$ , corresponding to a singlet bonding model. For  $(CO)_4M=CH_2$  ( $M=$ group 6) the indirect relativistic effects are important (220 kJ/mol!) for  $W$ . These indirect effects arise from the increase in d-orbital energy that accompanies contraction (and stabilization) of s-orbitals and, to a smaller degree, p-orbitals. The higher d-orbital energy strengthens the  $\pi$ -backbonding. The double-bond energies in  $(CO)_4M=M(CO)_4$  ( $M=$ group 8) show the trend  $Fe < Ru < Os$ . Unlike the main-group elements, the interatomic repulsion is decreasing here. For  $n=3$  ( $Fe$ ), the  $3s-3p$  core is about the same size as the valence  $3d$  orbitals; this is less of a problem for larger  $n$  (i.e.,  $Ru$  and  $Os$ ). For metal carbonyls, the first bond energies  $(CO)_{n-1}M-(CO)$  ( $M=$ groups 6, 8, 10) again need relativity to get the trends correct. As before, the indirect effects raise the d-orbital energies for  $W$  to make it a better  $\pi$ -donor but weaker  $\sigma$ -acceptor. 4d-metals have the weakest bonds to  $CO$  and 3d have the strongest. Without relativity, the usual bond strength trend is  $3d \gg 4d > 5d$ . With relativity, this changes to  $3d \gg 4d < 5d$ . For dihydrogen complexes  $M(L)_3H_2(H_2)$  ( $M=$ group 8) relativity is important for  $Os$ . Prof. Ziegler summarized optimistically, saying that bond energies can be calculated reasonably well even for difficult transition-metal systems, and that the bonding trends can be analyzed using available tools.

**Keith Harrison** (Univ. of So. Alabama) described the popular computer program CHETAH ("chemical thermodynamics and energy release") for property prediction using Benson's group methods. The program is developed and maintained as a volunteer project of ASTM Committee E-27 (hazard potential of chemicals). It was first released in 1974; the current version is number 7.0 and was released in 1994. It is also distributed as NIST Special Database 16. A new version with a graphical interface is now under development. It was developed to predict chemicals' potential for deflagration or detonation before the compounds are actually synthesized. It can predict flammability, enthalpy of combustion, equilibrium constants and combustion products, but its greatest use has been for predicting thermochemical constants ( $\Delta_f H^\circ$ ,  $\Delta G$ ,  $S^\circ$ ,  $C_p$ ). The predictions are mostly for the gas phase, but some solid salts can be accommodated using the Wilcox-

Brandenburg-Bromley group method. Benson's is a second-order group method (neighboring groups are also considered). CHETAH uses Benson's method but with a modified notation, which the user must learn. This bottleneck will be removed with the next, graphical version of the software (similar to current status of the NIST database). Prof. Harrison spent the rest of the time providing examples of applications. For example, requesting a "reactive hazard" evaluation of TNT correctly predicts a high energy release potential, based upon possible decomposition products and also upon a group additivity method for hazard. Hazards of mixtures such as the electropolishing solution water/HClO<sub>4</sub>/acetic anhydride can be predicted, categorizing various mixture ratios as "safe" or "explosive." In this case there are some discrepancies with experiment, illustrating that CHETAH is a good guide but is not completely reliable. Tables of thermodynamics data can be printed in JANAF format. Environmental applications include predicting the bio-degradation pathways of chlorinated aromatics (as described earlier by Jeff Madura). Often, CHETAH's predictions are accurate enough for plant design, since engineers always include safety margins in their designs. The major strength of the program is that values can be obtained rapidly for almost any compound. Although CHETAH has about 725 groups, more than any other similar program, sometimes a needed group is lacking. In this case a backup method is available for ideal-gas  $C_p$ , based upon atom additivity. Alternatively, one can estimate the value or compromise on the secondary, neighboring groups, such as by using a C-(C<sub>b</sub>)<sub>4</sub> group value instead of C-(C<sub>b</sub>)<sub>3</sub>(C). These approximations give reduced accuracy but are often quite good. Also note that distant groups have no effect on enthalpies of reaction, so reactions can often be modeled accurately using analogous systems with different side-chains. Progress is still needed (1) to obtain more Benson groups and (2) to get better predictions for inorganics. For inorganics, only 75 cations and 42 anions are currently available in the program. Sometimes the method gives very poor results (outliers).

**George Petersson** (Wesleyan Univ.) described the latest progress with his analytic basis set extrapolation theory. It's based upon the asymptotic convergence of pair-natural-orbital expansions and is correct to second order. It was first developed for a two-electron atom in the limit of  $Z \rightarrow \infty$ . Surprisingly, this model problem predicted behavior of real systems very well. It's much worse for molecules because one can't ever reach the HF limit. Nonetheless, for reaction energies one generally achieves a four-fold increase in accuracy without the usual factor of 4<sup>6</sup> (= 4096) increase in cost! In practical implementations it's important to balance the computational load to avoid bottlenecks, as Carl Melius mentioned earlier in his talk. There is always a trade-off in the accuracy obtained and the number of heavy atoms that can be accommodated. For example, CBS-4, CBS-Q, and CBS-QCI/APNO are in order of increasing cost. Isodesmic bond additivity corrections help tremendously for levels of CBS-4 and upward. Prof. Petersson also described a new "IRC-max" method for predicting absolute rates for gas-phase reactions. He began by noting that UHF and MP2 barriers and transition-state geometries disagree badly for the reaction  $H_2 + OH \rightarrow H + H_2O$ , but that the reaction paths themselves agree well. The agreement is because motions perpendicular to the reaction coordinate are just like vibrations of stable molecules. So there are great savings by calculating the reaction path (using the common "intrinsic reaction coordinate" method, IRC) at a low level of theory and then using a more expensive level of theory to look on that pathway for the transition state. Since little or no experimental information is available for transition states, he calibrates calculations against transition-state geometries and energies calculated at the CBS-QCI/APNO level by manual optimization. This was done for ten different reactions. The results for geometry and energy are quite good but worse than for stable molecules, as one would expect. There was good convergence as the method level was increased, but with an apparently constant

error in energy. This systematic error is removed by adding another correction of the “HLC” type, to make a total of three corrections: one to get H atom correct, one to get H<sub>2</sub> correct, and now a third one to get H<sub>3</sub> correct. The final method correctly predicts the rates to within a factor of two of experiment, which corresponds to an activation energy accuracy of 1.3 kJ/mol. An Eckart function is used to calculate the transmission coefficient, accounting for tunneling. The width parameter is not very sensitive to the level of theory used. ZPE effects also vary along the IRC and their changes must be included. No curvature (“corner-cutting”) effects were necessary in the variational transition-state (VTST) calculations to reproduce the experimental rates to within experimental error for all the rates calculated.

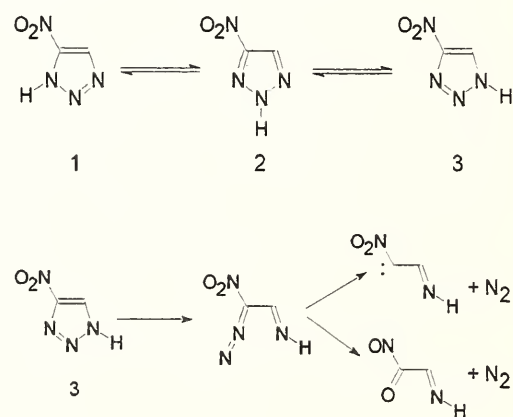
**Bill Schneider** (Ford Motor Co.) presented some of his work on the fate of CF<sub>3</sub> radicals in the troposphere.  $\Delta_f H^\circ(\text{CF}_3)$  can be determined using isogyric reactions and the energetics for the reactions  $\text{CF}_3\text{O}_x + \text{H}_2 \rightarrow \text{CF}_3\text{O}_x\text{H} + \text{H}$  and also  $\text{CF}_3\text{O}_x + \text{H}_2 \rightarrow \text{CF}_3\text{H} + \text{HO}_x$  were computed. Since test calculations for the analogous CH<sub>3</sub>O<sub>x</sub> species gave agreement with experiment of better than 4 kJ/mol (except for 8 kJ/mol for CH<sub>3</sub>O<sub>2</sub>, which may reflect experimental problems), the calculations for the CF<sub>3</sub> analogues were expected to be reliable. But they disagreed strongly with experimental data, especially for CF<sub>3</sub>O (disagreement > 25 kJ/mol). It turned out that the literature bond strength  $D(\text{CF}_3\text{O-H})$  was from a bond-additivity estimation; the ab initio result is much different and implies that CF<sub>3</sub>OH will form easily from CF<sub>3</sub>O + RH or H<sub>2</sub>O. A more recent experimental value appears somewhat too high. Rate constants can be estimated, or at least bounded, by combining calculated equilibrium constants with experimental limits on reverse reaction rates. The O-H bond strengths are 435 kJ/mol in CH<sub>3</sub>OH and 498 kJ/mol in CF<sub>3</sub>OH, despite approximately equal O-H distances. But the C-O distances are quite different and move in opposite directions when the H atom is removed. These differences can be explained readily in terms of negative hyperconjugation; it's important to identify the trends in the data obtained computationally and to understand their origins. Literature data yield a reaction enthalpy  $\Delta_f H^\circ = -15$  kJ/mol for CF<sub>3</sub>OH decomposition to CF<sub>2</sub>O + HF. His revised value for CF<sub>3</sub>OH implies +3 kJ/mol. High-level calculations lead to  $\Delta_f H^\circ = +29$  kJ/mol. The literature value for CF<sub>3</sub>O depended upon three good experiments, but these were anchored upon the value for CF<sub>2</sub>O, which has been significantly revised based upon many calculations. Consequently, values for CF<sub>3</sub>OF and CF<sub>3</sub>OOCF<sub>3</sub> have been revised also. Some radicals are especially challenging computationally. For example, FCO<sub>2</sub> treated at the HF level suffers from symmetry-breaking problems, and vibrational calculations show that the G2 ZPE is wrong by 9.6 kJ/mol. In related work, Dr. Schneider briefly described studies of the decomposition and other reactivity of peroxy radicals, such as CCl<sub>x</sub>H<sub>3-x</sub>O<sub>2</sub> + HO<sub>2</sub>. Unfortunately, he found that there is no relation between the enthalpies of reaction  $\Delta_f H^\circ$  and the observed branching ratios. He concluded with the generalization that discrepancies between theory and experiment are important and must be resolved; experimentalists and theoreticians can learn from each other.

**Michele Foisy**, from **Michael Mavrovouniotis'** group (Northwestern Univ.), presented progress in the estimation of infinite-dilution activity coefficients. The earlier UNIQUAC, “universal quasi-chemical,” method was developed by Abrams and Prausnitz in 1975. A subsequent refinement was UNIFAC, “Uniquac functional group activity coefficients,” which is used for modeling separation processes. It's based on the expression  $\ln \gamma_1^\infty = \ln \gamma_1^C + \ln \gamma_1^R$ , where the superscripts refer to “configuration” and “residual” contributions. The *C* contributions are related to size differences and considerations of the 10 nearest neighbors in a lattice model of the liquid. The *R* contributions reflect differences in intermolecular forces. The group contributions are not of the Benson type but refer to functional groups such as CH<sub>3</sub>, CH<sub>2</sub>, and OH. Some problems

with this group approach are that isomers are not distinguished (since they have the same groups) and that the basis for defining groups is not always clear. The approach at Northwestern is to consider “conjugation operators” that generate all the idealized formal resonance forms of a molecule. For example,  $\text{H-CH}_2\text{-CH}_2\text{-CH}_3 \sim \text{H}^- \text{CH}_2=\text{CH}_2 \text{CH}_3^+$ . Atom and bond contributions are then included for each of the conjugate forms. There are only a small number of adjustable parameters in the method. The user must count the conjugation operators for input to the program. The residual (group interaction) parameters also depend upon the charges in the conjugate forms. Parameters have been developed for alkanes, alcohols, amines and ethers. Initially 18 parameters were fit by regression, but nine were small and were deleted, leaving only nine parameters. Although this modified UNIFAC gave better fits for everything but alcohol-alcohol pairs, their set was too small for them to be confident statistically, and a larger set will be tested in the future. Partial charges are currently estimated by the methods of Constantinou (1994), but a better method is needed, perhaps from ion data. Further work will also extend the method to olefins.

**Peter Politzer** (Univ. of New Orleans) described calculations of the energetics and reaction mechanisms of high-energy molecules. He uses the density-functional method BP86/6-31G\*\* with empirical corrections to estimate  $\Delta_f H^\circ_{298}$ . Generally good agreement with experiment is obtained, despite the large size of the molecules and their heavy nitrogen content. Most of the compounds that he has investigated have not been synthesized; his calculations serve to screen candidate molecules. Solid-phase  $\Delta_f H^\circ$  values are much more useful, but require enthalpies of sublimation. These are estimated pretty well based upon the molecular electrostatic potential on the 0.001 contour surface of the total electron density. About 20 physical properties can be predicted acceptably well, even for large, polar, and unsaturated molecules. For bond strengths, consider  $\text{CH}_3\text{NO}_2 \rightarrow \text{CH}_3 + \text{NO}_2$  as a model. B3 DFT methods work fairly well; B3P86 appears best among them, in comparison with experiment.  $(\text{CH}_3)_2\text{NNO}_2 \rightarrow (\text{CH}_3)_2\text{N} + \text{NO}_2$  is another test, but the magnitude ordering of DFT methods is the opposite of the  $\text{CH}_3\text{NO}_2$  case. For a set of 27 bonds of relevant type, the overall accuracy is  $\text{B3P86} > \text{B3PW91} > \text{B3LYP}$ . Sensitivity is very important for all energetic materials, since it determines whether the time of detonation can be controlled!

Decomposition routes must be determined to address this issue. For example, tautomer 2 is the most stable (at right). But ring-opening of tautomer 3 is the key step in this molecule's decomposition, unlike the typical C-NO<sub>2</sub> bond rupture usually found. This low-energy route explains the high sensitivity of this compound. The decomposition of methyl nitrate ( $\text{CH}_3\text{ONO}_2$ ) is another interesting example. One can imagine three routes: (1) loss of NO<sub>2</sub> followed by isomerization to CH<sub>2</sub>OH, (2) scission to CH<sub>3</sub> + NO<sub>3</sub>, and (3) a cyclic transition state leading to H<sub>2</sub>C=C=O + HONO. The third is the only exothermic route and the computed activation energy is  $E_a \approx 170$  kJ/mol. But this doesn't explain the well-known amine-sensitization of methyl nitrate. One can explain this by noting that its tautomer H<sub>2</sub>C=NO<sub>2</sub>H has an acidity comparable to phenol. Its conjugate base is the nitronate ion, H<sub>2</sub>C=NO<sub>2</sub><sup>-</sup>. If ammonia is the sensitizing amine, it's calculated that displacement of NO<sub>2</sub><sup>-</sup> (to give CH<sub>3</sub>NH<sub>2</sub>) is exoergic,  $\Delta E = -59$  kJ/mol, and that the S<sub>N</sub>2 transition state lies at  $E_a \approx 138$  kJ/mol, much less than the typical C-NO<sub>2</sub> bond strength of 250 kJ/mol. Since the product is another amine, the reaction can propagate and the sensitization appears to be explained.



**Hussein Afeefy** (Univ. of Md. Baltimore Co.) described a new database of organic thermochemistry (no salts or polymers) that is nearing completion at NIST. It will have about 12,000 entries, making it far larger than earlier compilations (e.g., Pedley's contains 3000). Many "hidden" data are being mined successfully from the less well-known literature, and NIST is fortunate to have Russian scientists who can read the Soviet literature critically and who know the authors. The data-gathering also makes use of data collected at non-standard temperatures and also enthalpies of reaction, not just  $\Delta_r H^\circ_{298}$ . Data must be evaluated carefully by reading the primary literature, as illustrated by the recent example of aspirin. In this case, the value in Pedley's compilation disagreed severely with a new measurement made by Duane Kirklin at NIST. It turned out that an auxiliary value, the enthalpy of hydrolysis, had been used with the wrong sign, leading to a gross error. Dr. Afeefy and his colleagues have found more data than they expected, but they need still more. Parameter development for empirical methods (such as Benson's) is hindered by a lack of data and quantum mechanical calculations are not feasible because of the large number of compounds. Data for phase changes are also needed often. He concluded by soliciting contributions of data and by polling the audience on the desire for a hard-copy edition of the database when it is released (several "yes" votes were acknowledged; "no" votes were not polled).

**Manuel Minas da Piedade** (Inst. Sup. Técn., Lisbon) pointed out that entropy can control the course of a reaction but is usually not available experimentally for complicated organometallic systems. Analogies with better-characterized organic systems are often very helpful. For example, there is a linear correlation between the  $\Delta_r H^\circ$ s of  $\text{Cp}_2\text{Zr}(\text{Cl})(\text{OR})$  compounds and the corresponding alcohols ROH. He found 99 reactions for which  $\Delta S$  is available from van't Hoff plots. Estimating all the values using organic analogues yields a mean error of 12 and a maximum error of 58 J/mol.K, not very good. But it works very well in some cases, such as using the analogy  $\text{CH}_3\text{CH}_3 + \text{CO} \rightarrow \text{CH}_3\text{COCH}_3$  to estimate the entropy change for  $\text{Cp}_2\text{HfMe}_2 + \text{CO} \rightarrow \text{Cp}_2\text{Hf}(\text{Me})(\text{COCH}_3)$ . In many of the bad cases there are obvious features that are not well-modeled by organic analogues. For a reduced set of 65 reactions, the average deviation in  $T\Delta S$  is only 4.1 kJ/mol. For the method to work, temperature and solvent effects must be minor. Failure of the method, for example in estimating  $\Delta S$  for the displacement of  $\text{H}_2$  by  $\text{H}_2\text{O}$  in the complex  $\text{W}(\text{CO})_3(\text{PCy}_3)_2(\text{H}_2)$ , can be interpreted as indicating that solvent effects are strong. Prof. Minas da Piedade concluded by stressing that the method must be applied with care but is extremely simple.

**Tim Lee** (NASA Ames Res. Ctr.) pointed out that energetics are sometimes needed with accuracies of  $\pm 2.1$  or even  $\pm 0.4$  kJ/mol for critical data or modeling. Such applications may include combustion, atmospheric chemistry, limiting data for potential energy surfaces, or as anchors for isodesmic reaction schemes when experimental data are lacking. He also noted that calculating atomization energies is very difficult. He uses large correlation-consistent (cc) or atomic natural-orbital (ANO) basis sets in his work. Isodesmic reactions vary in quality, and simple connectivity is not enough. For example, the energy for the reaction  $\text{HOOH} + \text{F}_2\text{O} \rightleftharpoons \text{H}_2\text{O} + \text{FOOF}$  is revealed to be a poor reaction by the criteria (1) large variation in  $\Delta_r E$  with treatment of electron correlation (e.g., MP2 vs. CCSD vs. CCSD(T)), and (2) a value of  $\Delta_r E$  far from zero. Halogen-oxygen bonds are also unusual in that a plot of stretching force constant vs. bond length is not monotonic. The force constant generally decreases with increasing distance but has a double peak, with a dip at small  $r$  corresponding to *cis*-XONO. Well-behaved isodesmic reactions are useful for checking experimental data. For example, the calculated  $\Delta E$  for  $\text{FNO} + \text{ClNO}_2 \rightleftharpoons \text{FNO}_2 + \text{ClNO}$  is 14 kJ/mol but the experimental data imply  $\Delta E = -3$  kJ/mol. Dr. Lee also uses homodesmic reactions, in which bonds involving X are replaced with bonds involving Y. Of course, this is only successful insofar

as X and Y are similar. For example, the reaction  $\text{FNO} + \text{HOCl} \rightleftharpoons \text{ClNO} + \text{HOF}$  is better than  $\text{FNO} + \text{HOH} \rightleftharpoons \text{ClNO} + \text{HOF}$  because there is a closer similarity between F and Cl than between F and H. Unfortunately, even when a discrepancy with experiment is discovered it's not obvious for which molecule(s) the  $\Delta_f H^\circ$  value is in error, and further work is necessary. He finds the semi-empirical scheme of J. M. L. Martin to be useful for calculating atomization energies. This scheme gave accurate results for 14 test molecules and consists of a correction based upon the number of  $\sigma$ -bonds,  $\pi$ -bonds, and electron pairs,  $\Delta E_c = a_\sigma n_\sigma + b_\pi n_\pi + c_{\text{pair}} n_{\text{pair}}$ . When applied to HNO, FCN and ClCN, the corrections are as large as  $\sim 20$  kJ/mol even for spd-fg-quality basis sets! As a check, one can use the exponential basis set extrapolations used at PNNL and based upon Dunning's cc-pVnZ series of basis sets. But the exponential form is merely empirical and lacks physical motivation. A more justifiable form was suggested by Martin based upon Schwartz's observation (in 1963) that the second-order energy depends upon the angular momentum as  $\Delta E(l) \approx A(l+1/2)^{-4} + B(l+1/2)^{-6} + O(l^{-8})$ . The corresponding extrapolation formula is then  $E_{\text{at}} = E_{\text{at}}^\infty + A(l+1/2)^{-4} + B(l+1/2)^{-6}$ . In practice this seems to work better than the exponential extrapolation. Note that the term in  $(l+1/2)^{-6}$  can be omitted, with some loss of accuracy, to reduce the number of required calculations to only two (as opposed to three for the more complete expression or for the exponential expression). Even with extrapolations, further corrections are still needed for core correlation (which is not usually included explicitly because of its cost) and the spin-orbit splitting at the atomic asymptote. Dr. Lee's calculations supported the experimental  $\Delta_f H^\circ$  value for ClCN but not for FCN or HNO. In reply, a new experiment was done for HNO that yielded good agreement with theory. Such "direct" calculations of  $\Delta_f H^\circ$  for small molecules can be done with an accuracy of 2-4 kJ/mol. Extrapolated isodesmic reaction energies are more reliable than atomization energies. The two-parameter Schwartz extrapolation works pretty well. Dr. Lee has revised  $\Delta_f H^\circ(\text{FNO})$  and also done calculations for many bromine compounds, including a revision of  $\Delta_f H^\circ(\text{BrNO})$ . For bromine-containing radicals, the spin-orbit splitting is needed to predict the temperature effects on the thermodynamic functions. He concluded by summarizing his most important conclusions. (1) The best approach is to combine experiment and theory, but this requires accurate experimental reference data. (2) Even when using reaction schemes, it's best to extrapolate to the basis set limit. (3) The "direct" calculation of  $\Delta_f H^\circ$  based upon atomization energies still requires highly-correlated calculations and empirical corrections or basis-set extrapolations. (4) Schwartz extrapolation works better than exponential extrapolation. (5) One can obtain accuracies of  $\pm 2$  kJ/mol.

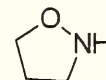
**Rajiv Berry** (Wright-Patterson AFB) described efforts to calculate thermochemistry and rate constants that are needed to model flame suppression by candidate halon replacements. Rate constants are needed as a function of temperature, and many of the important reactions involve H-atom abstraction by OH radicals. The simplest test case is  $\text{CH}_4 + \text{OH} \rightarrow \text{CH}_3 + \text{H}_2\text{O}$ . Using MP2 theory to calculate the height and width of the barrier leads to poor results at low temperature. Instead, it's much better to use Petersson's IRC-max method, which involves an Eckart model potential and yields a much wider barrier. A test on  $\text{CH}_3\text{F} + \text{OH}$  also does quite well, even at low temperature. This includes the 5.2 kJ/mol barrier lowering required to get correct results for  $\text{H}_3$ . Note that the changes in ZPE along the reaction coordinate should also be included. They would like to have rates  $k(T)$  for the reactions of OH radicals with all the  $\text{C}_1$  and  $\text{C}_2$  hydrofluorocarbons (HFCs), as well as the  $\Delta_f H^\circ(T)$  values. A variety of methods, such as G2 and G2(MP2), were tested for their predictions for fluoromethanes,  $\Delta_f H^\circ(\text{CH}_x\text{F}_{4-x})$ . Most had systematic errors proportional to the number of fluorine atoms. This suggested making simple bond-additivity corrections, which turned out to yield very good agreement with experimental data. Fluoroethanes were then tested to

probe the transferability of the  $\Delta_{C-F}$  correction. Results were generally good, but there are some outliers that suggest problems with the experimental data. For the chloromethanes, it was also necessary to include a nearest-neighbor correction as in Melius' BAC-MP4 theory. Such BAC parameters were developed for G2, G2(MP2), and CBS-Q theories. Work on bromine and iodine compounds is in progress.

**Per Siegbahn** (Stockholm Univ.) discussed transition-metal thermochemistry with an emphasis on his PCI-80 approach. For such systems one generally would like the predictions to have errors smaller than about 20 kJ/mol. This appears to have been achieved, for example using the B3LYP, PCI-80, and G2 schemes for the small molecules  $MH^+$  and  $MCH_3^+$ . He noted that it is much harder to predict thermochemistry for first-row (3d) transition metal compounds than for the heavier transition metals. The correlation energy in all cases is larger than might have been expected but is systematic. If one is to treat realistically large systems, a DZP basis is about the limit. It is easy to see that purely *ab initio* approaches are not going to work, and that some kind of empirical correction is needed. The ideas of multiplicative (e.g., Truhlar) and additive (e.g., G2 theory) scaling of correlation effects are old. Accordingly, Prof. Siegbahn finds that *ab initio* calculations typically recover a constant fraction of the correlation effects, often 80%. This is the origin of his method's name, PCI-80. One simply calculates the correlation effect on the energetics of interest and then divides it by 0.80. In some cases, mostly involving very short bonds, a HF-limit correction is also needed. The scaling procedure improves results from a typical error of 80 kJ/mol to one of 8 kJ/mol. Comparison with experiments of Weisshaar and of Armentrout shows that uncorrected MCPF is very poor but that the correction leads to good agreement. Adding more parameters makes little difference so he doesn't bother. Similarly, Herzberg's book includes very difficult molecules but a six-fold improvement is obtained by the scaling procedure. It's surprising that the same scaling factor applies to both transition metal compounds and to methane! The results are pretty stable to changes in basis set and calculational level. The results and philosophy are completely different from G2 theory and resemble DFT more. The second-row (4d) compounds are the easiest to predict. Third-row (5d) compounds are harder because some relativistic effects must be included. He generally treats spin-orbit effects using experimental data for atoms and assuming complete quenching in molecules. First-row (3d) compounds are the hardest because of the severe near-degeneracy effects. Another application is to Schwarz' experimental studies of the  $Pt^+$ -catalyzed oxidation of  $CH_4$  by  $O_2$ . Lots of calculations were required to characterize the catalytic cycle. B3LYP and PCI-80 both worked pretty well, with PCI-80 a little better. Likewise for the first-row  $MH^+$  and  $MCH_3^+$  molecules. For the first-row  $MCH_2^+$  molecules, there appears to be a problem for Cr and maybe V. This is disturbing because the  $T_1$  diagnostic is largest for Mn, not for Cr, and the largest MCSCF and MCPF coefficients don't flag Cr either.  $Cr^+$  does have nearly degenerate  $s^0$  and  $s^1$  states. There may be a similarity with the case  $Ni + CH_4 \rightarrow H-Ni-CH_3$ . There are no experimental data for this reaction, but there are big discrepancies among calculations only for Ni. The situation is even worse for the barrier height. The suspected physical origin of the discrepancies is near-degeneracy among the  $d^8$ ,  $d^9$ , and  $d^{10}$  states of the Ni atom. A related reaction is the insertion of  $CpM(CO)$  into methane ( $M$ =group 9). For the Co compound (triplet ground state), but not for Ir or Rh, there is a big discrepancy between B3LYP and PCI-80. Prof. Siegbahn trusts the B3LYP method more than PCI-80 for such problematic first-row compounds. He summarized by saying that PCI-80 is usually a little more accurate than B3LYP but sometimes fails, whereas B3LYP seems to fail very rarely.

**Andy Holder** (of both Univ. Missouri-Kansas City and Semichem Inc.) described recent progress in extending semi-empirical methods to the transition metals. He began by noting that semi-empirical methods were designed with only s- and p-orbitals in mind, and that a change is needed in the philosophy for parameter generation. Since chemically-relevant quantities such as ionization energies and electron affinities are periodic, semi-empirical quantum mechanical constants should be periodic too. Such quantities include the O'Leary one-center integrals  $G$  and  $H$ , which proliferate badly when d-functions are introduced. In AM1 theory, the  $\zeta_s$  orbital exponents are also periodic. Including gaussian functions disrupts the balance between electronic effects and core repulsion. These gaussians were originally included to fix the results for certain bond types but they are now overused. Sometimes they can even dominate the chemistry, overwhelming the quantum mechanics with empirical functions. Prof. Holder is trying to eliminate these gaussians as much as possible by using a better theoretical treatment. For example, two-electron repulsion integrals are now treated using an *ab initio* formula instead of a multipole expansion. The parameters are split into two sets, with special "atomic zetas"  $\zeta_A$  used for the one-center/two-electron integrals; parameter values are based primarily upon spectroscopic data. For organic compounds, both the mean error and standard deviation are better with the new SAM1 method than with AM1 or PM3 (e.g., 16.6 kJ/mol for  $\Delta_f H^\circ$ ). There are also big improvements in ionization energies (now 0.33 eV). Results for Zn are starting to look much better, but this work is still in progress.

**Dave Frurip** (Dow Chemical Co.) summarized that accurate thermochemistry is needed for the safe and successful scaleup of chemical processes. Hazards such as flammability and energy release must be known and controlled. Process heats are needed to determine engineering design parameters for components such as vessel jackets and heat exchangers. One must always consider the worst-case scenario of adiabatic temperature rise, since the container can be quite large. Both experiments and predictions are used. Experiments are better for complex systems with poorly-defined chemistry, when thermodynamic data are unavailable, or when the *rate* of heat release must be known. Predictions are better when the experiments are hard, for small-scale processes, or when only modest accuracy is needed. Agreement between experiment and theory is always reassuring, and disagreement means that more scrutiny is needed. The first few steps used at Dow for estimating reaction heats are (1) identify the reaction, (2) simplify the reaction by considering analogs that omit unneeded parts of the molecule, and (3) estimate  $\Delta_f H$ . Bond-energy approaches are poor for estimating  $\Delta_f H$  but can be used to estimate  $\Delta_f H$  as long as all the data are from the same source. Benson's group contribution methods are useful for estimating a number of properties. Sometimes groups were unavailable, and *ab initio* theory appeared too cumbersome, expensive and untested. But with the introduction of G2 theory and related methods, quantum mechanical methods are demonstrating cost savings and are becoming "mainstream" methods. For example, a problem could arise if water were added to a big tank as follows:  $i\text{-PrOCOC}l + \text{H}_2\text{O} \rightarrow i\text{-PrOH} + \text{CO}_2 + \text{HCl}$ . A value for  $\Delta_{rx} H$  was needed but the Benson group was missing. G2 theory was invoked for the smaller analog  $\text{HOCOC}l$  to get a group value for  $\text{CO}(\text{Cl},\text{O})$  of -182 kJ/mol. A later experiment was within 20 kJ/mol of the prediction. In another example, an explosion occurred while distilling the compound shown at right. A G2 calculation was done and the results imported into CHETAH, which predicted an energy release potential 80% of that of TNT! At this point they have compared 23 G2/G2(MP2) predictions with (unevaluated) experimental data. G2 is most useful for supplying missing Benson groups. For example, the group value for the central carbon in  $\text{H}_2\text{C}=\text{C}(\text{NH}_2)_2$  derived from G2 calculations agrees with Dr. Frurip's estimate





based upon conventional procedures. He concluded that computational techniques are useful for quick data and that G2 has been successfully “mainstreamed,” but that further development is still needed.

**Klavs Jensen** (MIT) described computational work toward providing thermochemistry for simulations of chemical vapor deposition (CVD). CVD modeling generally involves computational fluid dynamics (CFD) but combining chemistry with CFD is very difficult. Moreover, the required understanding of the chemistry in the gas phase, at the surface, and between the two is often lacking. Both thermochemistry and kinetics are necessary and must be available for the complete reaction model for the simulation to proceed. The accuracy of predictions must be known and sensitivity analysis must be performed. In his work, MP2, MP4, BLYP, and B3LYP quantum methods have been tested on silicon hydrides. B3LYP gives good  $\Delta_f H^\circ$  values (from atomization energies), and applying BAC-type corrections yields acceptable accuracy. For germane,  $\text{GeH}_4$ , B3LYP does well at geometry and vibrational frequencies, so it was presumed reliable for the lower hydrides  $\text{GeH}_x$ . Transition state geometries and activation energies were also calculated for  $\text{GeH}_4$  decomposition. As usual, MP2 energies were high, DFT were lower. The initial  $\text{GeH}_2$  product can insert into  $\text{GeH}_4$  and the metastable adduct decays to form  $\text{Ge}_2\text{H}_4$ ; the chemistry is similar to that of silane. RRKM calculations of the rate of  $\text{GeH}_4$  decomposition, using the UNIMOL program from QCPE, yield rates that agree well with experiment. The rates of surface reactions are very important and are taken from published experimental data. The material growth rate is calculated as a function of temperature, and a sensitivity analysis done. This shows that gas-phase chemistry becomes important for  $T > 1000$  K and that the gas-phase kinetics has a strong effect on the growth rate. They have also studied the mixed systems  $\text{SiH}_4/\text{GeH}_4$  and  $\text{GaMe}_3/\text{AsH}_3/\text{H}_2$ . Arsine decomposes by loss of H atoms or  $\text{H}_2$  molecules. The B3LYP vibrational frequencies of  $\text{AsH}_x$  agree well with those observed in argon matrices. They calculated the Ga- $\text{CH}_3$  bond strengths in  $\text{GaMe}_3$ , which were previously unknown, and found that the second methyl is very weakly bound ( $\text{GaCH}_3$  is a singlet), in contrast to the three similar bond strengths in  $\text{AsMe}_3$ . The decomposition of  $\text{AsH}_3$  was investigated using MP2//HF and B3LYP transition state calculations, with the conclusion that  $\text{AsH} + \text{H}_2$  is the favored route both thermodynamically and kinetically. A few activation energies are critical, but for the remaining processes good thermochemistry is enough. Prof. Jensen’s group has also studied silicon doping of GaAs.  $\text{Si}_2\text{H}_6$  is much better than  $\text{SiH}_4$  for this application. They proposed that  $\text{SiH}_3\text{AsH}_2$  is the important intermediate and calculated the kinetics for  $\text{SiH}_2 + \text{AH}_3$  reactions ( $\text{A}=\text{N,P,As}$ ) by B3LYP. The insertion barrier is calculated to be high for  $\text{NH}_3$ , lower for  $\text{PH}_3$ , and still lower (28 kJ/mol) for  $\text{AsH}_3$ . Once again, it’s important that the accuracy of the calculations be quantified along with calibration methods. The biggest immediate challenge is to develop models for the surface chemistry, such as clusters or slabs; the surface structures and energetics are badly needed.

## POSTER ABSTRACTS

**James Chickos** (Univ. of Missouri-St. Louis). A group additivity method reported recently for the estimation of the total phase change entropy  $\Delta S_{\text{tpce}}$  from the solid at 0 K to the isotropic liquid have been expanded and modified to encompass a broad spectrum of organic compounds that include a variety of compounds of interest because of their environmental and pharmaceutical importance. The current data base includes over a thousand compounds. Fusion enthalpies are calculated from the experimental melting point and  $\Delta S_{\text{tpce}}$ . The standard error associated with the estimation of  $\Delta S_{\text{tpce}}$  is approximately  $15 \text{ J mol}^{-1} \text{ K}^{-1}$ . Details of the estimations as well as a summary of the results will be made available.

**Albert Davydov** (Univ. of Florida). Recently, several breakthroughs have been made in the processing of compound semiconductors into useful devices. Processes used in fabricating such devices involve contact of a liquid or gas phase with the solid compound semiconductor. Since the analysis of these processes often specifies an equilibrium boundary condition, knowledge of the phase diagram is important for thin film and bulk crystal growth. In this paper we report the application of the CALPHAD (calculation of phase diagrams) procedures to the assessment of experimental thermochemical and phase diagram data in binary and multicomponent systems, in which compound semiconductors are formed. Self-consistent model representation of all available thermodynamic properties in these systems gives us the ability to calculate reasonable estimates of missing properties, such as temperatures and entropies of melting for compounds as well as their sublimation parameters. This information can further be used in directing and optimizing growth conditions for device structures.

**Maricel Torrent** (Univ. of Girona, Spain). One of the steps involved in the so-called Dötz reaction concerns the C-C bond formation between an alkyne and an  $\alpha,\beta$ -unsaturated pentacarbonyl carbene complex. The attack of the alkyne is suggested to require a vacant coordination site at the metal, which may be provided by the loss of a CO ligand. In pentacarbonyl compounds, one can distinguish four *cis* CO ligands and one *trans* CO group. The alkyne is assumed to occupy a position *cis* to the carbene ligand; therefore, a preference for the elimination of a *cis* CO ligand is desired. In principle, all five CO groups are expected to go through fast exchange with free CO molecules, but only a facial coordination will let the cocyclization occur. Density functional calculations have been performed for this decarbonylation step, special attention being focused on the energetics involved in each of the five possibilities of the CO loss.



