Improved Models for Thermodynamic Properties of Industrially Important Fluids

Contact: E.W. Lemmon (Physical and Chemical Properties Div., Boulder)

Chemical and Allied Products

Context and Approach:

Equations of state are used to represent the thermodynamic properties of fluids over ranges of state conditions; accurate and consistent information on such properties as density, heat capacities, phase boundaries, compressibility, sound speed, enthalpy, etc. are required to use the fluids in thermodynamic cycles or other applications, for commerce in these fluids, and in their production or use as a feedstock. Beyond the ideal gas law, equations of state for pure fluids have progressed over the years from simple cubic and virial equations of state to Beattie-Bridgeman, Benedict-Webb-Rubin (BWR), modified BWR (mBWR) equations and most recently to fundamental equations of state that are explicit in the Helmholtz energy. Most modern, high-accuracy equations of state for pure fluid properties are explicit in the Helmholtz energy as a function of density and temperature. All single-phase thermodynamic properties can be calculated as derivatives of the Helmholtz energy. During the past 20 years or so, many of the industrial equations of state for fluids such as nitrogen, carbon dioxide, R-134a, and water have been replaced with high-accuracy fundamental equations of state developed from fitting state-of the-art experimental measurements for density, phase equilibrium, heat capacity, and speed of sound. There are about 30 fluids that fit into this category; most of these fluids are fully characterized over the entire fluid surface by extensive experimental measurements. However, many fluids of industrial interest do not have an extensive set of experimental data to facilitate the development of high-accuracy, fundamental equations of state.

Major Accomplishment(s):

Our work here addresses the need for "technical equations of state" suited for advanced technical applications, such as process calculations including caloric properties, for fluids where state-of-the-art measurements are not available and for which very low uncertainties are not required. Unlike highly accurate equations, which generally use 20 to 50 fluid-specific terms to describe densities to the order of (0.01 to 0.1) %, we developed "technical equations" that are shorter and often use fixed functional forms to characterize the fluid properties. Although the new technical equations suffer from a slight loss in accuracy, their shorter forms allow for faster computations--from 2 to 10 times faster depending on the number of terms in other equations and on the use of special terms for increasing the accuracy. In this work, we developed technical equations of state for 20 industrially important fluids, shown in Table 1. For several fluids, the new equations presented here are the first attempt to depict the fluid's properties with multiparameter equations. For other fluids, the new equations replace older ones, many of which showed unphysical behavior such as calculated negative heat capacities at low temperatures, or that exhibited physically impossible behavior in accessible single-phase regions of the fluid surface. Typical

uncertainties of properties calculated with the new equations of state are 0.2 % in density, 1-2 % in heat capacity and liquid phase speed of sound, and 0.2 % in vapor pressure. We also have implemented the correlations into NIST Standard Reference Database 23 (NIST REFPRPOP), where users can easily generate graphical output useful in engineering applications such as the pressure-enthalpy diagram shown in Figure 1.

Impact: These correlations are important to both our infrastructural work (information will be disseminated through such standard reference databases as NIST REFPROP), and to the immediate needs of customers. The article describing the work was the most accessed manuscript for the second quarter of 2006 for the Journal of Chemical and Engineering Data.

Future Plans: The ongoing program of advanced equation of state development will continue to focus on both immediate and specific demands for improved property information for important industrial fluids.

Publications, Patents:

These correlations have been published recently (Lemmon, E.W. and Span, R., "Short Fundamental Equations of State for 20 Industrial Fluids, J. Chem. Eng. Data **2006**; *51*(3); 785-850)

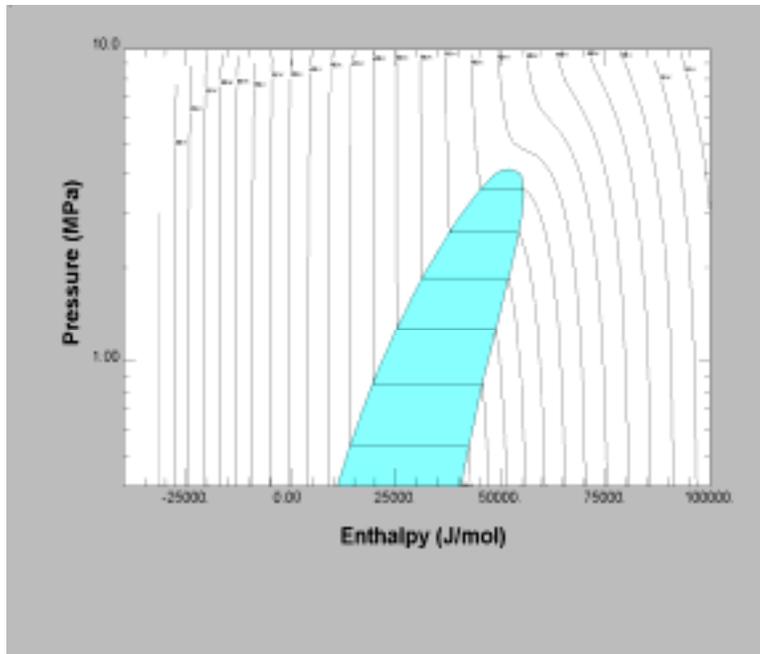| **Table 1. Summary of industrial fluids covered in the present study**			
acetone	carbon monoxide	carbonyl sulfide	decane
hydrogen sulfide	isopentane	neopentane	isohexane
krypton	nitrous oxide	nonane	sulfur dioxide
toluene	xenon	hexafluoroethane	octafluoropropane
R-141b	R-142b	R-245fa	R-41

Figure 1. Pressure-enthalpy diagram for toluene

THERMOInternational 2006 Conference held in Boulder

Contacts: D.G. Friend and M. Frenkel (Physical and Chemical Properties Div., Boulder)

Data and Informatics

A major international forum of about 700 leading technical researchers in thermodynamics and related fields, THERMOInternational 2006, was held at the University of Colorado in Boulder from July 31 to August 4. It was organized and sponsored jointly by NIST, the American Institute of Chemical Engineers (AIChE), the American Society of Mechanical Engineers (ASME), the International Union of Pure and Applied Chemistry (IUPAC), and the International Association of Chemical Thermodynamics. THERMOInternational 2006 comprised three historically distinct conferences: the 16th Symposium on Thermophysical Properties, the 19th International Conference on Chemical Thermodynamics (ICCT), and the 61st Calorimetry Conference. Because each of the conferences was slated for North America this year, they were joined to create the unique event. There were approximately 1000 presentations on the program, representing 750 speakers from 54 countries and 2000 authors from 62 countries.

Much of the research presented at the conference was related to NIST's mission in measurement science, standards, and technology, but the participation was global, with about 75% of the attendees from outside the U.S. The featured work impacts some of the larger themes and policy issues of our time: energy efficiency/self- sufficiency, a hydrogen-based economy and alternative fuels, global warming, ozone depletion, informatics, simulation, biophysics, etc. CSTL staff played prominent roles in the conference, giving many presentations as well as being the primary organizers. Willie May, Director of CSTL, was among those giving opening remarks. William (Mickey) Haynes (Scientist Emeritus in CSTL) headed the organizing committee, and Michael Frenkel (Group Leader in Physical and Chemical Properties Division) and Dan Friend (Acting Chief of the Physical and Chemical Properties Division) served as vice presidents and chaired the distinct components.

The Touloukian Award, a major honor within ASME for outstanding achievement in thermophysical properties research, was presented to NIST Fellow Emeritus Anneke Sengers. The Touloukian Award was awarded "for advancing the theoretical understanding of thermodynamics of pure fluids and mixtures near critical points, and for applying that understanding to improve practical predictions and correlations of the thermodynamic properties for industrial processes and electrical power cycles." Before the award was presented, Nobel Laureate Carl Wieman gave the Touloukian Memorial Lecture, "A Scientific Approach to Teaching Science." The Rossini Award of the Calorimetry Conference was presented to Alex Navrotsky from U. C. Davis, and the Huffman Award, of the ICCT, went to Earl Woolley of BYU.

Advanced Distillation curves for Characterizing Complex Fluids

Contact: T.J. Bruno (Physical and Chemical Properties Div., Boulder)

Energy and Fuels

Context and Approach:

One of the most important properties measured for complex fluid mixtures is the distillation curve. Simply stated, this curve is a graphical depiction of the boiling temperature of a fluid mixture plotted against volume fraction distilled, and such curves are critical in assessing the properties of any complex fluid mixture. It has been impossible to rely on current measurement methods because of large uncertainties and biases, resulting in poor design and specification criteria costing millions of dollars. In response to industry and DOD needs, several significant advances in the measurement of distillation curves were introduced by NIST.

Major Accomplishments:

We have recently introduced an advanced approach to the measurement of the distillation curve that takes the information content to a much higher level than is available with classical approaches[1-8]. The new metrology provides:

- a composition explicit data channel for each distillate fraction (for both qualitative and quantitative analysis),
- temperature measurements that are true thermodynamic state points,
- the energy content of each distillate fraction,
- consistency with a century of historical data,
- trace chemical analysis of each distillate fraction,
- corrosivity assessment of each distillate fraction.

Publications:

1.   Bruno, T.J., *Method and apparatus for precision in-line sampling of distillate.* Sep. Sci. Technol., 2006. 41(2): p. 309-314.
2.   Bruno, T.J., *Improvements in the measurement of distillation curves - part 1: a composition-explicit approach.* Ind. Eng. Chem. Res., 2006. 45: p. 4371-4380.
3.   Bruno, T.J., Smith, B.L., *Improvements in the measurement of distillation curves - part 2: application to aerospace/aviation fuels RP-1 and S-8.* Ind. Eng. Chem. Res., 2006. 45: p. 4381-4388.
4.   Smith, B.L., Bruno, T.J., *Advanced distillation curve measurement with a model predictive temperature controller.* Int. J. Thermophys., 2006. in press**.**
5.   Smith, B.L., Bruno, T.J., *Improvements in the measurement of distillation curves: part 3 - application to gasoline and gasoline + methanol mixtures.* Ind. Eng. Chem. Res., in press.

6.	Smith, B.L., Bruno, T.J., *Improvements in the measurement of distillation curves: part 4- application to the aviation turbine fuel Jet-A.* Ind. Eng. Chem. Res., 2006. in press**.**

7.	Bruno, T.J., Smith, B.L., *Enthalpy of combustion of fuels as a function of distillate cut: application of an advanced distillation curve method.* Energy and Fuels, 2006, 20, 2109-2116**.**

8.	Bruno, T.J., Huber, M.L., Laesecke, A, Lemmon, E.W., Perkins, R.A., *Thermochemical and thermophysical properties of JP-10, NIST-IR 6640, National Institute of Standards and Technology (U.S.),.* 2006.
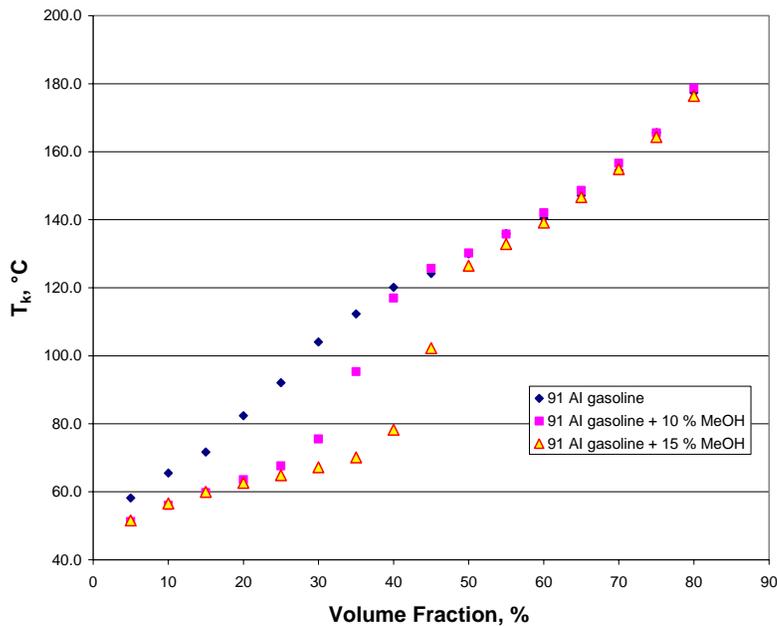
Distillation curve of a 91 AI gasoline and mixtures of the gasoline with methanol.

Property Models and Data in Support of the Hydrogen Economy

Contact: M.L Huber (Physical and Chemical Properties Div., Boulder)

Energy and Environmental Technologies

Abstract:
NIST work on the properties of hydrogen has been incorporated into two new standards that will support the development of hydrogen as a fuel in vehicle applications. A new, simplified equation for the density of hydrogen gas was developed for the draft SAE (Society of Automotive Engineers) procedure J2572. The NIST REFPROP equations for the thermodynamic and transport properties of hydrogen form the basis for the new ASTM International Standard D7265-06. We have also begun an experimental program, which will generate the data necessary to develop improved models for hydrogen and hydrogen-containing mixtures.

Context and Approach:
One of the challenges in the use of hydrogen in vehicle applications is the seemingly trivial matter of measuring fuel consumption. Consumers and industry are accustomed to the high accuracy that is easily achievable by the volumetric metering of gasoline and other liquid fuels. But hydrogen, used either as a compressed gas or a cryogenic liquid, presents significant metering challenges. Vehicle manufacturers have proposed a number of different protocols and algorithms for their hydrogen-fueled vehicles, and the EPA approached NIST to help provide a baseline for an appropriate standard. To evaluate the consumption of gaseous hydrogen fuel in motor vehicles, determination of the temperature and pressure before and after usage within a storage tank of known, and essentially fixed, volume is one of three methods recognized in the SAE draft procedure J2572 "Recommended Practice for Measuring the Fuel Consumption and Range of Fuel Cell Powered Electric and Hybrid Electric Vehicles Using Compressed Gaseous Hydrogen." Compared to the alternative methods of continuously metering the flow of hydrogen to the fuel cell or weighing the fuel tank before and after a test, the use of pressure, volume, and temperature measurements has the potential to be the most robust and economical method in terms of instrumentation costs and ongoing personnel test resources, as well as for measurement precision, repeatability, accuracy, and lab-to-lab reproducibility. The EPA National Vehicle Fuel and Emission Laboratory (NVFEL) is currently evaluating this method for fuel cell (FC) and internal combustion (IC) engines.

Major Accomplishment(s):
In support of SAE J2572, we developed a new equation for hydrogen gas densities using a truncated virial-type equation for the density of hydrogen in terms of pressure and temperature. The new equation reproduces the available experimental data to within 0.2 % (combined uncertainty with a coverage factor of two) and is recommended for use over the temperature range 220 to 400 K at pressures up to 45 MPa. This equation will standardize calculations of hydrogen fuel consumption in vehicles, and represents an initial facet of the NIST effort to help build the infrastructure for a hydrogen-based economy.

In a related effort, the hydrogen property values from the NIST REFPROP database form the basis for the new ASTM International standard D7265-06 "Standard Specification for Hydrogen Thermophysical Property Tables." This standard specification was written by NIST staff who also led the specification through the approval process within ASTM Committee D03 on Gaseous Fuels.

The SAE and ASTM documents (as well as the hydrogen properties in the current version of REFPROP) are based on an existing NIST Standard Reference Data formulation dating from 1982. This formulation is "old-fashioned" by current standards and must certainly be improved, but the immediate needs of industry demanded that we provide interim data. In anticipation of the development of more modern equations of state and property formulations, we have collaborated with researchers from the University of Idaho to survey the available thermophysical properties data for hydrogen over the entire fluid surface (not just the compressed gas). This survey will help to identify regions where additional data may be needed to support new applications.

We also initiated in FY'06 what is expected to be an ongoing experimental effort on hydrogen systems with measurements of the heat capacity and thermal conductivity of a methane plus hydrogen blend. The measurements extended from 140 to 350 K with pressures up to 20 MPa. This blend is used as a fuel, but the primary motivation was to provide data for developing models of hydrogen-containing mixtures.

Impact: The correlations developed this year will satisfy immediate needs of our customers, and will be disseminated through standard reference databases, such as NIST REFPROP. Before this NIST effort, there was no generally accepted industry-wide method of calculation for hydrogen-fueled vehicles.

Future Plans: A significant initiative on "codes and standards for the hydrogen economy" is being considered for funding in the FY'07 NIST budget. Under this initiative, we would significantly expand our experimental efforts on hydrogen and hydrogen-containing mixtures, which would be important in the production of hydrogen (e.g. process streams in coal gasification) or as fuels in their own right (e.g. hydrogen-methane blends). We will develop modern equations of state and other property formulations for hydrogen. We will work to satisfy the immediate data needs of industry, delivering interim data and models when necessary, while working towards the broader, long-term goal of developing comprehensive, high-accuracy property models for hydrogen and hydrogen-containing systems.

Project Team: E.W. Lemmon, M.L. Huber, D.G. Friend, D.G. Archer, M.O. McLinden, R. Perkins, J. Magee, (Physical and Chemical Properties Div.); R.T Jacobsen and J.W. Leachman (University of Idaho); C. Paulina (EPA)

Publications
D7265-06 Standard Specification for Hydrogen Thermophysical Property Tables, ASTM International, 2006.

E. W. Lemmon, M.L. Huber, D.G. Friend, and C. Paulina, "Standardized Equation for Hydrogen Gas Densities for Fuel Consumption", Paper 2006-01-0434, Proceedings SAE World Conference, April 3-6, 2006 Detroit, MI

R.T Jacobsen, J.W. Leachman, S.G. Penoncello, and E.W. Lemmon, "Current Status of Thermodynamic Properties of Hydrogen", Int. J. Thermophysics, submitted Aug 2006.

J.W. Leachman, R.T Jacobsen, S.G. Penoncello, and M.L. Huber, "Current Status of Transport Properties of Hydrogen", Int. J. Thermophysics, submitted Aug 2006.

Quote to highlight in a separate box (assuming the same format as FY'05 TARs):
"Before this NIST effort, there was no industry-wide method of calculating the fuel consumption of hydrogen-fueled vehicles."

Figure 1. "Hydrogen fuel-cell powered vehicle undergoing tests of fuel economy "

Bringing a Greater *in vivo* Relevance to *in vitro* Measurements

Contacts: DG. Archer (838), F.P. Schwarz (831)

**Biosciences; systems biology**

Context and Approach:

Thermodynamic measurements of biochemical reactions are typically conducted in very dilute solutions, but this does not reflect actual *in vivo* conditions where other, nonreacting molecules are present. There are *in vivo* system effects, which influence the thermodynamic properties of biochemical reactions, but which are not quantified at all in most *in vitro* biochemical studies of thermodynamics and kinetics. Providing a framework that allows prediction of the system effects on biochemical reactions is the goal of this project. We have developed new facilities in biological calorimetry to study such "macromolecular crowding phenomena." In FY '06 we set up a new laboratory and carried out initial measurements on the denaturation of lysozyme protein. Addressing these needs is well matched to NIST's expertise in both biochemical measurements and solution physics.

Major Accomplishment(s):

NIST has begun a program to facilitate the use of existing thermodynamic data for prediction of *in vivo* processes. Thermodynamic measurements of biochemical reactions are conducted *in vitro*, in very dilute solutions. Actual biochemical reactions occurring in physiological milieu find themselves in the presence of a very high volume percentage of biomolecules that are non-specific to the reaction being studied, a situation very unlike the *in vitro* conditions of the thermodynamic measurements. Awareness is growing that these additional molecules affect quite significantly the biochemical reactions. The presence of these additional molecules is theoretically predicted to influence the reactions by pushing them towards a state with smaller molecular sizes. In other words, a chemical reaction that reduces significantly the size of the product molecules, compared to the reactants, should be pushed further towards completion by the presence of a large volume fraction of non-reacting biomolecules than would be the same reaction in a very dilute solution, as in the *in vitro* measurement. Similarly, reactions that increase the molecular sizes of products over that of reactants should be inhibited by the presence of the additional, non-reacting biomolecules. These general effects are referred to as "macromolecular crowding."

Systematic studies are essential to improve and test theories of the effects of "crowding" on biochemical reactions and to facilitate the use of the large archive of biochemical data. The Physical and Chemical Properties Division has developed new facilities in biological calorimetry and has initiated experimental studies on macromolecular crowding phenomena. The Division has obtained a biological differential scanning calorimeter and an isothermal titration calorimeter, and has placed those instruments into operation. They are now being used to answer fundamental questions about macromolecular crowding. Calorimetry is an essential tool in separating entropic effects from specific energetic (enthalpic) effects. Calorimetry can usually provide more accurately the entropy and enthalpy of a reaction than is obtained by other means. As such, calorimetry can uncover smaller specific enthalpies of interaction – that invalidate the application of theory – than can other measures of biochemical reactions.

There are several open questions regarding the measurements of "crowding" in the first place.
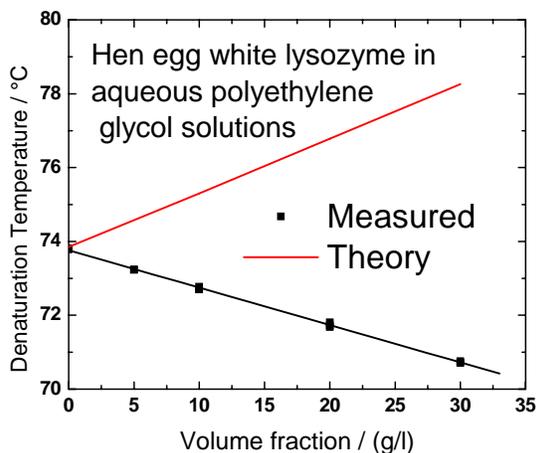
One of these is the concept of a perfect simulator molecule that can mimic physiological crowding. A molecule that simulates crowding should have no specific interaction with any of the components of the biochemical reaction that is being measured. Several different materials have been used to mimic physiological crowding. One of the most popular crowding simulators is polyethylene glycol, which has been used in more than half of the biochemical studies investigating the effects of crowding on individual biochemical reactions. We examined the effect of polyethylene glycol on the thermal denaturation of a protein, lysozyme. Denaturation of a protein occurs often with a significant change in the protein's gyration radius and therefore should be affected by macromolecular crowding. The temperature at which the protein denatured decreased with increasing concentration of polyethylene glycol, completely opposite to the prediction of theory (see figure). This result shows that polyethylene glycol is not free of interaction with at least some proteins and so it may be of questionable use as a crowding simulator. Studies of the effect of other crowding simulators on protein stability will be pursued in FY07.

Advances in Transport Property Modeling

Contacts: A. Laesecke and M.L. Huber (Physical and Chemical Properties Div., Boulder)

Chemical and Allied Products (Chem)

Context and Approach:
Methanol (CH_3-OH) is a widely used fluid in the chemical and process industries. It is also an important compound for healthcare as well as medical and pharmaceutical applications. The oldest use of methanol is in the conversion of biomass. This process is gaining importance because it produces a fuel that does not cause a net increase of carbon dioxide in the Earth's atmosphere. Even more prominent is the role of methanol as a hydrogen-rich fuel for electrochemical energy converters such as fuel cells. The development of accurate thermophysical property formulations for methanol will aid engineers involved in process design in these and other fields. In addition, since it is the first member of the homologous series of alkanols, its physical properties will help to characterize the properties of the series as a whole. Methanol is one of the most polar molecules, and its size, shape, and charge distribution determine its macroscopic properties. Although a reference-quality equation of state has been developed for the thermodynamic properties of methanol, a comprehensive study of the viscosity of methanol has not yet been carried out. This work fills this gap by selecting the most reliable measurements as the basis for a new reference correlation for the viscosity of methanol that is valid over the entire fluid region for vapor, liquid, and supercritical states. Our approach is to use an advanced residual concept for the correlation where the viscosity is expressed as a function of density and temperature that contains a zero-density limit term, a linear-in-density term and a third virial coefficient for the quadratic density term, besides higher-density terms for the compressed fluid region. The objective of this work is to apply kinetic theory to the dilute gas, Rainwater-Friend theory and the third viscosity virial coefficient for the moderately dense gas, and the Enskog dense hard-sphere theory to obtain a correlation for the viscosity of methanol for the entire fluid state that reproduces the most reliable data sets to within their estimated uncertainties and describes the phenomenological behavior of the viscosity of methanol from the triple point to 630 K at pressures up to 8 GPa.

Major Accomplishment(s):
Literature data sources for the viscosity of methanol were collected to a depth that exceeds all previous compilations. More than 250 references were located and screened for viscosity data, the earliest results having been published in 1861. With this many studies, methanol is after water and ethanol the third most measured liquid in viscometry. The temperatures of all data were converted to the ITS-90 temperature scale and their densities were calculated with an equation of state. Thus, all viscosity data of methanol have been transposed on a common basis to make them comparable. This transposition made it also possible to assess the uncertainty of all data even in the numerous cases where no uncertainty was reported in original accounts.

This body of data and assigned uncertainties was used here to develop a new reference-quality correlation for the viscosity of methanol that is valid over the entire fluid region, including vapor, liquid and metastable phases. To describe the zero-density viscosity with kinetic theory for polar gases, a new expression for the collision integral of the Stockmayer potential was introduced. The initial density dependence was based on the Rainwater-Friend theory. For the first time, a correlation for the temperature dependence of the third viscosity virial coefficient was developed from experimental data and applied in such a viscosity model. The high-density contribution to the viscosity was based on the Chapman-Enskog theory and includes a new expression for the hard-sphere diameter that is a function of both temperature and density. The resulting correlation is applicable for temperatures from the triple point to 630 K at pressures up to 8 GPa. The estimated uncertainty of the resulting correlation (with a coverage

factor of two) varies from 0.6 % in the dilute-gas phase between room temperature and 630 K, to less than 2 % for the liquid phase at pressures up to 30 MPa at temperatures between 273 and 343 K, 3 % for pressures from 30 to 100 MPa, 5 % for the liquid from 100 to 500 MPa, and 10 % between 500 MPa and 4 GPa. At very high pressures, from 4 to 8 GPa, the correlation has an estimated uncertainty of 30 % and can be used to indicate the qualitative behavior of the viscosity of methanol.

Impact: Correlations of this type strengthen our knowledge infrastructure of thermophysical properties (disseminated through such standard reference databases as NIST REFPROP) and the worldwide lead of NIST in this area. They also serve the immediate needs of customers. Because of its inclusiveness of the available data, their thorough uncertainty assessment, and the advanced modeling methodology, the correlation is a standard reference benchmark in various aspects. It directs experimenters in the planning of additional viscosity measurements to close data gaps or reconcile differing results for methanol. It provides also orientation to plan viscosity measurements of other fluids that are chemically similar to methanol. The novel components of the viscosity model are transferable to viscosity correlations for other fluids thus raising the level of viscosity correlation methodology in general.

Future Plans: The ongoing program of integrated modeling and experimental work will continue to focus on both immediate and specific demands for transport property information and on longer-term efforts to improve our predictive capabilities of important industrial fluids.

Publications, Patents : So far, one publication in an archival journal resulted from this work:
H.W. Xiang, A. Laesecke and M.L. Huber, "A New Reference Correlation for the Viscosity of Methanol", J. Phys. Chem. Ref. Data 2006, 35(4):1597-1620.
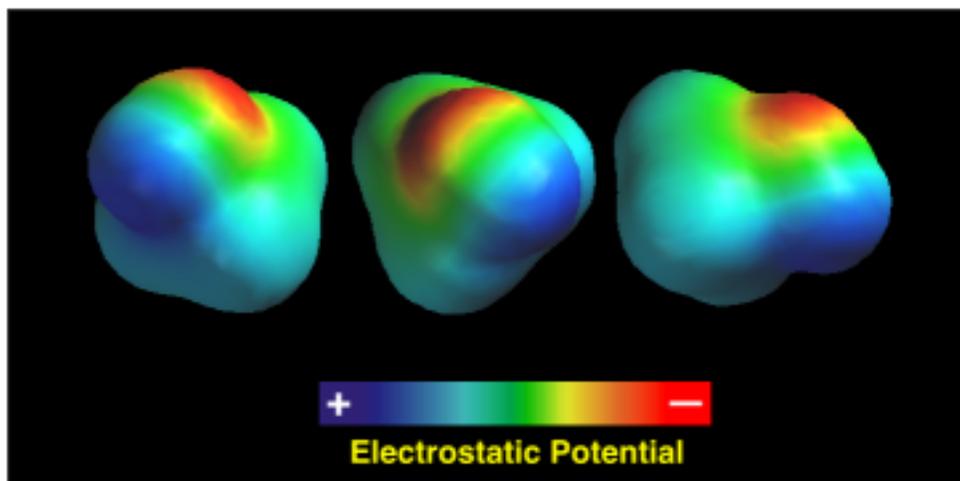

Figure 1:
Visualization of the methanol molecule in terms of an isosurface of the electron density with the electrostatic potential indicating the charge distribution and polarity color-mapped onto it. The chosen electron density level of 0.22×10^{-10} $C \cdot m^{-3}$ (0.002 $e^- \cdot bohr^{-3}$) represents about 98% of the molecule. The molecule is shown in three orientations to give a better impression of its size, shape, charge distribution, and polar centers.

Analysis of Force Transmission Errors in Magnetic Suspension Densimeters

Instruments employing the Archimedes (bouyancy) technique in combination with a magnetic suspension coupling for the measurement of fluid density are capable of very high accuracies, but they are subject to systematic uncertainties—known as force transmission errors—due to the effects of magnetic materials, including the magnetic characteristics of the fluid being measured. A new analysis of force transmission errors in magnetic suspension densimeters has been carried out, significantly reducing the measurement uncertainties in this type of instrument.

M.O. McLinden (838); R. Kleinrahm and W. Wagner (Ruhr Universität Bochum, Germany)

The pressure-density-temperature (p-ρ-T) properties of fluids are key thermophysical property data needed for the development of accurate equations of state. For measurements over wide ranges of temperature, pressure, and density, instruments combining the Archimedes technique with a magnetic suspension coupling have been highly successful in recent years. The Archimedes technique involves weighing a "sinker" of known volume while it is immersed in the fluid of interest; the difference between the apparant mass of the immersed sinker and its true mass yields the fluid. The magnetic suspension coupling transmits the weight of the sinkers to the balance, and separates the fluid, which may be at extremes of temperature and/or pressure, from the balance, which is at ambient conditions. It consists of two magnets on opposite sides of a pressure barrier: an electromagnet hung from the balance and a permanent magnet, from which the sinker is suspended, that is immersed in the fluid.

The efficiency of the force transmission across the magnetic coupling is nearly one, but the coupling will be slightly influenced by nearby magnetic materials, external magnetic fields, and the fluid being measured. These give rise to a "force transmission error" (FTE) that must be accounted for to realize the full accuracy of this technique. The FTE can be divided into two parts. The first arises from the magnetic characteristics of the apparatus itself; it is determined by comparing the known mass of the sinker with a weighing of the sinker using the magnetic coupling. The apparatus effect is well-known, and this correction is routinely made. The magnetic susceptibility of the fluid surrounding the magnets also influences the FTE, but

these effects are often ignored. Even when this "fluid specific effect" is included, the details of the correction have, to date, been murky, at best.

In a paper presented at the 16th Symposium on Thermophysical Properties (held August, 2006 in Boulder) and submitted to the *International Journal of Thermophysics*, NIST, together with researchers from Ruhr Universität in Bochum, Germany, has presented a thorough analysis of the force transmission error. The analysis introduces a parameter termed the "coupling factor", which characterizes the efficiency of the force transmission by the magnetic coupling. The NIST densimeter has two sinkers (of similar mass, but different volumes) and two external masses, which are used to calibrate the balance. Thus, a density determination involves four separate weighings. The forces on the balance for each of the weighings are written out in detail to yield a system of four equations in four unknowns that are solved for the fluid density, a balance calibration factor, a parameter related to the balance loads which do not change, and the coupling factor, which is directly realted to the force transmission error. The data from measurements of seven different fluids show that the coupling factor is the sum of an apparatus contribution (which is nearly constant) and a fluid-specific contribution, which is proportional to the fluid density and the magnetic susceptibility of the fluid.

The NIST densimeter is the only two-sinker densimeter in the world that has the external masses that allow the determination of the force transmission error at each measured density. There are only a handful other two-sinker densimeters, and these instruments have typically compensated only apparatus effect and have ignored the effects of the fluid. The fluid effect was found to be on the order of 20 ppm for most fluids, but it was as high as 800 ppm for air at high pressures. (While most fluids are weakly diamagnetic, the oxygen in air is strongly paramagnetic.) With the new analysis, the force transmission error is completely compensated for with an uncertainty of about 5 ppm.

The more common type of magnetic suspension densimeter has only a single sinker. This type of instrument does not benefit from some of the error-canceling effects present in two-sinker instruments, and force transmission errors are about an order of magnitude higher. But, the simulataneous analysis of

FTE is not possible—there are not enough equations to solve for the unknowns. Our analysis also considered single-sinker instruments. First, data from the NIST two-sinker densimeter were reanalyzed as a virtual one-sinker instrument (*i.e.* the weighing data for one of the sinkers were ignored); by comparing the resulting "single-sinker densities" to the two-sinker results, the magnitude of the force transmission errors were quantified. Then, we demonstrated how an apparatus constant (determined by comparing a set of measurements with a similar set made with a different sinker installed), combined with knowledge of the magnetic susceptibility of the fluid could be used to correct for the FTE in single-sinker densimeters. This allows single-sinker instruments to approach the accuracy of a two-sinker densimeter.

Sinkers and magnetic suspension coupling at the heart of the NIST densimeter. The electromagnet at the top hangs from the balance and levitates the permanent magnet below it. In this photo, the top sinker is being weighed, while the bottom sinker sits on its rest. The permanent magnet and sinkers are normally contained within a pressure vessel.

Properties of Real Fuels for Aviation, Aerospace and Vehicular Transportation

Contact: T.J. Bruno (Physical and Chemical Properties Div., Boulder)

Energy and Fuels

Context and Approach:
Liquid fuels have long been the most convenient on-board fuel source for all sectors of transportation. This includes aircraft, rockets, and land based vehicles (such as cars, trucks, locomotives and military vehicles). The design and specification of such fuels includes engine operability, environmental and supply considerations. For example, in recent years, oxygenates have been added to diesel fuel (to reduce particulates released into the environment), and aviation fuels have been formulated from coal and natural gas (to avoid supply problems). All of these reformulations change the properties of these fuels, and as such these fluids cannot be used without an understanding of how the physical and chemical properties are affected. A major effort is underway to experimentally characterize representative fluids, and to model them with an accurate equation of state.

Technical Approach:
A cooperative effort of experimental measurements and modeling is applied to the study of fuels. Typically, the measurements include density, viscosity and thermal conductivity (as a function of temperature and pressure), and the advanced distillation curve and speed of sound (at ambient pressure). The density is measured with several vibrating tube instruments that can measure at ambient pressure and in the compressed liquid region. We also have a dual sinker densimeter for reference quality measurements. The viscosity is measured with a Stabinger viscometer at ambient pressure, while the compressed liquid viscosities are measured with a torsional crystal viscometer. For the speed of sound a pulse-echo device is used at ambient pressure. The thermal conductivity is measured with a transient hot wire instrument. The distillation curves are measured with a new approach that provides a composition-explicit data channel in addition to the usual temperature and volume channels.

Our recent comprehensive measurements have included the missile fuel, JP-10, and a mixture of aviation fuels Jet-A and S-8 (the later being a fuel made from natural gas by the Fischer Tropsch process)[1, 2].

Publications:
1.   Bruno, T.J., Huber, M.L., Laesecke, A, Lemmon, E.W., Perkins, R.A., *Thermochemical and thermophysical properties of JP-10, NIST-IR 6640, National Institute of Standards and Technology (U.S.),.* 2006.

2.   Bruno, T.J., Arno Laesecke, Stephanie L. Outcalt, Hans-Dieter Seelig, Beverly L. Smith, *Properties of a 50/50 Mixture of Jet-A + S-8, NIST-IR, in press.* 2006.

Project Team:
T.J. Bruno, M.L. Huber, A. Laesecke, E.W. Lemmon, S.L. Outcalt, R.D. Perkins, H.-D. Sealing, B.L. Smith, Div. 838
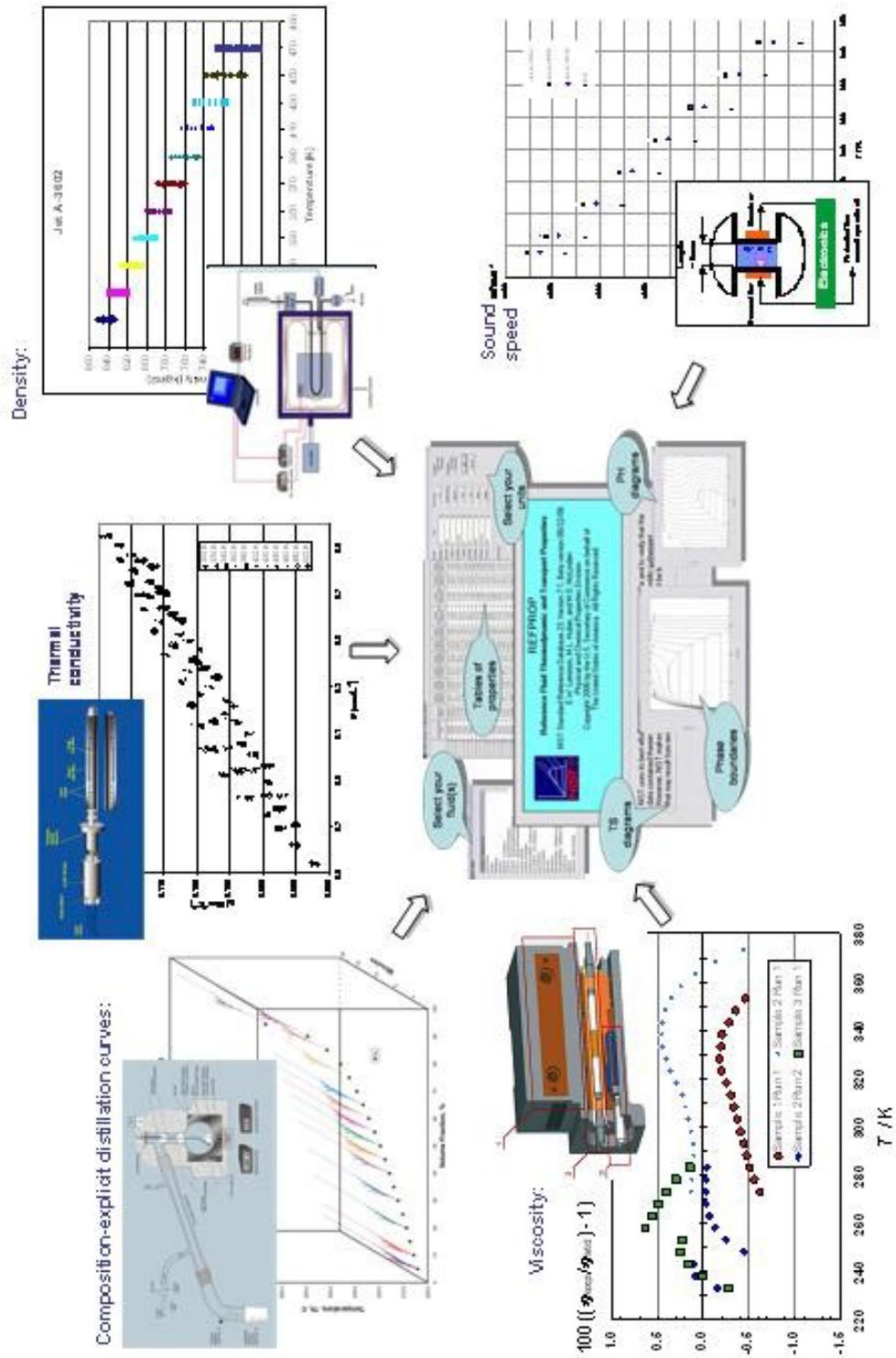
A schematic depiction of how experimental inputs are used to obtain a mathematical model that describes the properties of a fluid, accessible with a user-friendly standard reference database.

Thermodynamic and Transport Properties Database for Ionic Liquids:
ILTHERMO

Contact: Q. Dong (Physical and Chemical Properties Div., Gaithersburg)

Data and Informatics

Context and Approach:
There has been an explosion of interest in the last few years (including a near-exponential growth in journal articles) in the synthesis and use of ionic liquids – salts that melt below the boiling point of water. With typical vapor pressures in the range of 10^{-10} Pa, ionic liquids have essentially no vapor emissions, and so look like excellent candidates for "green solvents" to replace hazardous, air-polluting organic solvents such as acetone and benzene. With dozens of anions and cations to choose from, they can be tailored to specific needs and may be particularly useful as solvents for biocatalysis. Progress in utilizing ionic liquids has been hampered by the lack of an open-access, public-domain, comprehensive data retrieval system scoped to cover information pertaining to ionic liquids. Development of such a database infrastructure encompasses a number of complex issues related to data submission, processing, mining, quality control, management, critical evaluation, and dissemination. In 2003, an International Union of Pure and Applied Chemistry established its Task Group to conduct the project 2003-020-2-100 "Ionic Liquids Database" with an objective to "create an open-access, free, on-line, comprehensive database for storage and retrieval of metadata and numerical data for ionic liquids, including their syntheses, structure, properties, and uses; lack of this information is impeding progress in a burgeoning field of significant current interest".

Major Accomplishment(s):
In July 2006, ILThermo (NIST Standard Reference Database 147) was released to the public as a result of the development initiated at NIST in close cooperation with the IUPAC. ILThermo is a free web research tool that allows users worldwide to access an up-to-date data collection from the publications on experimental studies of thermodynamic and transport properties of ionic liquids, as well as binary and ternary mixtures containing ionic liquids. The inaugural version contains information on more than 200 ions and 300 ionic liquids. The collected data cover the relevant literature from 1982 to 2006 for pure ionic liquids as well as for binary and ternary systems containing ionic liquids. The experimental data stored in the database include those for phase transitions, transport, volumetric, and thermal properties as well as electrical conductivity, surface tension, refractive index, speed of sound, vapor pressure, and activity coefficients. The database presently contains more than 18000 experimental data points. The information available from ILThermo includes chemical identification, numerical property values, details of experimental methods, sample purity, and uncertainty. The database is made available to the public with the use of a Web-Oracle data dissemination system recently established at NIST. In order to increase efficiency and enforce strict data quality control for the data collection process of the thermophysical properties of ionic liquids, the NIST TRC Group modified its Guided Data Capture software to incorporate the elements of metadata specifically important for

ionic liquids. The data generated by ILThermo can easily be downloaded from the Web and propagated into a variety of the chemical engineering software applications.

Project Team: A. F. Kazakov, C. Muzny, R. D. Chirico, J. A. Widegren, V. Diky, J. W. Magee, M. Frenkel (Physical and Chemical Properties Div., Boulder); K. N. Marsh (University of Canterbury, Christchurch, New Zealand)

ThermoData Engine: New Generation Expert System for Thermodynamic Data Critical Evaluation

Contact: M. Frenkel (Physical and Chemical Properties Div., Boulder)

Data and Informatics

Context and Approach:
Traditionally, critical data evaluation is an extremely time- and resource-consuming process, which includes extensive use of labor in data collection, data mining, analysis, fitting, *etc*. Because of this, it must be performed far in advance of a need within an industrial or scientific application. In addition, it is quite common that by the time the critical data-evaluation process for a particular chemical system or property group is complete (sometimes after years of data evaluation projects involving highly skilled data experts), it must be reinitiated because significant new data have become available. This type of slow and inflexible critical data evaluation can be defined as 'static.' These shortcomings have become magnified dramatically within the last 5 to 10 years due to the significant increase in the rate of publication of experimental and predicted thermodynamic data which need to be analyzed during the critical data evaluation process.

Major Accomplishment(s):
To address the weaknesses of 'static" evaluations, the concept of a dynamic data evaluation system was developed at the NIST TRC Group. This concept requires large electronic databases capable of storing essentially all experimental data known to date with detailed descriptions of relevant metadata and uncertainties. The combination of these electronic databases with expert-system software, designed to automatically generate recommended data based on available experimental data, leads to the ability to produce critically evaluated data dynamically or 'to order' (see figure 1). This concept contrasts sharply with static critical data evaluation, which must be initiated far in advance of a particular need. The dynamic data evaluation process dramatically reduces the effort and costs associated with anticipating future needs and keeping static evaluations current.

Critically evaluated data produced by the deployment of the dynamic data-evaluation concept can rigorously be characterized with their quality assessments. That, in turn, provides the ability to propagate reliable data-quality limits to all aspects of chemical process design. Implementation of the dynamic data evaluation concept consists of the solution of a number of major tasks: (1) design and development of a comprehensive database system structure based on the principles of physical chemistry and capable of supporting a large-scale data entry operation for the complete set of thermophysical, thermochemical, and transport properties for chemical systems including pure compounds, binary mixtures, ternary mixtures, and chemical reactions; (2) development of software tools for automation of the data-entry process with robust and internally-consistent mechanisms for automatic assessments of data uncertainty; (3) design and development of algorithms and software tools to assure quality control at all stages of

data entry and analysis; (4) development of algorithms and computer codes to implement the stages of the dynamic data-evaluation concept; (5) development of algorithms to implement, target, and apply prediction methods depending on the nature of the chemical system and property, including automatic chemical structure recognition mechanisms; and (6) development of procedures allowing generation of output in a format suitable for application in major commercial simulation engines for chemical-process design.

The ThermoData Engine (TDE) software incorporates all major stages of the concept implementation including data retrieval, grouping, normalization, sorting, consistency enforcement, fitting, and prediction. The SOURCE data system is used in conjunction with TDE as the comprehensive storage facility for experimental thermophysical and thermochemical property data. In addition the NIST/TRC Ideal Gas Database is used as a source of thermodynamic property data in the ideal-gas state.

In FY05, version 2.0 was released *via* NIST Standard Reference Data Office (NIST Standard Reference Database 103). This version is limited to thermodynamic properties of pure compounds. The software architecture emphasizes enforcement of consistency

Fig. 1. Functional comparison of static and dynamic data evaluation

between related properties (including those obtained from predictions), assumes an imperfect source of original data, provides for flexibility in selection of default data models depending on the particular data scenario, incorporates a large variety of models for secondary fitting, and allows saving of critically evaluated data in the ThermoML format. The latter assures compatibility of the TDE software with any engineering application equipped with a ThermoML software 'reader.' New version has two new principal features: (1) the ability to generate equations of state (EOS) on-demand depending on the data scenario (there are 4 types of EOS incorporated) and (2) provision for periodic updates of the local TDE-SOURCE database that "feeds" TDE via the Web.
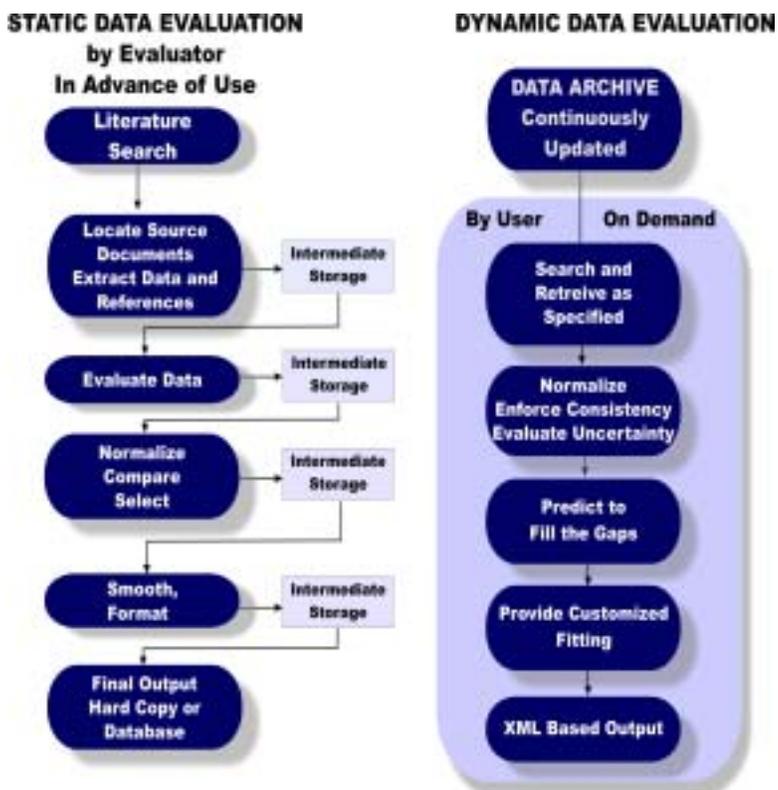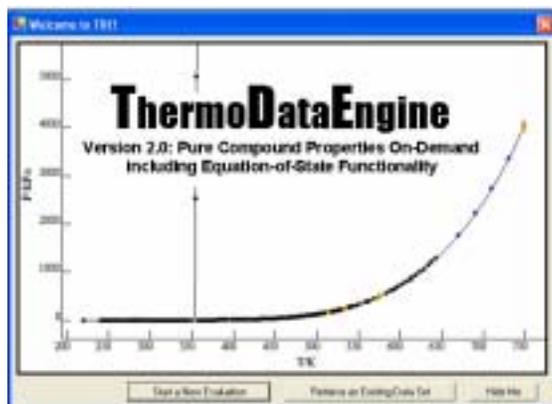
NIST has reached an agreement with a number of industrial organizations about 'bundling' TDE, version 2.0, to their chemical process engineering software, so that TDE would serve as a source of critically evaluated thermophysical property data for pure compounds in the 'bundled' products (see figure 3) providing principally new opportunities of chemical process design on-demand. One example of such cooperation is a new technology incorporated in the 2006 version of the ASPEN PLUS^{TM} process simulation engine developed as a result of the cooperative effort between NIST and Aspen Technology, Inc. In 2006 this technology has become available to 45,000 corporate users. It is currently planned to further expand this cooperation to other industrial and governmental organizations.
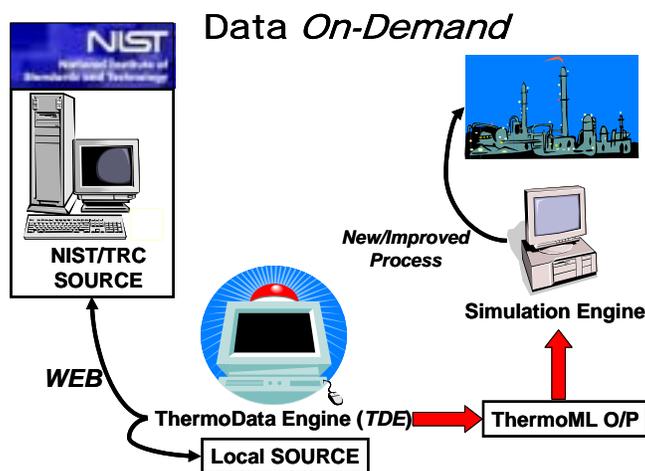
It is anticipated that the TDE 3.0 beta-version will be developed in FY07. This version will include incorporation expansion of TDE to critical data evaluation for binary mixtures as well thermodynamic properties of the formation reactions.

Project Team: M. Frenkel, R. D. Chirico, V. Diky, C. Muzny, E. W. Lemmon, (838-B); X. Yan, Q. Dong (838-G)

ThermoML: New IUPAC Standard for Thermodynamic Data Communications

Contact: M. Frenkel (Physical and Chemical Properties Div., Boulder)

Data and Informatics

Context and Approach:
Thermodynamic property data represent a key resource for development and improvement of all chemical process technologies. However, rapid growth in the number of custom-designed software tools for engineering applications has created an interoperability problem between the formats and structures of thermodynamic data files and required input/output structures for the software applications. Establishment of efficient means for communication of thermodynamic data is absolutely critical for provision of solutions to such technological challenges as elimination of data processing redundancies and data collection process duplication, creation of comprehensive data storage facilities, and rapid data propagation from measurement to data-management system and from data-management system to engineering application. Taking into account the diversity of thermodynamic data and numerous methods of their reporting and presentation, standardization of thermodynamic data communications is very complex.

Major Accomplishment(s):
Since 2002, the NIST Thermodynamics Research Center of the Physical and Chemical Properties Division has played a key role in the development and establishment of the international standard, leading the efforts of the Task Group for IUPAC project 2002-055-3-024, "XML-based IUPAC Standard for Experimental and Critically Evaluated Thermodynamic Property Data Storage and Capture". This project has recently been successfully completed with the establishment of the new IUPAC standard for thermodynamic data communications, ThermoML. The full description of ThermoML was published in the March issue of Pure and Applied Chemistry (Vol. 78, No. 3, pp. 541–612, 2006), and the ThermoML namespace has been established on the IUPAC Web site.

The role of ThermoML in thermodynamic data communications is reflected in making a significant impact for development of efficient chemical engineering applications, including chemical process design, providing new capabilities for data delivery from 'data producers' to 'data users', establishing software 'tools' designed to improve the quality of published experimental data, as well as new mechanisms for development of a variety of data products serving the scientific and engineering communities. The development of ThermoML and the software infrastructure supporting it has lead to a unique cooperation between five major journals in the field (Journal of Chemical and Engineering Data, The Journal of Chemical Thermodynamics, Fluid Phase Equilibria, Thermochimica Acta, and International Journal of Thermophysics), three scientific publishers (ACS, Elsevier, and Springer), NIST, IUPAC, and leading chemical process design companies.

Recently, the symposium, "ThermoML: Purpose, Structure, and Applications" was organized by the American Chemical Society (ACS) at its Spring 2006 National Meeting in Atlanta, Georgia. Speakers from the United States, Canada, UK, Germany, Netherlands, and New Zealand represented all components of the global data delivery process based on ThermoML. The announcement of ThermoML as a new IUPAC standard was made on March 27, 2006 by IUPAC President Prof. Bryan R. Henry during a special "ThermoML" reception co-sponsored by IUPAC, NIST, FizChemie Berlin (Germany), and Elsevier (Netherlands).

Future plans include expansion of the ThermoML for thermodynamic and transport properties of biomaterials.

Project Team: M. Frenkel, R. D. Chirico, V. Diky, C. Muzny (838-B); S. E. Stein, Q. Dong (838-G); K.N. Marsh (University of Canterbury, Christchurch, New Zealand); J. H. Dymond (University of Glasgow, Glasgow, UK); W. A. Wakeham (University of Southampton, Highfield, Southampton, UK); E. Königsberger (Murdoch University, Murdoch, WA, Australia); A. R. H. Goodwin (Schlumberger Technology Corporation, Sugar Land, TX, USA)

M. Frenkel opens the ACS ThermoML Symposium. Honorable guests (right to left): Daniel Friend, Acting Chief of the Physical and Chemical Properties Division, NIST, Bryan R. Henry, IUPAC President (University of Guelph, Canada), and Catherine T. Hunt, President-Elect of ACS (Rohm and Haas Company).

Propane as a Thermodynamic Reference Fluid

New measurements of the heat capacity and pressure-density-temperature (p-ρ-T) properties of propane, combined with carefully selected literature data, have been used to develop a new equation of state (EOS) of unprecedented accuracy. This new EOS will form the basis for future equations of state and other thermodynamic models developed in the Division.

Contacts: M. McLinden, E. Lemmon (Physical and Chemical Properties Div., Boulder)

Propane C_3H_8 is an important industrial fluid—it is a significant component of many natural gases and a major fuel in its own right. But propane was selected for the present work because it is representative of the wide range of light hydrocarbons found in natural gas and liquid fuels. It is chemically stable and has an experimentally-accessible critical temperature of 369 K. It has the practical advantages of low toxicity, and commercial availability at very high purity The ratio of its critical temperature to triple-point (freezing) temperature is among the highest of any fluid. These attributes make it an excellent reference fluid for the development of thermodynamic models.

Computerized databases, such as NIST REFPROP, employ equations of state—a mathematical representation of all the thermodynamic properties of a fluid. An EOS allows calculations at any condition and also provides values for properties, such as entropy, which cannot be measured directly. While some EOS are based on theory, the most accurate ones covering wide ranges of temperature and pressure are empirical and must be fitted to extensive experimental data.

The pressure-density-temperature $(p\text{-}\rho\text{-}T)$ properties are the most essential data for fitting an EOS. In the current work, we have employed a new two-sinker densimeter to measure the $(p\text{-}\rho\text{-}T)$ properties of high-purity (99.999 %) propane at temperatures from 265 K to 500 K with pressures to 36 MPa. The measurements extend from low-density vapor to compressed liquid states, and the near-critical region was covered extensively. The uncertainty $(k = 2)$ in density is $\pm(0.0025\ \% + 0.0006\ \text{kg/m}^3)$ at 293 K, increasing to $\pm(0.02\ \% + 0.001\ \text{kg/m}^3)$ at 500 K. These represent some of the most accurate data available (see figure).

Data for the heat capacity are also vital. While it is possible, in principle, to compute the heat capacity from an equation of state fitted only to $(p\text{-}\rho\text{-}T)$ data,

this involves the evaluation of multiple first and second derivatives; any small errors in the $(p\text{-}\rho\text{-}T)$ data would be greatly magnified. We have measured the heat capacity at constant volume C_V, using the same propane sample, from the triple point temperature of 85.48 K to 345 K with pressures to 35 MPa. Measurements of the heat capacity at saturation C_{sat} were also used to determine the vapor pressure from 85 K to the boiling point at 231 K.

The equation of state is expressed in terms of the Helmholtz free energy as a function of the temperature and density. The Helmholtz energy A is equal to $U - TS$, and this form is useful because calculation of all the other thermodynamic properties (including enthalpy, heat capacity, speed of sound, vapor pressure, etc.) require only the Helmholtz energy or its temperature or density derivatives. The general form of the equation is well established, but new terms were used to represent the liquid-vapor critical region. The new EOS has 20 terms—about one-half the number compared to other equations of comparable accuracy. Non-linear fitting techniques were used, and this approach demands high quality data, otherwise it is possible to fit systematic errors in the data rather than the real fluid behavior. The new experimental data, combined with carefully selected literature data, were vital for obtaining a successful fit. The EOS not only fits the experimental data within its error, it also displays qualitatively correct extrapolation behavior down to 40 K in temperature and to extremely high temperatures and pressures. This was achieved by placing theoretically based constraints on the non-linear fit.

This new EOS will serve as a reference in several different ways. First, the functional form will serve as the starting point for the fitting of other high-accuracy EOS; it has already been applied in this way to the refrigerants R227ea and R365mfc. The new propane EOS will serve as the reference fluid in extended corresponding states (ECS) models. The corresponding states concept is based on the observation that broad classes of fluids have similar properties when scaled by the critical temperature and density; ECS models add an additional parameter, the shape factor, which is based on the molecular size and shape. The wide two-phase range of propane makes it an excellent choice for this application. ECS models are particularly useful for fluids with very limited data. Finally, the new EOS will be used to develop simplified equations of state for fluids with intermediate data—those with not enough

data to develop a full high-accuracy EOS, but with more than the very limited data suited to an ECS model.

This work was presented in three talks at the 16th Symposium on Thermophysical Properties held August, 2006 in Boulder. The results will be submitted for publication in the Journal of Chemical and Engineering Data

Project Team: M. McLinden, E. Lemmon, R. Perkins, and J. Magee (838); J. Ochoa-Sanchez (Instituto Politecnico Nacional, Mexico)
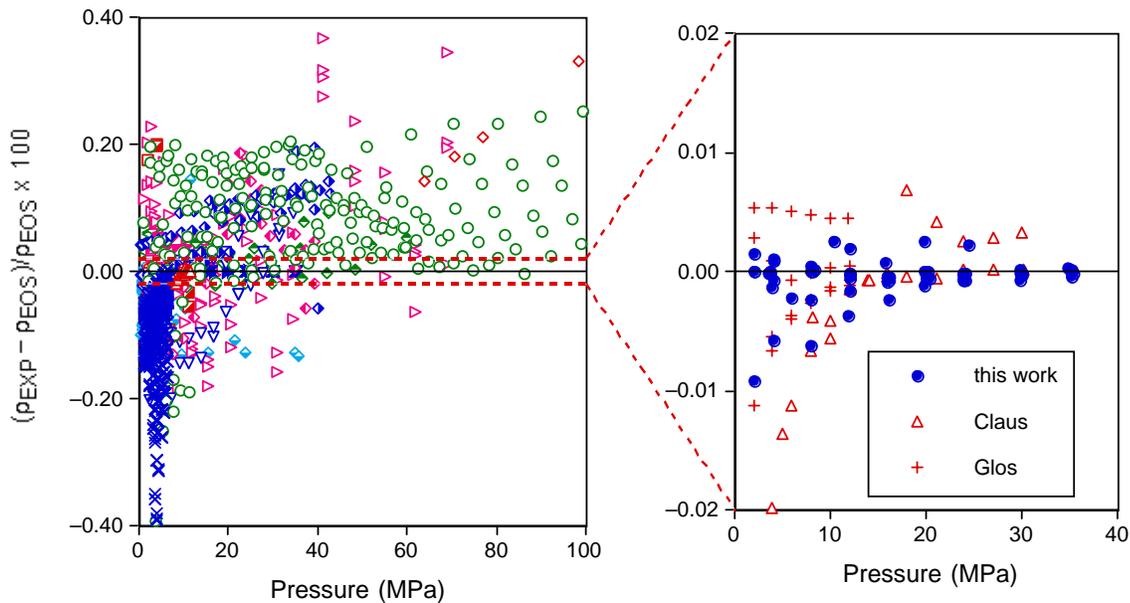
Measured densities for compressed liquid propane compared to the new equation of state. Although numerous literature sources exist (each symbol represents a different source), most exhibit very large scatter and uncertainties; the new data (blue circles) have uncertainties that are a factor of 20 lower than most previous data, and they are in excellent agreement with two sets of recent, high-accuracy measurements.

Development of Theoretical Models for the Study of Magnetic Properties in Metallic Nanoparticles

Contact: C.A. Gonzalez (Physical and Chemical Properties Div., Gaithersburg)

Purpose: To build a theoretical infrastructure for the modeling and understanding of experimentally observed ferromagnetic and other collective phenomena (such as plasmon excitations and binding energy shifts) that have a dependence on the number of atoms in metallic nanoparticles. These phenomena therefore depend on the particle size distribution and its deviation from bulk lattice behavior. This work will lead to the development, implementation, and validation of reliable theoretical tools that will enable researchers to understand the underlying physics governing the magnetic behavior exhibited by metallic nanoparticles.

Scientific Objectives. Metallic nanoparticles have been receiving considerable attention by both experimental and theoretical researchers. A driving motivation in this research is the interest in industrial applications of the emerging nanotechnologies. Magnetic properties of nanoparticles, such as ferromagnetism in transition metals like iron and cobalt and as recently reported in gold, have been of interest as materials for data storage applications. Other applications are in the health and medicine sectors where gold nanoparticles have seen usage as cancer cell killing agents, medical imaging contrast (MRI), hyperthermia treatment, drug delivery systems, and signal enhancement (Raman, Plasmon) factors.

Some of the interesting features of gold nanoparticles are shared by nanoparticles composed of different types of materials. As the nanoparticle size is made smaller, the accompanying large surface-to-volume ratio and lattice contraction induce interesting quantum effects such as electron localization, binding energy shifts, novel surface reactivity, surface ferromagnetic effects, and surface collective charge excitations as seen in plasmon absorption spectra. These effects have been experimentally reported as functions of the number of gold atoms in the nanoparticle. Despite significant progress in this area, the design and fabrication of these particles with desired physico-chemical properties remains more of an art than a science, involving tedious and expensive trial-and-error approaches. In addition, important metrology issues related to the characterization of size, shape and composition have not been resolved. A proper theoretical modeling of the fundamental physics and chemistry behind these effects will undoubtedly aid in the rational design and characterization of metallic nanoparticles with novel properties. Despite being diamagnetic in the bulk lattice (large scale), gold's emergent ferromagnetism observed at the nanoscale is interesting from a theoretical and technological point of view; a study of this type of emergent behavior will also shed light on the other and related novel quantum effects of lattice contraction and plasmon excitations as well as binding energy shifts at the intermediate nanoparticle sizes we are investigating. With reasonably accurate models, scientists will be able to predict beyond what has already been experimentally reported, testing the possibility of combined effects such magnetization and plasmon oscillations coexisting at some desired particle scale.

The CSTL team of researchers in the Computational Chemistry Group is working on an integrated approach that will lead to the development of a novel and reliable theoretical model based on *ab initio* molecular orbital calculations and a many-body quantum mechanical approach to explain all the quantum effects previously discussed.

Major Accomplishments:

Ab Initio Molecular Orbital Calculations of Magnetism in Bare Gold Nanoclusters. A particularly interesting question is how a ferromagnetic moment arises in the gold nanoparticle given its diamagnetic behavior in the bulk. Specifically, how does the symmetry of directed spins occur, at what size scales, and what are the accompanying electron charge effects. In order to answer these questions, we have applied spin-dependent density functional theory (DFT) in the scalar relativistic pseudo-potential formalism to study the energetic tendency for the gold clusters to spin polarize. The computed electronic structures of various gold nano-clusters reveal that permanent size-dependent spin-polarization appears without geometry relaxation for bare clusters even though bulk gold is diamagnetic. The polarized ground states for clusters are favorable due to the hybridization of the *s* and *d* orbitals, indicating that gold clusters are intrinsically magnetic due to the hybridization of the atomic orbitals. The magnetism is localized to a mono-layer on the outside of these clusters with the interior atoms remaining mostly diamagnetic. The results indicate that bare octahedral clusters are expected to be magnetic for cluster sizes of approximately 38 atoms and larger. As these clusters grow in size, the diamagnetic core dominates as soon as the surface-to-volume ratio becomes small and the core diamagnetism prevails. Thus, magnetism in gold clusters is a primarily a size-dependent effect. The results provided by the DFT calculations in this work are consistent with the experimental evidence that suggests that gold magnetism is strongest when weakly interacting capping agents are used. Strongly interacting capping agents would likely quench this magnetic behavior. That the magnetism of gold nano-particles can be tunable in certain size regimes strongly supports their use in many applications.

Mean Field Model for the Study of Size Dependence Of Ferromagnetism In Gold Nanoparticles. The DFT calculations discussed above provided the basis for a theoretical model to describe the size dependence of magnetism in gold nanoparticles. In this part of the project, a simple spin-spin Ising interaction model for the surface ferromagnetism was combined with the bulk gold diamagnetic response due to the effective surface interaction field to model the size dependence of the magnetization of the Au nanoparticle. This surface magnetic field is the Weiss field and acts as an effective applied field as seen by the core electrons. Using the Maximum Entropy formalism, we obtained the average temperature dependent magnetization within a mean-field model. It was found that the model reproduces the experimentally observed peak in magnetization reported in the literature for the case of gold nanoparticles. The results of this project indicate that the size dependence observed in the case of gold nanoparticles is a general effect and that the model should be applicable in all core-shell nanoparticles of diamagnetic metals.

Impact: The interest in nanotechnology and theoretical modeling and simulation tools has been increasing steadily. This interest is shared by governmental and academic

research institutions as well as industrial sectors that are interested in new technical innovations. Nanotechnology promises to revolutionize our infrastructure in the mid to longterm. A detailed examination of paradigmic nanoscale systems that exhibit novel quantum effects such as gold nanoparticles can produce theoretical modeling tools and insight that will be of value to any research effort into nanoscale materials characteristics. This project is creating the necessary infrastructure that could help researchers in industry develop more robust and cost-effective theoretical models for prediction and control of currently existing industrial nanoscale applications, as well as those proposed for the future as envisioned by scale plans in computing and electronics.

Future Plans: In the near future, this work will be extended to develop similar robust theoretical tools that will enable scientists obtain a fundamental understanding of the physics governing the size dependency of non-linear optics as well as electron transport properties in metallic nanoparticles. The development of these tools is critical for the rational design of novel devices (such as biomarkers, memory devices, sensors, etc.) at the nanoscale that can take advantage of these three physical properties simultaneously.

Project Team: C.A. Gonzalez, R. Magyar, F. Michael, V. Mujica (838-G).

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Fig.1 Total energy for the Au$_{38}$ nanocluster as a function of the spin imbalance described via different DFT methods.

QuickTime™ and a
TIFF (Uncompressed) decompressor
are needed to see this picture.

Fig2. Magnetic moment per atom as a function of nanoparticle diameter.

Evaluation and Improvement of Molecular Simulation Methods

Contact: R. Mountain and V.K. Shen (Physical and Chemical Properties Div., Gaithersburg)

Context and Approach:

Molecular simulation has been identified as a "breakthrough technology" for the estimation of fluid properties of industrial importance. It would improve the productivity of data generation and enhance the computational capabilities of industrial researchers. There are obstacles to realizing the breakthrough. The obstacles are the lack of molecular interaction models that contain the necessary physics for the molecules of interest and the availability of simulation codes that can adequately sample the fluid system in an acceptable time frame. The interest in molecular simulation is to determine fluid thermal properties for systems/conditions where data are lacking and direct measurement would be expensive and/or hazardous such as highly flammable materials or high temperature, high pressure states.

Major Accomplishment(s):

NIST is addressing the obstacles in two distinct ways. The development of improved interaction models for industrially interesting fluids is being encouraged through leadership in the activities of the Industrial Fluid Properties Simulation Collective (IFPSC), an association of industrial and government scientists interested developing and applying molecular simulation to industrially significant fluid property determination problems. The operation of the series of Industrial Fluid Property Simulation Challenges is the major activity of the IFPSC. The results of the first two challenges have identified situations where simulation methods can be improved to make them more robust and computationally efficient. These are the determination of fluid-fluid phase boundaries and the determination of the shear viscosity and thermal conductivity transport coefficients. NIST is actively involved in both problems and have found and evaluated simulation methods that satisfy the robust and efficient conditions.

The Grand Canonical Ensemble Transfer Matrix Monte Carlo method has been extended to enable a simulation to determine liquid-vapor and liquid-liquid phase boundaries for mixtures with greater computational efficiency than existing methods. The relevant codes are being incorporated into a public access Monte Carlo package, "Towhee", that is currently maintained by Sandia Labs. This method has also been applied to simple models of protein denaturation.

The estimation of transport coefficients through molecular simulation tends to be a computationally intensive business because of the large amount of sampling required to obtain converged results. The reversed-perturbation molecular dynamics method has been evaluated and shown to be computationally more efficient than the other available methods for estimating the shear viscosity and thermal conductivity of liquids as shorter computation times are needed for convergence. It is not useful for gases as very large systems are necessary to have it work for gases.

Publications, Patents:

1. J. R. Errington and V. K. Shen, J. Chem. Phys. **123**, 164103 (2005)
2. V. K. Shen, *et al*, J. Biophysical Journal **90**, 1949 (2006)
3. R. D. Mountain, J. Chem. Phys. **124**, 104109 (2006)

Peptide Ion Fragmentation for Proteomics Applications

Contact: K.K. Irikura (Physical and Chemical Properties Div., Gaithersburg)

Purpose: Tandem mass spectrometry is a central technology in the interdisciplinary, health-related field of proteomics. Great initial progress has been made by using data mining techniques to discern the most common patterns of peptide ion fragmentation. This knowledge has made it possible to infer sequence from the observed spectra. However, about half the information in the mass spectra is currently discarded as indecipherable. The goal of this project is to understand, at a predictive level, the chemical reactions responsible for peptide ion cleavage. This will improve the reliability and speed of proteomics measurements.

Scientific Objectives: Current "informatics" approaches (statistical pattern recognition) have been effective for developing rules for interpreting the tandem mass spectra of peptides. The most important rules are pervasive and were discovered by using modest databases. Finding additional rules will make it possible to extract more sequence information from the same mass spectra. Unfortunately, these lower-frequency rules cannot be discovered empirically without enormous databases, which are not yet available. Thus, we have adopted a different strategy in this project.

Tandem mass spectrometry is essentially a pyrolytic process. A mass-selected peptide ion is subjected to multiple collisions with an unreactive gas. The collisions increase the ion's internal energy until it finally dissociates. The observed mass spectrum is determined by the branching among the various unimolecular dissociation reactions. In this project, we study individual peptide ions to understand their fragmentation patterns in quantitative, chemical detail. By understanding the underlying organic chemistry, we will be able to extract fragmentation rules directly, without massive databases.

Our technical approach involves integrated theory and experiment. The initial energy deposition process is modeled theoretically as a collisional energy transfer process. Reaction mechanisms, thermochemistry, and rate constants are obtained from *ab initio* electronic structure theory. Rate constants will also be measured experimentally using a novel instrument (still under construction) designed to attain complete ion thermalization, as required for comparison with statistical theories. Peptides designed specifically to reveal the underlying chemistry are synthesized and their tandem mass spectra are measured in the laboratory using standard instrumentation. Peptides are selected to answer questions raised by our theoretical calculations, and calculations are designed to answer questions raised by our experiments. After the theoretical thermal rate constants are validated experimentally, non-thermal rate constants will be computed by using the theoretical energy deposition functions. The resulting branching fractions will be compared directly with spectra collected under a variety of experimental conditions.

Major Accomplishments:

Synthesis and Measurements of Homologous Series of Peptides. A database of about 40,000 tandem mass spectra of yeast-derived peptides has been assembled so far. We search it continually for patterns of reactivity that have not been previously documented or explained. Suspected patterns are then investigated by designing and synthesizing a series of peptides that will best reveal the underlying chemistry. This is important because the peptides selected by *S. cerevisiae* are not expected to be particularly well suited for revealing reactivity patterns.

One such pattern is related to loss of neutral H_2O or NH_3 from peptides with an N-terminal glutamine (Gln) residue. This was initially recognized among the statistics for the yeast peptides. A series of Gln-initiated peptides was prepared with systematic variations of basicity and chain length. Tandem mass spectra of the synthetic peptides confirmed the suspected pattern and placed it on a quantitative basis. Another neutral loss detected in library search is the loss of 62 Da from peptide ions containing two threonines at the amino end. By examining the tandem mass spectra of a series of synthetic peptides as a function of collision energy it became clear that the loss takes place in two steps and involves an initial loss of water from the penultimate threonine followed by loss of acetaldehyde from the terminal one. The effect of aspartic acid on peptide fragmentation is being studied.

Reaction Mechanisms, Thermochemistry, and Rate Constants. Quantum chemistry calculations are traditionally applied to much smaller molecules. Our initial work has been to develop software tools to address the particular problems that arise for polypeptides. A procedure has been developed that combines both force-field and *ab initio* methods to find the most important conformational isomers for arbitrary gas-phase peptide ions. Applications to small models for the Gln chemistry described above gave only partial agreement with the experimental results. The next step is to run additional experiments to determine the most efficient, chemically appropriate molecules for theoretical modeling.

The most popular chemical model in this field involves the "mobile proton hypothesis." Surprisingly little is known about the proton mobilization process. We are using both 1D lattice Coulomb models and *ab initio* polyglycine models to investigate the peculiar fragmentation of triply-charged peptide ions. Electro-static confinement of mobile protons is predicted to be an important effect; this will be tested experimentally. An unexpected trend in backbone proton affinities (see figure) was discovered for linear polyglycine. This biases fragmentation toward the N-terminus and can mask Coulombic effects beyond about 15 residues. Similar results have been obtained recently in aqueous solution; intramolecular electrostatic effects on half-wave potentials are damped beyond distances of about 1 nm.
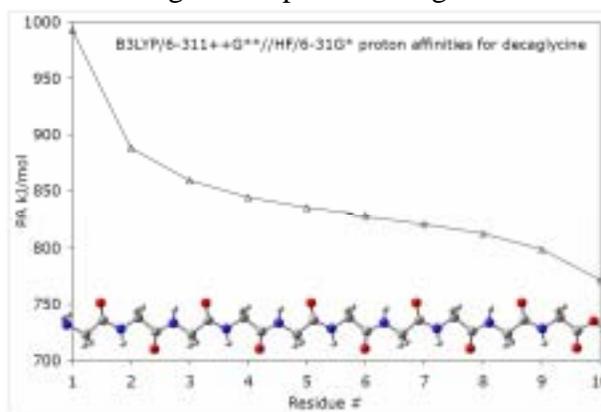
There is little consensus about the chemical reaction mechanisms involved in peptide ion fragmentation. In addition to investigating mechanisms that have already been suggested (such as the mobile proton mechanism), we are using the NIST technique of isopotential searching (IPS) to discover novel mechanisms. Since peptide ions are large molecules, the existing IPS software requires weeks of computer time to yield results. To accelerate these calculations, the software is now being completely rewritten to take advantage of the large NIH cluster computer. We expect near-linear scaling, which should permit us to obtain results 10 to 50 times faster than now possible.

Experimentally, the construction and design of the thermalized peptide machine continue to proceed as planned. The rate constants for thermal (not collisional) peptide fragmentation will be measured. Their internal energy distribution will be known (Maxwell-Boltzmann), which is necessary to make corresponding theoretical predictions of unimolecular rate constants. Thus, this experiment makes it possible to test the quality of the *ab initio* derived rate constants.

Collisional Energy Transfer. Quasiclassical trajectory simulations are being performed to understand how translational energy is converted to vibrational and rotational energies. The results of trajectory calculations are used in Monte Carlo modeling of the resultant vibrational and rotational energy distributions in ions. Statistical rate theories are then used to evaluate the rates of ion fragmentation. Early results indicate that the modeling results are in semi-quantitative agreement with experiment. Energy distributions resulting from trajectory calculations are very sensitive to the details of the short-range repulsive part of the ion-collider potential energy surface. Surprisingly, the best agreement with experiment is obtained when hard-spheres collision model is used, with more realistic "softer" potentials resulting in lower ion dissociation rates.

Impact: The chemical reactions studied so far are fairly specific, so impact on proteomics is correspondingly minor. More general conclusions are anticipated as additional reactions are discovered. Insights about the fragmentation of triply-charged ions will be analytically useful.

Future Plans: All aspects of this project appear to be progressing well and will be continued: the size effect on activation energies will be investigated; atomistic molecular dynamics trajectories will be used to model collisional energy deposition; *ab initio* methods will be used to further characterize proton shuttling, backbone reactivity, and neutral losses; the isopotential-searching software will be parallelized; development will continue on the unique instrument for measuring thermal peptide fragmentation kinetics; more spectra will be added to the peptide mass spectral database.

Project Team: K.K. Irikura, J. K. Merle, S. E. Stein, P. Neta, V. D. Knyazev,
J. W. Hudgens (838-G)

References
1. P. Neta, Q.-L. Pu, L. Kilpatrick, X. Yang, S. E. Stein *Dehydration vs. Deamination of N-Terminal Glutamine in Collision-Induced Dissociation of Protonated Peptides*; *J. Am. Soc. Mass Spectrom.*, in press.
2. P. Neta, Q.-L. Pu, X. Yang, S. E. Stein *Consecutive Neutral Losses of H2O and C2H4O from N-Terminal Thr-Thr and Thr-Ser in Collision-Induced Dissociation of Protonated Peptides. Position Dependent Water Loss from Single Thr or Ser*; *Int. J. Mass Specrom.*, submitted.
3. C. R. Kinsinger, K. K. Irikura *Efficient Conformational Searching of Protonated Peptides. How Good Are Force Fields for Gas-Phase Protonated Peptides?*; *J. Comput. Chem.*, submitted.
4. J. K. Merle, K. K. Irikura *Coulombic Confinement in the Mobile Proton Model*, Manuscript in preparation.

PAH/Soot Formation Chemistry in Combustion Systems

Contact: W. Tsang (Physical and Chemical Properties Div., Gaithersburg)

Autom. & Aerospace, Energy and Environ. Tech., Data & Informatics

Context and Approach:
The incomplete combustion of organic fuels can lead to the formation of polyaromatic hydrocarbons (PAH) and particulate matter. These materials are known to have deleterious health effects. A longstanding aim in combustion research is to develop methods for their amelioration through combustion controls. This has become particularly important in recent years due to the stringent regulations that have or are being imposed. Thus the PM 2.5 standard can have severe effects on the type of combustors that can be used or the type of filters and catalytic converters that must be fitted into the exhaust of the combustors. More recent interest in biofuels has led to similar concerns regarding their effect on particulate matter emissions. Equally important is the need to understand the health effects of specific PAHs.

For a number of years, CSTL and BFRL with the support of SERDP/DOD have carried out an experimental and analytical program designed to develop sufficient understanding of the combustion process so that formation of PAHs and soot can be predicted for a fuel such as heptane on the basis of simulations. The general process for formation of PAH/SOOT involves the breakdown of the fuel to small unsaturated fragments, followed by a subsequent build-up to form larger multi-ring aromatics and ultimately particulate matter. The nature of the small unsaturated compounds determines the quantity and presumably the nature of the soot that is formed. Oxidation reactions will interrupt the progression leading to the formation of water and carbon dioxide.

Accomplishments:
The two elements in the NIST program are: 1) development of a database to allow modeling of the combustion of heptane (a common reference fuel used in simulations) and, 2) to obtain experimental reference data to characterize the formation of PAH/SOOT. The first aspect of this program is near completion. The nature of the small unsaturated compounds has now been deduced on the basis of experiments with our heated single-pulse shock tube, evaluation of literature data, and application of theory with a NIST developed program, so that all combustion conditions can be covered.

Work on PAH/SOOT formation mechanisms is being carried out. This is an intrinsically harder problem, since different compounds are being created as the reaction proceeds. In contrast, the breakdown of a fuel during combustion leads to fewer compounds and hence removes mechanistic difficulties. To characterize PAH/SOOT formation, a well stirred/plug flow reactor has been constructed. A sketch of the apparatus can be found in Figure 1. The well stirred aspect of the reactor simply stretches out the flame front and leads directly to generation of reactive species and ultimately soot. The plug flow reactor gives time for the observation of the interesting chemistry preceding and during PAH and SOOT formation. Particularly interesting is the capability of introducing particular reaction intermediates into a stream of gas that has not yet began to soot, thereby inducing sooting behavior. We are now beginning to obtain information on the temporal behavior of the chemical constituents of the precursors of larger PAH and ultimately soot itself. Figure 2 summarizes experimental results on the light and smaller ring PAHs that have now been detected upon the addition of small quantities of benzene into the reaction stream. This represents a comparison of yields for the case where no benzene is added. It can be seen that

for the smaller unsaturates, changes are small. However the changes increase in magnitude with ring size. Clearly we are seeing an accelerating rate of ring growth. We have also made a determination through DMA size distribution. It is clear that there are important differences upon the addition of benzene. We also collect on a filter the particulate matter that is formed. These samples are being analyzed by the Analytical Chemistry Division. Larger multi-ring aromatics have now been detected. We will also be subjecting the solid soot to TEM analysis.

Impact: The most important short term impact is from our heptane breakdown work. Practically all existing combustion models focus on stoichiometric or near stoichiometric mixtures of fuel and oxidant. The present information should permit simulations of rich mixtures. This may be particularly important for various auto-thermal processes where the heat of reaction in a rich mixture is used to drive parallel processes during chemical conversion. It is of course also of great importance for soot formation modeling beginning with realistic fuels.

A particularly important and unexpected aspect of the work is the capability of generating well characterized soot samples with known amounts of PAHs. At the present time soot is treated as a generic material. This is probably not the case. There are significant uncertainties regarding the effects of "early" and "late" soot and controversy about the effect of various PAHs on health. The facility that we have developed has the capability of generating multi-gram quantities of such well defined material.

Future Plans: We are currently modeling the experimental results and testing the accuracy of the models that are used for making predictions regarding soot yields. Another direction of interest is to understand the effects of compounds distinct to biofuels on PAH/SOOT formation. Thus probing the behavior of phenols and methoxy ethers in a manner similar to that described for benzene will give direct information on the sooting propensity of biofuels.

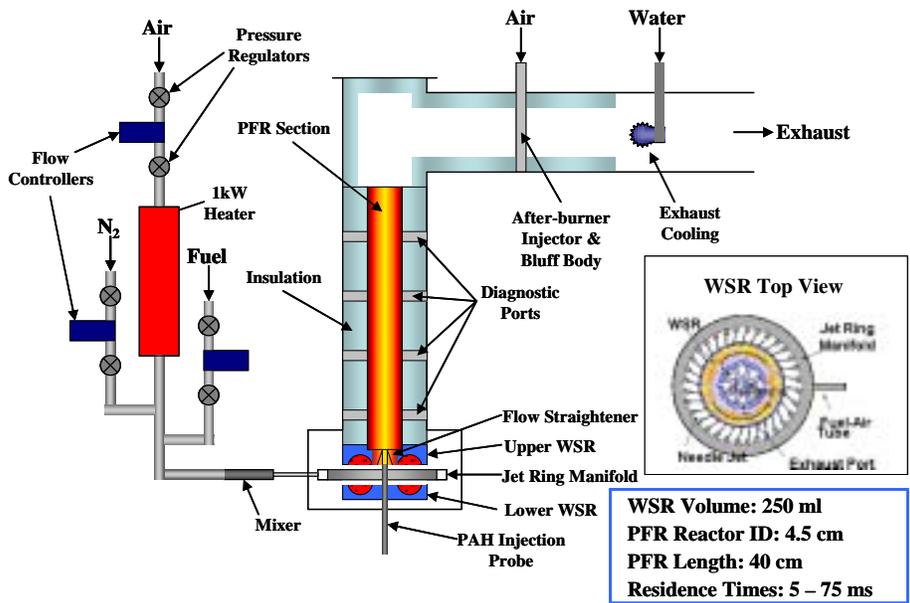**Project Team:** W. Tsang, C. Stroud, and J. A. Manion (838-G); S. Manzello and D. Lenhert (BFRL)

Figure 1:Schematic of the NIST WSR/PFR experimental apparatus displaying the WSR/PFR assembly and detail of WSR jet ring.
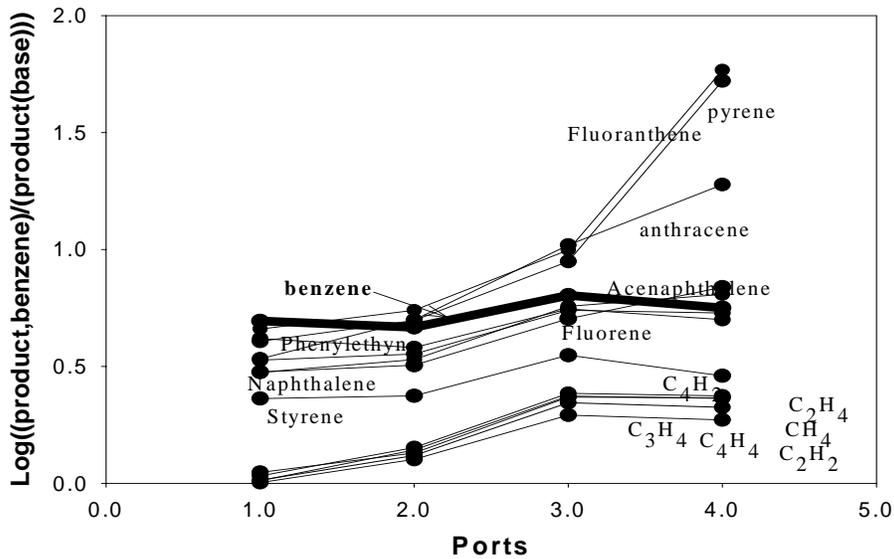

Figure 2: Comparison of comparisons at four ports in plug flow reactor of some compounds found when 725 ppm benzene is added to system with base ethylene-air combustion. Total time between ports 1 to 4 is 45 microseconds. Temperature is 1300 K. Pressure is 1 bar. Dark line is for benzene additive.

Properties and Processes for Cryogenic Refrigeration

Contact: R. Radebaugh (Physical and Chemical Properties Div., Boulder)

Context and Approach:

Cryogenic temperatures are required for many technology areas, including infrared sensors for surveillance and atmospheric studies; superconducting electronics, magnets, and power systems; to create clean vacuums in semiconductor fabrication processes; for liquefaction of industrial gas, and many other existing and potential applications. This project addresses issues associated with cryogenic refrigerators, i.e. cryocoolers, and materials performance that are limiting growth in all these technology areas. Proper measurements need to be established to characterize losses within cryocoolers and models need to be developed to optimize the design of such systems. Material properties at cryogenic temperatures also are needed by industry for the design of cryogenic equipment, but the data are difficult to find and interpret.

Major Accomplishment(s):

In the data and methodology area, the layout of our web site has been redesigned to provide clear access to references for the database on cryogenic material properties and to add a graph of each property as a function of temperature between about 4 K and 300 K. The web site is in the process of being updated with many of the graphs being added to materials already on the web site. We will continue adding new information. An invited paper on our Cryogenic Materials Property Database was presented at the International Cryogenic Materials Conference in Prague in June by Peter Bradley. Two invited short courses on cryocoolers were presented by Ray Radebaugh this year. The first was a one-day course given on June 13, the day before the International Cryocooler Conference, in Annapolis, MD. The second was a half-day short course emphasizing superconducting applications and was given the day before the Applied Superconductivity Conference in August.

Measurements of the resonance effect in inertance tubes used for phase shifting between flow and pressure in pulse tube refrigerators were carried out this year and reported on at the International Cryocooler Conference in June by Mike Lewis. The real and imaginary parts of the impedance were measured for frequencies from 40 to 70 Hz. The results agreed fairly well with our model predictions, although they begin to diverge at the highest frequency of 70 Hz. Further measurements at high frequencies are therefore warranted at this time. This resonance effect can be utilized to provide for a faster cooldown time of pulse tube refrigerators.

In January we were notified that we received the first DARPA contract as part of their new MEMS program on Micro Cryogenic Coolers (MCC). The first phase is 18 months and the second phase is 18 months. The University of Colorado is the prime contractor in this program and the NIST Electromagnetic Division is also collaborating on this program. The goal of the program is to cool a high-T_c superconducting terahertz detector to 77 K with 3 mW of cooling and have the entire system occupy less than 4 cm^3. This year we have identified a few optimized refrigerant mixtures in collaboration with Marcia Huber of the Theory and Modeling of Fluids Group to use in the Joule-Thomson cryocooler. The use of glass microchannel plates with square pores of 20 μm on a side is planned for the heat exchanger. The University of Colorado is currently testing various manifolding techniques for the heat exchanger.

Project Team: R. Radebaugh, P. Bradley, M. Lewis (838-B); J. Gary and A. O'Gallagher (ITL)

The NIST Chemistry WebBook

Contacts: P.J. Linstrom and W.G. Mallard (Physical and Chemical Properties Div., Gaithersburg)

Data and Informatics

Context and Approach:
The NIST Chemistry WebBook is a site on the Internet (http://webbook.nist.gov/) that distributes physical and chemical property data on well defined chemical species and reactions. The purpose of this activity is to improve the nation's science and technology efforts by providing convenient access to NIST Standard Reference Data[®‡] to scientists, engineers, and the general public.

The data in the WebBook is from databases developed by NIST and outside contributors. Such data were traditionally published only in hardcopy format. With the widespread introduction of personal computers SRD products were distributed in the form of databases on floppy disks and CD-ROMs. The development of the Internet provided a third avenue for distribution of NIST data. The NIST Chemistry WebBook was one of the first efforts to distribute this data over the Internet. This has several advantages for NIST's customers: rapid access to data from any location with Internet access, updates and corrections made without any action required by the user, and better integration of data and relevant documentation. The NIST Chemistry WebBook is a service, so it is available on a continuing basis. Users worldwide can (and do) access the site 24 hours a day.

The greatest technical challenge in this effort is the integration of databases from multiple sources into a single collection of data. Because of the lack of standards for physical and chemical properties, the data collections included in the site use diverse formats and conventions for reporting data. Problems with integrating these data sets are exacerbated by traditional chemical informatics problems associated with the identification of chemical species: inconsistent nomenclature, ambiguous or erroneous references to 3rd party registries, and inconsistent conventions for chemical formulas. Systematically resolving these issues provides a common infrastructure for searching, identifying, and providing reliable information. Additionally, the WebBook employs internationally recognized developments in nomenclature, data exchange standards, and chemical identifiers, such as InChI - see next section.

Major Accomplishments:
The site was established in 1996 and has grown to encompass a wide variety of thermochemical, ion energetics, physical, solubility, spectroscopic, and chromatographic data. The site also includes two special features: a set of interactive physical property models developed at the NIST Boulder labs and a group additivity based estimator for gas phase thermodynamic properties. The physical property models provide thermodynamic and transport property data at user supplied conditions for a number of industrially important fluids. The group-additivity based property starts with user supplied structures using Benson increments. The NIST Chemistry WebBook is part of the NIST Standard Reference Data program and is the result of the work of many contributors. These are identified in the "credits" section of the web page at http://webbook.nist.gov/chemistry/. The site is regularly updated with new functionality and new and expanded data set. The history of major changes in the site can be found at http://webbook.nist.gov/chemistry/history.html.

As a result of work sponsored by the European Union and efforts at NIST, the site is now available in Spanish, Portuguese, French, and Czech. These translations are invisible to users that primarily speak

---

[‡] cf. United States Congress in the Standard Reference Data Act (United States Code, title 15, chapter 7A, section 290).

English; they only show up for users with web browsers configured to prefer a language other than English.

In addition to displaying data, the NIST Chemistry WebBook also provides links other, relevant, on-line databases at NIST. This is done at the chemical species level, so the WebBook is effectively an index of the other sites. Linked databases include the gas phase kinetics database, the computational chemistry comparison and benchmark database, the NIST atomic spectra database, the NIST microwave spectra database, the electron-impact ionization cross section database, and the thermophysical properties of hydrogen web site.

Many of the recent improvements in the site have focused on the use of the IUPAC International Identifier (InChI) to improve interoperability with other data resources. The site has supported display and searching based on the identifier since shortly after its release. As illustrated in figure 1, InChI is used by the site to identify related chemical species such as optical and geometric isomers. In the past year the search algorithm used by the site was enhanced to better handle identifiers which are generated with a fixed hydrogen layer. This was done to improve interoperability with PubChem web site operated by National Center for Biotechnology Information (NCBI) at the National Institutes of Health (NIH). The IUPAC identifier is used to link species in the NCBI site to the NIST Chemistry WebBook.

Other significant work carried out in the past year included updates to the electronic spectra, chemical structure, and retention index databases.

Impact:

The NIST Chemistry WebBook is the most widely used SRD product developed by NIST. It is used by scientists and engineers in industry and academia and by students at the high school, college and graduate school levels. As noted in figure 2, usage of the site has grown over time. The site now regularly receives more than two million page views per month.

Future Plans:

This is an ongoing activity. Enhancements to the sites data collections and functionality occur continually over time. The following items are expected for FY 2007:

1. Updates to the ion and ion-cluster thermochemical databases.
2. Improved support cases where a search for a species based on its InChI fails. This change should also make it easier for others to link to specific species in the site.
3. Infrastructure enhancements to better support the storage and display of data for chemical reactions.

Publications:

The site itself is a NIST publication:

> P.J. Linstrom and W.G. Mallard, Eds., NIST Chemistry WebBook, NIST Standard Reference Database Number 69, June 2005, National Institute of Standards and Technology, Gaithersburg MD, 20899 (http://webbook.nist.gov).

Figure 1: Portion of web page showing display and use of the IUPAC International Chemical Identifier
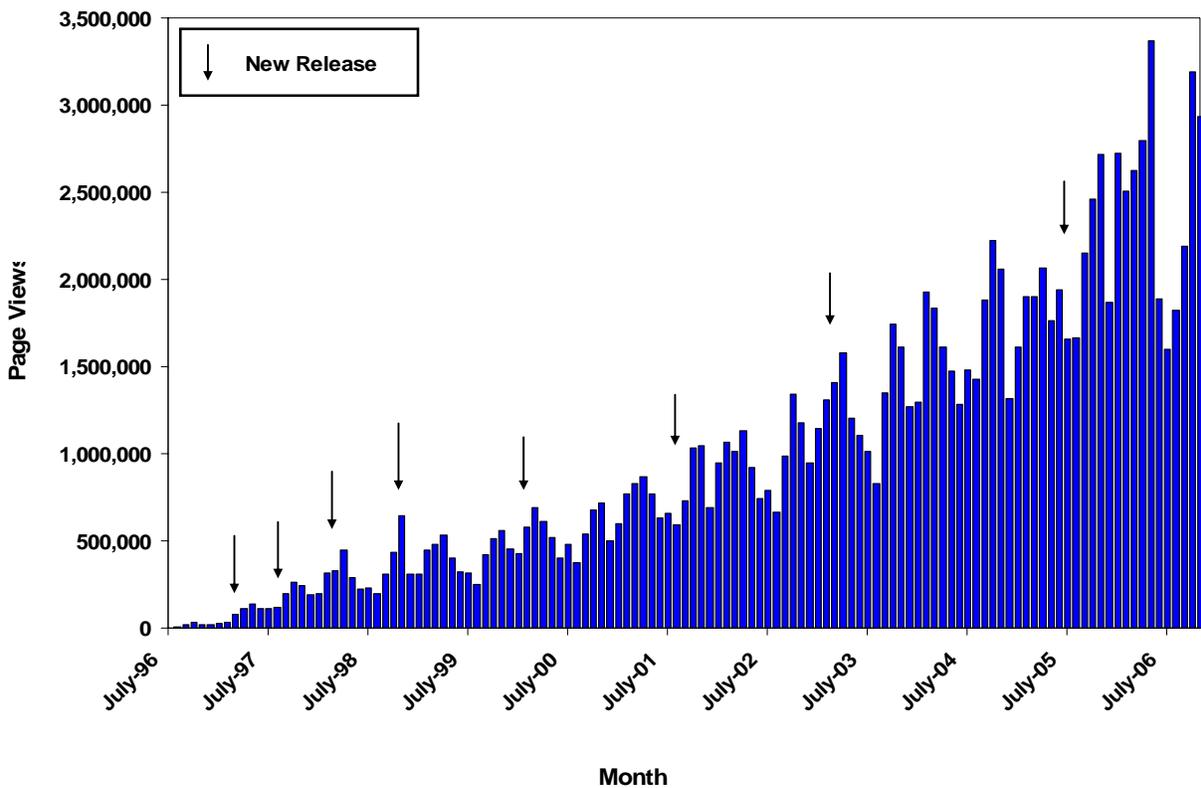
Figure 2: Usage history of the NIST Chemistry WebBook in page views per month.

Automatic Mass Spectral Deconvolution and Identification Software (AMDIS)

Contacts: G. Mallard and O. Toropov (Physical and Chemical Properties Div., Gaithersburg)

Health and Analytical Instruments

Context and Approach:
The problems of establishing chemical identity for trace species in complex matrices breaks down into two distinct pieces – isolating a single species and then analyzing the species so isolated. The tool of choice for many such complex matrices is the Gas Chromatograph coupled to the Mass Spectrometer. The data needed for identification of the species by mass spectrometry and NIST's role in providing such data is discussed elsewhere. The problem addressed here is extraction of the data for a single species. As mixtures become more complex or the concentration of a component decreases, the chromatographic process can not fully separate the component, and thus there are often multiple chemicals coming out "together" so that the mass spectral signatures are not clean enough to allow identification. The disentanglement of the multiple signatures is called deconvolution and is accomplished by AMDIS using a series of algorithms developed at NIST. The initial work was funded by the Defense Department in order to aid in the compliance with the Chemical Weapons Convention. The broader utility of the tool for analysis of complex mixtures (such as surface waters for pollutants, food for contamination, and complex fluids for medical applications) is only now being realized.

The fundamental algorithms that make up AMDIS are very stable and only minor changes have been needed. However, AMDIS has evolved to handle new instruments, to better couple to instrumentation software and to deal with types of analyses for which it was not originally designed. These efforts have been coordinated with a number of instrumentation companies to make the use of AMDIS easier for customers. In addition, we have been working directly with laboratories that are developing tools to respond to possible terrorism events where very large numbers of samples will need to be analyzed at possibly very sensitive levels. This latter effort has been the focus of the last year. AMDIS has been used to reduce the noise in reference spectra taken for the analysis of pesticides, toxins, poisons, and industrial chemicals. The effect of acquiring improved data of this type is to provide more sensitivity for the analysis of these chemicals without a corresponding increase in the risk of false positive identification.

Major Accomplishments:
The release of AMDIS 2.64 occurred in December 2005. During the year a major effort was devoted to developing libraries that could be of use to the laboratory emergency response networks. We directly worked with the FDA's Forensic Chemistry Center in developing these libraries both from their data and the data from the FDA Pesticide Residue Laboratory in Seattle, WA. Experimental data from both of these libraries, which had been analyzed using conventional analysis techniques, was reanalyzed using AMDIS. This reanalysis automatically provided the retention index data as well as substantially better spectra. The data from the two laboratories as well as data on chemical weapons from the Organization for the Prohibition of Chemical Weapons database was combined with additional data from the NIST mass spectral library to create a library of toxic industrial chemicals, pesticides and chemical weapon related compounds that was provided to a number of state laboratories as part of a training conducted jointly by personnel from NIST and the FDA Forensic Chemistry Center in August 2006. Work is ongoing on improvements in the database.

Impact:
AMDIS provides the first broadly available method for automating GC/MS analysis. With AMDIS an analyst can find, with high confidence, components in complex mixtures that would otherwise be missed as well as having greater confidence in all analytical results. The productivity of analysts in the laboratory is significantly increased using AMDIS since the first pass analysis of the data is in minutes - without direct intervention, rather than hours - with full involvement of the analyst. AMDIS can also aid in providing statistically valid confidence measures for analysis.

AMDIS usage is steadily increasing. It is used as part of the proficiency exams for laboratories world wide which participate in the analysis for compliance with the Chemical Weapons Convention. Increasingly it is being used for metabolite studies in complex matrices such a whole plant and whole cell analysis. Finally, the adoption of AMDIS as a part of the new DRS system of Agilent has meant that a far larger number of users will use it.

Future Plans:
Specific changes in AMDIS are driven by the user community. Minor modifications in the algorithm to take into account varying scan ranges as are often used in the instrumentation, improved access to the algorithm using the DLL (dynamic link library) in order to facilitate the use by instrument companies within their own data systems have been made in the past year.

The use of AMDIS to aid in the development of specific libraries for problems of national interest has become an increasing part of the effort associated with AMDIS. These library developments are coupled to specific protocols for analysis. In addition, training courses in the use of AMIDS, again tied to specific protocols for analysis will be important parts of the future efforts to make AMDIS more widely used.

Publications and Presentations:
Mallard, W.G., "Extending AMDIS – Interactions with Instrument Software, Concentration Calculations and Increased Sensitivity", Agilent Technologies, Wilmington, DL, November 17, 2005.
Mallard, W.G., "FERN GC-MS Chemistry Training Course LB508", FDA Forensic Chemistry Center, Cincinnati, OH, August 15-17, 2006.

The IUPAC International Chemical Identifier (InChI)

**Contact: D. Tchekhovskoi (Physical and Chemical Properties Div., Gaithersburg**

Chemical and Allied Products; Data and Informatics

Context and Approach:

The question of clearly identifying a chemical has been present almost since the beginnings of modern chemistry. As the number of chemicals grew, the need for systematic naming produced a number of results. The most widely adopted of these is that of the International Union of Pure and Applied Chemistry (IUPAC). But for many chemicals, the resulting names are complex and so common names are still widely used. For most chemists, the graphical structure is the best method for identifying a chemical. The structure provides graphic information that can allow a rapid understanding of the properties of the chemical that a long text name can never provide. However, the phenomena of tautomerism and (de)protonation change chemical structure, thus hiding the chemical identity of compounds. In addition, the same chemical structure may be drawn in such different ways that it is hard to establish equivalence of the drawings.

The need for a uniform and open standard that could be adopted by the entire chemical community generated the NIST/IUPAC project to develop a chemical identifier – The IUPAC-International Chemical Identifier (InChI). The aim of the project was not to create another naming system, or at least not a naming system that would be usable in normal communication. The goal was to create a naming system that would allow computers to uniquely identify a chemical, regardless of how it is drawn based entirely on the connectivity of the molecule – that is what atoms are connected to what other atoms.

To accomplish this, a lot of what is normally viewed as "chemical information" was discarded, and the molecules were trimmed to the minimum information needed to differentiate one from the other. In addition, a layered approach was developed to deal with some of the more complex issues of chemical structure. For example, the two molecules on the right below differ only in

unspecified   *trans*   *cis*
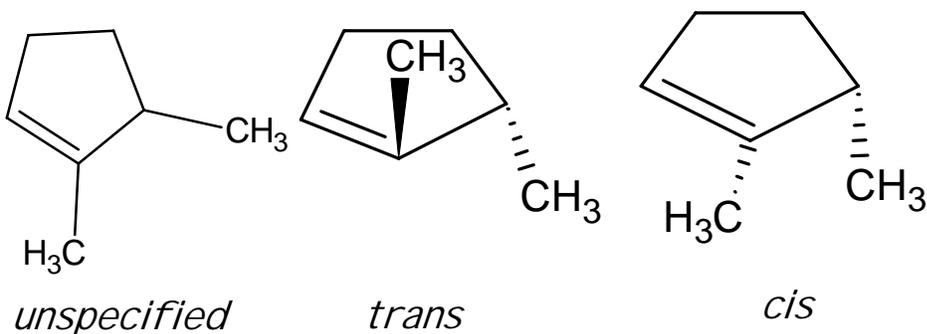
that one has the two methyl groups on the same side relative to the plane of the ring (on the right – called *cis*) and the other has the two methyl groups on opposite sides of the ring (in the center – called *trans*). On the left, is a diagram that can be used to represent either of the molecules. The left diagram shows only the connectivity and does not specify if the molecule is *cis* or *trans*. The

problem encountered prior to InChI is that the data retrieval was often dependent upon the way the molecule had been drawn. There is often a need to distinguish between the *cis* and *trans* form, and often a need to search for all possible forms, including cases where the configuration of the molecule was not known or it was known that a mixture was present.

The approach taken in developing InChI is a layered approach. This allowed as much information as was known to be specified, the search could be performed only on the information known and the search could be stopped with less than full information. Thus in the case above, a search for the *cis* isomer could be allowed to stop when it matched the connectivity or continued to find only the molecules that matched the geometric isomer.

Major Accomplishments:

The IUPAC International Chemical Identifier has been released by IUPAC in April 2005.

An InChI validation protocol was released in August 2006. It was designed to verify that a 3rd party InChI software or a ported to other platforms InChI software would still generate valid InChI strings. Simultaneously was released a minor software update. It includes a facility that allows reconstructing of a chemical structure out of InChI.

InChI has been adopted/used by a variety of entities:

PubChem, a major resource for medical research at the National Institues of Health (NIH) has adopted InChI as a standard for identifying and searching for compounds: For example, see: http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=6986.

It has been integrated by ACD Labs in their widely used commercial drawing program, ChemSketch as well as the freeware version that is distribute for home and student use, see: http://www.acdlabs.com/download/chemsk.html.

The Chemistry WebBook had adopted InChI and displays the identifier for all data in the WebBook thus making it possible for the very large audience that uses the WebBook to gain access to the identifier, see: http://webbook.nist.gov/cgi/cbook.cgi?Name=Horse&Units=SI,

The Environmental Protection Agency has adopted InChI for use in their Distributed Structure-Searchable Toxicity (DSSTox) Public Database Network, see: http://www.epa.gov/nheerl/dsstox/MoreonINChI.html

The Web of Science, (Thomson Scientific) one of the most widely used information sources for searching the scientific literature has adopted the identifier.

There are many other uses of the identifier – and since the software is distributed free of charge we do not even know if we are aware of all of them. One example is the Compendium of Pesticide Common Names (Alan Wood, UK), a site that allows users to find the common name as well as the structure for a pesticide, see: http://www.hclrss.demon.co.uk/. The wide number of adoptions has been made easy by a range of software solutions developed as a part of the project. Many of these are available from IUPAC.

Impact:

The identifier project is an example of open source software. Since it is a standard, it is not wise or reasonable to modify the code, but the code is available for any who want to use it. In addition, a number of tools to allow users to easily analyze their own structures using the freely available compiled code. As InChI becomes more widely adopted, it is expected that it will

enable a standardized referencing and search for chemical structures both over the Internet and in proprietary databases.

Future Plans:

We are examining the best way to extend InChI to include polymers, phase, and excited states. In addition, we are examining ways to make separate parts of InChI algorithm – chemical structure normalization, canonicalization, and serialization – available to 3rd party software through InChI API (Application Programming Interface.)

Project Team: D. Tchekhovskoi, S. Stein, S. Heller (838-G)

Publications and Presentations:

- The IUPAC International Chemical Identifier: InChl — A New Standard for Molecular Informatics, Alan McNaught, *Chemistry International*, November-December 2006, **8**(6), 12-14.
- Using InChI, Jeremy Frey, *Chemistry International*, November-December 2006, **8**(6), 14-15.
- Chemical 'Naming' Method Unveiled, *Chem. & Eng. News*, 22 Aug 2005 [link to *CEN*]
- Analysis of a Set of 2.6 Million Unique Compounds gathered from the Libraries of 32 Chemical Providers, A. Monge, A. Arrault, C. Marot and L. Morin-Allory, presented at the *10th Electronic Computational Chemistry Conference*, April 2005
- International chemical identifier goes online, *Chem. World*, 16 May 2005 [link to *CW*]
- Application of InChI to Curate, Index, and Query 3-D Structures, M.D. Prasanna, J. Vondrasek, A. Wlodawer and T.N. Bhat, *Proteins: Structure, Function, and Bioinformatics*, 2005, **60**, 1-4
- Enhancement of the chemical semantic web through the use of InChI identifiers, S.J. Coles, N.E. Day, P. Murray-Rust, H.S. Rzepa and Y. Zhang, *Org. Biomol. Chem.*, 2005, **3**(10), 1832-1834
- InChI FAQ, by Nick Day (Unilever Centre for Molecular Informatics, Cambridge University)
- Representation and Use of Chemistry in the Global Electronic Age, P. Murray-Rust, H.S. Rzepa, S.M. Tyrrell and Y. Zhang, *Org. Biomol. Chem.*, 2004, 3192-3203 [www.ch.ic.ac.uk/rzepa/obc/]
- That InChI feeling, *Reactive Reports*, issue 40, Sep 2004
- Unique labels for compounds, *Chem. & Eng. News*, 2 Dec 2002
- Chemists synthesize a single naming system, *Nature*, 23 May 2002
- That InChI feeling ... *The Alchemist*, 24 Apr 2002
- What's in a Name? *The Alchemist*, 21 Mar 2002

Reference Library of Peptide Ion Fragmentation Spectra

Contact: S.E. Stein (Physical and Chemical Properties Div., Gaithersburg)

Context and Approach:
Mass spectrometric measurements on proteins and peptides have attracted increasing interest over the last decade as the field of proteomics became dominant in many biochemical and biological investigations. However, current methods for peptide identification use only a fraction of the information in the mass spectra of peptide ions. Identifying peptides by matching their MS/MS spectra to reference spectra is an effective way of using all spectral information leading to faster and more reliable peptide and protein identifications than current sequence-based methods.

Current peptide identification methods, illustrated below, match each measured MS/MS spectrum against 'theoretical' spectra of all possible peptide sequences. Since relative abundances, neutral losses from parent and product ions, and ratios of products having different charge states are not predictable, this rich, peptide-specific information is not effectively used for establishing identity. Also, prior occurrence information is ignored – each search identifies the peptide as if for the first time. Therefore, for more rapid and reliable identifications, a library of peptide ion fragmentation spectra is being developed and tested. In addition, as seen above, the same spectrum can have very different scores from different search engines.

Our approach has been to use collision-induced dissociation (CID) mass spectra of peptide ions are collected in our laboratory with the different mass spectrometers as well as large numbers of spectra are also obtained from other laboratories for analysis of complex clinical specimens. The measurements from our laboratory on pure peptides and well-defined mixtures provide benchmark measurements under carefully controlled conditions. The resulting data from all sources is processed to provide spectra for inclusion into the library: 1) acquire and organize 'Shotgun' proteomics data files from diverse sources, 2) identify peptides with available sequence search engines, 3) create a 'consensus spectrum' from all replicate spectra and find best single spectrum for each peptide ion, 4) derive reliability measures for each spectrum, 5) remove ambiguities and build the library.

The library is based on both a 'Consensus spectrum' and the best experimental spectra from all spectra matching a given peptide ion. To find the consensus spectrum we first align m/z peaks from all spectra, rejects outliers and accept only ions that are present in a majority of the spectra that might have generated the peak. In addition, the best replicate spectrum based on search engine scores and spectrum quality is retained.

We then calculate a measure of reliability by examining a) spectrum/sequence consistency (to match theoretical spectrum, based on relative dissociation rates of adjacent amino acids), b) peptide sequence, and c) peptide class, including those with from semitryptic and other irregular cleavages. The resulting spectra are used to build a library, creating annotated spectra for consensus and best matching single spectra, and

resolve problems of similar spectra generating multiple peptides (homologies, small peptides, etc.).

Major Accomplishments:
The first version of the peptide library was completed and distributed to several laboratories for testing and the library made available through a public resource site (http:www.proteomecommons.org/tools.asp). It includes data collected in our laboratory and obtained from other laboratories. Automated quality control methods have been developed to deal with the inherent variability of energy, hence spectral patterns, of this class of mass spectra.

All spectra in this library originated from peptide ions generated by electrospray ionization in LC-MS/MS experiments. Most spectra were acquired with ion trap mass spectrometers, with a smaller number from low energy collision cell instruments (qtof-class). Three types of spectra are provided: 1) 'consensus' spectra - derived from multiple (replicate) identifications of a peptide ion, 2) best 'replicate' spectra and 3) high confidence 'single' spectrum identifications. The library is divided into five sub-libraries:

      1) Yeast – 35,135 Consensus, 35,684 Replicate, and 2,462 Single spectra (Saccharomyces Cerevisiae, or Baker's Yeast). This is the most complete collection. It was derived from 2509 data files from 13 laboratories and has served as the principal test-bed for library development.

      2) D. Radiodurans – 8,273 Consensus, 8,477 Replicate, and 287 Single spectra – All spectra for this radiation-resistant bacterium were derived from a web-based collection of ion trap spectra provided by PNNL/NCRR [http://ncrr.pnl.gov/data/].

      3) M. Smegmatis – 3,569 Consensus, 3,578 Replicate, and 133 Single spectra – All spectra for this bacterium originated from ion trap data files provided in the Open Proteomics Database [http://bioinformatics.icmb.utexas.edu/OPD/].

      4) Individual proteins – 3,938 Consensus, 3,922 Replicate, and 15 Single spectra – NIST-measured spectra from separate digests of 19 different proteins on ion trap instruments.

      5) Human – 45,377 Consensus, 46,053 Replicate, and 1,940 Single spectra (Homo sapiens) – These spectra were derived from a range of available sources, including labs involved in the Human Proteome Organization plasma proteome project. In addition, it contains spectra from peptides originating from a specific sample types, including leukocytes, hair, and saliva.

      This collection is intended primarily to demonstrate the utility of peptide ion fragmentation libraries. Emphasis has been placed on quality control methods, since it has been found that erroneous spectra, especially those containing significant impurity peaks, can lead to false positive results. In order to minimize false negative results (peptide ion not in library), extensive extraction methods were employed. These methods improved the separation of true and false positive results and combined results of several sequence search engines for initial peptide identification. Spectra were annotated to aid spectrum library scoring as well as to document the origin of the spectrum. It is important to note that none of these collections, especially the human library, are complete.

Additional measurements will substantially increase both the coverage and the quality of the collection.

Impact:

The library is already used for direct peptide identification and for sensitive detection of internal standards, biomarkers, targets proteins with capability to subtract a component from a mixture spectrum. It is sensitive, reliable, fast, and comprehensive, and it can competently search all spectra against library. The library confirms/rejects peptides identified by sequence search programs by comparing to reference spectra. It also links peptides between runs for later processing and identification. It is being incorporated into several proteomics 'pipelines' for integration into practical proteomics analysis.

Future Plans:

The libraries, especially those from human samples and standard proteins, are being actively expanded and collaborative work with proteomics centers will increase. We are also providing tested algorithms to search the library and demonstrating its practical applications. The ultimate goal is to have the library installed on and integrated within the data systems of all relevant mass spectrometers and data analysis systems.

Project Team: S.E. Stein, Y. Mirokhin, P. Neta, Q.-L. Pu, J. Roth, P.A. Rudnick, M.A. Wingerd, X. Yang (838-G); L. Kilpatrick (838, HML)

Presentations:

54th American Society for Mass Spectrometry Conference on Mass Spectrometry (2006). Human Proteome Organization (HUPO) 5th Annual World Congress (2006).