

THERM: A Computer Code for Estimating Thermodynamic Properties for Species Important to Combustion and Reaction Modeling

EDWARD R. RITTER[†]

Department of Chemical Engineering, Chemistry, and Environmental Science, New Jersey Institute of Technology, Newark, New Jersey 07102

Received January 14, 1991

A computer package has been developed called THERM, an acronym for **T**hermodynamic property **E**stimation for **R**adicals and **M**olecules. THERM is a versatile computer code designed to automate the estimation of ideal gas phase thermodynamic properties for radicals and molecules important to combustion and reaction-modeling studies. Thermodynamic properties calculated include heat of formation and entropies at 298 K and heat capacities from 300 to 1500 K. Heat capacity estimates are then extrapolated to above 5000 K, and NASA format polynomial thermodynamic property representations valid from 298 to 5000 K are generated. This code is written in Microsoft Fortran version 5.0 for use on machines running under MSDOS. THERM uses group additivity principles of Benson and current best values for bond strengths, changes in entropy, and loss of vibrational degrees of freedom to estimate properties for radical species from parent molecules. This ensemble of computer programs can be used to input literature data, estimate data when not available, and review, update, and revise entries to reflect improvements and modifications to the group contribution and bond dissociation databases. All input and output files are ASCII so that they can be easily edited, updated, or expanded. In addition, heats of reaction, entropy changes, Gibbs free-energy changes, and equilibrium constants can be calculated as functions of temperature from a NASA format polynomial database.

INTRODUCTION

Detailed reaction mechanisms are widely used in simulations of processes such as combustion and chemical vapor deposition (CVD). Such models, based upon fundamental thermodynamic and kinetic principles, can offer insight into the controlling chemistry of such complex phenomena. In developing detailed reaction mechanisms, researchers must supply accurate thermodynamic data over a wide temperature range for all stable and radical species considered in a mechanism. The THERM computer code was developed to aid in estimating this data.

THERM relies upon the principles of group additivity as developed by Benson and co-workers¹ to estimate thermodynamic properties for species where literature data are not available. Benson's group additivity is referred to as a second-order estimation technique² since it incorporates non nearest neighbor corrections and steric effects. Less sophisticated estimation techniques include first-order or bond additivity methods and zero-order or atomic contribution methods. It should be noted that with the exception of a few properties such as the molecular weight, zero-order atomic contributions are of little value. First-order methods work well for simple molecules such as normal hydrocarbons. These, however, are less accurate for more complex molecules. Second-order group contribution techniques incorporate important corrections for cyclization, gauche interactions, steric effects, repulsive and attractive effects for aromatic substituents, etc. In principle, there is no limit to the number of interaction groups which can be included, nor to the accuracy which can be obtained when these effects are taken into consideration. There are two limitations to this approach, however. First, there are only limited thermodynamic data available to determine the interaction contributions. Secondly, one must recognize the interactions of importance a priori, or resulting estimates will be less accurate than anticipated. Nevertheless, this method of group contributions has been embraced as the best all-around method for estimating ideal gas thermodynamic properties.²⁻⁵

THERM includes groups for hydrocarbons, oxygen, nitrogen, halogen, and sulfur-containing species. Ring correction

and other interaction groups are also included to correct for such things as optical isomers. One aspect that distinguishes THERM from other computer implementations of group additivity is the bond dissociation group data file. These data are included for the estimation of radical species important to high-temperature reaction-modeling and combustion studies. Groups include bond dissociation energy, entropy change, and heat capacity change reflecting loss of an H atom from a parent stable molecule.

Other computer codes^{2,6} have been developed to use Benson's group additivity method. Most notable is the CHETAH program distributed by the ASTM.^{2,7,8} CHETAH is designed specifically for use in chemical reaction hazard assessment, which is not addressed by THERM. THERM was designed specifically to automate several routine tasks encountered in reaction-modeling studies. THERM outputs data directly in a format that can be used with such industry-standard modeling codes as CHEMKIN,⁹ HTC,¹⁰ and STANJAN.¹¹ In addition, THERM automates the update of molecules to reflect changes in group contributions. It should be emphasized that some group contributions have been estimated based upon other group contributions due to the lack of available thermodynamic property data. It is desirable to easily recalculate thermodynamic property estimates to allow new data to be incorporated as it becomes available. This is especially true for bond-dissociation energies, where estimates and measurements are continually being refined.

Group Contribution Database. Group contributions were derived by Benson and co-workers and various other researchers^{1,4,12-21} by breaking similar groups of molecules with known thermodynamic properties into their constituent groups and then performing multivariable linear regression to find group contributions that gave the best fit to available experimental data.

Sources for the group database are presented elsewhere.⁵ Group data files contain group identification (name) followed by contributions to ΔH_f° and S° at 298 K and C_p 's at 300, 400, 500, 600, 800, 1000, and 1500 K (if available). This data is arranged in a tabular format, with a reference included for all groups not directly from Benson. The reference information is on the same line as the group data, following the C_p entry for 1500 K. THERM allows an on-line review of the group database including references entries.

[†] Present address: Department of Chemical Engineering, Villanova University, Villanova, PA 19085.

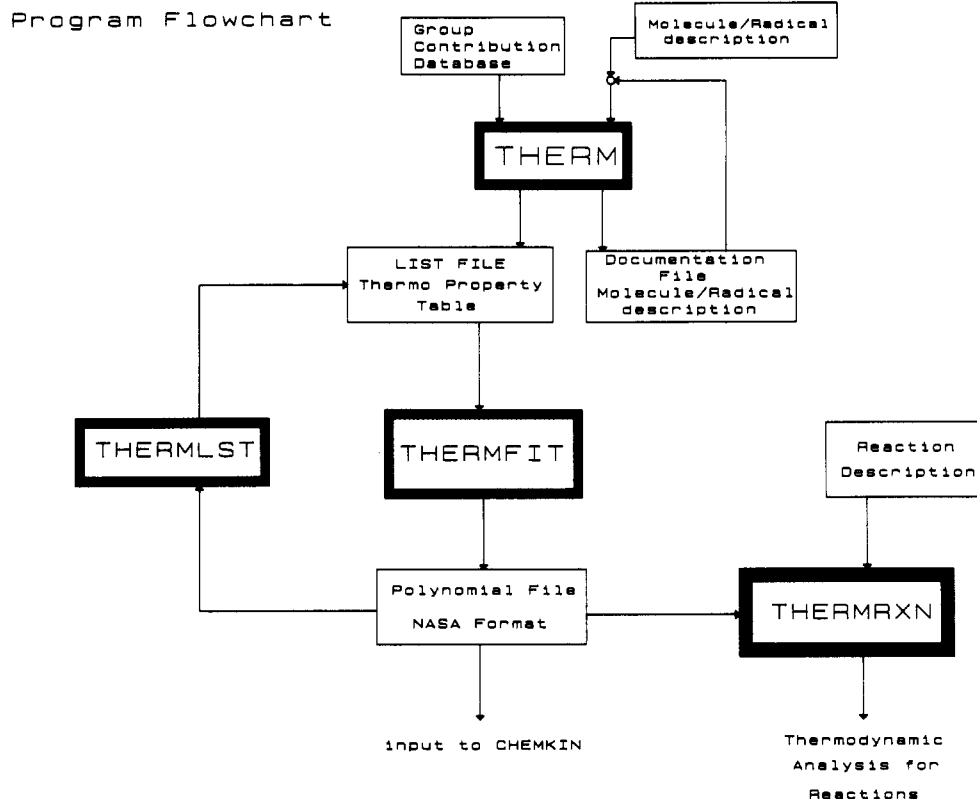


Figure 1. Functional diagram for the THERM ensemble of programs. Emphasized boxes are used to distinguish programs from file input and output. THERM consists of four programs: THERM, THERMFIT, THERMLST, and THERMRXN.

The group database is subdivided into seven files, currently containing more than 300 entries. A user may choose to change the number of files or the file names by changing information in the configuration file, THERM.CFG. When THERM is running, all group data, including reference information, are memory resident. As a result, searches for group information and calculation of properties are nearly instantaneous. This large memory requirement limits the number of groups in memory to 400. If a database is expanded beyond 400 entries, the program will prompt the user to choose which files are needed for the current session when the program is first started.

Group data files supplied with THERM contain primarily, group values of Benson¹ and Stein and Fahr.¹⁷ Additional groups that were derived as part of this work have been made consistent with Benson group values. Other hydrocarbon group data have recently been published by researchers such as Cohen¹² and Domalski;⁴ however, these do not form a comprehensive database. Although the revised hydrocarbon group values are based upon more recent experimental data, Benson group values have been selected for distribution with THERM. Mixing revised hydrocarbon groups with other Benson groups, such as those for halogens, may result in erroneous estimates. This may occur since halogen and other Benson groups were derived with Benson's hydrocarbon group values and not the revised hydrocarbon group values. Group data files can be edited, however, so that a user can employ any groups he chooses. Group data files, like all of THERM's files, are ASCII and can be easily altered by using any convenient text editor to change group values or add new groups to the data files.

THE THERM COMPUTER PACKAGE: STRUCTURE AND FUNCTIONS

Figure 1 presents a functional diagram which shows the relation of the THERM ensemble of computer programs to one another. Individual executable computer programs are

emphasized with heavy boxes to distinguish them from interactive or file input and output. The parent program is THERM, which uses a group contribution database in addition to a molecule/radical description to estimate the thermodynamic properties for species of interest. THERM generates both a tabular listing or "LIST" file and a documentation or "DOC" file which contain the information about the species that were estimated or entered to the database. The LIST file contains the heat of formation (ΔH_f°) and entropy (S°) at 298 K (ΔH_f°), heat capacities (C_p) between 300 and 1500 K, creation data, phase (gas), number of internal rotors, and elemental composition in a 132-column table format (see Table I). The DOC file contains this information along with the groups which were considered, species molecular formula, and symmetry correction to the entropy. DOC file entries for radical species include the bond-dissociation energy and entropy correction for electron spin in addition. Examples of this format are presented in Figures 3b and 4b. As shown in Figure 1, the molecule description either can be input from the keyboard by the user or can be read from a DOC file. An entire DOC file can be processed automatically, allowing database entries (DOC and LIST files) to be recalculated to reflect changes to the group contribution database. This function is particularly useful when one wishes to incorporate new literature data for bond energies or to examine the effect of uncertainties in group values. This allows the database to remain flexible to new input, eliminating a significant barrier to the use of new data when it becomes available. Data in Table I and Figure 3b are shown with units of kJ/mol (ΔH_f°) and J/(mol·K) (S and C_p); however, units can easily be toggled between these units and kcal/mol (ΔH_f°) and cal/(mol·K) (S° and C_p). THERM can read, write, and translate files in either set of units.

The THERM parent programs can spawn the three child processes THERMFIT, THERMLST, and THERMRXN. The subdivision of the package into these programs was necessitated by memory limitations encountered under MSDOS.

Table I. Example Thermodynamic Property Table (Output as a LST File)^a

species	$\Delta H_f^\circ(298)$	$S^\circ(298)$	CP300	CP400	CP500	CP600	CP800	CP1000	CP1500	date	ref	elements											
CO	-26.40	47.20	6.82	6.98	7.15	7.31	7.60	7.86	8.38		121286	C	1	O	1	0	0	0	G	0			
CO2	-94.01	51.00	8.92	9.83	10.61	11.26	12.26	12.95	13.86		121286	C	1	O	2	0	0	0	G	0			
H2O	-57.80	45.10	8.23	8.41	8.64	8.91	9.51	10.14	11.50		20387	H	2	O	1	0	0	0	G	0			
HCO	10.40	53.60	8.44	8.90	9.38	9.87	10.82	11.66	13.03		121286	H	1	C	1	O	1	0	G	0			
HO2	2.50	54.70	8.36	8.91	9.42	9.88	10.69	11.33	12.37		20387	H	1	O	2	0	0	0	G	0			
ClO	24.10	54.10	7.55	7.90	8.15	8.34	8.57	8.69	8.81		JANAF	Cl	1	O	1	0	0	0	G	0			
HF	-65.11	41.50	6.94	6.98	7.03	7.08	7.22	7.38	7.77		JANAF	H	1	F	1	0	0	0	G	0			
H	52.10	27.30	4.90	4.90	4.90	4.90	4.90	4.90	4.90			H	1		0	0	0	0	G	0			
HCl	-22.00	44.60	6.86	6.92	7.00	7.09	7.28	7.48	7.97		JANAF	H	1	Cl	1	0	0	0	G	0			
C2F4	-157.41	71.71	19.18	21.83	23.89	25.48	27.59	28.80	30.31		JANAF	C	2	F	4	0	0	0	G	0			
C2F2	5.00	58.31	13.70	15.03	16.11	16.97	18.21	18.98	19.89		JANAF	C	2	F	2	0	0	0	G	0			
C2Cl2	50.10	65.01	16.53	17.47	18.22	18.81	19.64	20.11	20.55		JANAF	Cl	2	C	2	0	0	0	G	0			
CF3CN	-118.41	71.31	18.72	21.39	23.48	25.11	27.32	28.62	30.19		JANAF	C	2	N	1	F	3	0	G	0			
C2Cl4	-2.90	82.01	22.53	24.49	26.01	27.17	28.72	29.61	30.71		JANAF	Cl	4	C	2	0	0	0	G	0			
C2HCl	51.10	57.81	13.42	14.64	15.67	16.51	17.78	18.62	19.68		JANAF	Cl	1	C	2	H	1	0	G	0			
C2HF	30.00	55.31	12.66	13.90	14.98	15.90	17.33	18.30	19.46		JANAF	C	2	H	1	F	1	0	G	0			
C2H2	54.10	48.00	10.48	11.80	12.92	13.87	15.36	16.44	18.13		J3/61	C	2	H	2	0	0	0	G	0			
C2H4	12.50	52.30	10.34	12.72	14.86	16.76	19.94	22.37	26.10		J9/65	C	2	H	4	0	0	0	G	0			
C2H6	-20.20	54.81	12.52	15.73	18.64	21.26	25.71	29.21	34.71		L5/72	C	2	H	6	0	0	0	G	1			
biphenyl	42.90	93.91	39.30	52.90	64.11	73.30	86.94	96.13	109.07	3/20/89	THERM	C	12	H	10	0	0	0	G	1			
propane	35.00	68.51	18.39	22.46	26.14	29.47	35.20	39.84	47.71	3/20/89	THERM	C	3	H	8	0	0	0	G	2			
cyc3ene	66.61	58.31	12.76	16.22	19.13	21.56	25.29	27.94	31.96	3/20/89	THERM	C	3	H	4	0	0	0	G	0			
cyc7	-28.80	81.91	31.77	44.12	54.66	63.62	77.63	87.60	101.90	3/20/89	THERM	C	7	H	14	0	0	0	G	0			
cyc7lc	43.70	75.61	25.61	34.24	41.49	47.55	56.83	63.30	72.65	3/20/89	THERM	C	7	H	8	0	0	0	G	0			
cyc8de15	13.00	83.81	33.25	45.16	55.20	63.62	76.55	85.56	98.28	3/20/89	THERM	C	8	H	12	0	0	0	G	0			
cyc8de13	18.60	85.91	33.34	44.77	54.46	62.65	75.35	84.35	97.39	3/20/89	THERM	C	8	H	12	0	0	0	G	0			

^a Units are in kcal.

```

THERMO
300., 1000., 5000.
C6H5O 3/20/89 THERMC 6H 50 1 OG 300.000 5000.000 1404.000 01
1.57324764E+01 1.54394760E-02-5.31383756E-06 8.28591116E-10-4.82238305E-14 2
-1.76799222E+03-6.20052317E+01-3.73072513E+00 6.71840729E-02-5.88121307E-05 3
2.61335718E-08-4.61968583E-12 4.31827564E+03 4.02213334E+01 4
CH3OCH3 3/20/89 THERMC 2H 60 1 OG 300.000 5000.000 1364.000 21
8.43631162E+00 1.30235626E-02-4.45885307E-06 6.93233262E-10-4.02757540E-14 2
-2.62777965E+04-2.24530764E+01 1.53263200E+00 2.37669572E-02-8.14177550E-06 3
-5.29451929E-10 5.96605679E-13-2.32831259E+04 1.66340885E+01 4
END

column numbers:

1 10 20 30 40 50 60 70 80
|23456789|123456789|123456789|123456789|123456789|123456789|123456789|123456789|
|<-name>|<-date>|<ref>|<---- elements ---->G|<- Tlow >|<- Thi->|<- Tbk->| 21
|<---- a1 ---->|<---- a2 ---->|<---- a3 ---->|<---- a4 ---->|<---- a5 ---->| 2
|<---- a6 ---->|<---- a7 ---->|<---- a8 ---->|<---- a9 ---->|<---- a10 ---->| 3
|<---- a11 ---->|<---- a12 ---->|<---- a13 ---->|<---- a14 ---->| 4

```

Figure 2. NASA format polynomial. Four lines per species, number of rotors appears in column 79 of first line. Line 1: species name column 1-9; date appears in columns 11-19; reference columns 20-24; elements appear in columns 26-44; phase column 45; low-temperature limit, high-temperature limit, break point temperature, number of rotors in columns 78-79; line number "1" for first line of species, column 80. Line 2: coefficients a_1 - a_5 columns 1-75; line number "2" for second line of species, column 80. Line 3: coefficients a_6 - a_{10} columns 1-75; line number "3" for third line of species, column 80. Line 4: coefficients a_{11} - a_{14} in columns 1-60; line number "4" for fourth (last) line for this species, column 80.

Each of these child programs fulfills a specific function. They can be run either from within the THERM main program or directly from the DOS prompt. THERMFIT is used to extrapolate C_p data contained in a LIST file from 1000 to 5000 K and create NASA format polynomials^{22,23} for direct use by CHEMKIN and other codes which use this format. The resulting polynomial representations of thermodynamic data are valid from 298 to 5000 K and are written to a file that contains polynomial coefficients.

Polynomial representations for heat capacity, enthalpy, and entropy functions use the following equations:

heat capacity

$$C_p(T) = R(a_1 + a_2T + a_3T^2 + a_4T^3 + a_5T^4) \quad T > T_{bk}$$

$$C_p(T) =$$

$$R(a_8 + a_9T + a_{10}T^2 + a_{11}T^3 + a_{12}T^4) \quad T < T_{bk} \quad (1)$$

enthalpy function

$$H^o(T) = R[a_1T + (a_2/2)T^2 + (a_3/3)T^3 + (a_4/4)T^4 + (a_5/5)T^5 + a_6] \quad T > T_{bk}$$

$$H^o(T) = R[a_8T + (a_9/2)T^2 + (a_{10}/3)T^3 + (a_{11}/4)T^4 + (a_{12}/5)T^5 + a_{13}] \quad T < T_{bk} \quad (2)$$

entropy

$$S^o(T) = R[a_1 \ln(T) + a_2T + (a_3/2)T^2 + (a_4/3)T^3 + (a_5/4)T^4 + a_7] \quad T > T_{bk}$$

$$S^o(T) = R[a_8 \ln(T) + a_9T + (a_{10}/2)T^2 + (a_{11}/3)T^3 + (a_{12}/4)T^4 + a_{14}] \quad T < T_{bk} \quad (3)$$

where R is the gas constant, T is the temperature in Kelvin, a_1 - a_{14} are the polynomial coefficients described in Figure 2, and T_{bk} is the break point temperature (a point of forced tangency between the two polynomials).

THERMFIT accepts input directly from the keyboard or from a LIST file and creates NASA format polynomials similar to that shown in Figure 2. The NASA format includes four lines for each species. The first contains the species name, creation date/reference, elemental composition, phase, lower and upper temperature limit, break point temperature, number of internal rotors in the species, and the index number 1. The

three subsequent lines contain the 14 polynomial coefficients as described in the caption to Figure 2. Thermodynamic properties are calculated from these coefficients by using eqs 1-3.

Since these polynomial coefficients cannot be easily reviewed for thermodynamic property data, the program THERMLST is included to create a LIST tabular format from polynomial files. This program can be used to list other NASA polynomial files such as those supplied with CHEMKIN.

THERM contains a program called THERMRXN. THERMRXN is used to examine heats of reaction. To use THERMRXN one simply enters a reaction expression from the keyboard such as the following:



Species are delimited by the operations + and =, and spaces are not required. In this reaction expression, H is an H atom, CY13PD is cyclopentadiene, C*CC. is allyl radical, and C2H2 is acetylene. $\Delta H_{\text{reaction}}$, $\Delta S_{\text{reaction}}$, $\Delta G_{\text{reaction}}$, and the equilibrium constant K_c are then calculated and displayed as functions of temperature as shown in Table II.

In this example one can see that equilibrium strongly favors the products allyl radical plus acetylene for addition of H atom to cyclopentadiene at temperatures above 800 K. Similar thermodynamic analyses for reaction can be obtained for any reaction involving species in the database. This can be useful for checking the thermodynamic consistency of reactions in a detailed reaction mechanism. THERMRXN also accepts and processes CHEMKIN reaction description input files to facilitate examination of larger systems of reactions.

Additional features worth noting include on-line help, editing capabilities, and file conversion utilities. The THERM main program comes with a significant amount of help accessed simply by entering a question mark "?" in response to a prompt for input. Errors or modifications to species that were estimated may be addressed by THERM's built-in species editor. This is not a sophisticated full-screen editor, but allows one to change groups, symmetry numbers, etc. without leaving the program or aborting the current molecule and starting again. In addition, one can make modifications to bond energies or calculated property values. All changes are recorded. Lastly, THERM and THERMRXN utilize formatted direct-access ASCII files for speed. This format requires that each line be the same length so that the program can access the file directly by line number. Since ASCII editors create sequential files

Table II. Example Thermodynamic Property Change for Reaction Output

Rx	H	+	CY13PD	=	C*CC.	+	C2H2
ΔH_f° (kcal/mol)	52.100		32.400		40.750		54.100
S (cal/mol·K)	27.300		64.100		63.010		48.000
dH_f (kcal/mol) (298 K) = 10.35				dH_f av (298.0, 1500.0 K) = 9.26			
dU (dE) (kcal/mol) (298 K) = 10.35				dU av (298.0, 1500.0 K) = 9.26			
dS_f (cal/mol·K) (298 K) = 19.61				dS_f av (298.0, 1500.0 K) = 18.77			
dG_f (kcal/mol) (298 K) = 4.50				dG_f av (298.0, 1500.0 K) = -7.61			
K_c (298 K) = 4.997E - 04				K_c av (298.0, 1500.0 K) = 7.095E + 01			
T (K)	dH (kcal/mol)	dU (kcal/mol)	dS (cal/mol·K)	K_c	dG (kcal/mol)		
300.00	1.035E+01	1.035E+01	1.963E+01	5.566E-04	4.467E+00		
400.00	1.054E+01	1.054E+01	2.016E+01	4.455E-02	2.473E+00		
500.00	1.059E+01	1.059E+01	2.028E+01	6.369E-01	4.482E-01		
600.00	1.053E+01	1.053E+01	2.018E+01	3.751E+00	-1.576E+00		
800.00	1.021E+01	1.021E+01	1.972E+01	3.326E+01	-5.570E+00		
1000.00	9.708E+00	9.708E+00	1.917E+01	1.169E+02	-9.460E+00		
1200.00	9.120E+00	9.120E+00	1.863E+01	2.579E+02	-1.324E+01		
1500.00	8.172E+00	8.172E+00	1.793E+01	5.343E+02	-1.872E+01		
2000.00	6.552E+00	6.552E+00	1.700E+01	9.974E+02	-2.744E+01		

(lines of different lengths), THERM comes with a file utility called STOD (sequential to direct) to accommodate file conversion after editing a DOC or polynomial file.

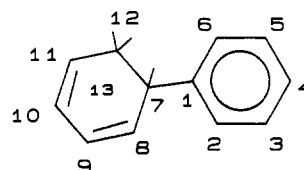
ESTIMATION OF MOLECULAR AND RADICAL THERMODYNAMIC PROPERTIES

To utilize the Benson approach to thermodynamic property estimation, one must first become familiar with the notation used to designate each group. Each molecule is broken down into its constituent groups by the user, and then groups are entered. There are a few rules used to describe group names. Groups for saturated carbons always have the four atoms bonded to this center represented in the group name. Groups for olefinic carbons, on the other hand, contain only the two atoms attached to this center by single bonds in the name; while nonfused ring aromatic carbons contain only the atoms which are attached to the ring by single bonds in the group name. For example, a terminal methyl group in a normal alkane is represented by the group C/C/H3, a terminal olefinic carbon is represented by CD/H2 (CD stands for double bonded carbon), and any carbon in benzene is represented by CB/H (CB stands for benzene-type carbon). Corrections for cyclic species (ring corrections) begin with the letters CY designating a cyclic correction. The number of carbon atoms follow, along with an "E" designating alkene or "DE" for diene if applicable. The ring corrections for cyclopropane, cyclopropene, cyclopentane, cyclopentene, cyclopentadiene, and cyclohexa-1,4-diene are represented by CY/C3, CY/C3E, CY/C5, CY/C5E, CY/C5/DE, and CY/C6/DE/14, respectively. Ring corrections for heterocyclic species such as furan, ethylene oxide, and pyrrole are represented by CY/C4O/DE, CY/C3O, and CY/C4N/DE, respectively. Benzene and fused aromatic rings require no additional ring corrections; however, they may require correction for substituent interactions.

Estimation of molecules is best illustrated with an example. Figure 3a shows 5-phenyl-1,3-cyclohexadiene and its associated groups. Group 1 (CB/C) represents a nonfused aromatic carbon (similar to benzene) that is attached to a saturated carbon. Groups 2-6 (CB/H) represent nonfused aromatic carbons attached to hydrogens. Group 7 (C/CB/CD/C/H) represents a saturated carbon that is attached to a benzene carbon (CB), an olefinic carbon (CD), another saturated carbon (C), and a hydrogen atom. Groups 8 and 11 (CD/C/H) represent olefinic carbons attached to saturated carbons and hydrogens. Groups 9 and 10 (CD/CD/H) represent olefinic carbons attached to a second olefinic carbon and a hydrogen. Group 12 (C/C/CD/H2) represents a saturated carbon attached to another saturated carbon, an olefinic

a

Phenyl-cyclohexadiene



Groups Considered

#	id	quantity
1	CB/C	1
2-6	CB/H	5
7	C/CB/CD/C/H	1
8+11	CD/C/H	2
9+10	CD/CD/H	2
12	C/CD/C/H2	1
13	CY/C6/DE/13	1

b

```

SPECIES
PHCHD
Thermo estimation for molecule
PHCHD
UNITS:KCAL
GROUPS 7
Gr # - GROUP ID - Quantity  Gr # - GROUP ID - Quantity
1 - CB/H - 5 | 5 - CD/C/H - 2
2 - CB/C - 1 | 6 - CD/CD/H - 2
3 - C/CB/CD/H - 1 | 7 - CY/C6/DE/13 - 1
4 - C/C/CD/H2 - 1
Hf S Cp 300 400 500 600 800 1000 1500
50.62 100.88 40.91 55.77 68.49 78.61 93.59 103.84 118.58
CPINF = 138.10
NROTORS: 1
SYMMETRY 2
CREATION DATE: 4/30/91
ENDSPECIES

```

Figure 3. (a) Structure for 5-phenyl-1,3-cyclohexadiene is shown decomposed into its constituent group. (b) Image of the documentation file output for 5-phenyl-1,3-cyclohexadiene.

carbon, and a hydrogen atom. The ring correction for a cyclohexadienyl ring system is represented by group 13 (CY/C6/DE/13). After entering these group identifications, the estimated thermodynamic properties are displayed as shown in Figure 3b.

There are a few differences between the group names listed elsewhere^{1,4,21} and those used with THERM. For example, all letters in a group name are capitalized. The ligands' identifications follow that for the central atom in alphabetical order separated by a slash "/". Therefore, Benson's C-(H)₃C becomes C/C/H3 and C-(Cl)₂(H)(C) becomes C/C/Cl2/H for the purposes of input to THERM. Alphabetic order for ligands was chosen, since this makes it much easier to re-

Table III

formula	species	$\Delta H_f^\circ(298\text{ K})$ (kcal/mol)			ref
		calcd	exp	δH_f°	
CH5N	methylamine	-5.28	-5.50	0.22	39
C2H3Cl	vinyl chloride	5.06	8.40	-3.34	38
C2H5O2N	nitroethane	-25.30	-24.45	-0.85	39
C2H7N	dimethylamine	-4.76	-4.45	-0.32	39
C3H7Cl	1-chloropropane	-31.93	-31.52	-0.41	39
C3H7Cl	2-chloropropane	-34.87	-34.63	-0.24	39
C3H8	propane	-25.33	-25.02	-0.31	39
C3H9N	trimethylamine	-5.84	-5.66	-0.17	39
C4H6	1-butyne	39.55	39.55	0.00	38
C4H6	2-butyne	34.70	34.97	-0.27	38
C4H11N	diethylamine	-18.20	-17.33	-0.87	39
C5H8	spiropentane	44.14	44.26	-0.12	39
C5H8	cyclopentene	7.82	8.10	-0.28	39
C5H10	cyclopentane	-18.74	-18.26	-0.48	39
C6H5Cl	chlorobenzene	12.30	12.43	-0.12	39
C6H5F	fluorobenzene	-27.86	-27.72	-0.14	39
C6H5Br	bromobenzene	25.10	25.19	-0.09	39
C6H8N2	phenyl hydrazine	48.30	48.49	-0.19	39
C6H12	cyclohexane	-29.58	-29.49	-0.09	39
C6H12	methyl cyclopentane	-25.91	-25.38	-0.53	39
C7H8O	anisole	-17.10	-16.23	-0.87	39
C8H10	m-xylene	3.82	4.13	-0.32	39
C9H12	mesitylene	-4.17	-3.80	-0.37	39
C10H7Cl	chloronaphthalene	28.59	28.63	-0.04	38
C12H8Cl2	4,4'-dichlorobiphenyl	28.01	28.94	-0.93	38
C14H12	1,1-diphenylethylene	58.64	58.71	-0.07	4
C16H10	pyrene	55.20	53.94	1.26	39

member the group names. This represents a significant improvement over the original notation presented in Benson¹ where the order appears somewhat arbitrary. In addition, the notation used for THERM requires fewer keystrokes and does not require the use of the shift key. These differences make data entry much more straight-forward. Group names can be changed to those of Benson if the database files are edited. This is left to the user if it is deemed important. Additional description of the groups used with this code is presented elsewhere.⁵

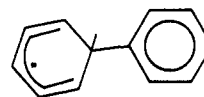
It should be pointed out that THERM makes no attempt to reduce a structure to groups, nor is there any plan to use a graphical interface to draw a molecule and have the program deduce groups. This is not the purpose for this code. The focus of this code is a variety of sophisticated extrapolation and fitting techniques, in addition to the automation of several file conversion and recalculation operations.

To estimate the thermodynamic properties for radical species, first estimate the parent molecule and then apply a bond dissociation increment to it to reflect the loss of an H atom. This is illustrated in Figure 4, parts a and b. Radicals may also be estimated directly by using Benson's groups for radical centers. The choice is left to the user.

Table III compares estimated ΔH_f° 's from THERM with literature data for several species. Data compared here agree within 1 kcal/mol in most cases. It should be pointed out that some classes of compounds do not fit as well as others. Classes such as polycyclic aromatic and chloroalkenes have larger errors in estimates from group additivity. Two examples are presented: an error of -3.34 kcal/mol for vinyl chloride and an error for pyrene of 1.26 kcal/mol. All other species shown are within 1 kcal/mol of the selected literature value. A comparison of several literature sources^{4,38,39} shows that reported experimental ΔH_f° 's may differ by as much as 1 kcal/mol for polycyclic hydrocarbons, while even larger uncertainties may exist for some chlorocarbons. Benson¹ gives typical error limits for group additivity ΔH_f° estimates to be ± 0.5 kcal/mol, but ranging as high as ± 3 kcal/mol for heavily substituted species. One must remain aware of these uncertainties.

a

Phenyl-cyclohexadienyl Radical



start with phchd parent

BD group: CHD

phchd. = phchd + chd

CHD = CYC6H7. - CY13HD

assumes loss of H atom
from the parent

b

```

SPECIES
PHCHD.
Thermo estimation for radical
PHCHD.
RADICAL BASED UPON PARENT PHCHD
PARENT FORMULA C6H5C6H7
PARENT SYMMETRY 2
UNITS:KCAL
GROUPS 8
Gr # - GROUP ID - Quantity Gr # - GROUP ID - Quantity
1 - CB/H - 5 5 - CD/C/H - 2
2 - CB/C - 1 6 - CD/CD/H - 2
3 - C/C/CB/CD/H - 1 7 - CY/C6/DE/13 - 1
4 - C/C/CD/H2 - 1 8 - CHD - 1
Hf S Cp 300 400 500 600 800 1000 1500
75.14 100.40 39.04 53.48 65.85 75.61 90.03 99.81 .00
CPINF = 132.14
NROTORS: 1
SYMMETRY 2
R ln(2) has been added to S to account
for unpaired electron
BOND 76.62
CREATION DATE: 4/30/91
ENDSPECIES

```

Figure 4. (a) Structure for phenylcyclohexadienyl radical is shown describing the bond-dissociation group. (b) Image of the documentation file output for phenylcyclohexadienyl radical.

EXTRAPOLATING THERMODYNAMIC PROPERTIES TO 5000 K WITH THERMFIT

As noted earlier, CHEMKIN and other computer codes utilize a modified NASA format polynomial representation for thermodynamic properties,²²⁻²⁵ as described by eqs 1-3 and Figure 2. This format utilizes two fourth-order polynomials to describe the heat capacity (C_p), enthalpy, and entropy functions from 298 to 5000 K. These polynomials have a forced tangency at a break point temperature. The low-temperature polynomial is utilized for temperatures below the break point, while the high-temperature polynomial is utilized for temperatures above the break point. In Figure 5 these polynomials are graphically compared with the data of Dorafeeva et al.²⁶ for the cycloheptatriene heat capacity function. A total of 10 parameters are required for this format. Group additivity, however, gives C_p estimates at only six temperatures covering only one-fifth the range of interest. Polynomials, therefore, cannot be generated directly from group additivity estimates. It is necessary to utilize more sophisticated models to extrapolate low-temperature C_p data to higher temperatures. Once a suitable extrapolation has been obtained, sufficient data points can be generated from this model to create the NASA format polynomials. The polynomial fitting method used by THERMFIT is presented elsewhere.⁵

THERMFIT includes three such more sophisticated C_p models allowing C_p estimates to be extrapolated to 5000 K. Two of these are now discussed. The preferred method (de-

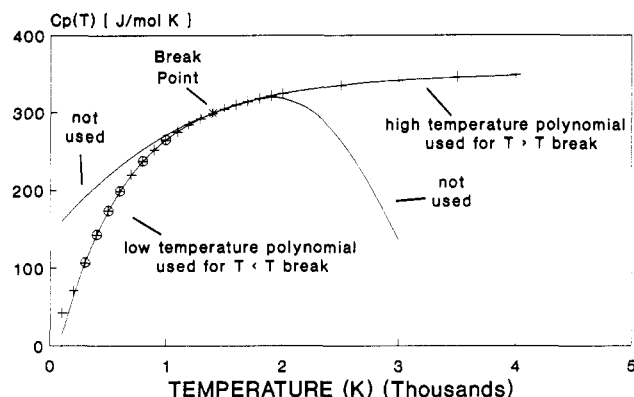


Figure 5. Fit of NASA format polynomials to the heat capacity function for cycloheptatriene. (O) THERM estimates used to generate the polynomials (solid lines), (+) literature data from Dorafieva et al., (*) break point temperature.

fault) used by this code is called the harmonic oscillator equation (HOE). This equation was developed by Ritter^{27,28} and requires a nonlinear regression. The HOE is based upon the harmonic oscillator model from statistical mechanics.²⁹⁻³¹ It is well-known that the heat capacity for an ideal gas is given by the following relations:

$$C_p = (7/2 + C_{vib})R \text{ linear molecule}$$

or

$$C_p = (4 + C_{vib})R \text{ nonlinear molecule} \quad (5)$$

where R is the gas constant and C_{vib} is given by the harmonic oscillator model as follows:

$$C_{vib} = \sum_i^s (\theta_{vi}/T)^2 \exp(\theta_{vi}/T) [\exp(\theta_{vi}/T) - 1]^{-2} \quad \theta_{vi} = \hbar c w_i / k \quad (6)$$

where \hbar is Planck's constant, c is the velocity of light, k is Boltzman's constant, w_i is the vibrational wavenumber in reciprocal centimeters, T is the temperature, and s is the number of vibrational degrees of freedom in the molecule.

The HOE replaces the summation over s terms with a three-term expression as follows:

$$C_{vib} = B_1 f(\theta_{v1}/T) + B_2 f(\theta_{v2}/T) + [s - (B_1 + B_2)] f(\theta_{v3}/T) \quad (7)$$

where

$$f(\theta_{vi}/T) = (\theta_{vi}/T)^2 \exp(\theta_{vi}/T) [\exp(\theta_{vi}/T) - 1]^{-2} \quad (8)$$

This expression is a five-parameter model for the heat capacity of an ideal gas. A slight variation on these expressions (eqs 5, 7, and 8) is employed by THERMFIT to extrapolate C_p estimates to high temperature for molecules with hindered internal rotations as well. This is demonstrated in Figure 6 for hexafluoroethane. The solid line is the THERM extrapolation using group additivity estimates, while the data points are from the JANAF tables.³² There is remarkable agreement between the extrapolated data and literature even for this molecule, which has a restricted internal rotor.

THERMFIT uses Marquardt's method³⁷ to determine the parameters B_1 , B_2 , θ_{v1} , θ_{v2} , and θ_{v3} . This equation is similar in principle to that of Lee et al.³³⁻³⁵ By constraining the coefficients (B_i) to sum to the number of vibrational modes (s), however, the HOE has fewer parameters than that of Lee. These parameters also retain physical significance.²⁷ The comparable version of Lee's equation (three-term) is designed to represent both vibrational and electronic contributions and contains seven parameters.³⁵ It should be noted that electronic contributions are typically negligible for most polyatomic species, and the term in Lee's equation that is designed to

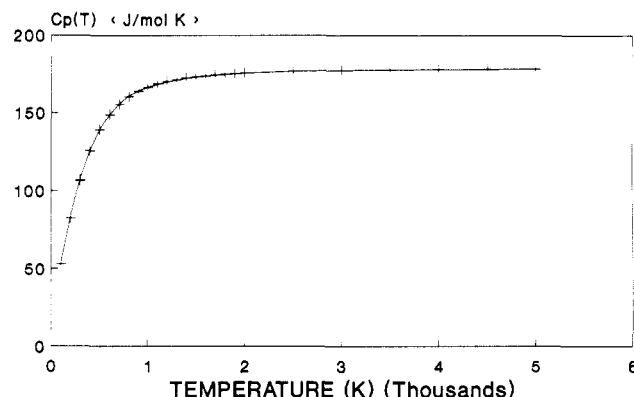


Figure 6. Comparison of THERM estimate for the heat capacity function of hexafluoroethane (solid lines) with data from the JANAF tables (+). Excellent agreement is attained by using the harmonic oscillator equation to extrapolate group additivity heat capacity estimates despite the hindered internal rotation about the C-C bond.

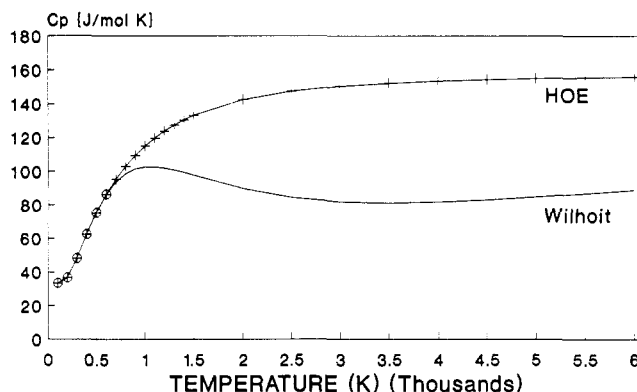


Figure 7. Comparison of HOE and Wilhoit extrapolations on heat capacity data for oxirane. Extrapolations are based upon data from 100 to 600 K (O). Solid lines are the two extrapolations; while (+) symbols are data from the JANAF tables.

represent electronic contributions actually represents vibrational contributions. This results in reduced physical significance for the parameters. Neither the HOE nor Lee's equation makes any attempt to fit or otherwise account for anharmonic contributions. It is recognized that these contributions are not easily quantified³⁶ and are typically neglected in most literature presentations of heat capacity data.^{32,36}

An additional method included with THERMFIT is the Wilhoit polynomial presented by Burcat and named after Randolph Wilhoit who developed the method.³⁶ This method generally gives a very good quality fit also; however, it does not extrapolate unconditionally. If given only five data points between 100 and 600 K, the Wilhoit polynomial may not extrapolate C_p adequately to high temperature. A comparison of such an extrapolation is presented in Figure 7. One can clearly see that the HOE method gives nearly a perfect extrapolation even with such limited data. This is due, in part, to the fact that the HOE method is based upon a fundamental model for the heat capacity of an ideal gas. The Wilhoit method, on the other hand, is a polynomial-based method. It should be noted that if reliable data were given up to 1500 or 2000 K, instead of only 600 K, the Wilhoit method would do an excellent job of extrapolation. This example is presented only to demonstrate how well the HOE method works, even when given so little data. The strength of Wilhoit's polynomial method lies in the fact that it requires only a linear regression (although a line-search method is used to solve for the steepness parameter³⁶). Regression is much more rapid; and since linear regressions have analytic solutions, Wilhoit parameters are obtained unconditionally. THERMFIT contains both techniques, and they are user-selectable. If the HOE fit

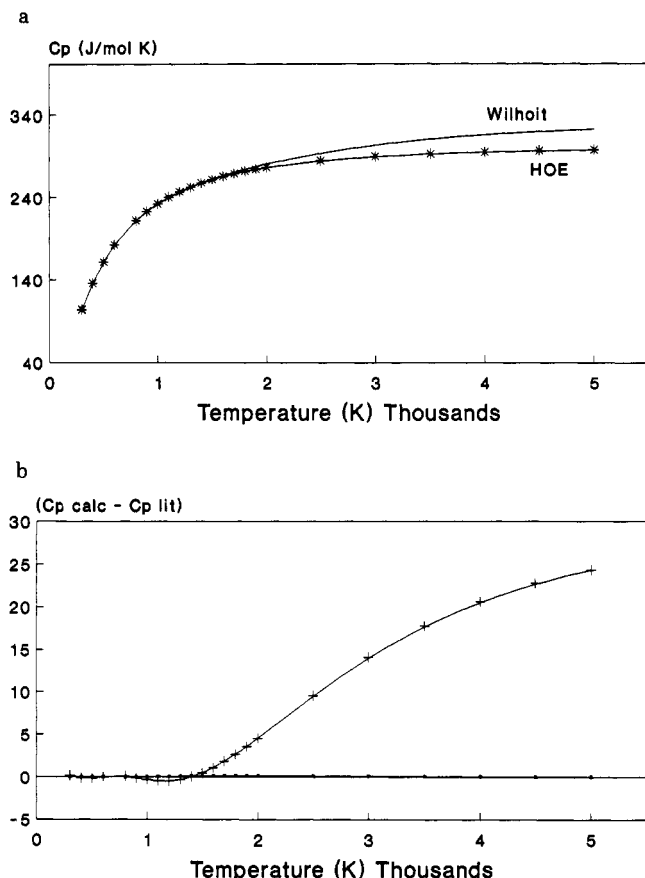


Figure 8. (a) Comparison of HOE and Wilhoit extrapolations for the heat capacity function of phenol. Solid lines are the extrapolations, while the symbols (+) are data from the JANAF tables. (b) Residuals in the extrapolation above show significant deviation by using Wilhoit parameters presented by Burcat,³⁶ while HOE extrapolations give little error.

should fail for any reason (it is a nonlinear regression), THERMFIT automatically defaults to the Wilhoit method.

Figure 8, parts a and b, compares a THERM estimate by using the HOE fit with data from Burcat for phenol.²² Also included is a fit using literature Wilhoit parameters.³⁶ The HOE extrapolation matches the data of Burcat much better than the extrapolation of the Wilhoit parameters. The HOE gives a maximum error of 1 J/(mol·K), while the Wilhoit data has a maximum error near 25 J/(mol·K). It should be noted that these Wilhoit parameters³⁶ have a stated temperature range from 298 to 1500 K, and therefore this extrapolation is well outside their range.

Accurate high-temperature thermodynamic data is critical to high-temperature kinetic modeling; however, its importance is often overlooked. Reaction modeling codes ensure thermodynamic consistency by utilizing thermodynamic data to calculate reaction rate constants for reverse reactions from equilibrium constants. Errors in heats of formation at elevated temperatures result in similar errors in the $\Delta G_{\text{reaction}}$ which translate into large errors in equilibrium constants. An error of 2 kcal/mole translates into a factor of 2.7 in the equilibrium constant at 1000 K, which introduces a significant error into the species concentrations. It is imperative that one have an accurate method to extrapolate heat capacity data to elevated temperatures. A rigorous test of the quality of an extrapolation is a plot of the residuals in the enthalpy function vs temperature. This represents the integrated error in a heat capacity fit. Figure 9a presents the enthalpy function $H^\circ(T)$ as predicted by using THERM for ethylene oxide, while Figure 9b presents the residuals of the THERM prediction with respect to data from the JANAF table.³² As one can see, the fit is

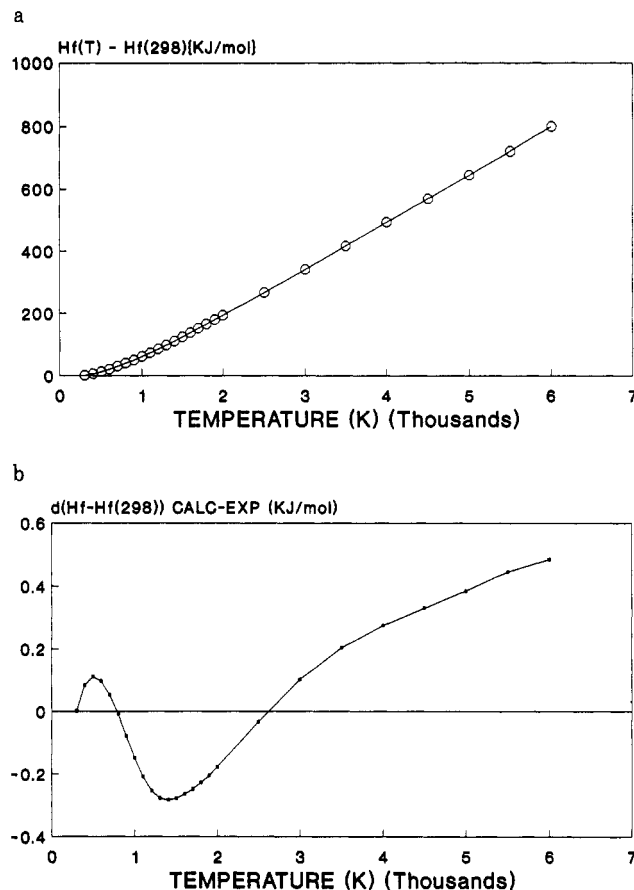


Figure 9. (a) Comparison of extrapolations of the enthalpy function for oxirane using the HOE. Data points (O) are data from the JANAF tables, while solid lines are obtained from THERM. (b) Residuals in the enthalpy function show insignificant deviations from JANAF data across the temperature range from 298 to 6000 K.

excellent. A maximum error is about 0.5 kcal/mol near 6000 K.

SUMMARY

A computer code has been developed that can be used to enter, estimate, and edit thermodynamic properties for ideal gas phase molecules and radicals. Some important aspects of this program are the heat capacity extrapolations to 5000 K, polynomial representations for thermodynamic properties, and calculation of property changes for chemical reaction. Predictions/estimates typically compare favorably with literature data.

Interested parties are encouraged to contact the author for details on receiving a copy of this program in addition to user documentation. The program is currently available in an executable format only, for use on a personal computer running under the MSDOS operating system.

ACKNOWLEDGMENT

I acknowledge Dr. Joseph W. Bozzelli of the Chemical Engineering Department at NJIT for his help in expanding the group database and his many suggestions on the development of this computer code.

REFERENCES AND NOTES

- (1) Benson, S. *Thermochemical Kinetics*; John Wiley and Sons: New York, 1976.
- (2) Freedman, E.; Seaton, W. H. *USA Ballistic Research Laboratories Memorandum Report No. 2320*; Aberdeen Proving Ground, Maryland, 1973.
- (3) Reid, R. C.; Prausnitz, J. M.; Sherwood, T. K. *Properties of Gases and Liquids*; McGraw Hill Book Co.: New York, 1977.

- (4) Domalski, E. S.; Hearing, E. D. *J. Phys. Chem. Ref. Data* **1988**, *17*, 4.
- (5) Ritter, E. R.; Bozzelli, J. W. Submitted to *Int. J. Chem. Kinet.*, Oct 1990.
- (6) Muller, C.; Scacchi, G.; Come, G. M. Presented at the AIChE 77th Annual Meeting, San Francisco, 1984.
- (7) Frurip, D. J.; Freedman, E.; Hertel, G. R. *Proc. Int. Symp. Runaway React.* **1989**.
- (8) Hoffmann, J. M.; Maser, D. C., Eds. *Chemical Hazard Process Review*; ACS Symposium Series 274; American Chemical Society: Washington, D.C., 1985.
- (9) Kee, R. J.; Miller, J. A.; Jefferson, T. H. Sanida Report No. SAND80-8003; 1980.
- (10) Lund, C. M. Lawrence Livermore National Laboratory Report No. UCRL-52504; 1978.
- (11) Reynolds, W. C. *STANJAN Version 3*; Stanford University: Stanford, 1986.
- (12) Cohen, N. Aerospace Report No. ATR-88 (7073)-2; Aerospace Corp.: El Segundo, CA, 1988.
- (13) Stein, S. E.; Golden, D. M.; Benson, S. W. *J. Phys. Chem.* **1977**, *81*, 4.
- (14) Shaw, R.; Golden, D. M.; Benson, S. W. *J. Phys. Chem.* **1977**, *81*, 18.
- (15) Stein, S. E.; Golden, D. M.; Benson, S. W. *J. Phys. Chem.* **1977**, *81*, 4.
- (16) Eigenmann, H. K.; Golden, D. M.; Benson, S. W. *J. Phys. Chem.* **1973**, *77*, 13.
- (17) Stein, S. E.; Fahr, A. *J. Phys. Chem.* **1985**, *89*, 3714.
- (18) Ritter, E. R.; Bozzelli, J. W. Manuscript in preparation, 1991.
- (19) This work.
- (20) Benson, S. W.; Buss, H. H. *J. Chem. Phys.* **1958**, *29*, 546.
- (21) Benson, S. W.; Cruickshank, F. R.; Golden, D. M.; Haugen, G. R.; O'Neal, H. E.; Rogers, A. S.; Shaw, R.; Walsh, R. *Chem. Rev.* **1969**, *69*, 279.
- (22) Burcat, A.; Zeleznik, F. J.; McBride, B. J. NASA Technical Memorandum No. 83800; 1985.
- (23) McBride, B. J.; Gordon, S. NASA Technical Memorandum, NASA TN-D-4097; 1967.
- (24) Hanson, R. J.; Haskell, K. H. Sandia National Laboratories Report, SAND77-0552; 1978.
- (25) Hanson, R. J.; Haskell, K. H. Sandia National Laboratories Report, SAND78-1290; 1979.
- (26) Dorafeeva, O. V.; Gurvich, L. V.; Jorish, V. S. *J. Phys. Chem. Ref. Data* **1986**, *15*, 2.
- (27) Ritter, E. R. Ph.D. Thesis, New Jersey Institute of Technology, 1989.
- (28) Ritter, E. R.; Bozzelli, J. W. *Chem. Phys. Proc. Combust.*; Paper 18, Eastern States Combustion Meeting, The Combustion Institute: Albany, 1989.
- (29) Rice, O. K. *Statistical Mechanics Thermodynamics and Kinetics*; W. H. Freeman and Company: San Francisco, 1967.
- (30) Soontag, R. E.; Van Wylen, G. J. *Fundamentals of Statistical Thermodynamics*; Robert E. Krieger Publishing Co.: Malabar, FL, 1985.
- (31) Herzberg, G. *Molecular Spectra and Molecular Structure II. Infrared and Raman Spectra of Polyatomic Molecules*; Van Nostrand Reinhold Co.: New York, 1945.
- (32) JANAF Thermochemical Tables. *J. Phys. Chem. Ref. Data* **1985**, *14*, 1.
- (33) Aly, F. A.; Lee, L. L. *Fluid Phase Equilib.* **1981**, *6*, 169.
- (34) Fakeeha, A.; Kache, A.; Rehman, Z. U.; Shoup, Y.; Lee, L. L. *Fluid Phase Equilib.* **1983**, *11*, 225.
- (35) Rehman, Z. U.; Lee, L. L. *Fluid Phase Equilib.* **1985**, *22*, 21.
- (36) Gardiner, W. C.; Burcat, In *Combustion Chemistry*, Gardiner, W. C., Jr., Ed.; Springer-Verlag: New York, 1984.
- (37) Reklaitis, G. V.; Ravindran, A.; Ragsdell, K. M. *Engineering Optimization*; John Wiley and Sons: New York, 1983.
- (38) Stull, D. R.; Westrum, E. F., Jr.; Sinke, G. C. *The Chemical Thermodynamics of Organic Compounds*; Robert E. Krieger Publishing Co.: Malabar, FL, 1987.
- (39) Pedley, J. B.; Naylor, R. D.; Kirby, S. P. *Thermochemical Data of Organic Compounds*; Chapman and Hall: New York, 1986.

An Integrated Approach to Three-Dimensional Information Management with MACCS-3D¹

OSMAN F. GÜNER,* DAVID W. HUGHES, and LISE M. DUMONT

Molecular Design Limited, 2132 Farallon Drive, San Leandro, California 94577

Received January 18, 1991

In the past decade, the scientific community has realized the value of three-dimensional (3D) structural information and '3D searching' has started to become an important new methodology for computer-aided drug design. During this time, molecular modeling information generated from various sources has proliferated due to the growing availability of software and hardware and the increasing use of crystallographic and spectroscopic techniques. This information needs to be organized to allow for its effective storage and retrieval. This paper presents an approach to address this problem with a recently introduced program, MACCS-3D. In particular, this approach utilizes MACCS-3D's capability of handling data specific for atoms and atom pairs. With this software, various biological, computational, and spectroscopic data can be merged, allowing scientists from different disciplines to access and use this information more efficiently.

INTRODUCTION

In the past decade, computational chemistry, molecular modeling, and spectroscopic and crystallographic methods of 3D structural elucidation have become standard techniques in new chemical and drug design and have found several applications in a variety of chemical research areas. New developments in technology also affect the methods of information transfer; activities that were once considered impossible are commonplace today. The ability to interface instruments to data-collecting devices allows for the easy generation of vast amounts of data. In addition, communication and interfacing technology allows research sites to bring together different types of data from various experiments, permitting the scientist to merge information from multiple sources.²

Baker analyzes scientific information barriers across international borders: "users demand more, new, timely, high-quality, and complete information services; developing technology induces increased efficiency and effectiveness in in-

formation flow; and economics force sharing of resources and bartering and exchange of information wherever possible".³ This analysis is also applicable to the information barriers across disciplines, even within the same organization.

For example, Brown explains that the development of a new human drug may involve the collaborative efforts of representatives from 30 to 50 distinct scientific disciplines. He then emphasizes the need for a multidisciplinary requirement for collecting, indexing, storing, retrieving, evaluating, and disseminating information over a decade of time between project definition and market introduction.⁴ There are, however, obvious technological limitations in integrating such information. For example, the programs that computational chemists use to determine and later store the 3D structures and data are highly specialized and unwieldy to most other researchers in an organization. And programs historically used for managing chemical information corporate-wide do not have the capability to handle data specific to certain parts of