В.И.Барановский

Квантовохимические расчеты повышенной точности

2015

Учебное пособие «Квантовохимические расчеты повышенной точности» было рассмотрено на заседании кафедры квантовой химии СПбГУ 31 марта 2015 г. и рекомендовано к публикации в электронной форме.

Введение

Данное учебное пособие предназначено, прежде всего, для аспирантов и магистрантов, осваивающих курсы «Квантово-химические методы в термохимии» и «Теория химических реакций». Оно является продолжением учебника «Лекции по квантовой химии» для студентов кафедры квантовой химии и для всех, интересующихся современными методами квантовохимических расчетов. С момента выхода этого учебного пособия появился целый ряд процедур для проведения высокоточных квантовохимическиъ расчетов, описанию которых посвящено данное руководство. Можно ожидать, что с развитием вычислительной техники описанные здесь методы станут рутинной процедурой, с которой весьма полезно познакомиться.

Общими для всех разделов руководства являются следующие учебники и монографии.

Барановский В.И. Квантовая механика и квантовая химия, 2008

Барановский В.И. Лекции по квантовой химии, 2009

J.Cioslovski (Ed.) Quantum-Mechanical Prediction of Thermochemical Data, 2001 (в тексте будут ссылки как на «Циословский»). Книгу можно скачать с сайта по адресу www.bookfi/org/book/54206

J.W.Ochterski, Thermochemistry in *Gaussian*, 2000, www.gaussian.com/g_whitepap/thermo.htm

При работе с данным руководством желательно иметь в своем распоряжении сборник статей под редакцией Циословского, ссылка на который приведена выше. Необходима также активная работа с приведенными в Приложении статьями, а также с работами, рекомендуемыми преподавателем.

Прежде всего, необходимо условиться, что мы понимаем под точностью расчета. Под химической точностью обычно понимают точность в расчете энергии порядка одной килокалории/моль (около 4 кДж/моль). Практика показывает, что эта точность достигается при использовании метода связанных кластеров уровня CCSD(T) и достаточно широкого базиса. Не исключено введение некоторых эмпирических поправок. Что же касается методов функционала плотности (DFT), то они применимы для достаточно точных расчетов геометрической структуры молекул (особенно их основных состояний). При расчетах термохимических величин погрешность составляет несколько килокалорий/моль.

Расчеты высокого уровня точности должны дать энергии с точностью порядка 0.1 ккал/моль (то есть лучше 0.5 кДж/моль). Эта точность достигается при учете большого числа факторов, которые будут рассмотрены в последующих главах. В настоящее время соответствующие методики применимы лишь к очень малым молекулам, содержащим не более трех-пяти «тяжелых» атомов (имеются в виду атомы элементов от бора до фтора, возможно и атомы элементов от алюминия до хлора).

Важное обстоятельство заключается в том, что необходимо соблюсти определенный баланс между базисным набором и методом расчета. Сказанное относится также и к расчетам кинетических параметров (констант равновесия, коэффициентов скоростей реакций), хотя в этих случаях требования к базису/методу расчета более жесткие. Так, расчеты, выполненные с «химической точностью» (не считая ряда других факторов), позволяют оценить константы скорости реакций с точностью до фактора 2.

Как известно, расчеты в приближении Хартри-Фока (эквивалентный термин – метод самосогласованного поля, ССП) не позволяет получить энергии и многие другие характеристики молекул с приемлимой точностью. Необходимость учета эффектов электронной корреляции была признана давно, фактически сейчас все расчеты проводятся с учетом этих эффектов. Следует отметить, что степень учета электронной корреляции зависит (через ссылочную или референтную функцию) от качества базисного набора. Однако, проведение расчетов с очень большими базисными наборами при использовании методов, учитывающих достаточно высокую долю энергии корреляции, в какой-то момент приводит к конфликту с возможностями вычислительной техники. Поэтому развитие методов расчета высокой точности в последнее время сочетается с разработкой техники экстраполяции на полный базисный набор (complete basis set, CBS).

Раньше проводилось различие между *статической* и *динамической* корреляцией. Сейчас понятие статической корреляции практически вышло из употребления, поскольку имелись в виду случаи, когда имело место вырождение или почти вырождение низших состояний, вследствие чего правильная волновая функция не может быть записана в виде одного детерминанта.

Мы начнем с краткого изложения (или напоминания) современных методов учета корреляционных эффектов, которые содержатся практически во всех программных комплексах, доступных нашим пользователям. Эти методы можно разделить на две большие группы: методы, основанные на использовании орбитального приближения, и метод функционала плотности (фактически это тоже группа методов, оперирующих различными корреляционными и обменными функционалами). Мы ограничимся рассмотрением только методов, относящихся к первой группе. Однако, прежде всего будут кратко рассмотрены базисные наборы, используемые в высокоточных расчетах.

Современные базисные наборы

Наиболее популярной базой данных базисных наборов является EMSL, Basis Set Exchange (http://bse.pnl.gov/bse/panel). Приведены практически все известные к настоящему моменту базисы с соответствующими литературными ссылками. При этом базисы представлены в форматах, используемых в известных программных комплексах, так что базисы могут быть без всяких затруднений вставлены во входной файл (типа .inp, .gjf, .com).

Современные базисы характерны тем, что они ориентированы на расчеты с учетом корреляционных эффектов. Первой серией такого рода гауссовых базисов был набор Даннинга cc-pvXz (N=2,3,4 или, иначе, D,T,Q) где X – число базисных функций валентной *sp*-группы. Присутствие поляризующих функций в базисе обязательно. Функции *sp*-группы берутся из расчетов методом ССП, а экспоненциальные параметры поляризующих функций оптимизируются в расчетах методом CISD. Для лучшей передачи корреляционных эффектов существенно включение в базис диффузных *sp*-функций. Такие базисы называются корреляционно согласованными (*correlation consistent,* отсюда префикс *cc*-). Общая структура таких базисных наборов видна из Табл. 1, в которой показано число примитивных и сжатых функций для каждого базиса.

Таблица 1.

	sp-группа		
Базис	Примитивные	Сжатые	Поляризующие
cc-pvDz	(9s4p)	[3s2p]	(1d)
cc-pvTz	(10s5p)	[4s3p]	(2d1f)
cc-pvQz	(12s6p)	[5s4p]	(3d2f1g)

Как видно, по мере увеличения параметра X растет и число поляризующих базисных функций, причем последовательно подключаются функции с большим орбитальным числом. В то время, когда выполнялась разработка *cc*-базисов, не существовало программ расчета молекулярных интегралов для l=5 (h-орбиталей), поэтому базисы *cc*-pv5z (с h-функциями) и *cc*-pv6z (с I-функциями) в указанной статье Даннинга отсутствуют. Однако, к настоящему моменту такие базисы существуют и включены в соответствующие базы данных. Кроме того, корреляционно согласованные базисы разработаны практически для всех остальных элементов.

Высокая точность расчетов предполагает учет корреляций внутри остовной оболочки и между электронами остова и валентными электронами. Соответствующие базисы (обозначаемые как cc-pcVXz) получают путем введения дополнительных базисных функций в cc-pVXz базисы. Структура этих базисных наборов представлена в Таблице 2.

Таблица 2. Корреляционно согласованные базисы для элементов первого периода (N - число функций в базисном наборе)

X	cc-pVXz	N	cc-pcVXz	N
D	3s2p1d	14	4s3p1d	18
T	4s3p2d1f	30	6s5p3d1f	43
Q	5s4p3d2f1g	55	8s7p5d3f 1g	84
5	6s5p4d3f2g1h	91	10s9p7d5f3g1h	145
6	7s6p5d4f3g2h1i	140	12s11p9d7f5g3h1i	230

Необходимость точных расчетов для анионных структур стимулировало разработку базисов, за основу которых приняты базисы *cc*-pvNz, дополненные (augmented) прежде всего диффузными функциями. Кроме того, еще дополнительно вводятся диффузные функции к соответствующие им поляризующим. Такие базисы символически обозначаются как aug-cc-pvXz. Эти базисы широко используются и при расчетах нейтральных молекул.

Появление методов расчета с явным учетом корреляционных эффектов (explicitly correlated) путем введения в расчет в явном виде межэлектронного расстояния R12 (или производной от него функции F12) привело к появлению соответствующих базисов (как правило, за счет расширения базисных функций cc-pvNz). Приведем в качестве примеров несколько базисных наборов различных типов.

Таблица 3. Базисные наборы cc-pvDz, cc-pvTz, aug-cc-pvDz и aug-cc-pvTz

cc-pvDz	cc-pvTz
C 0	C 0
S 8 1.00	S 8 1.00
6665.0000000 0.0006920	8236.0000000 0.0005310
1000.0000000 0.0053290	1235.0000000 0.0041080
228.0000000 0.0270770	280.8000000 0.0210870
64.7100000 0.1017180	79.2700000 0.0818530
21.0600000 0.2747400	25.5900000 0.2348170
7.4950000 0.4485640	8.9970000 0.4344010
2.7970000 0.2850740	3.3190000 0.3461290
0.5215000 0.0152040	0.3643000 -0.0089830
S 8 1.00	S 8 1.00
6665.0000000 -0.0001460	8236.0000000 -0.0001130
1000.0000000 -0.0011540	1235.0000000 -0.0008780
228.0000000 -0.0057250	280.8000000 -0.0045400
64.7100000 -0.0233120	79.2700000 -0.0181330
21.0600000 -0.0639550	25.5900000 -0.0557600
7.4950000 -0.1499810	8.9970000 -0.1268950
2.7970000 -0.1272620	3.3190000 -0.1703520
0.5215000 0.5445290	0.3643000 0.5986840
S 1 1.00	S 1 1.00
0.1596000 1.0000000	0.9059000 1.0000000
P 3 1.00	S 1 1.00
9.4390000 0.0381090	0.1285000 1.0000000
2.0020000 0.2094800	P 3 1.00
0.5456000 0.5085570	18.7100000 0.0140310
P 1 1.00	4.1330000 0.0868660
0.1517000 1.0000000	1.2000000 0.2902160
D 1 1.00	P 1 1.00
0.5500000 1.0000000	0.3827000 1.0000000
	P 1 1.00
	0.1209000 1.0000000
	D 1 1.00
	1.0970000 1.0000000
	D 1 1.00
	0.3180000 1.0000000
	F 1 1.00
	0.7610000 1.0000000

aug-cc-pvDz	aug-cc-pvTz
C 0 S 8 1.00 6665.0000000 0.0006920 1000.0000000 0.0053290 228.0000000 0.0270770 64.7100000 0.1017180 21.0600000 0.2747400 7.4950000 0.4485640 2.7970000 0.2850740 0.5215000 0.0152040 S 8 1.00 6665.0000000 -0.0001460 1000.0000000 -0.0011540 228.0000000 -0.0057250 64.7100000 -0.0233120 21.0600000 -0.0639550 7.4950000 -0.1499810 2.7970000 -0.1272620 0.5215000 0.5445290 S 1 1.00 0.1596000 1.0000000 S 1 1.00 0.0469000 1.0000000 P 3 1.00 9.4390000 0.0381090 2.0020000 0.2094800 0.5456000 0.5085570 P 1 1.00 0.1517000 1.0000000 P 1 1.00 0.0404100 1.0000000 D 1 1.00 0.5500000 1.0000000 D 1 1.00 0.1510000 1.0000000	C 0 S 8 1.00 8236.0000000 0.0005310 1235.0000000 0.0041080 280.8000000 0.0210870 79.2700000 0.0818530 25.5900000 0.2348170 8.9970000 0.4344010 3.3190000 0.3461290 0.3643000 -0.0089830 S 8 1.00 8236.0000000 -0.0001130 1235.0000000 -0.0008780 280.8000000 -0.0045400 79.2700000 -0.0181330 25.5900000 -0.0557600 8.9970000 -0.1268950 3.3190000 -0.1703520 0.3643000 0.5986840 S 1 1.00 0.9059000 1.0000000 S 1 1.00 0.1285000 1.0000000 S 1 1.00 0.0440200 1.0000000 P 3 1.00 18.7100000 0.0140310 4.1330000 0.0868660 1.2000000 0.2902160 P 1 1.00 0.3827000 1.0000000 P 1 1.00 0.1209000 1.0000000 P 1 1.00 0.0356900 1.0000000 D 1 1.00 1.0970000 1.0000000 D 1 1.00 0.3180000 1.0000000 D 1 1.00 0.1000000 1.0000000 F 1 1.00 0.7610000 1.0000000 F 1 1.00 0.2680000 1.0000000

Практически все современные процедуры, направленные на получение высокоточных значений энергии, содержат в качестве своей составляющей экстраполяцию на полный базис. Детальное описание этой процедуры содержится в Гл. 1 книги Циословского. Кроме того, в данном руководстве эти методы описаны в разделах, посвященным конкретным расчетным процедурам.

Методы расчета энергии электронной корреляции

Под энергией корреляции понимают разность между точной энергией молекулы и энергией Хартри-Фока, рассчитанной в полном базисе. Очевидно, что такое определение недостаточно практично. Фактически энергию корреляции определяют для конкретного базиса и конкретного метода:

$$E_{\text{корр.}}(\text{базис/метод}) = E(\text{базис/метод}) - E(\text{базис/ХФ})$$

Поскольку предполагается, что метод полного конфигурационного взаимодействия (см. ниже) дает точное значение энергии корреляции для выбранного базиса, то можно ставить вопрос о том, какой процент от корреляционной энергии для данного базиса дает тот или иной метод расчета.

Метод конфигурационного взаимодействия

Этот метод не используется в расчетах повышенной точности. Это обусловлено медленной сходимостью ряда КВ и нарушением требования однородности (size-extensive) при обращении к системам разного размера. Тем не менее, описание метода КВ полезно для лучшего понимания метода связанных кластеров.

Будем исходить из того, что имеется набор одноэлектронных спин-орбиталей $\{\psi_i\}$. Волновая функция N-электронной системы записывается в виде линейной комбинации слэтеровских детерминантов

$$\Psi_Q(1,2,...N) = \sum_K C_{KQ} \Phi_K(1,2,...N) \qquad (17.1)$$

где

$$\Phi_K(1,2,...N) = \det|\psi_{1K}\ \psi_{2K}\ \psi_{3K}\ ...\psi_{NK}|. \quad (17.2)$$

Здесь индексы при спин-орбиталях в слэтеровских детерминантах – это их номера в множестве индексов спин-орбиталей $\{\psi_i\}$, а индекс “K” указывает на то, что данная спин-орбиталь входит в детерминантную функцию $\Phi_K(1,2,...N)$. Эти детерминантные функции обычно называют *конфигурациями*, поскольку каждая из них соответствует определенному размещению электронов по спин-орбиталям (или орбиталям). Набор функций $\Phi_K(1,2,...N)$ образует *конфигурационный базис*. Функция (17.1) будет точной для полного конфигурационного базиса (которого можно отождествить с бесконечной длиной разложения, хотя очевидно, что на практике возможны лишь разложения конечной длины). Это приводит к проблеме отбора конфигураций, включенных в расчет.

В функции (17.1) вариационными параметрами являются коэффициенты разложения C_{KQ}. Но вообще говоря нет гарантии, что спин-орбитали, использованные для построения конфигураций, будут оптимальными для данной многоконфигурационной функции. Действительно, обычно в многоконфигурационных расчетах используют орбитали, полученные в результате решения уравнений Хартри-Фока с однодетерминантной волновой функцией. Эти функции «порождены» полем, которое они сами (или, точнее, соответствующее им распределение электронной плотности) и создают. При переходе к многоконфигурационной функции (17.1) это поле изменится, что должно привести и к изменению орбиталей $\{\psi_i\}$ – они больше не будут удовлетворять условиям самосогласованности. Методы, основанные на одновременной оптимизации (варьировании) коэффициентов разложения многоэлектронной функции по конфигурациям и коэффициентов разложения молекулярных орбиталей по базисным функциям (атомным орбиталям) называют *многоконфигурационными методами самосогласованного поля* (МКССП). Метод, основанный на варьировании только коэффициентов при конфигурациях, называют методом *конфигурационного взаимодействия* (КВ или CI). Варианты метода КВ определяются принципами построения *конфигурационного базиса.*

Вернемся к проблеме отбора и классификации конфигураций. Прежде всего выбирают конфигурацию, которую называют *ссылочной* (или референтной, *reference*). В качестве ссылочной конфигурации удобно принять однодетерминантную функцию метода Хартри-Фока

$$\Phi_0 = \frac{1}{\sqrt{N!}} | \psi_1 \psi_2 ... \psi_N |.$$

Пусть одноэлектронные функции $\{\psi_i\}$ являются решениями уравнения Хартри-Фока. Будем обозначать орбитали, входящие в Φ_0 (занятые орбитали), индексами i, j,… (i,j=1,2,…N) а все остальные (то есть виртуальные) орбитали - индексами a, b,…(a,b=N+1,…2M, где M – размер базиса в расчете ССП). Построение детерминантов, входящих в конфигурационный базис, можно осуществить путем замены одной или нескольких спин-орбиталей ψ_i, ψ_j,…в Φ_0 на такое же число виртуальных спин-орбиталей ψ_a, ψ_b, … .Такие конфигурации рассматривают как результат возбуждения электронов с орбиталей ψ_i, ψ_j,… на орбитали ψ_a, ψ_b, …, а число замен занятых орбиталей на виртуальные – *кратностью* возбуждения. Конфигурации с однократными возбуждениями обозначают Φ_i^a, с двукратными возбуждениями - Φ_{ij}^{ab}, с тройными - Φ_{ijk}^{abc} и так далее.

Варианты метода КВ, как правило, классифицируют по кратностям возбужденных конфигураций, включенных в базис. Так, если в конфигурационном базисе содержатся только двукратные возбуждения, то говорят о CID (Configuration Interaction, Doubles), если и однократные и двукратные возбуждения – CISD (Configuration Interaction, Singles and Doubles), сокращение CISDTQ означает, что учтены все возбуждения вплоть до четырехкратных. Так, в варианте CISD пробная функция имеет вид

$$\Psi_{CISD} = C_0\Phi_0 + \sum_{i}^{\text{зан.}} \sum_{a}^{\text{вирт.}} C_i^a \Phi_i^a + \sum_{i}^{\text{зан.}} \sum_{<j}^{\text{зан.}} \sum_{a}^{\text{вирт.}} \sum_{<b}^{\text{вирт.}} C_{ij}^{ab} \Phi_{ij}^{ab} \,.$$

Особое место занимает так называемое *полное КВ* (*full CI, FCI*) в котором рассматриваются все возможные возбуждения. Иначе говоря, строится набор детерминантов, которые соответствуют всем возможным размещениям N электронов по 2M спин-орбиталям. Для выбранного орбитального базисного набора метод FCI дает самую точную волновую функцию и самую низкую энергию, которую можно рассматривать как *точную энергию в данном базисе*. Это позволяет проводить тестирование приближенных методов учета электронной корреляции. При расширении базиса волновая функция и энергия приближаются к точному решению нерелятивистского уравнения Шредингера. Разность энергий метода FCI и Хартри-Фока дает полную корреляционную энергию для данного орбитального базиса

$$E_{corr.} = E(FCI) - E(HF)\,.$$

На практике прибегают с сокращению конфигурационного базиса. Первый (описанный выше) способ такого сокращения – ограничение кратности возбуждений (например, двукратными (CISD) или четырехкратными (CISDTQ)).

Другой путь сокращения размера конфигурационного базиса заключается в ограничении *активного пространства*. Все орбитали делятся на активные и неактивные. Неактивные орбитали входят или во все детерминанты конфигурационного базиса (их заселенность при этом равна двум – как правило, это орбитали остова), или ни в один из них (заселенность таких виртуальных орбиталей равна нулю). Остальные орбитали и образуют *активное пространство* расчета КВ. Оно включает в себя те занятые орбитали, с которых будут производиться возбуждения, и те виртуальные орбитали, на которые будут «приходить» возбуждаемые электроны. Наконец, можно использовать в качестве базисных не отдельные детерминанты, а *симметризованные* (прежде всего по спину) функции – конфигурационные функции состояний, CSF. Сочетая эти приемы, можно получить существенное сокращение числа многоэлектронных функций базиса КВ.

В расчетах методом КВ волновая функция записывается в виде линейной комбинации детерминантов, построенных из спин-орбиталей. Перед построением матрицы оператора Гамильтона производится преобразование двуэлектронных интегралов от атомного к молекулярно-орбитальному базису. Это преобразование обычно занимает большую часть времени расчета.

Многоконфигурационные методы самосогласованного поля (МКССП, MCSCF).

Методы МКССП предполагают одновременную оптимизацию как молекулярных орбиталей, так и коэффициентов разложения по конфигурациям. На заключительном этапе расчета, как правило, производится расчет методом КВ или по теории возмущений с новыми, оптимизированными МО.

Метод МКССП не используется в расчетах повышенной точности. Однако, в принципе он может служить исходной точкой для более точных расчетов.

Метод связанных кластеров

Метод *связанных кластеров* (*CC*) фактически представляет собой вариант метода КВ, который характеризуется специальной методикой построения конфигурационного базиса. В методе КВ возбужденные конфигурации получают действием на ссылочную функцию операторов возбуждения $\hat{C}_i$, где i – кратность возбуждения:

$$\hat{C}_1 = \sum_{i,a} c_i^a \hat{t}_i^a, \hat{C}_2 = \sum_{i>j}\sum_{a>b} c_{ij}^{ab} \hat{t}_{ij}^{ab}, \ldots \quad (17.6)$$

Результат действия этих операторов на ссылочную функцию Φ_0 можно записать следующим образом:

$$(\hat{C}_1 + \hat{C}_2 + \ldots)\Phi_0 = \sum_{i,a} c_i^a \Phi_i^a + \sum_{i>j}\sum_{a>b} c_{ij}^{ab}\Phi_{ij}^{ab} + \ldots . \quad (17.7)$$

Напомним, что коэффициенты при конфигурациях являются параметрами, определяемыми вариационным методом.

В методе связанных кластеров оператор возбуждения записывают в экспоненциальной форме

$$|\Psi_{CC}\rangle = e^{\hat{T}} |\Phi_0\rangle = (1 + \hat{T} + \frac{1}{2!}\hat{T}^2 + \frac{1}{3!}\hat{T}^3 + \ldots)|\Phi_0\rangle,$$

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \ldots,$$

$$\hat{T}_1 = \sum_{i,a} t_i^a \hat{t}_i^a, \quad \hat{T}_2 = \sum_{i>j}\sum_{a>b} t_{ij}^{ab} \hat{t}_{ij}^{ab}, \ldots \quad .$$

Операторы возбуждения $\hat{T}_i$ действуют на детерминантные функции точно так же, как операторы $\hat{C}_i$. Обычно параметры $t_i^a, \hat{t}_{ij}^{ab} \ldots$называют амплитудами. Как и в методе КВ, классификация проводится по кратностям возбуждений, включаемых в оператор $\hat{T}$. Так, если $\hat{T} = \hat{T}_1 + \hat{T}_2$то говорят о методе связанных кластеров с учетом однократных и двукратных возбуждений (CCSD). Если включены только двукратные возбуждения, то такой вариант обозначается как CCD и так далее. Однако, в отличие от метода КВ, вследствие экспоненциальной записи оператора $\hat{T}$ реально функция метода CC содержит конфигурации, соответствующие возбуждениям более высокой кратности, чем включенные в оператор $\hat{T}$. Так, для $\hat{T} = \hat{T}_1 + \hat{T}_2$

$$
\begin{aligned}
&|\Psi_{CC}\rangle = e^{\hat{T}}\,|\Phi_0\rangle = (1+\hat{T}_1+\hat{T}_2+\frac{1}{2}(\hat{T}_1+\hat{T}_2)^2+\ldots)\,|\Phi_0\rangle = \\
&= (1+\hat{T}_1+\hat{T}_2+\frac{1}{2}(\hat{T}_1\hat{T}_1+2\hat{T}_1\hat{T}_2+\hat{T}_2\hat{T}_2)+\ldots)\,|\Phi_0\rangle = \\
&=|\Phi_0\rangle+\sum_{i,a}t_i^a\,|\Phi_i^a\rangle+\sum_{i>j}\sum_{a>j}t_{ij}^{ab}\,|\Phi_{ij}^{ab}\rangle+\frac{1}{2}\sum_{i,a}\sum_{j,b}t_i^a t_j^b\,|\Phi_{ij}^{ab}\rangle+ \\
&+\frac{1}{2}\sum_{i,a}\sum_{j>k,b>c}t_i^a t_{jk}^{bc}\,|\Phi_{ijk}^{abc}\rangle+\ldots.
\end{aligned}
\quad . (17.8)
$$

Как видно, помимо одно- и двукратно возбужденных конфигураций, в функцию метода CCSD входят и конфигурации, получаемые путем трехкратных, четырехкратных и более высоких возбуждений. Очевидно, максимальная кратность возбуждений ограничена числом электронов в системе.

В отличие от метода КВ (за исключением полного КВ) метод связанных кластеров дает однородное (size-extensive) описание сложных систем, что обеспечивает аддитивность энергий подсистем при их разделении.

Метод связанных кластеров не является вариационным. Нахождение энергий и амплитуд достигается путем решения системы уравнений, в которые в качестве коэффициентов при неизвестных $(t_i^a, t_{ij}^{ab}, \ldots)$ входят разности энергий МО и комбинации трансформированных двухэлектронных интегралов. Уравнения решают итерационно до достижения заданной точности. Достаточно подробно, хотя и в несколько ином контексте, эта методика описана в статье J.Pople, M.Head-Gordon, K.Raghavachari, J.Chem.Phys., **87** (1987) 5968. На практике часто используют метод связанных кластеров в приближении CCSD, однако, как правило, стандартным является расчет, в котором тройные возбуждения учитывают методом теории возмущений (такой вариант записывают как CCSD(T)). В настоящее время доступны программы, реализующие приближение CCSDTQ.

Вариантом метода связанных кластеров является метод BD («бракнеровских дублетов»). В этом приближении в качестве одноэлектронных спин-орбиталей используются такие, в базисе которых амплитуды t_i^a (соответствующие однократным возбуждениям) равны нулю. Построение этих орбиталей осуществляется итерационным методом путем вращений в орбитальном пространстве. Этот метод используется в процедуре W1BD (см. ниже).

Расчет по методу связанных кластеров, как правило, используется на том или ином этапе процедуры, направленной на получение высокой точности.

Метод многочастичной теории возмущений

Один из способов учета электронной корреляции – это использование теории возмущений. Описанный в данном параграфе метод носит название метода Мёллера-Плессе (MPn, где n – наивысший порядок теории возмущений, использованный в расчете).

Прежде всего необходимо выбрать функцию и энергию нулевого приближения и оператор возмущения. В качестве функции нулевого приближения берут функцию, полученную методом Хартри-Фока, Ψ_{HF}. При указанном выборе волновой функции нулевого приближения естественно принять в качестве оператора возмущения $\hat{W}$ флуктуационный потенциал $F(r)$, который удобно записать в виде

$$\hat{W} = \hat{H} - \sum_{i}^{занят.} \hat{F}_i ,$$

а гамильтониан невозмущенной задачи как

$$\hat{H}_0 = \sum_{i}^{занят.} \hat{F}_i .$$

Все детерминантные функции Ψ_Q, полученные из хартри-фоковской путем замены орбиталей в Ψ_{HF} являются собственными функциями $\hat{H}_0$

$$\hat{H}_0 \Psi_Q = E_Q \Psi_Q .$$

Энергия нулевого приближения равна сумме орбитальных энергий

$$E^{(0)} = \sum_{i}^{занят.} \varepsilon_i .$$

Сумма энергии нулевого приближения и поправки к энергии первого порядка равна энергии, поученной в расчете методом Хартри-Фока. Поправка второго порядка к энергии рассчитывается по обычной формуле теории возмущений

$$E^{(2)} = -\sum_{Q} \frac{|\langle \Psi_{HF} | \hat{W} | \Psi_{Q} \rangle|^2}{E_Q - E_{HF}} \quad .$$

Все матричные элементы, соответствующие однократным возбуждениям, согласно теореме Бриллюена равны нулю. В знаменателе все члены положительны, так как функции Ψ_Q получены путем замены занятых орбиталей в Ψ_{HF} на виртуальные. Поэтому поправка второго порядка к энергии всегда отрицательна, и расчет с использованием метода МР2 всегда приводит к понижению энергии. Метод теории возмущений Мёллера-Плессе – невариационный. На практике при оптимизации геометрии с учетом корреляции используют метод МР2, а уточнение энергии после оптимизации геометрии производят методом МР4.

Разработан также ряд вариантов метода МР2 для расчета с многоконфигурационными функциями нулевого приближения. В реализованных методиках многодетерминантные функции нулевого приближения получают методом CASSCF. Оператор $\hat{H}_0$ строится как сумма некоторых одноэлектронных операторов, после чего путем прямой минимизации полной энергии получают волновые функции первого порядка, с которыми и рассчитывают поправки второго порядка к энергии. Таким образом, можно рассчитать в рамках теории возмущений одновременно энергии нескольких состояний, что очень удобно, в частности, для расчета электронных спектров молекул.

Методы расчета с явным учетом корреляционных эффектов

В последние годы появился ряд методов расчета, основанных на впервые предложенной Хиллераасом в 1929 г. при расчете атома гелия идее введения в волновую функцию межэлектронного расстояния. Этот метод был использован Джеймсом и Кулиджем в середине тридцатых годов при расчетах молекулы H_2, Методы, разработанные для использования волновых функций такого типа, обозначают как MP2-R12, CCSD(T)-R12 и так далее. Несколько более точные результаты получают, если вместо R12 используют функцию $F12 = \exp(-\gamma r_{12})$, где γ – варьируемый параметр, близкий к единице. В этом случае к методу добавляется составная часть -F12. Эти методы реализованы в программе Molpro, имеющейся в ВЦ СПбГУ.

Методы, в которых корреляция учитывается с помощью введения в функцию межэлектронного расстояния называют explicitly correlated методами.

В качестве примера приведем (Таблица 4, Рис. 1) сравнение результатов расчета молекулы H_2 методом связанных кластеров с результатами расчета Колоса и Рутана, полученные с 50-членной волновой функцией для R=0.741324 Å (W.Kolos, C.C.J.Roothaan, Rev. Modern Phys., 1060, **32**, 219). Этот расчет до сих пор является одним из самых точных, проведенных для этой молекулы.

Таблица 4. Сравнение результатов расчетов методами Полного КВ, CCSD(T)-F12 и CCSD(T) (компонента F12 в гамильтониане отсутствует).

Базис	Full CI	CCSD-F12/CCSD	E(CCSD-F12)-E(KR), kJ/mol
cc-pvDz-f12	-1.170671	-1.173463/ -1.170669	2.59
cc-pvTz-f12	-1.172609	-1.174045/ -1.172607	1.06
cc-pvQz-f12	-1.173746	-1.174336/ -1.173744	0.29
cc-pv5z	-1.174163	-1.174396/ -1.174223	0.14
CBS предел		-1.174462	-0.04
aug-cc-pvDz		-1.171183/ -1.164621	8.57
aug-cc-pvTz		-1.173778/ -1.172634	1.76
aug-cc-pvQz		-1.174254/ -1.173867	0.51
aug-cc-pv5z		-1.174409/ -1.174252	0.10
CBS предел		-1.174522	-0.19
Kolos/Roothaan 50-членная функция	-1.174448		0.

Примечание. В расчете FULLCI с cc-pv5z взято всего 100 активных виртуальных орбиталей вместо 139.

Как видно, значения энергии, полученные методом полного КВ, практически совпадают с полученными методом связанных кластеров без включения членов, связанных с F12, в гамильтониане. Наилучшая сходимость получается в расчетах методом CCSD(T)-F12, причем отличие от точного значения порядка десятых кДж/моль достигается уже при X=5. Интересно отметить, что экстраполяция на полный базис дает более разумное значение в случае базисов cc-pvXz-f12, чем для базисов типа aug-cc-pvXz.

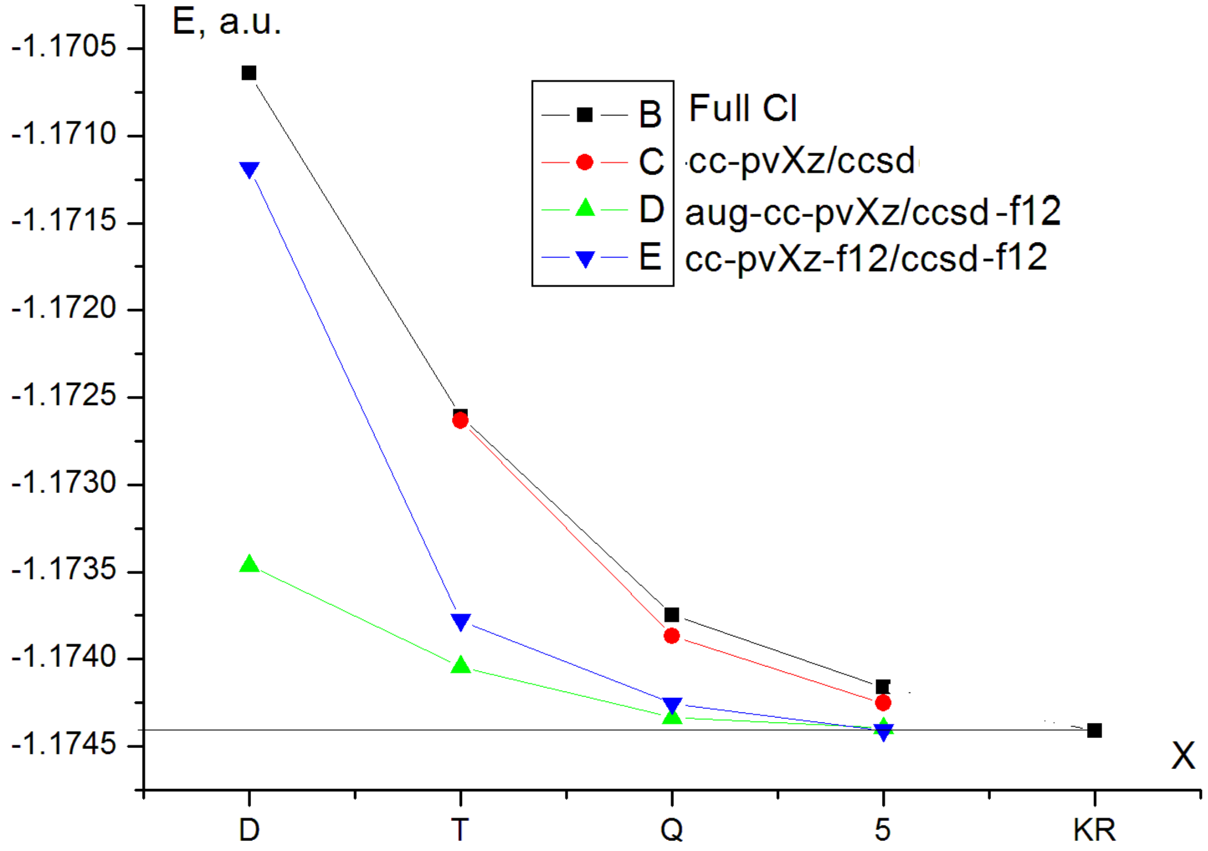


Рис. 1. Сходимость различных вариантов метода связанных кластеров в расчетах молекулы водорода к энергии, полученной Колосом и Рутаном.

Примеры использования explicitly correlated методов в квантовохимическихт расчетах можно найти в статьях:

J.Zhang, E,E.Valeev, Prediction of Reaction Barriers and Thermochemical Properties with Explicitly Correlated Coupled-Cluster Methods: A Basis Set Assesment, J. Chem.Theory Comp., 2012, **8**, 3175

W.Klopper et al., Atomization energies from coupled-cluster calculations augmented with explicitly-correlated perturbation theory, Chem. Phys, 2009, 356, 14

J.M.L.Martin, M.K.Kesharwani, Assesment of CCSD(T)-F12 Approximations and Basis Sets for Harmonic Vibrational Frequencies, , J. Chem.Theory Comp., 20114, **10**, 2085

D.H.Bross et al., Explicitly correlated composite thermochemistry of transition metal species, J. Chem. Phys., 2013, **139**, 09432

Подведем некоторые итоги рассмотрения методов учета корреляционных эффектов.

1. Размер конфигурационного базиса метода КВ чрезвычайно быстро растет с увеличением числа базисных функций. Особенно хорошо это видно в случае полного КВ. Для молекул с числом атомов, большим 4 – 5, этот метод практически не может быть использован в рутинных расчетах, а низшие приближения метода не обеспечивают достаточной степени учета корреляций. Поэтому в настоящее время этот метод не используется в расчетных методах повышенной точности.
2. Наиболее подходящим методом для расчета энергии корреляции является метод связанных кластеров, который дает значения энергии корреляции, весьма близкие к полученным в расчете полного КВ или из эксперимента. При разработке новых высокоточных процедур все чаще используются только метод связанных кластеров (в отличие от более ранних методов, в которых широко использовались методы MP2 и MP4). Однако, метод связанных кластеров требует больших затрат времени и ресурсов памяти, что ограничивает его использование относительно небольшими молекулами.
3. Широко используемые в расчетах базисы умеренных размеров (6-31G*, 6-311G*) дают всего около 50 – 60 % полной энергии корреляции. Поэтому в настоящее время в расчетах используют специально разработанные для расчетов с учетом корреляционных эффектов базисы, рассмотренные выше.
4. К методам, учитывающим эффекты электронной корреляции, относится также метод функционала плотности (DFT). Этот метод позволяет хорошо описать термохимические свойства основного состояния и геометрическую структуру молекул. Поэтому в расчетах с повышенной точностью метод DFT используется только для оптимизации геометрической структуры и получения термохимических параметров, что необходимо для учета температурных эффектов и перехода от электронной энергии к энтальпии. В последние годы для этой цели чаще используют расчет методом связанных кластеров.
5. Методы с явным учетом корреляции представляют собой весьма эффективный инструмент учета корреляционных эффектов и

следует ожидать постепенного введения их в практику квантовохимических расчетов.

Композитные методы расчета

Несмотря на развитие вычислительных методов, позволяющих учесть значительную долю энергии корреляции, постепенно к концу девятьсот восьмидесятых годов стало ясно, что один какой-либо метод, взятый в отдельности, не может обеспечить необходимую для ряда прикладных задач точность расчета. Особое внимание уделялось методам, призванным обеспечить высокую точность расчета теплот образования молекул (или их энергий атомизации). Знание этих величин необходимо при изучении энергетики и скоростей химических реакций, которые с трудом поддаются экспериментальным исследованиям, или которые необходимы для проверки фундаментальных основ теории химических процессов, или важны при изучении экологических проблем.

В основе современных высокоточных методов расчета лежит идея, что полная энергия системы может быть записана в виде суммы нескольких вкладов, каждый из которых соответствует некоторому физически объяснимому эффекту. Таким образом, разрабатываемая процедура должна состоять из ряда последовательных расчетов. Их совокупность обычно рассматривают как м*одельную химическую теорию*, в которой каждый шаг проводится с использованием совершенно определенного метода и базисного набора. Это обеспечивает одинаковый подход к расчету любых изучаемых молекул.

К настоящему времени разработан ряд таких модельных теорий (их называют также композитными методами). Мы остановимся только на некоторых из них, и прежде всего на тех, которые в программном комплексе GAUSSIAN оформлены в виде замкнутых в себе процедур (типа «черного ящика»), запуск которых осуществляется с помощью соответствующего «ключевого слова» (см. разделы “CBS Methods”, “G1-G4 Methods” и “W1 Methods” в списке Gaussian 09 Keywords).

Группы процедур CBS и Gaussian-n (G-n, n=1-4) были разработаны практически одновременно. Мы начнем с рассмотрения методов группы G-n, как наиболее широко используемых на практике.

Создатели программы GAUSSIAN поставили целью разработать процедуру, в которой полная электронная энергия была бы представлена как сумма ряда составляющих, каждая из которых может быть рассчитано при относительно небольших затратах вычислительных мощностей. Соответствующие программные блоки содержат ряд последовательно выполняемых заданий, переходы между которыми заданы в программе и происходят без участия пользователя. Эти процедуры получили наименование Gaussian-n (Gn, n=1,2,3,4). Точность результатов расчетов с течением времени (то есть с увеличением порядкового номера **n)** постоянно возрастало, хотя общая структура программы сохранялась. Поэтому для понимания подхода, заложенного в этих процедурах, достаточно рассмотреть один из существующих вариантов.

Литература

N.J.DeYonker, Th.R.Cundari, A.K.Wilson, The correlation consistent composite approach(ccCA): An alternative to the Gaussian-n methods, J.Chem.Phys. 2006, **124**, 114104

L.A.Curtiss, P.C.Redfern, K.Raghavachary, Gaussian-4 theory, J.Chem.Phys., 2007, **126**, 084108

J.A.Montgomery, Jr., M.J.Frisch, J.W.Ochterski, J.A.Petersson, A complete basis set model chemistry. VI. Use of density functional geometries and frequencies, J.Chem.Phys., 1999, **110**, 2822

E.C.Barnes, G.A.Petersson, J.A.Montgomery, M.J.Frisch, J.M.L.Martin, Unrestricted Coupled Cluster and Brueckner Doubles Variations of W1 Theory, J. Chem. Theory Comput., 2009, **5**, 2687

Метод G4.

Проследим последовательность шагов процедуры G4, будут отмечены также отличия от предыдущих версий теории.

1. Первый шаг заключается в оптимизации геометрической структуры молекулы методом B3LYP/6-31G(2df,p). Для систем с синглетным основным состоянием используется ограниченный, а в остальных случаях – неограниченный методы ССП. Было установлено, что метод B3LYP при расчете геометрии дает лучшие результаты, чем метод MP2, который использовался в ранних вариантах теории. С

использованием полученной геометрии молекулы и того же базиса производится расчет гармонических колебательных частот, которые затем умножаются на масштабирующий множитель, равный 0.9854. С исправленными частотами рассчитывается поправка на нулевые колебания (ZPE).

2. Серия корреляционных расчетов при фиксированной геометрической структуре, начиная с полного расчета MP4 в базисе 6-31G(d). К результату добавляется ряд корректирующих поправок, получаемых в результате серии дополнительных расчетов:

   (a) поправка на вклад диффузных функций

   ΔE(+)=E[MP4/6-31+G(d)] – E[MP4/6-31G(d)],

   (b) поправка на поляризующие функции с более высокими значениями *l* :

   ΔE(2df,p)=E[MP4/6-31G(2df,p] – E[MP4/6-31G(d)],

   (c) поправка на корреляционные эффекты более высоких порядков, которая рассчитывается с использованием метода связанных кластеров:

   ΔE(CC)=E[CCSD(T)/6-31G(d) – E[MP4/6-31G(d)],

   (d) поправка на эффект от расширения базиса и возможной неаддитивности, связанной с предположением о независимости вкладов от диффузных и поляризующих функций:

   ΔE(G3LargeXP)=E[MP2(full)/G3LargeXP] - E[MP2/6-31G(2df,p]- E[MP2/6-31+G(d)] – E[MP2/6-31G(d)].

   Здесь G3LargeXP – специально сконструированный базис, содержащий повышенное число поляризующих функций.

3. Два ССП расчета с модифицированными базисами aug-cc-pvQz и aug-cc-pv5z. Результаты этих расчетов используются для экстраполяции на полный хартри-фоковский базис. Разность между полученной предельной энергией и энергией, рассчитанной с базисом G3LargeXP, обозначается как ΔE(HF).

4. Отдельно определяется вклад от спин-орбитального взаимодействия, ΔE(SO). Этот вклад берется из экспериментальных данных или из имеющихся в литературе точных расчетов.

5. Наконец, рассчитанная электронная энергия записывается как E[MP4/6-31G(d)] и рассмотренных выше поправок:

   E(combined)=E[MP4/6-31G(d)] + ΔE(+) + ΔE(2df,p) + ΔE(CC) + ΔE(G3LargeXP) + ΔE(HF) + ΔE(SO).

Полная $E_e(G4)$ получается путем добавления к E(combined) эмпирической поправки HLC:

$HLC = -An_\beta$ для молекул с замкнутыми оболочками

$= - A'n_\beta - B(n_\alpha - n_\beta)$ для молекул с открытыми оболочками

$= -Cn\beta - D(n_\alpha - n_\beta)$ для атомов.

Здесь n_α, n_β – число валентных электронов со спинами α и β, причем предполагается, что $n_\alpha \geq n_\beta$. Вводится также поправка, учитывающая число электронных пар.

Наконец, полная электронная энергия G4(0 K) при 0 получается путем добавления энергии нулевых колебаний:

$$E_0(G4) = E_e(G4) + E(ZPE).$$

Добавление температурных поправок к $E_0(G4)$ для температуры 298.15 К дает “ G4 Energy”.

Как видно, расчеты высокого уровня с использованием умеренных базисов сочетаются с расчетами более низкого уровня, но с расширенными базисами. Приведем образец вывода результатов расчета молекулы CF.

```
Temperature=          298.150000 Pressure=              1.000000
 E(ZPE)=                0.003019  E(Thermal)=            0.005389
 E(CCSD(T))=         -137.447249  E(Empiric)=           -0.038081
 DE(Plus)=             -0.013390  DE(2DF)=              -0.082032
 E(Delta-G3XP)=        -0.153552  DE(HF)=               -0.016859
 G4(0 K)=            -137.748144  G4 Energy=          -137.745774
 G4 Enthalpy=        -137.744830  G4 Free Energy=     -137.768361
```

Здесь использованы (в слегка измененном виде) обозначения описанных ранее вкладов в полную энергию. Подробное описание распечатки результатов содержится в Инструкции к программе GAUSSIAN09 (раздел G1-G4 Methods).

Аналогичную структуру имеют и методы типа CBS (CBS-4M, CBS-QB3, CBS-APNO). Обе группы методов содержат в качестве составных частей эмпирические поправки, что несомненно является недостатком этих модельных теорий. Кроме того, эти методики позволяют достичь лишь «химической» точности (1-2 ккл.моль).

Методы W1 (W1U, W1BD, W1RO) и W2

Разработчики методов W1 (W1U, W1BD, W1RO) и W2 поставили перед собой следующие задачи:

Разработать процедуру расчета энергий, которая позволяет, как правило, получить «субхимическую» точность (1 кДж/моль)

В худшем случае ошибка не должна превосходить 1 ккал/моль

Процедура должна позволить проводить расчеты на современных настольных компьютерах

Эмпирические поправки должны быть полностью исключены.

Рассмотрим кратко только варианты протокола W1, включенного в программный комплекс GAUSSIAN09.

Оптимизация геометрической структуры молекулы производится методом B3LYP в базисах умеренного размера. После этого выполняется серия расчетов методом CCSD (W1U) или BD (W1BD) в базисах aug-cc-pvXz (X= d, t, q), дополненными d и f-функциями с умеренно большими значениями экспоненциальных параметров. Производится экстраполяция хартри-фоковских и корреляционных энергий на полный базис. Наконец, выполняются два расчета типа CCSD(T) в специальном базисе, содержащем только примитивные гауссианы. Цель расчетов – оценить вклад внутренних оболочек в корреляцию. Наконец, оценивается вклад скалярных релятивистских эффектов и спин-орбитального взаимодействия.

Конечные результаты расчета выдаются в виде таблицы, похожей на таблицу метода G-4. Отметим, что расчет в приближении CCSD(T) позволяет учесть в неявном виде вклады от более высоких возбуждений (Q и так далее).

Несмотря на то, что рассмотренные композитные методы задуманы как работающие по принципу черного ящика, имеется определенная возможность вмешательства пользователя в их работу. Например, на некоторых этапах расчет производится с использованием нестандартных базисных наборов. Однако, возможно получить распечатку этих базисов путем дополнения ключевого слова, запускающего соответствующую процедуру, ключевым словом gfinput;

G4 gfinput.

Имеется возможность расчета не только для конфигурации, соответствующей минимуму на потенциальной поверхности системы, но и для произвольной точки на пути реакции. В этом случае строка инструкций будет выглядеть, например, следующим образом:

```
# W1U opt(modredundant)

 ……….

 B 1 2 = F
```

Это означает, что оптимизация геометрии будет производиться при фиксированном расстоянии между атомами с номерами 1 и 2. Эта возможность может быть использована, например, для оценки констант скорости реакций методом Вариационной теории переходного состояния.

Другие методики высокоточных расчетов

Помимо рассмотренных выше, известен ряд методик высокоточных расчетов энергий молекулярных структур, которые не оформлены в виде замкнутых процедур. Ниже будет кратко рассмотрен один из таких подходов – HEAT: High accuracy extrapolated *ab initio* thermochemistry то есть Высокоточная неэмпирическая экстраполяционная термохимия. Метод предназначен для расчета энергий атомизации и энтальпий образования молекул и энтальпий реакций. Авторы заранее учитывали возможность взаимной компенсации ошибок при расчете реагентов и продуктов реакций. Отмечается также желательность (а в ряде случаев и необходимость) использования экспериментальных данных теплот образования атомов из простых веществ при расчетах теплот образования молекул.

Оптимизация геометрической структуры и расчете энергии нулевых колебаний используют расчет методом CCSD(T) в базисе cc-pVQz. Формула, по которой вычисляется энергия молекулы в рамках протокола HEAT, содержит восемь членов:

$$E_{HEAT} = E_{HF}^{\infty} + \Delta E_{CCSD(T)}^{\infty} + \Delta E_{CCSDT} + \Delta E_{CCSDTQ}$$

$$+\Delta E_{REL} + \Delta E_{ZPE} + \Delta E_{DBOC} + \Delta E_{SO}$$

Первые два слагаемых представляют собой экстраполированные на полный базис энергии ССП и корреляции, рассчитанной в приближении CCSD(T) (но не полную энергию метода CCSD(T)!). Расчет хартри-фоковской энергии производится в базисах aug-c c-pcVXz (x=T(3), Q(4) и 5), экстраполяция проводится по формуле

$$E_{HF}^{X} = E_{HF}^{\infty} + a\exp(-bX).$$

Очевидно, что для определения неизвестных E_{HF}^{∞}, a и b необходимо выполнить три расчета (например, с X=T, Q, 5).

Для экстраполяции энергии корреляции применяют двухпараметрическую формулу

$$E_{\mathrm{CCSD}(T)}^{X} = E_{CCSD(T)}^{\infty} + a/X^3.$$

Тогда $E_{CCSD(T)}^{\infty} = \frac{E_{CCSD(T)}^{X}X^3 - E_{CCSD(T)}^{Y}Y^3}{X^3 - Y^3}$.

Для определения $E_{CCSD(T)}^{\infty}$ достаточно провести два расчета с «соседними» значениями X и Y.

Третий и четвертый члены предназначены для оценки вкладов от полного учета тройных возбуждений и возбуждений более высоких порядков. Учтены также вклады от скалярной релятивистской энергии (ΔE_{REL}), диагональной поправки Борна-Оппенгеймера на конечную массу атомных ядер (ΔE_{DBOC}) и спин-орбитальное взаимодействие (ΔE_{SO}).

Очевидно, что эта описанная процедура применима к очень малым системам, включающим не более трех «тяжелых» атомов (то есть атомов первого периода).

В литературе описаны и другие методики точных расчетов энергий атомизации и энтальпий образования молекул. При этом можно проследить следующие тенденции.

1. Полный отказ от использования методов многочастичной теории возмущений в пользу метода связанных кластеров.
2. Большое внимание к определению энергии нулевых колебаний, прежде всего учет эффектов ангармонизма.
3. Отказ от разбиения энергии корреляции на вклады от валентных и остовных электронов, производится расчет этой величины для всей электронной системы.
4. Включение в схему расчета методы явного учета межэлектронного расстояния (R12 или F12) в гамильтониан и базисный набор.

В качестве примеров можно рекомендовать следующие работы.

W.Klopper, B.Ruscic, D.P.Tew, F.A.Bischoff, S. Wolfsegger, Atomization energies from coupled-cluster theory calculations augmented with explicitly-correlatrd perturbation theory, Chemical Physics, 2009, **356**, 14.

J.Zhang, E.F.Valeev, Prediction of Reaction Barriers and Thermochemical Properties with Explicitly Correlated Coupled-Cluster Methods: A Basis Set., J. Chem. Theory Comput., 2012, **8**, 3175.

B.Nagy, B.Csontos, J.Csontos, P.Szakas, M.Kallay, High-Accuracy Theoretical Thermochemistry of Fluoroethanes, J. Phys. Chem. A, 2014, **118**, 4824

Сравнение с экспериментом

Тестирование высокоточных методов расчетов проводится путем сравнения с экспериментальными данными по энтальпиям образования или атомизации. Обычно рассматриваются также полученные значения энтропии при нормальных условиях. Дополнительно анализируются данные по межатомным расстояниям и валентным углам и колебательным частотам. Но понятно, что результаты расчетов этих параметров являются лишь одной из составляющих значений термохимических величин, поэтому высокая точность их расчета не гарантирует высокого качества расчетов энтальпий образования и атомизации. Поэтому основное внимание должно быть уделено оценке качества расчета электронной энергии системы

Как правило, сравнение результатов расчета и эксперимента проводится для специально подобранных групп из нескольких десятков молекул, для которых имеются достаточно надежные данные по термохимическим характеристикам. Качество метода оценивается путем сравнения средних (и прежде всего, среднеквадратичных) отклонений расчетных величин от экспериментальных.

Источником сведений об экспериментальных данных служат разнообразные справочные таблицы. Ссылки на большое число доступных (в частности, в интернете) таблиц по термохимии содержатся в статье B.Ruscic et al., Introduction to Active Thermochemical Tables; Several “Key” Enthalpies of Formation Revisited, J. Phys. Chem. A, 2004, **108**, 9979. Обзор данных о точности расчетов разнообразных свойств молекул неэмпирическими методами квантовой химии можно найти в разделе 3.6.3 учебного пособия: В.Г.Цирельсон «Квантовая химия. Молекулы, молекулярные системы и твердые тела», М. БИНОМ. Лаборатория знаний, 2010.

Большое число разнообразных экспериментальных данных можно найти на сайте Национального института стандартов (США)

http://webbook.nist.gov/chemistry

Чрезвычайно полезные сведения по сопоставлению экспериментальных и расчетных (квантовохимических) данных содержатся на том же сайте

http://cccbdb.nist.gov

К сожалению, в работах западных авторов нет ссылок на фундаментальное издание «Термодинамические свойства индивидуальных веществ», опубликованное (под редакцией В.П.Глушко) в 1978-1979 годах. В связи с тематикой данного руководства особый интерес представляют тома 1 и 2. Справочник имеется в библиотеке Института химии СПбГУ.

В последние годы в качестве «эталонных» значений термохимических данных все чаще используются данные, полученные в проекте ATcT (Активные термохимические таблицы). Подробное описание методики построения таблиц содержится в статье B.Rustcic et al., Intriduction to Active Thermochemical Tables: Several “Key” Enthalpies of Formation Revisited, J. Phys. Chem. A, 2004, **108**, 9979. Краткое изложение идей, реализованных в протоколе ATcT, можно найти в статье B.Ruscic, J.V. Michael, P.C.Redfern, L.A.Curtiss, K.Raghavachari, Simultaneous Adjustment of Experimentally Based Enthalpies of Formation of CF_3X, X=nil, H, Cl,Br, I, CF_3, CN, and a Probe of G3 Theory, J. Phys. Chem. A, 1998, **102**, 10889. Мы будем следовать этой публикации.

Термохимическая сеть представляет собой набор ***известных***, экспериментально определенных теплот химических реакций, которые связывают между собой элементы блока, содержащего ***искомые*** теплоты образования молекул. Вообще говоря, некоторые из этих энтальпий образования могут быть также известны с высокой точностью. На Рис. 2 представлена локальная сеть, включающая фториды углерода и некоторые их производные.

Рассмотрим реакцию AB+C□ AC+D. Присвоим этой реакции индекс i, пусть $\Delta_r H_{298}(i) = A_i \pm \Delta_i$, где A_i и Δ_i – определенные экспериментально энтальпия реакции и погрешность измерения, заявленная авторами эксперимента. Теплоты образования C и D будем рассматривать как постоянные, определенные из эксперимента. Энтальпии образования AB и

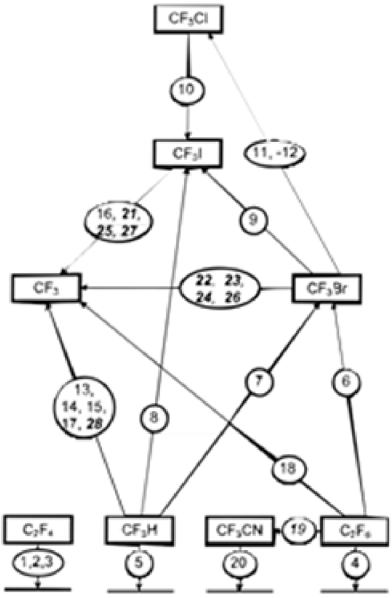


Рис. 2 Локальная термохимическая сеть, предназначенная для определения теплоты образования фторуглеродов и их производных

АС являются неизвестными. Имея сказанное в виду, после ряда преобразований получаем

$$\Delta_f H_{298}(AC) - \Delta_f H_{298}(AB) = (A_i \pm \Delta_i) + \Delta_f H_{298}(C) - \Delta_f H_{298}(D) = a_i \pm \delta_i.$$

Таким образом, сеть, содержащая m «звеньев» (энтальпий реакций) и n неизвестных (энтальпий образования), порождает систему m линейных уравнений с n неизвестными

$$\sum_j c_{ij}\, x_j = a_i \pm \delta_i, \quad \text{j=1, 2, … n, i=1, 2, … m.}$$

Как правило, m > n, так что система является переопределенной, кроме того, матрица C является очень разреженной, так как в каждой строке содержатся обычно лишь один – два ненулевых коэффициента c_{ij}, отражающих стехиометрию реакции. Решение системы производится методом наименьших квадратов.

Оценка точности результатов, полученных разными методами

В Приложениях мы воспроизводим тексты ряда статей, которые позволяют более подробно ознакомиться с расчетными процедурами, кратко описанными в предшествующем тексте. Кроме того, эти статьи дают возможность ближе познакомиться с величинами разных вкладов в конечный результат (например, теплоту образования молекулы), а также со значениями отклонений от экспериментальных данных для конкретных соединений. Эти работы позволяют проследить путь от методов, частично использующих эмпирические поправки и методы типа B3LYP, к процедурам, построенным (не считая ССП расчета) исключительно на расчетах методом связанных кластеров, в том числе включающих explicitly correlated варианты этого метода.

В этих статьях можно найти статистический анализ отклонений расчетных данных от эксперимента не только для энтальпий образования или атомизации, но и для потенциалов ионизации, сродству к электрону и протону.

Из данных, приведенных в указанных или упомянутых в тексте работах, видно, что с точки зрения точности расчета энтальпий образования молекул методы G3/G4 и W1 следует отнести к методам химической точности. Так, для методов G3/G4 среднеквадратичные отклонения (270 соединений) составляют 1.67 и 1.19 ккал/моль соответственно. Для более трудоемкого метода W1 эта величина (97 соединений) равна приблизительно 0.6±0.5 ккал/моль. По-видимому, более поздние варианты метода дают более высокую точность.

Методы, описанные в работах Зонтаса и др. и HEAT, позволяют получать теплоты образования с погрешностью порядка 0.5 кДж/моль. Следует отметить, что эти методы весьма трудоемки и требуют использования не всем доступных программ и расширенных вычислительных возможностей.

Приложения

THE JOURNAL OF CHEMICAL PHYSICS **126**, 084108 (2007)

Gaussian-4 theory

Larry A. Curtiss
Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439; Chemistry Division, Argonne National Laboratory, Argonne, Illinois 60439; and Center for Nanoscale Materials, Argonne National Laboratory, Argonne, Illinois 60439

Paul C. Redfern
Chemistry Division, Argonne National Laboratory, Argonne, Illinois 60439

Krishnan Raghavachari
Department of Chemistry, Indiana University, Bloomington, Indiana 47401

(Received 16 October 2006; accepted 3 January 2007; published online 28 February 2007)

The Gaussian-4 theory (G4 theory) for the calculation of energies of compounds containing first- (Li–F), second- (Na–Cl), and third-row main group (K, Ca, and Ga–Kr) atoms is presented. This theoretical procedure is the fourth in the Gaussian-*n* series of quantum chemical methods based on a sequence of single point energy calculations. The G4 theory modifies the Gaussian-3 (G3) theory in five ways. First, an extrapolation procedure is used to obtain the Hartree-Fock limit for inclusion in the total energy calculation. Second, the *d*-polarization sets are increased to 3*d* on the first-row atoms and to 4*d* on the second-row atoms, with reoptimization of the exponents for the latter. Third, the QCISD(T) method is replaced by the CCSD(T) method for the highest level of correlation treatment. Fourth, optimized geometries and zero-point energies are obtained with the B3LYP density functional. Fifth, two new higher level corrections are added to account for deficiencies in the energy calculations. The new method is assessed on the 454 experimental energies in the G3/05 test set [L. A. Curtiss, P. C. Redfern, and K. Raghavachari, J. Chem. Phys. **123**, 124107 (2005)], and the average absolute deviation from experiment shows significant improvement from 1.13 kcal/mol (G3 theory) to 0.83 kcal/mol (G4 theory). The largest improvement is found for 79 nonhydrogen systems (2.10 kcal/mol for G3 versus 1.13 kcal/mol for G4). The contributions of the new features to this improvement are analyzed and the performance on different types of energies is discussed.
[DOI: 10.1063/1.2436888]

I. INTRODUCTION

A number of approaches, based on quantum chemical methods, have been developed over the past decade and a half to make accurate predictions of thermochemical data. The Gaussian-*n* (G*n*) theories ($n=1,2,3$),[1–3] which we have developed, employ a set of calculations with different levels of accuracy and basis sets with the goal of approaching the exact energy. In the G*n* approach, a high level correlation calculation [e.g., QCISD(T) and CCSD(T)] with a moderate sized basis set is combined with energies from lower level calculations (e.g., MP4 and MP2) with larger basis sets to approximate the energies of more expensive calculations. In addition, several molecule-independent empirical parameters [higher level correction (HLC) terms] are included to estimate remaining deficiencies, assuming that they are systematic. An alternate and more accurate approach for the calculation of thermochemical data is based on CCSD(T) calculations using very large correlation consistent basis sets extrapolated to the complete basis set limit with addition of corrections for some smaller effects not included in the calculations such as core-valence effects, relativistic effects, and atomic spin-orbit effects.[4–11] This type of approach is limited to smaller molecules because of the use of very large basis sets. An intermediate approach referred to as correlation consistent composite approach (ccCA) that uses correlation consistent basis sets with no parametrization has recently been introduced.[12,13] Other composite techniques related to the Gaussian-*n* methods have also been introduced. These include the complete basis set methods of Montgomery, Jr. *et al.*[14,15] and the multicoefficient methods of Truhlar and co-workers.[16,17]

Concurrent with the development of the Gaussian-*n* theories, we have compiled a series of data sets of accurate experimental data, which have been used in the assessment of theoretical methods for energy calculations. The first in this series was the G2 test set of 125 energies.[2] This was followed by the G2/97 (301 energies),[18] G3/99 (376 energies),[19] and G3/05 (454 energies)[20] test sets. Each succeeding test set included energies from the preceding test sets and additional species of larger sizes and of different types. The test sets contain thermochemical data such as enthalpies of formation, ionization potentials, electron affinities, and proton affinities chosen based on a listed accuracy of ±1 kcal/mol or better in critical compilations. The latest test set, G3/05,[20] contains 270 enthalpies of formation, 105 ionization energies, 63 electron affinities, 10 proton affinities, and 6 hydrogen-bonded complexes. The expansion from G3/99 to G3/05 was done by including 14 new enthalpies of formation of nonhydrogen species, 58 energies of molecules containing third-row elements, and 6 hydrogen bonded complexes. The nonhydrogens were added because these have generally been the most difficult type of molecule for the G*n* methods to handle. The energies of the third-row elements and hydrogen bonded complexes were added to the test set for the first time in G3/05. This new test set provides a more rigorous database with which to evaluate quantum chemical methods than the previous ones.

When the G3 theory was originally published, it was assessed on the G2/97 test set and was found to have an average absolute deviation of 1.02 kcal/mol from experiment.[3] The two succeeding test sets, G3/99 and G3/05, gave average absolute deviations of 1.07 and 1.13 kcal/mol, respectively, for the G3 theory.[19,20] The increase in the error is partially due to the poor performance of the method on the subset of first- and second-row nonhydrogen species with the deviation of that subset increasing from 1.66 kcal/mol (G2/97) to 2.10 kcal/mol (G3/99) to 2.34 kcal/mol (G3/05). Each succeeding nonhydrogen subset contains, on average, larger species, which is also responsible for the increase in the average error. (The average absolute deviation given above for the nonhydrogen subset of G3/05 does not include 18 species from the third-row entries which are smaller and, when included in the statistics, reduce the error to 2.10 kcal/mol.) The other contributor to the increase in deviation from the G2/97 test set is an increase in the error in the substituted hydrocarbon and inorganic hydride subsets, although these are still under 1 kcal/mol. Thus, molecules with no hydrogens are the most challenging for the G3 theory.

In this paper, we describe modifications to the G3 theory that help to improve its overall performance as well as its performance on the nonhydrogen species. We refer to the resulting method as the G4 theory. This methodology is consistent with the G*n* approach in past publications: i.e., it is a composite technique aimed at getting accurate energies without requiring extensive computer resources. This approach depends on a cancellation of errors as well as well-defined parameters to achieve this. The modifications included in the G4 theory include (1) an extrapolation procedure to obtain the Hartree-Fock (HF) limit for inclusion in the total energy calculation, (2) increase of the *d*-polarization sets to 3*d* on the first-row atoms and to 4*d* on the second-row atoms, with reoptimization of the exponents for the 4*d* set, (3) the replacement of the QCISD(T) method by CCSD(T), (4) geometries and zero-point energies obtained at the B3LYP/6-31G(2*df*,*p*) level, and (5) two new higher level correction parameters to account for deficiencies in radicals and in species having only one electron pair in the valence space. The HF energy extrapolation used here eliminates any error that may be present due to the incompleteness of the basis set in the HF energy, thus reducing the sources of error in the calculation. This was partially implemented in a version of the G3 theory published in 2001, referred to as the G3X theory,[21] in which a *g* function was added to the second-row atoms. The use of density functional geometries and zero-point energies was also included in the G3X method. The details of these five modifications are described in Sec. II. It is shown in Sec. III that the G4 theory gives a significant overall improvement on the G3/05 test set, particularly for nonhydrogens. In Sec. III, we also discuss reasons for the remaining outliers, i.e., molecules with errors larger than 2 kcal/mol. We present a "complete" version of the G4 theory that is based on a single calculation using the full basis set. This method, G4 (complete), corrects a problem with the G3 theory, where calculations without additivity approximation have a significantly larger average absolute deviation than the version with additivity. Conclusions are given in Sec. IV.

II. DESCRIPTION OF THE GAUSSIAN-4 THEORY

The Gaussian-4 theory is a composite technique in which a sequence of well-defined *ab initio* molecular orbital calculations is performed to arrive at a total energy for a given molecular species. The steps in the G4 theory and the differences with the G3 theory[3] are as follows.

(1) The equilibrium structure is obtained at the B3LYP/6-31G(2*df*,*p*) level. Spin-restricted theory is used for singlet states and spin-unrestricted theory for others. This differs from the G3 theory in which the geometries are calculated at the MP2(full)/6-31G* level. As shown in our paper on the G3X theory,[21] the use of the B3LYP density functional theory[22] for geometries leads to an improvement in overall results compared to using the MP2 theory, which was used in the formulation of G1, G2, and G3 theories.

(2) The B3LYP/6-31G(2*df*,*p*) equilibrium structure is used to calculate harmonic frequencies, which are then scaled by a factor of 0.9854 (Ref. 21) to take account of known deficiencies at this level. These frequencies give the zero-point energy $E(\mathrm{ZPE})$ used to obtain E_0. This is a change from the G3 theory, in which HF/6-31G* was used for zero-point energies and the scale factor was appropriate for vibrational frequencies.[23] In most cases this modification should be more reliable.

(3) The Hartree-Fock energy limit $E(\mathrm{HF/limit})$ is calculated. This is a new step that was not included in previous methods. The Hartree-Fock basis set limit is determined using a linear two-point extrapolation scheme[24,25] and Dunning's aug-cc-pV*n*Z basis sets,[26–28]

$$E_{\mathrm{HF/aug\text{-}cc\text{-}pV}n\mathrm{Z}} = E_{\mathrm{HF/limit}} + B\exp(-\alpha n), \qquad (1)$$

where n is the number of contractions in the valence shell of the basis set and α is an adjustable parameter. The above formula yields a set of two linear equations with two unknowns from which the Hartree-Fock limit can be determined analytically, $E_{\mathrm{HF/limit}}(n,n+1) = (E_{\mathrm{HF}/n+1} - E_{\mathrm{HF}/n}\exp(-\alpha))/(1-\exp(-\alpha))$. We investigated various pairs of $n,n+1$ values and α values. We found that calculating the Hartree-Fock limit using $n=4$, $n+1=5$ (aug-cc-pVQZ and aug-cc-pV5Z basis sets), and $\alpha=1.63$ gave nearly converged values for a set of large molecules from the G3/05 test set. In order to reduce the computational time required, we modified the standard aug-cc-pVQZ and aug-cc-pV5Z basis sets by reducing the number of diffuse functions on heavy atoms and by reducing the hydrogen basis set as described in Appendix A. These basis set modifications saved significant computer time, without reducing the accuracy.

(4) A series of single point correlation energy calculations is then carried out. The first is based on the complete fourth-order Moller-Plesset perturbation theory with the 6-31G(d) basis set, MP4/6-31G(d). This is modified by corrections from additional calculations.

(a) a correction for diffuse functions,

$$\Delta E(+) = E[\text{MP4/6-31+G}(d)] - E[\text{MP4/6-31G}(d)], \tag{2}$$

(b) a correction for higher polarization functions,

$$\Delta E(2df,p) = E[\text{MP4/6-31G}(2df,p)] - E[\text{MP4/6-31G}(d)], \tag{3}$$

(c) a correction for correlation effects beyond a fourth-order perturbation theory using a coupled cluster theory,

$$\Delta E(\text{CC}) = E[\text{CCSD(T)/6-31G}(d)] - E[\text{MP4/6-31G}(d)], \tag{4}$$

(d) a correction for larger basis set effects and for the nonadditivity caused by the assumption of separate basis set extensions for diffuse functions and higher polarization functions,

$$\begin{aligned}\Delta E(\text{G3LargeXP}) = E[&\text{MP2(full)/G3LargeXP}] \\ &- E[\text{MP2/6-31G}(2df,p)] \\ &- E[\text{MP2/6-31+G}(d)] \\ &+ E[\text{MP2/6-31G}(d)].\end{aligned} \tag{5}$$

The basis sets in (a)–(c) are the same as those in the G3 theory.[3] In step (d) the G3LargeXP basis is used instead of the G3Large basis set[3] to correct for some additivity problems discussed in Sec. III. The difference between the two basis sets is added d-polarization functions in G3LargeXP (with XP standing for extra polarization functions). The new exponents and a description of how they were obtained are given in Appendix A. The $2df$ polarization set in G3Large on the first row is replaced by a $3df$ set in G3LargeXP, the $3d2f$ polarization functions on the second row (Al–Cl) are replaced by $4d2f$, and no changes are made for Na, Mg, K, Ca, and Ga–Kr. The other difference in step 4 is that the QCISD(T)/6-31G* calculation is replaced by CCSD(T)/6-31G*. This is done because in some cases the QCISD(T) method has rather dramatic failures, which does not occur for the CCSD(T) method.[29,30] The MP4 and CCSD(T) calculations are done in the frozen core approximation, while the MP2 calculation with the large basis set is done with all electrons correlated. The electrons included in the frozen core for the G4 theory are the same as those for the G3 theory; i.e., the $3d$ on Ga–Kr and $3s$ and $3p$ on K and Ca are included in the correlation space and the $2s,2p$ on Na and Mg are included in the valence space. This has been defined elsewhere[31] as the "small core" treatment of correlation.

(5) The MP4/6-31G(d) energy and the four correlation corrections from step 4 are combined in an additive manner along with a correction for the HF limit (step 3) and a spin-orbit correction, ΔE(SO)

$$\begin{aligned}E(\text{combined}) = E[&\text{MP4/6-31G}(d)] + \Delta E(+) \\ &+ \Delta E(2df,p) + \Delta E(\text{CC}) \\ &+ \Delta E(\text{G3LargeXP}) + \Delta E(\text{HF}) \\ &+ \Delta E(\text{SO}).\end{aligned} \tag{6}$$

The ΔE(HF) is calculated as the difference between E(HF/G3LargeXP) calculated in step 4 and E(HF/limit) calculated in step 3 [i.e., $\Delta E(\text{HF}) = E(\text{HF/limit}) - E(\text{HF/G3LargeXP})$]. A spin-orbit term is included for atoms, as described previously for the G3 theory for the first- and second-row[3] and for the third-row main group elements.[32–34] This is taken from experiment[35] where available and accurate calculations[3,36] elsewhere. The atomic spin-orbit corrections are listed in Table I. A spin-orbit correction is also included for all molecules with first order corrections ($^2\Pi$ diatomics). In this case the spin-orbit correction is taken from accurate theoretical calculations and are given in Ref. 32.

(6) A HLC is added to take into account remaining deficiencies in the energy calculations,

$$E_e(\text{G4}) = E(\text{combined}) + E(\text{HLC}). \tag{7}$$

The form of the HLC is the same as for the G3 theory[3] except that two additional parameters (A' and E) have been added. The HLC parametrization used in the G4 theory is $-An_\beta$ for closed shell molecules, $-A'n_\beta - B(n_\alpha - n_\beta)$ for open shell systems, and $-Cn_\beta - D(n_\alpha - n_\beta)$ for atoms (including atomic ions). The n_β and n_α are the number of β and α valence electrons, respectively, with $n_\alpha \geqslant n_\beta$. The number of valence electron pairs corresponds to n_β. The A' parameter has been added to the G4 theory to account for deficiencies in pairs of electrons in radical molecular species including ions. In addition, we have added a parameter, E, that corrects for the energy of pairs of electrons in molecular and atomic species whose valence electrons consist only of one pair of s electrons (not including systems having one or more $1s$ electrons). These *single electron pair* species represent a special case for which the basis set requirements are not as great. The A, A', B, C, and D values are chosen to give the smallest average absolute deviation from experiment for the whole G3/05 test set. The value of E is determined by the minimization of the root mean square deviation of the energies involving the subset of 13 species from the G5/03 test set that involve single pairs of electrons (see Table IV for this subset). For the G4 theory, $A = 6.947$ mhartree, $B = 2.441$ mhartree, $C = 7.116$ mhartree, $D = 1.414$ mhartree, $A' = 7.128$ mhartree, and

TABLE I. Total G4 energies (in hartrees) of atomic species and spin-orbit corrections (in mhartrees).

Atomic species	Atomic species	E_0(G4)	ΔE(SO)[a]	Atomic species	E_0(G4)	ΔE(SO)[a]
H (^2S)	H (^2S)	−0.501 42	0.0	F$^+$ (^3P)	−99.066 11	−0.67
He (^1S)	He (^1S)	−2.904 91	0.0	Ne$^+$ (^2P)	−128.108 67	−1.19
Li (^2S)	Li (^2S)	−7.466 36	0.0	Na$^+$ (^1S)	−161.928 92	0.0
Be (^1S)	Be (^1S)	−14.657 65	0.0	Mg$^+$ (^2S)	−199.630 07	0.0
B (^2P)	B (^2P)	−24.646 65	−0.05	Al$^+$ (^1S)	−242.001 35	0.0
C (^3P)	C (^3P)	−37.834 17	−0.14	Si$^+$ (^2P)	−288.937 90	−0.93
N (^4S)	N (^4S)	−54.573 67	0.0	P$^+$ (^3P)	−340.749 63	−1.43
O (^3P)	O (^3P)	−75.045 50	−0.36	S$^+$ (^4S)	−397.601 63	0.0
F (^2P)	F (^2P)	−99.704 98	−0.61	Cl$^+$ (^3P)	−459.540 26	−1.68
Ne (^1S)	Ne (^1S)	−128.900 99	0.0	Ar$^+$ (^2P)	−526.822 78	−2.18
Na (^2S)	Na (^2S)	−162.117 89	0.0	K$^+$ (^1S)	−599.553 54	0.0
Mg (^1S)	Mg (^1S)	−199.912 04	0.0	Ca$^+$ (^2S)	−677.139 71	0.0
Al (^2P)	Al (^2P)	−242.221 07	−0.34	Ga$^+$ (^1S)	−1924.129 16	0.0
Si (^3P)	Si (^3P)	−289.237 04	−0.68	Ge$^+$ (^2P)	−2076.150 05	−5.37
P (^4S)	P (^4S)	−341.134 63	0.0	As$^+$ (^3P)	−2234.951 93	−8.04
S (^3P)	S (^3P)	−397.980 18	−0.89	Se$^+$ (^4S)	−2400.620 74	0.0
Cl (^2P)	Cl (^2P)	−460.015 05	−1.34	Br$^+$ (^3P)	−2573.151 78	−6.71
Ar (^1S)	Ar (^1S)	−527.400 45	0.0	Kr$^+$ (^2P)	−2752.713 29	−8.16
K (^2S)	K (^2S)	−599.712 19	0.0	Li$^-$ (^1S)	−7.490 42	0.0
Ca (^1S)	Ca (^1S)	−677.362 61	0.0	B$^-$ (^3P)	−24.655 71	−0.03[b]
Ga (^2P)	Ga (^2P)	−1924.350 57	−2.51	C$^-$ (^4S)	−37.879 08	0.0
Ge (^3P)	Ge (^3P)	−2076.440 70	−4.41	O$^-$ (^2P)	−75.098 47	−0.26[b]
As (^4S)	As (^4S)	−2235.312 07	0.0	F$^-$ (^1S)	−99.833 64	0.0
Se (^3P)	Se (^3P)	−2400.977 97	−4.3	Na$^-$ (^1S)	−162.139 76	0.0
Br (^2P)	Br (^2P)	−2573.585 37	−5.6	Al$^-$ (^3P)	−242.235 93	−0.28[b]
Kr (^1S)	Kr (^1S)	−2753.225 82	0.0	Si$^-$ (^4S)	−289.286 57	0.0
He$^+$ (^2S)	He$^+$ (^2S)	−2.001 39	0.0	P$^-$ (^3P)	−341.159 86	−0.45[b]
Li$^+$ (^1S)	Li$^+$ (^1S)	−7.267 61	0.0	S$^-$ (^2P)	−398.055 13	−0.88[b]
Be$^+$ (^2S)	Be$^+$ (^2S)	−14.313 78	0.0	Cl$^-$ (^1S)	−460.146 71	0.0
B$^+$ (^1S)	B$^+$ (^1S)	−24.343 23	0.0	K$^-$ (^1S)	−599.730 00	0.0
C$^+$ (^2P)	C$^+$ (^2P)	−37.421 83	−0.2	Ge$^-$ (^4S)	−2076.487 18	0.0
N$^+$ (^3P)	N$^+$ (^3P)	−54.040 65	−0.43	Br$^-$ (^1S)	−2573.710 12	0.0
O$^+$ (^4S)	O$^+$ (^4S)	−74.547 31	0.0			

[a] Spin-orbit corrections are from Ref. 33, except where noted.

[b] Calculated value, Ref. 34.

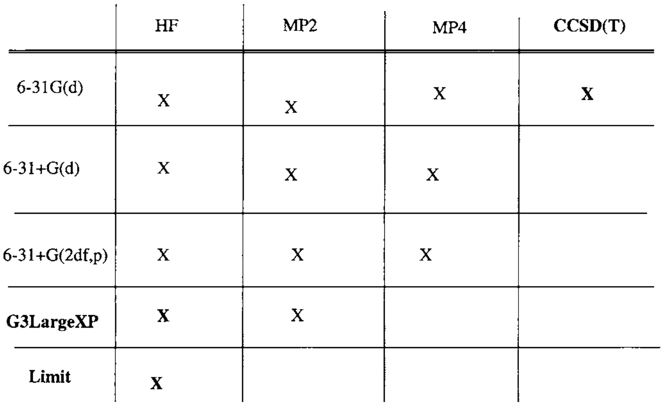
E=2.745 mhartree. The effect of adding the two new parameters is discussed in more detail in the next section. Note that in the case of the third-row species (K, Ca, and Ga–Kr) the HLC is based only on the valence electrons, i.e., 4*s* and 4*p*, although more orbitals are included in the correlation calculation (see step 4 above). This is slightly different from what was done in the G3 theory for the third row[32] where the HLC for K and Ca included 3*s* and 3*p* as valence electrons for the calculation of the HLC.

(7) Finally, the total energy at 0 K is obtained by adding the zero-point energy, obtained from the frequencies of step 2, to the total energy,

$$E_0(\mathrm{G4}) = E_e(\mathrm{G4}) + E(\mathrm{ZPE}). \tag{8}$$

The energy E_0 is referred to as the “G4 energy.”

The single point entry calculations in the G4 theory are summarized in Fig. 1. The final total energy is effectively at the CCSD(T,full)/G3LargeXP+HFlimit level if the different additivity approximations work well. The validity of such approximations is discussed in the next section. All calculations in this paper were done with the GAUSSIAN 03 computer program.[37] All of the basis sets in the G4 theory, with the exception of G3LargeXP and the third-row basis sets, are standard in GAUSSIAN 03. The latter basis sets are available on the web[38] and in the supplementary information.[39]



	HF	MP2	MP4	**CCSD(T)**
6-31G(d)	X	X	X	**X**
6-31+G(d)	X	X	X	
6-31+G(2df,p)	X	X	X	
G3LargeXP	**X**	X		
Limit	**X**			

FIG. 1. Single point energies used in the G4 theory (bold entries are new or modified from the G3 theory).

TABLE II. Comparison of average absolute deviations and root mean square deviations for G3 and G4 theories for the G3/05 test.

	Average absolute deviation (kcal/mol)		Root mean square deviation (kcal/mol)	
Type of energy	G3	G4	G3	G4
Enthalpies of formation (270)	1.19	0.80	1.78	1.10
Nonhydrogens (79)	2.10	1.13	2.85	1.49
Hydrocarbons (38)	0.69	0.48	0.93	0.70
Substitutes hydrocarbons (100)	0.82	0.68	1.12	0.94
Inorganic hydrides (19)	0.95	0.92	1.19	1.12
Radicals (34)	0.83	0.66	0.97	0.79
Ionization energies (105)	1.09	0.91	1.61	1.45
Electron affinities (63)	0.97	0.83	1.35	1.03
Proton affinities (10)	1.14	0.84	1.29	1.04
Hydrogen bonded complexes (6)	0.60	1.12	0.77	1.53
All (454)	1.13	0.83	1.67	1.19

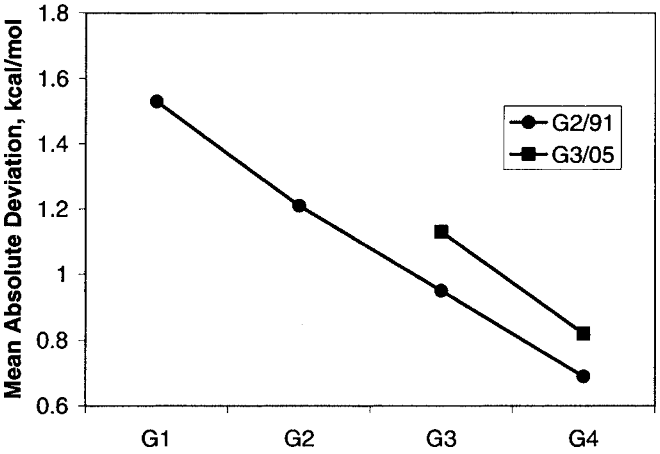


FIG. 2. Comparison of accuracies of Gn methods for the G2/91 (also referred to as the G2 test set) and G3/05 test sets.

III. ASSESSMENT OF THE G4 THEORY ON THE G3/05 TEST SET

The G3/05 test set[20] contains 454 energies including enthalpies of formation of neutrals, atomization energies, ionization potentials, electron affinities, proton affinities, and hydrogen bond energies. This test set includes energies for species containing first-, second-, and third- and (K, Ca, and Ga–Kr) row elements. The enthalpies of formation at 298 K were calculated as in Ref. 18. The ionization potentials, electron affinities, atomization energies, and proton affinities were calculated at 0 K.[18]

The G4 theory was used to calculate the energies of atoms, molecules, and ions in the G3/05 test set. Table I contains the G4 total energies of the atomic species and the spin-orbit corrections $\Delta E(\mathrm{SO})$, which are included in the total energies. The G4 total energies for the molecules and their geometries are available elsewhere,[38,39] as are the G4 deviations from experiment of all 454 energies. Table II contains a summary of the average absolute deviations and root mean square deviations of the G4 theory from experiment, with results for G3 included for comparison.

A. Comparison of G3 and G4 theories

The results in Table II indicate that for the 454 energies, the average absolute deviation from experiment at the G4 level is 0.83 kcal/mol, which is a significant improvement over 1.13 kcal/mol for the G3 theory. The root mean square deviation of the G4 theory (1.19 kcal/mol) also significantly improves compared to that of the G3 theory (1.67 kcal/mol). The G4 theory especially improves for enthalpies of formation (1.19 kcal/mol for G3 versus 0.80 kcal/mol for G4), with the largest improvement occurring for nonhydrogens, although all of the other enthalpy types (hydrocarbons, substituted hydrocarbons, inorganic hydrides, and radicals) also show improvement due to the changes made in the G4 theory. The average absolute deviation for enthalpies of formation of 79 nonhydrogens decreases from 2.10 kcal/mol (G3) to 1.13 kcal/mol (G4). The improvement for the remaining species is smaller since the G3 theory is already well within the target accuracy of 1 kcal/mol. The next largest improvement occurs for the 38 hydrocarbons for which the average absolute deviation decreases from 0.69 kcal/mol (G3) to 0.48 kcal/mol (G4). The average absolute deviation for the 100 substituted hydrocarbons decreases from 0.82 to 0.68 kcal/mol, while that for the 34 radicals decreases from 0.83 to 0.66 kcal/mol. Finally, the enthalpies for 19 inorganic hydrides improve slightly from 0.95 to 0.92 kcal/mol.

The G4 theory also has significant improvements over the G3 theory for ionization energies electron affinities, and proton affinities. These three quantities have average absolute deviations of 0.91, 0.83, and 0.84 kcal/mol, respectively, at the G4 level compared to 1.09, 0.97, and 1.14 kcal/mol at the G3 level. The only type of energy for which the accuracy decreases is hydrogen bond energies, which have an average absolute deviation of 1.12 kcal/mol at the G4 level compared to 0.60 kcal/mol at the G3 level. This increase is due to poor results for the water and hydrogen fluoride dimers where the B3LYP/6-31G$(2df,p)$ geometries fail. This will be discussed in more detail later.

The G3X theory,[21] which we previously published, is an extension of the G3 theory that has two of the new features that are included in the G4 theory. The average absolute deviation of G3X from experiment is 1.01 kcal/mol for the G3/05 test set. The G4 theory has an additional three new features that make a substantial improvement over G3X to reduce the average absolute deviation to 0.83 kcal/mol. First, the HF limit step in G4 is more rigorous than the inclusion of a single g function on second-row atoms for the G3X theory. While the inclusion of a g function is an improvement, it can still miss significant HF energy. For example, for SF_6 the HF basis set in G3X still misses 1.8 kcal/mol of the HF binding energy compared to the extrapolated limit. Second, the expanded d-polarization set in the G4 theory corrects a problem in the complete (i.e., nonadditive) version of G3 and G3X theories, as discussed in Sec. III C. Third, the higher level correction parameter set is increased from four to six. A comparison of the accuracy of the Gn methods, $n=1-4$, is shown in Fig. 2.

TABLE III. Contributions of different modifications of G3 theory to the performance of G4 theory on the G3/05 test set. [G3//B2*df* uses B3LYP/6-31G(2*df*,*p*) optimized geometries and scaled zero-point energies instead of MP2/6-31G* geometries and HF/6-31G* scaled zero-point energies; G3(CC)//B2*df* replaces QCISD(T)/6-31G* with CCSD(T)/6-31G*; G3(CC,HF)//B2*df* adds the estimated HF limit to the energy where $\Delta E(\mathrm{HF}) = E(\mathrm{HF/limit}) - E(\mathrm{HF/G3Large})$; G3(CC,HF,XP)//B2*df* includes extra polarization functions on the G3Large basis set, i.e, G3LargeXP. See text for more details of the modifications. In each case the HLC as defined for the G4 theory was reoptimized using the G3/99 test set.]

	Average absolute deviation (kcal/mol)					
	All (454)	Neutral enthalpies[a]	IPs	EAs	PAs	H-bond complexes
G3	1.13	1.19(2.10,0.69,0.82,0.95,0.83)	1.09	0.97	1.14	0.60
G3//B2*df*	1.06	1.09(1.88,0.55,0.82,0.87,0.75)	1.04	0.99	1.09	1.06
G3(CC)//B2*df*	1.06	1.08(1.85,0.55,0.81,0.83,0.79)	1.06	1.01	1.05	1.06
G3(CC,HF)//B2*df*	0.92	0.87(1.26,0.61,0.68,0.88,0.78)	1.07	0.93	0.83	1.13
G3(CC,HF,XP)//B2*df*	0.90	0.84(1.27,0.47,0.69,0.89,0.68)	1.01	0.95	0.84	1.12
G3(CC,HF,XP,HLC5)	0.87	0.83(1.27,0.48,0.68,0.87,0.65)	0.92	0.93	0.83	1.12
G3(CC,HF,XP,HLC6)=G4	0.83	0.80(1.13,0.48,0.68,0.92,0.66)	0.91	0.83	0.84	1.12

[a] Values in parentheses are for the molecules in the subsets in the following order: nonhydrogens, hydrocarbons, substituted hydrocarbons, inorganic hydrides, and radicals. The neutral enthalpy subset here includes results for atomization energies for third-row species that were added to the test set in Ref. 18.

B. Assessment of the new features of the G4 theory

We calculated the contributions of the five new features in the G4 theory to the average absolute deviation from experiment relative to the G3 theory in a stepwise manner, including reoptimization of the HLC at each step. The results for the G3/05 test set are summarized in Table III and are discussed in this section.

1. Geometries and zero-point energies (G3//2df)

The use of B3LYP/6-31G(2*df*,*p*) geometries and zero-point energies (steps 1 and 2) reduces the overall average absolute deviation from 1.13 to 1.06 kcal/mol with the improvement largely coming from the enthalpies for formation. The use of density functional geometries is most important for the nonhydrogens where the average absolute deviation of the subset decreases from 2.10 to 1.88 kcal/mol. This has been noted in our paper on the G3X theory,[21] with examples such as PF_5, SO_2, SO_3, and SF_6, where the density functional theory gives improved geometries. On the other hand, as noted above, the B3LYP/6-31G(2*df*,*p*) method does poorly on geometries of several hydrogen bonded dimers (hydrogen fluoride and water dimers), which increases the average absolute deviation for this set.

2. CCSD(T) [G3(CC)//2df]

The use of the CCSD(T) method [step (4)(c)] results in no change in the average absolute deviation for the G3/05 test set, as expected based on previous studies.[29,30] The change from QCISD(T) to CCSD(T) is expected to help in specific cases where the former fails. Such cases are not included in the test set, but may be encountered in applications of the methods.

3. HF limit [G3(CC,HF)//2df]

Of the five modifications, the inclusion of the HF limit energy in the formulation of the G4 theory (steps 3 and 5) has the largest effect, as it reduces the overall average absolute deviation from 1.06 to 0.92 kcal/mol. The improvement largely comes from the enthalpies of formation where it decreases the average absolute deviation from 1.08 to 0.87 kcal/mol. In particular, the inclusion of the HF limit is most important for the nonhydrogens where the HF/G3LargeXP energy is farthest from the HF limit. The average absolute deviation of the 79 molecules in the nonhydrogen set decreases from 1.85 to 1.26 kcal/mol. Most of the other subsets show some improvement from this new feature, but not as large.

The use of smaller basis sets than those described in step 3 to obtain the extrapolation to the HF limit was investigated. We tried using the aug-cc-pVTZ and aug-cc-pVQZ basis sets for the extrapolation in Eq. (1), but this pair of basis sets did not converge adequately for larger nonhydrogen species. We also investigated extrapolations based on the G3Large basis set with variation of the polarization sets and found that this approach was also not adequate for obtaining the HF limit energy. Finally, Martin[40] reported that tight polarization functions had a significant effect on the atomization energy of SO_2 at the HF level. We have investigated the inclusion of tight polarization functions in the basis sets used for the SO_2 HF extrapolation and found that the proposed extrapolation gives essentially the same result whether or not tight polarization functions are added.

4. Basis set change [G3(CC,HF,XP)//2df]

The addition of more *d* functions to the G3Large basis set [step (4)(d)], i.e., G3LargeXP, for use in the MP2(full) calculation leads to a slight improvement in the average absolute deviation from 0.92 to 0.90 kcal/mol. The largest improvement is for hydrocarbon enthalpies and ionization potentials (see Table III). This change also corrects some deficiencies in the complete calculation, i.e., without the use of any additivity approximations, as discussed in Sec. III C.

5. Higher level corrections

The addition of the two new HLC parameters reduces the average absolute deviation from 0.90 to 0.83 for the whole test set, with both parameters contributing to this improvement. The addition of the added A' parameter for mol-

TABLE IV. Comparison of deviations for energies involving species having a single valence electron pair in the G3/05 test set. (Not including those with one or more 1*s* electrons.)

	Deviation (kcal/mol)	
	G3(CC,HF,XP,HLC5)[a]	G4
ΔH_f(BeH)	−1.01	1.81
ΔH_f(Li_2)	2.33	−0.30
ΔH_f(Na_2)	3.74	1.11
ΔH_f(LiNa)	3.21	0.58
ΔH_f(BeF_2)	−3.45	−0.82
ΔH_f($MgCl_2$)	−0.72	1.91
D_0(K_2)	−2.27	0.36
IP(Be)	−3.54	−0.86
IP(Mg)	−0.30	−0.62
IP(Ca)	−1.65	1.03
EA(Li)	−3.52	−0.84
EA(Na)	−3.77	−1.09
EA(K)	−2.33	0.35

[a] Results without inclusion of the HLC parameter (E) for the single electron pair species.

ecules provides for a different parameter for paired electrons in the closed shell species (A) compared to open shell species (A'), including radical ions and neutrals. This is important since spin polarization in the latter case can lead to differences in the correction needed compared to closed shell systems. Note that a similar parameter could be added for atoms, but since there are few closed shell atomic systems in the test set (He, Ne, Ar, Kr, F^-, Cl^-, B^+, and Al^+), this is not as important. The closed shell atoms Be, Mg, and Ca are covered by the new E parameter. The most significant improvement from the addition of the A' parameter occurs for ionization potentials (IPs) whose average absolute deviation decreases from 1.01 to 0.92 kcal/mol. The neutral radical enthalpies also improve from 0.68 to 0.65, and the electron affinities improve slightly from 0.95 to 0.93 kcal/mol. The overall average absolute deviation for the G3/05 test set decreases from 0.90 to 0.87 kcal/mol with the addition of this parameter. The addition of the HLC parameter for single pairs (E) further reduces the average absolute deviation to 0.83 kcal/mol, with the largest effect on molecules and atoms having a single valence electron pair. A summary of the energies involving these species is given in Table IV, showing the overall improvement when this HLC parameter is added. The E parameter is smaller by about 4 mhartree than the value for A, indicating that the reason for doing this was valid, i.e., that the basis set requirement and hence the correction required for these systems are much smaller.

Finally, we note that when the HLC is not included in G4 theory, the average absolute deviation increases to 8.6 kcal/mol, which is slightly lower than the value of 9.1 kcal/mol for the G3 theory. A detailed table of the deviations without the HLC included in the G4 theory is given in the supplementary tables.[39]

6. Timings

The cost of a G4 calculation is increased compared to a G3 theory calculation due largely to the two HF calculations used to extrapolate to the HF limit in step 3. The ratios of computer times for a G4 calculation compared to a G3 calculation for four representative molecules, benzene, hexane, octane, and silicon tetrachloride, are given in Table V. Also given are the relative times for a G2 calculation for benzene and silicon tetrachloride. For these molecules the G4 calculation takes two to three times more CPU time than G3, so that the savings that was gained in G3 compared to G2 is eliminated. However, since this increase is a result of HF calculations, which scale only as $\sim n^{3-4}$, the increase is not a serious problem for these calculations. The CPU increase is less for nonhydrogens (e.g., $SiCl_4$) compared to hydrocarbons due to the basis sets on the hydrogens in the HF calculations.

TABLE V. Relative CPU times used in G2, G3, and G4 single point energy calculations.

Method	$SiCl_4$	Benzene	Hexane	Heptane
G2	2.4	1.9	⋯	⋯
G3	1.0	1.0	1.0	1.0
G4	1.9	3.0	3.0	2.5

C. Assessment of additivity approximation in G4 theory

The additivity approximations in the Gaussian-*n* approach to the computation of molecular energies are essential to reducing computer resource requirements. These approximations can be assessed by calculating energies at the highest correlation level [QCISD(T) or CCSD(T)] with the largest basis set of the specific method and reoptimization of the HLC. Ideally, the geometries and zero-point energies from this level of theory would also be used, but due to the cost we have not done this. The additivity approximations in the G2 theory were previously investigated by calculating QCISD(T)/6-311+G($3df$,$2p$)//MP2(full)/6-31G* energies, scaled HF/6-31G* zero-point energies, and a reoptimized HLC.[41] This type of calculation has been referred to as G2(complete). A comparison of the G2(complete) results for the G2 test set of 125 energies found the approximations to work well, with only two energies differing by more than 1 kcal/mol (the electron affinities of NH_2 and OH). The average absolute deviation of the full basis set calculation was very close to that using the additive approximation (1.17 kcal/mol vs 1.21 kcal/mol).

In the current work we have investigated the additivity approximations of the G4 theory along with those of the G3 theory, which were not previously investigated. Results using analogous G3(complete) and G4(complete) methods have been obtained for the G2 test set of 125 energies. The G4(complete) method uses CCSD(T,full)/G3LargeXP//B3LYP/6-31G($2df$,p) energies along with scaled B3LYP/6-31G($2df$,p) zero-point energies, Hartree-Fock extrapolation, atomic spin-orbit corrections, and reoptimized HLC parameters. The G3(complete) method uses QCISD(T,full)/G3Large//B3LYP/6-31G($2df$,p) energies along with scaled HF/6-31G* zero-point energies, atomic spin-orbit corrections, and reoptimized HLC parameters. A

TABLE VI. Comparison of G3 and G4 methods with and without additivity approximation on the G2 test set of 125 energies.[a,b]

			Average absolute deviation (kcal/mol)			
Breakdown	G2	G2(complete)	G3	G3(complete)	G4	G4(complete)
All (125)	1.21	1.17	0.95	1.07	0.72	0.79
ΔH_f			0.87	0.99	0.65	0.78
IP			0.95	0.90	0.70	0.84
EA			1.01	1.45	0.82	0.70
PA			1.32	1.34	1.01	1.00

[a] G3(complete) is a QCISD(T,full)/G3Large//MP2(FU)/6-31G* energy calculation with scaled HF/6-31G* zero-point energies. G4(complete) is a CCSD(T,full)/G3LargeXP//B3LYP/6-31G(2*df*,*p*) plus HF/limit extrapolation calculation with scaled B3LYP/6-31G(2*df*,*p*) zero-point energies.

[b] All methods have HLC parameters individually optimized for the G2/91 test set. HLC parameters for the G4 theory: $A=8.669$ mhartree, $B=3.126$ mhartree, $C=8.723$ mhartree, $D=2.337$ mhartree, $A'=8.949$ mhartree, and $E=3.216$ mhartree. HLC parameters for G4(complete): $A=9.815$ mhartree, $B=3.727$ mhartree, $C=9.534$ mhartree, $D=2.872$ mhartree, $A'=9.871$ mhartree, and $E=4.904$ mhartree.

summary of the results is given in Table VI. Previous results for the G2 theory are also included. The results in Table VI are surprising because G3 and G4 theories both perform better than the corresponding versions without the additivity approximations. The average absolute deviation of G3 is 0.12 kcal/mol smaller than G3(complete) and that of G4 is 0.07 kcal/mol smaller than G4(complete). In each case about 11 species differ by more than 1 kcal/mol. These species fall into two categories. The first is enthalpies involving silicon and phosphorus (specifically, SiH_2, SiH_3, SiH_4, PH_2, PH_3, SiO, and Si_2H_6) and the second is the electron affinities of O, F, OH, and NH_2.

The problem with the enthalpies of formation of the Si and P species is largely due to the correlation treatment of core electrons. Martin *et al.*[42] have pointed out that the MP2(full)/G3Large level of theory gives a reasonable account of core correlation due to a cancellation of errors from using MP2 instead of CCSD(T) (that typically leads to an overestimation) and a relatively small basis set (that leads to an underestimation). From our calculations we find that the evaluation of core correlation at the MP2 level versus the QCISD(T) [or CCSD(T)] level is responsible for most of the difference between G*n* and G*n*(complete), $n=3,4$, for the enthalpies of the seven Si and P species listed above. The use of MP2(full)/G3Large for core correlation in the additive methods gives better agreement with experiment due to the cancellation of errors.

There are two sources for the nonadditivity of the G3 and G4 methods in the calculation of electron affinities. The first source, which has been corrected in the G4 theory, is a deficiency in the *d*-polarization functions and the second, which is not corrected in the G4 theory, is due to nonadditivity of the effects of *d*-polarization and diffuse functions at the higher correlation levels. The *d*-polarization function deficiency is caused by the use of two types of basis sets, i.e., 6-31G(*d*) in QCISD(T) and MP4 calculations and 6-311G(*d*) basis in the large basis set (G3Large) for MP2 calculations. The 6-311G(*d*) basis has a much tighter *d* function (e.g., 1.75 for F) than 6-31G(*d*) (e.g., 0.8 for F). This causes a problem in G3(complete) because the QCISD(T)/G3Large calculation uses a 2*df* polarization set[3] having a small *d* exponent (e.g., $1.75/2=0.875$ for F) that is not diffuse enough for electron affinities (EAs) of species such as F and OH. This is not a problem in the G3 theory because it includes a MP4/6-31G(2*df*,*p*) calculation that has a more diffuse *d*-function exponent (e.g., $0.8/2=0.4$ for F). The first nonadditivity problem is corrected in the G4 theory by the G3LargeXP basis set, which has an expanded polarization set of 3*df* on the first row and of 4*d*2*f* on the second row. Use of the expanded polarization set results in a dramatic improvement for electron affinities with the average absolute deviation of G4(complete) being 0.70 kcal/mol compared to 1.45 kcal/mol for G3(complete) for the G2 test set. The second source of error, i.e., nonadditivity due to the separation of *d*-polarization and diffuse functions in the higher correlation calculations, is still present in the G4 theory. As a result, the G4(complete) electron affinities of F, O, OH, and NH_2 differ by more than 1 kcal/mol with the G4 theory, but G4(complete) is now in better agreement with experiment than the G4 theory in these cases.

The average absolute deviation of the G4 theory (0.72 kcal/mol) is still smaller than that of the G4(complete) theory (0.79 kcal/mol) for the G2 test set. The poorer agreement with experiment for the complete method is largely due to the problem with the core correlation calculation for the seven Si and P species. If these seven species are not included in the assessment, the average absolute deviations of the two methods are about the same.

D. Analysis of problem energies

While the G4 theory is significantly improved relative to the G3 theory on the G3/05 test set as discussed above, there are still some problem energies. Of the 454 energies only 35 have errors greater than 2 kcal/mol. These are listed in Table VII. In this section we discuss reasons for the larger errors in cases where there exists an apparent explanation.

1. Enthalpies of formation of neutrals

Eighteen of the 270 enthalpies of formation (or atomization energies in the case of third-row species) in the G3/05 test deviate from experiment by more than 2 kcal/mol. Three

TABLE VII. Outliers in the G3/05 test set that are greater than 2 kcal/mol. (Units are kcal/mol. Experimental values are from Refs. 16–18).

Property	Species	Expt.	Expt.- G4 (kcal/mol)	Property	Species	Expt.	Expt.- G4 (kcal/mol)
ΔH_f(298 K)	BF_3	−271.4	−2.8	IP	CH_4	291.0	−2.4
	COS	−33.1	2.5		BF_3	358.8	−2.4
	CS_2	28.0	3.0		BCl_3	267.5	−2.6
	C_2F_4	−154.4	3.2		B_2F_4	278.3	8.9
	C_2Cl_4	−3.0	3.1		see-C_3H_7	170.0	−2.9
	$CH_2{=}CHCl$ (vinyl chloride)	8.9	3.6		CN	313.6	−3.7
	$C_4H_4N_2$ (pyrimidine)	46.8	2.5		CH_3F	287.6	−2.6
	$CH_3{-}C({=}O){-}CCH$	15.6	−2.6		B_2H_4	223.7	2.3
	$Si(CH_3)_4$ (tetramethylsilane)	−55.7	−3.4		Si_2H_5	175.3	−2.5
	PF_5	−38.1	−2.3		NaBr	191.6	−4.7
	$POCl_3$	−133.8	−2.4	EA	F	78.4	−2.3
	Cl_2O_2S	−84.8	−2.2		CH_3	1.8	2.2
	PCl_3	−69.0	−2.8		C_2	75.4	2.2
	AlF	−63.5	2.3		CH_2NC	24.4	−2.3
	Al_2Cl_6	−309.7	5.2		HS	54.4	2.3
	$ClFO_3$	−5.1	−4.2	HB	H_2O dimer, ΔH_f	−3.6	−2.3
D_0	KF	117.6	2.3		HF dimer, D_0	−2.97	−2.6
	GeH_4	270.5	−2.5				

species (C_2F_4, C_2Cl_4, and CH_2CHCl) may have problems with the experimental values. This has been discussed elsewhere.[19,43] Of the remaining 15 energies, 12 are nonhydrogens. Only three of these have errors greater than 3 kcal/mol (CS_2, Al_2Cl_6, and $ClFO_3$). The difficulties with these nonhydrogens are probably due to the slow convergence of correlation energy with basis set and errors in core correlation effects that cannot be accounted for adequately by the current higher level correction.[44] In addition, relativistic effects are not included in the G4 theory, and for some nonhydrogen systems these are probably not accounted for by the higher level correction. The remaining three energies in this group are for substituted hydrocarbons [$Si(CH_3)_4$, $C_4H_4N_2$, CH_3COC_2H].

2. Ionization potentials

Ten of the 105 ionization potentials from the G4 theory deviate by more than 2 kcal/mol from experiment. The IPs of CH_4, BF_3, and BCl_3 are too large because the B3LYP method fails for the Jahn-Teller distortions in the ions, as discussed previously.[45] These errors are corrected when MP2(full)/6-31G* geometries are used, as shown in Table VIII. The large error for the IP of B_2F_4 is probably due to errors in the experimental IP, as discussed previously.[19] The error for CN is probably related to the fact that CN^+ is isoelectronic with C_2, which is known to be a challenging molecule. The large error (−4.9 kcal/mol) in the IP of NaBr is similar to the error found for the G3 theory.[32] In that case we investigated the need to include the Na 2*s* and 2*p* electrons in the correlation treatment and found that this did not improve the results. In addition, the calculation of this IP with G4 (complete) gives about the same result. We note that the error is much less at the G2 level of theory (−2.4 kcal/mol)[33] so the large error may be due to a basis set effect. The reasons for the >2 kcal/mol errors in the remaining four IPs in Table VI (C_3H_7, Si_2H_5, CH_3F, and B_2H_4) are unclear.

3. Electron affinities

Five of the 63 electron affinities from the G4 theory deviate from experiment by more than 2 kcal/mol. The F electron affinity is too large due to the additivity problem of the basis sets in the G4 theory, as discussed in Sec. III C. This is corrected when done without additivity approximations. The error for C_2 is likely due to the highly correlated nature of the neutral C_2 molecule that is known to be a challenge to describe accurately. The CH_3 radical weakly binds an electron (EA=1.8 kcal/mol), so it is not surprising that it is underbound by 2.2 kcal/mol due to the lack of enough

TABLE VIII. Effect of use of improved geometry on H-bonded dimers and ionization potentials involving cations with Jahn-Teller distortions.

	Expt. Theor. (kcal/mol)	
	G4	G4 with improved geometry[a]
H_2O dimer, D_0	−2.28	−0.48
CH_3OH dimer, D_0	0.02	0.46
$CH_3(CO)CH_3$ dimer, D_0	−1.36	−1.33
HCl dimer, D_0	−0.25	−0.24
CH_3COOH dimer, D_0	0.19	0.15
HF dimer, D_0	−2.61	−0.15
CH_4, IP	−2.40	−0.93
BF_3, IP	−2.38	0.28
BCl_3, IP	−2.57	−0.92
AAD (kcal/mol)	1.56	0.55

[a] For hydrogen bonded complexes the improved optimized geometries are from B3LYP/6-31+G(2*df*,*p*) calculations, and for the Jahn-Teller species they are from MP2(full)/6-31G* calculations.

TABLE IX. Basis sets used in single point HF energy calculations for the G4 theory.

	aug-cc-pVQZ		aug-cc-pV5Z	
Atoms	Literature[a]	G4 modified[b]	Literature[a]	G4 modified[b]
H,He	$4s3p2d1f$+diffuse $spdf$	$4s2pd$	$5s4p3d2f1g$+diffuse $spdfg$	$5s3p2d$
Li–Ne	$5s4p3d2f1g$+diffuse $spdfg$	$5s4p3d2f1g$+diffuse sp	$6s5p4d3f2g1h$+diffuse $spdfgh$	$6s5p4d3f2g1h$+diffuse sp
Na, Mg	$6s5p3d2f1g$	$6s5p3d2f1g$	$7s6p4d3f2g1h$	$7s6p4d3f2g1h$
Al–Ar	$6s5p3d2f1g$+diffuse $spdfg$	$6s5p3d2f1g$+diffuse sp	$7s6p4d3f2g1h$+diffuse $spdfgh$	$7s6p4d3f2g1h$+diffuse sp
K,Ca	$7s6p4d2f1g$	$7s6p4d2f1g$	$8s7p5d3f2g1h$	$8s7p5d3f2g1h$
Ga–Kr	$7s6p4d2f1g$+diffuse $spdfg$	$7s6p4d2f1g$+diffuse sp	$8s7p5d3f2g1h$+diffuse $spdfgh$	$8s7p5d3f2g1h$+diffuse sp

[a] See Appendix A for references.

[b] Modified basis set used in the G4 theory (see Appendix A).

diffuse functions. This was also true for the G3 theory. The reason for the errors of greater than 2 kcal/mol for the remaining two systems (HS and CH_2NC) is unclear, but they are only slightly more than 2 kcal/mol.

4. Hydrogen bonded dimers

As discussed earlier, two hydrogen bonded dimers, the water dimer and the hydrogen fluoride dimer, have errors of >2 kcal/mol due to the need to include diffuse functions in the basis set used for geometry optimization,[20] which is not included in the G4 theory. Results for the G4 energies for all hydrogen bonded dimers when a diffuse function is included in the basis set (6-31+G($2df,p$)) are given in Table VIII. The errors in the water and hydrogen fluoride dimers are reduced to under 1 kcal/mol and the average absolute deviation is reduced from 1.12 kcal/mol to 0.47 kcal/mol.

We note that an analysis of the overall absolute deviation for the G3/05 test set using improved geometries in the case of the hydrogen bonded complexes and Jahn-Teller systems gives a value of 0.81 kcal/mol. In addition, if the four energies with probable problems with the experimental data (enthalpies of C_2F_4, C_2Cl_4, and CH_2CHCl and ionization potential of B_2F_4) are excluded, the average absolute deviation is reduced to 0.78 kcal/mol.

IV. CONCLUSIONS

In this paper we have presented Gaussian-4 theory (G4 theory) for the calculation of energies of molecular species containing first-row (Li–F), second-row (Na–Cl), and third-row main group elements. The G4 theory modifies the Gaussian-3 (G3) theory in five ways, including an estimate of the HF energy limit, an expanded polarization set for the large basis set calculation, use of CCSD(T) energies, use of DFT geometries and zero-point energies, and two added higher level correction parameters. The overall average absolute deviation for the 454 energies in this test set is 0.83 kcal/mol, a significant improvement over the G3 theory. The largest improvement occurs for enthalpies of formation of nonhydrogen species, which are reduced to 1.13 kcal/mol. With the exception of hydrogen-bonded complexes the other types of energies in the G3/05 test set are also improved in the G4 theory. The inclusion of the HF limit energy in the G4 method contributes to the largest improvement over the G3 theory. The addition of the two new higher level correction parameters, one for paired electrons in open shell species and the other for species with only one pair of valence electrons, also contributes to a significant improvement. Finally, the expanded d-polarization set corrects a problem with nonadditivity that is present in the G3 theory for some anions.

ACKNOWLEDGMENTS

The authors acknowledge grants of computer time at the National Energy Research Supercomputer Center and the Laboratory Computing Resources Center at Argonne National Laboratory. They thank Dr. K. Peterson for providing them with the potassium basis set. This work was supported by the U.S. Department of Energy, Division of Materials Sciences under Contract No. DE-AC-02-06CH11357 and by a NSF grant, CHE-0616737, at Indiana University.

APPENDIX A: BASIS SETS FOR HF LIMIT EXTRAPOLATION

These basis sets are based on aug-cc-pVQZ or aug-cc-pV5Z basis sets[26–28] available from the EMSL database (http://www.emsl.pnl.gov/forms/basisform.html) with the exception of potassium.[46] In most cases the basis sets were modified to save CPU time. The basis set compositions are summarized in Table IX. The G4 quadruple zeta basis sets for H and He use the s part of cc-pVQZ combined with the $2pd$ polarization functions from the smaller cc-pVTZ basis set. They have no diffuse functions. The G4 quadruple zeta basis sets for the other atoms use the standard aug-cc-pVQZ basis set, but include only s and p diffuse functions (no d, f, or g diffuse functions). The only exceptions to this are Na, Mg, K, and Ca, which have no diffuse functions added. The 5Z basis set for H and He uses the s part of cc-pV5Z combined with $3p2d$ polarization functions taken from cc-pVQZ. They have no diffuse functions. The other atoms use the aug-cc-pV5Z basis set, but include only s and p diffuse functions (no d, f, g, or h diffuse functions). The only exceptions to this are Na, Mg, K, and Ca, which have no diffuse functions added. The basis sets are included in the supplementary information[39] and are available on the web.[38]

APPENDIX B: EXPANDED *d*-POLARIZATION BASIS

In the expanded d-polarization set in the G3LargeXP basis set, the $2df$ polarization set in G3Large on the first row

TABLE X. Polarization exponents used in the G3LargeXP basis set.

Atom	$4d$	d	$d/4$	4th d	MP2(full)/G3LargeXP
LI	0.8	0.2	0.05		−7.464 22
Be	1.02	0.255	0.0638		−14.635 86
B	1.604	0.401	0.1003		−24.614 20
C	2.504	0.626	0.1565		−37.798 67
N	3.652	0.913	0.2283		−54.538 46
O	5.168	1.292	0.323		−74.998 55
F	7.0	1.75	0.4375		−99.650 43
Ne	9.126	2.304	0.576		−128.843 11
Na	0.7	0.175	0.0438		−162.102 97
Mg	0.7	0.175	0.0438		−199.890 69
Al	2.16	0.54	0.135	0.39	−242.189 16
Si	3.04	0.76	0.19	0.47	−289.198 96
P	3.68	0.92	0.23	0.61	−341.093 58
S	4.08	1.02	0.26	0.59	−397.926 05
Cl	4.8	1.2	0.3	0.72	−459.951 86
Ar	5.6	1.4	0.35	0.85	−527.332 84

is replaced by a $3df$ with a $4:1:1/4$ geometrical progression, and the $3d2f$ polarization functions on the second row (Al–Cl) are replaced by $4d2f$. We investigated the need to reoptimize the first-row exponents and found that it was not necessary, but that optimization of the $4d$ exponents was required. This was done by optimizing the $3d$ set and a fourth d exponent sequentially until converged. The resulting d exponents for both first and second rows are listed in Table X. The species used in the Al through Kr optimizations were Al, AlF_3, Si, SiH_4, SiF_4, P, PH_3, PF_3, S, SO_2, SH_2, Cl, HCl, CCl_4, and Ar. Values for a given atom type were averaged. Note that the fourth exponent falls approximately midway between the two smaller exponents of the $3d$ set. A similar type of exponent spread has been found recently by Dunning and co-workers[47,48] to be necessary for an accurate description of the second row. The $3d2f$ polarization set is still used for Na and Mg in the second row. The d functions in the $3d2f$ polarization set were not changed for the third row (K, Ca, Ga–Kr) from their values in G3Large because of the filled d shell. The f exponents, diffuse exponents, and tight polarization functions in the G3LargeXP basis set remain the same as in G3Large.

[1] J. A. Pople, M. Head-Gordon, D. J. Fox, K. Raghavachari, and L. A. Curtiss, J. Chem. Phys. **90**, 5622 (1989); L. A. Curtiss, C. Jones, G. W. Trucks, K. Raghavachari, and J. A. Pople, *ibid.* **93**, 2537 (1990).

[2] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991).

[3] L. A. Curtiss, K. Raghavachari, P. C. Redfern, V. Rassolov, and J. A. Pople, J. Chem. Phys. **109**, 7764 (1998).

[4] T. H. Dunning, J. Phys. Chem. A **104**, 9062 (2000).

[5] D. Feller and D. A. Dixon, J. Chem. Phys. **115**, 3484 (2001).

[6] D. Feller and D. A. Dixon, J. Phys. Chem. A **104**, 3048 (2000).

[7] J. M. L. Martin and G. De Oliveira, J. Chem. Phys. **111**, 1843 (1999).

[8] A. Boese, M. Oren, O. Atasoylu, J. M. L. Martin, M. Kállay, and J. Gauss, J. Chem. Phys. **120**, 4129 (2004).

[9] D. Feller and K. A. Peterson, J. Chem. Phys. **108**, 154 (1998).

[10] D. Feller and K. A. Peterson, J. Chem. Phys. **110**, 8384 (1999).

[11] A. Tajti, P. G. Szalay, A. G. Csaszar, M. Kallay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vazquez, and J. F. Stanton, J. Chem. Phys. **121**, 11599 (2004).

[12] N. J. DeYonker, T. R. Cundari, and A. K. Wilson, J. Chem. Phys. **124**, 114104 (2006).

[13] N. J. DeYonker, T. Grimes, S. Yockel, A. Dinescu, B. Mintz, T. R. Cundari, and A. K. Wilson, J. Chem. Phys. **125**, 104111 (2006).

[14] J. A. Montgomery, Jr.,M. J. Frisch, J. W. Ochterski, and G. A. Petersson, J. Chem. Phys. **110**, 2822 (1999).

[15] J. A. Montgomery, Jr., M. J. Frisch, J. W. Ochterski, and G. A. Petersson, J. Chem. Phys. **112**, 6532 (2000).

[16] P. L. Fast, M. L. Sanchez, and D. G. Truhlar, Chem. Phys. Lett. **306**, 407 (1999).

[17] B. J. Lynch and D. G. Truhlar, J. Phys. Chem. A **107**, 3898 (2003).

[18] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **106**, 1063 (1997); L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, *ibid.* **109**, 42 (1998).

[19] L. A. Curtiss, K. Raghavachari, P. C. Redfern, and J. A. Pople, J. Chem. Phys. **112**, 7374 (2000).

[20] L. A. Curtiss, P. C. Redfern, and K. Raghavachari, J. Chem. Phys. **123**, 124107 (2005).

[21] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **114**, 108 (2001).

[22] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[23] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, Chem. Phys. Lett. **270**, 419 (1999).

[24] A. Halkier, T. Helgaker, P. Jorgensen, W. Klopper, and J. Olsen, Chem. Phys. Lett. **302**, 437 (1999).

[25] D. Feller, J. Chem. Phys. **98**, 7059 (1993).

[26] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[27] D. E. Woon and T. H. Dunning, Jr., J. Chem. Phys. **98**, 1358 (1993).

[28] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning, Jr., J. Chem. Phys. **110**, 7667 (1999).

[29] L. A. Curtiss, K. Raghavachari, P. C. Redfern, A. G. Baboul, and J. A. Pople, Chem. Phys. Lett. **314**, 101 (1999).

[30] L. A. Curtiss, P. C. Redfern, K. Raghavachari, and J. A. Pople, Chem. Phys. Lett. **359**, 390 (2002).

[31] V. A. Rassolov, J. A. Pople, P. C. Redfern, and L. A. Curtiss, Chem. Phys. Lett. **350**, 573 (2001).

[32] L. A. Curtiss, P. C. Redfern, V. Rassolov, G. Kedziora, and J. A. Pople, J. Chem. Phys. **114**, 9287 (2001).

[33] L. A. Curtiss, M. P. McGrath, J.-P. Blaudeau, N. E. Davis, and R. Binning, J. Chem. Phys. **103**, 6104 (1995).

[34] J.-P. Blaudeau, M. P. McGrath, L. A. Curtiss, and L. Radom, J. Chem. Phys. **107**, 5016 (1997).

[35] C. Moore, Natl. Bur. Stand. (U.S.) Circ. No. 467 (U.S. GPO, Washington, D.C., 1952).

[36] J.-P. Blaudeau (private communication).

[37] M. J. Frisch, G. W. Trucks, H. B. Schlegel *et al.*, GAUSSIAN 03, Gaussian, Inc., Wallingford, CT, 2004.

[38] chemistry.anl.gov/compmat/comptherm.htm

[39] See EPAPS document no. E-JCPSA6-126-310707 for supplementary information including G4 total energies, both E_e and E_o values, for all species in the G3/05 test set, the G4 deviations from experiment, the

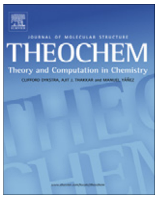
G3LargeXP basis set, and the modified aug-cc-pVQZ or aug-cc-pV5Z basis sets. This document can be reached via a direct link in the online article's HTML reference section or via the EPAPS homepage (http://www.aip.org/pubservs/epaps.html).
[40] J. M. L. Martin, J. Chem. Phys. **108**, 2791 (1998).
[41] L. A. Curtiss, J. E. Carpenter, K. Raghavachari, and J. A. Pople, J. Chem. Phys. **96**, 9030 (1992).
[42] J. M. L. Martin, A. Sunderman, P. L. Fast, and D. G. Truhlar, J. Chem. Phys. **113**, 1348 (2000).
[43] L. Pollack, T. L. Windus, W. A. Windus, and D. A. Dixon, J. Phys. Chem. A **109**, 6934 (2005).
[44] K. E. Gutowski and D. A. Dixon, J. Phys. Chem. A **110**, 12044 (2006).
[45] A. G. Baboul, L. A. Curtiss, P. C. Redfern, and K. Raghavachari, J. Chem. Phys. **110**, 7650 (1999).
[46] K. Peterson (private communication).
[47] T. H. Dunning, Jr., K. A. Peterson, and A. K. Wilson, J. Chem. Phys. **114**, 9244 (2001).
[48] A. K. Wilson and T. H. Dunning, Jr., J. Phys. Chem. A **108**, 3129 (2004).

Journal of Molecular Structure: THEOCHEM 953 (2010) 47–48

Contents lists available at ScienceDirect

Journal of Molecular Structure: THEOCHEM

journal homepage: www.elsevier.com/locate/theochem



Performance of Gaussian-3 and Gaussian-4 level theoretical methods in estimating gas phase enthalpies of formation for representative C_1 and C_2 chlorofluorocarbons and hydrochlorofluorocarbons

Sierra Rayne [a,*], Kaya Forest [b]

[a] *Ecologica Research, Penticton, British Columbia, Canada V2A 8J3*
[b] *Department of Chemistry, Okanagan College, Penticton, British Columbia, Canada V2A 8E1*

ARTICLE INFO

Article history:
Received 2 April 2010
Received in revised form 25 April 2010
Accepted 26 April 2010

Keywords:
Chlorofluorocarbons
Hydrochlorofluorocarbons
Enthalpy of formation
Ab initio post-Hartree–Fock composite methods
Atomization approach
G3/G4

ABSTRACT

The Gaussian-3 (G3) and Gaussian-4 (G4) ab initio post-Hartree–Fock composite methods were employed to estimate the gas phase standard state (298.15 K, 1 atm) enthalpies of formation ($\Delta_f H^\circ_{(g)}$) for 38 representative $C_1(n = 15)$ and $C_2(n = 23)$ chlorofluorocarbons (CFCs) and hydrochlorofluorocarbons (HCFCs). Using the atomization approach, good agreement was found between experimental $\Delta_f H^\circ_{(g)}$ and corresponding G3 and G4 estimates. Where significant differences between G3/G4 and experimental $\Delta_f H^\circ_{(g)}$ exist, the errors may be due to problems with the experimental data rather than deficiencies in the theoretical methods.

Chlorofluorocarbons (CFCs) and hydrochlorofluorocarbons (HCFCs) are refrigerants, propellants, solvents, and coolants [1,2]. Their thermodynamic properties are of interest [3–5] for use in environmental modeling efforts, the design of new industrial materials, and for benchmarking theoretical methods on small halogenated molecules.

Application of the atomization enthalpy approach to calculations at the G3 [6] and G4 [7] levels on 38 representative $C_1(n = 15)$ and $C_2(n = 23)$ CFCs and HCFCs using Gaussian 09 [8] yields the estimated gas phase standard state (298.15 K, 1 atm) enthalpies of formation ($\Delta_f H^\circ_{(g)}$) given in Table 1. Full G09 archive entries are provided in the Supplementary Materials with optimized geometries, energies at each stage of the G3/G4 calculation process, and frequency information for thermochemical and spectroscopic analysis.

A review of experimental $\Delta_f H^\circ_{(g)}$ data from Refs. [5,9,10] reveals there is substantial uncertainty for a number of these compounds, often ranging up to 15–20 kJ/mol (and sometimes higher) for a single member. For example, in the NIST database, CH_2ClCCl_3 has two $\Delta_f H^\circ_{(g)}$ values of -135.6 and 152.3 ± 2.4 kJ/mol. Similarly, the $\Delta_f H^\circ_{(g)}$ of -100 ± 20 kJ/mol for CCl_4 reported in the NIST database is an average of six individual literature datapoints that range over 31 kJ/mol [from -94 ± 2 to -125.0 ± 4.6 kJ/mol]. $CFCl_2CF_2Cl$ has a $\Delta_f H^\circ_{(g)}$ of -726.8 ± 4.3 kJ/mol in the NIST database; the corresponding experimental $\Delta_f H^\circ_{(g)}$ given in Ref. [5] is -706.3 kJ/mol. This variability is not surprising, given the potential lack of sample purities and difficulty in ensuring complete combustion where applicable. In short, the CFCs and HCFCs effectively lack a coherent, consistent, and high accuracy experimental thermochemistry database.

Clearly some experimental $\Delta_f H^\circ_{(g)}$ values given in Table 1 are still unsatisfactory due to their absence in the NIST database but presence in other compendia, and large deviation from the G4 calculations (which should achieve consistent near chemical accuracy for these small compounds). For example, we doubt the $\Delta_f H^\circ_{(g)}$ of fluoroethane is -263.2 kJ/mol; it is likely closer to -274.8 kJ/mol based on the G4 results. Otherwise, if one assumes that the error between G3/G4 estimates and experimental values should increase with increasing numbers of C–F and C–Cl bonds on a given hydrocarbon framework (a prerequisite for reliable application of bond additive correction (BAC) methods, as has been attempted for these compounds), there appears no rational way to explain the low accuracy for fluoroethane, but high accuracy obtained for 1,1-difluoroethane (the G4 $\Delta_f H^\circ_{(g)}$ differs from the experimental value by only 0.4 kJ/mol), reasonable agreement for CHF_2CH_2F (deviation

---

* Corresponding author. Tel.: +1 250 487 0166.
E-mail address: rayne.sierra@gmail.com (S. Rayne).

0166-1280/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2010.04.031

Table 1
Comparison between experimental and G3/G4 calculated $\Delta_f H^{\circ}_{(g)}$ for 38 CFCs and HCFCs. Values are in kJ/mol and deviations from the experimental dataset at each level of theory are given in parentheses.

	Expt.	G3	G4
CH_4	−74.5	−75.6 (−1.1)	−74.4 (0.1)
CH_3F	−234.3	−237.2 (−2.9)	−235.4 (−1.1)
CH_3Cl	−82.0	−81.1 (0.8)	−81.1 (0.8)
CH_2F_2	−450.7	−452.5 (−1.8)	−450.3 (0.4)
CH_2FCl	−261.9	−264.2 (−2.3)	−262.2 (−0.2)
CH_2Cl_2	−95.1	−93.0 (2.1)	−92.5 (2.6)
CHF_3	−695.4	−698.1 (−2.7)	−695.1 (0.3)
CHF_2Cl	−481.6	−484.3 (−2.7)	−480.4 (1.2)
$CHFCl_2$	−283.3	−285.9 (−2.7)	−282.2 (1.1)
$CHCl_3$	−102.7	−102.5 (0.2)	−100.4 (2.3)
CF_4	−933.2	−935.4 (−2.2)	−931.7 (1.5)
CF_3Cl	−707.9	−713.0 (−5.1)	−707.6 (0.3)
CF_2Cl_2	−491.6	−499.2 (−7.5)	−492.2 (−0.6)
$CFCl_3$	−288.7	−295.1 (−6.4)	−288.3 (0.4)
CCl_4	−100.0	−102.5 (−2.5)	−98.1 (1.9)
C_2H_6	−83.8	−84.6 (−0.8)	−82.8 (1.0)
C_2H_5F	−263.2	−273.8 (−10.6)	−270.6 (−7.4)
C_2H_5Cl	−109.0	−111.1 (−2.1)	−109.7 (−0.7)
CH_3CHFCl	−313.4	−313.0 (0.4)	−308.6 (4.8)
CH_2FCH_2F	−433.9	−450.9 (−17.0)	−446.7 (−12.8)
CH_3CHF_2	−500.8	−506.2 (−5.4)	−501.2 (−0.4)
CH_2ClCH_2Cl	−132.0	−132.6 (−0.6)	−131.3 (0.7)
CH_3CHCl_2	−132.5	−135.4 (−2.9)	−133.0 (−0.5)
CH_3CF_3	−749.0	−757.2 (−8.2)	−750.3 (−1.3)
CHF_2CH_2F	−664.8	−672.3 (−7.5)	−666.1 (−1.3)
CH_3CCl_3	−145.0	−150.1 (−5.1)	−146.1 (−1.1)
$CHCl_2CH_2Cl$	−148.0	−149.9 (−1.9)	−147.1 (0.9)
CH_2FCF_3	−895.8	−914.5 (−18.7)	−906.4 (−10.6)
CHF_2CHF_2	−877.8	−887.4 (−9.6)	−879.3 (−1.5)
$CHCl_2CHCl_2$	−156.7	−161.9 (−5.2)	−157.1 (−0.4)
CH_2ClCCl_3	−152.3	−160.4 (−8.1)	−155.3 (−3.0)
CHF_2CF_3	−1104.6	−1121.2 (−16.6)	−1111.5 (−6.9)
C_2HCl_5	−155.9	−169.0 (−13.1)	−161.7 (−5.8)
C_2F_6	−1342.7	−1353.7 (−11.0)	−1342.1 (0.6)
CF_2ClCF_3	−1118.8	−1145.1 (−26.3)	−1131.5 (−12.7)
CF_2ClCF_2Cl	−930.0	−937.3 (−7.3)	−921.9 (8.1)
$CFCl_2CF_2Cl$	−726.8	−739.9 (−13.1)	−724.3 (2.5)
C_2Cl_6	−148.2	−172.2 (−24.0)	−161.6 (−13.4)

of 1.3 kJ/mol between G4 and experimental), and high accuracy obtained for hexafluoroethane (0.6 kJ/mol difference between the G4 and experimental).

As noted by others in recent works on a wide range of hydrocarbons (including hydrofluorocarbons), [4,11,12] where there is large and otherwise unusual disagreement between Gaussian-3 (and now Gaussian-4) $\Delta_f H^{\circ}_{(g)}$ and experimental data, one may quite reasonably distrust the experimental data to an equal – and possibly greater – degree than the high level calculations (assuming the lowest energy conformation is obtained – typically not a major issue for C_1 and C_2 derivatives). Additional experimental $\Delta_f H^{\circ}_{(g)}$ data of major concern include those for CH_2FCH_2F, CH_2FCF_3, CHF_2CF_3, C_2HCl_5, CF_2ClCF_3, CF_2ClCF_2Cl, and C_2Cl_6.

Using our experimental $\Delta_f H^{\circ}_{(g)}$ values, we obtain MAD/RMSD/MSD of 2.9/3.5/−2.5 kJ/mol and 9.4/11.8/−9.3 kJ/mol at the G3 level for the C_1 and C_2 compounds, respectively, and an overall G3 MAD/RMSD/MSD of 6.8/9.4/−6.6 kJ/mol. At the G4 level, our MAD/RMSD/MSD drop to only 1.0/1.3/0.7 kJ/mol (C_1), 4.3/6.1/−2.7 kJ/mol (C_2), and 3.0/4.8/−1.3 kJ/mol (overall). The acquisition of more recent and reliable experimental $\Delta_f H^{\circ}_{(g)}$ values for these compounds appears likely to result in even higher accuracy for the G3/G4 methods, given how many of the current $\Delta_f H^{\circ}_{(g)}$ values in the experimental database (particularly for the C_2 CFCs and HCFCs) vary widely and make it effectively impossible to reliably benchmark high level theoretical methods on this class of compounds. Thus, the G4 $\Delta_f H^{\circ}_{(g)}$ reported herein may be useful surrogates for presently uncertain experimental data until discrepancies are satisfactorily resolved.

Acknowledgements

This work was made possible by the facilities of the Western Canada Research Grid (WestGrid: www.westgrid.ca; project 100185), the Shared Hierarchical Academic Research Computing Network (SHARCNET: www.sharcnet.ca; project sn4612), and Compute/Calcul Canada.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.theochem.2010.04.031.

References

[1] M. Khalil, Ann. Rev. Energy Environ. 24 (1999) 645.
[2] D. Good, J. Francisco, Chem. Rev. 103 (2003) 4999.
[3] M. Zachariah, P. Westmoreland, D. Burgess, W. Tsang, C. Melius, J. Phys. Chem. 100 (1996) 8737.
[4] N. Haworth, M. Smith, G. Bacskay, J. Mackie, J. Phys. Chem. A 104 (2000) 7600.
[5] J. Zhang, W. Zhou, Y. Li, H. Gao, Z. Zhou, J. Fluor. Chem. 131 (2010) 606–611.
[6] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, J. Chem. Phys. 109 (1998) 7764.
[7] L.A. Curtiss, P.C. Redfern, K. Raghavachari, J. Chem. Phys. 126 (2007) 84108.
[8] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, et al., Gaussian 09, Revision A.02, Gaussian, Inc., Wallingford, CT, USA, 2009.
[9] D. Lide, CRC Handbook of Chemistry and Physics, Taylor and Francis, Boca Raton, FL, USA, 2009.
[10] H. Afeefy, J. Liebman, S. Stein, Neutral Thermochemical Data: NIST Chemistry WebBook, NIST Standard Reference Database Number 69, National Institute of Standards and Technology, Gaithersburg, MD, USA, 2010.
[11] D. Bond, J. Org. Chem. 72 (2007) 7313.
[12] E. Taskinen, J. Phys. Org. Chem. 22 (2009) 632.

JOURNAL OF CHEMICAL PHYSICS VOLUME 121, NUMBER 23 15 DECEMBER 2004

HEAT: High accuracy extrapolated *ab initio* thermochemistry

Attila Tajti
Department of Theoretical Chemistry, Eötvös Loránd University, P.O. Box 32, H-1518 Budapest, Hungary

Péter G. Szalay
Department of Theoretical Chemistry, Eötvös Loránd University, P.O. Box 32, H-1518 Budapest, Hungary and Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712

Attila G. Császár
Department of Theoretical Chemistry, Eötvös Loránd University, P.O. Box 32, H-1518 Budapest, Hungary

Mihály Kállay and Jürgen Gauss
Institut für Physikalische Chemie, Universität Mainz, D-55099 Mainz, Germany

Edward F. Valeev
School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, Georgia 30332-0400

Bradley A. Flowers, Juana Vázquez, and John F. Stanton
Institute for Theoretical Chemistry, Department of Chemistry and Biochemistry, The University of Texas at Austin, Austin, Texas 78712

(Received 9 August 2004; accepted 9 September 2004)

A theoretical model chemistry designed to achieve high accuracy for enthalpies of formation of atoms and small molecules is described. This approach is entirely independent of experimental data and contains no empirical scaling factors, and includes a treatment of electron correlation up to the full coupled-cluster singles, doubles, triples and quadruples approach. Energies are further augmented by anharmonic zero-point vibrational energies, a scalar relativistic correction, first-order spin–orbit coupling, and the diagonal Born–Oppenheimer correction. The accuracy of the approach is assessed by several means. Enthalpies of formation (at 0 K) calculated for a test suite of 31 atoms and molecules via direct calculation of the corresponding elemental formation reactions are within 1 kJ mol^{-1} to experiment in all cases. Given the quite different bonding environments in the product and reactant sides of these reactions, the results strongly indicate that even greater accuracy may be expected in reactions that preserve (either exactly or approximately) the number and types of chemical bonds. © *2004 American Institute of Physics.* [DOI: 10.1063/1.1811608]

I. INTRODUCTION

At a time when the majority of practitioners have essentially abandoned the field in favor of density-functional theory, traditional quantum chemistry (comprising what have come to be known—somewhat pejoratively in some circles—as "wave function methods") has evolved so that methods are now available that are capable of determining molecular properties at a very high level of accuracy. Instead of providing rough estimates of quantities such as vibrational frequencies and structural parameters, routinely applicable modern techniques are capable of $\approx$10 cm^{-1} accuracy for *fundamental* vibrational frequencies[1,2] and $\approx$0.002–0.003 Å accuracy in equilibrium bond distances.[3] With the seemingly constant improvements made in computer hardware technology, both the level of accuracy and the scope of systems suitable for treatment at a given level of accuracy will continue to grow.

One area where extremely high accuracy is generally useful and has significant impact is in the determination of thermochemical parameters. Enthalpies of formation, heat capacities, and standard entropies of molecular species entirely determine their thermodynamic fate. Experimental methods based on calorimetry, kinetics, spectroscopy, and various ion cycles have been used for decades to determine these important quantities.[4,5] As a result of this intense area of research, relatively tight bounds ($<$10 kJ mol^{-1}) have been established for enthalpies of formation for many molecules that are stable enough to be studied easily in the laboratory.[6] However, in two fields that are strongly dependent on accurate thermochemical information—combustion and atmospheric chemistry—many, and perhaps a majority, of the most important compounds are transient species. Accordingly, error bars on the enthalpies of formation for these radicals and other reactive molecules tend to be significantly larger than those for simple closed-shell species. As an example, the enthalpy of formation for the hydroperoxy radical (HO_2) has only recently been determined to better than 1 kJ mol^{-1}.[7]

The extent to which modern high-level *ab initio* ("wave function") calculations can be competitive with experiment in the precise determination of thermodynamic parameters depends to a large extent on the size of the molecular species in question. For the smallest molecules (a dozen or fewer electrons), there is little question that theory can provide very accurate total electronic energies, irrespective of how "difficult" the molecule is. For molecules that can be easily studied experimentally, theory can offer competitive precision for thermochemical parameters; for transient species not amenable to experimental characterization, theory is arguably better. A case in point are very recent studies of CH and CH_2,[8,9] where accuracies estimated to be $<0.5\ \mathrm{kJ\,mol^{-1}}$ were achieved for the enthalpies of formation, as well as others of nearly comparable accuracy.[7,10–20]

Of course, it is not the total energies of molecules that are relevant for thermochemistry, but rather appropriately defined energy differences. Hence, it is not necessary for a method to obtain a given level of accuracy (say $1\ \mathrm{kJ\,mol^{-1}}$) in total energies in order to achieve the same level of accuracy for energy differences. Instead, one can benefit from error cancellation; deficiencies of the total energy calculations for the various species have many common sources. Total molecular energies, of course, are those relative to the separated atoms of the molecule in question, all completely ionized with the electrons at rest and the nuclei in their ground states. For example, the total electronic energy of carbon monoxide is $E(\mathrm{C}^{6+})+E(\mathrm{O}^{8+})+E(14e^-)-E(\mathrm{CO})$. However, it is a much simpler task to calculate the bond energy of CO, viz. $E[\mathrm{C}(^3\mathrm{P})]+E[\mathrm{O}(^3\mathrm{P})]-E[\mathrm{CO}(^1\Sigma)]$ since the core electrons—which make the largest contribution to the total energy due to strong nuclear attraction forces—of C and O are only slightly perturbed in the molecular environment. The more "similar" the A and B species involved in the energy difference $E(A)-E(B)$, the less demanding is the calculation needed to achieve a specified level of accuracy. This was realized long ago; so-called "isodesmic" reactions[21]—those in which the number and qualitative "types" of bonds in A and B are the same—are known to be those in which calculated energy differences tend toward the highest accuracy.[22]

The preceding paragraph, while elementary and straightforward, is important to one of the themes of this paper. Most (perhaps all) "theoretical model chemistries"[23–25] (those in which all species are treated at a consistent and well-defined level of approximation) use atomization energies as the basis for thermochemistry. In this approach, total molecular energies are calculated at some level of theory. Atomization energies, defined for the molecule $M\equiv A_aB_b\cdots Z_z$ as

$$AE(M)=aE(A)+bE(B)+\cdots+zE(Z)-E(M) \tag{1}$$

are then calculated using total atomic energies obtained at the same level of theory. The enthalpy of formation for M at 0 K is then given by Hess's law as

$$\Delta_f H^\circ = a\Delta_f H^\circ(A)+b\Delta_f H^\circ(B)+\cdots+z\Delta_f H^\circ(Z)-AE(M), \tag{2}$$

where the atomic enthalpies of formation are set to literature values. With the notable exceptions of carbon and fluorine,[12] enthalpies of formation are known quite precisely for atoms in the first two rows of the periodic table, so the inherent error in this approach is often localized almost entirely in the atomization energies. However, it is very difficult to calculate atomization energies. By definition, all bonds in the molecule M are destroyed in the reaction that is used as the basis for the calculated quantity; it is as far from an isodesmic reaction as possible. Another approach, rarely used in practice, would be to calculate the enthalpies of formation *directly*, meaning that the molecular energy of M and those of its constituent elements in their most stable form are used. For example, if M were water, the reaction considered would be

$$H_2+\tfrac{1}{2}O_2\rightarrow H_2O. \tag{3}$$

The enthalpy of formation for water is given, by definition, as the difference of ground state energies in the reactants and products in Eq. (3), and its calculation therefore does not make any use of experimental quantities. This method does have some serious limitations, however. For example, any organic molecule would require an (impossible) calculation of graphite; sulfur would necessitate calculations on a molecule with 128 electrons (S_8), chlorine would require *ab initio* calculations of the liquid and the enthalpy of vaporization, etc. Moreover, while chemical bonds can indeed be found on both sides of Eq. (3), it is clearly not an isodesmic reaction, or even nearly so. While it might be easier to calculate this energy difference accurately than that of the atomization reaction $H_2O\rightarrow 2H+O$, the benefits of doing so are marginal. The atomic enthalpies of formation for H and O are also known so precisely ($<0.01\ \mathrm{kJ\,mol^{-1}}$) that there is really no advantage in eliminating the use of this experimental information. Hence, due to simplicity, the well-defined nature of the process, and the generally straightforward problems posed by atoms to theoretical treatment, the atomization energy approach has been the method of choice for quantum chemical practitioners in estimating enthalpies of formation for molecules with various model chemistries.

One purpose of this work is to present a new theoretical model chemistry. Unlike previous efforts (the only possible exception being the recent definition of W3 theory by Martin and collaborators[25]), we propose a method that cannot be generally applied to all "small" (loosely defined here as those having five or fewer non-hydrogen atoms) molecules. Rather, it is our intent to define an approach that is essentially the best that can be done—with current computer technology—for very small molecules (four or fewer *total* atoms), and to assess the level of accuracy achieved. It is our belief that this is not an academic exercise with only a narrow practical benefit. First, there are still molecules of this size where accurate ($<1\ \mathrm{kJ\,mol^{-1}}$) enthalpies of formation are not available[6] [for example, NH, NH_2 and, until recently, OH (Ref. 13) and HO_2 (Ref. 7)]. Second, the accuracy that theory can achieve for small systems is of intrinsic interest, because it is useful to know how large a role is played by some usually neglected effects (coupling of core and valence correlation, relativistic corrections, the diagonal Born–Oppenheimer correction, anharmonic contributions to zero-point vibrational energies, spin–orbit coupling, etc.) in overall accuracy. Third, and most important, since the methods used in our work are "size extensive"[26,27] (meaning that the quality of the energy calculation is not degraded by the size of the molecule described within a given one-particle basis set), the accuracy achieved for the benchmark systems stud-

TABLE I. Contributions to the HEAT total energies for the 31 species studied in this work. All values are in atomic units. Conversion factor used 2625.4976 kJ mol^{-1}=1 E_h.

Species	E_{HF}^{∞}	$\Delta E_{\mathrm{CCSD(T)}}^{\infty}$	$\Delta E_{\mathrm{CCSDT}}$	$\Delta E_{\mathrm{CCSDTQ}}$	ΔE_{REL}	ΔE_{ZPE}	ΔE_{DBOC}	ΔE_{SO}	Total
N_2	−108.993 257	−0.549 274	0.000 507	−0.001 457	−0.058 658	0.005 379	0.003 982	0.000 000	−109.592 778
H_2	−1.133 661	−0.040 911	0.000 000	0.000 000	−0.000 010	0.009 892	0.000 460	0.000 000	−1.164 230
F_2	−198.774 570	−0.756 425	0.000 100	−0.001 536	−0.174 461	0.002 095	0.005 175	0.000 000	−199.699 622
O_2	−149.691 925	−0.635 217	0.000 112	−0.001 854	−0.104 607	0.003 641	0.004 711	−0.000 012	−150.425 151
C	−37.693 774	−0.151 041	−0.000 466	−0.000 030	−0.015 090	0.000 000	0.001 660	−0.000 144	−37.858 885
F	−99.416 800	−0.318 033	−0.000 199	−0.000 116	−0.087 268	0.000 000	0.002 591	−0.000 574	−99.820 399
H	−0.500 022	0.000 000	0.000 000	0.000 000	−0.000 007	0.000 000	0.000 272	0.000 000	−0.499 757
N	−54.404 657	−0.184 700	−0.000 344	−0.000 042	−0.029 435	0.000 000	0.002 007	0.000 000	−54.617 171
O	−74.819 232	−0.248 562	−0.000 338	−0.000 078	−0.052 459	0.000 000	0.002 366	−0.000 312	−75.118 615
CO	−112.790 997	−0.535 543	0.000 097	−0.000 951	−0.067 285	0.004 945	0.004 000	0.000 000	−113.385 734
C_2H_2	−76.855 684	−0.480 411	0.000 223	−0.000 911	−0.029 760	0.026 253	0.003 674	0.000 000	−77.336 616
CCH	−76.183 645	−0.428 923	−0.001 147	−0.000 928	−0.029 761	0.013 842	0.003 503	0.000 000	−76.627 059
CH_2	−38.941 051	−0.207 841	−0.000 458	−0.000 082	−0.014 865	0.017 167	0.002 161	0.000 000	−39.144 969
CH	−38.284 553	−0.194 296	−0.000 628	−0.000 076	−0.015 029	0.006 463	0.002 063	−0.000 067	−38.486 123
CH_3	−39.581 308	−0.254 300	−0.000 408	−0.000 120	−0.014 835	0.029 573	0.002 404	0.000 000	−39.818 994
CO_2	−187.725 719	−0.876 563	0.000 567	−0.001 753	−0.119 227	0.011 580	0.006 314	0.000 000	−188.704 801
H_2O_2	−150.852 930	−0.711 472	0.000 227	−0.001 280	−0.104 313	0.026 192	0.005 078	0.000 000	−151.638 498
H_2O	−76.067 761	−0.371 594	0.000 033	−0.000 453	−0.052 040	0.021 228	0.002 710	0.000 000	−76.467 877
HCO	−113.304 223	−0.553 349	−0.000 125	−0.000 943	−0.067 113	0.012 960	0.004 336	0.000 000	−113.908 457
HF	−100.071 316	−0.389 256	0.000 056	−0.000 392	−0.086 955	0.009 391	0.002 735	0.000 000	−100.535 737
HO_2	−150.253 106	−0.661 501	−0.000 447	−0.001 116	−0.104 469	0.014 110	0.004 996	0.000 000	−151.001 533
NO	−129.309 786	−0.589 993	−0.000 094	−0.001 291	−0.081 577	0.004 364	0.004 351	−0.000 275	−129.974 301
OH	−75.428 343	−0.310 311	−0.000 289	−0.000 259	−0.052 261	0.008 461	0.002 619	−0.000 297	−75.780 680
HNO	−129.850 244	−0.634 499	0.000 253	−0.001 469	−0.081 448	0.013 680	0.004 732	0.000 000	−130.548 995
CN	−92.242 929	−0.477 290	−0.002 005	−0.001 378	−0.044 284	0.004 858	0.003 629	0.000 000	−92.759 399
HCN	−92.915 916	−0.517 656	0.000 413	−0.001 230	−0.044 175	0.015 898	0.003 819	0.000 000	−93.458 847
CF	−137.239 487	−0.551 585	−0.000 435	−0.000 525	−0.102 084	0.003 002	0.004 239	−0.000 178	−137.887 053
NH_2	−55.592 445	−0.287 829	−0.000 368	−0.000 224	−0.029 194	0.018 882	0.002 564	0.000 000	−55.888 614
NH_3	−56.225 187	−0.339 348	−0.000 104	−0.000 316	−0.029 045	0.034 069	0.002 609	0.000 000	−56.557 322
NH	−54.986 522	−0.235 129	−0.000 454	−0.000 127	−0.029 323	0.007 412	0.002 353	−0.000 001	−55.241 791
OF	−174.211 642	−0.674 172	−0.000 938	−0.000 994	−0.139 570	0.002 426	0.004 951	−0.000 414	−175.020 353

ied here will be the same as that for larger molecules that will be amenable to the treatment in the future.

Finally, the systems studied here become the initial members of a database of compounds that can be used for any number of thermochemical studies. Specifically, we believe that the *best* way to calculate a molecular enthalpy of formation is to use approaches other than atomization enthalpy and direct elemental reaction strategies summarized above. In an ideal world, an isodesmic reaction can be designed in which all participants other than the target molecule M have enthalpies of formation that are known precisely from experiment. Then, if total electronic energies from the high-level theoretical model chemistry defined in this paper are available for all species, the reaction energy can presumably be obtained with negligible theoretical error. Adjustment of the reaction energy to the enthalpy of formation of M (by appropriate addition and subtraction of experimental enthalpies of formation for the other species), should then give $\Delta_f H^\circ$ for M with extraordinary precision. However, it is recognized that this will not usually be possible, either because of inability to design a truly isodesmic reaction (radicals can be difficult in this regard) or the lack of precise thermochemical knowledge about some of the species in an appropriate reaction. Then, alternative strategies can be followed, using reactions that are not isodesmic (but clearly superior to atomization schemes!) but involve at least an approximate conservation of bond types. Such an approach was recently used by us to determine the enthalpy of formation of HO_2 to an accuracy of ≈0.5 kJ mol^{-1}, nearly an order of magnitude better than the definition of ''chemical accuracy'' (1 kcal mol^{-1} or 4.184 kJ mol^{-1}), and by Schuurman *et al.* in a study of HNCO isomers.[10]

The next few sections define the theoretical model chemistry that we have named HEAT. This is an acronym for ''high accuracy extrapolated *ab initio* thermochemistry,'' which emphasizes two things—apart from energy extrapolation schemes, the approach involves no empirical scaling factors or adjustments, and that the principal area of application that we envision for HEAT will be in the area of thermochemistry. After defining the method, and discussing the various theoretical approaches used to determine the total energies that are the ''bottom line'' of HEAT for any atom or molecule, the approach will be applied to a test suite of atoms and molecules. While quantities such as atomization energies and enthalpies of formation calculated from them (using the approaches discussed above) as well as from elemental reactions (excepting carbon) will be presented and discussed, we emphasize that it is the total energies obtained by the HEAT protocol that are the most important numbers documented in this work. Hopefully, the total energies for 31 atoms and molecules found later in this paper (Table I)—and straightforwardly calculable (at least in principle) for other species—will form the initial members of a database that can be used by experimentalists and theorists alike to calculate enthalpies of formation using suitable *reaction-based* approaches.

II. DEFINITION OF HEAT MODEL CHEMISTRY

In order to determine standard enthalpies of formation at 0 K,[28] it is always necessary to know the ground state energy of the target species. Within the Born–Oppenheimer approximation, the ground state energy may be partitioned into electronic and vibrational contributions. The former is given by the lowest eigenvalue of the electronic Hamiltonian (electronic energy) at the equilibrium geometry; the latter by the lowest eigenvalue of the nuclear Hamiltonian containing a potential described by electronic energies as a function of position. In some cases, the lowest rotational state is prohibited by nuclear spin statistics, but we ignore this here.

The electronic part, as usually calculated, involves use of the nonrelativistic electronic Hamiltonian. The three simplest improvements upon this are to include ‘‘scalar’’ relativistic effects[29] using perturbation theory (which is believed to be entirely adequate for atoms in the first two rows of the periodic table[30]), splitting of the energy—and lowering of the ground state energy—by spin–orbit interactions, and to compute the so-called diagonal Born–Oppenheimer correction.[31–34] The latter is given by the expectation value of the nuclear kinetic energy operator over the electronic wave function (that which diagonalizes the electronic Hamiltonian), and is a first-order correction to the simple Born–Oppenheimer approximation that does not spoil the concept of a potential energy surface.[35]

In practice, all of the contributions above need to be calculated approximately. First, finite basis sets must be used. Even if the treatment of correlation was complete (full configuration interaction, or FCI), the resulting energies would be compromised by limitations of the one-particle basis set. It is nonetheless impossible to do FCI calculations for all but the smallest molecules—even then in necessarily small basis sets—so approximate measures for treating correlation are called for. In HEAT, as in all other model chemistries that we know of, size-extensive many-body methods are used in the treatment of electron correlation.

The total energy defined by the HEAT protocol may be expressed by a formula that contains eight terms

$$E_{\mathrm{HEAT}}=E_{\mathrm{HF}}^{\infty}+\Delta E_{\mathrm{CCSD(T)}}^{\infty}+\Delta E_{\mathrm{CCSDT}}+\Delta E_{\mathrm{CCSDTQ}} +\Delta E_{\mathrm{REL}}+\Delta E_{\mathrm{ZPE}}+\Delta E_{\mathrm{DBOC}}+\Delta E_{\mathrm{SO}}. \tag{4}$$

In Eq. (4), E_{HF}^{∞} and $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ are the HF-SCF (Hartree–Fock self-consistent field) and correlation energies, the latter given by the coupled-cluster singles and doubles method[36] with a perturbative treatment of triple excitations,[37] both extrapolated to the basis set limit. The next term is intended to account for deficiencies in the treatment of triple excitations in CCSD(T), the fourth term to account, approximately, for differences between the CCSDT (Ref. 38) and FCI correlation energies, where the latter is approximated by the CCSDTQ method.[39] ΔE_{ZPE} is the zero-point vibrational energy. The remaining terms remedy shortcomings of the simple, nonrelativistic Born–Oppenheimer approximation: ΔE_{DBOC} is the diagonal Born–Oppenheimer correction, ΔE_{SO} is the spin–orbit correction, and ΔE_{REL} is the scalar relativistic contribution to the energy.

We now describe how the individual terms in Eq. (4) are calculated.

A. Molecular geometries

The geometries[40] of species are taken from optimizations carried out at the CCSD(T) level of theory with the correlation-consistent cc-pVQZ basis sets.[41] A recent benchmark study[3] has demonstrated that this level of theory gives equilibrium geometries that are accurate to <0.003 Å (bond lengths) and $<0.5°$ (angles) of experimentally inferred values, where the latter are available. This level of accuracy is comparable to the best that can be achieved solely through analysis of experimental data, since experimental measurements of geometrical parameters never correspond to equilibrium distances and angles, but rather some type of averaged quantities.[42] Moreover, we correlate *all* electrons in the geometry optimizations, not just the valence electrons. While this is clearly not as good as using a properly defined (and much larger) core correlation basis (cc-pCVXZ),[43] it is considerably cheaper and the aforementioned benchmark study[3] demonstrated that the approach used here gives geometries that differ only negligibly for molecules containing first- and second-row atoms from those obtained in full-blown CCSD(T)/cc-pCVQZ optimizations. These geometries are then used in all subsequent calculations of quantities contributing to the HEAT energy, and the vibrational problem required for ΔE_{ZPE} is solved with CCSD(T)/cc-pVQZ at this geometry, as well. For closed-shell molecules, the restricted Hartree–Fock (RHF) orbitals are used in all calculations. For open-shell molecules, there are two obvious choices: unrestricted or restricted open-shell Hartree–Fock (UHF and ROHF, respectively). It was our intent to define the HEAT strategy in terms of UHF-based calculations. There are at least two cogent reasons for this choice. At very high levels of theory such as CCSDT, differences between UHF- and ROHF-based total energies are usually very small, so that the choice of reference function should have no impact on the energy calculations. Second, ROHF methods are more prone to symmetry-breaking and related effects,[44] and its use very often gives rise to nonsensical vibrational frequencies. However, in the course of this work, a rather curious problem was noted for the diatomic NO: UHF-based calculations give absurd parameters for the anharmonic force field; these severely degrade the quality of the calculated vibrational zero-point energy. This problem has been analyzed, and our findings can be found in a separate publication.[45] Hence, for the moment, the HEAT method will use UHF orbitals as the default for HF-SCF and CCSD(T) calculations on open-shell molecules; systems for which ROHF orbitals turn out to be more appropriate will be so designated.

B. CCSD(T) total energy

Together, the E_{HF}^{∞} and $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ terms give the estimated exact nonrelativistic electronic energy within the simple Born–Oppenheimer approximation, using the well-known CCSD(T) method to account for electron correlation effects. Following a relatively common convention, HF-SCF and CCSD(T) correlation energies have been obtained in a hierarchical series of basis sets, and then extrapolated *separately* to obtain estimates of the corresponding basis set limits. For closed-shell molecules, the restricted Hartree–Fock method has been used. Open-shell molecules were treated using the UHF approach.

For the HF-SCF energy, calculations were carried out using the augmented correlation consistent basis sets aug-cc-pCVXZ [$X=T(3)$, $Q(4)$ and 5] (Ref. 46) which are designed to treat core correlation effects properly. These three energies were then extrapolated with the formula advocated by Feller,[47]

$$E_{\mathrm{HF}}^{X}=E_{\mathrm{HF}}^{\infty}+a\exp(-bX), \tag{5}$$

where E_{HF}^{X} is the HF-SCF energy obtained with the aug-cc-pCVXZ basis set. The parameters, a, b, and the extrapolated HF-SCF energy E_{HF}^{∞}, are determined uniquely from the three energies.

For the correlation energy, a formula motivated by the atomic partial wave expansion is used,[48] viz.[49]

$$\Delta E_{\mathrm{CCSD(T)}}^{X}=\Delta E_{\mathrm{CCSD(T)}}^{\infty}+\frac{a}{X^3}, \tag{6}$$

where $\Delta E_{\mathrm{CCSD(T)}}^{X}$ is the CCSD(T) correlation energy [*not* the total CCSD(T) energy, which includes the HF-SCF contribution] obtained with the aug-cc-pCVXZ basis set. Here, there are two parameters, a and the estimated complete basis set limit CCSD(T) correlation energy $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$. These are uniquely determined by two correlation energies; the aug-cc-pCVQZ and aug-cc-pCV5Z energies are used in the HEAT protocol.

At this point, it is important to note a distinct difference between the HEAT approach and other model chemistries as well as most isolated efforts to obtain very accurate thermochemical parameters. We *do not* attempt to separate valence correlation effects from those arising from correlation of the core electrons. All of the calculations carried out in determining the extrapolated CCSD(T) energy use basis sets that are designed to treat core correlation as well as valence correlation, and no electrons are dropped from the correlation treatment in the individual CCSD(T) calculations. While it is true that one can obtain similar extrapolated valence-only estimates of the correlation energy and then add a correction for core correlation effects calculated with significantly smaller basis sets, we have chosen not to make the assumption of separation. Again, the point of this work is to do the best calculations possible in a common current computational environment with as few approximations as possible. However, we do recognize that this approach of combining core and valence correlation effects runs counter to the common practice of many of our colleagues, but point out that our strategy is undeniably more rigorous.[50]

C. Higher level correlation effects

Despite the never-ending success story that is the CCSD(T) method, it must not be forgotten that the (T) correction[37] is based on perturbation theory.[51] For cases where triples corrections are large, or alternatively (but not entirely independently) when there are severe problems associated with the reference function, there is cause to investigate the extent to which CCSD(T) differs from the complete treatment of triple excitations defined by the CCSDT approximation.[38] However, it is not possible to perform full CCSDT calculations using the large basis sets met with in the extrapolated CCSD(T) energies defined above. Due to both the relatively small differences expected and the fact that it appears that correlation effects beyond CCSD(T) can be estimated with smaller basis sets, we have chosen to estimate this contribution with the cc-pVTZ and cc-pVQZ basis sets and to correlate only the valence electrons. Implicitly assumed here is that effects due to diffuse functions and core correlation are already given sufficiently well by the extrapolated HF-SCF and CCSD(T) energies. Our formula for the CCSDT-CCSD(T) energy difference ($\Delta E_{\mathrm{CCSDT}}$) is

$$\Delta E_{\mathrm{CCSDT}}=E_{\mathrm{CCSDT}}^{\mathrm{TQ}}(\mathrm{fc})-E_{\mathrm{CCSD(T)}}^{\mathrm{TQ}}(\mathrm{fc}), \tag{7}$$

where TQ denotes that the corresponding contribution has been obtained by the correlation energy extrapolation formula [Eq. (6)] using the frozen-core CCSDT and CCSD(T) energies obtained with the cc-pVTZ and cc-pVQZ basis sets. For radicals, both the CCSDT and CCSD(T) calculations were performed using UHF reference functions.

Despite its computational complexity and cost, even CCSDT does not give correlation energies that are sufficiently accurate for the most demanding applications.[7,25,47,52,53] However, coupled-cluster calculations beyond CCSDT have only been generally possible (for small molecules, of course) with the development of general coupled-cluster codes.[54–59] Recently, Ruden and collaborators studied the impact of connected quadruples on atomization energies in double- and triple-ζ quality basis sets for six molecules.[60] The authors found that the contribution of quadruple excitations, as measured by the CCSDTQ CCSDT energy difference, converges rapidly with basis set size. In absolute terms, changes are largely negligible when going beyond a polarized valence double-ζ basis. In a more recent study focusing on a larger sample of molecules, Boese and co-workers[25] reached similar conclusions.

Based on these findings, the effects of higher-level correlation effects (those beyond CCSDT) are estimated in the HEAT protocol by subtracting the CCSDT and CCSDTQ correlation energies obtained with the cc-pVDZ basis set in the frozen-core approximation:

$$\Delta E_{\mathrm{CCSDTQ}}=E_{\mathrm{CCSDTQ}}^{\mathrm{cc\text{-}pVDZ}}(\mathrm{fc})-E_{\mathrm{CCSDT}}^{\mathrm{cc\text{-}pVDZ}}(\mathrm{fc}). \tag{8}$$

Due to program limitations that existed while the data were being compiled for this research, the $\Delta E_{\mathrm{CCSDTQ}}$ correction is based on ROHF reference functions for the radical species. There is an implicit assumption here—that the CCSDT total energies for radicals are independent of the reference function—but this seems to be justified.[61]

This approximation, which is intended to account for the difference between CCSDT and an exact treatment of correlation, is clearly important and is further discussed in the appendix.

D. Zero-point vibrational energy

Zero-point vibrational energies for all species (apart from atoms, where it vanishes) were determined from anharmonic force fields calculated at the (all electron) CCSD(T) level of theory with the cc-pVQZ basis sets, using RHF orbitals for closed-shell systems and UHF orbitals for most of the open-shell molecules.[62] Using standard spectroscopic rovibrational perturbation theory,[63] the vibrational energy levels are given by the expression

$$E(v)=G_0+\sum_i \omega_i\left(v_i+\frac{1}{2}\right)+\sum_{i\geqslant j} x_{ij}\left(v_i+\frac{1}{2}\right)\left(v_j+\frac{1}{2}\right), \tag{9}$$

where ω_i are the harmonic frequencies. Explicit equations for the anharmonicity constants x_{ij} are given, for example, in Ref. 63. Hence, the ground state energy is given by

$$E_{\mathrm{ZPE}}=G_0+\sum_i \frac{\omega_i}{2}+\sum_{i\geqslant j}\frac{x_{ij}}{4} \tag{10}$$

within this model. The second term is the familiar harmonic approximation, while addition of the third and (especially) first terms are rarely included in quantum chemical investigations. Work over the past six years in our laboratories has led to the development of analytic second derivative methods for the CCSD(T) method,[64,65] and parallel developments have provided the avenue toward accurate and efficient evaluation of quartic force fields via numerical differentiation of analytic second derivatives. In fact, it is now possible to calculate the cubic and quartic force fields, many physical quantities that depend upon them, and the second and third terms of Eq. (10) by simply "pushing a button."[66] The first term is an oft-forgotten constant (does not depend on the vibrational state) term[67,68] which contributes to the vibrational energy. Recent efforts in our laboratory[69] as well as others[10,70] have led to the development of explicit formulas for G_0 in terms of quantities calculable from the quartic force field. However, our efforts in this direction have only provided equations that apply to asymmetric tops, although we are working on the required modifications to treat symmetric tops, spherical tops, and polyatomic linear molecules. Because of our inability to calculate G_0 for all molecules, we have chosen to neglect this contribution in the zero-point energies. However, we take some solace from the recent work of Schuurman *et al.*,[10] where this term was included in an exhaustive study of the enthalpies of formation for HNCO and its isomers (which, notably, used a reaction scheme of the sort that we advocate rather than being based strictly on atomization energies). They found that G_0 was typically less than 10 cm^{-1} in magnitude, similar to what we have found in pilot applications,[69] and below that which is inherent in the calculation of the other two (and numerically more significant) contributions to Eq. (10). Hence, neglect of this constant term is not expected to cause significant errors, although it should be checked in selected cases (cf. H_2O_2 in the following section) and eventually included when general formulas are available. However, for now, the HEAT zero-point vibrational energy contribution will be defined as

$$\Delta E_{\mathrm{ZPE}}=\sum_i \frac{\omega_i}{2}+\sum_{i\geqslant j}\frac{x_{ij}}{4}. \tag{11}$$

E. Diagonal Born–Oppenheimer correction

It is not always appreciated that the electronic energy, as obtained from traditional electronic structure calculations, is not equivalent to the expectation value of the molecular Hamiltonian over the electronic (clamped-nucleus) wave function. The difference lies in the contribution of the nuclear kinetic energy operator, which can be viewed as a first-order correction to the usual electronic energy. Nonetheless, the simple potential energy surface picture of a molecule is not lost, as each geometry continues to be associated with a specific value of the energy (for a given electronic state) although this surface becomes mass dependent. For atoms, this correction—known as the diagonal Born–Oppenheimer correction (DBOC)—accounts for the finite mass of the nucleus. This diagonal Born–Oppenheimer correction[31–34,71–73] is calculated by the expectation value

$$\Delta E_{\mathrm{DBOC}}=\langle\Psi_e(\mathbf{r};\mathbf{R})|\hat{T}_n|\Psi_e(\mathbf{r};\mathbf{R})\rangle, \tag{12}$$

where $\hat{T}_n$ is the nuclear kinetic energy operator and $\Psi_e(\mathbf{r};\mathbf{R})$ is the normalized electronic wave function obtained at the set of nuclear positions parametrized by $\mathbf{R}$.

Despite its deceptively simple form, it is not straightforward to calculate the DBOC; efforts in this direction have mostly been (for an exception, see Ref. 72) limited to the HF-SCF or multiconfigurational SCF (MCSCF) level of theory. The landmark paper of Handy, Yamaguchi, and Schaefer[32] was the first which reported DBOC energies for a number of polyatomic systems. Recent studies[71,72] indicate that the use of correlated electronic wave functions has only a modest effect on the DBOC correction. Thus we believe that DBOC corrections calculated with HF-SCF wave functions are sufficiently accurate for the purposes of this study. Therefore, it has been chosen for the HEAT protocol. The aug-cc-pVTZ basis set has been used to calculate ΔE_{DBOC} at the HF-SCF (RHF and ROHF for closed and open-shell systems, respectively) level. This particular choice for the basis set is motivated by a previous study[72] where the DBOC correction was found to converge relatively rapidly to the one-particle basis limit, provided that diffuse (low-exponent) functions are included in the basis set.

F. Spin–orbit correction

Calculations performed within the framework of a nonrelativistic Hamiltonian give a weighted average over the energies of various states involving different coupling of spin and orbital angular momentum. Consideration[74] of this shortcoming of nonrelativistic theory is necessary for some of the species considered in this work. The relative energies of various states split by the spin–orbit interaction can be calculated from a Hamiltonian that includes the spin–orbit operator. The calculated energy lowering of the lowest spin–orbit state (which is of course the ground state of interest) with respect to the averaged state obtained in a nonrelativistic calculation can be used to adjust the ground state energy.

Since the magnitude of this effect is relatively small for first- and second-row atoms, these calculations need not be performed at levels of theory rivaling those used in the energy extrapolations.

The calculations of ΔE_{SO} (defined here as the energy difference between the ground state and the statistically weighted average of all spin–orbit states—which is not equivalent to the spin–orbit coupling constant!) have been performed with a spin–orbit configuration interaction (CI) procedure. For a detailed description, see Ref. 75. In summary, the core electrons are described by relativistic effective core potentials (RECP) including spin–orbit terms that allow a straightforward calculation of the spin–orbit interaction integrals. The CI wave functions are constructed by considering all single and double excitations out of a valence complete active space reference function. To reduce the computational effort, the double-group symmetry that often facilitates relativistic quantum calculations can be used. The cc-pVDZ basis set developed by Pitzer[76] together with the corresponding RECPs (Ref. 77) were used in the calculations.

It should be noted that we consider only first-order spin-orbit interactions in this work. Hence, the only molecules for which the ΔE_{SO} contribution is nonzero are radicals in degenerate ground states. Second-order spin–orbit effects, which involve coupling of the ground state with excited states of different spin through the spin–orbit operator, are not included and are expected to be of negligible import for thermochemistry.

One could of course use measured spin–orbit splittings to calculate the correction. It is clearly the contribution appearing in Eq. (4) for which experimental determination is clearly superior to computational estimation. However, since one might want to investigate species where the spin–orbit splitting is not available, we prefer the computational approach. It is also in keeping with the spirit of the HEAT method, in which appeals to experiment and empiricism are kept to a minimum. A quick check reveals that experimentally measured and calculated values of ΔE_{SO} for the molecules in this work differ by less than 10 cm^{-1} in all cases.

G. Scalar relativistic effects

The effect of so-called scalar relativistic contributions[29] to the HEAT total energy (ΔE_{REL}) are included by contracting the one-particle density matrix obtained at the CCSD(T)/aug-cc-pCVTZ level with (one-electron) Darwin and mass-velocity terms. As discussed by Davidson, Ishikawa and Malli,[78] this is a reasonable approximation for relativistic effects when first-row elements are considered. Recently, Boese *et al.*[25] compared the sum of Darwin and mass-velocity contributions obtained at high levels of theory to the second-order Douglas–Kroll contributions. While the comparison there is complicated a bit by the lack of a common basis set or method, the results suggest that negligible error is incurred with the simpler first-order treatment *for first- and second-row atoms*.

H. Computational details

Calculations of the stabilization of the lowest spin–orbit level relative to the weighted average energy were performed with COLUMBUS,[79] while most of the calculations of ΔE_{DBOC} were carried out with the PSI3.2 electronic structure program suite.[80] CCSDTQ calculations were done with a string-based many-body code.[57] All HF-SCF, CCSD, CCSD(T), and CCSDT energy calculations as well as evaluations of ΔE_{REL} and some ΔE_{DBOC} calculations come from a local version of ACES2.[81] Anharmonic force fields used for ΔE_{ZPE} were calculated with ACES2 using the algorithm described in Ref. 82. All calculations, some of which involved more than 500 basis functions, were performed on personal computers running the LINUX operating system.

III. RESULTS

Total HEAT energies, as well as the individual contributions defined in the preceding section, are listed for 31 atoms and molecules in Table I. Species in the test suite contain only hydrogen, carbon, oxygen, nitrogen, or fluorine atoms; there are 12 closed-shell systems and 19 are doublets, triplets, and quartets. The complete documentation of the individual energy contributions should be of use to others interested in expanding the HEAT database, as they can be used to provide a check on calculations. However, the most important numbers in Table I—and indeed in the entire manuscript—are those in the rightmost column: the HEAT total energies for each species. In the following sections, we illustrate use of the HEAT energies in calculating atomization energies and standard enthalpies of formation at 0 K.

A. Atomization energies

Atomization energies calculated from HEAT energies are given in Table II for the 27 molecular species in the test suite, along with the individual contributions. Among other things, the first two columns of the table show the well-known anomalous behavior of fluorine (F_2) and also OF, which are both unbound at the UHF level and owe their stability as molecules to electron correlation. There are also several other species (O_2, HNO, HO_2, NO) for which correlation effects account for more than half of the atomization energy; Hartree–Fock and correlation contributions are essentially equal for N_2, H_2O_2, and CN. It is not particularly surprising that the magnitude of higher-level correlation contributions—defined here as those beyond CCSD(T)—correlates strongly with the overall correlation contributions identified above. The largest values of $\Delta E_{\mathrm{CCSDTQ}}$ are found for N_2, F_2, O_2, CO_2, H_2O_2, NO, HNO, CN, and HCN, all but HCN mentioned above in the context of having large overall correlation contributions to the atomization energy. For each of these species, the effect of correlation beyond CCSDT—a tract of the quantum chemical landscape that is rarely trod upon in practice—is to increase atomization energies by more than 3 $\mathrm{kJ\,mol}^{-1}$. Given that the standard definition of chemical accuracy is 4.184 $\mathrm{kJ\,mol}^{-1}$, one realizes that this time-honored goal of quantum chemistry is still not easily obtained, at least in the context of atomization energies.

TABLE II. Contributions to atomization enthalpies for the molecules in the test set. All values are in kJ mol^{-1}.

Species	E_{HF}^{∞}	$\Delta E_{\mathrm{CCSD(T)}}^{\infty}$	$\Delta E_{\mathrm{CCSDT}}$	$\Delta E_{\mathrm{CCSDTQ}}$	ΔE_{REL}	ΔE_{ZPE}	ΔE_{DBOC}	ΔE_{SO}	Total
N_2	482.94	472.26	−3.14	3.61	−0.56	−14.12	0.08	0.00	941.07
H_2	350.81	107.41	0.00	0.00	−0.01	−25.97	0.22	0.00	432.46
F_2	−154.98	316.00	−1.31	3.42	−0.20	−5.50	0.02	−3.01	154.44
O_2	140.36	362.56	−2.07	4.46	−0.82	−9.56	0.06	−1.61	493.39
CO	729.86	356.91	−2.37	2.21	−0.69	−12.98	0.07	−1.20	1071.82
C_2H_2	1228.97	468.20	−3.03	2.23	−1.14	−68.93	0.50	−0.76	1626.06
CCH	777.34	333.02	0.56	2.28	−1.12	−36.34	0.23	−0.76	1075.23
CH_2	649.11	149.13	−0.02	0.14	−0.63	−45.07	0.11	−0.38	752.39
CH	238.28	113.57	0.43	0.12	−0.18	−16.97	−0.34	−0.20	334.70
CH_3	1017.30	271.11	−0.15	0.24	−0.72	−77.64	0.19	−0.38	1209.93
CO_2	1033.08	599.66	−4.49	4.11	−2.05	−30.40	0.20	−2.02	1598.10
H_2O_2	562.96	562.77	−2.37	2.95	−1.63	−68.77	0.52	−1.64	1054.81
H_2O	652.40	323.02	−0.97	0.99	−1.14	−55.73	0.53	−0.82	918.26
HCO	764.53	403.66	−1.78	2.19	−1.16	−34.03	−0.10	−1.20	1132.11
HF	405.62	187.00	−0.67	0.72	−0.84	−24.66	0.34	−1.51	566.01
HO_2	300.93	431.57	−0.60	2.52	−1.20	−37.05	0.02	−1.64	694.56
NO	225.52	411.50	−1.54	3.07	−0.83	−11.46	0.06	−0.10	626.22
OH	286.41	162.12	−0.13	0.48	−0.54	−22.21	0.05	−0.04	426.14
HNO	331.69	528.35	−2.45	3.54	−1.19	−35.92	−0.23	−0.82	822.97
CN	379.38	371.64	3.14	3.43	−0.63	−12.75	0.10	−0.38	743.92
HCN	833.50	477.62	−3.21	3.04	−0.94	−41.74	0.32	−0.38	1268.20
CF	338.46	216.63	−0.60	1.00	−0.72	−7.88	0.03	−1.42	545.50
NH_2	492.92	270.76	0.06	0.48	−0.67	−49.57	−0.03	0.00	713.95
NH_3	841.38	406.03	−0.63	0.72	−1.08	−89.45	0.56	0.00	1157.53
NH	214.88	132.40	0.29	0.22	−0.31	−19.46	−0.19	0.00	327.83
OF	−64.04	282.44	1.05	2.10	−0.41	−6.37	0.02	−1.24	213.56

The overall difference between the exact correlation contributions (as estimated by HEAT) and those associated with $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ also includes $\Delta E_{\mathrm{CCSDT}}$, which is intended to remedy deficiencies in the CCSD(T) treatment of triple excitation effects. And here, one can only say that CCSD(T) comes through again. While the $\Delta E_{\mathrm{CCSDTQ}}$ contributions are uniformly positive, those associated with $\Delta E_{\mathrm{CCSDT}}$ are generally negative, indicating that a HEAT thermochemistry based on CCSD(T) energies instead of estimated exact correlation energies would outperform one based on CCSDT energies.[83] When these two contributions are combined [thereby providing an estimate of the difference between CCSD(T) and FCI], excellent cancellation is found in many cases. Exceptions, where the net effect exceeds 1 kJ mol^{-1}, are (kJ mol^{-1} in descending order): CN (6.57), OF (3.15), CCH (2.84), O_2 (2.39), F_2 (2.12), HO_2 (1.92), NO (1.53), and HNO (1.09), most being radicals where the perturbative nature of the (T) correction to the energy is most suspect.[84] The particularly large value seen for CN also reflects differences in CCSD(T) energies calculated with UHF and ROHF reference functions; it is quite likely that this value would be smaller if the ROHF-based CCSD(T) method[85,86] was used in determining the $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ contribution.

Continuing to the right in Table II, we see that scalar relativistic effects systematically reduce the atomization energies by as much as 2 kJ mol^{-1}, corrections to the simple Born–Oppenheimer approximation generally—but not always—increase the atomization energy by no more than 0.5 kJ mol^{-1}. Stabilization of the lowest spin–orbit level can amount to as much as 2 kJ mol^{-1}. Here, it is important to consider the scope of the molecules in the test suite. Scalar relativistic effects will increase with the total number of electrons and the number of bonding electrons (NH, NH_2, and NH_3 form an interesting sequence here, as does OH and H_2O—it seems to indeed be well-approximated by "bond contributions" that can be inferred from the table), and will become more important as one moves down the periodic table. The same is true for spin–orbit effects. The DBOC, on the other hand, is most important for light atoms as can clearly be seen by noting its essentially negligible magnitude for molecules that do not contain hydrogen atoms.

Zero-point vibrational energy contributions to the atomization energies are obviously negative in all cases. Enthalpies of formation calculated from atomization energies and experimental atomic enthalpies of formation will be discussed subsequently.

B. Formation from the elements

Most theoretical model chemistries present atomization energies as the primary thermochemical data (we remind the reader that the total HEAT energies in Table I are *the* primary data of this paper), but one could also base the calculations on different elemental reference compounds. One choice would be the elements in their standard state, in which case the reaction energies determined would be equivalent to standard enthalpies of formation. However, standard states as defined for many elements are not amenable to computation (carbon, chlorine, boron, to name a few) and this approach has not been followed in the literature of model chemistries. Nevertheless, we will have a go at it here.

Table III lists reaction energies calculated from HEAT data for the formation of the test suite species from molecular reference compounds. For H, N, O, and F, the standard

TABLE III. Contributions to reaction enthalpies for the formation of the test compounds from the elemental reactions, as defined in the text. All values are in kJ mol^{-1}.

Species	E_{HF}^{∞}	$\Delta E_{\mathrm{CCSD(T)}}^{\infty}$	$\Delta E_{\mathrm{CCSDT}}$	$\Delta E_{\mathrm{CCSDTQ}}$	ΔE_{REL}	ΔE_{ZPE}	ΔE_{DBOC}	ΔE_{SO}	Total
C	729.86	356.91	−2.37	2.21	−0.69	−12.98	0.07	−1.20	1071.82
F	−77.49	158.00	−0.65	1.71	−0.10	−2.75	0.01	−1.51	77.22
H	175.41	53.71	0.00	0.00	−0.01	−12.99	0.11	0.00	216.23
N	241.47	236.13	−1.57	1.80	−0.28	−7.06	0.04	0.00	470.54
O	70.18	181.28	−1.03	2.23	−0.41	−4.78	0.03	−0.80	246.69
C_2H_2	581.57	353.03	−1.70	2.19	−0.26	16.99	−0.14	−1.64	950.04
CCH	857.79	434.51	−5.30	2.15	−0.27	−2.61	0.01	−1.64	1284.64
CH_2	431.57	315.19	−2.34	2.08	−0.08	6.12	0.18	−0.82	751.89
CH	666.99	297.05	−2.79	2.09	−0.52	−9.00	0.52	−1.00	953.35
CH_3	238.79	246.92	−2.21	1.98	0.02	25.70	0.21	−0.82	510.58
CO_2	−162.86	119.81	0.05	2.56	0.54	7.86	−0.08	−0.79	−32.90
H_2O_2	−71.79	−92.80	0.30	1.51	0.80	33.24	−0.24	0.03	−128.96
H_2O	−231.40	−34.33	−0.06	1.24	0.72	24.98	−0.28	0.02	−239.11
HCO	210.92	188.24	−1.62	2.25	0.06	3.28	0.31	−0.80	402.63
HF	−307.71	24.71	0.02	0.99	0.74	8.92	−0.22	0.00	−272.55
HO_2	14.83	−15.30	−1.47	1.94	0.38	14.50	0.14	0.03	15.05
NO	86.13	5.91	−1.06	0.96	0.15	−0.38	0.01	−0.71	91.01
OH	−40.83	72.87	−0.91	1.75	0.12	4.45	0.09	−0.76	36.78
HNO	155.37	−57.23	−0.15	0.49	0.50	11.09	0.41	0.02	110.49
CN	591.96	221.40	−7.07	0.59	−0.34	−7.29	0.01	−0.82	798.44
HCN	313.24	169.13	−0.72	0.98	−0.04	8.71	−0.09	−0.82	490.38
CF	313.91	298.28	−2.42	2.93	−0.07	−7.85	0.05	−1.29	603.54
NH_2	99.36	72.78	−1.63	1.32	0.38	16.54	0.30	0.00	189.05
NH_3	−73.69	−8.78	−0.94	1.08	0.79	43.43	−0.19	0.00	−38.30
NH	202.00	157.43	−1.86	1.58	0.03	−0.59	0.35	0.00	358.94
OF	56.73	56.84	−2.74	1.84	−0.09	−1.16	0.02	−1.07	110.36

definition of the corresponding diatomic molecule is used, so that the tabulated reaction energies for molecules containing only these elements are equivalent to the standard enthalpies of formation. For carbon, where the elemental standard state is graphite, we use carbon monoxide in a way most easily communicated by example. For acetylene, the reaction used is

$$2\mathrm{CO}+\mathrm{H}_2\rightarrow \mathrm{C}_2\mathrm{H}_2+2\mathrm{O}, \tag{13}$$

while

$$\mathrm{O}_2+\mathrm{H}_2\rightarrow \mathrm{HOOH} \tag{14}$$

is used for the stoichiometrically similar hydrogen peroxide. In the former case, the calculated reaction energy is not equivalent to the enthalpy of formation, differing from it by twice the difference of the enthalpies of formation of CO and O.

All in all, the relative magnitudes of the Hartree–Fock and correlation contributions to the reaction energies is not entirely dissimilar to that found for atomization energies. Correlation corrections dominate in magnitude in a few cases, and others exhibit comparable contributions.

Higher-level correlation contributions are also comparable in magnitude using the atomization and elemental reaction calculations. While the $\Delta E_{\mathrm{CCSDTQ}}$ contributions are decidedly smaller for the latter (no contributions above 3 kJ mol^{-1} while 8 of the 26 atomization energy contributions exceed this value), the mean absolute post-CCSD(T) correlation contributions (the sum of $\Delta E_{\mathrm{CCSDT}}$ and $\Delta E_{\mathrm{CCSDTQ}}$) are about the same: 0.82 kJ mol^{-1} for the atomization energies and 0.90 kJ mol^{-1} for the elemental reactions.

The remaining contributions are also similar in magnitude in the two cases, except for ΔE_{ZPE} since now there are vibrational modes in all species in the chemical equation.

C. Enthalpies of formation

In Table IV, enthalpies of formation at 0 K ($\Delta_f H_0^\circ$) calculated from HEAT energies are given for all members of the test suite. The calculations were done by two different procedures. In the first (I), HEAT atomization energies were corrected to standard enthalpies of formation according to Eq. (1),[87] using experimental enthalpies of formation for the atoms. Procedure II is based on the elemental reactions summarized in the preceding section. For all molecules not containing carbon, values of $\Delta_f H^\circ$ obtained by procedure II are equal to the reaction energies given in Table III. This underscores one relative advantage of the elemental reaction approach, specifically that it is based on reference compounds whose enthalpy of formation is precisely zero by definition. This is advantageous only in extremely accurate calculations such as those presented here, since it avoids errors associated with atomic enthalpies of formation for species such as C and F.[88] Since one cannot do a HEAT calculation on a chunk of graphite, however, an alternative approach was used. The reference for carbon used in approach II was carbon monoxide, as described in the preceding section. CO was chosen since its experimentally determined enthalpy of formation [-113.81 ± 0.17 from Ruscic's Active Thermochemical Tables (ATcT) Refs. 89–91] is not tied to that of the carbon atom[92] and should therefore be a relatively stable reference. Given the CO value as well as that for O ($\Delta_f H^\circ = 246.84$

TABLE IV. Standard enthalpies of formation (in kJ mol^{-1}), calculated from HEAT atomization energies and experimental atomic enthalpies of formation (I) and the elemental reaction approach (II). Experimental values are from the ATcT (Ref. 89) (superscript a) where available. Remaining values come from the NIST-JANAF compendium. Fixed values are in parenthesis.

Species	$\Delta_f H^\circ$(I)	$\Delta_f H^\circ$(II)	Expt.
N_2	0.11	0.00	0.00±0.00
H_2	−0.39	0.00	0.00±0.00
F_2	−0.02	0.00	0.00±0.00
O_2	0.3(0)	0.00	0.00±0.00
C	(711.79)	711.17	711.79±0.21[a]
F	(77.21)	77.22	77.21±0.24[a]
H	(216.03)	216.23	216.03±0.00[a]
N	(470.59)	470.54	470.59±0.05[a]
O	(246.84)	246.69	246.84±0.00[a]
CO	(−113.18)	−113.81	−113.81±0.17[a]
C_2H_2	229.59	228.74	228.20±0.64[a]
CCH	564.39	563.34	563.32±0.65[a]
CH_2	391.47	391.24	390.65±0.54[a]
CH	593.12	592.70	593.19±0.36[a]
CH_3	149.96	149.93	149.94±0.11[a]
CO_2	−392.63	−393.55	−393.11±0.01[a]
H_2O_2	−129.05	−128.96	−129.82±0.08[a]
H_2O	−239.35	−239.11	−238.92±0.04[a]
HCO	42.55	41.98	42.09±0.38[a]
HF	−272.76	−272.55	−272.73±0.24[a]
HO_2	15.16	15.05	14.96±0.64[a]
NO	91.22	91.01	90.53±0.09[a]
OH	36.74	36.78	37.09±0.05[a]
HNO	110.50	110.49	102.50±0.42[b]
CN	438.47	437.79	436.80±10.00[b]
HCN	130.21	129.73	135.53±8.40[b]
CF	243.50	242.89	251.60±8.00[b]
NH_2	188.71	189.05	193.25±6.30[b]
NH_3	−38.84	−38.30	−38.91±0.40[b]
NH	358.80	358.94	376.51±16.70[b]
OF	110.50	110.36	108.00±10.00[b]
Mean absolute error[c]	0.37	0.24	⋯
Mean signed error[c]	−0.28	−0.06	⋯
RMS error[c]	0.56	0.35	⋯
Maximum error[c]	1.39	0.86	⋯

[a] From active thermochemical tables.
[b] From NIST-JANAF compilation.
[c] Based only on active thermochemical tables data.

±0.002), elemental reaction energies—as defined earlier—need to be adjusted by 360.65 kJ mol^{-1} per carbon atom to give enthalpies of formation.

The results are startling. For compounds that have well-established and self-consistent enthalpies of formation, as determined by Ruscic's ATcT approach,[89,91] the HEAT values determined by method II are within 1 kJ mol^{-1} in all cases! Indeed, for only one example—H_2O_2 (−128.96 versus −129.82±0.08 kJ mol^{-1})—does the calculated HEAT value fall more than 0.5 kJ mol^{-1} outside the range estimated by the ATcT approach. Using method I, both C_2H_2 and H_2O_2 fall more than 0.5 kJ mol^{-1} outside the ATcT estimate. For set I, 7 of the 16 HEAT values are within the ATcT error bars, while 8 set II values fall within the estimated bounds. The statistical analysis shown at the bottom of Table IV gives further support that enthalpies of formation determined by the atomization energy approach (I) are not quite as good, although the performance difference is decidedly small.

The astonishing accuracy of HEAT energies in predicting the enthalpies of formation for the well-characterized compounds in Table IV suggests that they are "better" estimates than those found in the NIST-JANAF database for all of the remaining compounds (except perhaps ammonia), in the sense that the error bar associated with the HEAT values (which we assign as 1 kJ mol^{-1} for all but difficult cases) is roughly an order of magnitude smaller. An interesting case that warrants further study is the HNO molecule, where the NIST-JANAF uncertainty is rather small (0.42 kJ mol^{-1}), while the HEAT value is about 8 kJ mol^{-1} higher.

The compounds that appear to be the most difficult cases for HEAT are C_2H_2 and H_2O_2. Calculated enthalpies of formation (0 K) for both are well outside (relatively speaking) the range of values from ATcT, which merits some discussion. For C_2H_2, the error appears to result from a neglect of *connected pentuple excitations* (see appendix), while the discrepancy for H_2O_2 is unclear. It is perhaps significant that H_2O_2 stands alone amongst the test suite molecules in having a low-frequency torsional mode. The way in which ΔE_{ZPE} is calculated in the HEAT protocol is ideally suited for semirigid molecules where both the harmonic approximation and second-order vibrational perturbation theory[63] are expected to work well. It may well break down for torsional modes such as that in H_2O_2, at least at the level of tenths of a kJ mol^{-1}. A detailed analysis of H_2O_2 is underway, and will be reported in a separate publication.[93]

IV. DISCUSSION

As continually stated throughout this paper, the total HEAT energies given in Table I represent the principal (and indeed, the only nonredundant) results of this research. It is hoped that the 31 examples studied here are just the beginning of a fairly significant database of systems for which HEAT results are available. To this end, we have already begun calculations on a few additional systems, and invite other members of the quantum chemical community to join the effort. Although the results presented in the preceding section for the enthalpies of formation are already extremely accurate (more so, in fact, than we had anticipated when this project began), even greater accuracy can be obtained if the theoretical results are used to calculate enthalpy changes for reactions in which the bonding environments of atoms on the left and right side of the chemical equation are similar. There is essentially a continuum between atomization energies (which are undoubtedly the most difficult quantities to calculate accurately) and isodesmic reactions in which the local environment of every atom is preserved on both sides of the equation. In isodesmic and nearly isodesmic reactions, one expects correlation effects to be less severe. Hence, the $\Delta E_{\mathrm{CCSDT}}$ and $\Delta E_{\mathrm{CCSDTQ}}$ contributions to the HEAT energy should be considerably smaller in magnitude. This is advantageous, because unlike E_{HF}^{∞} and $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$, these higher-order correlation contributions are not extrapolated as throughly nor are they free of the assumption of negligible core correlation effects.

Some evidence that this is the case can be found in Table V, where HEAT thermochemistry values are given for sev-

TABLE V. Individual contributions from HEAT energies to reaction enthalpies at 0 K. All values in kJ mol^{-1}.

Reaction	E_{HF}^{∞}	$\Delta E_{\mathrm{CCSD(T)}}^{\infty}$	$\Delta E_{\mathrm{CCSDT}}$	$\Delta E_{\mathrm{CCSDTQ}}$	ΔE_{REL}	ΔE_{ZPE}	ΔE_{DBOC}	ΔE_{SO}	Total
$H+H_2O_2 \rightarrow HO_2+H_2O$	−103.95	−29.70	−0.92	−0.08	0.17	1.80	0.02	0.78	−131.88
$CH+C_2H_2 \rightarrow CCH+CH_2$	40.80	99.62	−3.15	−0.06	0.43	−4.48	−0.19	0.18	133.14
$CH_2+C_2H_2 \rightarrow CCH+CH_3$	83.44	13.20	−3.47	−0.14	0.08	−0.01	0.19	0.00	93.29
$HCO+C \rightarrow CH+CO$	−203.62	−66.82	0.16	−0.14	−0.29	−4.07	0.18	0.20	−274.40
$N+NH_3 \rightarrow NH_2+NH$	133.58	2.86	−0.98	0.02	−0.10	−20.41	0.79	0.00	115.75
$2NH \rightarrow N+NH_2$	−63.16	−5.96	0.51	−0.03	0.04	10.65	−0.36	0.01	−58.29
$HCN+C \rightarrow CH+CN$	215.84	−7.59	−6.77	−0.51	−0.13	−12.02	0.56	0.20	189.59
$CH \rightarrow C+H$	238.28	113.57	0.43	0.12	−0.18	−16.97	−0.34	−0.20	334.70
$CH_2 \rightarrow CH+H$	410.83	35.56	−0.45	0.02	−0.45	−28.10	0.46	−0.18	417.69
$CH_3 \rightarrow CH_2+H$	368.19	121.98	−0.13	0.10	−0.10	−32.57	0.08	0.00	457.54
$NH \rightarrow N+H$	214.88	132.40	0.29	0.22	−0.31	−19.46	−0.19	0.00	327.83
$NH_2 \rightarrow NH+H$	278.04	138.36	−0.23	0.25	−0.36	−30.11	0.16	0.00	386.12
$NH_3 \rightarrow NH_2+H$	348.46	135.26	−0.69	0.24	−0.41	−39.87	0.60	0.00	443.58
$OH \rightarrow O+H$	286.41	162.12	−0.13	0.48	−0.54	−22.21	0.05	−0.04	426.14
$H_2O \rightarrow OH+H$	365.98	160.90	−0.85	0.51	−0.60	−33.52	0.48	−0.78	492.12
$HF \rightarrow F+H$	405.62	187.00	−0.67	0.72	−0.84	−24.66	0.34	−1.51	566.01

eral chemical reactions. The first group of reactions are isodesmic or nearly so; the types (partners and bond order) of chemical bonds are preserved in the reactions, although free atoms appear in the last four reactions and the seventh reaction involves a dramatic difference in the location of an unpaired electron. Nonetheless, the $\Delta E_{\mathrm{CCSDTQ}}$ corrections are less than 1 kJ mol^{-1} in all cases, in contrast to the atomization and elemental reactions documented in Tables II and III, where $\Delta E_{\mathrm{CCSDTQ}} > 1$ kJ mol^{-1} is the norm. Hence, enthalpies calculated from HEAT energies that exclude the $\Delta E_{\mathrm{CCSDTQ}}$ contribution which is—at least for the larger species—the most expensive calculation met with in the HEAT protocol, should be considerably more accurate than atomization or elemental formation reactions based on the same approximations. We note in passing that some of the $\Delta E_{\mathrm{CCSDT}}$ contributions are large for these reactions, specifically those involving the isoelectronic CCH and CN radicals for which UHF-CCSD(T) performs poorly due to spin-contamination effects.[94]

Also interesting is the second set of reactions in Table V, all of which involve breaking of an X-H bond, where X is a first-row element. In all of these, there is just one difference in the bonding of products and reactants (there is an extra X-H bond in the former), and these should benefit from cancellation of errors. And indeed, higher-order correlation effects are quite small, with both $\Delta E_{\mathrm{CCSDT}}$ and $\Delta E_{\mathrm{CCSDTQ}}$ always below 1 kJ mol^{-1}. Even beyond this, there is a striking regularity in the $\Delta E_{\mathrm{CCSDTQ}}$ contributions to these reaction energies. The magnitude of this contribution is always positive, and increases systematically with the electronegativity of X. Alternatively, it increases with decreasing X-H distance. Moreover, it is remarkably constant for the sequences $XH \rightarrow X+H$, $XH_2 \rightarrow XH+H$, etc. This can be used advantageously in HEAT studies in which this expensive contribution is excluded. For example, if one was interested in determining the enthalpy of formation of an alkoxy radical from the corresponding alcohol, it would seem pragmatic to skip the CCSDTQ calculation and simply assume that the difference in $\Delta E_{\mathrm{CCSDTQ}}$ energies of the alcohol and alkoxy radical is 0.5 kJ mol^{-1}. The behavior of $\Delta E_{\mathrm{CCSDT}}$ is considerably less systematic, but this is less of a problem since it is comparatively cheap to calculate.

An example of how we feel the HEAT energies are best used is provided by the NH and NH_2 molecules. Uncertainties in NIST-JANAF values for $\Delta_f H^\circ$ for these molecules are rather large (16.7 and 6.3 kJ mol^{-1}, respectively). While the values in Table IV are likely to be within 1 kJ mol^{-1}, there is another approach which is potentially superior. The calculated HEAT reaction energies for $NH_3 \rightarrow NH_2+H$, $NH_2 \rightarrow NH+H$ and $NH \rightarrow N+H$ in Table V are surely more precise than enthalpies of formation calculated from the relatively difficult atomization energies or elemental formation energies in Table III. These can be combined with the rather precisely known $\Delta_f H^\circ$ value of H and NH_3 (216.03 and -38.91 ± 0.4 kJ mol^{-1}) to yield values of 188.64 kJ mol^{-1} and 358.73 kJ mol^{-1} for $\Delta_f H^\circ$ of NH_2 and NH, respectively. It should be noted that both of these numbers are also in good agreement with those in Table IV.

Several follow-up studies are in order. First of all, we recognize the computationally demanding nature of the HEAT procedure. The calculations of $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$ and $\Delta E_{\mathrm{CCSDTQ}}$, in particular, are the most arduous. Simplification of these steps would certainly result in a theoretical model chemistry that would be more widely applicable. But just how accurate would it be? For $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$, there are some alternatives that come to mind. First of all, the extrapolated energy could be based on aug-cc-pCVTZ and aug-cc-pCVQZ calculations, which would avoid the aug-cc-pCV5Z calculations that involve, for example, 543 basis functions for a molecule as small as CO_2. Another course of action would be to invoke separation of core and valence correlation. One could even dispense with diffuse functions,[95] and calculate extrapolated CCSD(T) energies with the cc-pVQZ and cc-pV5Z (or cc-pVTZ and cc-pVQZ) basis sets in dropped core calculations. Effects of core correlation could then be estimated from relatively small basis sets (cc-pCVDZ or cc-pCVTZ) in all electron calculations. These values, when combined, would represent an approximation to $\Delta E_{\mathrm{CCSD(T)}}^{\infty}$. Ultimately, basis set extrapolations are empirical in nature. Explicitly correlated coupled-cluster methods such as the linear R12 methods[96] could be used to evaluate $\Delta E^{\infty}_{\mathrm{CCSD(T)}}$ more accurately and reliably than by the extrapolation-based schemes used here. For $\Delta E_{\mathrm{CCSDTQ}}$, estimation of quadruple effects by a noniterative approximation would be desirable. Although there is no shortage of such noniterative approaches that have been suggested in the literature,[97] little has been done in the way of testing them for nontrivial systems. Recent work by some of us[98] has led to the derivation and implementation of a noniterative correction to CCSDT that corresponds—in the sense of perturbation theory—to the (T) correction to CCSD. It is also distinct from any other approach that has been advocated for correcting the CCSDT energy. One of the potential application areas for this ‘‘CCSDT(Q)’’ method will be in the area of thermochemistry. Alternatively, it is also sensible to dispense completely with the quadruples correction *if thermochemical parameters are calculated from (nearly) isodesmic reactions*, especially if accuracies better than 1 kJ mol^{-1} are not required. Other areas worthy of study are: correlation effects in the diagonal Born–Oppenheimer approximation[71,72] (where we have also recently focused some theoretical effort[99]), the importance of anharmonicity in the zero-point vibrational correction, and the importance of the thus far ignored constant G_0 contribution to ΔE_{ZPE}. Also of interest would be the use of CC calculations based on ROHF reference functions. Although we believe that differences in the final energies will be negligible, the partitioning of $\Delta E^{\infty}_{\mathrm{CCSD(T)}}$, $\Delta E_{\mathrm{CCSDT}}$, and $\Delta E_{\mathrm{CCSDTQ}}$ will undoubtedly change in interesting ways, which might render more systematic behavior in the individual contributions.

Studies along the lines suggested above will be carried out. The results, however, are predictable. Use of simplified HEAT strategies will likely result in a level of accuracy that is comparable to that obtained in other sophisticated theoretical model chemistries such as G3, CBS-Q, W2, and W3, which have been worked out, benchmarked, and documented in excellent detail in the literature. We are not interested in proliferating the population of roughly equivalent methods, so the community will not be subjected to follow-up papers reporting HEAT2, HEAT3, etc., methods. However, investigation of the approximations suggested above should be interesting, especially with regard to the utility of the noniterative approximation for quadruple excitation effects and for any light that it might shed on the magnitude and systematics of the error introduced by assuming that core and valence correlation effects can be separated.

ACKNOWLEDGMENTS

This work grew from a discussion involving three of the authors (A.G.C., P.G.S., and J.F.S.) in Budapest during a meeting of the IUPAC task force on free radical thermochemistry. We therefore thank IUPAC for providing the stimulus for this excursion into computational thermochemistry. Financial support for this work comes from the U.S. National Science Foundation and the Robert A. Welch Foundation (J.V., B.A.F., and J.F.S.), the Fonds der Chemischen Industrie (M.K. and J.G.), the Hungarian Research Foundation under OTKA Grant No. T047182 (P.G.S. and M.K.), and OTKA Grant No. T047185 (A.G.C.). During the latter stages of this research, P.G.S. was supported by the Fulbright Foundation during a sabbatical at the University of Texas at Austin. The research presented in this paper is part of a current and future work by a Task Group of the International Union of Pure and Applied Chemistry Grant No. (2000-013-2-100) to determine structures, vibrational frequencies, and thermodynamic functions of free radicals of importance in atmospheric chemistry. We also thank Branko Ruscic (Argonne National Laboratory) for his enthusiastic support of our research as well as guidance in the subtle world of high-accuracy thermochemistry.

APPENDIX: COMMENTS ON CALCULATIONS USING CCSDTQ AND BEYOND

We begin with a few comments on the cost of these calculations. The largest CCSDTQ calculation, which was carried out for the HCO radical using the cc-pVTZ basis set, involves the simultaneous solution of roughly 622×10^6 nonlinear equations. The most demanding CCSDTQP (which includes complete treatment of single, double, triple, quadruple, and pentuple excitations) calculation was carried out for H_2O_2, where the dimension of the nonlinear system was slightly greater than 1.3×10^9. These calculations required several weeks of computer time, and are clearly beyond the scope of anything that could remotely be characterized as a routine application. Despite this, none of the CCSDTQ calculations with the cc-pVDZ basis set (those specified in the HEAT protocol) required more than three days on LINUX-based personal computers.

To test the suitability of the approximation used for $\Delta E_{\mathrm{CCSDTQ}}$, additional CCSDTQ calculations were performed for a subset of molecules (the smaller ones, of course) using cc-pVTZ and cc-pVQZ basis sets, and at the CCSDTQP level with the cc-pVDZ basis. The effects of these on values of $\Delta_f H^\circ$ calculated according to method II (elemental reactions) are documented in Table VI. From this, it can be seen that the cc-pVDZ basis set (the only member of the cc-pV*XZ* hierarchy that can really be used for $\Delta E_{\mathrm{CCSDTQ}}$ for all of the molecules) gives results that are quite suitable for the present purposes. In fact, from an inspection of Table VII, it appears that the cc-pVDZ basis set is superior to the cc-pVTZ basis set, since it gives CCSDTQ - CCSDT differences that are closer to those based on cc-pVTZ/cc-pVQZ extrapolations for systems small enough to permit CCSDTQ/cc-pVQZ calculations. These calculations also indicate (see Table VI) that the cc-pVDZ contribution to the elemental reaction enthalpies are not very sensitive to further extension of the basis. The fact that the magnitude of the cc-pVDZ CCSDTQ - CCSDT energies appears to be slightly too small with respect to cc-pVTZ, cc-pVQZ, and extrapolated results was in fact the basis for an empirical scaling of quadruple excitation effects by Boese *et al.* in their W3 model.[25] However, we have not chosen to scale our calculated quadruple corrections.

Of perhaps greater interest is the effect of connected pentuple excitations. While CCSDTQP calculations are clearly at least a few decades from becoming routinely ap-

TABLE VI. Contribution of quadruple and pentuple excitations to enthalpies of formation in different basis sets. All values are in kJ mol^{-1}.

Species	CCSDTQ-CCSDT cc-pVDZ	CCSDTQ-CCSDT cc-pVTZ	CCSDTQP-CCSDTQ cc-pVDZ
C_2H_2	−2.27	−2.52	−0.54
C	−0.01	0.08	−0.11
CCH	−2.31	−2.62	−0.54
CF	0.70	0.83	0.04
CH_2	−0.15	−0.02	−0.12
CH	−0.14	−0.05	−0.11
CH_3	−0.25	−0.12	−0.12
CN	−1.64	−2.03	−0.40
CO_2	0.33		
F	1.71	1.66	0.08
H_2O_2	1.51		0.28
H_2O	1.24	1.52	0.21
HCN	−1.25	−1.40	−0.29
HCO	0.02	0.04	−0.02
HF	0.99	1.27	0.07
HNO	0.49		0.13
HO_2	1.94		0.27
N	1.80	1.94	0.24
NH_2	1.32	1.54	0.21
NH_3	1.08	1.30	0.19
NH	1.58	1.77	0.23
NO	0.96	0.95	0.12
O	2.23	2.22	0.24
OF	1.84	1.96	0.18
OH	1.75	1.92	0.23

plicable, their importance in accurate thermochemistry is apparently nonvanishing. With the cc-pVDZ basis set, it is interesting to note that the magnitude of the largest contributions of pentuples to enthalpies of formation calculated by procedure II (the elemental reaction approach) are roughly 0.5 kJ mol^{-1}. These large contributions found for acetylene, CCH, and CN can be applied for ‘‘superHEATed’’ reaction enthalpies of these species. Adding these contribution to those found in Table IV, one finds $\Delta_f H^\circ(C_2H_2) = 228.20$ kJ mol^{-1} versus the ATcT value of 228.20 ± 0.64 kJ mol^{-1} (!); $\Delta_f H^\circ(CCH) = 562.78$ kJ mol^{-1} versus the ATcT value of 563.32 ± 0.65 kJ mol^{-1}; and $\Delta_f H^\circ(CN) = 437.39$ kJ mol^{-1} versus the NIST-JANAF value of 436.8 ± 10 kJ mol^{-1}. Of some interest here, however, is Ruscic's recommendation[92] of $\Delta_f H^\circ(CN) = 438.5 \pm 4.0$ kJ mol^{-1}, based on relatively recent experiments.[100] Hence, it appears that pentuple excitation effects are responsible for the relatively poor performance of HEAT for acetylene, which is one of the two prominent outliers seen in Table IV (the other being H_2O_2).[101] Fortunately, there is ample evidence that effects of hextuple and higher excitations can be safely ignored. Some recent work[57,58,60,102] has indicated that contributions to molecular properties from connected excitations decrease by an order of magnitude upon each increase in the excitation level. This, together with the results documented in this appendix, suggests that neglect of higher than pentuple excitations results in errors of <0.05 kJ mol^{-1} (which, after all, is less than 5 cm^{-1}), at least for states that are relative free of strong nondynamic correlation effects. Examples such as ozone would be an interesting test, but unfortunately are too large to be calculated beyond CCSDTQ even with the cc-pVDZ basis set.

TABLE VII. Difference of CCSDTQ and CCSDT total energies (E_{CCSDT} -E_{CCSDTQ}) in different basis sets and with basis set extrapolation (μE_h).

	cc-pVDZ	cc-pVTZ	cc-pVQZ	DT extrapolation	TQ extrapolation
C	−30	−43	−44	−49	−45
F	−115	−60	−90	−37	−112
N	−42	−43	−57	−43	−67
O	−78	−51	−75	−40	−92
CH	−76	−93	−105	−99	−114
NH	−127	−107	−129	−98	−145
OH	−259	−167	−206	−128	−234
HF	−392	−209	−258	−132	−294
CH_2	−82	−80	−96	−80	−108
NH_2	−224	−196	−230	−185	−254
Error[a]	29	42	18	58	⋯

[a] Mean absolute error with respect to TQ extrapolated values.

[1] See, for example, the special issues of Spectrochim. Acta **58**, 599–898 (2002).

[2] For a representative and chemically relevant example, see E. R. Jochnowitz, M. R. Nimlos, M. E. Varner, G. B. Ellison, and J. F. Stanton, J. Am. Chem. Soc. (to be published).

[3] K. L. Bak, J. Gauss, P. Jørgensen, J. Olsen, T. Helgaker, and J. F. Stanton, J. Chem. Phys. **114**, 6548 (2001).

[4] For a review, see S. J. Blanksby and G. B. Ellison, Acc. Chem. Res. **36**, 255 (2003).

[5] B. Ruscic, J. E. Boggs, A. Burcat *et al.*, J. Phys. Chem. Ref. Data (in press).

[6] Well-known tabulations of molecular enthalpies of formation include: M. W. Chase, Jr., in *NIST-JANAF Thermochemical Tables*, 4th ed. Journal of Physical Chemical Reference Data, Monograph Vol. 9 (1998), p. 1; J. D. Cox, D. D. Wagman, and V. A. Medvedev, *CODATA Key Values for Thermodynamics* (Hemisphere Publishing Corp., New York, 1989); L. V. Gurvich, I. V. Veyts, and C. B. Alcock, *Thermodynamic Properties of Individual Substances*, 4th ed. (Hemisphere Publishing Company, New York, 1989).

[7] B. A. Flowers, P. G. Szalay, J. F. Stanton, M. Kállay, J. Gauss, and A. G. Császár, J. Phys. Chem. A **108**, 3195 (2004), and references therein.

[8] A. G. Császár, P. G. Szalay, and M. L. Leininger, Mol. Phys. **100**, 3879 (2002).

[9] A. G. Császár, M. L. Leininger, and V. Szalay, J. Chem. Phys. **118**, 10631 (2003).

[10] M. S. Schuurman, S. R. Muir, W. D. Allen, and H. F. Schaefer, J. Chem. Phys. **120**, 11586 (2004).

[11] S. Parthiban and J. M. L. Martin, J. Chem. Phys. **114**, 6014 (2001).

[12] Before the development of the Active Thermochemical Tables, atomic enthalpies of formation for O and N were also not known precisely. However, uncertainties in these values have been dramatically reduced in the ATcT approach, B. Ruscic (private communication).

[13] B. Ruscic, A. F. Wagner, L. B. Harding *et al.*, J. Phys. Chem. A **106**, 2727 (2002).

[14] A. L. L. East and W. D. Allen, J. Chem. Phys. **99**, 4638 (1993).

[15] D. A. Dixon, D. Feller, and G. Sandrone, J. Phys. Chem. A **103**, 4744 (1999).

[16] C. W. Bauschlicher, J. M. L. Martin, and P. R. Taylor, J. Phys. Chem. A **103**, 7715 (1999).

[17] J. M. L. Martin and G. de Oliveira, J. Chem. Phys. **111**, 1843 (1999).

[18] T. Helgaker, W. Klopper, A. Halkier, K. L. Bak, P. Jørgensen, and J. Olsen, *Quantum Mechanical Prediction of Thermochemical Data* (American Chemical Society, Washington DC, 1998).

[19] K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and W. Klopper, J. Chem. Phys. **112**, 9229 (2000).

[20] *Quantum Mechanical Prediction of Thermochemical Data*, edited J. Cioslowski (Kluwer, Dordrecht, 2001).

[21] W. J. Hehre, L. Radom, P. V. R. Schleyer, and J. A. Pople, *Molecular Orbital Theory* (Wiley, New York, 1986), p. 298.

[22] The reader is reminded that isodesmic reactions are those which would be

thermoneutral in the limit that additivity of ‘‘bond energies’’ gave a precise picture of chemical thermodynamics.

[23] Notable in this regard are the G(aussian) n ($n=1$, 2, and 3) family of methods developed by Pople and co-workers [K. Raghavachari and L. A. Curtiss in Ref. 20]; the W(eizmann) n ($n=1$, 2, and 3) methods of Martin and co-workers [Ref. 17 (W1); Ref. 11 (W2); Ref. 25 (W3)]; and the CBS methods of Petersson [G. A. Petersson, in Ref. 20]. A related approach is the focal-point strategy advocated by Allen and collaborators (Ref. 24); see Ref. 10 for a recent example.

[24] A. G. Császár, W. D. Allen, and H. F. Schaefer III, J. Chem. Phys. **1998**, 108 (9751).

[25] A. D. Boese, M. Oren, O. Atasolyu, J. M. L. Martin, M. Kállay, and J. Gauss, J. Chem. Phys. **120**, 4129 (2004).

[26] R. J. Bartlett, Annu. Rev. Phys. Chem. **32**, 359 (1981).

[27] An exception is our calculation of spin–orbit coupling, which uses a non-size-extensive approach.

[28] Throughout the text, the term ‘‘enthalpy of formation’’ refers to the difference in ground-state energies, i.e., the enthalpy of formation at 0 K.

[29] R. D. Cowan and M. Griffin, J. Opt. Soc. Am. **66**, 1010 (1976); R. L. Martin, J. Phys. Chem. **87**, 750 (1983).

[30] As is customary in the quantum chemistry world, the second row of the periodic table is taken to be that containing Li, Be, B, C, N, O, F, and Ne.

[31] H. Sellers and P. Pulay, Chem. Phys. Lett. **103**, 463 (1984).

[32] N. C. Handy, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. **84**, 4481 (1986).

[33] N. C. Handy and A. M. Lee, Chem. Phys. Lett. **252**, 425 (1996).

[34] W. Kutzelnigg, Mol. Phys. **90**, 909 (1997).

[35] It should be noted that inclusion of these effects shifts the position of the equilibrium internuclear distance. Nonetheless, we ignore this complication, and base all of our calculations on equilibrium geometries determined with CCSD(T)/cc-pVQZ. The issue of equilibrium geometry is not of numerical significance. In the vicinity of the true minimum, the energy changes only in second order when the nuclei are displaced. Taking a typical (and large) value for a bond stretching force constant (10 aJ Å^{-2}), an error of 0.005 Å in the equilibrium bond length translates to a calculated energy error of 0.08 kJ mol^{-1}.

[36] G. D. Purvis and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

[37] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).

[38] J. Noga and R. J. Bartlett, J. Chem. Phys. **86**, 7041 (1987); G. E. Scuseria and H. F. Schaefer, Chem. Phys. Lett. **132**, 382 (1988); J. D. Watts and R. J. Bartlett, J. Chem. Phys. **93**, 6104 (1993).

[39] S. A. Kucharski and R. J. Bartlett, Theor. Chim. Acta **80**, 387 (1991); J. Chem. Phys. **97**, 4282 (1992); N. Oliphant and L. Adamowicz, *ibid.* **94**, 1229 (1991).

[40] Optimized all-electron CCSD(T)/cc-pVQZ geometries used in all calculations: C_2H_2—$r_{CC}=1.204\,10$ Å, $r_{CH}=1.061\,26$ Å; C_2H—$r_{CC}=1.205\,30$ Å, $r_{CH}=1.061\,90$ Å; CF—$r_{CF}=1.270\,89$ Å; CH_2—$r_{CH}=1.074\,36$ Å, $\theta_{HCH}=133.73°$; CH_3—$r_{CH}=1.075\,11$ Å; CH—$r_{CH}=1.116\,67$ Å; CN—$r_{CN}=1.167\,39$ Å; CO—$r_{CO}=1.128\,91$ Å; CO_2—$r_{CO}=1.160\,13$ Å; F_2—$r_{FF}=1.411\,12$ Å; H_2—$r_{HH}=0.741\,86$ Å; H_2O—$r_{OH}=0.956\,23$ Å, $\theta_{HOH}=104.25°$; H_2O_2—$r_{OH}=0.960\,96$ Å, $r_{OO}=1.449\,52$ Å, $\theta_{HOO}=100.00°$, $\phi_{HOOH}=112.44°$; HCN—$r_{CN}=1.153\,78$ Å, $r_{CH}=1.064\,49$ Å; HCO—$r_{CO}=1.175\,86$ Å, $r_{CH}=1.116\,17$ Å, $\theta_{HCO}=124.57°$, HF—$r_{HF}=0.915\,16$ Å; HNO—$r_{NO}=1.208\,16$ Å, $r_{NH}=1.050\,95$ Å, $\theta_{HNO}=108.09°$; HO_2—$r_{HO}=0.968\,67$ Å, $r_{OO}=1.327\,76$ Å, $\theta_{HOO}=104.30°$; N_2—$r_{NN}=1.098\,09$ Å; NH—$r_{NH}=1.034\,72$ Å; NH_2—$r_{NH}=1.022\,59$ Å, $\theta_{HNH}=102.89°$; NH_3—$r_{NH}=1.010\,14$ Å, $\theta_{HNH}=106.36°$; NO—$r_{NO}=1.147\,88$ Å; NO_2—$r_{NO}=1.192\,85$ Å, $\theta_{NOO}=134.26°$; O_2—$r_{OO}=1.205\,77$ Å; OF—$r_{OF}=1.350\,11$ Å; OH—$r_{OH}=0.967\,98$ Å.

[41] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

[42] K. Kuchitsu, in *Accurate Molecular Structures*, edited by A. Domenicano and I. Hargittai (Oxford University Press, Oxford, 1992), pp. 14–46.

[43] R. A. Kendall and T. H. Dunning, J. Chem. Phys. **96**, 6796 (1992).

[44] See, for example, J. F. Stanton and J. Gauss, Adv. Chem. Phys. **125**, 101 (2003).

[45] P. G. Szalay, C. S. Simmons, J. Vázquez, and J. F. Stanton, J. Chem. Phys. **121**, 7624 (2004).

[46] D. E. Woon and T. H. Dunning, J. Chem. Phys. **103**, 4572 (1995).

[47] D. Feller, J. Chem. Phys. **96**, 6104 (1992).

[48] It can be shown that the atomic correlation energy converges to the exact result according to Eq. (6), if X is defined such that it corresponds to a basis set in which sets of orbitals associated with all angular momenta up to and including l ($=0$ for s, *1*, for p, etc.) are saturated. Use of this formula for molecules stands upon a rather flimsy theoretical foundation in the first place, due to the lack of spherical symmetry. It is however an even more extreme (albeit widely used) approximation to treat a basis set such as aug-cc-pCVTZ as though it is saturated with f functions for first-row atoms, since it has but two f functions. Nevertheless, extrapolation schemes based on the formula work well in practice and seem to be as good as any others that have been advocated for this purpose.

[49] T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997).

[50] To test the significance of the combined treatment of core and valence correlation, the total energy of C, O, CO, and CH_2 has been calculated at the frozen core CCSD(T) level extrapolated from aug-cc-pVQZ and aug-cc-pV5Z results. The core correlation effects have been calculated by comparing the extrapolated (Q5) frozen-core and all electron energies at the CCSD(T) level, the latter using aug-cc-pCVXZ basis sets. In all cases, this procedure gives lower total energies compared to the direct inclusion of core correlation effect as used in the HEAT protocol (0.23, 0.52, 0.17, and 0.50 kJ mol^{-1} for C, O, CH_2, and CO, respectively). From these data the enthalpy of formation of CO changes significantly (by 0.21 kJ mol^{-1}); that of CH_2 shifts by less than 0.1 kJ mol^{-1}.

[51] J. F. Stanton, Chem. Phys. Lett. **281**, 130 (1997).

[52] D. Feller, J. Chem. Phys. **111**, 4373 (1999).

[53] D. Feller and J. A. Sordo, J. Chem. Phys. **113**, 485 (2000).

[54] M. Kállay and J. Gauss, J. Chem. Phys. **120**, 6841 (2004).

[55] S. Hirata and R. J. Bartlett, Chem. Phys. Lett. **321**, 216 (2000).

[56] J. Olsen, J. Chem. Phys. **113**, 7140 (2000).

[57] M. Kállay and P. R. Surján, J. Chem. Phys. **115**, 2945 (2001).

[58] M. Kállay and P. R. Surján, J. Chem. Phys. **113**, 1359 (2000).

[59] S. Hirata, J. Phys. Chem. A **107**, 9887 (2003).

[60] T. A. Ruden, T. Helgaker, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **371**, 62 (2003).

[61] Test calculations were carried out for several species, including the CN and CCH radicals that are plagued by significant spin contamination at the UHF level. In no cases did the absolute energies obtained with ROHF and UHF reference functions at the CCSDT level differ by more than 0.1 kJ mol^{-1}.

[62] For NO, CCH, CF, CN, and FO, the $\Delta E_{\rm ZPE}$ contribution was obtained at the ROHF/CCSD(T) level with the cc-pVQZ basis set, at the corresponding equilibrium internuclear distances of 1.150 40; 1.208 39 (CH), and 1.062 14 (CO); 1.270 90, 1.171 85, and 1.351 09 Å, respectively]. All other contributions to the HEAT energy were obtained with UHF reference functions.

[63] I. M. Mills, in *Modern Spectroscopy: Modern Research*, edited by K. N. Rao and C. W. Matthews (Academic, New York, 1972), pp. 115–140.

[64] J. Gauss and J. F. Stanton, Chem. Phys. Lett. **276**, 70 (1997).

[65] P. G. Szalay, J. Gauss, and J. F. Stanton, Theor. Chim. Acta **100**, 5 (1998).

[66] J. F. Stanton and J. Gauss, Int. Rev. Phys. Chem. **19**, 61 (2000).

[67] H. H. Nielsen, in *Handbuch der Physik*, edited by S. Flügge (Springer, Berlin, 1959), Vol. 37, Part 1, p. 171. H. H. Nielsen, Rev. Mod. Phys. **23**, 90 (1951).

[68] D. G. Truhlar, J. Comput. Chem. **12**, 266 (1991).

[69] J. Vázquez and J. F. Stanton (unpublished). For asymmetric tops, G_0 is given in lowest order, by the expression

$$G_0=-\frac{1}{4}\sum_{\alpha} B_e^{\alpha}+\frac{1}{64}\sum_{k}\phi_{kkkk}-\frac{7}{576}\sum_{k}\frac{\phi_{kkk}^2}{\omega_k}+\frac{3}{64}\sum_{k\neq l}\frac{\omega_l\phi_{kkl}^2}{4\omega_k^2-\omega_l^2}$$

$$-\frac{1}{4}\sum_{k\neq l\neq m}\frac{\phi_{klm}^2\omega_k\omega_l\omega_m}{D_{klm}}-\frac{1}{2}\sum_{k\neq l}\sum_{\alpha}B_e^{\alpha}(\zeta_{kl}^{\alpha})^2,$$

where B_e^{α} is the equilibrium rotational constant corresponding to the αth inertial axis, ϕ_{ijk} and ϕ_{ijkl} are cubic and quartic force constants in the dimensionless normal coordinate representation, ω_i is the harmonic frequency of normal mode i, ζ_{kl}^{α} is the Coriolis coupling constant between modes k and l with respect to the αth inertial axis, and D_{klm} is defined as $(\omega_k+\omega_l+\omega_m)(\omega_k-\omega_l-\omega_m)(\omega_k+\omega_l-\omega_m)(\omega_k-\omega_l+\omega_m)$. This equation has been checked by a direct summation procedure, in which the rovibrational Hamiltonian is constructed explicitly and $E_0^{(2)}\equiv\langle 0^{(1)}|H_1|0^{(0)}\rangle+\langle 0^{(0)}|H_2|0^{(0)}\rangle$ (H_n and $0^{(n)}$ are the nth/order Hamiltonian and ground state wave function, respectively) has been evaluated

numerically. Our result also agrees with the unpublished result of Allen and co-workers [W. D. Allen (private communication)]. It should be noted that this equation differs from that published in Ref. 70, although the numbers in that paper are consistent with our expression rather than that which was published. We have also derived higher-order terms contributing to G_0, but they are of no significance for the hydrogen peroxide example addressed later in this paper.

[70] V. Barone, J. Chem. Phys. **120**, 3059 (2004).

[71] D. W. Schwenke, J. Phys. Chem. A **105**, 2352 (2001).

[72] E. F. Valeev and C. D. Sherrill, J. Chem. Phys. **118**, 3921 (2003).

[73] O. L. Polyansky, A. G. Császár, S. V. Shirin, N. F. Zobov, P. Barletta, J. Tennyson, D. W. Schwenke, and P. J. Knowles, Science **299**, 539 (2003).

[74] P. Pyykkö, Adv. Quantum Chem. **11**, 353 (1978); P. Pyykkö, Chem. Rev. (Washington, D.C.) **88**, 563 (1988); K. Balasubramanian, *Relativistic Effects in Chemistry, Part A: Theory and Techniques and Part B: Applications* (Wiley, New York, 1997).

[75] J. L. Tilson, W. C. Ermler, and R. M. Pitzer, Comput. Phys. Commun. **128**, 128 (2000).

[76] R. M. Pitzer (private communication), see also http://www.chemistry.ohio-state.edu/pitzer

[77] W. C. Ermler, R. B. Ross, and P. A. Christiansen, Adv. Quantum Chem. **19**, 139 (1988), and references therein.

[78] E. R. Davidson, Y. Ishikawa, and G. L. Malli, Chem. Phys. Lett. **84**, 226 (1981).

[79] H. Lischka, R. Shepard, I. Shavitt *et al.*, COLUMBUS, an *Ab Initio* Electronic Structure Program, release 5.9, 2003.

[80] T. D. Crawford, C. D. Sherrill, E. F. Valeev *et al.*, Psi 3.2, 2003. Freely available at http://www.psicode.org/.

[81] J. F. Stanton, J. Gauss, J. D. Watts, W. J. Lauderdale, and R. J. Bartlett, Int. J. Quantum Chem. **S26**, 879 (1992).

[82] J. F. Stanton, C. L. Lopreore, and J. Gauss, J. Chem. Phys. **108**, 7190 (1998).

[83] For a discussion of CCSDT in the context of atomization energies, see: K. L. Bak, P. Jørgensen, J. Olsen, T. Helgaker, and J. Gauss, Chem. Phys. Lett. **317**, 116 (2000).

[84] Especially when there is strong spin contamination or symmetry breaking (equivalent to unphysical spin localization when appropriate symmetry elements are not present). In this context, it is not surprising to note that CN has by far the largest post-CCSD(T) correction, as it is the most significantly spin-contaminated species in the study. An ROHF-based $E^{\infty}_{\mathrm{HF}}+\Delta E^{\infty}_{\mathrm{CCSD(T)}}$ analysis would be interesting in this context, but is beyond the scope of this initial presentation and study of the HEAT approach.

[85] J. Gauss, W. J. Lauderdale, J. F. Stanton, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **182**, 207 (1991).

[86] J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. **98**, 8718 (1993).

[87] Atomic enthalpies of formation (0 K) were taken from Ruscic's Active Thermochemical Tables (ATcT): H ($216.034\pm0.000\ \mathrm{kJ\,mol^{-1}}$); C ($711.79\pm0.21\ \mathrm{kJ\,mol^{-1}}$); N ($470.592\pm0.045\ \mathrm{kJ\,mol^{-1}}$); O ($246.844\pm0.002\ \mathrm{kJ\,mol^{-1}}$); F ($77.21\pm0.24\ \mathrm{kJ\,mol^{-1}}$). These values have considerably smaller error bars than those from the CODATA compilation, and have been chosen in this work for that reason.

[88] Enthalpies of formation for carbon and fluorine atoms given in the NIST-JANAF database (see Ref. 6) are $711.19\pm0.46\ \mathrm{kJ\,mol^{-1}}$ and $77.28\pm0.30\ \mathrm{kJ\,mol^{-1}}$). Note that these uncertainties are larger than the mean absolute errors in the enthalpies of formation determined with HEAT. Recent work by some of us (B. Ruscic, A. G. Császár, and J. F. Stanton, to be published) has focused on an improved estimate for $\Delta_f H^\circ(C_{\mathrm{gas}})$. Note the relatively large difference between the NIST-JANAF value and that from the active tables (711.79 vs $711.19\ \mathrm{kJ\,mol^{-1}}$) for carbon.

[89] The idea of the Active Thermochemical Tables is presented in detail in: G. von Laszewski, B. Ruscic, P. Wagstrom *et al.*, in *Lecture Notes in Computer Science*, edited by M. Parashar (Springer, Berlin, 2002), Vol. 2536, pp. 25–38.

[90] The ATcT results presented in this paper were privately communicated by B. Ruscic, and were obtained from Active Thermochemical Tables, ver. 1.20 operating on the Core (Argonne) Thermochemical Network ver. 1.032 (2004). A small subset of these values (H, O, H_2O, CO_2, F, and HF) will appear in: B. Ruscic, R. E. Pinzon, M. L. Morton, G. V. Laszevski, S. J. Bittner, S. G. Nijsure, K. A. Amin, M. Minkoff, and A. F. Wagner, J. Phys. Chem. A. **108**, 9979 (2004).

[91] A very basic description of the Active Thermochemical Tables is also scheduled to appear as: B. Ruscic, in *Active Thermochemical Tables*, McGraw-Hill 2005 Yearbook of Science and Technology (McGraw-Hill, New York, 2004).

[92] B. Ruscic (private communication).

[93] For H_2O_2, the constant contribution to the zero-point energy [G_0 in Eq. (10)] appears to be important. For the elemental formation reaction, $H_2+O_2\rightarrow H_2O_2$, the calculated contribution (J. Vázquez and J. F. Stanton, unpublished) of G_0 at the CCSD(T)/DZP level of theory is $-32.7\ \mathrm{cm}^{-1}$, or about $-0.4\ \mathrm{kJ\,mol^{-1}}$. When this is added to the HEAT value of $\Delta_f H^\circ\ (H_2O_2)$, agreement with experiment is improved significantly. Note also that an independent variational estimate [M. Mladenovic, Spectrochim. Acta **58A**, 809 (2002)] of the ZPE is $22.4\ \mathrm{cm}^{-1}$ below the value in Table I.

[94] For a discussion concerning CCH, see P. G. Szalay, L. S. Thøgersen, J. Olsen, M. Kállay, and J. Gauss, J. Phys. Chem. A **108**, 3030 (2004).

[95] One could even question why diffuse functions are used in the HEAT approach. If indeed the extrapolation formulas gave the correct result for any sequence of hierarchical basis sets (cc-pCVXZ or aug-cc-pCVXZ), then the use of diffuse functions in the HEAT method would offer no benefit whatsoever. However, the extrapolation formulas are essentially empirical in nature (see Ref. 48). Also, if diffuse functions are of importance, then the sequence of basis sets containing them should converge to the exact result more rapidly than that which omits them.

[96] J. Noga, W. Klopper, and W. Kutzelnigg, in *Recent Advances in Coupled-Cluster Methods* (World Scientific, Singapore, 1997), p. 1.

[97] S. A. Kucharski and R. J. Bartlett, Chem. Phys. Lett. **158**, 550 (1989); J. Chem. Phys. **108**, 9221 (1998).

[98] Y. J. Bomble, M. Kállay, J. Gauss, and J. F. Stanton (unpublished).

[99] A. Tatji, M. Kállay, P. G. Szalay, J. Gauss, and J. F. Stanton (unpublished).

[100] J. Berkowitz, G. B. Ellison, and D. Gutman, J. Phys. Chem. **98**, 2744 (1994).

[101] It should be noted that calculation of $\Delta_f H^\circ$ for C_2H_2 using the second and third isodesmic reaction in Table V together with ATcT values of H_f^0 for CCH, CH_2, and CH yields 227.70 and $229.32\ \mathrm{kJ\,mol^{-1}}$, respectively, which clearly differ appreciably. The discrepancy is removed to some degree by using $\Delta_f H^\circ$ values for CH and CH_2 from Császár and co-workers (Refs. 8 and 9) (592.47 and $390.45\ \mathrm{kJ\,mol^{-1}}$, respectively), with the former notably outside the ATcT range, which gives $\Delta_f H^\circ\ (C_2H_2) = 228.16\ \mathrm{kJ\,mol^{-1}}$ (second reaction in Table V) and $229.52\ \mathrm{kJ\,mol^{-1}}$ (third reaction). Error in these calculations arises almost exclusively from uncertainities in the values used in the subtraction procedure; the calculated reaction enthalpies are unlikely to be in error by more than $0.5\ \mathrm{kJ\,mol^{-1}}$. This example illustrates the pros and cons of using reaction-based approaches, where the theoretically calculated contribution can be excellent but the experimental uncertainty annoying, and the elemental formation reaction where there is absolutely no uncertainty in the "experimental" numbers, but the theoretical calculations are more demanding. Note that pentuple excitations, essential for the calculation of $\Delta_f H^\circ\ (C_2H_2)$ from elemental reactions, is not at all important with the isodesmic reaction.

[102] M. Kállay, P. G. Szalay, and J. Gauss, J. Chem. Phys. **119**, 2991 (2003).

J. Chem. Theory Comput. **2009,** *5,* 2687–2693

Journal of Chemical Theory and Computation

Unrestricted Coupled Cluster and Brueckner Doubles Variations of W1 Theory

Ericka C. Barnes,[†] George A. Petersson,[*,†] John A. Montgomery, Jr.,[‡,¶]
Michael J. Frisch,[‡] and Jan M. L. Martin[¶,§]

Hall-Atwater Laboratories of Chemistry, Wesleyan University, Middletown, Connecticut 06459-0180, Gaussian, Incorporated, 340 Quinnipiac Street, Building 40, Wallingford, Connecticut 06492-4050, and Department of Organic Chemistry, Weizmann Institute of Science, IL-76100 Reovot, Israel

Received May 20, 2009

Abstract: Unrestricted coupled cluster spin contamination corrected [UCCSD(T)] and unrestricted Brueckner doubles [UBD(T)] variations of the Weizmann-1 theory (W1), denoted as W1U, W1Usc, and W1BD, respectively, are compared with the restricted open-shell W1 theory [W1(RO)]. The performances of the four W1 variants are assessed with 220 total atomization energies, electron affinities, ionization potentials, and proton affinities in the G2/97 test set, for consistency with the error analysis of the original W1(RO) study. The root-mean-square deviations from the experiment of W1U (0.65 ± 0.48 kcal/mol), W1Usc (0.57 ± 0.48 kcal/mol), W1BD (0.62 ± 0.48 kcal/mol), and W1(RO) (0.57 ± 0.48 kcal/mol) show that the four methods are virtually indistinguishable. This error analysis excludes the "singlet biradicals," C_2 and O_3, since single determinantal methods are not really adequate for these strongly multireference systems. The unrestricted W1 variants perform poorly for such highly spin-contaminated and multireference species (the largest deviation from experiment for W1Usc is −4.2 ± 0.1 kcal/mol for the O_3 EA). W1(RO) performs much better than its unrestricted counterparts for these pathological cases (the deviation from experiment is reduced to −1.5 ± 0.1 kcal/mol for the O_3 EA), though the errors are significantly larger than those for the overall test set. The examples of C_2, O_3, and the F_2 potential energy curve indicate that an advantage to using W1BD is that the error in $\langle S^2 \rangle$ correlates with the magnitude of the error in energy, whereas W1(RO) loses accuracy without such a warning.

I. Introduction

Advances in computational methods and computer hardware have made possible the accurate ab initio calculation of energies for small- and medium-size molecules. Combined with Pople's model chemistry concept, these calculations provide reliable thermochemical predictions, which are a significant achievement of modern computational chemistry.[1]

A "theoretical model chemistry" is a complete algorithm for the calculation of the energy of any molecular system.[2,3] It cannot involve subjective decisions in its application. It must be size extensive, giving energies that are additive for separated systems, so that the energy of every molecular species is uniquely defined. A model chemistry is useful if for some class of molecules it is the most accurate calculation we can afford to do. A number of "black-box" computational methods have emerged in the past two decades, through the development of composite theoretical model chemistry methods such as the complete basis set (CBS) model chemistries of Petersson et al.,[4−8] the Gaussian-*n* methods of Pople and co-workers,[9−12] the Weizmann-*n* (W*n*) theories of Martin and co-workers,[13−17] the high-accuracy extrapo-

---

[*] Corresponding author. E-mail: gpetersson@wesleyan.edu.

[†] Wesleyan University.

[‡] Gaussian, Incorporated.

[¶] Weizmann Institute of Science

[§] Current address: Department of Physics, University of Connecticut, 2152 Hillside Road, U-3046, Storrs, CT 06269-3046

10.1021/ct900260g CCC: $40.75

lated ab initio thermochemistry (HEAT) protocol of the Gauss and Stanton groups,[18−20] and the correlation-consistent composite approach (ccCA) of Wilson and co-workers,[21,22] to name a few.

The W*n* computational protocols of Martin and co-workers offer a sequence of models of increasing cost and accuracy,[23−26] the converging hierarchy of which currently ranges from W1 to W4.4. W1 theory, the most computationally accessible member of the evolving W*n* family, is often employed as a benchmark for more approximate methods in the absence of accurate experimental data. Some of the key accomplishments of W1 theory are:
i. It achieves 0.44 kcal/mol mean absolute deviation (0.56 kcal/mol rms deviation) for 220 total atomization energies (TAEs), electron affinities (EAs), ionization potentials (IPs), and proton affinities (PAs) of the G2/97 set;
ii.As a properly defined theoretical model chemistry, it is applicable in a "black-box" manner by a nonspecialist; and
iii. It is completely devoid of parameters adjusted to fit experimental data.

The present study compares the performance of several unrestricted variants of W1 theory: unrestricted W1 (W1U), W1U with a spin contamination correction (W1Usc), and unrestricted Brueckner doubles W1 (W1BD). Standard W1 will be explicitly referred to as W1(RO) in this study, in order to avoid confusion in the comparison.

II. Restricted vs Unrestricted Reference

One of the major issues in computational studies of open-shell species is the selection of a restricted or an unrestricted reference wave function. Each has its well-known advantages and disadvantages. For example, restricted open-shell coupled cluster singles and doubles[27−31] with perturbative triples,[32−35] [ROCCSD(T)], dissociate to the wrong energy limit, while the unrestricted coupled cluster [UCCSD(T)] wave function dissociates to the correct energy limit, but the wave function becomes significantly spin contaminated as a bond dissociates. Several approaches have been used through the years to alleviate the spin contamination problem. For example, the spin correction term in W1Usc was introduced for this purpose.[36]

Whether one chooses a restricted or an unrestricted reference determinant, it is best to be consistent. Although it is common practice in the application of unrestricted methods to treat many unrestricted Hartree−Fock (UHF)-unstable species (e.g., F_2, alkenes, or polyenes) with a restricted reference, problems arise in *reactions* involving these "closed-shell" molecules. For example, the methyl C−H bond dissociation energy (BDE) of UHF-unstable 1-butene gives a radical product, $H_2C{=}CHCH_2\dot{C}H_2$, that would be described by an unrestricted determinant. Employing a restricted reference for 1-butene and an unrestricted reference for the 1-buten-4-yl radical would create consistency problems between reactant and product energies and, thus, generate spurious BDE contributions. No matter how distant a radical center is, an unrestricted treatment of the radical will induce spin polarization of the UHF-unstable π-bond. These problems are eliminated if one employs a restricted open-shell Hartree−Fock (ROHF) reference for the radical or a UHF reference for the parent alkene.

Another alternative to unrestricted Hartree–Fock (HF) and coupled cluster is the use of charge coupled device (CCD) calculations with Brueckner orbitals,[37−42] coined as "Brueckner doubles" (BD) by Handy et al.[41] This method employs a reference configuration, BDRef (from which the singles coefficients are zero for the coupled cluster wave function truncated at the doubles level), in place of the HF reference typically employed in the coupled cluster ansatz. The Brueckner condition implies that the corresponding reference determinants give the best overlap of a one-configuration approximation of the wave function with the exact wave function.[39] Such a reference was introduced by its namesake Brueckner in 1954, as part of a self-consistent method in nuclear physics,[37] and was first employed by Nesbet[38] for use in the configuration interaction (CI) expansion of a wave function. In the field of computational chemistry, the resurgence of interest in Brueckner orbitals in the past two decades was prompted by Chiles and Dykstra[40] and later by Handy and co-workers.[41] Similar studies that demonstrate the robustness of BD as an alternative to its coupled cluster counterpart[42−45] have been carried out over the years. Since the species considered in the previous paragraph (i.e., F_2 and alkenes) are not UBD-unstable, the problems of consistency between the reactant and the radical product are also eliminated by using a BD-based method.

III. Computational Details

A. Components of W1. The justification for the selection of each component calculation in W1(RO) has been presented in detail[14,15] and will not be repeated here. The W1U and W1BD methods retain the essential features of standard W1(RO) but replace the sequence of ROHF, ROCCSD, and ROCCSD(T) calculations with their spin-unrestricted counterparts for W1U for calculations involving open-shell species and with the BDRef, BD, and BD(T)[46] sequence for W1BD theory. All calculations were carried out with Gaussian 09,[47] which uses the ROCCSD(T) definition of UCCSD(T) in a basis of semicanonicalized ROHF orbitals,[35,48] with the semicanonicalization carried out before the integral transformation. The triples contribution to BD(T) is evaluated with semicanonical Brueckner orbitals.

The UB3LYP/cc-pVTZ+1d geometry optimization and frequency calculations are retained from the original W1(RO) implementation. Optimized geometries and energy components for all species are available in the accompanying "geometries.txt" and "components.pdf" files as Supporting Information. In the original implementation of W1(RO), scalar relativistic corrections were obtained with the Martin−Taylor small (MTsmall) basis sets[49,50] as one-electron Darwin and mass−velocity terms[51,52] from averaged coupled pair wave functions,[53] while corresponding components in this study were obtained with Douglas−Kroll−Hess (DKH) second-order scalar relativistic calculations[54−58] using a Gaussian nuclear model[59] (also employing the MTsmall basis sets). Spin−orbit calculations were taken directly from ref 14. The spin correction term in W1Usc,

$$\Delta E(\text{spin}) = -6.28mE_h \times \Delta\langle S^2\rangle_{\text{UHF}} \quad (1)$$

minimizes the difference between ROCCSD(T) and UCCSD(T) energies for some highly spin-contaminated species.[36]

B. Implementation of Brueckner Doubles. The BD algorithm employed throughout this study involves macroiterations to update the orbitals, wherein each macroiteration involves an integral transformation and a CCD calculation. Thus, BD calculations appear, at first glance, to be significantly more expensive than the corresponding CCSD calculations. In practice, however, W1BD calculations are only slightly more expensive than the corresponding ones in W1U or W1(RO). The two main reasons for this are as follows:

i. Since W1 involves a sequence of calculations, the converged orbitals and amplitudes from each step can be used to start the next. As a result, most of the extra CC iterations required are performed in the first BD(T) calculation, which uses the smallest basis set in the sequence of calculations. For the largest basis set (i.e., the BD/aug-cc-pVQZ+2df step), typically only two or three BD macroiterations are required, and the second and later macroiterations require only a few CC iterations.

ii. Calculations involving even three heavy atoms spend a significant amount of time in the (T) steps, (the only $O(N^7)$ parts of W1), and calculations on systems having four or more heavy atoms are dominated by these steps. The triples calculations have the same cost for CCSD(T) and BD(T).

The result is that the extra cost of doing W1BD over W1U or W1(RO) is in the range of 20−40% for two heavy-atom systems, about 20% for three heavy-atom systems, and going down further to less than 20% for systems with more than three heavy atoms and/or those involving second-row atoms and beyond. (W1 calculations involving two or more second-row or heavier atoms are dominated by the last calculation, which includes triples and correlates core electrons. As noted previously, the triples part of this step has the same cost for CCSD(T) and BD(T), and the BD iterations converge quickly since they use the orbitals, amplitudes, and basis set from the preceding frozen-core calculation as an initial guess.)

All but one of the correlation energy calculations in W1 use the frozen-core (FC) approximation, which substantially speeds up the calculations as compared to correlating all electrons (Full). Previous papers on the BD method have not discussed the issue of frozen-core with this model, and some programs, such as Gaussian 03, freeze the core orbitals at their initial values during BD iterations. This means that the converged BD(FC) energy varies with different initial guess orbitals. For example, in CN, the UHF solution is highly spin contaminated, while the BD(Full) reference determinant has very minimal spin contamination. Hence, a BD(FC) calculation starting from the UHF orbitals will freeze a core that is more spin polarized than that of a BD(Full) calculation and produce a different energy than, say, a BD(FC) calculation that starts with (also much less spin contaminated) B3LYP orbitals. The consistent approach to frozen-core BD calculations is to update all orbitals, including the core, during the BD iterations but to restrict the amplitudes in the CCD calculations to those involving only valence orbitals. The results are then independent of the initial orbitals, and the core is spin polarized only to the extent that the BD valence is. This approach is used in all BD results reported here.[47]

IV. Results and Discussion

Since the four flavors of W1 theory are virtually the same for closed-shell species, we shall focus on their performance for open-shell systems. The extent to which the ROHF and UHF reference determinants differ can be measured by the error in $\langle S^2\rangle$ for the reference configuration.

A. Spin-Contaminated Species. The restricted and unrestricted variations of W1 theory have been evaluated with bond dissociation enthalpies (BDEs), EAs, and IPs of some radicals and biradicals (Table 1). The sample of 11 reactions was selected on the basis of sizable errors in $\langle S^2\rangle_{\text{UHF}}$ and the availability of reliable experimental data.[60−67]

The rms deviation from experiment (Table 1) is not surprisingly the largest for W1U (1.3 ± 0.3 kcal/mol), in the absence of any rectification of the spin contamination problem. The spin correction in W1Usc indeed reduces the rms error to 0.6 ± 0.3 kcal/mol but is not always reliable. For example, ΔE(spin) in eq 1 overestimates the correction for the first C−H bond dissociation in acetylene by 0.8 ± 0.02 kcal/mol and worsens the deviation with experiment of the first C−H bond dissociation of propene by 0.8 ± 0.4 kcal/mol (Table 1). The rms deviations for W1U, W1Usc, W1BD, and W1(RO) are 1.3 ± 0.3, 0.6 ± 0.3, 0.8 ± 0.3, and 0.6 ± 0.3 kcal/mol, respectively, demonstrating the comparable accuracy of W1Usc, W1BD, and W1(RO) in handling these UHF spin-contaminated species. Note that both W1BD and W1(RO) attain these results in the *absence* of the empirical spin correction in W1Usc.

The following interesting observations are made for the three most spin-contaminated species of the G2/97 test sets, C_2, O_3, and CS^+, selected on the basis of $\Delta\langle S^2\rangle_{\text{UHF}}$ values greater than 0.6:[68]

i. First is the existence of multiple solutions, which complicates the use of "black-box" methods. To compound the problem of multiple solutions for both restricted and unrestricted versions of both HF and Brueckner determinants for C_2, the relative energies of the solutions are switched between the HF vs the CCSD(T) levels of theory for both restricted and unrestricted reference configurations. That a single reference method produces several solutions comes as no surprise in these cases, since C_2 and O_3 are known to possess a significant multiconfigurational character in their wave functions. If there are indications of near degeneracies of configurations (such as a large error in $\langle S^2\rangle$ for UHF or UBDRef or a negative eigenvalue for a virtual orbital as in neutral C_2), then it is necessary to explore multiple solutions in addition to checking for the correct number of imaginary frequencies. A more thorough discussion is presented as Supporting Information (see "Problem_species.pdf" file), which includes an example of solutions crossing between the BD vs BD(T) levels of theory.

ii. Second, the reactions involving these severely spin-contaminated systems indicate that $\Delta\langle S^2\rangle_{\text{UHF}}$ values do not correlate as well with the energy error in a W1U calculation

Table 1. Deviations (Experiment−theory) of Calculated BDEs and adiabatic IPs and EAs from Experiment of Selected Spin-Contaminated Species, in kcal/mol

		W1U $\Delta\langle S^2\rangle$		W1BD $\Delta\langle S^2\rangle$		deviation (experiment−theory)				experiment
species[a]	reaction	reactant	product	reactant	product	W1U	W1Usc	W1BD	W1RO[b]	
		σ-Radicals								
BDE ΔH_{298}	$\mathrm{H{-}CN \rightarrow H + \cdot C{\equiv}N}$	0.00	0.37	0.00	0.01	−1.9	−0.4	−1.0	−0.8	126.29 ± 0.2[c]
	$\mathrm{H{-}C{\equiv}CH \rightarrow H + \cdot C{\equiv}CH}$	0.00	0.35	0.00	0.03	−0.6	0.8	0.0	0.1	133.46 ± 0.02[d]
	$\mathrm{H{-}CH{=}CH_2 \rightarrow H + \cdot CH{=}CH_2}$	0.00	0.18	0.00	0.01	−0.4	0.3	−0.1	−0.1	110.71 ± 0.6[c]
IP ΔE_0	$\mathrm{C{=}O \rightarrow e^- + C{=}O^+}$	0.00	0.21	0.00	0.02	−0.9	−0.1	−0.6	−0.2	323.17 ± 0.01[e]
	$\mathrm{N{\equiv}C{-}C{\equiv}N \rightarrow e^- + N{\equiv}C{-}C{\equiv}N^+}$	0.00	0.35	0.00	0.03	−1.8	−0.4	−1.3	−1.0	308.42 ± 0.2[f]
EA ΔE_0	$\mathrm{CH_2{=}C{=}CH^- \rightarrow e^- + CH_2{=}C{=}CH}$	0.00	0.21	0.00	0.03	−0.7	0.1	−0.4	−0.3	20.59 ± 0.1[f]
	$\mathrm{C{=}N \rightarrow e^- + C{=}N}$	0.00	0.37	0.00	0.01	−1.8	−0.3	−1.0	−0.7	89.06 ± 0.1[e]
		π-Radicals								
BDE ΔH_{298}	$\mathrm{HCH_2{-}CH{=}CH_2 \rightarrow H + \cdot CH_2{-}CH{=}CH_2}$	0.00	0.20	0.00	0.04	0.5	1.3	0.8	0.9	88.79 ± 0.4[g]
EA ΔE_0	$\mathrm{CH_2CHCH_2^- \rightarrow e^- + CH_2CHCH_2}$	0.00	0.20	0.00	0.04	−0.5	0.3	−0.3	−0.1	11.09 ± 0.2[f]
		Singlet Biradicals								
BDE ΔH_{298}	$\mathrm{F{-}F \rightarrow F + F}$	0.30	0.01	0.00	0.01	1.8	0.6	1.1	1.0	38.00 ± 0.2[f]
		Triplet Biradicals								
EA ΔE_0	$\mathrm{H{-}C{-}C{\equiv}N^- \rightarrow e^- + H{-}C{-}C{\equiv}N}$	0.12	0.36	0.02	0.07	−0.5	0.5	−0.3	−0.2	46.20 ± 0.3[h]
MAD						1.1	0.5	0.6	0.5	± 0.2
rms						1.3	0.6	0.8	0.6	± 0.3
LD						−2.1	1.3	−1.3	−1.0	

[a] Geometries are optimized at the UB3LYP/cc-pVTZ+1d level and available as Supporting Information. [b] W1(RO) energies were calculated with Gaussian 09 (ref 47). [c] ref 60. [d] ref 61. [e] ref 62. [f] ref 63. [g] ref 64. [h] ref 65.

as $\Delta\langle S^2\rangle_{\mathrm{UBDRef}}$ values do for W1BD errors (vide infra). A sizable $\Delta\langle S^2\rangle_{\mathrm{UBDRef}}$ value appears to be a useful warning of when to be skeptical of W1BD results (refer to Supporting Information Table S−III and discussion on pages 11−12 of "Problem_species.pdf" file for details).

B. G2/97 Test Set. The G2−1[9−11] and the G2−2[69,70] test sets, collectively referred to as the G2/97 data set of Curtiss et al. were employed to calibrate the accuracy of W1(RO) for EAs, IPs, and PAs, while the G2−1 test set and a subset of the G2−2 data set (26 out of 93) of heats of formation were used in the calibration of total atomization energies (TAEs).[14] The selection of TAEs, EAs, IPs, and PAs is retained in assessing the performance of W1U, W1Usc, and W1BD, in order to facilitate comparisons with W1(RO) values in the literature.[14,15] A thorough discussion on the notable discrepancies with experiment for some problematic cases has already been presented by one of the authors.[14] A comprehensive breakdown of the G2−1 and the G2−2 TAE, EA, IP, and PA energetic components and error statistics for W1U, W1Usc, W1BD, and W1(RO) are available as Supporting Information. The error analyses for 220 reactions from the G2/97 data set are summarized in Table 2. The reported uncertainties in the deviations from experiment represent the uncertainties in the experimental data.[36]

The overall G2/97 rms errors for W1U, W1Usc, W1BD, and W1(RO) (excluding C_2 and O_3) are 0.65, 0.57, 0.62, and 0.57 ± 0.48 kcal/mol, respectively. Comparison of the calculated energies with experiment is rather problematic due to the large experimental uncertainties (± 0.65 for TAEs, ± 0.32 for EAs, and ± 0.35 kcal/mol for IPs, Table 2). The experimental uncertainties are of greater magnitude than the differences in rms errors between the methods themselves, making W1 variants virtually indistinguishable from one another for the G2/97 test set (Table 2).

C. F_2 Potential Energy Curve. Although the W*n* methods are intended for the thermochemistry of molecular systems at equilibrium geometries, the potential energy curve for the F_2 $^1\Sigma_g^+$ ground-state dissociation provides insight into the differences between the UCCSD(T), the spin-corrected UCCSD(T), the UBD(T), and the ROCCSD(T) levels of theory. Deviations of the potential energy curves from the exact solution in the cc-pVDZ basis are shown in Figure 1. A constant geometry-independent shift of the energy from the full configuration interaction (FCI) reference would display as zero error throughout the potential energy curve.

Neither restricted nor unrestricted CCSD(T) give a reliable description of bond-breaking reactions. These qualitative features of restricted and unrestricted methods along a potential energy curve for bond dissociation are well-known.[71] The spin correction successfully reduces the UCCSD(T) error in energy but shows fluctuation and dissociates improperly, reminiscent of problems encountered with a restricted reference (Figure 1). Furthermore, Ochterski et al. pointed out that the spin contamination error increases linearly with $\Delta\langle S^2\rangle_{\mathrm{UHF}}$ for $\Delta\langle S^2\rangle_{\mathrm{UHF}}$ less than 0.6.[68] This is consistent with our observations for C_2, O_3, and CS^+ (see "Problem_species.pdf" file given as Supporting Information) and strongly suggests that eq 1 is also not applicable for $\Delta\langle S^2\rangle_{\mathrm{UHF}} > 0.6$ on the F_2 potential energy curve ($\Delta\langle S^2\rangle_{\mathrm{UHF}} = 0.6$ is marked by the vertical dashed line at 1.5 Å in Figure 1).

The Brueckner doubles approach possesses compensating advantages over both the RHF- and the UHF-based methods. First, there is very good agreement between the UBD(T) and the ROCCSD(T) potential energy curves in the vicinity of the equilibrium geometry, where UCCSD(T) is already contaminated with unwanted contributions from higher spin multiplicities. This efficacy of Brueckner orbitals in reducing spin contamination over a wider range of geometries compared to UHF-based methods is also well documented.[72] Furthermore, the Brueckner doubles curve is practically indistinguishable from that of ROCCSD(T), up to the point where the UBD(T) energy errors exhibit extreme sensitivity

Table 2. Error Analysis (kcal/mol) for TAEs (298 K), EAs, IPs, and PAs of Select Molecules in the G2/97 Test Set for the Different Variations of W1 Theory[a]

G2/97 subset[b]	method	MAD	rms	LD	species/LD
TAE_{298K} (81)[c]					
	W1U	0.60 ± 0.41	0.75 ± 0.65	1.72 ± 0.10	ClNO
	W1Usc	0.56 ± 0.41	0.70 ± 0.65	1.79 ± 0.10	ClNO
	W1BD	0.61 ± 0.41	0.77 ± 0.65	1.97 ± 0.10	ClNO
	W1(RO)	0.55 ± 0.41	0.67 ± 0.65	1.75 ± 0.10	ClNO
EA (55)[d]					
	W1U	0.46 ± 0.22	0.60 ± 0.32	−1.96 ± 0.60	CH_2NC
	W1Usc	0.41 ± 0.22	0.53 ± 0.32	−1.64 ± 0.60	CH_2NC
	W1BD	0.43 ± 0.22	0.55 ± 0.32	−1.81 ± 0.60	CH_2NC
	W1(RO)	0.43 ± 0.22	0.53 ± 0.32	−1.74 ± 0.60	CH_2NC
IP (76)[e,f]					
	W1U	0.41 ± 0.16	0.57 ± 0.35	−2.10 ± 0.23	CS
	W1Usc	0.33 ± 0.16	0.42 ± 0.35	1.39 ± 0.05	P_2
	W1BD	0.37 ± 0.16	0.49 ± 0.35	−1.48 ± 0.18	N_2 ($^2\Sigma$ cation)
	W1(RO)	0.35 ± 0.16	0.46 ± 0.35	1.42 ± 0.05	P_2
PA (8)[g]					
	W1U	0.42	0.48	−0.83	C_2H_2
	W1BD	0.42	0.49	−0.88	C_2H_2
	W1(RO)	0.43	0.49	−0.83	C_2H_2
Total (220)					
	W1U	0.49 ± 0.27	0.65 ± 0.48		
	W1Usc	0.44 ± 0.27	0.57 ± 0.48		
	W1BD	0.48 ± 0.27	0.62 ± 0.48		
	W1(RO)	0.44 ± 0.27	0.57 ± 0.48		

[a] C_2 and O_3 are excluded in all statistics. Optimized geometries, total energies, and energy changes of species and reactions in the G2/97 sets are given in the Supporting Information “geometries.txt”, “components.pdf”, and “Test_set.pdf” files, respectively. The uncertainties in the deviations from experiment represent the uncertainties in the experimental data (see “Test_set.pdf”). [b] The number of reactions considered in each subset is indicated as the number in the parentheses in the first column. [c] See Supporting Information Tables S−III and S−IV for individual TAEs. [d] See Supporting Information Tables S−V and S−VI for individual EAs. [e] See Supporting Information Tables S−VII and S−VIII for individual IPs. [f] Excludes CN ($^3\Pi$), CN ($^1\Sigma^+$), B_2H_4, sec−C_3H_7, and Si2H6 (see ref 14). [g] See Supporting Information Table S–IX for individual PAs. Calculated proton affinities are considered converged at the W1 level (ref 14), agreeing well with experiment.

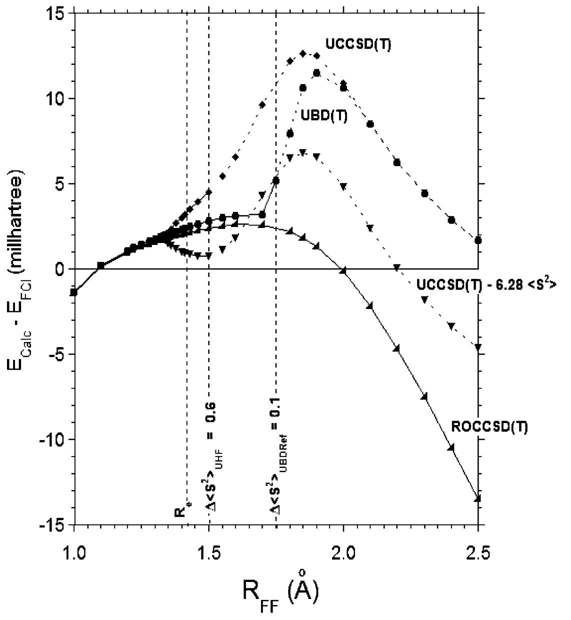


Figure 1. The error (mE_h) in the UCCSD(T), the spin-corrected UCCSD(T), the UBD(T), and the ROCCSD(T)/cc-pVDZ potential energy curves for the F_2 $^1\Sigma_g^+$ ground state. W1U/W1Usc and W1BD results beyond $\Delta\langle S^2\rangle_{UHF} = 0.6$ (1.5 Å) and $\Delta\langle S^2\rangle_{UBDRef} = 0.1$ (1.75 Å), respectively, (indicated by dashed rather than solid curves) are considered unreliable. Spin contamination errors cease to have a linear relationship with energetic errors beyond these cut-offs.

to geometry changes and the restricted coupled cluster begins to break down (1.2R_e or 1.75 Å in Figure 1). This RBD(T)/UBD(T) instability point is an unambiguous demarcation between the absence and the onset of significant BD energy errors, holding promise as a diagnostic tool for assessing the reliability of BD calculations. The good correlation between the W1BD $\Delta\langle S^2\rangle$ values and the energy errors is consistent with the results for CS^+ and O_3. The potential energy curve for the restricted wave function proceeds toward the wrong energy limit without such a warning.

V. Conclusions

The unrestricted flavors of W1 theory presented in this study are viable alternatives to W1(RO). W1U benefits significantly from a spin correction term for the 11 spin-contaminated reactions (Table 1), reducing the rms error from 1.3 to 0.6 ± 0.4 kcal/mol. The accuracy of the W1Usc, the W1BD, and the W1(RO) theories are indistinguishable from one another when evaluated with the moderately spin-contaminated data set (Table 1) and the 220 total atomization energies, electron affinities, ionization potentials, and proton affinities in the G2/97 test set (Table 2). Unlike W1U, the W1BD method demonstrates good correlation between the $\Delta\langle S^2\rangle_{UBDRef}$ and the energetic errors and, thus, gives a clear indication of the onset of energetic errors associated with spin contamination. W1(RO) gives no such straightforward warning when its single determinant reference wave function is inadequate. We, therefore, recommend the use of W1BD, rather than W1Usc, as an unrestricted alternative to W1(RO) in studies of potentially pathological cases. Multiple solutions were obtained for both the restricted and the unrestricted methods for C_2 and O_3, demonstrating the need for exploring multiple solutions with single-reference wave functions even in the "black-box" context. Of course, testing for the correct number of imaginary vibrational frequencies is always necessary.

Acknowledgment. This research was supported by Gaussian, Inc. and by the National Science Foundation through TeraGrid resources provided by NCSA (Grant number: TG-CHE080015).[73] We also acknowledge Wesleyan University for computer time supported by the NSF (Grant number CNS-0619508).

Supporting Information Available: Optimized geometries, energy components, and chemical energy differences for all species in this study, as well as additional discussion on the problematic species C_2, O_3, and CS^+, are available in the accompanying "geometries.txt", "components.pdf", "Test-_set.pdf", and "Problem_species.pdf" files, respectively. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) For recent reviews on accurate computational thermochemistry, see for example:Helgaker, T.; Klopper, W.; Halkier, A.; Bak, K. L.; Jørgensen, P. Olsen, J. In *Quantum-Mechanical Prediction of Thermochemical Data*, Cioslowski, J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2001; pp. 1−30; Martin, J. M. L.; Parthiban, S. In *Quantum-Mechanical Prediction of Thermochemical Data*, Cioslowski, J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2001; pp. 31−65; Raghavachari, K.; Curtiss, L. A. In *Quantum-Mechanical Prediction of Thermochemical Data*, Cioslowski, J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2001; pp. 67−98; Petersson, G. A. In *Quantum-Mechanical Prediction of Thermochemical Data*, Cioslowski, J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2001; pp. 99−130; Henry, D. J.; Radom, L. In *Quantum-Mechanical Prediction of Thermochemical Data*, Cioslowski, J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2001; pp. 161−197.

(2) Pople, J. A. Theoretical Models for Chemistry. *Proceedings of the Summer Research Conference on Theoretical Chemistry, Energy Structure and Reactivity*, 1973; p 51.

(3) (a) Pople, J. A.; Binkley, J. S.; Seeger, R. *Int. J. Quantum Chem. Symp.* **1976**, *10*, 1–19. (b) Krishnan, R.; Pople, J. A. *Int. J. Quantum Chem.* **1978**, *14*, 91–100.

(4) Petersson, G. A.; Bennett, A.; Tensfeldt, T. G.; Al-Laham, M. A.; Shirley, W. A.; Mantzaris, J. *J. Chem. Phys.* **1988**, *89*, 2193–2218.

(5) Montgomery, J. A., Jr.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1994**, *101*, 5900–5909.

(6) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1996**, *104*, 2598–2619.

(7) Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1999**, *110*, 2822–2827.

(8) Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **2000**, *112*, 6532–6542.

(9) Pople, J. A.; Head-Gordon, M.; Fox, D. J.; Raghavachari, K.; Curtiss, L. A. *J. Chem. Phys.* **1989**, *90*, 5622–5629.

(10) Curtiss, L. A.; Jones, C.; Trucks, G. W.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1990**, *93*, 2537–2545.

(11) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221–7230.

(12) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764–7776.

(13) Martin, J. M. L.; Oliveira, G. D. *J. Chem. Phys.* **1999**, *111*, 1843–1856.

(14) Parthiban, S.; Martin, J. M. L. *J. Chem. Phys.* **2001**, *114*, 6014–6029.

(15) Martin, J. M. L.; Parthiban, S. In *W1 and W2 theory and their variants: thermochemistry in the kJ/mol accuracy range*; Cioslowski, J., Ed.; Kluwer Academic: Dordrecht, The Netherlands, 2001; pp 1−30.

(16) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129–4141.

(17) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108−1144108−17.

(18) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599–11613.

(19) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108−1−064108−8.

(20) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111−1114111−15.

(21) DeYonker, N. J.; Cundari, T. R.; Wilson, A. K. *J. Chem. Phys.* **2006**, *124*, 11410−1114104−17.

(22) DeYonker, N. J.; Grimes, T.; Yockel, S.; Dinescu, A.; Mintz, B.; Cundari, T. R.; Wilson, A. K. *J. Chem. Phys.* **2006**, *125*, 104111−1104111−15.

(23) Karton, A.; Ruscic, B.; Martin, J. M. L. *J. Mol. Struct. (THEOCHEM)* **2007**, *811*, 345–353.

(24) Karton, A.; Taylor, P. R.; Martin, J. M. L. *J. Chem. Phys.* **2007**, *127*, 064104−1−064104−11.

(25) Karton, A.; Martin, J. M. L. *J. Phys. Chem. A* **2007**, *111*, 5936–5944.

(26) Karton, A.; Martin, J. M. L. *Mol. Phys.* **2007**, *105*, 2499–2505.

(27) Roothaan, C. C. J. *Rev. Mod. Phys.* **1951**, *23*, 69–89.

(28) Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* **1978**, *14*, 561–581.

(29) Rittby, M.; Bartlett, R. *J. Phys. Chem.* **1988**, *92*, 3033–3036.

(30) Lauderdale, W. J.; Stanton, J. F.; Gauss, J.; Watts, J. D.; Bartlett, R. J. *Chem. Phys. Lett.* **1991**, *187*, 21–28.

(31) Bartlett, R. J. In *Coupled-Cluster Theory: An Overview of Recent Developments*; Yarkony, D. R., Ed.; World Scientific, Singapore, 1995; pp 1047−1131.

(32) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(33) Pople, J. A.; Head-Gordon, M.; Raghavachari, K. *J. Chem. Phys.* **1987**, *87*, 5968–5975.

(34) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(35) Watts, J. D.; Gauss, J.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 8718–8733.

(36) Wood, G. P. F.; Radom, L.; Petersson, G. A.; Barnes, E. C.; Frisch, M. J.; Montgomery, J. A., Jr. *J. Chem. Phys.* **2006**, *125*, 094106−1−094106−16.

(37) Brueckner, K. A. *Phys. Rev.* **1954**, *96*, 508–516.

(38) Nesbet, R. K. *Phys. Rev.* **1958**, *109*, 1632–1638.

(39) Larsson, S.; Smith, V. *Phys. Rev.* **1969**, *178*, 137–152.

(40) Chiles, R. A.; Dykstra, C. E. *J. Chem. Phys.* **1981**, *74*, 4544–4556.

(41) Handy, N. C.; Pople, J. A.; Head-Gordon, M.; Raghavachari, K.; Trucks, G. W. *Chem. Phys. Lett.* **1989**, *164*, 185–192.

(42) Stanton, J. F.; Gauss, J.; Bartlett, R. J. *J. Chem. Phys.* **1992**, *97*, 5554–5559.

(43) Hampel, C.; Peterson, K. A.; Werner, H. J. *Chem. Phys. Lett.* **1992**, *190*, 1–12.

(44) Barnes, L. A.; Lindh, R. *Chem. Phys. Lett.* **1994**, *223*, 207–214.

(45) Watts, J. D.; Bartlett, R. J. *Int. J. Quantum Chem.* **1994**, *52*, 195–203.

(46) Scuseria, G. E. *Chem. Phys. Lett.* **1994**, *226*, 251–256.

(47) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M. Knox, J. E. Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09, Revision A.2,* Gaussian, Inc.: Wallingford, CT, 2009.

(48) Knowles, P. J.; Andrews, J. S.; Amos, R. D.; Handy, N. C.; Pople, J. A. *Chem. Phys. Lett.* **1991**, *186*, 130–136.

(49) Martin, J. M. L.; Taylor, P. R. *Chem. Phys. Lett.* **1994**, *225*, 473–479.

(50) Martin, J. M. L. *Chem. Phys. Lett.* **1995**, *242*, 343–350.

(51) Cowan, R. D.; Griffin, D. C. *J. Opt. Soc. Am.* **1976**, *66*, 1010–1014.

(52) Martin, R. L. *J. Phys. Chem.* **1983**, *87*, 750–754.

(53) Gdanitz, R. J.; Ahlrichs, R. *Chem. Phys. Lett.* **1988**, *143*, 413–420.

(54) Douglas, M.; Kroll, N. M. *Ann. Phys.* **1974**, *82*, 89–155.

(55) Hess, B. A. *Phys. Rev. A* **1985**, *32*, 756–763.

(56) Hess, B. A. *Phys. Rev. A* **1986**, *33*, 3742–3748.

(57) Barysz, M.; Sadlej, A. J. *J. Mol. Struct. (THEOCHEM)* **2001**, *573*, 181–200.

(58) de Jong, W. A.; Harrison, R. J.; Dixon, D. A. *J. Chem. Phys.* **2001**, *114*, 48–53.

(59) Visscher, L.; Dyall, K. G. *At. Data and Nucl. Data Tables* **1997**, *67*, 207–224.

(60) Ervin, K. M.; DeTuri, V. F. *J. Phys. Chem. A* **2002**, *106*, 9947–9956.

(61) Mordaunt, D. H.; Ashfold, M. N. R. *J. Chem. Phys.* **1994**, *101*, 2630–2631.

(62) Afeefy, H. Y.; Liebman, J. F.; Stein, S. E. NIST Chemistry Webbook. In *NIST Standard Reference Database*; Mallard, W. G., Linstrom, P. J., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 1998; Vol. 69.

(63) *CODATA Key Values for Thermodynamics*; Cox, J. D., Wagman, D. D., Medvedev, V. A., Eds.; Hemisphere Pub. Corp: New York, 1989.

(64) Ellison, G. B.; Davico, G. E.; Bierbaum, V. M.; DePuy, C. H. *Int. J. Mass Spectrom. Ion Process.* **1996**, *156*, 109–131.

(65) Nimlos, M. R.; Davico, G.; Geise, C. M.; Wenthold, P. G.; Lineberger, W. C.; Blanksby, S. R.; Hadad, C. M.; Petersson, G. A.; Ellison, G. B. *J. Chem. Phys.* **2002**, *117*, 4323–4339.

(66) Blanksby, S. J.; Ellison, G. B. *Acc. Chem. Res.* **2003**, *36*, 255–263.

(67) Wenthold, P. G.; Squires, R. R. *J. Am. Chem. Soc.* **1994**, *116*, 6401–6412.

(68) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1995**, *104*, 2598–2619.

(69) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063–1079.

(70) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 42–55.

(71) Schlegel, H. B. *J. Chem. Phys.* **1986**, *84*, 4530–4534.

(72) See, for example Watts, J. D.; Bartlett, R. J. *Int. J. Quantum Chem.* **1994**, *52*, 195–203.

(73) Catlett, C. In *TeraGrid: Analysis of Organization, System Architecture, and Middleware Enabling New Types of Applications*; Grandinetti, L., Ed.; Advances in Parallel Computing Series; IOS Press: Amsterdam, The Netherlands, 2007.

CT900260G

Chemical Physics 356 (2009) 14–24

Contents lists available at ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

Atomization energies from coupled-cluster calculations augmented with explicitly-correlated perturbation theory

Wim Klopper [a,*], Branko Ruscic [b], David P. Tew [a], Florian A. Bischoff [a], Sandra Wolfsegger [a]

[a] *Lehrstuhl für Theoretische Chemie, Institut für Physikalische Chemie, Universität Karlsruhe (TH), D-76128 Karlsruhe, Germany*
[b] *Chemical Sciences and Engineering Division, Argonne National Laboratory, Argonne, IL 60439, USA*

ARTICLE INFO

Article history:
Received 15 August 2008
Accepted 10 November 2008
Available online 17 November 2008

Dedicated to Prof. Dr. W. Kutzelnigg on the occasion of his 75th birthday.

Keywords:
Thermochemistry
Atomization energy
Coupled-cluster theory
Explicitly-correlated theory

ABSTRACT

The atomization energies of the 105 molecules in the test set of Bakowies [D. Bakowies, J. Chem. Phys. 127 (2007) 084105] have been computed with an estimated standard deviation (from the values compiled in the Active Thermochemical Tables) of ±0.1 kJ/mol per valence electron in the molecule. Equilibrium geometries and harmonic vibrational frequencies were calculated at the all-electron CCSD(T)/cc-pCVTZ level, that is, at the level of coupled-cluster theory with singles, doubles and non-iterative triples in a correlation-consistent polarized core–valence triple-zeta basis. Single-point energy calculations were performed at the all-electron CCSD(T) level in a correlation-consistent polarized core–valence quadruple-zeta basis (cc-pCVQZ), and several corrections were added: (i) a correction for the basis-set truncation error, obtained from second-order perturbation theory using Slater-type geminals (MP2-F12 theory), (ii) a correction for the effect of anharmonicity on the zero-point vibrational energy, (iii) a relativistic correction, (iv) a correction for the difference between the full CCSDT model (coupled-cluster theory with singles, doubles and triples) and the CCSD(T) approximation, and (v) a correction for connected quadruple excitations obtained from CCSDT(Q) calculations. The correction for the basis-set truncation error was obtained from MP2-F12 calculations by scaling the MP2 basis-set truncation error by an empirically optimized "interference factor" of $f_{\text{int}} = 0.78$. The reference values from the Active Thermochemical Tables for 73 molecules in the test set, the equilibrium geometries, the harmonic vibrational frequencies, and all of the energy corrections represent valuable data for performance assessments of additivity schemes that will be developed in the future, in which the basis-set truncation error will be calculated at the level of coupled-cluster theory using Slater-type geminals (CC-F12 theory). Such a scheme will be free of empirical corrections and scaling factors.

1. Introduction

The purpose of the present article is to provide a set of benchmark data in anticipation of rigorous assessments of various explicitly-correlated coupled-cluster R12 and F12 methods (CC-R12 and CC-F12).

These CC-R12 and CC-F12 methods have been developed since the publication of the pioneering work by Kutzelnigg on the He atom in 1985 [1], in which the conventional He-atom configuration-interaction expansion in terms of orbital products was augmented with one extra two-electron function that was linear in the interelectronic distance r_{12}. Furthermore, in this 1985 paper, it was proposed to use basis functions of the form

$$|\chi_{mn}\rangle = \widehat{Q}_{12} f(r_{12})|mn\rangle \qquad (1)$$

in correlated calculations on many-electron atoms and molecules, where $|mn\rangle$ is an antisymmetrized product of two spin orbitals that are occupied in the Hartree–Fock reference wave function, where $\widehat{Q}_{12} = (1-\widehat{O}_1)(1-\widehat{O}_2)$, with $\widehat{O} = \sum_m |m\rangle\langle m|$, is a projection operator taking care of strong orthogonality [1–4], and where $f(r_{12}) = r_{12}$. (It has recently become common practice to speak of R12 methods when $f(r_{12}) = r_{12}$ and of F12 methods when a particular function of r_{12} is used.) For many-electron systems, first R12 results were published at the level of Møller–Plesset perturbation theory (MP2) in Ref. [2] and the general theory for all matrix elements needed in configuration-interaction with singles and doubles (CISD), as well as second- and third-order Møller–Plesset perturbation (MP3) theories, was published in Ref. [3]. First R12 results obtained at the configuration-interaction with doubles (CID) and MP3 levels were published in Ref. [5].

A few years later, Noga et al. developed the CC-R12 methods [6]. In a landmark paper, Noga and Kutzelnigg presented a comprehensive diagrammatic derivation of the CC-R12 equations in the so-called "standard approximation B" at the levels of coupled-cluster

---

[*] Corresponding author. Tel.: +49 721 6087263; fax: +49 721 6083319.
E-mail address: klopper@chem-bio.uni-karlsruhe.de (W. Klopper).

0301-0104/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.chemphys.2008.11.013

theory with singles, doubles, and triples (CCSDT) as well as at the level of fourth-order Møller–Plesset (MP4) perturbation theory [7]. Fifth-order non-iterative corrections were also discussed.

An integral-direct CC-R12 program denoted DIRCCR12-OS was developed soon thereafter [8,9], both for closed- and open-shell species [10]. Recent examples of the performance of the DIRCCR12-OS program can be found in Refs. [11–13].

In the years 2002–2004, two important developments triggered renewed interest in the R12 methods. In 2002, Klopper and Samson introduced an auxiliary basis set for the resolution-of-the-indentity (RI) approximation that is used to avoid the calculation of three- and more-electron integrals [14]. In the aforementioned standard approximation (SA), the same basis set was used for the orbital expansion *and* the RI insertion, which simplified the equations dramatically [7], but which required a very large basis set of atomic orbitals to achieve reliable results, as can be seen in Refs. [11–13]. This problem was alleviated by the introduction of the auxiliary basis set.

In 2004, Ten-no proposed to use Slater-type geminals (STGs) of the form $f(r_{12}) = \exp(-\gamma r_{12})$ in place of the linear r_{12} terms [15]. The use of STGs was a significant improvement on the linear r_{12} terms in particular in calculations with small basis sets, which had become possible by using an auxiliary basis set for the RI approximation.

Since 2004, a number of approximate CC-F12 methods have been developed using an auxiliary basis set and STGs.

Fliegl et al. have introduced an approximation to the CCSD-F12 method with singles, doubles, and STGs, which was denoted CCSD(F12) [16]. The corresponding model CCSD(T)(F12) includes a correction for non-iterative connected triple excitations [17]. Tew et al. showed that this CCSD(T)(F12) approach can provide quintuple-ζ quality CCSD(T) correlation energies in just a triple-ζ basis set, not only when optimizing all STG amplitudes [18] but also when keeping these amplitudes fixed [19] at their values dictated by the electron–electron cusp conditions (for a recent discussion of the cusp conditions, cf. Ref. [20]), as first done by Ten-no at the MP2-F12 level [21,22].

Adler et al. have introduced a much simpler approximation to the CCSD-F12 model [23], and in a series of papers, Valeev and co-workers [24–26] have developed the approximate model $\text{CCSD}(2)_{\overline{\text{R12}}}$, which treats the R12- or F12-dependent terms through Löwdin perturbation theory. Also the fixed-amplitudes approach and corrections for connected triple excitations have been included in the $\text{CCSD}(2)_{\overline{\text{R12}}}$ model.

Based on an automated implementation by means of computerized symbolic algebra, Shiozaki et al. have implemented a full CCSD-F12 method that relies neither on the SA, nor on truncations or other approximations such at the generalized (GBC) and extended (EBC) Brillouin conditions [27,28]. Instead, these autors use multiple RI insertions. In contrast to this, Bokhan and co-workers have implemented a fixed-amplitude CCSD(T)-F12 method on the basis of the SA, without auxiliary basis [29].

The present work aims at providing benchmark data for future work on performance assessments of the CCSD-F12 method and its approximations.

In the next section, we shall briefly introduce the additivity scheme that is used to calculate the atomization energies of a test set of 106 molecules. This is the same test set that was investigated by Bakowies [30]. It contains 105 closed-shell molecules, plus dihydrogen. The latter was added because it often occurs in assessments of reaction enthalpies.

It is important to note that the additivity scheme given below is not meant as a new "model chemistry" or "composite" or "multicoefficient" method. Rather, the purpose of the present work is to provide accurate equilibrium geometries and a number of energy corrections such that, when added to CCSD-F12 energies, it should be possible to obtain calculated atomization energies to within a standard deviation of ≈0.1 kJ/mol per valence electron from the ATcT reference values (cf. Section 3.1).

Various additivity schemes and model chemistries already exist in the literature (a recent brief review can be found in Ref. [31]). Among these are the Gaussian-*n* ($n = 2,3,4$) theories of Curtiss et al. [32–34], the correlation-consistent composite approach (ccCA) of DeYonker et al. [35,36], the complete basis set (CBS) methods of Petersson and co-workers [37,38], the focal-point analysis (FPA) approach of Allen and co-workers [39,40], the multicoefficient correlation methods (MCCMs) of Fast et al. [41–43], the HEAT protocol of Tajti et al. [44], Bomble et al. [45] and Harding et al. [46], and the Weizmann-*n* ($n = 1$–4) theories of Martin and co-workers [47–50]. Many of these model chemistries contain extrapolations, empirical corrections, and empirical scaling factors. Also Feller, Dixon and co-workers (cf. Refs. [51–56] and references therein) have developed an approach to calculate molecular thermodynamic properties based on fc-CCSD(T) coupled-cluster theory using correlation-consistent basis sets, extrapolations to the basis-set limit, and inclusion of a number of corrections such as core–valence interactions, scalar and spin–orbit relativistic effects and zero-point vibrational energies (ZPVEs).

As already said, the additivity scheme given below, which in its present form contains an empirical scaling factor, should not be misunderstood as a new black-box tool with broad applicability in computational thermochemistry. Rather, the hope is expressed that an accurate and reliable additivity scheme can be designed in the future based on the CCSD(T)(F12)/def2-QZVPP level, for example, free of extrapolations, empirical corrections, and empirical scaling factors. Such a scheme will be studied in future work, using the data compiled in the present work.

The present article is organized as follows: In Section 2, we give a brief overview of the additivity scheme used. More detailed information on the calculation of the individual contributions to the additivity scheme is given in Section 3. Results are presented in Section 4, including a discussion of the remaining sources of error in Section 4.2. Our conclusions are collected in Section 5.

2. Overview of the additivity scheme

In the following, we shall give a brief overview of the additivity scheme used before we turn to the full computational details presented in Section 3.

We shall denote total electronic energies as E and electron-correlation contributions as δE. For example,

$$E_{\text{fc-CCSD(T)/cc-pVTZ}} = E_{\text{Hartree–Fock/cc-pVTZ}} + \delta E_{\text{fc-CCSD(T)/cc-pVTZ}}. \tag{2}$$

The total electronic energy is obtained by adding various corrections to the fc-CCSD(T)/cc-pCVQZ//ae-CCSD(T)/cc-pCVTZ energy,

$$\begin{aligned} E_{\text{total}} = {} & E_{\text{fc-CCSD(T)/cc-pCVQZ}} + \Delta E_{\text{CV}} + \Delta E_{\text{ZPVE}} + \Delta E_{\text{Anh}} + \Delta E_{\text{F12}} \\ & + \Delta E_{\text{MVD}} + \Delta E_{\text{SO}} + \Delta E_{\text{T}} + \Delta E_{\text{(Q)}}. \end{aligned} \tag{3}$$

All of these corrections except ΔE_{Anh} were computed at the ae-CCSD(T)/cc-pCVTZ equilibrium geometry of the molecule in question (cf. Section 3.3). ΔE_{CV} is a correction for core–valence correlation effects, which was obtained as the difference between the ae-CCSD(T)/cc-pCVQZ and fc-CCSD(T)/cc-pCVQZ energies. ΔE_{ZPVE} is the harmonic zero-point vibrational energy calculated at the ae-CCSD(T)/cc-pCVTZ level. ΔE_{Anh} is a correction to the harmonic ZPVE due to anharmonic effects, calculated at the fc-MP2/cc-pVDZ level. ΔE_{F12} is a correction for the basis-set truncation error. It was calculated as follows:

$$\delta E_{\text{F12}} = f_{\text{int}}\left(\delta E_{\text{fc-MP2-F12}} - \delta E_{\text{fc-MP2/cc-pCVQZ}}\right), \tag{4}$$

$$\Delta E_{\text{F12}} = E_{\text{Hartree–Fock/def2-QZVPP}} - E_{\text{Hartree–Fock/cc-pCVQZ}} + \delta E_{\text{F12}}, \tag{5}$$

where f_{int} is an empirical "interference" factor [57–60]. In the course of the present work, an optimization of this factor by minimizing the mean deviation from the ATcT reference values yielded $f_{\text{int}} = 0.78$, but we also report the statistical analysis of the calculated atomization energies obtained with $f_{\text{int}} = 0.0$ and $f_{\text{int}} = 1.0$. Concerning Eq. (5), we note that the Hartree–Fock energy in the def2-QZVPP basis was always lower than in the cc-pCVQZ basis.

ΔE_{MVD} is a correction for scalar-relativistic effects (one-electron Darwin and mass–velocity terms) calculated at the ae-CCSD(T)/cc-pCVTZ level [61,62]. For the atoms C, O, and F, the spin–orbit corrections to the total electronic energy amount to $\Delta E_{\text{SO}} = -0.35399$, -0.93278, and -1.61153 kJ/mol, respectively [63].

Finally, a full correction for connected triple excitations and a perturbative correction for connected quadruple excitations were computed as

$$\Delta E_{\text{T}} = E_{\text{CCSDT/cc-pVTZ}} - E_{\text{CCSD(T)/cc-pVTZ}}, \tag{6}$$

$$\Delta E_{(\text{Q})} = E_{\text{CCSDT(Q)/cc-pVDZ}} - E_{\text{CCSDT/cc-pVDZ}}. \tag{7}$$

Note that ΔE_{T} does *not* refer to the total contribution of connected triples excitations but only to the difference between the CCSDT and CCSD(T) models.

3. Computational details

3.1. Active Thermochemical Tables

Accurate, reliable, and internally consistent thermochemical values were taken from the Active Thermochemical Tables (ATcT) [64–67]. As opposed to the traditional sequential approach, the ATcT derive their results from a thermochemical network (TN) using all available knowledge. The thermochemical values used in the present work have been obtained from the latest version of the Core (Argonne) Thermochemical Network, C(A)TN, which is currently under development [68] and describes ca. 900 species interconnected by ca. 10,000 experimental and theoretical determinations.

3.2. Computer programs

All coupled-cluster and second-order Møller–Plesset (MP2) calculations were carried out with the Mainz–Austin–Budapest 2005 version of the ACES II program [69] and with the MRCC program [70,71]. The coupled-cluster calculations performed with ACES II were carried out at the level of coupled-cluster theory with singles and doubles (CCSD, cf. Refs. [72,73]) including a non-iterative correction for connected triple excitations (CCSD(T), cf. Refs. [74–78]). Kállay's MRCC program was used for coupled-cluster calculations with singles, doubles, and triples including a non-iterative correction for connected quadruple excitations (CCSDT(Q), cf. Refs. [79,80]).

Explicitly-correlated calculations [81,82] were carried out at the level of second-order perturbation theory (MP2-F12) with the TURBOMOLE program [83].

3.3. Geometries and zero-point vibrational energies

All molecular equilibrium geometries were optimized at the all-electron CCSD(T) level [ae-CCSD(T)] in the correlation-consistent polarized core–valence triple-zeta basis set (cc-pCVTZ) of Dunning [84] and Woon and Dunning [85]. Using the cc-pCVTZ basis implies that this basis is used for C, N, O, and F in conjunction with a correlation-consistent polarized valence triple-zeta basis set (cc-pVTZ) for H [84]. Harmonic vibrational frequencies and harmonic ZPVEs were computed at the same level, that is, at the ae-CCSD(T)/cc-pCVTZ level.

The ae-CCSD(T) calculations of the closed-shell molecules were performed using a restricted Hartree–Fock (RHF) reference determinant.

3.4. Anharmonic corrections

MP2 calculations were carried out to obtain anharmonic corrections to the ZPVEs. The cubic force field and those parts of the quartic force field that are required for the determination of anharmonic effects were obtained by means of numerical differentiation of analytical Hessians about the fc-MP2/cc-pVDZ equilibrium structure, as implemented in the Mainz–Austin–Budapest 2005 version of the ACES II program [69,86]. The MP2 calculations of the closed-shell molecules were done using an RHF reference, in the frozen-core approximation, and in the correlation-consistent polarized valence double-zeta (cc-pVDZ) basis set of Dunning [84]. Thus, harmonic ZPVEs were obtained at the ae-CCSD(T)/cc-pCVTZ level and anharmonic corrections to these were obtained at the fc-MP2/cc-pVDZ level.

3.5. Coupled-cluster single-point energies

The CCSD(T) single-point energy calculations of the closed-shell molecules were performed using an RHF reference determinant. The corresponding calculations of the atoms C, N, O, and F were done using a restricted open-shell Hartree–Fock (ROHF) reference as implemented in ACES II, employing semi-canonical orbitals [78].

The single-point energy CCSD(T) calculations were performed in the correlation-consistent polarized core–valence quadruple-zeta (cc-pCVQZ) basis set (cc-pVQZ for H), both in an all-electron treatment [ae-CCSD(T)/cc-pCVQZ] and in a frozen-core treatment [fc-CCSD(T)/cc-pCVQZ]. The difference between these two calculations is referred to as core–valence (CV) contribution.

The single-point energy CCSDT and CCSDT(Q) calculations were performed in the correlation-consistent polarized triple-zeta (cc-pVTZ) and double-zeta (cc-pVDZ) basis sets [84], respectively, in the frozen-core approximation. The corresponding calculations of the atoms C, N, O, and F were done using an unrestricted Hartree–Fock (UHF) reference as implemented in MRCC. The (Q) contribution was obtained as the difference between the fc-CCSDT(Q)/cc-pVDZ and fc-CCSDT/cc-pVDZ energies, using RHF and UHF reference determinants for the molecules and atoms, respectively. Similarly, the full-triples-minus-(T)-triples correction was obtained as the difference between the fc-CCSDT/cc-pVTZ and fc-CCSD(T)/cc-pVTZ energies.

3.6. Explicitly-correlated perturbation theory

The MP2-F12 calculations were carried out using a Slater-type geminal with exponent $\gamma = 1.4\ a_0^{-1}$, represented by a linear combination of six Gaussians with exponents and coefficients taken from Ref. [87].

The implementation of the MP2-F12 method in the RICC2 module [88,89] of TURBOMOLE is described in detail in Refs. [90,91]. The robust density-fitting technique of Manby was used to avoid the computation of four-index integrals [92], and a complementary auxiliary basis set (CABS) [93] was used for the resolution-of-the-indentity (RI) approximation of explicitly-correlated theory. The approach 2B of explicitly-correlated theory [90,91] was used and the amplitudes were optimized in an orbital-invariant manner [94]. The matrix representation of the core Hamiltonian, $\widehat{T} + \widehat{V}$, was used for the commutator approximation that avoids the two-electron integrals over the operator $[\widehat{T}, f(r_{12})]$ [90,95]. For the open-shell MP2-F12 calculations of the atoms, an ROHF reference function and semi-canonical orbitals were used. The core orbitals

(1s for C, N, O, and F) were not included in the MP2-F12 approach (fc-MP2-F12).

The fc-MP2-F12 calculations were performed in the def2-QZVPP basis [96]. For the density-fitting approximation, the aug-cc-pwCV5Z MP2 fitting basis of Hättig was used (aug-cc-pV5Z for H) [97]. The def2-QZVPP Hartree–Fock exchange fitting basis was used in two different manners. On the one hand, it was used as CABS, and on the other hand, it was used for computing matrix elements of the Fock operator using the RI-*JK* approximation [98].

3.7. Relativistic corrections

Darwin and mass–velocity energy corrections [99] were computed analytically [78,100] as first-order molecular properties at the ae-CCSD(T)/cc-pCVTZ level, using RHF and ROHF reference wave functions for the molecules and atoms, respectively.

4. Results and discussion

4.1. Final results

The computed atomization energies are reported in Table 1 and are compared with the ATcT values where available. The deviation of the computed atomization energies from the ATcT values are given in the second-last column. In Table 1, the reported CCSD(T) values refer to the fc-CCSD(T)/cc-pCVQZ//ae-CCSD(T)/cc-pCVTZ level, which is the baseline of the additivity scheme. The values presented in the other columns are the corrections described in Eq. (3).

In the last column of Table 1, the D1 diagnostic of Jansen and Nielsen is given [101], which is an indicator for the multireference character of the molecule's electronic ground state. Since all of the calculations performed in the present work are of the single-reference type, we expected larger deviations for the molecules with large D1 values than for those with small D1 values. To investigate this, we plot in Fig. 1 the deviations of the computed atomization energies from the ATcT values as a function of the molecule's D1 value (from 0 to 0.06). If the F12 correction is omitted, there are large deviations when D1 is large, but there are also large errors for relatively small D1 values (e.g., for cyclopropane, where the error amounts to −29.6 kJ/mol while D1 is only 0.015). Hence, for the molecules under study and in contrast to our expectations, the error is not dominated by the high-level excitations (beyond the CCSDT(Q) model) of coupled-cluster theory, which become more important with increasing multireference character. After adding the F12 correction from MP2-F12 theory (but without scaling with the interference factor f_{int}), we observe a very weak trend of increasing error with increasing D1 diagnostic, but after scaling with the empirical factor of $f_{\text{int}} = 0.78$, the remaining deviations appear to be independent of D1. Hence, the final deviations between the calculated atomization energies and the ATcT reference values are independent of D1 up to values of $D1 = 0.06$. The deviations for N_2O_3 and N_2O_4 (with D1 > 0.07, not shown in Fig. 1) are quite large, however. For $f_{\text{int}} = 0$, the deviations are −42 and −52 kJ/mol, and for $f_{\text{int}} = 1$, they are 13 and 14 kJ/mol, respectively.

The errors reported in Table 1 are statistically analyzed in Table 2. For $f_{\text{int}} = 0.78$, we observe that the mean deviation between calculation and ATcT amounts to −0.12 kJ/mol, with a mean absolute deviation of 0.9 kJ/mol and a root-mean-square (RMS) deviation of 1.2 kJ/mol. Table 2 also reports these errors for calculations without F12 corrections from MP2-F12 theory ($f_{\text{int}} = 0.0$) as well as for calculations with F12 correction from MP2-F12 theory but without empirical scaling ($f_{\text{int}} = 1.0$). In both cases, the errors are substantially larger. In the cases $f_{\text{int}} = 0.0$ and $f_{\text{int}} = 1.0$, the errors are about 20 and 5–6 times larger, respectively, than for $f_{\text{int}} = 0.78$. This means that already adding the unscaled F12 contribution represents an improvement over the fully uncorrected results, reducing the error by roughly a factor of 3–4. However, a satisfactory agreement between computed and ATcT values is only obtained for $f_{\text{int}} = 0.78$.

In Section 4.2, we shall analyze the uncertainties of the individual contributions of the additivity scheme together with the possible errors due to neglected terms. As we shall see, these uncertainties and errors are of about the same order of magnitude and add up (in terms of a Gaussian error propagation) to the observed deviation between the calculated atomization energies and the ATcT reference values. In Section 4.2, we shall focus on the RMS errors *per valence electron* when we discuss the errors of the individual contributions. We do this for two reasons. Firstly, as Harding et al. [46] have pointed out, it is a necessary consequence of size extensivity that the characteristic relative error in atomization energies will remain constant while the absolute error will grow linearly with the size of the system. We have chosen to take the number of electrons in the valence shells (1 for H, 4 for C, 5 for N, 6 for O, 7 for F) as a measure of the size of the system. Secondly, for the propagated estimate of the error to be a legitimate estimate of a standard deviation σ, all of the components also must be a σ or best estimates thereof (e.g., RMS or experience-based-95%-confidence-interval/2, but nothing based on mean absolute errors).

Figs. 2 and 3 show that not only the deviations per molecule (Fig. 2) but also the deviations per valence electron (Fig. 3) appear to form normal distributions (Gaussian distributions). In these figures, the points represent the number of molecules with an error within the corresponding interval (e.g., 16 molecules have a total deviation between −1.5 and −0.5 kJ/mol), and the bell curves are simple non-linear fits to these points.

4.2. Error estimation of individual contributions

4.2.1. Hartree–Fock contributions

Concerning the Hartree–Fock level, our additivity scheme is based on Hartree–Fock calculations in the def2-QZVPP basis [cf. Eq. (5)]. This basis yields Hartree–Fock contributions to the atomization energies closer to the limit of a complete basis than the cc-pCVQZ basis, but the basis-set truncation error is still not negligible. To estimate this error, we have computed the Hartree–Fock contribution in the cc-pCV5Z basis for the 18 molecules shown in Table 3. For these molecules, the RMS deviation between the def2-QZVPP and cc-pCV5Z contributions to the atomization energies amounts to 1.0 kJ/mol per molecule or 0.08 kJ/mol per valence electron. We adopt this RMS error of 0.08 kJ/mol per valence electron for our overall error analysis (cf. Section 4.2.7).

4.2.2. Core–valence contributions

In Table 3, the core–valence contributions to the atomization energies are presented for a selection of 18 molecules, obtained at the CCSD(T) level in the cc-pCVQZ and cc-pCV5Z basis sets as well as from a two-point X^{-3} extrapolation [102] based on the cc-pCVQZ and cc-pCV5Z results, which is denoted cc-pCV(Q5)Z. For the final results presented in Table 1, the CCSD(T)/cc-pCVQZ data were used, and the accuracy of these data can be estimated from the difference between the cc-pCVQZ and cc-pCV(Q5)Z core–valence contributions. The mean absolute and RMS deviations for these 18 molecules amount to 0.36 and 0.46 kJ/mol, respectively. Expressed in terms of error per valence electron, the mean absolute and RMS deviations are 0.033 and 0.042 kJ/mol.

In Ref. [49], Karton et al. report core–valence contributions for the 14 molecules H_2O, C_2H_2, CH_4, CO_2, CO, F_2, FH, N_2, H_3N, N_2O, O_3, C_2H_4, CH_2O (formaldehyde), and HNO, as obtained in W4 theory, that is, at the extrapolated CCSD(T)/aug-cc-pwCV(TQ)Z level.

Table 1
Atomization energies in kJ/mol[a].

Nr.[b]	Molecule[c]		CCSD(T)	CV	ZPVE	Anh.	F12	MVD	SO	T	(Q)	Total	ATcT	Error	D1[d]
1	CFN	Cyanogen fluoride	1252.9	6.9	−26.0	0.1	19.1	−1.8	−2.0	−3.1	4.9	1250.9	1250.3 ± 1.7	0.6	0.027
2	CFN	Isocyanogen fluoride	956.1	5.2	−21.1	0.1	18.0	−1.6	−2.0	−2.4	4.7	957.0	959.2 ± 2.7	−2.2	0.035
3	CF_2	Singlet difluoromethylene	1065.6	1.9	−18.6	0.1	14.5	−1.5	−3.6	−1.1	2.6	1060.0	1059.1 ± 0.8	0.9	0.045
4	CF_2O	Carbonyl fluoride	1732.6	5.9	−37.5	0.2	22.3	−3.0	−4.5	−3.1	4.3	1717.2	1718.4 ± 0.9	−1.2	0.041
5	CF_4	Tetrafluoromethane	1975.6	4.6	−46.1	0.0	25.6	−3.9	−6.8	−3.2	3.6	1949.5	1947.9 ± 0.6	1.6	0.027
6	CHF	Singlet fluoromethylene	878.1	1.8	−32.6	0.6	10.6	−1.0	−2.0	0.0	1.5	857.0			0.038
7	CHFO	Formyl fluoride	1664.8	5.5	−55.2	0.5	19.0	−2.1	−2.9	−2.2	3.5	1631.0	1631.4 ± 0.9	−0.4	0.038
8	CHF_3	Trifluoromethane	1898.2	4.5	−67.8	0.0	22.1	−3.1	−5.2	−2.3	2.8	1849.2	1848.7 ± 0.9	0.5	0.026
9	CHN	Hydrogen cyanide	1287.5	6.6	−41.9	0.3	16.1	−0.9	−0.4	−2.5	3.7	1268.5	1268.3 ± 0.2	0.2	0.027
10	CHN	Hydrogen isocyanide	1225.3	5.8	−40.8	0.3	15.7	−1.1	−0.4	−1.6	2.4	1205.7	1207.0 ± 0.6	−1.3	0.029
11	CHNO	Cyanic acid	1683.1	8.2	−56.3	0.3	23.8	−2.2	−1.3	−3.7	5.2	1657.1	1657.2 ± 1.0	−0.1	0.029
12	CHNO	Isocyanic acid	1784.7	8.6	−55.8	0.2	24.5	−2.3	−1.3	−3.0	4.9	1760.4	1761.0 ± 0.4	−0.6	0.043
13	CHNO	Formonitrile oxide	1490.7	8.9	−50.1	0.1	24.6	−2.4	−1.3	−4.7	8.1	1473.9	1474.1 ± 1.2	−0.2	0.051
14	CHNO	Isofulminic acid	1434.6	6.6	−53.2	0.4	22.9	−2.1	−1.3	−3.1	4.7	1409.4	1410.2 ± 1.0	−0.8	0.033
15	CH_2	Singlet methylene	749.7	1.6	−43.8	0.6	6.7	−0.4	−0.4	0.9	0.3	715.4	714.9 ± 0.2	0.5	0.009
16	CH_2F_2	Difluoromethane	1810.3	4.5	−87.3	1.0	18.5	−2.4	−3.6	−1.5	2.0	1741.5	1741.7 ± 0.8	−0.2	0.024
17	CH_2N_2	Cyanamide	1993.6	9.8	−89.7	1.0	28.0	−2.2	−0.4	−3.5	4.8	1941.5			0.030
18	CH_2N_2	3H-Diazirine	1817.0	7.1	−87.4	1.2	27.2	−1.7	−0.4	−4.5	5.8	1764.3			0.028
19	CH_2N_2	Diazomethane	1854.6	9.3	−82.8	1.4	26.8	−2.1	−0.4	−3.5	6.5	1809.9			0.038
20	CH_2O	Formaldehyde	1547.1	5.2	−70.2	1.0	15.2	−1.4	−1.3	−1.5	2.5	1496.6	1495.8 ± 0.2	0.8	0.034
21	CH_2O	Hydroxymethylene	1330.2	3.5	−70.5	1.0	15.0	−1.4	−1.3	−0.7	2.0	1277.9	1277.8 ± 1.1	0.1	0.038
22	CH_2O_2	Dioxirane	1686.3	4.8	−85.8	1.1	23.2	−1.8	−2.2	−3.8	6.0	1627.9	1629.6 ± 1.7	−1.7	0.027
23	CH_2O_2	Formic acid	2070.1	6.7	−89.4	1.2	23.7	−2.6	−2.2	−2.7	4.0	2008.7	2008.4 ± 0.3	0.3	0.041
24	CH_2O_3	Performic acid	2209.6	6.7	−96.2	1.2	30.7	−2.9	−3.2	−3.8	6.8	2148.9			0.042
25	CH_3F	Fluoromethane	1750.8	4.7	−103.9	0.5	14.6	−1.6	−2.0	−0.8	1.2	1663.4	1665.1 ± 0.6	−1.7	0.019
26	CH_3N	Methanimine	1812.7	6.4	−104.9	1.4	19.0	−1.5	−0.4	−1.6	2.2	1733.6	1733.5 ± 1.0	0.1	0.027
27	CH_3NO	Formamide	2339.9	8.7	−119.5	2.1	28.3	−2.8	−1.3	−2.4	3.6	2256.6			0.040
28	CH_3NO_2	Methyl nitrite	2464.8	6.8	−127.7	1.8	34.0	−2.6	−2.2	−3.7	7.8	2379.0			0.052
29	CH_3NO_2	Nitromethane	2472.8	8.3	−131.4	2.2	35.3	−3.2	−2.2	−4.5	8.1	2385.3			0.054
30	CH_4	Methane	1744.9	5.0	−117.9	0.3	9.5	−0.8	−0.4	−0.2	0.3	1640.9	1642.2 ± 0.1	−1.3	0.011
31	CH_4N_2O	Urea	3098.9	11.6	−168.7	3.0	40.7	−3.9	−1.3	−3.6	4.5	2981.0			0.041
32	CH_4O	Methanol	2124.4	5.8	−135.5	1.8	18.8	−2.0	−1.3	−1.3	1.6	2012.4	2012.7 ± 0.2	−0.3	0.017
33	CH_5N	Methylamine	2407.3	7.2	−169.0	2.6	22.3	−1.9	−0.4	−1.1	1.4	2268.4	2269.0 ± 0.5	−0.6	0.014
34	CO	Carbon monoxide	1071.6	3.8	−12.9	0.0	11.0	−0.7	−1.3	−1.6	2.7	1072.5	1072.1 ± 0.1	0.4	0.035
35	CO_2	Carbon dioxide	1605.2	7.0	−30.4	0.1	19.7	−2.1	−2.2	−3.3	5.1	1599.1	1598.2 ± 0.1	0.9	0.042
36	C_2F_2	Difluoroacetylene	1582.1	11.4	−34.6	0.2	21.6	−3.2	−3.9	−3.5	4.7	1574.8	1577.0 ± 1.7	−2.2	0.027
37	C_2F_4	Tetrafluoroethylene	2425.5	10.7	−57.3	0.5	33.3	−4.7	−7.2	−4.0	5.3	2402.2	2405.2 ± 1.0	−3.0	0.037
38	C_2HF	Fluoroacetylene	1638.5	10.4	−52.5	0.4	17.6	−2.2	−2.3	−3.0	3.6	1610.7	1612.3 ± 1.0	−1.6	0.024
39	C_2HF_3	Trifluoroethylene	2409.4	10.3	−77.6	0.7	29.2	−3.9	−5.5	−3.2	4.3	2363.5			0.034
40	C_2H_2	Acetylene	1671.4	9.6	−69.4	0.9	14.6	−1.2	−0.7	−2.3	2.6	1625.5	1626.2 ± 0.2	−0.7	0.025
41	$C_2H_2F_2$	1,1-Difluoroethylene	2420.2	10.0	−96.8	1.1	24.4	−3.0	−3.9	−2.7	3.4	2352.7			0.029
42	C_2H_2O	Ketene	2198.9	10.8	−82.8	0.9	21.1	−2.0	−1.6	−2.5	3.9	2146.7	2147.3 ± 0.2	−0.6	0.039
43	C_2H_2O	Oxirene	1871.7	9.5	−75.7	1.0	24.8	−2.0	−1.6	−3.7	4.8	1828.8			0.026
44	$C_2H_2O_2$	Glyoxal	2616.9	10.5	−97.3	1.3	29.2	−2.8	−2.6	−3.7	5.7	2557.1	2555.3 ± 0.6	1.8	0.040
45	C_2H_3F	Fluoroethylene	2370.4	9.6	−115.7	1.4	20.7	−2.2	−2.3	−1.9	2.5	2282.5	2278.4 ± 1.7	4.1	0.024
46	C_2H_3FO	Acetyl fluoride	2919.4	10.6	−129.3	1.6	28.2	−3.0	−3.3	−3.1	4.0	2825.1			0.039
47	C_2H_3N	Acetonitrile	2542.8	11.6	−118.9	0.6	24.5	−1.8	−0.7	−3.4	4.3	2459.2			0.027
48	C_2H_3N	Methyl isocyanide	2441.3	10.5	−118.8	0.6	24.2	−1.9	−0.7	−2.8	3.2	2355.6			0.028
49	C_2H_4	Ethylene	2334.3	9.3	−133.7	2.1	16.2	−1.4	−0.7	−1.4	1.5	2226.3	2225.9 ± 0.2	0.4	0.025
50	C_2H_4O	Acetaldehyde	2802.0	10.2	−146.1	1.8	24.0	−2.2	−1.6	−2.3	3.1	2688.9	2688.9 ± 0.4	0.0	0.036
51	C_2H_4O	Oxirane	2691.0	9.9	−151.5	2.2	25.4	−2.4	−1.6	−3.0	3.1	2573.0	2573.9 ± 0.5	−0.9	0.023
52	$C_2H_4O_2$	Acetic acid	3320.2	11.7	−163.0	2.1	32.7	−3.4	−2.6	−3.7	4.5	3198.5	3199.3 ± 1.5	−0.8	0.042
53	$C_2H_4O_2$	Methyl formate	3249.9	11.0	−163.9	2.1	32.4	−3.4	−2.6	−3.8	4.7	3126.5	3125.2 ± 0.6	1.3	0.043
54	C_2H_5F	Fluoroethane	2987.0	9.5	−179.5	2.5	23.0	−2.4	−2.3	−1.6	1.7	2838.0	2838.5 ± 1.9	−0.5	0.021

55	C_2H_5N	Aziridine	2974.5	11.4	−185.3	2.8	28.8	−2.5	−0.7	−3.0	2.8	2828.7			0.018
56	C_2H_6	Ethane	2956.5	9.6	−196.6	0.9	18.0	−1.6	−0.7	−1.0	0.9	2786.0	2787.2 ± 0.2	−1.2	0.012
57	C_2H_6O	Dimethyl ether	3307.2	10.1	−210.4	3.3	26.8	−2.8	−1.6	−2.2	2.2	3132.6	3132.4 ± 0.5	0.2	0.021
58	C_2H_6O	Ethanol	3357.5	10.6	−210.9	2.9	27.2	−2.7	−1.6	−2.2	2.2	3183.0	3182.8 ± 0.3	0.2	0.018
59	C_2N_2	Cyanogen	2054.5	13.4	−41.2	0.2	29.5	−2.0	−0.7	−6.3	9.1	2056.5	2055.8 ± 0.5	0.7	0.029
60	C_3H_3N	Acrylonitrile	3144.0	16.1	−132.3	1.5	30.8	−2.4	−1.1	−4.9	6.0	3057.7			0.031
61	C_3H_4	Allene	2908.3	14.3	−144.3	1.0	22.4	−2.0	−1.1	−2.6	2.9	2798.9	2800.9 ± 0.5	−2.0	0.023
62	C_3H_4	Cyclopropene	2815.6	14.2	−147.0	2.0	23.4	−2.2	−1.1	−3.6	3.1	2704.3	2705.1 ± 1.0	−0.8	0.024
63	C_3H_4	Propyne	2916.5	14.8	−145.6	1.7	19.7	−2.0	−1.1	−3.3	3.4	2804.1	2805.6 ± 0.5	−1.5	0.025
64	C_3H_6	Cyclopropane	3533.8	14.7	−214.5	1.1	25.3	−2.4	−1.1	−2.8	2.1	3356.1	3359.7 ± 0.6	−3.6	0.015
65	C_3H_6	Propene	3565.7	14.2	−209.2	3.0	24.8	−2.2	−1.1	−2.3	2.3	3395.1	3395.0 ± 0.4	0.1	0.024
66	C_3H_8	Propane	4177.0	14.3	−272.1	4.0	26.6	−2.4	−1.1	−2.0	1.6	3945.9	3944.6 ± 0.4	1.3	0.013
67	C_3O_2	Carbon suboxide	2730.2	16.9	−55.8	0.4	32.1	−3.3	−2.9	−6.6	10.4	2721.3			0.053
68	C_4H_4	Butatriene	3489.4	19.1	−156.4	2.3	28.5	−2.6	−1.4	−3.9	5.1	3380.1			0.031
69	C_4H_4	Cyclobutadiene	3383.7	17.5	−159.0	2.0	30.9	−2.7	−1.4	−4.3	5.6	3272.3			0.031
70	C_4H_4	Tetrahedrane	3268.6	20.0	−156.2	2.6	32.1	−3.3	−1.4	−6.4	4.3	3160.4			0.015
71	C_4H_4	Vinylacetylene	3524.1	19.1	−158.9	1.9	29.2	−2.6	−1.4	−4.8	5.0	3411.6			0.029
72	C_4N_2	Dicyanoacetylene	3258.9	23.5	−67.4	0.5	42.0	−3.3	−1.4	−10.3	13.4	3256.1			0.030
73	FH	Hydrogen fluoride	585.3	0.8	−25.0	0.3	6.6	−0.8	−1.6	−0.4	0.8	565.9	566.0 ± 0.0	−0.1	0.013
74	FHO	Hypofluorous acid	648.8	0.6	−36.5	0.4	12.4	−1.0	−2.5	−1.2	3.3	624.3	624.0 ± 0.4	0.3	0.029
75	FHO_2	Fluoroperoxide	853.8	0.7	−47.4	0.6	20.1	−1.3	−3.5	−2.2	6.6	827.4			0.039
76	FH_2N	Monofluoroamine	1052.9	1.9	−72.7	1.2	17.4	−1.4	−1.6	−0.8	2.1	999.0			0.026
77	FH_3N_2	Fluorohydrazine	1678.7	4.4	−119.2	2.1	30.4	−2.5	−1.6	−1.7	3.5	1594.1			0.034
78	FNO	Nitrosyl fluoride	873.6	1.3	−19.2	0.1	21.2	−1.1	−2.5	−1.6	7.0	878.7			0.050
79	F_2	Difluorine	153.7	−0.3	−5.5	0.0	6.4	−0.2	−3.2	−1.0	3.9	153.8	154.6 ± 0.2	−0.8	0.024
80	F_2N_2	Difluorodiazene (cis)	1019.9	2.2	−31.0	0.3	28.8	−1.7	−3.2	−2.8	7.2	1019.7			0.040
81	F_2N_2	Difluorodiazene (trans)	1013.8	2.3	−30.2	0.3	28.5	−1.8	−3.2	−2.8	6.8	1013.5			0.036
82	F_2O	Difluorine monoxide	372.8	−0.4	−13.7	0.1	14.7	−0.6	−4.2	−1.5	6.2	373.5	373.3 ± 0.7	0.2	0.032
83	F_2O_2	Perfluoroperoxide	605.6	−0.4	−21.5	−0.2	22.8	−0.8	−5.1	−1.7	13.6	612.3	609.7 ± 0.8	2.6	0.048
84	F_3N	Trifluoroamine	835.0	0.1	−27.7	0.2	24.3	−1.7	−4.8	−1.7	5.2	828.9			0.033
85	HNO	Nitrosylhydride	840.7	1.9	−36.4	0.7	17.0	−1.2	−0.9	−1.7	4.2	824.3	823.6 ± 0.1	0.7	0.035
86	HNO_2	Nitrous acid (cis)	1276.2	2.5	−53.2	0.7	24.9	−1.8	−1.9	−2.5	6.7	1251.7	1251.5 ± 0.4	0.2	0.047
87	HNO_2	Nitrous acid (trans)	1276.9	2.6	−53.2	0.7	25.1	−1.8	−1.9	−2.2	6.9	1253.0	1253.3 ± 0.1	−0.3	0.045
88	HNO_2	Nitrous acid, H–NO_2	1243.6	3.7	−57.7	0.5	26.1	−2.4	−1.9	−3.2	7.5	1216.4			0.052
89	HNO_3	Nitric acid	1585.1	4.4	−69.8	0.7	33.3	−3.2	−2.8	−5.0	9.2	1552.0	1551.6 ± 0.2	0.4	0.053
90	HN_3	Hydrogen azide	1348.5	6.3	−55.9	0.3	29.5	−2.1	0.0	−4.3	8.5	1330.8	1329.7 ± 0.6	1.1	0.043
91	H_2N_2	Diazene (cis)	1192.4	3.2	−72.8	1.1	21.2	−1.3	0.0	−1.8	3.3	1145.4	1143.5 ± 0.9	1.9	0.028
92	H_2N_2	Diazene (trans)	1214.2	3.3	−74.4	1.1	21.8	−1.4	0.0	−1.8	3.3	1166.1	1165.8 ± 0.7	0.3	0.028
93	H_2N_2	Diazene (iso)	1109.1	3.8	−70.9	1.6	22.5	−1.6	0.0	−1.0	2.5	1066.1	1065.1 ± 0.9	1.0	0.038
94	H_2N_2O	Nitrosamide	1580.4	5.0	−85.9	1.6	30.2	−2.4	−0.9	−2.4	5.8	1531.4			0.048
95	H_2O	Water	961.9	1.6	−56.5	0.9	11.3	−1.1	−0.9	−0.6	1.1	917.6	917.8 ± 0.1	−0.2	0.014
96	H_2O_2	Hydrogen peroxide	1106.2	1.7	−69.8	1.3	17.3	−1.6	−1.9	−1.6	3.4	1055.0	1055.2 ± 0.1	−0.2	0.019
97	H_3N	Ammonia	1230.0	2.7	−90.7	1.3	14.2	−1.1	0.0	−0.3	0.8	1156.9	1157.3 ± 0.1	−0.4	0.011
98	H_3NO	Ammonia oxide	1370.3	3.1	−107.9	0.9	25.8	−1.9	−0.9	−1.3	2.3	1290.5			0.022
99	H_3NO	Hydroxylamine	1480.0	3.1	−106.4	1.8	22.0	−2.0	−0.9	−1.3	2.5	1398.7	1398.7 ± 0.5	0.0	0.017
100	H_4N_2	Hydrazine	1803.1	4.8	−141.2	2.7	26.0	−2.2	0.0	−1.2	2.0	1694.0	1695.6 ± 0.2	−1.6	0.013
101	N_2	Dinitrogen	933.0	3.4	−14.1	0.1	17.2	−0.6	0.0	−2.4	4.3	940.8	941.1 ± 0.1	−0.3	0.026
102	N_2O	Nitrous oxide	1100.3	5.0	−28.7	0.1	24.8	−1.9	−0.9	−5.0	9.2	1102.8	1102.0 ± 0.1	0.8	0.046
103	N_2O_3	Dinitrogen trioxide	1587.7	4.5	−44.8	0.3	41.5	−2.9	−2.8	−6.7	15.6	1592.4	1591.1 ± 0.2	1.3	0.071
104	N_2O_4	Dinitrogen tetraoxide	1912.7	6.6	−61.1	0.3	49.2	−4.3	−3.7	−9.4	17.8	1908.0	1908.5 ± 0.2	−0.5	0.072
105	O_3	Ozone	583.1	0.2	−17.6	−0.1	19.8	−1.1	−2.8	−4.2	17.6	594.9	596.1 ± 0.1	−1.2	0.057
106	H_2	Dihydrogen	456.6	0.0	−26.4	0.3	2.2	0.0	0.0	0.0	0.0	432.7	432.1 ± 0.0	0.6	0.005

[a] The individual contributions are explained in the text, see Section 2.

[b] Same number and same molecule as in Ref. [30] except for dihydrogen.

[c] The ae-CCSD(T)/cc-pCVTZ equilibrium geometry was used for each molecule.

[d] D1 diagnostic of Jansen and Nielsen [101] at the fc-MP2/def2-QZVPP level.

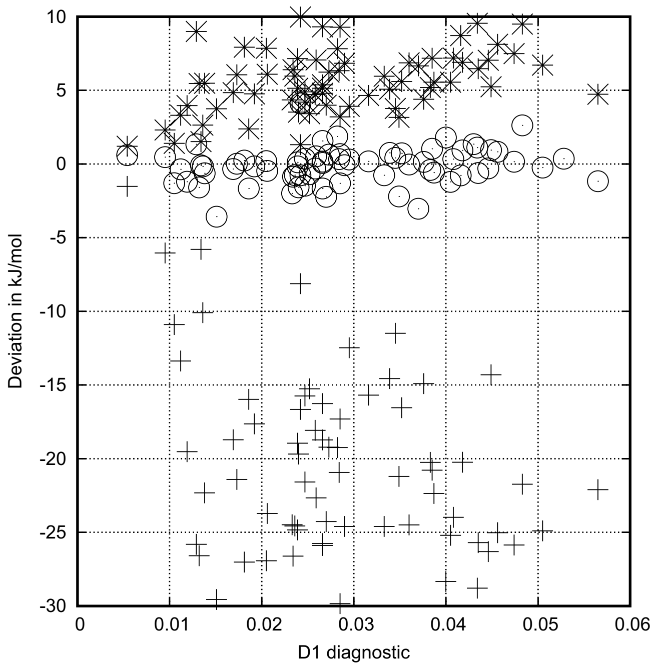


Fig. 1. Deviation of the calculated atomization energy (in kJ/mol) from the ATcT reference value, as a function of the D1 diagnostic. Results are shown for the additivity scheme using either unscaled (∗) or scaled (○) F12 contributions. The deviations are also shown for the additivity scheme *without* F12 correction (+).

Our CCSD(T)/cc-pCVQZ data agree to within 0.36 kJ/mol (RMS error) with the W4 values. Expressed in terms of error per valence electron, the RMS deviation between our values and the W4 values amounts to 0.032 kJ/mol.

We adopt an RMS error of 0.04 kJ/mol per valence electron for the core–valence contribution.

4.2.3. Zero-point vibrational energies

For the same 14 molecules that were discussed in the previous section, Karton et al. [49] report the accurate ZPVE used in W4 theory. For these molecules, our anharmonic zero-point vibrational energies agree with those given in Ref. [49] to within a mean absolute deviation of 0.27 kJ/mol (the RMS deviation amounts to 0.47 kJ/mol). On a per-valence-electron basis, the mean absolute and RMS deviations are 0.028 and 0.055 kJ/mol, respectively.

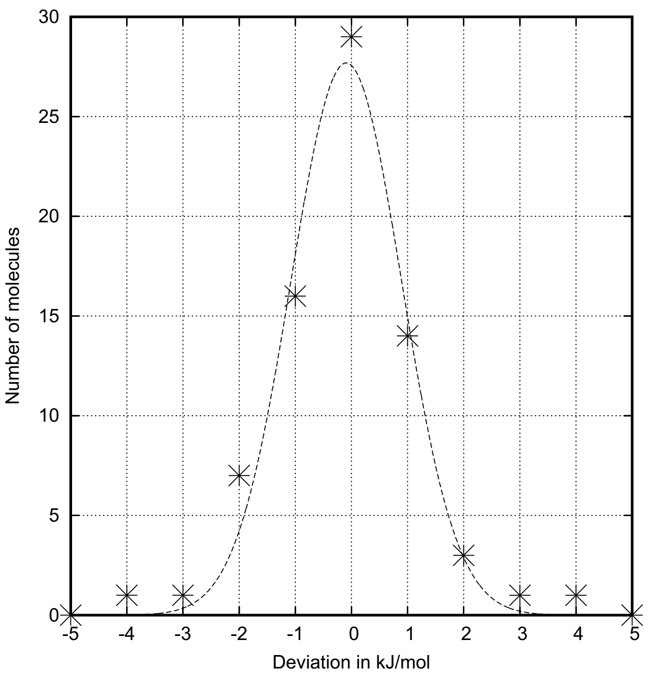


Fig. 2. Distribution of the errors per molecule of 73 molecules (in kJ/mol). Molecules are counted with errors in the same 1 kJ/mol wide interval.

Harding et al. [46] report ZPVEs for the molecules N_2, H_2, F_2, CO, C_2H_2, CO_2, H_2O_2, H_2O, CHN (cyanic acid), FH, HNO, and H_3N. Our values deviate from these ZPVEs with an RMS error of 0.24 kJ/mol (0.023 kJ/mol RMS per valence electron).

We also note that, in our calculation of the anharmonic ZPVE, we have neglected the constant G_0 term in the expression [45]

$$E_{\mathrm{ZPVE}} = G_0 + \sum_i \frac{\omega_i}{2} + \frac{1}{4}\sum_{i\leqslant j} x_{ij}. \tag{8}$$

In Ref. [44], it was found that neglecting the constant G_0 term leads to errors of the order to a few tenths of a kJ/mol in the enthalpies of formation (at 0 K) of molecules/radicals such as C_2H_2, CH_2, HCO, and HO_2.

In view of the agreement with the HEAT (Ref. [46]) and W4 (Ref. [49]) data, and in view of the missing G_0 contribution, we expect that the zero-point vibrational energies reported in Table 1 are accurate to within ±0.06 kJ/mol per valence electron (RMS error).

Table 2
Statistics of the deviations of the computed values from the ATcT reference data (all deviations in kJ/mol).

F12 scaling	N[a]	δ_{ave}[b]	δ_{mad}[c]	δ_{rms}[d]	95% [e]	δ_{max}[f]	Molecule[g]
Errors per molecule							
$f_{\mathrm{int}} = 0.0$	73	−21.7	21.7	23.1	46.3	−52.4	N_2O_4
$f_{\mathrm{int}} = 0.78$	73	−0.12	0.90	1.22	2.44	4.1	C_2H_3F
$f_{\mathrm{int}} = 1.0$	73	5.98	5.98	6.52	13.0	14.1	N_2O_4
Errors per valence electron[h]							
$f_{\mathrm{int}} = 0.0$	73	−1.33	1.33	1.36	2.72	−1.9	H_4N_2
$f_{\mathrm{int}} = 0.78$	73	−0.01	0.06	0.08	0.16	0.2	C_2H_3F
$f_{\mathrm{int}} = 1.0$	73	0.37	0.37	0.38	0.77	0.7	H_2N_2 (cis-diazene)

[a] Number of molecules in assessment.
[b] Mean error.
[c] Mean absolute error.
[d] Root-mean-square error.
[e] 95% confidence limit.
[f] Maximum deviation.
[g] Molecule with largest error.
[h] Statistics of the error per valence electron. For each molecule, the deviation is divided by the number of valence electrons.

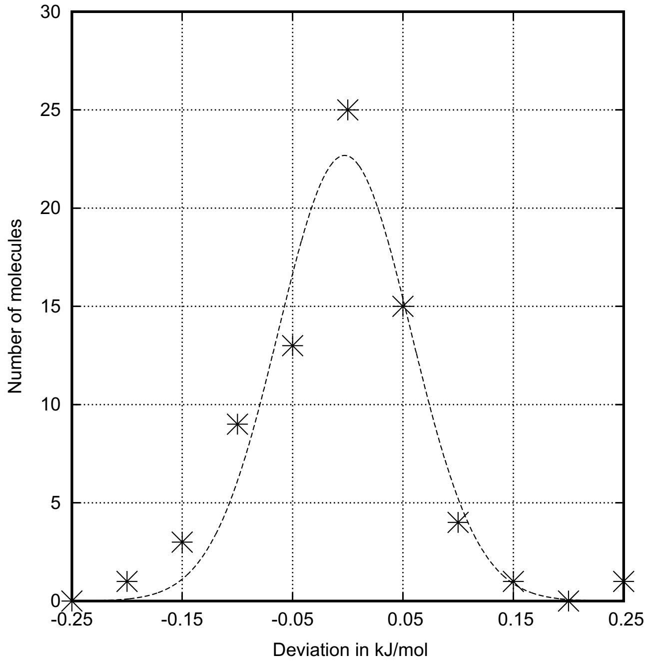


Fig. 3. Distribution of the errors per valence electron of 73 molecules (in kJ/mol). Molecules are counted with errors in the same 0.05 kJ/mol wide interval.

Table 3
Basis-set convergence of the core–valence contribution (kJ/mol) as obtained at the CCSD(T) level.

Nr.[a]	Molecule[b]		cc-pCVQZ	cc-pCV5Z	cc-pV(Q5)Z
1	CFN	Cyanogen fluoride	6.92	7.26	7.61
9	CHN	Hydrogen cyanide	6.59	6.95	7.32
10	CHN	Hydrogen isocyanide	5.78	6.06	6.35
15	CH_2	Singlet methylene	1.59	1.65	1.71
20	CH_2O	Formaldehyde	5.21	5.43	5.65
30	CH_4	Methane	4.99	5.19	5.40
34	CO	Carbon monoxide	3.76	3.96	4.17
35	CO_2	Carbon dioxide	7.00	7.33	7.68
40	C_2H_2	Acetylene	9.60	10.10	10.62
73	FH	Hydrogen fluoride	0.77	0.77	0.77
74	FHO	Hypofluorous acid	0.64	0.62	0.60
79	F_2	Difluorine	−0.29	−0.32	−0.35
92	H_2N_2	Diazene (trans)	3.34	3.44	3.56
95	H_2O	Water	1.61	1.63	1.66
97	H_3N	Ammonia	2.71	2.80	2.88
101	N_2	Dinitrogen	3.38	3.55	3.74
102	N_2O	Nitrous oxide	5.03	5.21	5.41
105	O_3	Ozone	0.24	0.17	0.09

[a] Same number and same molecule as in Ref. [30].
[b] The ae-CCSD(T)/cc-pCVTZ equilibrium geometry was used for each molecule.

4.2.4. Relativistic corrections

For the same 14 molecules that were discussed in Section 4.2.2, the scalar-relativistic corrections (MVD) given in Table 1 agree with their W4 counterparts of Ref. [49] to within a mean absolute deviation of 0.026 kJ/mol (the RMS deviation amounts to 0.037 kJ/mol or 0.0025 kJ/mol per valence electron). In W4 theory, the scalar-relativistic corrections are obtained at the second-order Douglas–Kroll–Hess CCSD(T)/DK-aug-cc-pV(Q+d)Z level. We feel that the two-electron Darwin term as well as higher-order terms such as the Breit interaction and second-order spin–orbit coupling can safely be neglected for the light molecules studied in the present work, and that both our values and the W4 values are accurate to within ±0.003 kJ/mol (RMS) per valence electron. Atomic spin–orbit coupling terms were taken from the experimental fine structure and their uncertainty is insignificant [63]. We adopt an RMS error uncertainty of 0.003 kJ/mol per valence electron for the relativistic corrections.

4.2.5. Full triples and perturbative quadruples

Table 4 shows post-CCSD(T) contributions to the atomization energies of a selected set of 18 molecules, obtained using the cc-pVDZ and cc-pVTZ basis sets at the fc-CCSDT(Q) level. In Table 4, the results shown for X = (DT) refer to the results obtained from the two-point extrapolation procedure of Helgaker et al. [102] using the cc-pVXZ contributions with X = D and T, which is the level of calculation used in W4 theory [49]. Indeed, the CCSDT–CCSD(T) contributions for the molecules N_2 (−3.26 kJ/mol), F_2 (−1.50 kJ/mol), CO (−2.35 kJ/mol), FH (−0.57 kJ/mol), and H_2O (−0.85 kJ/mol) completely agree with the data presented in Ref. [50]. Karton and co-workers not only report CCSDT–CCSD(T) contributions at the (DT) level but also at the extrapolated (TQ) and (Q5) levels, and it seems that the (DT) extrapolation yields useful estimates. The mean and RMS deviations between the (DT) and (Q5) extrapolated data for the 16 molecules studied in Ref. [50] amount to 0.09 and 0.2 kJ/mol, respectively. This accuracy is quite remarkable in view of the large differences between the cc-pVDZ and cc-pVTZ results (Table 4).

Although the (DT) extrapolated data for the CCSDT–CCSD(T) contributions computed by Karton et al. appear to be astonishingly accurate, we nevertheless feel that the difference between the cc-pVDZ and cc-pVTZ basis sets is too large to use them as a basis for a trustworthy extrapolation. Therefore, we have decided to use the plain fc-CCSDT/cc-pVTZ energies. For the same 14 molecules that were discussed in Section 4.2.2, the fc-CCSDT–fc-CCSD(T) contributions in the cc-pVTZ basis deviate from the W4 data (i.e., the (DT) extrapolated values) by 0.7 kJ/mol on average, with the cc-pVTZ data always underestimating the magnitude of the corresponding W4 corrections. The RMS deviation amounts to 0.8 kJ/mol. Hence, we expect that our CCSDT–CCSD(T) contributions may be a few tenths of a kJ/mol up to 1.5 kJ/mol too small in magnitude. Fortunately, this error is partly cancelled by the CCSDT(Q) contribution.

Concerning the CCSDT(Q) contributions for connected quadruple excitations, Karton et al. note that the (DT) extrapolation does more harm than good [49]. Therefore, these authors prefer to use the fc-CCSDT(Q)/cc-pVTZ results scaled by an empirical factor of 1.1. Thus, in W4 theory, the sum of the contributions denoted T and (Q) in Table 1 is computed as

$$\Delta E_{\mathrm{T+(Q)}}(\mathrm{W4}) = E_{\text{fc-CCSDT/cc-pV(DT)Z}} - E_{\text{fc-CCSD(T)/cc-pV(DT)Z}} + 1.1 \times \left(E_{\text{fc-CCSDT(Q)/cc-pVTZ}} - E_{\text{fc-CCSDT/cc-pVTZ}}\right). \tag{9}$$

In the present work, however, we compute this sum as

$$\Delta E_{\mathrm{T+(Q)}}(\text{present work}) = E_{\text{fc-CCSDT/cc-pVTZ}} - E_{\text{fc-CCSD(T)/cc-pVTZ}} + E_{\text{fc-CCSDT(Q)/cc-pVDZ}} - E_{\text{fc-CCSDT/cc-pVDZ}}. \tag{10}$$

For the 18 molecules of Table 4, the difference between Eqs. (9) and (10) can be computed from the data reported. For these molecules, the mean deviation between Eqs. (9) and (10) amounts to 0.1 kJ/mol, with mean absolute and RMS deviations of 0.40 and 0.60 kJ/mol, respectively. Measured per valence electron, the mean, mean absolute, and RMS deviations amount to 0.015, 0.033, and 0.044 kJ/mol.

We therefore feel that Eq. (10) is sufficiently accurate for our present purposes, partly because there is some fortuitous error compensation between the T and (Q) terms. We adopt an RMS error of 0.05 kJ/mol per valence electron for the combined full-triples-and-perturbative-quadruples contribution.

Table 4
Basis-set convergence of the CCSDT–CCSD(T) and CCSDT(Q)–CCSDT contributions (kJ/mol), obtained in cc-pVXZ basis sets with X = D, T, and (DT).

Nr.[a]	Molecule[b]		ΔE_T			$\Delta E_{(Q)}$			$\Delta E_T + \Delta E_{(Q)}$		
			D	T	(DT)	D	T	(DT)	D	T	(DT)
1	CFN	Cyanogen fluoride	−0.72	−3.14	−4.16	4.91	5.02	5.06	4.19	1.88	0.90
9	CHN	Hydrogen cyanide	−0.64	−2.48	−3.26	3.67	4.03	4.18	3.03	1.55	0.92
10	CHN	Hydrogen isocyanide	0.13	−1.64	−2.39	2.42	2.78	2.93	2.55	1.14	0.55
15	CH_2	Singlet methylene	1.06	0.91	0.84	0.35	0.41	0.44	1.41	1.32	1.28
20	CH_2O	Formaldehyde	0.11	−1.54	−2.23	2.50	2.46	2.44	2.61	0.92	0.21
30	CH_4	Methane	0.21	−0.19	−0.36	0.33	0.30	0.29	0.54	0.11	−0.07
34	CO	Carbon monoxide	0.10	−1.62	−2.35	2.66	2.74	2.77	2.76	1.12	0.42
35	CO_2	Carbon dioxide	−0.76	−3.30	−4.37	5.10	4.91	4.83	4.34	1.60	0.45
40	C_2H_2	Acetylene	−0.60	−2.31	−3.02	2.59	2.99	3.16	1.99	0.69	0.14
73	FH	Hydrogen fluoride	0.04	−0.39	−0.57	0.79	0.46	0.32	0.83	0.07	−0.25
74	FHO	Hypofluorous acid	0.26	−1.16	−1.76	3.30	3.12	3.05	3.56	1.96	1.29
79	F_2	Difluorine	0.31	−0.96	−1.50	3.89	3.82	3.79	4.20	2.86	2.29
92	H_2N_2	Diazene (trans)	0.07	−1.79	−2.58	3.27	3.55	3.66	3.35	1.75	1.08
95	H_2O	Water	0.13	−0.56	−0.85	1.10	0.80	0.68	1.23	0.24	−0.17
97	H_3N	Ammonia	0.37	−0.28	−0.56	0.79	0.69	0.65	1.16	0.41	0.09
101	N_2	Dinitrogen	−0.49	−2.44	−3.26	4.31	4.57	4.67	3.81	2.13	1.41
102	N_2O	Nitrous oxide	−2.04	−5.04	−6.30	9.23	9.47	9.58	7.19	4.43	3.28
105	O_3	Ozone	−0.59	−4.21	−5.74	17.63	18.72	19.18	17.04	14.51	13.44

[a] Same number and same molecule as in Ref. [30].

[b] The ae-CCSD(T)/cc-pCVTZ equilibrium geometry was used for each molecule.

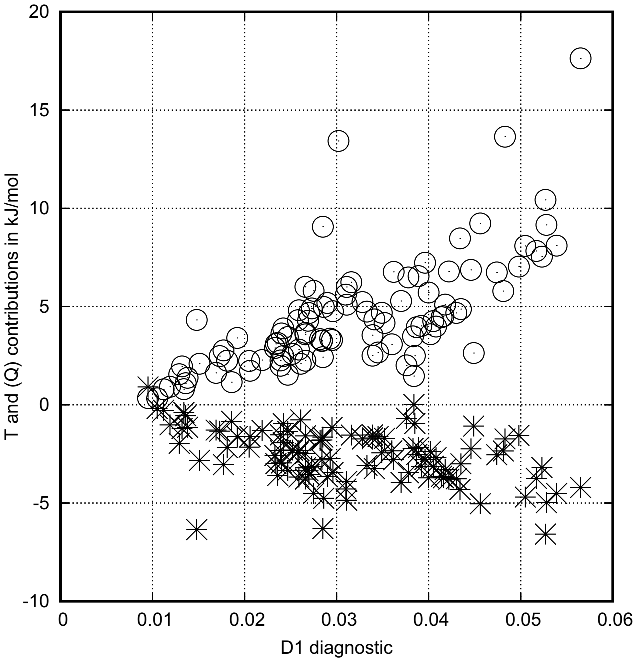


Fig. 4. $\Delta E_T = E_{CCSDT/cc\text{-}pVTZ} - E_{CCSD(T)/cc\text{-}pVTZ}$ (∗) and $\Delta E_{(Q)} = E_{CCSDT(Q)/cc\text{-}pVDZ} - E_{CCSDT/cc\text{-}pVDZ}$ (○) contributions (in kJ/mol) as a function of the D1 diagnostic.

We furthermore note that the T, (Q), and higher excitation terms are likely to become more important for molecules with distinct multireference character. Indeed, Fig. 4 shows that the E_T and $E_{(Q)}$ terms tend to increase in magnitude with increasing D1 diagnostic [101]. In other words, the errors discussed here may not be applicable to molecules with very large D1 diagnostics, for which a single-reference correlation treatment is inadequate.

4.2.6. Neglected contributions

In our additivity scheme, a number of contributions have so far not been taken into account. Among these are valence-shell electron-correlation effects beyond the CCSDT(Q) level, core–valence correlation effects beyond the CCSD(T) level, and non-Born–Oppenheimer effects such as the diagonal Born–Oppenheimer correction (DBOC). What errors can be expected due to neglecting these terms? In Ref. [50], Karton et al. find that in the cc-pVQZ basis, the frozen-core CCSDTQ–CCSDT(Q) contribution amounts to −0.69, −0.41, −0.06, and −0.09 kJ/mol for the molecules N_2, CO, FH, and H_2O, respectively. In the cc-pVDZ basis, the valence–shell correlation contribution due to connected quintuple excitations amounts to 0.48, 0.13, 0.01, and 0.03 kJ/mol, respectively, for the same four molecules. In view of the opposite signs of the CCSDTQ–CCSDT(Q) and connected quintuples contributions, we expect that neglecting these two terms will not give rise to errors significantly larger than ±0.02 kJ/mol per valence electron (RMS error).

Also, core–valence correlation effects beyond the CCSD(T) level are expected to be small (estimated at about 0.03 kJ/mol per valence electron RMS). For the molecules N_2, F_2, CO, FH, H_2O, and C_2H_2, Karton et al. [50] report CCSDT(Q)–CCSD(T) core–valence contributions of 0.14, 0.08, 0.16, 0.02, 0.03, and 0.13 kJ/mol.

The DBOC may also contribute a few tenths of a kJ/mol to the atomization energy of our molecules, in particular to those with X–H bonds (X = C, N, O, F). For a molecule such as *trans*-butadien, for example, a DBOC of the order of 0.45 kJ/mol was computed by Gauss and co-workers at the coupled-cluster level [103]. Using perturbation theory, values of 0.6 and 1.1 kJ/mol were obtained for benzene and naphthalene [104]. Hence, corrections of the order of 0.1 kJ/mol per X–H bond seem quite reasonable, which implies errors of up to a few tenths of a kJ/mol for the 106 molecules of our test set due to neglecting the DBOC. In Ref. [46], Harding et al. have calculated the DBOC of 26 molecules. In their work, neglecting these contributions would have led to an RMS error of 0.14 kJ/mol per molecule, or 0.02 kJ/mol per valence electron. We adopt this latter RMS error as an estimate of our error due to neglecting the DBOC.

4.2.7. Total statistical uncertainty

If we assume that the errors discussed above are statistical in nature, then we may compute the expected accuracy of our computed atomization energy by Gaussian error propagation. For this, we use the following RMS errors per valence electron: ±0.08 kJ/mol for the Hartree–Fock contribution, ±0.04 kJ/mol for the core–valence contribution, ±0.06 kJ/mol for the zero-point vibrational energy, ±0.003 kJ/mol for the relativistic correction, ±0.05 kJ/mol for full triples and perturbative quadruples, ±0.02 kJ/mol for valence-shell post-CCSDT(Q) effects, ±0.03 kJ/mol for core–valence post-CCSD(T) effects, and ±0.02 kJ/mol for the DBOC. The result of the error propagation may be expressed as

$$\sigma = n_{\text{valence electrons}} \times 0.13\ \text{kJ/mol}, \tag{11}$$

where $n_{\text{valence electrons}}$ is the number of electrons in the valence shell. For molecules such as methane, ethane, and propane, Eq. (11) yields total uncertainties of $\sigma = 1.0$, 1.8, and 2.6 kJ/mol, respectively. For diatomics such as N_2 and CO, the uncertainty amounts to $\sigma = 1.3$ kJ/mol. We feel that these are very reasonable estimates of the accuracy of our additivity scheme. Only six of the 105 molecules of the Bakowies test set (H_2 is a somewhat special case) for which reference ATcT values are available, show a deviation from these values larger than the uncertainty given by Eq. (11). None of the molecules except H_2 show a deviation larger than 2σ, which can be interpreted as an estimate of the 95%-confidence limit.

5. Conclusions

The atomization energies of the 105 molecules in the test set of Bakowies [30] have been computed with an estimated standard deviation from the ATcT values of ±0.1 kJ/mol per electron in the valence shell of the molecule. This accuracy has been achieved by adding an empirically scaled MP2-F12 correction for the basis-set truncation error of the cc-pCVQZ basis, in which the ae-CCSD(T) calculations were carried out. Without adding such a correction, the standard deviation would have been as large as 1.3 kJ/mol per valence electron. Hence, the errors were reduced by more than an order of magnitude by the F12 corrections.

In the present work, we have introduced an empirical scaling factor of $f_{\text{int}} = 0.78$ to account for the interference effect in the basis-set truncation error, that is, for the fact that at the level of second-order perturbation theory, the basis-set truncation error is significantly larger (ca. 25%) than at the full configuration-interaction level. However, even without resorting to such an interference factor, that is, by adding 100% of the MP2-F12 correction, the errors in the computed atomization energies would have been reduced already by a factor of 3–4.

In future work, we shall investigate the performance of various coupled-cluster CC-F12 methods to see how these methods could be used to replace the scaled MP2-F12 corrections in the present additivity scheme. When using CC-F12 methods, no empirical factors will be needed, and an additivity scheme using F12 methods and no empirical factors is a very appealing prospect—from the point of view of both theory and efficiency.

Concerning the latter, we notice that our scheme, which includes the ΔE_{F12} correction for the basis-set truncation error, is more efficient than the schemes that are based on CCSD(T) calculations in very large basis sets such as aug-cc-pCVQZ and aug-cc-pCV5Z followed by extrapolation (e.g., in the HEAT345 approach [46]). To demonstrate this, let us take fluoroperoxide (FHO_2) as an example. On our hardware, the fc-CCSD(T) calculations of this molecule in the cc-pCVQZ and aug-cc-pCVQZ basis sets took 4 and 23 h, respectively, on a single processor. The corresponding calculation in the aug-cc-pCV5Z basis has not been carried out but we estimate that the computation time for this calculation would have amounted to about 180 hours. The Hartree–Fock and MP2-F12 calculations (correlation energy only) in the def2-QZVPP basis each took only 6 min while the corresponding fc-CCSD(F12)/def2-QZVPP calculation took 5 h. In other words, by using explicitly-correlated theory, the computation time needed for the fc-CCSD(T) part of the composite scheme can be reduced by a factor of 35–45, from about 180 to 4 or 5 h, for fluoroperoxide. Nevertheless, considerable computation time is still required for the fc-CCSDT/cc-pVTZ and fc-CCSDT(Q)/cc-pVDZ calculations, and for the ae-CCSD(T)/cc-pCVTZ calculation of the harmonic vibrational frequencies. For fluoroperoxide, these calculations took 28, 13 and 40 h, respectively. Of course, as soon as very large basis sets are no longer needed at the CCSD(T) level by virtue of using explicitly-correlated theory, other contributions will become the computationally most demanding steps in the additivity scheme.

Acknowledgements

W.K. gratefully acknowledges support by the DFG through Grant No. KL 721/2-2 and additional financial support by the Fonds der Chemischen Industrie. D.P.T. gratefully acknowledges support by the DFG through Grant No. TE 644/1-1. F.A.B. gratefully acknowledges support by the Fonds der Chemischen Industrie through a Chemiefonds stipend. The work at Argonne National Laboratory was performed under the auspices of the U.S. Department of Energy, Division of Chemical Sciences, Geosciences, and Biosciences of the Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. Portions of the research presented in this paper were conducted within the framework of the Task Group of the International Union of Pure and Applied Chemistry, “Selected Free Radicals and Critical Intermediates: Thermodynamic Properties from Theory and Experiment” (IUPAC Project 2003-024-1-100).

References

[1] W. Kutzelnigg, Theor. Chim. Acta 68 (1985) 445.
[2] W. Klopper, W. Kutzelnigg, Chem. Phys. Lett. 134 (1987) 17.
[3] W. Kutzelnigg, W. Klopper, J. Chem. Phys. 94 (1991) 1985.
[4] D.P. Tew, W. Klopper, F.R. Manby, J. Chem. Phys. 127 (2007) 174105.
[5] W. Klopper, R. Röhse, W. Kutzelnigg, Chem. Phys. Lett. 178 (1991) 455.
[6] J. Noga, W. Kutzelnigg, W. Klopper, Chem. Phys. Lett. 199 (1992) 497.
[7] J. Noga, W. Kutzelnigg, J. Chem. Phys. 101 (1994) 7738.
[8] J. Noga, W. Klopper, T. Helgaker, P. Valiron, DIRCCR12-OS, an integral-direct explicitly-correlated coupled-cluster program. See <http://www-laog.obs.ujf-grenoble.fr/~valiron/ccr12>.
[9] W. Klopper, J. Noga, J. Chem. Phys. 103 (1995) 6127.
[10] J. Noga, P. Valiron, Chem. Phys. Lett. 324 (2000) 166.
[11] D.P. Tew, W. Klopper, M. Heckert, J. Gauss, J. Phys. Chem. A 111 (2007) 11242.
[12] J. Aguilera-Iparraguirre, A.D. Boese, W. Klopper, B. Ruscic, Chem. Phys. 346 (2008) 56.
[13] J. Aguilera-Iparraguirre, H.J. Curran, W. Klopper, J.M. Simmie, J. Phys. Chem. A 112 (2008) 7047.
[14] W. Klopper, C.C.M. Samson, J. Chem. Phys. 116 (2002) 6397.
[15] S. Ten-no, Chem. Phys. Lett. 398 (2004) 56.
[16] H. Fliegl, W. Klopper, C. Hättig, J. Chem. Phys. 122 (2005) 084107.
[17] H. Fliegl, C. Hättig, W. Klopper, Int. J. Quantum Chem. 106 (2006) 2306.
[18] D.P. Tew, W. Klopper, C. Neiss, C. Hättig, Phys. Chem. Chem. Phys. 9 (2007) 1921.
[19] D.P. Tew, W. Klopper, C. Hättig, Chem. Phys. Lett. 452 (2008) 326.
[20] D.P. Tew, J. Chem. Phys. 129 (2008) 014104.
[21] S. Ten-no, J. Chem. Phys. 121 (2004) 117.
[22] S. Ten-no, J. Chem. Phys. 126 (2007) 014108.
[23] T.B. Adler, G. Knizia, H.-J. Werner, J. Chem. Phys. 127 (2007) 221106.
[24] E.F. Valeev, Phys. Chem. Chem. Phys. 10 (2008) 106.
[25] M. Torheyden, E.F. Valeev, Phys. Chem. Chem. Phys. 10 (2008) 3410.
[26] E.F. Valeev, T.D. Crawford, J. Chem. Phys. 128 (2008) 244113.
[27] T. Shiozaki, M. Kamiya, S. Hirata, E.F. Valeev, Phys. Chem. Chem. Phys. 10 (2008) 3358.
[28] T. Shiozaki, M. Kimaya, S. Hirata, E.F. Valeev, J. Chem. Phys. 129 (2008) 071101.
[29] D. Bokhan, S. Ten-no, J. Noga, Phys. Chem. Chem. Phys. 10 (2008) 3320.
[30] D. Bakowies, J. Chem. Phys. 127 (2007) 084105.
[31] T. Helgaker, W. Klopper, D.P. Tew, Mol. Phys. 106 (2008) 2107.
[32] L.A. Curtiss, K. Raghavachari, G.W. Trucks, J.A. Pople, J. Chem. Phys. 94 (1991) 7221.
[33] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, J. Chem. Phys. 109 (1998) 7764.
[34] L.A. Curtiss, P.C. Redfern, K. Raghavachari, J. Chem. Phys. 126 (2007) 084108.
[35] N.J. DeYonker, T.R. Cundari, A.K. Wilson, J. Chem. Phys. 124 (2006) 114104.
[36] N.J. DeYonker, T. Grimes, S. Yockel, A. Dinescu, B. Mintz, T.R. Cundari, A.K. Wilson, J. Chem. Phys. 125 (2006) 104111.
[37] J.A. Montgomery Jr., M.J. Frisch, J.W. Ochterski, G.A. Petersson, J. Chem. Phys. 110 (1999) 2822.

[38] J.A. Montgomery Jr., M.J. Frisch, J.W. Ochterski, G.A. Petersson, J. Chem. Phys. 112 (2000) 6532.
[39] A.L.L. East, W.D. Allen, J. Chem. Phys. 99 (1993) 4638.
[40] A.G. Császár, W.D. Allen, H.F. Schaefer III, J. Chem. Phys. 108 (1998) 9751.
[41] P.L. Fast, M.L. Sánchez, D.G. Truhlar, Chem. Phys. Lett. 306 (1999) 407.
[42] P.L. Fast, J. Corchado, M.L. Sánchez, D.G. Truhlar, J. Phys. Chem. A 103 (1999) 3139.
[43] P.L. Fast, J.C. Corchado, M.L. Sánchez, D.G. Truhlar, J. Phys. Chem. A 103 (1999) 5129.
[44] A. Tajti, P.G. Szalay, A.G. Császár, M. Kállay, J. Gauss, E.F. Valeev, B.A. Flowers, J. Vázquez, J.F. Stanton, J. Chem. Phys. 121 (2004) 11599.
[45] Y.J. Bomble, J. Vázquez, M. Kállay, C. Michauk, P.G. Szalay, Attila G. Császár, J. Gauss, J.F. Stanton, J. Chem. Phys. 125 (2006) 064108.
[46] M.E. Harding, J. Vázquez, B. Ruscic, A.K. Wilson, J. Gauss, J.F. Stanton, J. Chem. Phys. 128 (2008) 114111. Note that the columns "ΔE_{SO}" and "ΔE_{DBOC} have been interchanged in Table 1 of this paper.
[47] J.M.L. Martin, O. de Oliveira, J. Chem. Phys. 111 (1999) 1843.
[48] A.D. Boese, M. Oren, O. Atasoylu, J.M.L. Martin, M. Kállay, J. Gauss, J. Chem. Phys. 120 (2004) 4129.
[49] A. Karton, E. Rabinovich, J.M.L. Martin, B. Ruscic, J. Chem. Phys. 125 (2006) 144108.
[50] A. Karton, P.R. Taylor, J.M.L. Martin, J. Chem. Phys. 127 (2007) 064104.
[51] D. Feller, K.A. Peterson, J. Chem. Phys. 108 (1998) 154.
[52] D. Feller, K.A. Peterson, J. Chem. Phys. 110 (1999) 8384.
[53] D. Feller, D.A. Dixon, J. Phys. Chem. A 103 (1999) 6413.
[54] D. Feller, D.A. Dixon, J. Phys. Chem. A 107 (2003) 9641.
[55] L. Pollack, T.L. Windus, W.A. de Jong, D.A. Dixon, J. Phys. Chem. A 109 (2005) 6934.
[56] M.T. Nguyen, M.H. Matus, W.A. Lester, D.A. Dixon, J. Phys. Chem. A 112 (2008) 2082.
[57] M.R. Nyden, G.A. Petersson, J. Chem. Phys. 75 (1981) 1843.
[58] G.A. Petersson, M.R. Nyden, J. Chem. Phys. 75 (1981) 3423.
[59] G.A. Petersson, S.L. Licht, J. Chem. Phys. 75 (1981) 4556.
[60] J.A. Montgomery Jr., J.W. Ochterski, G.A. Petersson, J. Chem. Phys. 101 (1994) 5900.
[61] G. Tarczay, A.G. Császár, W. Klopper, H.M. Quiney, Mol. Phys. 99 (2001) 1769.
[62] S. Coriani, T. Helgaker, P. Jørgensen, W. Klopper, J. Chem. Phys. 121 (2004) 6591.
[63] J.E. Sansonetti, W.C. Martin, J. Phys. Chem. Ref. Data 34 (2005) 1559. Also available at: <http://physics.nist.gov/PhysRefData/Handbook>.
[64] B. Ruscic, R.E. Pinzon, M.L. Morton, G. von Laszewski, S. Bittner, S.G. Nijsure, K.A. Amin, M. Minkoff, A.F. Wagner, J. Phys. Chem. A. 108 (2004) 9979.
[65] B. Ruscic, Active Thermochemical Tables, in: 2005 Yearbook of Science and Technology (Annual Update to McGraw-Hill Encyclopedia of Science and Technology), McGraw-Hill, New York, 2004, p. 3.
[66] B. Ruscic, R.E. Pinzon, G. von Laszewski, D. Kodeboyina, A. Burcat, D. Leahy, D. Montoya, A.F. Wagner, J. Phys. Conf. Ser. 16 (2005) 561.
[67] B. Ruscic, R.E. Pinzon, M.L. Morton, N.K. Srinivasan, M.-C. Su, J.W. Sutherland, J.V. Michael, J. Phys. Chem. 110 (2006) 6592.
[68] B. Ruscic, unpublished results obtained from Active Thermochemical Tables (ATcT) ver. 1.36 using the Core (Argonne) Thermochemical Network ver. 1.070, 2008.
[69] J.F. Stanton, J. Gauss, J.D. Watts, P.G. Szalay, R.J. Bartlett, A.A. Auer, D.B. Bernholdt, O. Christiansen, M.E. Harding, M. Heckert, O. Heun, C. Huber, D. Jonsson, J. Juselius, W.J. Lauderdale, T. Metzroth, C. Michauk, D.P. O'Neill, D.R. Price, K. Ruud, F. Schiffmann, A. Tajti, M.E. Varner, and J. Vázquez, ACES II, (Mainz–Austin–Budapest 2005 version), a Quantum Chemical Program Package for High-Level Calculations of Energies and Properties, containing the integral packages MOLECULE (J. Almlöf and P.R. Taylor), PROPS (P.R. Taylor) and ABACUS (T. Helgaker, H.J.Aa. Jensen, P. Jørgensen, J. Olsen). For the current version, see <http://www.aces2.de>.
[70] MRCC, a string-based quantum chemical program suite written by M. Kállay. See <http://www.mrcc.hu>.
[71] M. Kállay, P.R. Surján, J. Chem. Phys. 115 (2001) 2945.
[72] G.D. Purvis III, R.J. Bartlett, J. Chem. Phys. 76 (1982) 1910.
[73] J.F. Stanton, J. Gauss, J.D. Watts, R.J. Bartlett, J. Chem. Phys. 949 (1991) 4334.
[74] K. Raghavachari, G.W. Trucks, J.A. Pople, M. Head-Gordon, Chem. Phys. Lett. 157 (1989) 479.
[75] R.J. Bartlett, J.D. Watts, S.A. Kucharski, J. Noga, Chem. Phys. Lett. 165 (1990) 513.
[76] J.F. Stanton, Chem. Phys. Lett. 281 (1997) 130.
[77] J.D. Watts, J. Gauss, R.J. Bartlett, Chem. Phys. Lett. 200 (1992) 1.
[78] J.D. Watts, J. Gauss, R.J. Bartlett, J. Chem. Phys. 98 (1993) 8718.
[79] Y.J. Bomble, J.F. Stanton, M. Kállay, J. Gauss, J. Chem. Phys. 123 (2005) 054101.
[80] M. Kállay, J. Gauss, J. Chem. Phys. 123 (2005) 214105.
[81] W. Klopper, F.R. Manby, S. Ten-no, E.F. Valeev, Int. Rev. Phys. Chem. 25 (2006) 427.
[82] H.-J. Werner, T.B. Adler, F.R. Manby, J. Chem. Phys. 126 (2007) 164102.
[83] TURBOMOLE v5.10, Turbomole GmbH, Karlsruhe, 2008. See <http://www.turbomole.com>.
[84] T.H. Dunning Jr., J. Chem. Phys. 90 (1989) 1007.
[85] D.E. Woon, T.H. Dunning Jr., J. Chem. Phys. 103 (1995) 4572.
[86] J.F. Stanton, C.L. Lopreore, J. Gauss, J. Chem. Phys. 108 (1998) 7190.
[87] D.P. Tew, W. Klopper, J. Chem. Phys. 123 (2005) 074101.
[88] F. Weigend, M. Häser, Theor. Chem. Acc. 97 (1997) 331.
[89] C. Hättig, F. Weigend, J. Chem. Phys. 113 (2000) 5154.
[90] F.A. Bischoff, S. Höfener, A. Glöß, W. Klopper, Theor. Chem. Acc. 121 (2008) 11.
[91] S. Höfener, F.A. Bischoff, A. Glöß, W. Klopper, Phys. Chem. Chem. Phys. 10 (2008) 3390.
[92] F.R. Manby, J. Chem. Phys. 119 (2003) 4606.
[93] E.F. Valeev, Chem. Phys. Lett. 395 (2004) 190.
[94] W. Klopper, Chem. Phys. Lett. 186 (1991) 583.
[95] S. Kedžuch, M. Milko, J. Noga, Int. J. Quantum Chem. 105 (2005) 929.
[96] F. Weigend, R. Ahlrichs, Phys. Chem. Chem. Phys. 7 (2005) 3297.
[97] C. Hättig, Phys. Chem. Chem. Phys. 7 (2005) 59.
[98] F. Weigend, J. Comp. Chem. 29 (2008) 167.
[99] R.D. Cowan, D.C. Griffin, J. Opt. Soc. Am. 66 (1976) 1010.
[100] G.E. Scuseria, J. Chem. Phys. 94 (1991) 442.
[101] C.L. Janssen, I.M.B. Nielsen, Chem. Phys. Lett. 290 (1998) 423.
[102] T. Helgaker, W. Klopper, H. Koch, J. Noga, J. Chem. Phys. 106 (1997) 9639.
[103] J. Gauss, A. Tajti, M. Kállay, J.F. Stanton, P.G. Szalay, J. Chem. Phys. 125 (2006) 144111.
[104] A. Tajti, P.G. Szalay, J. Gauss, J. Chem. Phys. 127 (2007) 014101.

J. Phys. Chem. A **2010,** *114,* 13093–13103

High-Accuracy Thermochemistry of Atmospherically Important Fluorinated and Chlorinated Methane Derivatives

József Csontos,* Zoltán Rolik, Sanghamitra Das, and Mihály Kállay

Department of Physical Chemistry and Materials Science, Budapest University of Technology and Economics, Budapest P. O. Box 91, H-1521 Hungary

Received: January 8, 2010; Revised Manuscript Received: September 24, 2010

High-precision quantum chemical calculations have been performed for atmospherically important halomethane derivatives including CF, CF_3, CHF_2, CH_2F, CF_2, CF_4, CHF, CHF_3, CH_3F, CH_2F_2, CCl, CCl_3, $CHCl_2$, CH_2Cl, CCl_2, CCl_4, CHCl, $CHCl_3$, CH_3Cl, CH_2Cl_2, CHFCl, CF_2Cl, $CFCl_2$, CFCl, $CFCl_3$, CF_2Cl_2, CF_3Cl, $CHFCl_2$, CHF_2Cl, and CH_2FCl. Theoretical estimates for the standard enthalpy of formation at 0 and 298.15 K as well as for the entropy at 298.15 K are presented. The determined values are mostly within the experimental uncertainty where accurate experimental results are available, while for the majority of the considered heat of formation and entropy values the present results represent the best available estimates.

Introduction

As stated in the Fourth Assessment Report of the Intergovernmental Panel on Climate Change (IPCC), "The observed widespread warming of the atmosphere and ocean, together with ice mass loss, support the conclusion that it is *extremely unlikely* that global climate change of the past 50 years can be explained without external forcing and *very likely* that it is not due to known natural causes alone."[1] Emission of greenhouse gases (GHGs) and ozone depletion agents as a byproduct of human activities is among the most important factors which drive climate change.

To facilitate the discussion of global warming, the concept of radiative forcing (RF) was introduced. It is used to measure how the energy balance of the Earth−atmosphere system is affected by a given factor, such as volcanic eruptions, solar irradiance, or GHG emission. The best estimate for the total net RF due to human activities is $+1.6$ W/m^2.[2] Most of the positive RF is originated from the emissions of GHGs, such as CO_2 (1.66 W/m^2), CH_4 (0.48 W/m^2), N_2O (0.16 W/m^2), and haloalkanes (0.34 W/m^2). Although, the concentration of haloalkanes is about a millionth of that of carbon dioxide in the atmosphere, haloalkanes have a noticeable 13% share in the positive RF. This quite large contribution can be explained by considering their higher global warming potentials, which are typically a couple of thousand times of that of the reference CO_2.

The other, previously mentioned, global environmental issue, ozone depletion, is also linked to anthropogenic haloalkane emission.[3] Halogen atoms produced from haloalkanes by ultraviolet light efficiently destroy ozone molecules in the stratosphere. Because of their importance, global warming and ozone depletion are covered by international treaties, such as the Kyoto[4] and Montreal[5] protocols, respectively. Chlorofluorocarbons (CFCs), hydrochlorofluorocarbons (HCFCs), and halons are subject of the Montreal protocol as ozone depletion substances, while hydrofluorocarbons (HFCs) and perfluorocarbons are covered by the Kyoto protocol. Nevertheless, global warming and ozone depletion are not independent processes. For instance, CFCs are primarily responsible for ozone depletion but they are the largest contributors to RF, too, or HCFC emission is controlled by the Montreal protocol but HCFCs are even more dangerous as GHGs since they have large global warming potentials and considerably smaller ozone depletion potential. Further connection is that ozone depletion eventually causes indirect negative RF. It deserves attention that the two, probably, greatest challenges for mankind in the 21st century are linked to the same chemical class of molecules, haloalkanes.[6]

To understand and predict how human activities alternate the climate, several climate models have been developed. Among them the so-called chemistry-climate models[3,7] take into account the chemical processes taking place in the atmosphere. The precise knowledge of the thermodynamic and kinetic properties of the atmospheric reactions as well as of the physical parameters of the corresponding species is required for the chemistry-climate models. Though several databases, including NIST-JANAF,[8] CODATA,[9] ATcT,[10] JPL,[11] and Burcat's Third Millennium Thermodynamic Database,[12] contain the relevant physicochemical parameters, many discrepancies exist and numerous data have considerably large error bars. The source of the latter uncertainty is mostly due to the fact that the chemistry of the troposphere and stratosphere is dominated by free radical reactions and experimental determination of physical and chemical parameters of radicals and radical reactions is still challenging. In general, experimental inaccuracy for open-shell systems is usually larger than that for closed-shell molecules.

In recent years computational chemistry evolved to a stage where the calculation of thermodynamic functions of small molecules are possible with experimental or even higher accuracy. Models, which can achieve this level of accuracy are usually based on the observation that different contributions to the energy reach convergence at different levels of theory. Furthermore, the additivity of the applied approximations is taken for granted, and empirical corrections, calibrated on a reference set of compounds, are frequently used as well. The Gaussian-*n* (G*n*) family of approaches of Pople and co-workers[13,14] was the first among these, so-called,

---

* To whom correspondence should be addressed. E-mail: jcsontos@mail.bme.hu.

10.1021/jp105268m

model chemistries, which primarily were designed for applications in thermochemistry. Pople's idea has served as a template for several other schemes. The complete basis set (CBS) approach of Petersson and associates[15,16] was developed to correct those errors which are originated from the truncation of one-electron basis sets. More accurate model chemistries introduced recently take aim at the ambitious 1 kJ/mol accuracy range, include also relativistic, post-Born−Oppenheimer, and high-order correlation effects, and furthermore avoid the use of empirical parameters. This group includes the focal point approach proposed by Allen and associates,[17,18] the W*n* theories of Martin et al.,[19−21] methods used by Feller and co-workers[22] and Helgaker et al.,[23] and the HEAT family of protocols.[24−26] Using the latter methods, it has been proven for several free radicals and reactions of importance in atmospheric chemistry that high-accuracy calculations can provide superior precision to experimental methods.[20,24,27−34]

In this study a model chemistry inspired by the W3 and HEAT protocols is used to determine accurate heats of formation and entropies for atmospherically important fluorinated and chlorinated methane derivatives, and the obtained results are compared to existing thermochemical data.

Theoretical Methods

Total Energies. In any theoretical model chemistry aimed at the calculation of thermochemical properties the basic quantities are the total energies of atoms and molecules. In the present study total energies were calculated invoking coupled-cluster theory (CC)[35] and Dunning's correlation consistent cc-pV*X*Z,[36,37] aug-cc-pV*X*Z,[38] and cc-pCV*X*Z[39] basis sets. Except the geometry optimizations, vibrational frequency, and core electron correlation calculations, the core electrons were kept frozen. In all calculations restricted and unrestricted Hartree−Fock (HF) orbitals were used for closed- and open-shell molecules, respectively.

The structures of the molecules were taken from geometry optimizations carried out by the coupled-cluster singles, doubles, and perturbative triples [CCSD(T)] method[40] with the cc-pVQZ basis set. This approach was demonstrated to give highly accurate geometries[41] and was also validated in several thermochemical studies including the HEAT project.[24]

In the energy calculations the additivity of the various contributions was assumed according to the following scheme

$$E = E_{\mathrm{HF}} + \Delta E_{\mathrm{CCSD(T)}} + \Delta E_{\mathrm{CCSDT}} + \Delta E_{\mathrm{CCSDT(Q)}} + \Delta E_{\mathrm{core}} + \Delta E_{\mathrm{ZPE}} + \Delta E_{\mathrm{DBOC}} + \Delta E_{\mathrm{REL}} \quad (1)$$

In eq 1 E_{HF} is the basis set limit Hartree−Fock (HF) self-consistent field energy. E_{HF} was obtained by extrapolating the aug-cc-pV*X*Z (X = T, Q, 5) HF energies using the three-point exponential formula of Feller.[42] The correlation energy of valence electrons was evaluated as a sum of three contributions, $\Delta E_{\mathrm{CCSD(T)}}$, $\Delta E_{\mathrm{CCSDT}}$, and $\Delta E_{\mathrm{CCSDT(Q)}}$. $\Delta E_{\mathrm{CCSD(T)}}$ is the correlation energy calculated by the CCSD(T) method extrapolated to the basis set limit. For the CCSD(T) correlation energy a two-point extrapolation formula put forth by Helgaker and co-workers was used[43] in conjunction with the aug-cc-pVQZ and aug-cc-pV5Z basis sets. The iterative triples and the perturbative quadruples contributions are defined by $\Delta E_{\mathrm{CCSDT}} = E_{\mathrm{CCSDT}} - E_{\mathrm{CCSD(T)}}$, and $\Delta E_{\mathrm{CCSDT(Q)}} = E_{\mathrm{CCSDT(Q)}} - E_{\mathrm{CCSDT}}$, respectively, where $E_{\mathrm{CCSD(T)}}$ stands for the CCSD(T) total energy, and E_{CCSDT} denotes the CC singles, doubles, and triples (CCSDT) total energy,[44] while $E_{\mathrm{CCSDT(Q)}}$ is the energy calculated by the CCSDT with a perturbative treatment of quadruples [CCSDT(Q)] approach.[45,46] The triples and quadruples contributions were calculated with the cc-pVTZ and cc-pVDZ basis sets, respectively, and not extrapolated.

To evaluate the core correlation energy, ΔE_{core}, frozen-core (FC), and all-electron (AE) CCSD(T) calculations were performed with the cc-pCVTZ and cc-pCVQZ basis sets. The difference of the AE and FC values was extrapolated to the basis set limit using the aforementioned two-point formula.

According to vibrational perturbation theory[47] the zero-point vibrational energy (ZPE) is given by

$$\Delta E_{\mathrm{ZPE}} = G_0 + \sum_i \frac{\omega_i}{2} + \sum_{i \geq j} \frac{x_{ij}}{4} \quad (2)$$

where G_0 is a constant term independent of the vibrational level, ω_i are the harmonic frequencies, x_{ij} are anharmonicity constants, and the summation runs through all vibrational modes. The ZPEs were determined from CCSD(T) calculations correlating all electrons. For harmonic frequencies the cc-pVQZ basis set and analytic second derivative techniques were used,[48,49] while the G_0 term and the anharmonicity constants were taken from cc-pVTZ cubic force fields obtained by numerical differentiation of analytic second derivatives.[25,50]

The deficiencies of the Born−Oppenheimer (BO) approximation were corrected by adjusting the energy with the diagonal BO correction (DBOC) calculated at the CCSD level with the cc-pCVTZ basis and the formalism of Gauss et al.[51]

The relativistic contributions (ΔE_{REL}) were taken into account by evaluating the expectation value of the mass-velocity and one- and two-electron Darwin operators at the CCSD(T)/cc-pCVTZ level. The energy lowering of the lowest spin−orbit (SO) state with respect to the energy evaluated within a nonrelativistic approximation was also included for particular species. For the carbon atom and the CCl molecule these contributions were calculated from the experimental fine-structure splittings available in the NIST Atomic Spectra Database,[52] and for the CF molecule the theoretical value of $-0.000178\ E_{\mathrm{h}}$ from ref 24 was taken.

CCSDT(Q) calculations were carried out with the MRCC suite of quantum chemical programs[54] interfaced to the CFOUR package,[55] while all other calculations came from CFOUR.

Thermodynamic Functions. Relying on the calculated total energies, harmonic frequencies, and rotational constants standard enthalpies (H_T°) and entropies (S_T°) were computed at $T = 0$ and 298.15 K at a pressure of 1 bar via the standard formulas of statistical thermodynamics (STD) within the ideal gas approximation.[56]

In the case of the systems where the electronic ground state splits due to spin−orbit interaction the spin−orbit states of the ground state were considered at the calculation of partition functions. For the other molecules excited states lie far above the ground state and were not taken into account for the calculation of thermodynamic functions.

For the rotational and vibrational degrees of freedom the rigid-rotor, harmonic oscillator (RRHO) approximation was invoked. Exceptions are those four-atom systems, where the inversion mode was treated by approximately determining the lowest eigenvalues of a one-dimensional effective Schrödinger equation of the inversion motion. This investigation was conducted on those molecules where (i) the barrier between the two minima of the large amplitude motion was low compared to the characteristic energy of the linear oscillator fitted to a single valley of the potential energy surface (PES) and (ii) the normal mode belonging to the inversion motion was well-defined.

Our approach is based on the work of Rush and Wiberg,[57] who used the zero-order rotational-large-amplitude-motion Hamiltonian[58] to describe the inversion motion of ammonia isotopomers. The reaction coordinate was defined as the distance between the central carbon atom and the plane determined by the ligands. For each investigated point of the reaction coordinate the internal coordinates were relaxed. To the calculated points of the minimal energy path of the inversion motion, an eight-order polynomial was fitted. The calculation of the vibrational−rotational **G** matrix requires the atomic coordinates in the center of mass system as a function of the reaction coordinate (and also the derivative of this function). These data were also fitted by eight-order polynomials. After the determination of the PES and the **G** matrix, the one-dimensional Schrödinger equation was solved numerically using the log derivative method.[59] The obtained vibrational energies were used in the calculation of the heats of formation and entropies. At the calculation of the ZPE via eq 2, (i) the contribution of the harmonic frequency of the inversion motion in the second term was replaced by the lowest solution of the one-dimensional Schrödinger equation and (ii) the diagonal elements of the anharmonicity matrix x_{ii} belonging to the inversion motion were dropped from the third term, while the off-diagonal elements describing the interactions of different normal modes were retained. For the evaluation of entropies as well as temperature corrections to enthalpies the partition functions were explicitly calculated for the inversion motion considering all the eigenvalues of the vibrational Hamiltonian which give a noticeable contribution at 298 K. The harmonic oscillator partition functions for the inversion modes were replaced with the latter partition functions.

Heats of Formation. Heats of formation can be obtained in several ways from the calculated absolute enthalpies. As is common practice in quantum chemistry heats of formation at temperature T ($\Delta_f H_T^\circ$) can be calculated from the corresponding atomization enthalpies. An alternative approach relies on the formation reaction of the species from the elemental reference compounds,[24] that is, for a halomethane $CH_xF_yCl_z$ the

$$C + \frac{x}{2}H_2 + \frac{y}{2}F_2 + \frac{z}{2}Cl_2 = CH_xF_yCl_z \qquad (3)$$

reaction. Since the elemental standard state of carbon is graphite, for which it is currently not possible to perform high-accuracy calculations, we use the gaseous carbon atom as a reference state. A straightforward derivation shows that the heat of formation of the halomethane compound can be evaluated as

$$\Delta_f H_T^\circ(CH_xF_yCl_z) = H_T^\circ(CH_xF_yCl_z) - H_T^\circ(C_{gas}) - \frac{x}{2}H_T^\circ(H_2) - \frac{y}{2}H_T^\circ(F_2) - \frac{z}{2}H_T^\circ(Cl_2) + \Delta_f H_T^\circ(C_{gas}) \qquad (4)$$

As it is pointed out in ref 24, the elemental reaction approach, which is pursued in this study, may be more advantageous than that based on atomization energies because eq 3 is closer to the ideal case of the isodesmic reaction than an atomization process, where all the chemical bonds are broken.

For the heat of formation of the carbon atom at 0 K [$\Delta_f H_0^\circ(C_{gas})$] the ab initio value of ref 30, 711.65 ± 0.32 kJ/mol, was adopted. To calculate the heat of formation at 298.15 K, the thermal corrections ($\Delta_f H_{298}^\circ - \Delta_f H_0^\circ$) were obtained from the NIST-JANAF tables resulting in $\Delta_f H_{298}^\circ(C_{gas}) = 717.13 \pm 0.32$ kJ/mol.

Results

Estimated Errors of the Protocol Used. To validate our computational scheme, we performed test calculations for a set of 26 molecules and atoms for which accurate experimental results (error bars smaller than 1 kJ/mol) are available. A group of 17 first-row species including F, N, O, CO, C_2H_2, CCH, CH_2, CH, CH_3, CO_2, H_2O_2, H_2O, HCO, HF, HO_2, NO, and OH was selected from the original HEAT test set.[24] Since the above species do not contain chlorine atoms, we also collected those chlorine compounds whose heat of formation is precisely known. Only 9 species have been found, namely, Cl, ClO, HCl, ClF, HOCl, CH_3Cl, CH_2Cl_2, $CHCl_3$, and CCl_4.[60] The heats of formation for the above species have been calculated by the present scheme and, for comparison, by the HEAT approach. For the HEAT calculations a slightly modified version of the HEAT345-(Q) protocol[25] was employed: the DBOC contributions were computed at the CCSD instead of HF level of theory, as well as the fine structure splittings for Cl and ClO were taken from experiment.[8] In the HEAT calculations the last four chlorinated methane derivatives were not considered since some of the calculations prescribed by the HEAT protocol are currently not feasible for these molecules.

The calculated heats of formation are compiled in Table 1 together with the statistical measures of error. Considering first-row species, the performance of our protocol is highly satisfactory. On average our calculated heats of formation deviate by about 0.5 kJ/mol from the experimental results, and only by about 0.1 kJ/mol less accurate than the much more expensive HEAT values. All of our calculated results are within 1 kJ/mol except CO_2, the only molecule containing three non-hydrogen atoms, where the deviation is 1.1 kJ/mol. Similar conclusions can be drawn for the chlorine compounds containing only one chlorine atom. The results are satisfactory; in avarge the difference with respect to the experiment and the HEAT results (see values in parentheses in Table 1) is not larger than 0.5 and 0.1 kJ/mol, respectively, and all values are within 1.1 kJ/mol of experiment. The errors for the remaining halomethane derivatives, CH_2Cl_2, $CHCl_3$, and CCl_4, are significantly larger; however, we should note that these are larger systems than the former ones, and obviously the error increases monotonically with the system size.

To reveal the sources of the significant errors for the chlorinated hydrocarbons, we analyzed further the approximations made. Since the size of the larger (hydro)chlorocarbons prohibits calculations which are more accurate than the present ones, the smallest chlorocarbons, CCl and CCl_2, were considered in these studies in addition to the above test suite, and we attempted to deduce conclusions relying on the comparison of the present scheme and the more rigorous but still feasible HEAT approach. Our computational scheme contains three major sources of error: i.e., (i) the core and valence separation and the basis-set error of the CCSD(T) correlation energy, (ii) the basis-set error of the SCF energy, and (iii) the iterative triples correction, which was determined with the relatively small cc-pVTZ basis set. First we considered these three approximations.

The separability of the core and valence correlations and the basis set error of the CCSD(T) contributions were recently

TABLE 1: Heats of Formation at 0 K (kJ/mol) Calculated Using the Present Scheme and the HEAT Protocol in Comparison with Experimental Values

species	this work	HEAT	experiment
F	77.41	77.34	77.11 ± 0.16[a]
N	470.82	470.86	470.59 ± 0.05[b]
O	246.80	246.77	246.84 ± 0.00[b]
CO	−114.18	−113.84	−113.82 ± 0.03[c]
C_2H_2	228.56	229.37	228.82 ± 0.30[c]
CCH	564.64	564.18	563.94 ± 0.31[c]
CH_2	391.28	391.64	390.96 ± 0.27[c]
CH	593.18	593.20	592.96 ± 0.25[c]
CH_3	150.74	150.30	149.97 ± 0.10[c]
CO_2	−394.17	−393.83	−393.11 ± 0.01[c]
H_2O_2	−129.45	−129.20	−129.78 ± 0.07[c]
H_2O	−239.30	−239.16	−238.91 ± 0.03[c]
HCO	41.05	42.06	41.92 ± 0.26[c]
HF	−272.78	−272.39	−272.73 ± 0.24[c]
HO_2	14.56	14.74	15.21 ± 0.25[c]
NO	90.03	90.89	90.59 ± 0.08[c]
OH	36.87	36.85	37.09 ± 0.05[c]
mean absolute error	0.43	0.33	
mean signed error	0.09	−0.13	
rms error	0.51	0.39	
maximum error	1.06	0.72	
	Chlorinated Species		
Cl	120.66	119.35	119.62 ± 0.01[d]
ClO	100.65	101.62	101.04 ± 0.13[e]
HCl	−92.07	−92.24	−91.99 ± 0.10[a]
ClF	−55.53	−55.23	−55.71 ± 0.42[e]
HOCl	−73.87	−73.58	−73.97 ± 0.13[f]
CH_3Cl	−74.35	–	−74.70 ± 0.35[a]
CH_2Cl_2	−86.53	–	−88.55 ± 0.74[a]
$CHCl_3$	−94.64	–	−98.35 ± 0.77[a]
CCl_4	−88.66	–	−93.20 ± 0.55[g]
mean absolute error	1.38 (0.36)[h]	0.39	
mean signed error	−1.28 (−0.17)[h]	−0.19	
rms error	2.10 (0.51)[h]	0.41	
maximum error	4.54 (1.04)[h]	0.58	

[a] ATcT values from ref 12. [b] ATcT values from ref. 24. [c] ATcT values from ref 25. [d] Experimental values from ref 8. [e] Experimental values from ref 67. [f] Experimental values from ref 34. [g] ATcT values from ref 12; temperature correction from ref 76. [h] Calculated excluding CH_3Cl, CH_2Cl_2, $CHCl_3$, and CCl_4.

analyzed by Harding et al.[26] It was demonstrated that the error caused by these approximations is not larger than 0.5 kJ/mol for the atomization energies of the molecules of the HEAT test set. Using the elemental reaction approach an even smaller error is expected. To gain some insight into the effect of the core−valence separation for chlorine-containing species tests have been conducted for our test set and for CCl and CCl_2. In the HEAT heats of formation the CCSD(T) contributions, which were originally determined using all-electron aug-cc-pCV(Q,5)Z-extrapolated CCSD(T) correlation energies, were replaced by the sum of the $\Delta E_{\mathrm{CCSD(T)}}$ and ΔE_{core} correction calculated according to the present protocol. We found that this separation of core and valence correlation changed the mean absolute error, the mean signed error, the root mean square (rms) error, and the maximum error, respectively, to 0.31, −0.12, 0.37, and 0.72 kJ/mol for first-row molecules and to 0.26, −0.16, 0.30, and 0.43 for chlorine compounds. That is, surprisingly the errors are somewhat smaller. Concerning the two chlorocarbon radicals, this approximation caused a slight deviation with respect to the HEAT results, 0.07 and 0.17 kJ/mol, respectively, for $\Delta_f H_0^\circ$(CCl) and $\Delta_f H_0^\circ$(CCl_2).

If the aug-cc-pCV(T,Q,5)Z-extrapolated self-consistent field (SCF) energies are replaced by the aug-cc-pV(T,Q,5)Z ones, the respective statistical errors for first-row compounds are slightly worsened to 0.35, −0.15, 0.42, and 0.80 kJ/mol, while for the chlorine derivatives they are surprisingly reduced to 0.23, −0.09, 0.27, and 0.44 kJ/mol. For $\Delta_f H_0^\circ$(CCl) and $\Delta_f H_0^\circ$(CCl_2) the deviations are again moderate, 0.12 and 0.26 kJ/mol, respectively. We note that the effect of the above two contributions seems to compensate each other for the species of the test set. If both the SCF energy and the CCSD(T) correction are replaced simultaneously in the HEAT energies, we arrive at statistical errors rather close to the original ones. However, for CCl and CCl_2 no such error cancellation occurs and differences of 0.19 and 0.42 kJ/mol are obtained.

To assess the impact of the small basis set used for the CCSDT calculations, we performed a similar statistical analysis. The extrapolated cc-pV(T,Q)Z T-(T) contributions to the total HEAT energies were replaced by the T−(T) increments calculated with the cc-pVTZ basis set. The mean absolute error, the mean signed error, the rms error, and the maximum error changed, respectively, to 0.42, 0.08, 0.52, and 1.47 for first-row and to 0.37, −0.31, 0.46, and 0.81 for chlorine-containing molecules. Using the cc-pVTZ basis set for CCSDT resulted in a value for $\Delta_f H_0^\circ$(CCl) and $\Delta_f H_0^\circ$(CCl_2) which differs by 0.23 and 0.58 kJ/mol, respectively, from the original one. These results suggest that the CCSDT contribution may be the largest error source.

The present computational scheme includes two further approximations with respect to the HEAT345-(Q) protocol; however, their effects are significantly smaller. First, in contrast to HEAT we evaluate the scalar relativistic contribution using the cc-pCVTZ basis set instead of aug-cc-pCVTZ. This simplification causes a negligible loss in accuracy, i.e., less than 0.02 kJ/mol. Second, a triple-ζ basis set is used for the calculation of the anharmonic contributions to the ZPE instead of quadruple-ζ basis sets. Our test calculations for the first-row test set revealed that the anharmonic contribution obtained with the cc-pVTZ basis set usually differs from the cc-pVQZ value only by less than 0.1 kJ/mol, while the effect is negligible for chlorine-containing species. We note that the strongly anharmonic molecules are treated separately and are discussed later. We also note that in contrast to the original HEAT approach[24] we evaluate the contribution of quadruple excitations at the CCSDT(Q) level instead of the parent CCSDTQ approach. This replacement was validated in ref 25 and has been justified in several further studies. The changes in the statistical measures of error are less than 0.1 kJ/mol for the HEAT test set when using CCSDT(Q), and, in fact, the results are usually somewhat better due to fortunate error cancellation.

Finally, if all the aforementioned approximations are introduced, the average loss of accuracy with respect to the more rigorous HEAT approach is about 0.1 kJ/mol for the molecules of the test suite (see above). For CCl and CCl_2 deviations of 0.40 and 0.12 kJ/mol are observed. This means that the errors of the individual contributions cancel each other to some extent both for the molecules of the test set and for the latter chloromethane models (note that the sums of the absolute deviations of the particular contributions are 0.4 and 1.0 kJ/mol for CCl and CCl_2). However, a closer inspection of Table 1 as well as the deviations for CCl and CCl_2 reveals that the error compensation is less pronounced for the chlorine compounds. This is probably a consequence of the relatively large error in the total energy of the chlorine molecule, which is needed for the calculations of the heat of formation of each

TABLE 2: Contributions to the Total Energies for the Species Studied in This Work (All Values in Atomic Units)

species	E_{HF}[a]	$\Delta E_{\mathrm{CCSD(T)}}$[b]	$\Delta E_{\mathrm{CCSDT}}$[c]	$\Delta E_{\mathrm{CCSDT(Q)}}$[d]	ΔE_{core}[e]	ΔE_{ZPE}[f]	ΔE_{DBOC}[g]	ΔE_{REL}[h]	total
CF	−137.239 546	−0.430 313	−0.000 578	−0.000 601	−0.122 843	0.003 004	0.004 383	−0.108 170	−137.894 665
CF_3	−336.302 698	−1.045 206	−0.000 295	−0.001 432	−0.256 846	0.012 362	0.009 677	−0.291 121	−337.875 558
CHF_2	−237.387 791	−0.763 623	−0.000 413	−0.000 985	−0.190 481	0.019 216	0.007 291	−0.199 403	−238.516 190
CH_2F	−138.477 109	−0.481 103	−0.000 474	−0.000 533	−0.124 124	0.024 022	0.004 905	−0.107 696	−139.162 111
CF_2	−236.781 267	−0.758 661	−0.000 402	−0.001 236	−0.189 495	0.007 014	0.007 050	−0.199 695	−237.916 692
CF_4	−435.863 760	−1.365 178	0.000 047	−0.001 819	−0.323 449	0.017 070	0.012 329	−0.382 742	−437.907 502
CHF	−137.826 261	−0.471 279	−0.000 646	−0.000 685	−0.123 053	0.012 316	0.004 797	−0.107 904	−138.512 715
CHF_3	−336.944 336	−1.084 595	−0.000 113	−0.001 417	−0.256 981	0.026 027	0.009 941	−0.291 070	−338.542 544
CH_3F	−139.110 578	−0.520 913	−0.000 321	−0.000 570	−0.124 266	0.038 573	0.005 161	−0.107 669	−139.820 583
CH_2F_2	−238.022 124	−0.802 776	−0.000 237	−0.000 997	−0.190 589	0.032 876	0.007 564	−0.199 364	−239.175 648
CCl	−497.268 428	−0.383 459	−0.001 212	−0.000 612	−0.493 986	0.002 002	0.007 802	−1.461 479	−499.599 371
CCl_3	−1416.398 966	−0.901 905	−0.001 253	−0.001 346	−1.370 717	0.007 379	0.019 956	−4.350 534	−1422.997 386
$CHCl_2$	−957.465 159	−0.665 302	−0.001 134	−0.000 901	−0.933 093	0.015 511	0.014 135	−2.905 584	−961.941 527
CH_2Cl	−498.525 055	−0.430 745	−0.000 874	−0.000 490	−0.495 410	0.023 347	0.008 319	−1.460 743	−500.881 651
CCl_2	−956.821 767	−0.662 933	−0.001 257	−0.001 229	−0.931 878	0.004 224	0.013 895	−2.906 100	−961.307 045
CCl_4	−1875.940 341	−1.179 214	−0.000 813	−0.001 848	−1.808 324	0.009 755	0.026 037	−5.795 691	−1884.690 439
CHCl	−497.863 344	−0.422 065	−0.001 165	−0.000 685	−0.494 279	0.011 242	0.008 208	−1.461 046	−500.223 135
$CHCl_3$	−1417.021 008	−0.938 922	−0.000 885	−0.001 365	−1.370 732	0.019 764	0.020 214	−4.350 622	−1423.643 557
CH_3Cl	−499.157 187	−0.468 918	−0.000 647	−0.000 517	−0.495 522	0.037 486	0.008 559	−1.460 758	−501.537 503
CH_2Cl_2	−958.092 545	−0.701 980	−0.000 822	−0.000 922	−0.933 116	0.029 327	0.014 390	−2.905 637	−962.591 305
CHFCl	−597.422 945	−0.714 416	−0.000 829	−0.000 928	−0.561 760	0.017 403	0.010 714	−1.552 506	−600.225 265
CF_2Cl	−696.329 470	−0.997 572	−0.000 707	−0.001 373	−0.628 070	0.010 537	0.013 106	−1.644 292	−699.577 841
$CFCl_2$	−1056.361 068	−0.949 912	−0.001 033	−0.001 349	−0.999 363	0.008 861	0.016 532	−2.997 428	−1061.284 760
CFCl	−596.796 330	−0.710 625	−0.000 927	−0.001 232	−0.560 629	0.005 494	0.010 474	−1.552 923	−599.606 697
$CFCl_3$	−1515.914 658	−1.226 025	−0.000 733	−0.001 789	−1.437 073	0.011 387	0.022 612	−4.442 437	−1522.988 716
CF_2Cl_2	−1155.893 930	−1.272 494	−0.000 554	−0.001 759	−1.065 842	0.013 372	0.019 186	−3.089 191	−1161.291 212
CF_3Cl	−795.877 299	−1.318 854	−0.000 288	−0.001 767	−0.694 635	0.015 357	0.015 758	−1.735 955	−799.597 683
$CHFCl_2$	−1056.989 368	−0.987 647	−0.000 720	−0.001 349	−0.999 463	0.021 697	0.016 793	−2.997 431	−1061.937 489
CHF_2Cl	−696.963 964	−1.036 108	−0.000 459	−0.001 363	−0.628 212	0.023 563	0.013 368	−1.644 246	−700.237 420
CH_2FCl	−598.054 148	−0.752 356	−0.000 577	−0.000 943	−0.561 844	0.031 046	0.010 978	−1.552 498	−600.880 343
C	−37.693 785	−0.095 914	−0.000 463	−0.000 021	−0.055 802	0.000 000	0.001 710	−0.016 425	−37.860 701
H_2	−1.133 661	−0.040 912	0.000 000	0.000 000	0.000 000	0.009 929	0.000 522	−0.000 013	−1.164 135
F_2	−198.774 680	−0.625 595	0.000 008	−0.001 694	−0.132 755	0.002 092	0.005 351	−0.183 912	−199.711 184
Cl_2	−919.011 008	−0.498 088	−0.000 814	−0.000 708	−0.875 347	0.001 264	0.012 186	−2.889 943	−923.262 458

[a] E_{HF} was obtained by extrapolating the aug-cc-pV*XZ* (X = T, Q, 5) HF energies. [b] $\Delta E_{\mathrm{CCSD(T)}}$ is the CCSD(T) valence correlation energy extrapolated to the basis set limit using the aug-cc-pV*XZ* (X = Q, 5) basis sets. [c] $\Delta E_{\mathrm{CCSDT}}$ is defined as $E_{\mathrm{CCSDT}} - E_{\mathrm{CCSD(T)}}$ using the cc-pVTZ basis set. [d] $\Delta E_{\mathrm{CCSDT(Q)}}$ is defined as $E_{\mathrm{CCSDT(Q)}} - E_{\mathrm{CCSDT}}$ using the cc-pVDZ basis set. [e] ΔE_{core} was obtained by extrapolating the difference of the frozen-core and all-electron CCSD(T) correlation energies to the basis set limit. The cc-pCVTZ and cc-pCVQZ basis sets were used. [f] The ZPEs were determined from all-electron CCSD(T) calculations. For harmonic frequencies and anharmonic contributions the cc-pVQZ and cc-pVTZ basis sets were used, respectively. [g] ΔE_{DBOC} was calculated at the CCSD/cc-pCVTZ level. [h] ΔE_{REL} was calculated at the CCSD(T)/cc-pCVTZ level, and it includes spin−orbit corrections which are −0.000 178, −0.000 307, and −0.000 135 E_h for CF, CCl, and C, respectively, and vanish for other species.

chlorinated molecule. Since the deviation for the heat of formation of the chlorine atom (see Table 1), i.e., half of the dissociation energy of Cl_2, is 1.3 kJ/mol, the largest for the test set, we have a good reason to suppose that the total energy of the chlorine molecule is affected by a relatively large error.

In conclusion, our observations suggest that (i) the errors of the individual contributions are moderate, but grow with the size of the system, and none of the contributions can be identified as the principal source of error; (ii) the sum of the errors of the individual contributions is significant, however, the errors cancel each other; and (iii) the remarkably larger errors for the chlorocarbon species seem to be caused by the lack of error compensation.

The comparison of the calculated and experimental thermochemical quantities is highly facilitated if sufficient and well-defined error bars are attached to both. While for the experimental results it is standard to give an error estimate, it is not customary to do so for theoretical values. The definition of error bars for theoretical results is a delicate issue. Basically there are two possibilities to estimate the errors of a calculation. On the one hand one can estimate the intrinsic errors of the methods employed and sum up these numbers. In this way usually a far too big error bar is determined since this approach does not account for the cancellation of the errors of the individual contributions. On the other hand one can benchmark the theoretically obtained quantities to experimental ones and derive some error estimate from this analysis. For this purpose the double of the rms deviation calculated for a test set has been proposed, which corresponds to a 95% confidence limit (approximately two standard deviations) used in experimental thermochemistry.[26,34] However, this error bar should be applied to molecules with size similar to those used in the test set since in general it does not seem reasonable to suppose that the size of the error is independent of the size of the molecule. For this reason we attempt to give another definition on a per-atom basis, which is thus intrinsically size-dependent. To that end, we calculated the error per atom for each first-row molecule in our test set (i.e., we divided the error with respect to the experiment by the number of atoms in the molecule) and calculated the rms value for those quantities. For the chlorine compounds, we performed the same procedure but, to be more conservative, we only weighted the errors by the number of chlorine atoms. In this way we obtained rms deviations of 0.20 and 0.76 kJ/mol; consequently we propose to take 0.4 and 1.5 kJ/mol for each first-row and chlorine atom, respectively, and sum up these atomic contributions to calculate the error bar for the given molecule. For instance, our error bar for CHF_2Cl is calculated as 4 × 0.4 kJ/mol + 1 × 1.5 kJ/mol = 3.1 kJ/mol. For all of the 26 molecules in our test set the difference between the current and the experimental results is not larger, usually much smaller, than our error bar.[61] This gives us confidence that our error estimates for the heats of formation are on the conservative side.

In the case of CF, CF_4, CHF, CHF_3, CH_2F_2, CCl, CCl_4, CHCl, $CHCl_3$, CH_3Cl, CH_2Cl_2, $CFCl_3$, CF_3Cl, CF_2Cl_2, and CH_2FCl, very accurate (the uncertainty is less than ±0.5 J/(K·mol)) entropy values are available in the NIST-JANAF database. In these cases the entropies determined with our protocol were compared with those listed in the NIST-JANAF compilation. The root-mean-square error was 0.6 J/(K·mol), and the largest deviation was 1.2 J/(K·mol). Therefore our conservative estimate for the error associated with our protocol is 1.5 J/(K·mol) for the computed S°_{298} data.

Corrections Specific to Four-Atom Radicals. Four-atom radicals with nonplanar equilibrium geometry investigated in this study can be classified into three groups on the basis of the characteristic features of the inversion motion.

(i) The energy barrier along the path of the inversion motion is large in comparison with the energy of the appropriate normal mode. Thus, the contribution of the inversion motion to the heat of formation at room temperature can be treated harmonically. CF_3, CF_2Cl, and $CFCl_2$ belong to this group with an energy barrier of 10 170, 5138, and 2150 cm^{-1}, respectively.

(ii) The energy barrier is low along the path of the inversion motion, and one normal mode gives the main contribution to the inversion motion. In these cases the contributions for $\Delta_f H^\circ_0$, $\Delta_f H^\circ_{298}$, and S°_{298} were approximated by solving the one-dimensional effective Schrödinger equation for the given reaction coordinate. CH_2F and CCl_3 belong to this group with an energy barrier of 99 and 564 cm^{-1}, respectively.

(iii) The energy barrier is low along the path of the inversion motion; however, two normal modes give significant contribution to the inversion motion. Therefore an accurate treatment of these molecules would require the mapping of a two-dimensional PES and the calculation of the eigenvalues of a two-dimensional Schrödinger equation. However, this is out of the scope of the present study; thus we simply used the harmonic frequencies and anharmonicity corrections as described above to calculate the thermodynamic properties of these systems. In turn, the error bars attached to the calculated thermodynamic functions were increased. The estimated increases in the uncertainties are based on the one-dimensional results obtained for CH_2F and CCl_3. In the case of CH_2F the corrections were 0.62, 1.02, and 3.4 J/(K·mol), respectively, for $\Delta_f H^\circ_0$, $\Delta_f H^\circ_{298}$, and S°_{298}, while for CCl_3 the corresponding values are 0.01, 0.07, and 0.46 J/(K·mol), respectively. To be conservative for the corresponding molecules, the error bars were increased by significantly larger numbers than the former values, that is, by 1.0, 2.0, and 5.0 J/(K·mol) for $\Delta_f H^\circ_0$, $\Delta_f H^\circ_{298}$, and S°_{298}, respectively. Radicals with one H atom, CHF_2, $CHCl_2$, and CHFCl, belong to this group, for which the barrier heights are 2490, 133, and 869 cm^{-1}, respectively.

Although CH_2Cl has planar geometry, its out of plane motion has a strong anharmonic character;[62] thus, its more accurate treatment was also required, and the one-dimensional effective Schrödinger equation was solved to calculate the vibrational levels.

TABLE 3: Temperature Corrections and Heats of Formation (kJ/mol) and Entropies (J/(K·mol)) for Species Studied in This Work

		heats of formation		
species	$H^\circ_{298} - H^\circ_0$ [a]	$\Delta_f H^\circ_0$	$\Delta_f H^\circ_{298}$	entropy[b]
CF	9.1	243.3	247.0	212.9
CF_3	11.4	−464.8	−467.6	264.1
CHF_2	10.6	−239.4	−243.0	255.8
CH_2F	10.9	−27.9	−31.2	232.5
CF_2	10.3	−193.6	−193.2	240.6
CF_4	12.6	−927.8	−933.8	260.9
CHF	10.0	148.9	149.0	223.1
CHF_3	11.5	−687.7	−694.9	259.2
CH_3F	10.1	−228.5	−236.9	222.5
CH_2F_2	10.6	−442.6	−450.5	246.3
CCl	9.4	429.6	433.3	224.3
CCl_3	14.2	73.7	73.1	301.3
$CHCl_2$	12.1	91.2	88.8	280.7
CH_2Cl	10.9	119.4	116.0	236.3
CCl_2	11.4	228.9	230.0	264.5
CCl_4	17.1	−88.7	−91.0	308.6
CHCl	10.1	320.1	320.3	234.6
$CHCl_3$	14.1	−94.6	−99.7	294.8
CH_3Cl	10.4	−74.3	−82.6	234.0
CH_2Cl_2	11.8	−86.5	−93.7	269.7
CHFCl	11.2	−64.6	−67.8	273.5
CF_2Cl	12.2	−272.2	−274.4	285.5
$CFCl_2$	13.1	−91.8	−93.3	297.4
CFCl	10.9	31.2	32.0	258.9
$CFCl_3$	15.9	−282.7	−286.0	308.6
CF_2Cl_2	14.8	−487.9	−492.1	300.2
CF_3Cl	13.7	−703.4	−708.6	284.7
$CHFCl_2$	13.2	−277.3	−283.1	292.4
CHF_2Cl	12.3	−475.7	−482.2	280.5
CH_2FCl	11.2	−256.3	−263.9	263.8
C	6.5	0.0	0.0	158.1
H_2	8.7	0.0	0.0	130.3
F_2	8.8	0.0	0.0	202.6
Cl_2	9.2	0.0	0.0	222.8

[a] Including translational, rotational, vibrational, and electronic contributions. Translational temperature corrections are 1.24 kJ/mol for each translational and rotational degree of freedom. Electronic contributions are 0.38, 0.55, and 0.34 kJ/mol for CF, CCl, and C, respectively, and vanish for other species. It also includes corrections for the inversion mode of CH_2F, CCl_3, and CH_2Cl, which are 0.40, 0.07, and −0.84 kJ/mol, respectively. [b] Including the following contributions. Electronic contributions are 11.4, 11.1, and 18.2 J/(K·mol) for CF, CCl, and C, respectively. Furthermore, for polyatomic molecules with doublet ground states the electronic contribution is 5.8 J/(K·mol), and it vanishes for other species. Inversion mode corrections for CH_2F, CCl_3, and CH_2Cl are 2.7, 6.2, and −6.2 J/(K·mol), respectively.

Total Energies and Thermodynamic Functions. As it is can be seen in Table 2, the increment in the CCSD(T) correlation energy due to the addition of hydrogens is small. It is noteworthy but expected that, since only valence electrons are correlated and the fluorine atom is more compact than the chlorine atom, the CCSD(T) contribution for a F atom is larger than that for a Cl atom. Consequently, as it can be seen from Table 2, the CCSD(T) contribution is always larger for a $CH_xF_yCl_z$ molecule than for the corresponding $CH_xCl_yF_z$ molecule if $y > z$ (for instance, CHF_2Cl vs $CHFCl_2$, or CF_4 vs CCl_4). In addition, several regularities can be recognized for fluorocarbons, chlorocarbons, HFCs, hydrochlorocarbons (HCCs), and CFCs. For fluorocarbons (CF_n) and chlorocarbons (CCl_n) the CCSD(T) contribution increases with increasing n. For HFCs, $CH_{4-n}F_n$ (n = 1, ..., 4) and $CH_{3\text{-}n}F_n$ (n = 1, ..., 3), the magnitude of the CCSD(T) contribution always increases with an increasing number of F atoms. [Please note that CF_4 is not a HFC but it can be regarded as a chemically interesting extension to the $CH_{4-n}F_n$ series. A similar convention is followed in other series

TABLE 4: Heats of Formation (kJ/mol) and Entropies (J/(K·mol)) for the Halomethane Derivatives Studied in This Work

species	$\Delta_f H^\circ_0$	$\Delta_f H^\circ_{298}$	S°_{298}	source[a]
CF	242.9 ± 0.6	246.6 ± 0.6		ref 24, this work[b]
		246.9 ± 0.7		ref 12, ATcT
	243.3 ± 0.8	247.0 ± 0.8	212.9 ± 1.5	this work
		246.4 ± 0.8		ref 22, CC
	251.6 ± 8.0	255.2 ± 8.0	213.0 ± 0.0	ref 8
CF_3	−464.8 ± 1.6	−467.6 ± 1.6	264.1 ± 1.5	this work
		−469.0 ± 1.7		ref 22,[c] CC
		−467.4 ± 2.0		ref 12, ATcT
		−465.7 ± 2.1		ref 64
	−467.5 ± 4.2	−470.3 ± 4.2	265.1 ± 4.2	refs 8, 63
		−470.2		ref 65, MP2
CHF_2	−239.4 ± 2.6	−243.0 ± 3.6	255.8 ± 6.5	this work
		−239.0 ± 4.0		ref 66
		−236.6 ± 10.0		ref 69, CCSD(T)
		−254.0	258.5	ref 12
		−241.2		ref 70, QCISD(T)
		−239.8		ref 70, G3
		−247.3		ref 68, BAC-MP4
CH_2F	−27.9 ± 1.6	−31.2 ± 1.6	232.5 ± 1.5	this work
		−28.0 ± 4.2		ref 71, CC
		−32.0 ± 8.0		ref 66
		−26.9 ± 10.0		ref 69, CCSD(T)
		−32.0	236.5	ref 12
		−29.0		ref 70, QCISD(T)
		−27.2		ref 70, G3
		−31.4		ref 68, BAC-MP4
CF_2	−193.6 ± 1.2	−193.2 ± 1.2	240.6 ± 1.5	this work
		−194.1 ± 1.3		ref 22, CC
	−191.7 ± 1.4	−191.3 ± 1.4		ref 12, ATcT
	−182.5 ± 6.3	−182.0 ± 6.3	240.8 ± 0.0	ref 8
CF_4	−927.2 ± 0.5	−933.4 ± 0.5	261.5	ref 12, ATcT
CF_4	−927.2 ± 1.3	−933.2 ± 1.3	261.4 ± 0.3	ref 8
		−933.2 ± 0.8	261.5	refs 9, 11, 108, and 109
	−927.8 ± 2.0	−933.8 ± 2.0	260.9 ± 1.5	this work
		−933.0 ± 4.2		ref 71, CC
		−930.2 ± 10.0		ref 69, CCSD(T)
		−930.5		ref 103, BAC-G2
		−930.7		ref 103, BAC-G2(MP2)
		−931.2		ref 103, BAC-CBS-4
		−933.7		ref 103, BAC-CBS-Q
		−918.7		ref 79, AAC-G2
		−936.9		refs 64 and 78, G3
CHF	148.9 ± 1.2	149.0 ± 1.2	223.1 ± 1.5	this work
		143.1 ± 12.0		ref 74
		108.8 ± 12.6		ref 72
		156.9 ± 18.0		ref 73
	125.2 ± 29.0	125.5 ± 29.0	223.4 ± 0.2	ref 8
		132.6		ref 74, G2
		131.7		ref 75, BAC-MP4
CHF_3	−687.7 ± 2.0	−694.9 ± 2.0	259.2 ± 1.5	this work
		−695.3 ± 2.0		ref 12, ATcT
		−692.9 ± 2.1		refs 11 and 64, ATcT
	−690.1 ± 3.3	−697.1 ± 3.3	259.7 ± 0.3	refs 8 and 63
		−697.1		ref 102, MP4
CH_3F	−228.5 ± 2.0	−236.9 ± 2.0	222.5 ± 1.5	this work
	−231.5 ± 2.7	−239.6 ± 2.7	222.8	ref 12, ATcT
		−236.4 ± 4.2		ref 71, CC
	−229.6 ± 8.0	−238.0 ± 8.0	222.8	refs 11 and 76
		−236.4 ± 10.0		ref 69, CCSD(T)
	−226.3 ± 33.0	−234.3 ± 29.0	222.8	ref 8
		−221.8		ref 77, MP4
		−234.4		ref 78, G3
		−234.5		ref 79, AAC-G2
		−233.7		ref 79, AAC-G2(MP2)
		−233.8		ref 68, BAC-MP4
CH_2F_2	−445.1 ± 0.8	−452.7 ± 0.8	246.6	refs 76 and 110
	−444.7 ± 1.0	−452.7 ± 1.0	246.3	ref 12 and 67
		−451.9 ± 1.2		ref 22, CC
	−443.0 ± 1.7	−450.7 ± 1.7	246.7 ± 0.0	ref 8
	−442.6 ± 2.0	−450.5 ± 2.0	246.3 ± 1.5	this work
		−450.6 ± 4.2		ref 71, CC
		−449.7 ± 10.0		ref 69, CCSD(T)
		−451.6		ref 78, G3
		−444.7		ref 79, AAC-G2
		−444.6		ref 79, AAC-G2(MP2)
		−451.1		ref 68, BAC-MP4
		−433.5		ref 77, MP4
CCl	430.0 ± 1.1	433.7 ± 1.1	224.3 ± 1.5	this work[d]
	431.4 ± 1.7			ref 82, CC
		443.1 ± 13.0		ref 81
	498.3 ± 20.0	502.1 ± 20.0	224.5 ± 0.4	ref 8
	428.9	433.6	224.6	refs 12 and 80, CCSD(T)
CCl_3		71.1 ± 2.5		ref 84
		70.6 ± 3.7		ref 83, ET
	73.7 ± 4.9	73.1 ± 4.9	301.3 ± 1.5	this work
	80.1 ± 8.0	79.5 ± 8.0	296.8 ± 6.3	ref 8
		79.1		ref 65, MP2
$CHCl_2$		89.0 ± 3.0	280.0 ± 7.0	refs 11 and 89
		92.2 ± 3.9		ref 83, ET
		98.3 ± 5.0		ref 86
	91.2 ± 4.8	88.8 ± 5.8	280.7 ± 6.5	this work
		93.3 ± 8.4		ref 88
		92.9 ± 8.4		ref 88
		96.7 ± 10.0		ref 69, CCSD(T)
	97.5	95.8	285.5	ref 12
		100.8		ref 87
		91.8		ref 70, G3
		81.3		ref 65, CBS-RAD
		90.4		ref 85, QCISD(T)
		91.1		ref 70, QCISD(T)
		104.8		ref 65, MP2
CH_2Cl	119.4 ± 2.7	116.0 ± 2.7	236.3 ± 1.5	this work
		117.3 ± 3.1	271.0 ± 7.0	refs 11 and 89
		121.8 ± 4.2		ref 86
		115.5 ± 4.2		ref 71, CCSD(T)
		121.7 ± 4.5		ref 83, ET
		112.5 ± 8.4		ref 88
		119.2 ± 8.4		ref 88
		118.7 ± 10.0		ref 69, CCSD(T)
	122.3	119.2	242.6	ref 12
		117.0		ref 70, QCISD(T)
		116.7		ref 85, QCISD(T)
		118.8		ref 70, G3
		114.6		ref 91, CBS-RAD
		126.9		ref 91, MP2
CCl_2	229.0 ± 1.9	230.1 ± 1.9	264.5 ± 1.5	this work[d]
		213.4 ± 8.4		ref 95
		230.1 ± 8.4		refs 11 and 74, ET
		239.3 ± 16.8		ref 94
	237.3 ± 21.0	238.5 ± 21.0	265.3	ref 8
	230.5	230.7	266.1	ref 12
CCl_4	−93.7 ± 0.6	−95.8 ± 0.6	309.9	ref 76
		−95.6 ± 0.6		ref 108
		−95.4 ± 0.6		ref 12, ATcT
	−93.8 ± 2.1	−96.0 ± 2.1	309.8 ± 0.0	ref 8
		−93.3 ± 4.2		ref 71, CC
	−88.7 ± 6.4	−91.0 ± 6.4	308.6 ± 1.5	this work
		−99.3 ± 10.0		ref 69, CCSD(T)
		−92.9	308.8	ref 108
		−103.0		ref 78, G3
		−84.6		ref 103, BAC-MP4
CHCl	320.1 ± 2.3	320.3 ± 2.3	234.6 ± 1.5	this work
		326.4 ± 8.4		ref 74, ET
		316.7 ± 20.1		ref 73
	334.0 ± 42.0	335.0 ± 42.0	234.9 ± 0.2	ref 8, NF
$CHCl_3$	−98.0 ± 0.8	−102.9 ± 0.8	295.5	refs 76, 108, and 110
	−98.4 ± 0.8	−103.3 ± 0.8	295.9	ref 12, ATcT
	−98.3 ± 1.3	−103.2 ± 1.3	295.6 ± 0.0	ref 8
	−94.6 ± 5.3	−99.7 ± 5.3	294.8 ± 1.5	this work
CH_3Cl		−82.6 ± 0.4		ref 12, ATcT
	−74.0 ± 0.6	−81.9 ± 0.6	234.2	refs 76, 108, and 111
		−82.8 ± 1.7		ref 22, CC
	−75.8 ± 2.1	−83.7 ± 2.1	234.4 ± 0.4	ref 8
	−74.3 ± 3.1	−82.6 ± 3.1	234.0 ± 1.5	this work
		−86.2 ± 4.2		ref 71, CC
		−83.1 ± 10.0		ref 69, CCSD(T)
		−77.6		ref 78, G3
		−83.4		ref 79, AAC-G2
		−81.8		ref 79, AAC-G2(MP2)
		−83.3		ref 77, MP4
		−84.1		ref 103, BAC-MP4
CH_2Cl_2		−95.4 ± 0.7	270.4	ref 12, ATcT
	−88.5 ± 0.8	−95.1 ± 0.8	270.3 ± 0.0	refs 76, 108, and 110
		−92.0 ± 2.1		ref 22, CC
	−86.5 ± 4.2	−93.7 ± 4.2	269.7 ± 1.5	this work
		−93.7 ± 4.2		ref 71, CC
		−93.1 ± 10.0		ref 69, CCSD(T)
		−91.1		ref 78, G3
		−94.0		ref 79, AAC-G2
		−92.9		ref 79, AAC-G2(MP2)

TABLE 4: Continued

species	$\Delta_f H^\circ_0$	$\Delta_f H^\circ_{298}$	S°_{298}	source[a]
		−94.1		ref 77, MP4
		−94.6		ref 103, BAC-MP4
CHFCl	−64.6 ± 3.7	−67.8 ± 4.7	273.5 ± 6.5	this work
		−59.7 ± 10.0		ref 69, CCSD(T)
		−60.7 ± 10.0	266.5	refs 11 and 86
	−80.0 ± 25.0			refs 12 and 67, NF
		−64.4		ref 70, G3
		−63.8		ref 70, QCISD(T)
CF_2Cl	−272.2 ± 2.7	−274.4 ± 2.7	285.5 ± 1.5	this work
		−279.3 ± 8.0		ref 97
	−273.0 ± 25.0	−275.0 ± 25.0	287.4	refs 12 and 67, NF
		−274.7		ref 70, QCISD(T)
		−272.1		ref 70, G3
		−283.6		ref 98, G2
		−276.2		ref 99, MP2
$CFCl_2$	−91.8 ± 3.8	−93.3 ± 3.8	297.4 ± 1.5	this work
		−89.1 ± 10.0		refs 11 and 86
		−100.4 ± 16.7		ref 100
		−105.0 ± 20.0	298.9	refs 12 and 67, NF
		−94.3		ref 70, QCISD(T)
		−93.4		ref 70, G3
		−105.7		ref 65, CBS-RAD
		−100.4		ref 65, MP2
CFCl	31.2 ± 2.3	32.0 ± 2.3	258.9 ± 1.5	this work
		31.0 ± 13.4		refs 11 and 74
		25.9 ± 30.0	259.2	refs 12 and 67, NF
		24.6		ref 74, G2
$CFCl_3$	−282.5 ± 1.7	−285.8 ± 1.7		ref 111
		−285.3 ± 1.7		ref 11
		−284.9 ± 1.7		ref 96
	−282.7 ± 5.3	−286.0 ± 5.3	308.6 ± 1.5	this work
	−285.5 ± 6.3	−288.7 ± 6.3	309.7 ± 0.2	ref 8
		−283.7	309.8	refs 12 and 67
		−296.0		ref 78, G3
		−290.3		ref 103, BAC-G2
		−292.9		ref 103, BAC−CBS-4
		−289.8		ref 103, BAC−CBS-Q
		−290.2		ref 103, BAC-G2(MP2)
		−287.9		ref 102, MP4
CF_2Cl_2	−490.5 ± 2.5	−494.7 ± 2.5		ref 111
		−494.1 ± 2.5	300.7	ref 11
		−493.3 ± 2.5		ref 96
	−487.9 ± 4.2	−492.1 ± 4.2	300.2 ± 1.5	this work
	−487.4 ± 8.0	−491.6 ± 8.0	300.1 ± 0.2	ref 8
		−490.8	300.9	refs 12 and 67
		−492.9		ref 102, MP4
		−495.3		ref 103, BAC-G2
		−495.0		ref 103, BAC-G2(MP2)
		−498.3		ref 103, BAC-CBS-4
		−493.4		ref 103, BAC-CBS-Q
CF_3Cl	−704.9 ± 2.2	−710.0 ± 2.2	285.4	ref 12, ATcT
		−709.2 ± 2.9	285.2	refs 11 and 64
	−703.4 ± 3.1	−708.6 ± 3.1	284.7 ± 1.5	this work
	−702.8 ± 3.3	−707.9 ± 3.3	285.4 ± 0.4	ref 8
		−707.9 ± 3.8		ref 96
		−714.3		refs 64 and 78, G3
		−709.4		ref 103, BAC-G2
		−709.1		ref 103, BAC-G2(MP2)
		−711.4		ref 103, BAC-CBS-4
		−709.2		ref 103, BAC-CBS-Q
$CHFCl_2$	−277.3 ± 4.2	−283.1 ± 4.2	292.4 ± 1.5	this work
		−284.9 ± 8.8	293.2	refs 12 and 96
		−285.0 ± 9.0	293.0	ref 11
	−277.7 ± 13.0	−283.3 ± 13.0	293.3 ± 0.8	ref 8
		−283.7		ref 103 BAC-G2
		−283.6		ref 103, BAC-G2(MP2)
		−282.3		ref 103, BAC-CBS-4
		−280.9		ref 103, BAC-CBS-Q
		−296.2		ref 102, MP4
		−282.4		ref 79, AAC-G2
		−285.8		ref 78, G3
CHF_2Cl	−475.7 ± 3.1	−482.2 ± 3.1	280.5 ± 1.5	this work
	−484.4 ± 2.3	−490.7 ± 2.3	280.9	refs 12 and 106, ATcT
		−483.7 ± 5.6		refs 11 and 105
	−475.3 ± 13.0	−481.6 ± 13.0	281.0 ± 0.8	refs 8 and 104
		−482.3		ref 103, BAC-G2
		−482.2		ref 103, BAC-G2(MP2)
		−481.2		ref 103, BAC-CBS-4
		−480.4		ref 103, BAC-CBS-Q
		−472.8		ref 102, MP4
		−476.4		ref 79, AAC-G2
		−484.5		ref 78, G3
CH_2FCl	−256.3 ± 3.1	−263.9 ± 3.1	263.8 ± 1.5	this work
		−262.5 ± 8.0	264.4	ref 12, G3B3
		−264.4 ± 8.4	264.3	refs 11 and 96, NF
		−262.4 ± 10.0		ref 69, CCSD(T)
	−254.7 ± 13.0	−261.9 ± 13.0	264.4 ± 0.4	refs 8 and 104
		−264.0		ref 103, BAC-G2
		−263.7		ref 103, BAC-G2(MP2)
		−260.6		ref 103, BAC-CBS-4
		−260.3		ref 103, BAC-CBS-Q
		−261.8		ref 79, AAC-G2
		−253.1		ref 102, MP4
		−262.9		ref 78, G3

[a] Unless otherwise noted, the data are obtained from experiment. If a composite scheme is used in a theoretical study, only the highest level method is indicated. For further details on the experimental setup or on the theoretical methods, please refer to the appropriate literature. ATcT, Active Thermochemical Tables; ET, combined experimental and theoretical investigation; NF, empirical result, no direct experimental findings; CC, coupled-cluster-based model chemistry. [b] $\Delta_f H^\circ_0$ is taken from ref 24; $\Delta_f H^\circ_{298}$ is calculated by adjusting $\Delta_f H^\circ_0$ with the temperature correction obtained in this study. [c] The error bar here given by Feller and associates is possibly too conservative in this specific case. [d] The core and valence correlations were not separated as well as the iterative triples contribution was determined with the cc-pVQZ basis set (see text).

as well.] Likewise, for HCCs, $CH_{4-n}Cl_n$ ($n = 1, ..., 4$) and $CH_{3-n}Cl_n$ ($n = 1, ..., 3$), the magnitude of the CCSD(T) contribution increases when the number of chlorine atoms increases in the given series. For CFCs, CF_nCl_{4-n} ($n = 0, ..., 4$) and CF_nCl_{3-n} ($n = 0, ..., 3$), the CCSD(T) contribution increases with increasing number of F atoms. This is the consequence of the compactness of F atom noted previously.

The CCSDT corrections are always negative with the notable exception of CF_4. For fluorocarbons, HFCs, and CFCs several regularities can be noticed. The magnitude of the CCSDT contribution decreases along the CF, CF_2, CF_3, and CF_4 series. Similarly for HFCs, $CH_{4-n}F_n$ ($n = 1, ..., 4$) and $CH_{3-n}F_n$ ($n = 1, ..., 3$), as the number of the attached fluorine atoms increases the magnitude of the CCSDT correction decreases. CFCs, $CCl_{4-n}F_n$ ($n = 0, ..., 4$) and $CCl_{3-n}F_n$ ($n = 0, ..., 3$), behave similarly; the absolute value of the CCSDT contribution decreases with increasing n. Regularity also can be found for HCC radicals, $CH_{3-n}Cl_n$ ($n = 1, ..., 3$), but, in contrast by HFC radicals, the CCSDT contribution increases with the increasing number of Cl atoms. It is also remarkable that HCC radicals follow the same pattern for both the CCSD(T) and CCSDT corrections; the magnitude of the correction increases with an increasing number of Cl atoms. All other molecule series mentioned above for CCSDT behave on the contrary for CCSD(T).

On the basis of the magnitudes of the CCSDT(Q) contributions, molecules can be classified into four disjoint sets. Molecules with one, two, three, and four heavy atoms (in addition to carbon) have CCSDT(Q) corrections in the range of −490, −685; −901, −1236; −1346, −1432; and −1759, −1848 μE_h, respectively. However, inside these ranges the magnitudes of the contributions seem to be independent of the type of the heavy atoms. It is interesting to observe that for first-row molecules the contribution of quadruple excitations is always larger than that of the iterative triples. However, neither this statement nor its contrary is true for molecules containing Cl.

Regarding the size of the core correlation energy, molecules can be grouped into five sets. Molecules without Cl atoms (HFCs) as well as molecules with one, two, three, and four Cl atoms (CCl_4) have core correlation energies in the range of −0.12, −0.33; −0.49, −0.70; −0.93, −1.07; −1.37, −1.44; and larger than −1.80 E_h, respectively. Inside these ranges the size of the contribution is proportional to the number of electrons of the molecule. These two features involve that in each of the CF_n, $CH_{4-n}F_n$ ($n = 1, ..., 4$), $CH_{3-n}F_n$ ($n = 1, ..., 3$), CCl_n, $CH_{4-n}Cl_n$ ($n = 1, ..., 4$), $CH_{3-n}Cl_n$ ($n = 1, ..., 3$), $CF_{4-n}Cl_n$ ($n = 0, ..., 4$), and $CF_{3-n}Cl_n$ ($n = 0, ..., 3$) molecule sets the core correlation energy increases with increasing n. It is also evident that the sum of the CCSD(T), CCSDT, CCSDT(Q), and core contributions increases with the number of electrons of the molecules.

As it is expected, the larger the number of the bonds is and the lighter the attached atoms in the molecule are, the larger the magnitude of the ZPE contribution is. The size of the DBOC contribution is in the range of that of ZPE, and it correlates with the molecular mass.

Considering the size of the relativistic contributions, the HFCs can be separated clearly from those molecules which contain Cl atoms. For HFCs the magnitude of the correction is in the range of −0.1, −0.4 E_h, and it is proportional to the number of F atoms. The average correction is close to −0.1 E_h per F atom. For Cl-containing molecules, similarly to HFCs, the magnitude of the contribution is proportional to the number of Cl atoms. However, the average correction is considerably larger than that for HFCs; it is close to 1.4 E_h per Cl atom.

Thermochemical Data. Our results including the heat of formation and entropy values, as well as temperature corrections and entropy contributions, are summarized in Table 3, while the best available $\Delta_f H^\circ_{298}$, $\Delta_f H^\circ_0$, and S°_{298} values are summarized in Table 4. It can be observed that the present study delivers estimates for the heats of formation of CF_3, CHF_2, CH_2F, CF_2, CHF, CHF_3, CH_3F, CCl, CHCl, CCl, CHCl, CHFCl, CFCl, CFCl, CFCl, CHFCl, CHFCl, and CH_2FCl, which are more accurate than or considerably differ from the previous experimental and theoretical results. The same holds for the entropy values of CF_3, CHF_2, CH_2F, CH_3F, CCl_3, $CHCl_2$, CH_2Cl, CCl_2, CHFCl, CF_2Cl, $CFCl_2$, and CFCl. In these cases our results might supersede the currently accepted reference values and the revision of the latter might be necessary.

In all cases but CF, CH_2F_2, CH_3Cl, and CH_2Cl_2 this study also presents the most accurate theoretical data to date for the investigated species. One of the exceptions, CF, was investigated by Tajti and his associates[24] by applying the original HEAT protocol that can be regarded as a more sophisticated parent version of our current protocol. Feller and associates[22] also calculated the heat of formation for CF, and although they utilized a heptuple-ζ basis for the calculation of the CCSD(T) contribution, the core and valence correlations were calculated separately in contrast to ref 24. In a recent study Harding et al.[26] calculated the atomization energy for this molecule using several approximations and found that the difference between the atomization energies based on aug-cc-CV(4,5)Z and aug-cc-CV(5,6)Z CCSD(T) calculations is only 0.02 kJ/mol, while the separated treatment of core and valence correlations resulted in an error of 0.2 kJ/mol. Since the core−valence separation causes a larger error than that originated from the basis set dependence of the CCSD(T) contribution, the $\Delta_f H^\circ_0$(CF) value of ref 24 is more accurate. According to our analysis the error bar for the latter is 0.6 kJ/mol,[61] which is smaller than that for the present calculation as well as for the most accurate experimental results. The other three molecules, CH_2F_2, CH_3Cl, and CH_2Cl_2, have been studied by Feller and his co-workers[22] using a CC-based composite approach, which is very similar to the protocol used here. Beyond the CCSD(T) level the effects of the triple and quadruple excitations were also considered, as well as relativistic and DBOC corrections were taken into account. The most notable difference, which probably makes their approach more accurate than ours, is the use of the sextuple-ζ basis sets in the CCSD(T) extrapolation.

Concluding Remarks

The well-known thermochemical databases contain heats of formation within an accuracy of several tens of kJ/mol for numerous small species consisting of two to five first- and second-row atoms. In contrast, modern quantum chemistry can provide such quantities in the kJ/mol accuracy range even for somewhat larger molecules. In this study benchmark theoretical calculations have been performed for the thermodynamic functions of 30 fluorinated and chlorinated methane derivatives important for atmospheric and combustion chemistry. The previous experimental and theoretical results have also been presented, and we have pointed out that the revision of the reference values might be necessary in many cases. These results underpin the importance of high-precision quantum chemical methods for calculating accurate thermochemical quantities.

Acknowledgment. Financial support has been provided by the European Research Council (ERC) under the European Community's Seventh Framework Programme (FP7/2007−2013), ERC Grant Agreement No. 200639, and by the Hungarian Scientific Research Fund (OTKA), Grant No. NF72194, M.K. acknowledges the Bolyai Research Scholarship of the Hungarian Academy of Sciences.

Supporting Information Available: Tables giving equilibrium geometries, rotational constants, harmonic frequencies, anharmonic corrections, and inversion barriers and vibrational energy levels for four-atom species. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*; Pachauri, R. K., Reisinger, A., Eds.; IPCC: Geneva, Switzerland, 2008.

(2) Forster, P.; Ramaswamy, V.; Artaxo, P.; Berntsen, T.; Betts, R.; Fahey, D.; Haywood, J.; Lean, J.; Lowe, D.; Myhre, G.; Nganga, J.; Prinn, R.; Raga, G.; Schulz, M.; Dorland, R. V. *Climate Change 2007: The Physical Science Basis*; Cambridge University Press: Cambridge, U.K., 2007.

(3) Nohende Ajavon, A.-L.; Albritton, D. L.; Watson, R. T. *Scientific Assessment of Ozone Depletion: 2006*; Global Ozone Research Monitoring Project, Report No. 50; World Meteorological Organization: Geneva, Switzerland, 2007.

(4) http://unfccc.int/resource/docs/convkp/kpeng.html.

(5) http://ozone.unep.org/Publications/MP_Handbook/Section _1.1_ The_Montreal_Protocol/.

(6) *Safeguarding the Ozone Layer and the Global Climate System: Issues Related to Hydrofluorocarbons and Perfluorocarbons*; Metz, B., Kuijpers, L., Solomon, S., Andersen, S. O., Davidson, O., Pons, J., de Jager, D., Kestin, T., Manning, M., Meyer, L., Eds.; Cambridge University Press: Cambridge, U.K., 2005.

(7) Austin, J.; Shindell, D.; Beagley, S. R.; Brühl, C.; Dameris, M.; Manzini, E.; Nagashima, T.; Newman, P.; Pawson, S.; Pitari, G.; Rozanov, E.; Schnadt, C.; Shepherd, T. G. *Atmos. Chem. Phys.* **2003**, *3*, 1.

(8) Chase, M. W., Jr. *NIST-JANAF Thermochemical Tables*, 4th ed.; Journal of Physical and Chemical Reference Data, Monograph Vol. 9; American Institute of Physics: College Park, MD, 1998.

(9) Cox, J. D.; Wagman, D. D.; Medvedev, V. A. *CODATA Key Values for Thermodynamics*; Hemisphere: New York, 1989. See also: http://www.codata.org/resources/databases/key1.html.

(10) Ruscic, B.; Pinzon, R. E.; Morton, M. L.; von Laszewski, G.; Bittner, S. J.; Nijsure, S. G.; Amin, K. A.; Minkoff, M.; Wagner, A. F. *J. Phys. Chem. A* **2004**, *108*, 9979.

(11) Sander, S. P.; Friedl, R. R.; Golden, D. M.; Kurylo, M. J.; Moortgat, G. K.; Wine, P. H.; Ravishankara, A. R.; Kolb, C. E.; Molina, M. J.; Finlayson-Pitts, B. J.; Huie, R. E.; Orkin, V. L. *Chemical Kinetics and Photochemical Data for Use in Atmospheric Studies*, Evaluation No. 15; NASA Jet Propulsion Laboratory Publication; California Institute of Technology: Pasadena, CA, 2006.

(12) Burcat, A.; Ruscic, B. Third Millennium Ideal Gas and Condensed Phase Thermochemical Database for Combustion with Updates from Active Thermochemical Tables, ANL-05/20 and TAE 960, Technion-IIT; Aerospace Engineering, and Argonne National Laboratory, Chemistry Division, September 2005; ftp://ftp.technion.ac.il/pub/supported/aetdd/thermodynamics mirrored at http://garfield.chem.elte.hu/Burcat/burcat.html (accessed in March 2010).

(13) Curtiss, L. A.; Raghavachari, K.; Trucks, G. W.; Pople, J. A. *J. Chem. Phys.* **1991**, *94*, 7221.

(14) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Rassolov, V.; Pople, J. A. *J. Chem. Phys.* **1998**, *109*, 7764.

(15) Montgomery, J. A., Jr.; Ochterski, J. W.; Petersson, G. A. *J. Chem. Phys.* **1994**, *101*, 5900.

(16) Ochterski, J. W.; Petersson, G. A.; Montgomery, J. A., Jr. *J. Chem. Phys.* **1996**, *104*, 2598.

(17) East, A. L. L.; Allen, W. D. *J. Chem. Phys.* **1993**, *99*, 4638.

(18) Császár, A.; Allen, W. D.; Schaefer, H. F., III. *J. Chem. Phys.* **1998**, *108*, 9751.

(19) Martin, J. M. L.; de Oliveira, G. *J. Chem. Phys.* **1999**, *111*, 1843.

(20) Boese, A. D.; Oren, M.; Atasoylu, O.; Martin, J. M. L.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *120*, 4129.

(21) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. *J. Chem. Phys.* **2006**, *125*, 144108.

(22) Feller, D.; Peterson, K. A.; Dixon, D. A. *J. Chem. Phys.* **2008**, *129*, 204105.

(23) Helgaker, T.; Ruden, T. A.; Jørgensen, P.; Olsen, J.; Klopper, W. *J. Phys. Org. Chem.* **2004**, *17*, 913.

(24) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599.

(25) Bomble, Y. J.; Vazquéz, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2006**, *125*, 064108.

(26) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. *J. Chem. Phys.* **2008**, *128*, 114111.

(27) Flowers, B. A.; Szalay, P. G.; Vazquéz, J.; Stanton, J. F.; Kállay, M.; Gauss, J.; Császár, A. G. *J. Phys. Chem. A* **2004**, *108*, 3195.

(28) Szalay, P. G.; Thøgersen, L. S.; Olsen, J.; Kállay, M.; Gauss, J. *J. Phys. Chem. A* **2004**, *108*, 3030.

(29) Willitsch, S.; Merkt, F.; Kállay, M.; Gauss, J. *Mol. Phys.* **2006**, *104*, 1457.

(30) Tasi, G.; Izsák, R.; Matisz, G.; Császár, A. G.; Kállay, M.; Ruscic, B.; Stanton, J. F. *ChemPhysChem* **2006**, *7*, 1664.

(31) Szalay, P. G.; Tajti, A.; Stanton, J. F. *Mol. Phys.* **2005**, *103*, 2159.

(32) Demaison, J.; Császár, A. G.; Dehayem-Kamadjeu, A. *J. Phys. Chem. A* **2006**, *110*, 13609.

(33) Karton, A.; Taylor, P. R.; Martin, J. M. L. *J. Chem. Phys.* **2007**, *127*, 064104.

(34) Karton, A.; Parthiban, S.; Martin, J. M. L. *J. Phys. Chem. A* **2009**, *113*, 4802.

(35) Gauss, J. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Jorgensen, W. L., Schaefer, H. F., III, Schreiner, P. R., Thiel, W., Eds.; Wiley: New York, 1998; p 615.

(36) Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007.

(37) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1993**, *98*, 1358.

(38) Kendall, R. A.; Dunning, T. H., Jr.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796.

(39) Woon, D. E.; Dunning, T. H., Jr. *J. Chem. Phys.* **1995**, *103*, 4572.

(40) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479.

(41) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. *J. Chem. Phys.* **2001**, *114*, 6548.

(42) Feller, D. *J. Chem. Phys.* **1992**, *96*, 6104.

(43) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. *J. Chem. Phys.* **1997**, *106*, 9639.

(44) Piecuch, P.; Kucharski, S. A.; Bartlett, R. J. *J. Chem. Phys.* **1999**, *110*, 6103.

(45) Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. *J. Chem. Phys.* **2005**, *123*, 054101.

(46) Kállay, M.; Gauss, J. *J. Chem. Phys.* **2008**, *129*, 144101.

(47) Mills, I. M. In *Molecular Spectroscopy: Modern Research*; Rao, K. N., Mathews, C. W., Eds.; Academic Press: New York, 1972.

(48) Gauss, J.; Stanton, J. F. *Chem. Phys. Lett.* **1997**, *276*, 70.

(49) Szalay, P. G.; Gauss, J.; Stanton, J. F. *Theor. Chim. Acta* **1998**, *100*, 5.

(50) Stanton, J. F.; Gauss, J. *Int. Rev. Phys. Chem.* **2000**, *19*, 61.

(51) Gauss, J.; Tajti, A.; Kállay, M.; Stanton, J. F.; Szalay, P. G. *J. Chem. Phys.* **2006**, *125*, 144111.

(52) Ralchenko, Y.; Kramida, A. E.; Reader, J.; NIST ASD Team, NIST Atomic Spectra Database (version 3.1.5), [Online]; National Institute of Standards and Technology, Gaithersburg, MD, 2008. Available at: http://physics.nist.gov/asd3 (accessed July 2009).

(53) Kállay, M.; Surján, P. R. *J. Chem. Phys.* **2001**, *115*, 2945.

(54) Kállay, M. mrcc, a string-based quantum chemical program suite. See also: Reference 53 as well as http://www.mrcc.hu/.

(55) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. Cfour, a quantum chemical program package [with contributions from Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Christiansen, O.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Jusélius, J.; Klein, K.; Lauderdale, W. J.; Matthews, D. A.; Metzroth, T.; O'Neill, D. P.; Price, D. R.; Prochnow, E.; Ruud, K.; Schiffmann, F.; Stopkowicz, S.; Vázquez, J.; Wang, F.; Watts, J. D.], the integral packages MOLECULE [Almlöf, J.; Taylor, P. R.], PROPS [Taylor, P. R.], and ABACUS [Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J.], and ECP routines [Mitin, A. V.; van Wüllen, C.]. For the current version, see: http://www.cfour.de.

(56) McQuarrie, D. A.; Simon, J. D. *Molecular Thermodynamics*; University Science Books: Sausalito, CA, 1999.

(57) Rush, D. J.; Wiberg, K. B. *J. Phys. Chem. A* **1997**, *101*, 3143.

(58) Hougen, J. T.; Bunker, P. R.; Johns, J. W. C. *J. Mol. Spectrosc.* **1985**, *34*, 136.

(59) Johnson, B. R. *J. Chem. Phys.* **1977**, *67*, 4086.

(60) We note that, in principle, the ClNO molecule also satisfies our criterion; i.e., the experimental error bar for its heat of formation is lower than 1 kJ/mol.[8] However, we have found that there is a significant discrepancy between the calculated and experimental values, indicating that the reference value might require revision. Thus, this molecule was not considered in our error analysis.

(61) We note that for the modified HEAT345-(Q) protocol emloyed for the test calculations the corresponding rms values are 0.16 and 0.41 kJ/mol for first-row and chlorine-containing species, respectively. Consequently we propose error bars of 0.3 and 0.8 kJ/mol per first-row and chlorine atoms, respectively, in the case of this protocol. The experimental results are within our error bars for each species of the test set using this definition.

(62) Levchenko, S.; Krylov, A. *J. Phys. Chem. A* **2002**, *106*, 5169.

(63) Syverud, A. N. *Quarterly Technical Report T-0009-2Q-69* under AFRPL Contract No. F04611-67-C-0009, July Technical Report; Dow Chemical Co., 1969.

(64) Ruscic, B.; Michael, J.; Redfern, P.; Curtiss, L.; Raghavachari, K. *J. Phys. Chem. A* **1998**, *102*, 10889–10899.

(65) Olleta, A. C.; Lane, S. I. *Phys. Chem. Chem. Phys.* **2001**, *3*, 811.

(66) Pickard, J. M.; Rodgers, A. S. *Int. J. Chem. Kinet.* **1983**, *15*, 569.

(67) Gurvich, L. V.; Veyts, I. V.; Alock, C. B. *Thermodynamic Properties of Individual Substances*, 4th ed.; Hemisphere: New York, 1992.

(68) Zachariah, M. R.; Westmorel, P. R.; Burgess, D. R., Jr.; Tsang, W.; Melius, C. F. *J. Chem. Phys.* **1996**, *100*, 8737.

(69) Lazarou, Y. G.; Papadimitriou, V. C.; Prosmitis, A. V.; Papagiannakopoulos, P. *J. Phys. Chem. A* **2002**, *106*, 11502.

(70) Schwartz, M.; Peebles, L. R.; Berry, R. J.; Marshall, P. *J. Chem. Phys.* **2003**, *118*, 557.

(71) Feller, D.; Peterson, K. A.; de Jong, W. A.; Dixon, D. A. *J. Chem. Phys.* **2003**, *118*, 3510.

(72) Lias, S. G.; Karpas, Z.; Lieman, J. F. *J. Am. Chem. Soc.* **1985**, *107*, 6089.

(73) Born, M.; Ingemann, S.; Nibbering, N. M. M. *J. Am. Chem. Soc.* **1994**, *116*, 7210.

(74) Poutsma, J. C.; Paulino, J. A.; Squires, R. R. *J. Phys. Chem. A* **1997**, *101*, 5327.

(75) Zachariah, M. R.; Westmoreland, P. R.; Burgess, D. R., Jr.; Tsang, W.; Melius, C. F. *J. Phys. Chem.* **1996**, *100*, 8737.

(76) Rodgers, A. S.; Chao, J.; Wilhoit, R. C.; Zwolinski, B. J. *J. Phys. Chem. Ref. Data* **1974**, *3*, 117.

(77) Ignacio, E. W.; Schlegel, H. B. *J. Phys. Chem.* **1992**, *96*, 5830.

(78) Ma, N. L.; Lau, K.-C.; Chien, S.-H.; Li, W.-K. *Chem. Phys. Lett.* **1999**, *311*, 275.

(79) Kondo, S.; Takahashi, A.; Tokuhashi, K. *J. Hazard. Mater. A* **2002**, *94*, 37.

(80) Kumaran, S.; Su, M.; Lim, K.; Michael, J.; Klippenstein, S.; DiFelice, J.; Mudipalli, P.; Kiefer, J.; Dixon, D.; Peterson, K. *J. Phys. Chem. A* **1997**, *101*, 8653.

(81) Jesinger, R.; Squires, R. *Int. J. Mass Spectrom.* **1999**, *185*, 745.

(82) Dixon, D. A.; Peterson, K. A. *J. Chem. Phys.* **2001**, *115*, 6327.

(83) Seetula, J. A. *Phys. Chem. Chem. Phys.* **2000**, *2*, 3807.

(84) Hudgens, J. W.; Johnson, R. D.; Timonen, R. S.; Seetula, J. A.; Gutman, D. *J. Phys. Chem.* **1991**, *95*, 400.

(85) Skorobogatov, G. A.; Dymov, B. P.; Tschuikow-Roux, E. P. *Russ. J. Gen. Chem.* **2003**, *73*, 218.

(86) Tschuikow-Roux, E.; Paddison, S. *Int. J. Chem. Kinet.* **1987**, *19*, 15.

(87) Price, S. J. W.; Trotman-Dickerson, A. F. *Trans. Faraday Soc.* **1958**, *54*, 1630.

(88) Holmes, J. L.; Lossing, F. P. *J. Am. Chem. Soc.* **1988**, *110*, 7343.

(89) Seetula, J. A. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 3069.

(90) Schwartz, M.; Peebles, L. R.; Berry, R. J.; Marshall, P. *J. Chem. Phys.* **2003**, *118*, 557.

(91) Mayer, P. M.; Parkinson, C. J.; Smith, D. M.; Radom, L. *J. Chem. Phys.* **1998**, *108*, 604.

(92) Grabowski, J. J. In *Advances in Gas Phase Ion Chemistry*, Vol. 1; Adams, N. G., Babcock, L. M., Eds.; JAI Press: Greenwich, CT, 1992.

(93) Kohn, D. W.; Robles, E. S. J.; Logan, C. F.; Chen, P. *J. Phys. Chem.* **1993**, *97*, 4936.

(94) Chen, S.; Wilhoit, R.; Zwolinski, B. *J. Phys. Chem. Ref. Data* **1976**, *5*, 571.

(95) Miyokawa, K.; Tschuikow-Roux, E. *J. Phys. Chem.* **1992**, *96*, 7328.

(96) Cheong, B.-S.; Cho, H.-G. *J. Phys. Chem. A* **1997**, *101*, 7901.

(97) Olleta, A. C.; Lane, S. I. *Phys. Chem. Chem. Phys.* **2001**, *3*, 811.

(98) Foon, R.; Tait, K. B. *J. Chem. Soc., Faraday Trans. 1* **1972**, *68*, 104.

(99) Ignacio, E. W.; Schlegel, H. B. *J. Phys. Chem.* **1992**, *96*, 5830.

(100) Berry, R. J.; Burgess, D. R., Jr.; Nyden, M. R.; Zachariah, M. R.; Melius, C. F.; Schwartz, M. *J. Phys. Chem.* **1996**, *100*, 7405.

(101) Edwards, J. W.; Small, P. A. *Ind. Eng. Chem. Fundam.* **1965**, *4*, 396.

(102) Cox, J.; Pilcher, G. *Thermochemistry of Organic and Organometallic Compounds*; Academic Press: London, 1970.

(103) Domalski, E. S.; Armstrong, G. T. *J. Res. Natl. Bur. Stand., Sect. A* **1967**, *71*, 105.

(104) Greenberg, E.; Hubbard, W. N. *J. Phys. Chem.* **1968**, *72*, 222.

(105) Neugebauer, C. A.; Margrave, J. L. *J. Phys. Chem.* **1958**, *62*, 1043.

(106) $\Delta_f H^\circ_{298}$ is calculated as described in ref 96, but recent auxiliary values were used, $\Delta_f H^\circ_{298}(CF_3Cl) = -710.0$ kJ/mol and $\Delta_f H^\circ_{298}(CCl_4) = -95.8$ kJ/mol. $\Delta_f H^\circ_0$ is calculated by adjusting $\Delta_f H^\circ_{298}$ with the temperature correction obtained in this study.

(107) Although the error bar attached to the ATcT result is, in principle, smaller than ours, there is a significant discrepancy between the ATcT and the JPL value. Since our results support the JPL value but it is more accurate than that, our calculation indicates some problem with the current ATcT value.

(108) Manion, J. A. *J. Phys. Chem. Ref. Data* **2002**, *31*, 123.

(109) Huybrechts, G.; Narmon, M.; Mele, B. V. *Int. J. Chem. Kinet.* **1996**, *28*, 27.

(110) Hu, A. T.; Sinke, G. C. *J. Chem. Thermodyn.* **1969**, *1*, 507.

(111) Fletcher, R. A.; Pilcher, G. *Trans. Faraday Soc.* **1971**, *67*, 3191.

JP105268M