Molpro: a general-purpose quantum chemistry program package

Hans-Joachim Werner,[1]* Peter J. Knowles,[2] Gerald Knizia,[1] Frederick R. Manby[3] and Martin Schütz[4]

Molpro (available at http://www.molpro.net) is a general-purpose quantum chemical program. The original focus was on high-accuracy wave function calculations for small molecules, but using local approximations combined with explicit correlation treatments, highly accurate coupled-cluster calculations are now possible for molecules with up to approximately 100 atoms. Recently, multireference correlation treatments were also made applicable to larger molecules. Furthermore, an efficient implementation of density functional theory is available. © 2011 John Wiley & Sons, Ltd.

INTRODUCTION

Molpro was founded by Wilfried Meyer and Peter Pulay in the late 1960s. At that time, Pulay developed the first analytical gradient code for Hartree–Fock (HF),[1–3] and Meyer his PNO-CEPA (pseudo-natural orbital coupled-electron pair approximation) methods.[4,5] Orbital optimization for multireference cases became possible in 1980 through a new state-averaged, quadratically convergent MCSCF (multiconfiguration self-consistent field) program by Werner and Meyer.[6] In 1980, Werner and Reinsch developed the first internally contracted multireference configuration interaction (IC-MRCI) program.[7] A new generation of programs was founded in 1984 when Werner and Knowles started to collaborate on a new CASSCF (complete active space SCF) program. They combined fast orbital optimization algorithms[6] with determinant-based full CI codes,[8] and additional, more general, unitary group configuration interaction (CI) codes. This resulted in a very quickly converging MCSCF/CASSCF code called MULTI,[9,10] which is still in use today. Subsequently, Knowles and Werner developed a new, more efficient, IC-MRCI method.[11,12] Extensions for accurate treatments of excited states using IC-MRCI followed somewhat later.[13] For brevity, henceforth IC-MRCI will be denoted as MRCI. These MCSCF and MRCI programs formed the basis of the modern Molpro. In the following years, coupled cluster, DFT, and many other programs were added; these will be summarized in the following sections. Many people have contributed to this, and a full author list can be found at http://www.molpro.net. Due to restrictions on length, not all references to original work can be included, and we refer to our cited earlier papers for more information.

QUANTUM CHEMICAL METHODS

Hartree–Fock

Molpro includes closed-shell and open-shell spin-restricted (RHF, ROHF) Hartree–Fock programs, as well as spin-unrestricted Hartree–Fock (UHF). All programs can run with precomputed integrals on disk, or in direct mode in which the integrals are computed on the fly. For both cases parallel implementations are available. The most efficient way to

*Correspondence to: werner@theochem.uni-stuttgart.de

[1]Institut für Theoretische Chemie, Universität Stuttgart, Stuttgart, Germany

[2]School of Chemistry, Cardiff University, Cardiff, UK

[3]School of Chemistry, University of Bristol, Cantocks Close, Bristol, UK

[4]Institut für Physikalische und Theoretische Chemie, Universität Regensburg, Regensburg, Germany

DOI: 10.1002/wcms.82

TABLE 1 | Elapsed times (in seconds on a single node with two Quad-Core Xeon W5590@3.33GHz CPUs) for density fitting Hartree–Fock (DF-HF), DF-MP2, and density fitting spin-restricted (DF-KS) calculations for pregnanediol $C_{21}O_2H_{36}$, using the cc-pVTZ orbital basis set (OBS) (1194 CGTOs) and cc-pVTZ/JKFIT (2897 GTOs, for HF/B3LYP) or TZVPP/JFIT (1523 GTOs, for BLYP)

Cores	DF-HF		DF-MP2	DF-KS/BLYP		DF-KS/B3LYP	
	Energy	Gradient	Energy	Energy	Gradient	Energy	Gradient
1	2701	530	1555	213	130	2860	567
4	754	172	383	54	41	830	184
8	462	103	206	45	26	518	113

In all DF-HF and DF-KS calculations 10 iterations were needed.

carry out HF calculations for large molecules is to use density fitting (DF). This is available for RHF and ROHF, and very efficiently implemented using the new AIC (adaptive integral core) integral routines[14] written by one of us (GK). Example timings are given in Table 1.

Optionally, local density fitting[15,16] can be used to speed up the treatment of exchange, using localized orbitals, and local subsets of fitting functions for each orbital product. In large molecules, this can significantly increase efficiency without much affecting accuracy. For example, the elapsed time for the calculation in Table 1 is roughly halved, with submicro-Hartree errors.

Density Functional Theory

Density functional theory (DFT) is available for spin-restricted (KS) and spin-unrestricted (UKS) cases, either using precomputed integrals, direct integrals or density fitting. Many different functionals with or without exact exchange are available. The functionals are expressed naturally in their mathematical form within the syntax of the Maple[17] symbolic algebra system, which is used to generate both executable Fortran code and documentation.[18] Recent improvements of the DF-KS code include faster integral evaluation, integral caching, and faster evaluation of the exchange-correlation potential. Ongoing efforts also include the development of special code for graphical processing units (GPUs), optionally using Poisson density fitting.[19–21] Table 1 contains a comparison of elapsed times for DF-HF and DF-KS calculations and the corresponding gradients. Unfortunately, for pure functionals the speedup with more than four processors is small, because nonparallelized parts such as diagonalization and direct inversion of the iterative subspace (DIIS) start to dominate. However, because the whole DFT calculation takes less than a minute—for a 59 atomic molecule with triple-zeta basis—this is of little relevance.

Standard Single Reference Correlation Methods

Most standard single reference correlation methods are available in Molpro. For closed shells, MP2, MP3, MP4, CCSD(T), QCISD(T), BCCD(T), CISD, and CEPA (various versions) are available.[22] In addition, singles configuration interaction (CIS) and equation of motion coupled-cluster (EOM-CCSD) can be used to compute excitation energies.[23] For open shells, UMP2 (based on UHF orbitals) and RMP2 (based on ROHF semicanonical orbitals[24]) are implemented. The standard spin-orbital-based RHF–UCCSD(T) method as well as the partially spin-restricted RHF–RCCSD(T) method[25,26] are implemented. All these methods are based on efficient matrix/tensor algorithms that extensively use matrix multiplications, and therefore run at highest possible speed on modern computers, and are parallelized (cf., section *Parallel Execution*). For MP2, the integrals can be computed most efficiently using density fitting approximations. Timing examples for DF-MP2 and CCSD(T) can be found in Tables 1 and 3, respectively.

The MRCC program of Kállay[27,28] is interfaced to Molpro, making it possible to carry out coupled-cluster calculations with higher excitations (CCSDT(Q), CCSDTQ, and many other variants).

MCSCF, CASSCF, and Full CI

Full CI calculations are possible using the determinant-based full CI program of Knowles and Handy.[8] The MCSCF/CASSCF program MULTI[9,10] can use either this algorithm (default for CASSCF), or configuration state functions for general MCSCF calculations. The optimization includes higher than quadratic terms in the orbital rotation parameters, which makes convergence very fast. The determinant-based CASSCF code allows state-averaged calculations even for states of different symmetry and spin. Dipole and transition moments as well as many other

properties can be computed (cf., section *Electrical and Vibrational Properties*). The final orbitals can be transformed to natural, pseudo-canonical, or localized form, either using state-averaged or state-specific density matrices (several different orbital sets can be stored for further use or visualization). The valence bond program CASVB of Cooper et al.[29–31] is also interfaced to MULTI.

Multireference Perturbation Theory

Single-state CASPT2 and RASPT2 Rayleigh–Schrödinger perturbation theory is available in two different implementations: the older RS2 program[32] is based on the MRCI code of Werner and Knowles[11,12] (cf., section *Multireference Configuration Interaction*). This program also allows CASPT3 and multistate (MS) CASPT2 calculations. Level shifts or the IPEA shift[33] can be used to avoid intruder-state problems. In the RS2C program developed by Celani and Werner,[34] the inactive orbitals are treated much more efficiently than before, and density matrices and coupling coefficients are computed only for the active orbital space. Furthermore, most classes of internal and singly external configurations are internally contracted. The resulting speedups make it possible to treat much larger molecules. Recently, explicitly correlated terms have been added to the RS2 program, which strongly reduces basis set truncation errors (cf., section *Explicit Correlation Methods*).

As an alternative to RS2 or RS2C, the NEVPT2 program of Cimiraglia et al.[35–37] is available in Molpro.

Multireference Configuration Interaction

The MRCI program of Werner and Knowles[11,12] uses internally contracted configurations. Their number depends only on the number of correlated orbitals, not on the number of reference configurations, so very large reference spaces can be used. Excited states can be computed using two options[13]: in the first case, all states are treated together, using the union of the contracted configurations from all reference states. This is the most accurate method and allows the proper description of avoided crossings and conical intersections, but the computational effort increases almost quadratically with the number of optimized states. This bottleneck can be avoided using a projection method,[13] in which each state is optimized in turn (each with only one reference state). Size consistency errors can be reduced either by using the Davidson or Pople corrections, or by using the averaged coupled-pair functional (ACPF) or one of its variants (e.g., AQCC).[38]

TABLE 2 | Root mean square errors of various properties computed with VTZ-F12 sets[40,44–46] relative to extrapolated basis set limits

Property	CCSD(T)	CCSD(T)-F12[1]
Atomization energies [kJ mol^{-1}]	26.7	1.3
Reaction energies [kJ mol^{-1}]	19.1	1.3
Electron affinities[1] [meV]	83.1	9.3
Ionization potentials [meV]	105.5	13.3
Equilibrium distances [pm]	0.59	0.05
Vibrational frequencies [cm^{-1}]	21.0	2.6

[1] Using the aug-cc-pVTZ basis set.

An entirely new MRCI program, CIC, was recently developed by Shamasundar et al.[39] using the same contraction scheme as the RS2C program.[34] The resulting method is highly complex but has been implemented using automated techniques described below (cf. section *The Integrated Tensor Framework*). Despite its complexity, the CIC program is much more efficient than the old MRCI. The savings are greatest for large molecules with many inactive orbitals, as typically encountered in transition metal clusters. For example, a calculation for the complex $O_2Cu_2(NH_3)_6$ (584 basis functions, 3540 reference configuration state functions (CSFs), 80 correlated electrons) took 41 min CPU time per iteration. Such calculations cannot currently be performed with any other program. CIC is still limited to single-state calculations, but an extension to the multistate case is under development. Another important recent extension is the addition of explicitly correlated terms to MRCI (section *Explicit Correlation Methods*).

Explicit Correlation Methods

One problem with conventional wave function methods is slow convergence of electron correlation energies with the size of the basis. This can be circumvented using wave functions that depend explicitly on r_{12}. Enormous progress has been made in recent years to develop so called F12-methods.[40,41] Molpro now contains a very wide range of F12 methods that are very efficient, robust, and easy to use. Currently, MP2-F12 (closed-shell),[42] RMP2-F12 (high-spin open-shell),[43] CCSD(T)-F12 (closed-shell[14,44] and open-shell[45,46]), CASPT2-F12,[47] and MS-CASPT2-F12, as well as MRCI-F12 for ground[48] and excited[49] states are available. Local variants[50–55] will be described in the section *Local Correlation Methods for Ground States*. The additional effort for the F12 corrections is small and becomes negligible for larger molecules. Table 2 summarizes the typical improvements for various properties, and Table 3

Molecule	OBS/RI	Symmetry	Iterations	CCSD(T)	CCSD(T)-F12[1]
Ethanol	276/408	C_1	9	41.1	42.5
Thiophene	331/506	C_{2v}	9	11.9	14.1
Benzene	414/642	D_{2h}	8	13.6	20.2

[1]The MP2-F12 part is computed using density fitting without symmetry.
Column OBS/RI shows the sizes of the orbital basis set (OBS) and the auxiliary basis set used for the resolution of the identity (RI).

shows typical elapsed times (the pure user-CPU times are 1%–3% shorter). As the MP2-F12 part currently does not use symmetry, while Abelian symmetry can be used in the CCSD(T) calculation, the fraction of time spent for the MP2-F12 is larger in cases with high symmetry (cf., benzene) than without symmetry (e.g., ethanol) (cf., Table 3).

Figure 1 shows one-dimensional cuts through the MRCI-F12 potential energy surface (PES) of ozone as a function of one bond distance R_2, with R_1 fixed at 2.4 a_0 and the bond angle at 116.8°.[49] Figure 1(a) shows conventional MRCI calculations for several basis sets, and Figure 1(b) shows the corresponding F12 calculations. While the conventional curves are significantly affected by the basis set (in particular the dissociation energies), the curves obtained with MRCI-F12 and the VDZ-F12 and VQZ-F12 basis sets can hardly be distinguished. Part of the improvement is due to a CABS singles correction of

the CASSCF reference energies.[49] For example, this reduces the difference of the MRCI-F12/VDZ-F12 and MRCI-F12/VQZ-F12 ground-state dissociation energies from 2.58 kcal mol^{-1} to 1.02 kcal mol^{-1}. If state-averaged orbitals are used as in Figure 1, the effect is even larger. The MRCI-F12 methods will be made available for public use soon.

Local Correlation Methods for Ground States

The steep scaling of the computational cost with molecular size is mainly due to the delocalized character of the canonical HF orbitals that are used in most conventional calculations. The short-range character of electron correlation can be exploited using local occupied and virtual orbitals. In Molpro, such methods have been developed over the last 15 years.[23,56–72] It is possible to achieve linear scaling
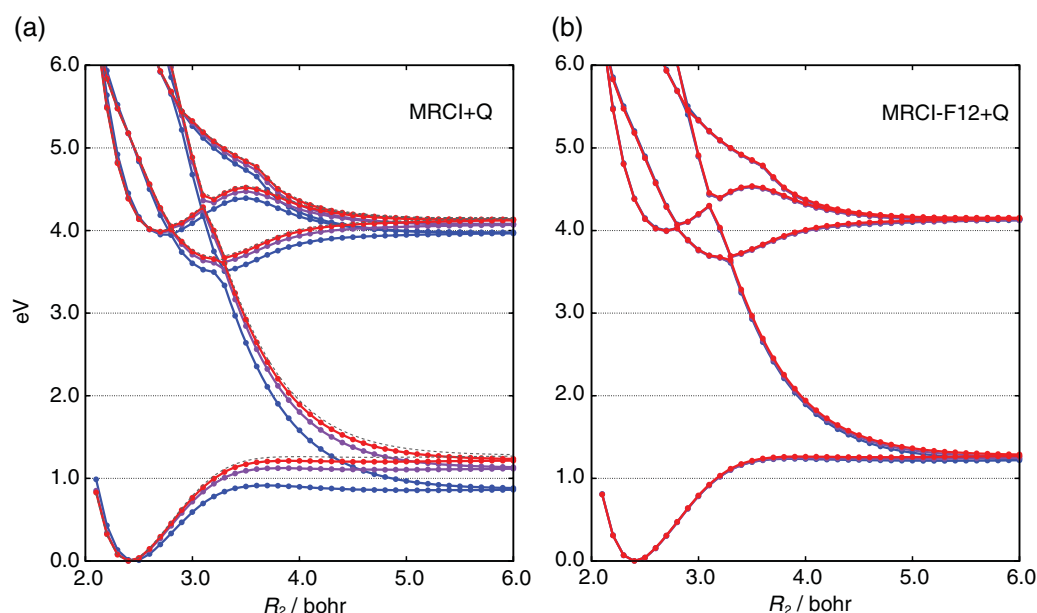

FIGURE 1 | PESs of low-lying singlet A' states of ozone computed by MRCI+Q (a) and MRCI-F12b+Q (b). One dimensional cuts for $R_1 = 2.4$ a_0 and $\alpha = 116.8°$ are shown. The aug-cc-pVDZ (blue), aug-cc-pVTZ (purple), and aug-cc-pVQZ (red) basis sets are used. Dotted lines in (a) are those computed by MRCI-F12b+Q with aug-cc-pVQZ (from Ref 49).

TABLE 4 | Reaction energies (in kJ mol^{-1}) for the two reactions shown in Figure 2, using triple-zeta basis sets

| Method	Reaction I[1]	Reaction II[2]	Reaction III[3]
DF-MP2/CBS	−113.1	−20.3	19.6
DF-MP2	−116.7	−26.2	29.6
DF-LMP2	−120.0	−21.3	16.8
DF-LCCSD(T0)	−126.9	−26.6	20.8
DF-MP2-F12	−114.6	−19.4	20.3
DF-LMP2-F12	−114.6	−18.2	19.2
DF-LCCSD(T0)-F12	−121.1	−23.7	23.5

[1] Reaction I: VTZ-F12.
[2] Reaction II: aug-cc-pVTZ.
[3] Reaction III: aug-cc-pVTZ, H=cc-pVTZ.
The MP2/CBS values have been obtained by extrapolating QZ/5Z results.

I: A possible step in the synthesis of biotin

II: Reaction of testosterone

III: Reaction step to synthesize androstendione

FIGURE 2 | Three reactions studied using LCCSD(T)-F12.

even for LCCSD(T)[61,70,71] (the local variant of a given method is denoted by prepending an L). The most efficient variants of the local methods use density fitting for evaluating and transforming all two-electron integrals. Currently, local single-reference methods are available for closed-shell [DF-LMP2 and DF-LCCSD(T)] as well as high-spin open-shell [DF-LRMP2, DF-LUCCSD(T)] (the open-shell methods will be made available soon). Analytical energy gradients are available for DF-LMP2.[68] DF-LMP2 and DF-LRMP2 calculations are easily possible with more than 4000 basis functions; LCCSD(T) calculations have been done with more than 2600 basis functions (using augmented triple-zeta basis sets). A noniterative approximation[70,71] [denoted as (T0)] is normally used for the perturbative triples correction (iterative variants are also available).[66,71] The CPU time for DF-LCCSD(T0) calculations is often comparable to the time needed for the preceding DF-HF calculation. These developments have made it possible to study enzyme reactions using LCCSD(T0) in QM/MM calculations.[73–75]

For a long time, a major problem of local correlation methods was the error caused by the domain approximation. The domain errors can be reduced by extending the domains,[69] but the CPU time increases with the fourth power of the average domain size. Recently, it has been shown that by adding suitably defined explicitly correlated terms to LMP2 or LCCSD wave functions the domain errors can be eliminated almost completely.[52–55] Benchmarks for over 100 reactions have shown that the accuracy of CCSD(T)-F12 and LCCSD(T)-F12 is virtually the same.

The local approximations strongly reduce the scaling of computational effort with molecular size, and rather large molecules can be treated (approximately 200 atoms using LMP2-F12 and 100 atoms
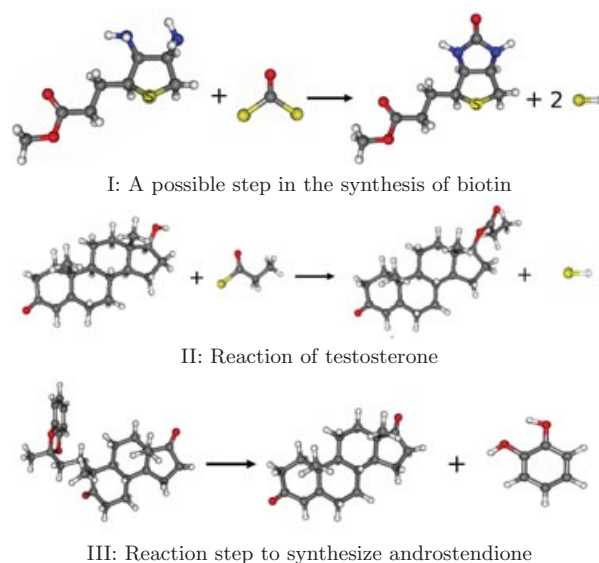
with LCCSD(T)-F12). As examples, we present in Table 4 results for the three reactions shown in Figure 2. Comparison of DF-MP2/CBS with DF-MP2 using triple-zeta basis sets (first two rows in the table) shows that there are significant basis set effects, for example, 10 kJ mol^{-1} for reaction III. Furthermore, comparing DF-MP2 with DF-LMP2 (rows 2 and 3) shows that there is a large domain error (12.8 kJ mol^{-1} for reaction III), which partly cancels with the basis set error [this may be due to the reduction of basis set superposition effects (BSSE) in the local methods]. In contrast, the MP2-F12 and LMP2-F12 results agree with each other and also with the extrapolated MP2/CBS limit within approximately 1 kJ mol^{-1}. Of course, there is no way to compare the DF-LCCSD(T0)-F12 calculations with canonical results, but basis-set and domain errors in LMP2-F12 and LCCSD(T0)-F12 calculations are known to be very similar, so it can be expected that the LCCSD(T0)-F12 results are accurate to approximately 2 kJ mol^{-1}.

Local Correlation Methods for Excited States

Singlet and triplet excited states of extended molecules can be calculated within the framework of time-dependent local coupled cluster (TD-LCC) response theory at the level of the second-order CC2 model.[76–79] Density fitting is used to decompose the electron repulsion integrals. Due to the importance of the singles amplitudes for TD-CC response theory, local approximations are applied to the doubles amplitudes only.[79] Multistate calculations with
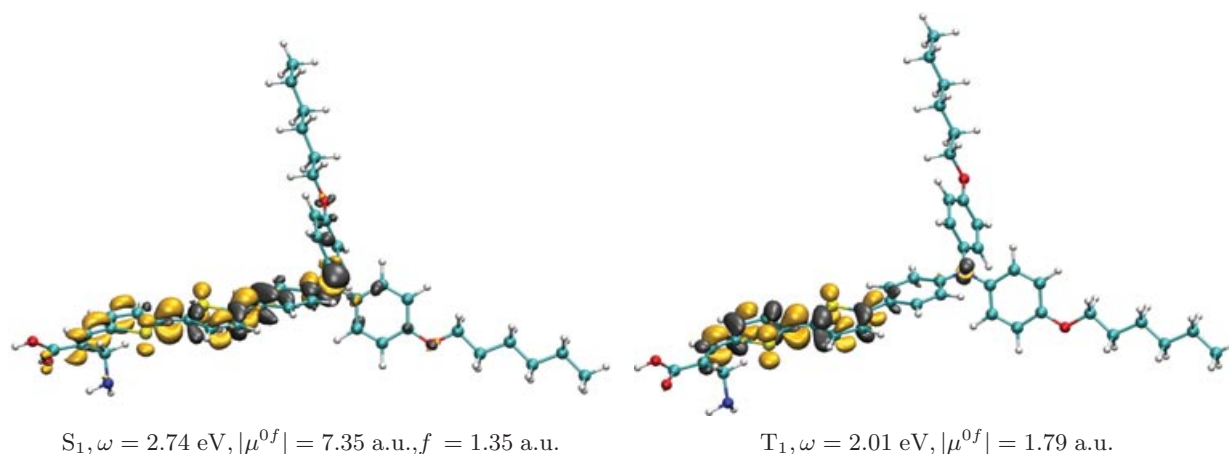
$S_1, \omega = 2.74$ eV, $|\mu^{0f}| = 7.35$ a.u., $f = 1.35$ a.u. $T_1, \omega = 2.01$ eV, $|\mu^{0f}| = 1.79$ a.u.

FIGURE 3 | Density difference of the lowest singlet S_1 and triplet T_1 states of the D21L6 molecule relative to the ground state. The yellow (bright) and dark gray iso-surfaces represent a value of $+0.002$ and -0.002, respectively. The excitation energies, the norms of the dipole difference vectors, and the oscillator strength (length gauge, singlet state only) are also given.

state-specific adaptive local approximations for the individual states are possible by virtue of the Laplace transform trick.[78,79] The deviations between the local and canonical excitation energies are well below 0.1 eV. Transition strength tensors for singlet excited states are calculated as the residues of the LCC2 linear response function. Excited-state first-order properties (dipole moments, etc.) are available for singlet and triplet excited states.

As an alternative to the TD-LCC2 response method also a local correlation variant of the second-order algebraic diagrammatic construction scheme ADC(2)[80,81] is available. Even though this method has an entirely different background, the working equations are closely related to those of TD-CC2 response.[82] However, the matrix to be diagonalized for the excitation energies is Hermitian, in contrast to the CC2 Jacobian. This has advantages near conical intersections (where the CC2 excitation energies become complex) and for calculating properties.

For example, LCC2 calculations have been carried out on the organic solar-cell sensitizer shown in Figure 3.[83] The molecule has 98 atoms, 262 correlated electrons, and 948 basis functions in the cc-pVDZ basis. Figure 3 displays the calculated density difference plots for the lowest singlet and triplet excited states. The corresponding excitation energies, dipole moment changes, and oscillator strengths are also given. In agreement with experiment,[83] a singlet charge transfer state with large oscillator strength is calculated as the lowest excited singlet state. The computed excitation energy of 2.74 eV is in very good agreement with the experimental value of 2.71 eV.[78] A pilot im-

plementation of local EOM-CCSD method[23] is also available.

Gradients and Geometry Optimization

Analytical energy gradients are currently available for Hartree–Fock and DFT (RHF, ROHF, UHF, KS, RKS, UKS, in all cases with or without density fitting), single-reference correlation methods, MP2, DF-MP2, LMP2,[84] DF-LMP2,[68] QCISD(T), LQCISD(T),[72] CCSD, as well as multiconfiguration methods MCSCF/CASSCF (including state-averaged MCSCF),[85] CASPT2,[86] MS-CASPT2. For any other methods gradients can be computed numerically using finite differences. In order to minimize the number of displacements in numerical gradient calculations, symmetrical displacement coordinates are used. Automatic geometry optimization for minima,[88] transition states, and reaction pathways[87] is possible for any method. The optimization can be carried out using cartesian coordinates, Z-matrix coordinates, or redundant internal coordinates. Various optimization algorithms and Hessian update schemes are implemented (for details, see Ref 88). Optionally, exact or approximate Hessians can be computed at intermediate optimization steps. There are facilities to linearly combine gradients, so that, for example, counterpoise corrected geometry optimizations are possible.

Electrical and Vibrational Properties

Many one-electron properties can be computed using Molpro, for example, multipole moments, electric

fields, electric field gradients, diamagnetic shielding tensors, angular momenta, velocity operators, Darwin and Cowan-Griffin relativistic corrections, and spin-orbit couplings. Second-order properties such as dipole polarizabilities can be computed analytically for HF and MP2. For other methods the finite field approach can be used. For magnetic properties, see section *Magnetic Properties*.

Analytical Hessians are available for closed-shell HF and MCSCF/CASSCF (in the latter case only without symmetry and without state averaging). In all other cases Hessians can be constructed automatically using finite difference methods (using analytical gradients or energies). Harmonic vibrational frequencies and thermodynamic data are computed in the standard way.

Anharmonic vibrational spectra can be computed using vibrational SCF (VSCF), VMP2, VCI, or VMCSCF procedures developed by Rauhut and coworkers.[89–92] The necessary many-dimensional potential energy surfaces around the equilibrium structures are computed automatically.[89–92] So-called many-mode approximations, in which coupling potentials are computed at lower lever than the diagonal potentials (along the normal coordinates) can be used to save computation time without much loss of accuracy. The PES calculation is perfectly parallelized using the Molpro MPPX scheme (cf., section *Parallel Execution*), where independent processes compute the energies for different geometries.

Magnetic Properties

Nuclear magnetic shielding tensors can presently be calculated at the level of DF-HF.[93] The gauge-origin problem is circumvented by the use of gauge-including atomic orbitals (GIAOs).[94,95] GIAOs cannot be used as fitting functions, since the explicit gauge origin dependence then cannot cancel (there is no corresponding `bra` to the `ket` function). Density fitting based on a set of ordinary Gaussians, however, works remarkably well; the dependence of the chemical shifts on the size of the fitting basis set is virtually negligible and the agreement with results from conventional calculations without density fitting is typically within 0.1 ppm or less. The program is quite efficient and molecules with over 100 atoms can be treated.[93] A program for correlated nuclear magnetic shielding tensors at the level of density fitted local MP2 is presently under development.

Special Relativity

Molpro is a Schrödinger-only code that deals only with ansätze based on one-component nonrelativis-

tic wave functions. However, it is possible to include some relativistic effects approximately. Associated with the library of basis sets, there is a corresponding library of effective core potentials (ECPs), some of which embed scalar relativistic effects. As well as effective one-electron potentials, there is also the possibility to introduce outer-product two-electron potentials, with the specific example of core polarization potentials for modeling core-valence correlation effects.[96]

It is also possible to calculate the spin-orbit part of the matrix element of the Breit–Pauli operator between CASSCF or MRCI wave functions, and these calculations are placed in an automatic framework that calculates all nonzero matrix elements arising in a given set of states, and then performs degenerate perturbation theory to obtain the spin-orbit shifts of the Schrödinger states.

As an alternative to the use of ECPs, approximate scalar relativistic effects can be included in all-electron calculations through perturbation theory or using the Douglas–Kroll–Hess (DKH) Hamiltonian.[97]

Density Functional Theory Symmetry-Adapted Perturbation Theory

A density functional theory symmetry-adapted perturbation theory (DFT-SAPT) program is also available in Molpro.[98] DFT-SAPT directly calculates the individual components of the intermolecular interaction energy between two closed-shell systems, that is, first-order electrostatic and exchange, second-order induction and dispersion. It provides very accurate interaction energies, in quality close to CCSD(T). The intramonomer correlation is treated at the level of DFT. The program employs density fitting and is very efficient. Some very recent applications comprise the H_2–C_{60} interaction[99] and destabilization of DNA Watson–Crick pairs in photodamaged DNA.[100]

TECHNICAL FEATURES

Input Script Language

Molpro's execution is controlled by an input file that contains a script written in a custom-designed language. Each statement of the script consists of a command, followed by one or more qualifying options. Some of the commands run program modules that implement a particular quantum chemistry method (e.g., RHF, CCSD(T), MRCI, etc.) and in most cases the name of the command is identical to the standard acronym

for the method. Syntax checking of the whole script is performed at the start of execution.

Other commands are built in to the language, and offer features such as branching, loops and data manipulation. The language also supports symbolic variables that are one-dimensional arrays of integer, floating-point, logical or string type, and scalars are implemented as a special case of these vectors. Variables can be assigned to arbitrary arithmetical expressions containing constants and other variables, that are evaluated element by element down the list. A typical use of variables is to store the results of a number of calculations, with possible postprocessing, followed by presentation (tabular and graphical) and export of these results.

Built-in Utilities
Molpro contains a number of utilities to support analysis and manipulation of results, molecular orbitals, densities, etc. These include EXTRAPOLATE, which automatically extrapolates the last energy calculation to the basis set limit according to one of the standard formulae. The MERGE and MATROP utilities perform matrix manipulations of objects holding molecular orbitals, densities and properties. Additionally, there are utilities for performing orbital localization, and Mulliken, natural bond orbital, and distributed multipole analyses.

Visualization and Interface to Other Programs
Molpro can produce several standard-format representations of geometry, molecular orbital, and electron density data. Built-in commands implement the writing of data for the molden viewer, which constructs pictures of orbitals and densities from the orbital basis set (OBS) and coefficients, and the writing of 'cube' files that tabulate values of orbitals on a cuboidal grid, suitable for manipulation by programs such as gOpenMol. An interface to the JMol interactive three-dimensional widget supports the construction of interactive web pages directly from Molpro.[101]

Simple two-dimensional tables can be constructed by specifying one or more variables that will form its columns. These tables can be exported to external files in a number of standard formats. Tables can also be presented graphically, through full integration with the Grace plotting system, producing PDF files directly from the Molpro run.

A recent development is a graphical application for preparing, running, and viewing Molpro jobs.

This graphical user interface (GUI) is built on the Qt system, making extensive use of molecular model display and other features in the Avogadro library, and runs on Macintosh, Windows, and Linux systems. The GUI integrates a simple editor for input, a viewer for the output stream, and a visual object display for molecular models and other graphical objects.

Symmetry
All methods make use of point-group symmetry (up to the highest abelian subgroup), except in the case of local methods, and where density fitting is used.

Parallel Execution
Molpro is parallelized using a multiple process model. At the start of a job, a specified number of identical instances of the executable is started, and in many parts of the program, work is distributed between the processors, with appropriate communication to share data and consolidate results. The support for these communication functions is provided by a software layer parallel programming interface for distributed data (PPIDD)[102,103] that itself can be built to sit on top of the Global Arrays (GA) toolkit, or any standard MPI library. PPIDD supports global data structures that can be accessed by any process without message exchange or other reference to other processes.

Normally (designated MPP), parallelization is fine-grained, for example, splitting the work for constructing a single Fock matrix. All functionality of the code works correctly in parallel, and reasonable performance on up to tens of processors is achieved, for example, in DFT, CCSD, CCSD(T), and MRCI calculations. Alternatively (designated MPPX), the program can be built so that certain large tasks, for example, entire energy calculations at displaced geometries for finite-difference forces, are each run in serial, but concurrently.

Molpro makes extensive use of storage of temporary quantities outside the memory space of a process, so that calculations need not be limited by constraints in that space. The abstraction layer that serves this data staging can be configured to choose from a number of different mechanisms. In some computing environments, it uses files that are completely shared between all processes, either through the use of a disk-based shared file system such as Lustre or GPFS, or, more efficiently, in the combined global memory of all nodes, and under these circumstances the appropriate functionality provided by GA or MPI is used. On systems where scratch disks are local to individual nodes, Molpro can take advantage of the large

aggregate I/O bandwidth available by managing its own data placement on node-local files, with the option to either replicate, and maintain synchronized, items of data on all nodes, or to distribute data across nodes. These features are important in achieving optimum performance on a variety of computer architectures.

The Integrated Tensor Framework

Novel hardware platforms are increasingly difficult to program efficiently, and electronic structure methods are becoming more and more complex. To address this situation, an abstraction layer between the quantum chemistry algorithms and the hardware platforms is being developed, the integrated tensor framework (ITF).

ITF is a framework for handling networks of tensor contractions in an efficient manner. Such networks can execute the core operations of many electronic structure methods; for example, we implemented multireference methods,[39] coupled cluster methods, CCSD analytic gradients, and F12 methods on top of it, each time reaching performance rivalling the hand-optimized implementation. When using ITF, the equations are not directly programmed in a native programming language, but rather in terms of an expressive high-level language, which is interpreted at runtime and which only encodes the very essential aspects of an evaluation algorithm. Implementation details are handled by the framework. This approach simplifies the programming of equation systems, and allows for automatic adoption of different computer platforms.

REFERENCES

1. Pulay P. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. I. Theory. *Mol Phys* 1969, 17:197–204.

2. Pulay P. Ab initio calculation of force constants and equilibrium geometries in polyatomic molecules. II. Force constants of water. *Mol Phys* 1970, 18:473–480.

3. Pulay P. Ab initio calculation of force constants and equilibrium geometries. III. Second-row hydrides. *Mol Phys* 1971, 21:219–339.

4. Meyer W. PNO-CI and CEPA studies of electron correlation effects. I. Configuration expansion by means of nonorthogonal orbitals, and application to the ground state and ionized states of methane. *J Chem Phys* 1973, 58:1017–1035.

5. Meyer W. PNO-CI and CEPA studies of electron correlation effects. II. Potential curves and dipole-moment functions of the OH radical. *Theor Chim Acta* 1974, 35:277–292.

6. Werner HJ, Meyer W. A quadratically convergent MCSCF method for the simultaneous optimization of several states. *J Chem Phys* 1981, 74:5794–5801.

7. Werner HJ, Reinsch EA. The self-consistent electron pairs method for multiconfiguration reference state functions. *J Chem Phys* 1982, 76:3144–3156.

8. Knowles PJ, Handy NC. A new determinant-based full configuration interaction method. *Chem Phys Lett* 1984, 111:315–321.

9. Werner HJ, Knowles PJ. A second order MCSCF method with optimum convergence. *J Chem Phys* 1985, 82:5053–5063.

10. Knowles PJ, Werner HJ. An efficient second order MCSCF method for long configuration expansions. *Chem Phys Lett* 1985, 115:259–267.

11. Werner HJ, Knowles PJ. An efficient internally contracted multiconfiguration reference CI method. *J Chem Phys* 1988, 89:5803–5814.

12. Knowles PJ, Werner HJ. An efficient method for the evaluation of coupling coefficients in configuration interaction calculations. *Chem Phys Lett* 1988, 145:514–522.

13. Knowles PJ, Werner HJ. Internally contracted multiconfiguration reference configuration interaction calculations for excited states. *Theor Chim Acta* 1992, 84:95–103.

14. Werner HJ, Knizia G, Manby FR. Explicitly correlated coupled cluster methods with pair-specific geminals. *Mol Phys* 2011, 109:407–417.

15. Werner HJ, Manby FR, Knowles PJ. Fast linear scaling second-order Møller–Plesset perturbation theory

(MP2) using local and density fitting approximations. *J Chem Phys* 2003, 118:8149–8160.

16. Polly R, Werner HJ, Manby FR, Knowles PJ. Fast Hartree–Fock theory using local density fitting approximations. *Mol Phys* 2004, 102:2311–2321.

17. Monagan MB, Geddes KO, Heal KM, Labahn G, Vorkoetter SM, McCarron J, DeMarco P. *Maple 10 Programming Guide*. Waterloo, ON: Maplesoft; 2005.

18. Strange R, Manby FR, Knowles PJ. Automatic code generation in density functional theory. *Comput Phys Commun* 2001, 136:310–318.

19. Manby FR, Knowles PJ. The Poisson equation in the Kohn–Sham Coulomb problem. *Phys Rev Lett* 2001, 87:163001.

20. Brown P, Woods C, McIntosh-Smith S, Manby FR. Massively multicore parallelization of Kohn–Sham theory. *J Chem Theory Comput* 2008, 4:1620–1626.

21. Woods CJ, Brown P, Manby FR. Multicore parallelization of Kohn–Sham theory. *J Chem Theory Comput* 2009, 5:1776–1784.

22. Hampel C, Peterson KA, Werner HJ. A comparison of the efficiency and accuracy of the quadratic configuration interaction (QCISD), coupled cluster (CCSD), and Brueckner coupled cluster (BCCD) methods. *Chem Phys Lett* 1992, 190:1–12.

23. Korona T, Werner HJ. Local treatment of electron excitations in the EOM-CCSD method. *J Chem Phys* 2003, 118:3006–3019.

24. Knowles PJ, Andrews JS, Amos RD, Handy NC, Pople JA. Restricted Møller–Plesset theory for open shell molecules. *Chem Phys Lett* 1991, 186:130–136.

25. Knowles PJ, Hampel C, Werner HJ. Coupled cluster theory for high spin open shell reference wavefunctions. *J Chem Phys* 1993, 99:5219–5227.

26. Deegan MJO, Knowles PJ. Perturbative corrections to account for triple excitations in closed and open shell coupled cluster theories. *Chem Phys Lett* 1994, 227:321–326.

27. Kállay M, Surján PR. Higher excitations in coupled-cluster theory. *J Chem Phys* 2001, 115:2945–2954.

28. Kállay M, Gauss J. Approximate treatment of higher excitations in coupled-cluster theory. *J Chem Phys* 2005, 123:214105.

29. Thorsteinsson T, Cooper DL, Gerratt J, Karadakov PB, Raimondi M. Modern valence bond representations of CASSCF wavefunctions. *Theor Chim Acta* 1996, 93:343–366.

30. Cooper DL, Thorsteinsson T, Gerratt J. Fully variational optimization of modern VB wave functions using the CASVB strategy. *Int J Quantum Chem* 1997, 65:439–451.

31. Cooper DL, Thorsteinsson T, Gerratt J. Modern VB representations of CASSCF wave functions and the fully-variational optimization of modern VB wave functions using the CASVB strategy. *Adv Quantum Chem* 1998, 32:51–67.

32. Werner HJ. Third-order multiconfiguration reference perturbation theory: the CASPT3 method. *Mol Phys* 1996, 89:645–661.

33. Ghigo G, Roos BO, Malmqvist PA. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chem Phys Lett* 2004, 396:142–149.

34. Celani P, Werner HJ. Multireference perturbation theory for large restricted and selected active space reference wave functions. *J Chem Phys* 2000, 112:5546–5557.

35. Angelia C, Cimiraglia R, Evangelisti S, Leininger T, Malrieu JP. Introduction of n-electron valence states for multireference perturbation theory. *J Chem Phys* 2001, 114:10252–10264.

36. Angeli C, Cimiraglia R, Malrieu JP. n-electron valence state perturbation theory: a spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J Chem Phys* 2002, 117:9138–9153.

37. Angeli C, Pastore M, Cimiraglia R. New perspectives in multireference perturbation theory: the n-electron valence state approach. *Theor Chem Acc* 2007, 117:743–754.

38. Werner HJ, Knowles PJ. A comparison of variational and non-variational internally contracted multiconfiguration-reference configuration interaction calculations. *Theor Chim Acta* 1990, 78:175–187.

39. Shamasundar KR, Knizia G, Werner HJ. A new internally contracted multi-reference configuration interaction (MRCI) method. *J Chem Phys*. In press doi: 10.1063/1.3609809.

40. Werner HJ, Adler TB, Knizia G, Manby FR. Efficient explicitly correlated coupled-cluster approximations. In: Čársky P, Paldus J, Pittner J, eds. *Recent Progress in Coupled Cluster Methods*. Dordrecht/Heidelberg/London/New York: Springer; 2010, 573–619.

41. Tew DP, Hättig C, Bachorz RA, Klopper W. Explicitly correlated coupled-cluster theory. In: Čársky P, Paldus J, Pittner J, eds. *Recent Progress in Coupled Cluster Methods*. Dordrecht/Heidelberg/London/New York: Springer; 2010, 535–572.

42. Werner HJ, Adler TB, Manby FR. General orbital invariant MP2-F12 theory. *J Chem Phys* 2007, 126:164102.

43. Knizia G, Werner HJ. Explicitly correlated RMP2 for high-spin open-shell reference states. *J Chem Phys* 2008, 128:154103.

44. Adler TB, Knizia G, Werner HJ. A simple and efficient CCSD(T)-F12 approximation. *J Chem Phys* 2007, 127:221106.

45. Knizia G, Adler TB, Werner HJ. Simplified CCSD(T)-F12 methods: theory and benchmarks. *J Chem Phys* 2009, 130:054104.

46. Werner HJ, Knizia G, Adler TB, Marchetti O. Benchmark studies for explicitly correlated perturbation- and coupled cluster theories. *Z Phys Chem* 2010, 224:493–511.

47. Shiozaki T, Werner HJ. Communication: second-order multireference perturbation theory with explicit correlation: CASPT2-F12. *J Chem Phys* 2010, 133:141103.

48. Shiozaki T, Knizia G, Werner HJ. Explicitly correlated multireference configuration interaction: MRCI-F12. *J Chem Phys* 2011, 134:034113.

49. Shiozaki T, Werner HJ. Explicitly correlated multireference configuration interaction with multiple reference functions: avoided crossings and conical intersections. *J Chem Phys* 2011, 134:184104.

50. Werner HJ, Manby FR. Explicitly correlated second-order perturbation theory using density fitting and local approximations. *J Chem Phys* 2006, 124:054114.

51. Manby FR, Werner HJ, Adler TB, May AJ. Explicitly correlated local second-order perturbation theory with a frozen geminal correlation factor. *J Chem Phys* 2006, 124:094103.

52. Werner HJ. Eliminating the domain error in local explicitly correlated second-order Møller–Plesset perturbation theory. *J Chem Phys* 2008, 129:101103.

53. Adler TB, Werner HJ, Manby FR. Local explicitly correlated second-order perturbation theory for the accurate treatment of large molecules. *J Chem Phys* 2009, 130:054106.

54. Adler TB, Werner HJ. Local explicitly correlated coupled-cluster methods: efficient removal of the basis set incompleteness and domain errors. *J Chem Phys* 2009, 130:241101.

55. Adler TB, Werner HJ. Local explicitly correlated coupled cluster methods for large molecules: LCCSD(T)-F12. Submitted for publication.

56. Hampel C, Werner HJ. Local treatment of electron correlation in coupled cluster theory. *J Chem Phys* 1996, 104:6286–6297.

57. Hetzer G, Pulay P, Werner HJ. Multipole approximation of distant pair energies in local MP2 calculations. *Chem Phys Lett* 1998, 290:143–149.

58. Schütz M, Hetzer G, Werner HJ. Low-order scaling local electron correlation methods. I. Linear scaling local MP2. *J Chem Phys* 1999, 111:5691–5705.

59. Hetzer G, Schütz M, Stoll H, Werner HJ. Low-order scaling local correlation methods II: Splitting the Coulomb operator in linear scaling local second-order Møller–Plesset perturbation theory. *J Chem Phys* 2000, 113:9443–9455.

60. Gauss J, Werner HJ. NMR chemical shift calculations within local correlation methods: the GIAO-LMP2 approach. *Phys Chem Chem Phys* 2000, 2:2083–2090.

61. Schütz M, Werner HJ. Low-order scaling local electron correlation methods. IV. Linear scaling local coupled-cluster (LCCSD). *J Chem Phys* 2001, 114:661–681.

62. Schütz M, Manby FR. Linear scaling local coupled cluster theory with density fitting. I: 4-external integrals. *Phys Chem Chem Phys* 2003, 5:3349–3358.

63. Mata R, Werner HJ. Calculation of smooth potential energy surfaces using local electron correlation methods. *J Chem Phys* 2006, 125:184110.

64. Mata R, Werner HJ. Local correlation methods with a natural localized molecular orbital basis. *Mol Phys* 2007, 105:2753–2761.

65. Mata R, Werner HJ, Schütz M. Correlation regions within a localized molecular orbital approach. *J Chem Phys* 2008, 128:144106.

66. Schütz M. Linear scaling local connected triples beyond local (T): local CCSDT-1b with O(N) scaling. *J Chem Phys* 2002, 116:8772–8785.

67. Schütz M. A new, fast, semi-direct implementation of linear scaling local coupled cluster theory. *Phys Chem Chem Phys* 2002, 4:3941–3947.

68. Schütz M, Werner HJ, Lindh R, Manby FR. Analytical energy gradients for local second-order Møller–Plesset perturbation theory using density fitting approximations. *J Chem Phys* 2004, 121:737–750.

69. Werner HJ, Pflüger K. On the selection of domains and orbital pairs in local correlation treatments. *Annu Rep Comput Chem* 2006, 2:53–80.

70. Schütz M, Werner HJ. Local perturbative triples correction (T) with linear cost scaling. *Chem Phys Lett* 2000, 318:370–378.

71. Schütz M. Low-order scaling local electron correlation methods. III. Linear scaling local perturbative triples correction (T). *J Chem Phys* 2000, 113:9986–10001.

72. Rauhut G, Werner HJ. Analytical energy gradients for local coupled-cluster methods. *Phys Chem Chem Phys* 2001, 3:4853–4862.

73. Claeyssens F, Harvey JN, Manby FR, Mata RA, Mulholland AJ, Ranaghan KE, Schütz M, Thiel S, Thiel W, Werner HJ. High-accuracy computation of reaction barriers in enzymes. *Angew Chem* 2006, 118:7010–7013.

74. Mata RA, Werner HJ, Thiel S, Thiel W. Towards accurate barriers for enzymatic reactions: QM/MM case study on *p*-hydroxybenzoate hydroxylase. *J Chem Phys* 2008, 128:025104.

75. Dieterich JM, Werner HJ, Mata RA, Metz S, Thiel W. Reductive half-reaction of aldehyde oxidoreductase toward acetaldehyde: ab initio and free energy

QM/MM calculations. *J Chem Phys* 2010, 132:035101.

76. Kats D, Korona T, Schütz M. Local CC2 electronic excitation energies for large molecules with density fitting. *J Chem Phys* 2006, 125:104106.

77. Kats D, Korona T, Schütz M. Transition strengths and first-order properties of excited states from local coupled cluster CC2 response theory with density fitting. *J Chem Phys* 2007, 127:064107.

78. Freundorfer K, Kats D, Korona T, Schütz M. Local CC2 response method for triplet states based on Laplace transform: excitation energies and first-order properties. *J Chem Phys* 2010, 133:244110.

79. Kats D, Schütz M. A multistate local coupled cluster CC2 response method based on the Laplace transform. *J Chem Phys* 2009, 131:124117.

80. Schirmer J. Beyond the random-phase approximation: a new approximation scheme for the polarization propagator. *Phys Rev A* 1982, 26:2395–2416.

81. Trofimov A, Schirmer J. An efficient polarization propagator approach to valence electron excitation spectra. *J Phys B: At Mol Opt Phys* 1995, 28:2299–2324.

82. Hättig C. Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2). *Adv Quantum Chem* 2005, 50:37–60.

83. Yum J, Hagberg D, Moon S, Karlsson K, Marinado T, Sun L, Hagfeldt A, Nazeeruddin M, Grätzel M. A light-resistant organic sensitizer for solar-cell applications. *Angew Chem Int Edit* 2009, 48:1576–1580.

84. ElAzhary A, Rauhut G, Pulay P, Werner HJ. Analytical energy gradients for local second-order Møller–Plesset perturbation theory. *J Chem Phys* 1998, 108:5185–5193.

85. Busch T, Esposti AD, Werner HJ. Analytical energy gradients for MCSCF wave functions with frozen core orbitals. *J Chem Phys* 1991, 94:6708–6715.

86. Celani P, Werner HJ. Analytical energy gradients for internally contracted second-order multi-reference perturbation theory (CASPT2). *J Chem Phys* 2003, 119:5044–5057.

87. Eckert F, Werner HJ. Reaction path following by quadratic steepest decent. *Theor Chem Acc* 1998, 100:21–30.

88. Eckert F, Pulay P, Werner HJ. Ab initio geometry optimization for large molecules. *J Comput Chem* 1997, 18:1473–1483.

89. Rauhut G. Efficient calculation of potential energy surfaces for the generation of vibrational wave functions. *J Chem Phys* 2004, 121:9313–9322.

90. Hrenar T, Werner HJ, Rauhut G. Accurate calculation of anharmonic vibrational frequencies of medium sized molecules using local coupled cluster methods. *J Chem Phys* 2007, 126:134108.

91. Rauhut G, Hrenar T. A combined variational and perturbational study on the vibrational spectrum of P_2F_4. *Chem Phys* 2008, 346:160–166.

92. Neff M, Rauhut G. Toward large scale vibrational configuration interaction calculations. *J Chem Phys* 2009, 131:124129.

93. Loibl S, Manby FR, Schütz M. Density fitted, local Hartree–Fock treatment of NMR chemical shifts using London atomic orbitals. *Mol Phys* 2010, 108:477–485.

94. London F. Theorie quantique des courants interatomiques dans les combinaisons aromatiques. *J Phys Radium* 1937, 8:397–409.

95. Ditchfield R. Self-consistent perturbation theory of diamagnetism I. A gauge-invariant LCAO method for N.M.R. chemical shifts. *Mol Phys* 1974, 27:789–807.

96. Müller W, Flesch J, Meyer W. Treatment of intershell correlation effects in ab initio calculations by use of core polarization potentials. Method and application to alkali and alkaline earth atoms. *J Chem Phys* 1984, 80:3297–3310.

97. Wolf A, Reiher M, Hess BA. The generalized Douglas–Kroll transformation. *J Chem Phys* 2002, 117:9215–9226.

98. Hesselmann A, Jansen G, Schütz M. Density-functional theory symmetry-adapted intermolecular perturbation theory with density fitting: a new efficient method to study intermolecular interaction energies. *J Chem Phys* 2005, 122:014103.

99. Korona T, Hesselmann A, Dodziuk H. Symmetry-adapted perturbation theory applied to endohedral fullerene complexes: a stability study of $H_2@C_{60}$ and $2H_2@C_{60}$. *J Chem Theory Comput* 2009, 5:1585–1596.

100. Sadeghian K, Bocola M, Schütz M. Intermolecular interactions in photodamaged DNA from density functional theory-symmetry-adapted perturbation theory. *ChemPhysChem* 2011, 12:1251–1254.

101. MolproView. Available at: http://www.molpro.net/molproView. (Accessed June 24, 2011).

102. Wang M, May AJ, Knowles PJ. Parallel programming interface for distributed data. *Comput Phys Commun* 2009, 180:2673–2679.

103. Wang M, May AJ, Knowles PJ. Improved version of the parallel programming interface for distributed data. *Comput Phys Commun* 2011, 182:1502–1506.