GAMESSPLUS:
A Module Incorporating Electrostatic Potential Hessians for Site-Site Representation of Electrostatic Embedding, Generalized Hybrid Orbital QM/MM Methods, the SM5.42, SM5.43, SM6, and SM6T Solvation Models, the CM2, CM3, and CM4 Charge Models, and the Löwdin and Redistributed Löwdin Population Analysis Methods into *GAMESS*

Users Manual

Version 2008
Date of finalization of this version of the software: Apr 24, 2008
Date of most recent change in this document: June 30, 2008

Masahiro Higashi, [a] Adam Chamberlin, [a] Jingzhi Pu,[a] Casey P. Kelly,[a]
Jason D. Thompson,[a] James D. Xidos,[a] Jiabo Li,[a] Tianhai Zhu,[a]
Gregory D. Hawkins,[a] Yao-Yuan Chuang,[a] Patton L. Fast,[a] Benjamin J. Lynch,[a]
Daniel A. Liotard,[b] Daniel Rinaldi,[c] Jiali Gao,[a]
Christopher J. Cramer,[a] and Donald G. Truhlar[a]

[a] *Department of Chemistry and Supercomputer Institute, University of Minnesota, Minneapolis, MN 55455-0431, U. S. A.*

[b] *Laboratoire de Physico-Chimie Theorique, Universite de Bordeaux 1, 351 Cours de la Liberation, 33405 Talence Cedex, France*

[c] *Laboratoire de Chimie Theorique, Universite de Nancy I, Vandoeuvre-Nancy 54506, France*

Distribution site: *http://comp.chem.umn.edu/gamessplus*

Contents

Executive Summary *4*

Extended Abstract *10*

Löwdin population analysis and redistributed Löwdin population analysis 10

Charge Model 2, Charge Model 3, and Charge Model 4 10

SM5.42, SM5.43, SM6, and SM6T Solvation Models 11

Allowed Combinations of Solvent Model, Electronic Structure Theory, and Basis Set. 15

NDDO and CM2 Specific Reaction Parameters (SRP) Models 18

Solubility Calculations 19

Soil Sorption Calculations 19

QM/MM Calculations at the *Ab Initio* HF Level with the GHO Boundary Treatment 20

Electrostatically Embedded QM Calculation with a Site-Site Representation of the QM/MM Electrostatic Interaction 20

GAMESSPLUS Reference *22*

Literature References *23*

Usage *30*

Notes on *GAMESS* Input 31

Namelists $GMSOL and $CM2 34

Namelist $CM2SRP 42

Namelist $NDDOSRP 43

***GAMESSPLUS* keywords Required for Running Standard SMx Calculations** 45

Namelist $EEQM 46

Input and Output Examples ***48***

Notes on Running GAMESSPLUS ***59***

The GAMESSPLUS test suite ***60***

Description of the test suite for CM2, CM3, CM4, SM5.42, and SM5.43. .... ***60***

Description of the test suite for SM6. .... ***64***

Description of the test suite for SM6T. ***67***

Description of the test suite for EEQM ***70***

Verifying the Installation of GAMESSPLUS Using the Test Suite Results ***72***

Special Notes on Basis Sets ***72***

MIDI! basis set 72

cc-pVDZ basis set in *Gaussian* 73

6-31G* and 6-31+G* basis sets in CM*x* (*x* = 2 or 3) and SM*x* (*x* = 5.42, 5.43, 6, or 6T) 74

SCF Schemes ***75***

Notation for the solvation models **78**

Solvent Parameters ***78***

Program Distribution ***79***

A Note on GAMESS Versions ***81***

Standard Method for Updating and Compiling GAMESSPLUS ***81***

Makepatch Method for Updating and Compiling GAMESSPLUS ***82***

Manually Updating and Compiling GAMESSPLUS ***83***

Platforms ***85***

GAMESSPLUS Solubility Utility ***87***

Executive Summary **87**

The SM5.42 and SM5.43 continuum solvation models **89**

Usage **90**

gamess/gamessplus Input 90

Input Specific to the GAMESSPLUS Solubility Utility 92

Input Options Specific to the $VAPOR namelist 95

Test Calculations **96**

Input 96

Output 98

Installing the Solubility Utility Program **100**

Installing the Solubility Utility Program **100**

Running the Solubility Utility Program **101**

GAMESSPLUS Soil Sorption Utility ***102***

Executive Summary **102**

The SM5.42 continuum solvation model for aqueous solution and soil **104**

Solvent descriptors for bulk soil **105**

Usage **106**

Input 106

Test Calculations **107**

Input 107

Output 112

Installing the GAMESSPLUS Soil Sorption Utility Program **115**

Running the Soil Sorption Utility Program **116**

GAMESSPLUS Revision History and Version Summaries ***117***

Executive Summary

GAMESSPLUS version 2008 is a module that incorporates the following methods into *GAMESS*:

- combined quantum mechanics and molecular mechanics (QM/MM) with the generalized hybrid orbital (GHO) boundary treatment
- Löwdin population analysis
- redistributed Löwdin population analysis (RLPA)
- CM2, CM3, and CM4 charge models
- SM*x* (*x* = 5.42, 5.43, and 6) solvation models
- SM6 with temperature dependence (SM6T)
- electrostatically embedded quantum mechanical (EEQM) energy and its first and second derivatives with respect to coordinates and electrostatic potentials with a site-site representation of the QM/MM electrostatic interaction

The SM*x* solvation models are based on the generalized Born method for electrostatics augmented with semiempirical surface tensions for non-bulk electrostatics. These models can calculate free energies of solvation using gas-phase geometries, as well as carry out geometry optimization in the liquid phase using analytical gradients.

The GHO QM/MM method is available by means of a *CHARMM/GAMESSPLUS* combination package for treating the QM sub-system at the *ab initio* Hartree-Fock level. The GHO analytical gradients are also available for QM/MM geometry optimizations.

To perform GHO QM/MM calculations, the user needs to obtain *GAMESS* from Iowa State University (March 24, 2007 (R6) version of *GAMESS*), *GAMESSPLUS* from the University of Minnesota, and *CHARMM* from Harvard University. The compilation of the *CHARMM/GAMESSPLUS* combination package as an integrated executable is supported by a utility package called *CGPLUS*, which is available at http://comp.chem.umn.edu/cgplus. The usage of the *CHARMM/GAMESSPLUS* combination package for carrying out GHO-AIHF calculations is covered in the *CGPLUS* manual (see the CGPLUS-v2008 User Manual). *CGPLUS* also provides a separate test suite for testing the GHO-AIHF functionality of the *CHARMM/GAMESSPLUS* combination package.

The EEQM energy calculations with a site-site representation of the QM/MM electrostatic interaction enable one to calculate the electronic energy in the presence of an external electrostatic potential such as the electrostatic potential from a solvent or a molecular mechanics region. In these calculations, the external electrostatic potential distribution is described as the collection of the values of the external electrostatic potential at the locations of the QM nuclei. The first and second derivatives of the EEQM energy with respect to coordinates and external electrostatic potentials can be calculated.

In order to make the following description of some of the capabilities of *GAMESSPLUS* more clear, we note that the following basis sets use Cartesian *d* functions:

MIDI!6D (also known as MIDIX6D)
6-31G*
6-31+G*
6-31+G**
DZVP

and the following basis sets use spherical harmonic *d* functions:

MIDI! (also known as MIDI!5D and MIDIX5D)
cc-pVDZ

GAMESSPLUS adds the following new capabilities to *GAMESS:*

- The B3LYP hybrid density functional theory method, as it is implemented in *Gaussian* and *HONDOPLUS* (i.e., using version III of the VWN correlation functional) has been added. This method can be used to obtain restricted and unrestricted wave functions and is requested with the DFTTYP=B3LYP3 keyword in the $DFT data group; see the section entitled *Notes on GAMESS input* below. (The DFTTYP=B3LYP5 keyword uses version V of the VWN functional, which is the non-standard form of the VWN functional).

- The MPW*X*, where *X* is the percentage of Hartree-Fock exchange, hybrid density functional theory method. This method can be used to obtain restricted and unrestricted wave functions and is requested with the DFTTYP=MPW*X* keyword in the $DFT data group; see the section entitled *Notes on GAMESS input* below.

- For all restricted and unrestricted HF, DFT, and hybrid DFT methods using basis sets containing functions up to *f* in angular momentum, gas-phase and liquid-phase Löwdin partial atomic charges (Class II charges) can be calculated. For calculations using the 6-31+G* and 6-31+G** basis sets, gas-phase and liquid-phase redistributed Löwdin population analysis (RLPA) partial atomic charges can be calculated for all restricted and unrestricted HF, DFT, and hybrid DFT methods available in *GAMESS*.

- Gas-phase and liquid-phase CM2 class IV charges can be determined for the following combinations of electronic structure theory and basis set (using either a restricted or an unrestricted formalism):

  AM1
  HF/MIDI!
  HF/MIDI!6D
  HF/6-31G*
  BPW91/MIDI!
  BPW91/MIDI!6D

  PM3
  B3LYP/MIDI!
  BPW91/6-31G*
  HF/6-31+G*
  HF/cc-pVDZ
  BPW91/DZVP

- Gas-phase and liquid-phase CM3 class IV charges can be determined for the following combinations of electronic structure theory and basis set (using either a restricted or an unrestricted formalism):

  AM1
  HF/MIDI!6D
  MPW*X*/MIDI!
  MPW*X*/6-31G*
  MPW*X*/6-31+G**
  B3LYP/MIDI!6D
  B3LYP/6-31+G*

  PM3
  HF/6-31G*
  MPW*X*/MIDI!6D
  MPW*X*/6-31+G*
  BLYP/6-31G*
  B3LYP/6-31G*

MPW*X* is a method that uses the mPW exchange functional of Adamo and Barone (Adamo, C.; Barone, V. *J. Chem. Phys.* **1998**, *108*, 664), the PW91 correlation functional (Perdew, J. P. Electronic Structure of Solids '91; Zieesche, P., Eshrig, H., Eds.; Akademie: Berlin, 1991) and a percentage of HF exchange, *X*. Note that MPW*X* includes the following special cases:

MPW0 ≡ mPWPW91
MPW6 ≡ MPW1S
MPW25 ≡ mPW1PW91
MPW42.8 ≡ MPW1K
MPW60.6 ≡ MPW1KK

For all of the MPW*X* methods listed above, CM3 has been parameterized for five specific values of *X*, namely 0, 25, 42.8, 60.6, and 99.9, and these parameter sets are available in *MN-GSM*. Every CM3 and CM4 parameter is a linear or a quadratic function of the percentage of HF exchange used in the mPW exchange functional. So, in addition to the specific CM3 and CM4 parameter sets (i.e. when *X* in MPW*X* is 0, 25, 42.8, 60.6, and 99) the CM3 and CM4 Charge Models are available for any value of *X* in MPW*X* between 0.0 and 100.0. Note that the CM3 and CM4 parameters were optimized using a corrected version of the modified Perdew-Wang density functional as implemented in *Gaussian*. The details of this correction are described fully in "The Effectiveness of Diffuse Basis Functions for Calculating Relative Energies by Density Functional Theory" by Lynch, B. J.; Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A*, **2003**, *107*, 1384.

The CM3 model for the BLYP and B3LYP methods uses a slightly modified mapping scheme for compounds that contain N and O. For more information, see "Parameterization of Charge Model 3 For AM1, PM3, BLYP, and B3LYP" by Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem.*, **2003**, *24*, 1291. We have also developed a special CM3 model for assigning partial atomic charges to high-energy materials. This model is called CM3.1, and it uses the same mapping scheme as the CM3 model for BLYP and B3LYP. This model has been parameterized for use with HF/MIDI!, and is described in "Accurate Partial Atomic Charges for High-Energy Molecules with the MIDI! Basis Set" by Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.*, **2005**, *113*, 133.

- Gas-phase and liquid-phase CM4 class IV charges can be determined for the following combinations of electronic structure theory and basis set (using either a restricted or an unrestricted formalism):

BLYP/MIDI!6D
BLYP/6-31G*
G96LYP/MIDI!6D
G96LYP/6-31G*
B3LYP/MIDI!6D
B3LYP/6-31G*
MPW*X*/MIDI!6D
MPW*X*/6-31G*

BLYP/6-31+G*
BLYP/6-31+G**
G96LYP/6-31+G*
G96LYP/6-31+G**
B3LYP/6-31+G*
B3LYP/6-31+G**
MPW*X*/6-31+G*
MPW*X*/6-31+G**

- Liquid-phase calculations based on gas-phase geometries can be performed with SM5.42 for the following restricted and unrestricted Hartree-Fock, DFT, and adiabatic-connection-method wave functions (i.e. hybrid DFT wave functions) that employ spherical harmonic or Cartesian *d* functions:

HF/MIDI!
HF/MIDI!6D
HF/6-31G*
BPW91/MIDI!
BPW91/MIDI!6D
B3LYP/MIDI!
BPW91/6-31G*
HF/6-31+G*
HF/cc-pVDZ
BPW91/DZVP

- Liquid-phase calculations based on gas-phase geometries can be performed with SM5.43 for the following restricted and unrestricted Hartree-Fock, DFT, and adiabatic-connection-method wave functions (i.e. hybrid DFT wave functions) that employ spherical harmonic or Cartesian *d* functions:

HF/6-31G*
MPW*X*/MIDI!
MPW*X*/6-31G*
MPW*X*/6-31+G**
B3LYP/6-31G*
MPW*X*/MIDI!6D
MPW*X*/6-31+G*

- Liquid-phase calculations based on gas-phase geometries can be performed with SM6 and SM6T for the following restricted and unrestricted DFT and adiabatic-connection-method wave functions (the four basis sets for which SM6 is parameterized use Cartesian *d* functions):

BLYP/MIDI!6D
BLYP/6-31G*
G96LYP/MIDI!6D
G96LYP/6-31G*
B3LYP/MIDI!6D
B3LYP/6-31G*
MPW*X*/MIDI!6D
MPW*X*/6-31G*
BLYP/6-31+G*
BLYP/6-31+G**
G96LYP/6-31+G*
G96LYP/6-31+G**
B3LYP/6-31+G*
B3LYP/6-31+G**
MPW*X*/6-31+G*
MPW*X*/6-31+G**

- Liquid-phase analytical gradients for SM6 are available for basis sets that use Cartesian *d* shells, namely:

MIDI!6D
6-31+G*
6-31G*
6-31+G**

Note that the B3LYP options in the list above should use the standard version III VWN functional, which is requested with the ‘DFTTYP=B3LYP3’ keyword in data group $DFT.

- Löwdin population analysis partial atomic charges can be used in conjunction with the generalized Born method to calculate the electrostatic contribution to the free energy of solvation using HF, DFT, and hybrid DFT and basis sets containing *s*, *p*, *d*, and *f* functions. For basis sets involving Cartesian *d* and *f* functions, analytic gradients of the generalized Born free energy are available, and they can be used for geometry optimizations and numerical Hessian and vibrational frequency calculations.

- Redistributed Löwdin population analysis charges can be used in conjunction with the generalized Born method to calculate the electrostatic contribution to the free energy of solvation using HF, DFT, and hybrid DFT and the 6-31G* and 6-31+G** basis sets. Analytic gradients of the generalized Born free energy are available, and they can be used for geometry optimizations and numerical Hessian and vibrational frequency calculations.

- CM2, CM3, and CM4 charges can be used in conjunction with the generalized Born method to calculate the electrostatic contribution to the free energy of solvation using any of the CM2, CM3, and CM4 methods detailed above. Liquid-phase geometry optimizations and Hessian and vibrational frequency analysis calculations are available for the CM2, CM3, and CM4 methods for which analytical gradients of the generalized Born solvation energy are available.

- The necessary modification of NDDO Hamiltonians to carry out AM1-SRP and PM3-SRP calculations has been implemented.

- *GAMESSPLUS* includes the *GAMESSPLUS* Solubility Utility for calculating the solubility of a given solute A in a given solvent B. This utility is described in a self-contained section of this manual. Therefore users who only want to calculate solubilities do not need to be familiar with the entire *GAMESSPLUS* manual.

- *GAMESSPLUS* includes the *GAMESSPLUS* Soil Sorption Utility for calculating the soil sorption coefficients. This utility is described in a self-contained section of this manual. Therefore users who only want to calculate soil sorption coefficients do not need to be familiar with the entire *GAMESSPLUS* manual.

- *GAMESSPLUS* can now be used with the *CHARMM* package through the *CHARMM/GAMESSPLUS* interface for combined QM/MM calculations with the QM/MM boundary treated by the generalized hybrid orbital (GHO) method at the *ab initio* HF level (GHO-AIHF). A parametrized version of GHO-AIHF is available for the MIDI! basis set.

- The QM energy in the presence of an external electrostatic potential with a site-site representation of the QM/MM electrostatic interaction energy can be calculated. The first and second derivatives with respect to coordinates and electrostatic potentials are available.

- The "rungmsplus" script allows *GAMESSPLUS* to be run conveniently even on computers where the script "rungms" provided in the *GAMESS* distribution fails.

Extended Abstract

Löwdin population analysis and redistributed Löwdin population analysis

Löwdin population analysis, like Mulliken analysis, provides class II atomic partial charges, but the Löwdin method has certain advantages. It has been implemented in *GAMESSPLUS* because Löwdin population analysis charges are used for obtaining CM2, CM3, and CM4 charges. However, there may be some independent interest in Löwdin analysis since it can be used with any basis set (whereas CM2, CM3, and CM4 cannot), and it is expected usually to yield more useful population analyses than Mulliken's method. Note that Löwdin and Mulliken charges are identical for AM1 and PM3 because overlap is neglected in these methods.

Partial atomic charges obtained from Löwdin population analysis can, however, be sensitive to basis set size, particularly for extended basis sets that include diffuse functions. We have developed and implemented a new method, called redistributed Löwdin population analysis (or RLPA), which alleviates some of this sensitivity to basis set size. For methods using diffuse basis sets (6-31+G* and 6-31+G**), RLPA charges are used for obtaining CM3 and CM4 charges.

Charge Model 2, Charge Model 3, and Charge Model 4

Charge Model 2 (CM2), Charge Model 3 (CM3), and Charge Model 4 (CM4) are our second-, third-, and fourth-generation models of class IV charges, respectively. Charge Model 3 has been parameterized with a larger training set than CM2 (398 data vs. 198 data), and it is available for different combinations of electronic structure theory and basis sets than CM2. Furthermore, it is parameterized for Li and for molecules that contain Si-O, Si-F, and Si-Cl bonds (CM2 is not). Charge Model 4 has been parametrized against the same training set that CM3 was, except that CM4 gives improved charges for aliphatic functional groups, which is important for modeling hydrophobic effects. Class IV charges are obtained by a semiempirical mapping of class II charges defined for a particular basis set. CM2, CM3, and CM4 use the same semiempirical mapping scheme. Class IV charges have the following advantages over class III charge models (*e.g.*, ChElPG and Merz-Kollman algorithms):

- Class III charges are unreliable for buried charges (this problem is widely recognized, as discussed in work by Kollman and Francl and their respective coworkers). Class IV charge models provide a practical and stable way to obtain reasonable charges for buried atoms.

- Class III charges are at best as good as the basis set and wave function used, whereas class IV charges represent extrapolation to full CI with a complete basis.

Class IV charges are useful for any purpose for which ChElPG or Merz-Kollman charges are useful, but we believe that they are better.

SM5.42, SM5.43, SM6, and SM6T Solvation Models

SM5.42, our earliest solvation model implemented in *GAMESSPLUS*, is a universal Solvation Model based on SM5 functional forms for atomic surface tensions (hence the first three characters in the name of the method are SM5), built on class IV point charges (hence .4 comes next) of the CM2 type (hence 2). A more recent model, called SM5.43, uses the same functional forms for atomic surface tensions as does SM5.42, but SM5.43 uses CM3 charges (hence the 3 in the name). Our most recent solvent model, called SM6, is based on SM6 functional forms for atomic surface tensions and uses class IV CM4 point charges.

The SM*x* solvation models provide a way to calculate electronic wave functions in liquid-phase solution and free energies of solvation. The SM5 models are universal, *i.e.*, parameterized for water and any organic solvent. With universal models, if desired, one can calculate solvation free energies for two different solvents (*e.g.*, water and 1-octanol) and use the results to calculate log *P*, where *P* is the partition coefficient. Presently, the SM6 model has only been parametrized for aqueous solvent.

For solvation calculations based on gas-phase geometries, the standard-state free energy of solvation $\Delta G_{\mathrm{S}}^{\mathrm{o}}(\mathbf{R})$ is given by two components:

$$\Delta G_{\mathrm{S}}^{\mathrm{o}}(\mathbf{R}) = \Delta G_{\mathrm{EP}} + G_{\mathrm{CDS}} \qquad (1)$$

where

$$\Delta G_{\mathrm{EP}} = \Delta E_{\mathrm{E}} + G_{\mathrm{P}} \qquad (2)$$

In equations (1) and (2), ΔG_{EP} is the bulk electrostatic component of the solvation free energy; it is the sum of the polarization energy G_{P} (representing favorable solute-solvent interactions and the associated solvent rearrangement cost) and the distortion energy ΔE_{E} (the cost of distorting the solute electronic charge distribution to be self-consistent with the solvent electric polarization). G_{CDS} accounts for first-solvation-shell effects.

ΔG_{EP} is determined by a self-consistent reaction field (SCRF) calculation, which allows the solvent-induced change in the solute electronic wave function to be optimized variationally. G_{CDS} is not a self-consistent term; it has no effect on the solute electronic wave function. In its simplest form G_{CDS} is defined as:

$$G_{\mathrm{CDS}} = \sum_{k} A_k \sigma_k \qquad (3)$$

where A_k is the exposed surface area of atom *k* (this depends on the solute's 3-D geometry and is calculated by the Analytical Surface Area (ASA) algorithm as described in reference *LH95* and as included in recent versions of *AMSOL*, and σ_k is the atomic surface tension of atom *k*. The atomic surface tension σ_k is itself a function of the solute's 3-D geometry and a small set of solvent descriptors. References *LH98*, *ZL98*, and *LZ99* present a more expanded form of G_{CDS} than what appears in equation (3):

$$G_{\mathrm{CDS}} = \sum_k \sigma_k^A A_k\left(\{r_k : r_{\mathrm{S}}^{\mathrm{CD}}\}\right) + \sigma^{\mathrm{CS}} \sum_k A_k\left(\{r_k ; r_{\mathrm{S}}^{\mathrm{CS}}\}\right) \qquad (4)$$

where *k* represents atoms, σ_k^A is the partial atomic surface tension for atom *k*, and $A_k\left(\{r_k ; r_{\mathrm{S}}^{CX}\}\right)$, with *X* = C or D, is the solvent-accessible surface area of atom k that is dependant upon a set of parameterized set or effective solute radii, $\{r_k\}$, and on one or another effective solvent radii, $r_{\mathrm{S}}^{\mathrm{CD}}$ and $r_{\mathrm{S}}^{\mathrm{CS}}$. Although $r_{\mathrm{S}}^{\mathrm{CD}}$ and $r_{\mathrm{S}}^{\mathrm{CS}}$ have different nonzero values for some solvation models, they are both set to zero in the SM5.42 model. In the SM5.43 and SM6 model, we use 0.40 Å for both of these values. Each σ_k^A is a function of one-, two-, and/or three-atom surface tension coefficients, $\tilde{\sigma} \in \{\tilde{\sigma}_{z_k}, \tilde{\sigma}_{z_k z_{k'}}, \tilde{\sigma}_{z_k z_{k'} z_{k''}}\}$, that depend on the atomic numbers z_k, $z_{k'}$, and $z_{k''}$ of atoms *k*, *k*′, and *k*″, respectively, and the switching functions, $f_{z_k z_{k'}}(R_{kk'})$ and $f_{z_k z_{k'} z_{k''}}(R_{kk'}, R_{k'k''},)$, that are functions of interatomic distances, $R_{kk'}$ and $R_{k'k''}$:

$$\sigma_k = \tilde{\sigma}_{z_k} + \sum_{k'} \tilde{\sigma}_{z_k z_{k'}} f_{z_k z_{k'}}(R_{kk'}) + \sum_{k',k''} \tilde{\sigma}_{z_k z_{k'} z_{k''}} f_{z_k z_{k'} z_{k''}}(R_{kk'}, R_{k'k''}) \qquad (5)$$

For water solvent, σ^{CS} is absorbed into σ_k^A, and a set of $\tilde{\sigma}_i \in \{\tilde{\sigma}_{z_k}, \tilde{\sigma}_{z_k z_{k'}}, \tilde{\sigma}_{z_k z_{k'} z_{k''}}\}$ is specifically determined constants (similar to $\bar{\sigma}_k^A$, which will be defined below).

For a particular organic solvent, each $\tilde{\sigma}_i$ (with *i* = z_k, $z_{k'}$, or $z_{k''}$) and σ^{CS} are defined to be linear functions of bulk solvent properties:

$$\tilde{\sigma}_i = \ddot{\sigma}_i^{(n)} n + \ddot{\sigma}_i^{(\alpha)} \alpha + \ddot{\sigma}_i^{(\beta)} \beta \qquad (6)$$

$$\sigma^{\mathrm{CS}} = \ddot{\sigma}^{(\gamma)} \gamma + \ddot{\sigma}^{(\beta^2)} \beta^2 + \ddot{\sigma}^{(\phi^2)} \phi^2 + \ddot{\sigma}^{(\psi^2)} \psi^2 \qquad (7)$$

where *n*, 〈, ®, ©, ϕ, and ⁄ refer to the index of refraction, Abraham's hydrogen bond acidity and hydrogen bond basicity, macroscopic surface tension, electronegative halogenicity, and aromaticity, respectively. The constants $\ddot{\sigma}_i^{(n)}$, $\ddot{\sigma}_i^{(\alpha)}$, $\ddot{\sigma}_i^{(\beta)}$, $\ddot{\sigma}^{(\gamma)}$, $\ddot{\sigma}^{(\beta^2)}$, $\ddot{\sigma}^{(\phi^2)}$, and $\ddot{\sigma}^{(\psi^2)}$ have been optimized using a large number of solvent/solute combinations for which experimental solvation energies are available.

The surface tension functional forms are the same in all SM5.42 and SM5.43 models. SM6 and SM6T use a different set of functional forms. The SM6 and SM6T functional forms are better for most purposes than those used in SM1–SM5.

Incorporating temperature dependence into the SM*x* models: SM6T

To account for the variation of the free energy of solvation as a function of temperature, the temperature dependence of both the bulk electrocstatic, ΔG_{EP}, and the non-bulk electrostatic, ΔG_{CDS}, contributions are included. The effect of temperature on the bulk-electrostatic contributions to the free energy of solvation is accounted for using a temperature dependent dielectric constant, $\varepsilon(T)$ which was computed using the following equation

$$\varepsilon(T)=249.21-.79T+.00072T^2 \tag{8}$$

where *T* is the temperature of the aqueous solvent. This is a empirically derived equation found in the *CRC Handbook of Chemistry and Physics 76th edition*, ed. Lide, D. R., 1995, CRC Press, New York. The variation of the free energy of solvation due bulk electrostatic contributions are quite small. The majority of the temperature dependence of aqueous free energies of solvation must by accounted for using ΔG_{CDS}.

In SM6T the ΔG_{CDS} term mimics the thermodynamic equation for the temperature dependence of free energies of solvation where the thermodynamic properties, the heat capacity and the entropy of solvation, have been replaced parameterized atomic surface tensions:

$$G_{\mathrm{CDS}}(T) = (T - 298)\sum_k A_k \sigma_k^B + \left(T - 298 - T \ln \frac{T}{298} \right) \sum_k A_k \sigma_k^C \tag{9}$$

where σ_k^B and σ_k^C are atomic surface tensions with identical functional forms to those of σ_k, but the parameters are different. Caution should be used in assigning any physical meaning to the atomic surface tensions shown above. While the sum $\sum_k A_k \sigma_k^B$ appears to be the solute's entropy of solvation and the sum $\sum_k A_k \sigma_k^C$ appears to be the solute's heat capacity, it must be pointed out that some of the temperature dependence of the free energy of solvation has been accounted for in the electrostatic term. Additionally the covariance between the two terms in the above equation and the relatively small number of points for each compound (on average 10 points were used) means that the actual numerical values of these two terms may vary significantly from experimental entropies and heat capacities of solvation while still reproducing experimental values with high accuracy. Note that the model has only been developed for C, H, and O containing compounds in aqueous solutions for the temperature 273 to 373 K.

A comment on using gas-phase geometries to calculate solvation free energies.

For SM1–4 and SM5.4, geometry optimization in solution was an essential part of the parameterization. SM5.42, SM5.43, SM6 and SM6T are parameterized in such a way that one fixes the geometry at a reasonable value (any reasonably accurate gas-phase geometry should be acceptable) and calculates the solvation energy without changing the geometry. Thus, geometry optimization in the presence of solvent is not required to obtain accurate solvation free energies. This method of obtaining solvation parameters based on gas-phase geometries was adopted for several reasons. First, previous experience has shown that the difference one gets from re-optimizing the geometry in the presence of solvent in *almost* all cases is small—less than the average uncertainty in the method or in any competing method. Second, for many solutes, less expensive (e.g. semiempirical or molecular mechanics methods) can yield accurate gas-phase geometries. Third, for other solutes, such as transition states, solutes with low-barrier torsions, multiple low-energy conformations, weakly bound complexes, and in cases where one or more solvent molecules are treated explicitly, more expensive levels of theory might be needed to yield accurate geometries. Finally, solvation energies obtained using gas-phase geometries can be added conveniently to gas-phase energies for separable-equilibrium-solvation dynamics calculations.

In some cases, geometry optimization in the presence of solvent is important. In these cases, one can also apply the SM5.42, SM5.43, or SM6 models at a solute geometry **R** that is not an approximation to an equilibrium gas-phase geometry. This type of calculation corresponds to the fixed-**R** solvation energy, which is still given by $\Delta G_{\mathrm{S}}^{\mathrm{o}}(\mathbf{R})$ of equation (1). Evaluation of this quantity for geometries that do not correspond to an equilibrium structure is useful for dynamics calculations because the potential of mean force is given by

$$W(\mathbf{R}) = V(\mathbf{R}) + \Delta G_{\mathrm{S}}^{\mathrm{o}}(\mathbf{R}) \tag{10}$$

where *V*I is the gas-phase potential energy surface (which is itself given by the sum of the gas-phase electronic energy and the gas-phase nuclear repulsion energy). If one applies the SM5.42, SM5.43, or SM6 models to a geometry optimized in solution and subtracts the gas-phase energy at a geometry optimized for the gas phase, one obtains the true solvation energy for the given method.

Note that both free energies of solvation and potentials of mean force depend on temperature; all of our solvation models are currently only parameterized for a single temperature, namely 298 K.

Furthermore $\Delta G_{\mathrm{S}}^{\mathrm{o}}(\mathbf{R})$ depends on standard state choices; the values given directly by the SM5, SM6, and SM6T models correspond to using the same molar density (*e.g.*, one mole per liter) in the gas phase and in the liquid-phase solution. Furthermore the liquid-solution standard state corresponds to an ideal dilute solution at that concentration. However, one may adjust the results to correspond to other choices of standard state by standard thermodynamic formulae. Note that changing the standard state corresponds to adding a constant to *W*I; thus the gradient of *W*I, which is used for dynamics, is not affected.

Allowed Combinations of Solvent Model, Electronic Structure Theory, and Basis Set.

When using *GAMESSPLUS*, the SM5.42 models should only be applied with the ten combinations of basis sets and density functionals for which parameters are currently available in *Gaussian*:

Standard notation	*GAMESS keyword(s)*	*notes*
AM1		
PM3		
HF/MIDI!	HF/MIDIX	
HF/MIDI!6D	HF/MIDIX 6D	
HF/6-31G*	HF/6-31G*	
BPW91/MIDI!	BPW91/MIDIX	
BPW91/MIDI!6D	BPW91/MIDIX 6D	
B3LYP/MIDI!	B3LYP/MIDIX	
BPW91/6-31G*	BPW91/6-31G*	
HF/6-31+G*	HF/6-31+G*	
HF/cc-pVDZ	HF/GEN 5D	*a*
BPW91/DZVP	BPW91/GEN 6D	*b*

[a] See *Special notes on basis sets* section

[b] Note that the DZVP basis set must be provided as a general basis.

Similarly, the SM5.43 models should only be applied with the seven combinations of basis sets and density functionals for which parameters are currently available in *Gaussian*:

Standard notation	*GAMESS keyword(s)*	*notes*
HF/6-31G*	HF/6-31G*	
B3LYP/6-31G*	B3LYP/6-31G*	
MPW*X*/MIDI!	MPW*X*/MIDIX	*a*
MPW*X*/MIDI!6D	MPW*X*/MIDIX 6D	*b*
MPW*X*/6-31G*	MPW*X*/6-31G*	*b*
MPW*X*/6-31+G*	MPW*X*/6-31+G*	*b*
MPW*X*/6-31+G**	MPW*X*/6-31+G**	*b*

[a] Analytic gradients not available for this method.

[b] The SM5.43 parameters developed for this method can be applied to wave functions that use any value *X* of percentage Hartree-Fock exchange.

The SM6 model has been tested against several different density functionals, and has been shown to retain its accuracy when different density functionals besides MPW*X* (the method against which the CM4 and SM6 parameters were originally developed). Thus, the SM6 model is only basis-set-dependent, and can be used with any good density functional. Currently, CM4 and SM6 parameters exist for the following four basis sets:

MIDI!6D
6-31G*
6-31+G*
6-31+G**

A list of density functionals that are available in *GAMESSPLUS* and that are recommended for use with CM4 and/or SM6 is in the section entitled "*Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS*".

The SM6T model is applicable to the same combinations of theory and basis set as SM6. The temperature dependence of the bulk-electrostatic term, ΔG_{ENP}, is quite small and so the variation due to the level of electronic structure theory is within the error of the model. Thus, parameterization of the temperature-dependent component of the model for multiple levels of theory and basis sets is unnecessary and only depends on the applicability of SM6.

Why Use SM5.42, SM5.43, or SM6?

- The semiempirical CDS terms make the above models more accurate than alternative models for absolute free energies of solvation of neutral solutes.
- SM5.42 and SM5.43 are universal models, *i.e.*, the semiempirical parameters are adjusted for water and for all solvents for which a small number of required solvent descriptors are known or can be estimated; this includes essentially any organic solvent.
- SM5.42, SM5.43, and SM6 use class IV charges to calculate the bulk electrostatic contribution to the solvation free energy; this is typically more accurate than calculating the charge distribution directly from the approximate wave function. This has two consequences:
    (1) The electrostatics are more accurate.
    (2) CM2, CM3, and CM4 yield very accurate charges both in the gas phase and in liquid-phase solutions, and this is useful for a qualitative understanding of solvent-induced changes in the solute.
- SM5 and SM6 parameterizations included an extremely broad range of solute functional groups, including molecules containing phosphorus, which are very hard to treat.
- SM5 and SM6 do not need to be corrected for outlying charge error, which can be large in some other methods.

Furthermore, our most recent model, SM6, has several advantages compared to earlier solvent models (e.g. SM5.42 and SM5.43) developed within our group:

- SM6 can be used with any of the density functional methods supported in *GAMESSPLUS*. For a list of recommended density functional methods that are supported in *GAMESSPLUS*, see the section entitled "*Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS*".

- SM6 significantly outperforms SM5.42, SM5.43, as well as all other competing continuum solvation models that it has been tested against for predicting aqueous solvation free energies of ions. This is important because aqueous solvation free energies of ions can be used in various thermodynamic cycles to calculate pK_a.
- When explicit water molecules are added to the calculation (which is sometimes required for accurately calculating aqueous solvation free energies of some solutes, in particular ions with concentrated regions of charge) the performance of SM6 significantly improves, whereas the performance of SM5.42, SM5.43, and most other competing continuum solvation models that SM6 has been tested against worsens. Thus, unlike other popular continuum solvation models, SM6 can be applied to solute-water clusters.
- SM6 uses an improved set of surface tension functionals; using this new set of surface tension functionals significantly improves the performance of the model for molecules containing peroxide functional groups.
- SM6 uses class IV CM4 charges, which give more realistic partial atomic charges for aliphatic groups than our previous class IV models; this is important for accurately modeling hydrophobic effects.

Why Use SM6T?

- SM6T is one of the few implicit models capable predicting aqueous free energies of solvation as a function of temperature.
- SM6T has been tested against a much broader range of compounds than prior models.
- SM6T is based on SM6, and thus shares its advantages**.**

Analytical gradients and geometry optimization

Analytical gradients have been implemented in *GAMESSPLUS* beginning with version 2.0. In particular, *GAMESSPLUS* contains analytical gradients for restricted and unrestricted wave functions for basis sets with Cartesian *d* shells. However, analytical gradients are not available for basis sets with spherical harmonic *d* functions (e.g., for HF/MIDI!, HF/cc-pVDZ), and methods using basis sets containing functions higher in angular momentum than *f*. Analytical gradients are also available when the AM1 and PM3 or method is used.

The availability of gradients allows for efficient geometry optimization in liquid-phase solution. This is necessary in many cases. For example, the transition state geometry for the S_N2 reaction of ammonia and chloromethane (the Menschutkin reaction) depends strongly on solvent. Other applications include the study of phase-dependent reaction mechanisms and molecular conformational preferences.

Equation (1) can be rewritten as:

$$\Delta G_S^{\mathrm{o}}(\mathbf{R}) = G^{\mathrm{o}}(l, R_e(g)) - G^{\mathrm{o}}(g, R_e(g)) \qquad (11)$$

where $\Delta G_\mathrm{S}^\mathrm{o}(\mathbf{R})$ is the difference in the standard-state free energy of a molecular species fixed in its gas-phase equilibrium geometry, $R_e(g)$, in liquid-phase solution, *l*, and in the gas phase, *g*. Furthermore,

$$G^\mathrm{o}(l, R_e(g)) = E(l, R_e(g)) + G_\mathrm{P}(R_e(g)) + G_\mathrm{CDS}(R_e(g)) \tag{12}$$

and

$$G^\mathrm{o}(g, R_e(g)) = E(g, R_e(g)) \tag{13}$$

where $E(l, R_e(g))$ is the SCRF energy in the presence of the solvent, while $E(g, R_e(g))$ is the gas-phase SCF energy.

Geometry optimization in solution yields the solute equilibrium geometry, $R_e(l)$. In the SM5.42, SM5.43, and SM6 models, the definition of $\Delta G_\mathrm{S}^\mathrm{o}$ is analogous to the definition of $\Delta G_\mathrm{S}^\mathrm{o}(\mathbf{R})$ in equation (9):

$$\begin{aligned} \Delta G_\mathrm{S}^\mathrm{o} &= G^\mathrm{o}(l,\ R_e(l)) - G^\mathrm{o}(g,\ R_e(g)) \\ &= E(l,\ R_e(l)) + G_\mathrm{P}(R_e(l)) + G_\mathrm{CDS}(R_e(l)) - E(g,\ R_e(g)) \end{aligned} \tag{14}$$

A full derivation of the analytical gradient is presented in the paper by T. Zhu *et al.* entitled “Analytical Gradients of a Self-Consistent Reaction-Field Solvation Model Based on CM2 Atomic Charges” (*J. Chem. Phys.* **1999**, *110*, 5503-5513).

SCRF calculations using the generalized Born model and CM2, CM3, or CM4 charges

A self-consistent-reaction-field method based on the generalized Born approximation and CM2, CM3, or CM4 charges has been implemented in the program. One can use the SCRF orbitals to calculate the polarization and distortion energies and solvated atomic partial charges and to investigate solvation effects on the atomic charge distribution, multipole moments, and other one-electron properties. In addition, the SCRF calculation is one of the two components of a calculation of the free energy of solvation.

NDDO and CM2 Specific Reaction Parameters (SRP) Models

GAMESSPLUS can use specific reaction parameters (i.e., nonstandard parameters) for the NDDO Hamiltonians of the AM1 and PM3 models in the gas-phase for the CM2/AM1 and CM2/PM3 methods and in the liquid-phase for the CM2/AM1, CM2/PM3, SM5.42/AM1, and SM5.42/PM3 methods.

AM1 and PM3 calculations in either the gas-phase or liquid-phase may be performed without using the arithmetic mean rule for the resonance parameters. In standard AM1 and PM3 calculations, the resonance parameter $\beta_{lxl'y}$ for interaction of an orbital with angular momentum *l* on an atom of element *x* and an orbital with angular momentum *l'* on an atom of element *y* is given by

$$\beta_{lxl'y} = (\beta_{lx} + \beta_{l'y})/2 \qquad (15)$$

where β_{lx} and $\beta_{l'y}$ are standard parameters. The user can now override eq. (9) by inputting specific values of the resonance parameter for one or more sets of *l*, *x*, *l'* and *y*. A reference for this general procedure is reference *CE95* in the *Literature References* section.

Solubility Calculations

The solubility of a given solute A, in a liquid solvent, B are calculated using a thermodynamic relationship between the solubility, free energy of solvation, and pure-substance vapor pressure of solute A, which is given by:

$$S = \left(P_{\mathrm{A}}^{\bullet} \middle/ P^{\mathrm{o}} \right) \exp\left[-\Delta G_{\mathrm{S}}^{\mathrm{o}} \middle/ RT \right] \qquad (16)$$

In this equation, *S* is the solubility of solute A in solvent B, $P_{\mathrm{A}}^{\bullet}$ is the equilibrium vapor pressure of solute A over a pure solution of A, P^{o} is the pressure of an ideal gas for a given standard-state (a 1 molar standard-state at 298 K is used in this calculation for all phases; therefore P^{o} is 24.45 atm), $\Delta G_{\mathrm{S}}^{\mathrm{o}}$ is the standard-state free energy of solvation of solute A in solvent B, *R* is the universal gas constant, and *T* is temperature. Full details are given in the section of this manual entitled *GAMESSPLUS Solubility Utility*.

Soil Sorption Calculations

For a given solute, the soil sorption coefficient (K_{OC}) is defined as

$$K_{\mathrm{OC}} = \frac{C_{\mathrm{soil}}/C_{\mathrm{soil}}^{\circ}}{C_{\mathrm{w}}/C_{\mathrm{w}}^{\circ}} \qquad (17)$$

where C_{soil} is the concentration of solute per gram of carbon in standard soil, C_{w} is the concentration of solute per volume of aqueous solution, and $C_{\mathrm{soil}}^{\circ}$ and C_{w}° are the standard state concentrations of organic carbon for soil and aqueous solution, respectively. Typically, a standard state of 1 μg of solute/g of organic carbon is used for $C_{\mathrm{soil}}^{\circ}$, and 1 mol/L is used for C_{w}°. K_{OC} may be calculated according to

$$K_{\mathrm{OC}} = \rho_{\mathrm{soil}} \left(\Delta G_{\mathrm{w}}^{\circ} - \Delta G_{\mathrm{soil}}^{\circ} \right) \qquad (18)$$

where ρ_{soil} is the density of soil (in g/mL), $\Delta G_{\mathrm{w}}^{\circ}$ is the standard state free energy associated with transferring a solute from the gas phase to aqueous solution, and $\Delta G_{\mathrm{soil}}^{\circ}$ is the standard state free energy associated with transferring a solute from the gas phase to soil. Full details are given in the section of this manual entitled *GAMESSPLUS Soil Sorption Utility*.

QM/MM Calculations at the *Ab Initio* HF Level with the GHO Boundary Treatment

GAMESSPLUS can be compiled into a *CHARMM/GAMESSPLUS* combination package for calculations that combine *ab initio* HF wave functions with molecular mechanics. For the QM/MM partition along a covalent bond, the generalized hybrid orbital (GHO) method is used to provide a smooth connection between the QM subsystem and the MM subsystem. In the GHO treatment, sp^3 carbons are often chosen as GHO boundary atoms, denoted by B. Atom B is both a QM atom and an MM atom. The QM atom bonded to B is called a QM frontier atom, denoted by A. The other three MM atoms directly bonded to B are denoted by X, Y, and Z. A set of generalized hybrid orbitals $\{\eta_\mathrm{B}, \eta_x, \eta_y, \eta_z\}$ is placed on the GHO boundary atom B, where the hybridization scheme is completely determined by the local geometry of the QM/MM boundary (atoms Q, B, X, Y, and Z). Among the four hybrid orbitals, one approximately pointing toward A (denoted by η_B) will participate in the SCF procedure with other QM basis functions and is therefore called an active hybrid orbital. The remaining three hybrid orbitals $\{\eta_x, \eta_y, \eta_z\}$ are called auxiliary orbitals, and they are excluded from the SCF procedure. With this restriction, on one hand, the active molecular orbitals (MOs) in GHO are only allowed to be expanded over the active basis functions (including η_B). On the other hand, each auxiliary hybrid orbital forms an auxiliary MO by itself, and it is occupied by a fixed auxiliary charge density. To distribute the MM point charge on B over the three auxiliary orbitals, the charge density for each auxiliary orbital is determined as $1 - q_\mathrm{B}/3.0$, where q_B denotes the MM point charge on B.

In *GAMESSPLUS*, the implementation of GHO at the *ab initio* HF level (GHO-AIHF) is based on algorithms described in the paper of J. Pu, J. Gao, and D. G. Truhlar (see Ref. PG04). The major features of this extension include: (i) The basis set on the GHO boundary B is chosen as an STO-3Gv set; the 1*s* core electrons are not explicitly present. (ii) The active basis functions are orthogonalized to the auxiliary orbitals to maintain the global MO orthonormal constraints. Four orthogonalization schemes are proposed and implemented. (iii) The GHO gradients are calculated analytically by incorporating additional forces due to the basis transformations of the GHO scheme. Further details of the *CHARMM/GAMESSPLUS* combination package are given in the *CGPLUS* user manual.

Electrostatically Embedded QM Calculation with a Site-Site Representation of the QM/MM Electrostatic Interaction

In the electrostatically embedded QM calculations with a site-site representation of the QM/MM electrostatic interaction, the sum of the QM electronic energy and the QM/MM electrostatic interaction energy is given by

$$V^{\mathrm{EEQM}}\left(\mathbf{R},\mathbf{\Phi}\right)=\left\langle\Psi\middle|\hat{H}_0+\hat{\mathbf{Q}}^{\mathrm{T}}\mathbf{\Phi}\middle|\Psi\right\rangle, \qquad (19)$$

where $\mathbf{R}$ stands for the collection of the coordinates $\mathbf{R}_a$ $\left(a = 1, 2, \cdots, N^{\text{QM}}\right)$ of atoms in the QM region, Ψ is the electronic wave function, $\hat{H}_0$ is the electronic Hamiltonian (including nuclear repulsion) of the QM region, $\hat{Q}_a$ is the population operator that generates the partial charge Q_a on QM atomic site a,

$$Q_a = \left\langle \Psi \middle| \hat{Q}_a \middle| \Psi \right\rangle , \tag{20}$$

and Φ_a is the electrostatic potential at atom a from the MM region. In *GAMESSPLUS*, one can choose the operator $\hat{Q}_a$ according to Löwdin population analysis, redistributed Löwdin population analysis, Charge Model 2, Charge Model 3, or Charge Model 4.

In *GAMESSPLUS*, one can calculate V^{EEQM} and its first and second derivatives with respect to $\mathbf{R}$ and $\mathbf{\Phi}$ for given $\mathbf{R}$ and $\mathbf{\Phi}$. The first derivative of V^{EEQM} with respect to a component of $\mathbf{\Phi}$ is given by

$$\frac{\partial V^{\text{EEQM}}}{\partial \Phi_a} = \left\langle \Psi \middle| \hat{Q}_a \middle| \Psi \right\rangle = Q_a \ . \tag{21}$$

Then the second partial derivatives of V^{EEQM} (first or second order in electrostatic potential) are

$$\frac{\partial V^{\text{EEQM}}}{\partial \Phi_a \partial R_b} = \frac{\partial Q_a}{\partial R_b} \equiv \kappa_{ab} \ . \tag{22}$$

and

$$\frac{\partial V^{\text{EEQM}}}{\partial \Phi_a \partial \Phi_b} = \frac{\partial Q_a}{\partial \Phi_b} \equiv \chi_{ab} \ , \tag{23}$$

These variables χ_{ab} and κ_{ab} are known as charge response kernels (CRKs). In *GAMESSPLUS*, the CRKs can be obtained by numerical differentiations of the charges.

V^{EEQM} and its first and second derivatives can be used as input for electrostatically embedded multiconfiguration mechanics calculation.

GAMESSPLUS Reference

Publications including work performed with *GAMESSPLUS*, version 4.7 should cite the software package in the following ways:

Journal of Chemical Physics or World Scientific style

M. Higashi, A. C. Chamberlin, J. Pu, C. P. Kelly, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.-Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar, GAMESSPLUS–version 2008, University of Minnesota, Minneapolis, 2008, based on the General Atomic and Molecular Electronic Structure System (GAMESS) as described in M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, *J. Comput. Chem.* **14**, 1347 (1993).

Elsevier style

M. Higashi, A. C. Chamberlin, J. Pu, C. P. Kelly, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.-Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, D. G. Truhlar, GAMESSPLUS–version 2008, University of Minnesota, Minneapolis, 2008, based on the General Atomic and Molecular Electronic Structure System (GAMESS) as described in M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, J. Comput. Chem. 14 (1993) 1347.

ACS style

Higashi, M.; Chamberlin, A. C.; Pu, J.; Kelly, C. P.; Thompson, J. D.; Xidos, J. D.; Li, J.; Zhu, T.; Hawkins, G. D.; Chuang, Y.-Y.; Fast, P. L.; Lynch, B. J.; Liotard, D. A.; Rinaldi, D.; Gao, J.; Cramer, C. J.; Truhlar, D. G. GAMESSPLUS–version 2008, University of Minnesota, Minneapolis, 2008, based on the General Atomic and Molecular Electronic Structure System (GAMESS) as described in Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery J. A. *J. Comput. Chem*. **1993**, *14*, 1347.

Theoretical Chemistry Accounts style

Higashi M,Chamberlin AC, Pu J, Kelly CP, Thompson JD, Xidos JD, Li J, Zhu T, Hawkins GD, Chuang Y-Y, Fast PL, Lynch BJ, Liotard DA, Rinaldi D, Gao J, Cramer CJ, Truhlar DG (2008) GAMESSPLUS–version 4.7, University of Minnesota, Minneapolis, 2008, based on the General Atomic and Molecular Electronic Structure System (GAMESS) as described in Schmidt MW, Baldridge KK, Boatz JA, Elbert ST, Gordon MS, Jensen JH, Koseki S, Matsunaga N, Nguyen KA, Su SJ, Windus TL, Dupuis M, Montgomery JA (1993) J. Comput. Chem. 14: 1347

In addition, as usual, the user should give literature references for any methods used. A convenient collection of literature references in provided in the following section.

Literature References

ASA algorithm

LH95 Liotard, D. A; Hawkins, G. D.; Lynch, G. C.; Cramer, C. J.; Truhlar, D. G. "Improved Methods for Semiempirical Solvation Models" *J. Comp. Chem.* **1995**, *16*, 422.

MIDI! Basis set

EG96 Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. "The MIDI! Basis Set for Quantum Mechanical Calculations of Molecular Geometries and Partial Charges" *Theor. Chim. Acta* **1996**, *93*, 281.

LC98 Li, J.; Cramer, C. J.; Truhlar, D. G. "MIDI! Basis Set for Silicon, Bromine, and Iodine" *Theor. Chem. Acc.* **1998**, *99*, 192.

TW01 Thompson, J. D.; Winget, P.; Truhlar, D. G. "MIDIX basis set for the lithium atom: Accurate geometries and partial atomic charges with minimal computational cost", *PhysChemComm*, **2001**, *4,* 72.

Corrected modified Perdew-Wang hybrid DFT method

LZ03 Lynch, B. J.; Zhao, Y.; Truhlar, D. G. "The Effectiveness of Diffuse Basis Functions for Calculating Relative Energies by Density Functional Theory", *J. Phys. Chem. A* **2003**, *107*, 1384.

MPW1K

LF00 Lynch, B. J.; Fast, P. L.; Harris, M.; Truhlar D. G. "Adiabatic Connection for Kinetics", *J. Phys. Chem. A* **2000**, *104*, 4811.

MPWPW91, mPW1PW91, and MPW*X*

AB98 Adamo, C.; Barone, V.; "Exchange functionals with Improved Long-Range Behavior and Adiabatic Connection Methods without Adjustable Parameters: The mPW and mPW1PW Models" *J. Chem. Phys.* **1998**, *108*, 664.

PT02 Pu, J.; Truhlar, D. G. "Parameterized Direct Dynamics Study of Rate Constants of H with CH4 from 250 to 2400K", *J. Chem. Phys.* **2002**, *116*, 1468.

Löwdin population analysis

LD50 Löwdin, P. –O. *J. Chem. Phys.* **1950**, *18*, 365. (b) Del Re, G.; In *Quantum Science*; Calais, J. –L., Goscinski, O., Linderberg, J., Öhrn, Y., Eds.;Plenum Press: New York, 1976; p 53.

Redistributed Löwdin population analysis

TX02 Thompson, J. D.; Xidos, J. D.; Sonbuchner, T. M.; Cramer, C. J.; Truhlar, D. G. "More Reliable Partial Atomic Charges when using Diffuse Basis Sets", *PhysChemComm*, **2002**, *5*, 117.

Original parameterizations of the CM2 Charge Model

LZ98 Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. "A New Class IV Charge Model for Extracting Accurate Partial Charges from Wave Functions" *J. Phys. Chem. A* **1998**, *102*, 1820.

Parameterization of the CM2 model for HF/cc-pVDZ

LX99 Li, J.; Xing, J.; Cramer, C. J.; Truhlar, D. G. "Accurate Dipole Moments from Hartree-Fock Calculations by Means of Class IV Charges" *J. Chem. Phys.* **1999**, *111*, 885.

Original parameterizations of the CM3 Charge Model

WT02 Winget, P.; Thompson J. D.; Xidos, J. D.; Cramer, C. J.; Truhlar, D. G. "Charge Model 3: A Class IV Charge Model based on Hybrid Density Functional Theory with Variable Exchange" *J. Phys. Chem. A*, **2002**, *106*, 10707.

Extension of CM3 to AM1, PM3, BLYP, and B3LYP

TC03 Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. "Parameterization of Charge Model 3 For AM1, PM3, BLYP, and B3LYP" *J. Comput. Chem.*, **2003**, *24*, 1291.

Extension of CM3 and CM3.1 to HF/MIDI!

KC05a Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. "Accurate Partial Atomic Charges for High-Energy Molecules with the MIDI! Basis Set" *Theor. Chem. Acc.*, **2005**, *113*, 133.

Parameterizations of the CM4 Charge Model

KC05b Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. "SM6: A Density Functional Theory Continuum Solvation Model for Calculating Aqueous Solvation Free Energies of Neutrals, Ions, and Solute-Water Clusters" *J. Chem . Theory Comput.* **2005**, *1*, 1133.

SM5.42/DFT method and ISCRF=1 option

ZL98 Zhu, T.; Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. "Density Functional Solvation Model Based on CM2 Atomic Charges" *J. Chem. Phys.* **1998**, *109*, 9117.

SM5.42/MIDI!6D method and more efficient version of ISCRF=1 option

LH98 Li, J.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. "Universal Reaction Field Model Based on *Ab Initio* Hartree-Fock Theory" *Chem. Phys. Lett.* **1998**, *288*, 293.

Additional SM5.42 parameterizations and ISCRF=2 option

LZ99 Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. "Extension of the Platform of Applicability of the SM5.42R Universal Solvation Model" *Theor. Chem. Acc.* **1999**, *103*, 9.

SM5.42 parameterization for Si

WT02 Winget, P.; Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. "Parameterization of a Universal Solvation Model for Molecules Containing Silicon" *J. Phys. Chem. B*, **2002**, *106*, 5160.

Parameterization of SM5.42 for soil

WC00 Winget, P.; Cramer, C. J.; Truhlar, D. G. "Prediction of Soil Sorption Coefficients Using a Universal Solvation Model" *Env. Sci. Technol.* **2000**, *34*, 4733.

SM5.43 parameterizations for HF/6-31G*, B3LYP/6-31G*, mPW1PW91/6-31G*, and mPW1PW91/6-31+G*

TC04 Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. "New Universal Solvation Model and Comparison of the Accuracy of the SM5.42R, SM5.43R, C-PCM, D-PCM, and IEF-PCM Continuum Solvation Models for Aqueous and Organic Solvation Free Energies and Vapor Pressures" *J. Phys. Chem. A* **2004**, *108*, 6532.

SM5.43 parameterizations for MPW*X*/MIDIX, MPW*X*/MIDI!6D, MPW*X*/6-31G*, MPW*X*/6-31+G*, and MPW*X*/6-31+G**

TC05 Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. "Density-Functional Theory and Hybrid Density-Functional Theory Continuum Solvation Models for Aqueous and Organic Solvents: Universal SM5.43 and SM5.43R Solvation Models for Any Fraction of Hartree-Fock Exchange" *Theor. Chem. Acc.* **2005**, *113*, 107.

SM6 aqueous model

KC05b Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. "SM6: A Density Functional Theory Continuum Solvation Model for Calculating Aqueous Solvation Free Energies of Neutrals, Ions, and Solute-Water Clusters" *J. Chem. Theory Comput.* **2005,** *1*, 1133.

SM6T aqueous model

AC06 Chamberlin, A. C.; Cramer, C. J.; Truhlar, D. G. "Predicting Free Energies of Solvation as Functions of Temperature" *J. Phys. Chem. B.* **2006**, *110,* 5665.

Analytical gradients of the SM5.42, SM5.43, and SM6 models

ZL99 Zhu, T.; Li, J.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. "Analytical Gradients of a Self-Consistent Reaction-Field Solvation Model Based on CM2 Atomic Charges" *J. Chem. Phys.* **1999**, *110*, 5503.

Overview article including SM5.42

HZ99 Hawkins, G. D.; Zhu, T.; Li, J.; Chambers, C. C.; Giesen, D. J.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. "Universal Solvation Models" in *Combined Quantum Mechanical and Molecular Mechanical Methods*, Gao, J., Thompson, M. A., Eds.; American Chemical Society: Washington DC; 1998; pp. 201-219.

Sample applications of SM5.42

TC98 Truhlar, D. G.; Cramer, C. J. "Solvent Effects on 1,3-Dipolar Addition Reactions" *Faraday Discussions* **1998**, *110*, 477.

LC99 Li, J.; Cramer, C. J.; Truhlar, D. G. "Application of a Universal Solvation Model to Nucleic Acid Bases. Comparison of Semiempirical Molecular Orbital Theory, *Ab Initio* Hartree-Fock Theory, and Density Functional Theory" *Biophys. Chem.* **1999**, *78*, 147.

WH00 Winget, P.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. "Prediction of Vapor Pressures from Self-Solvation Free Energies Calculated by the SM5 Series of Universal Solvation Models" *J. Phys. Chem. B*, **2000**, *104*, 4726.

WW00 Winget, P.; Weber, E. J.; Cramer, C. J.; Truhlar, D. G. "Computational Electrochemistry: Aqueous Oxidation Potentials for Substituted Anilines," *Phys. Chem. Chem. Phys.,* **2000**, *2***,** 1231.

CR99 Chuang, Y.-Y.; Radhakrishnan, M. L.; Fast, P. L.; Cramer, C. J.; Truhlar, D. G. "Direct Dynamics for Free Radical Kinetics in Solution: Solvent Effect on the Rate Constant for the Reaction of Methanol with Atomic Hydrogen" *J. Phys. Chem. A* **1999**, *103*, 4893.

ST01 Sicinska, D.; Truhlar, D. G.; Paneth, P. "Solvent-Dependent Transition states for Decarboxylations", *J. Am. Chem. Soc.* **2001**, *123*, 7863.

SP02 Sicinska, D.; Paneth, P.; Truhlar, D. G. "How Well Does Microsolvation Represent Macrosolvation? A Test case: Dynamics of Decarboxylation of 4-Pyridylacetic Acid Zwitterion" *J. Phys. Chem. B* **2002**, *106*, 2708.

AM1-SRP and PM3-SRP models

GO91 Gonzàlez-Lafont, A.; Truong, T. N.; Truhlar D. G. “Direct Dynamics Calculations with Neglect of Diatomic Differential Overlap Molecular Orbital Theory with Specific Reaction Parameters,” *J. Phys. Chem.* **1991**, *95*, 4618.

CE95 Corchado, J. C.; Espinosa-Garcia, J.; Hu, W.-P.; Rossi, I.; Truhlar, D. G. “Dual-Level Reaction-Path Dynamics (The /// Approach to VTST with Semiclassical Tunneling). Application to OH + $NH_3 \rightarrow H_2O + NH_2$,” *J. Phys. Chem.* **1995**, *99*, 687.

CR99 Chuang, Y.-Y.; Radhakrishnan, M. L.; Fast, P. L.; Cramer, C. J.; Truhlar, D. G. “Direct Dynamics for Free Radical Kinetics in Solution: Solvent Effect on the Rate Constant for the Reaction of Methanol with Atomic Hydrogen,” *J. Phys. Chem. A* **1999**, *103*, 4893.

Solubility

TC03 Thompson, J. D.; Cramer, C. J.; Truhlar, D. G. “Predicting Aqueous Solubilities From Aqueous Free Energies of Solvation and Experimental or Calculated Vapor Pressures of Pure Substances,” *J. Chem. Phy*s. **2003**, *119*, 1661-1670.

QM/MM with GHO-AIHF boundary treatment

PG04 Pu, J.; Gao, J.; Truhlar, D. G. “Generalized Hybrid Orbital (GHO) Method for Combining *Ab Initio* Hartree-Fock Wave Functions with Molecular Mechanics,” *J. Phys. Chem. A* **2004**, *108*, 632-650.

Charge response kernel

MK97 Morita, A.; Kato, S. “Ab Initio Molecular Orbital Theory on Intramolecular Charge Polarization: Effect of Hydrogen Abstraction on the Charge Sensitivity of Aromatic and Nonaromatic Species,” *J. Am. Chem. Soc.* **1997**, *119*, 4021-4032.

LY04 Lu, Z.; Yang, W. “Reaction path potential for complex systems derived from combined *ab initio* quantum mechanical and molecular mechanical calculations,” *J. Chem. Phys.* **2004**, *121*, 89-100.

HT08 Higashi, M.; Truhlar, D. G. “Electrostatically Embedded Muticonfiguration Molecular Mechanics Based on the Combined Density Functional and Molecular Mechanical Method,” *J. Chem. Theory Comput.* **2008**, in press.

Quick index to literature

The references for the methods supported by *GAMESSPLUS-v2008* are as follows:

Method	Reference
Löwdin charges	*LD50*
RLPA charges	*TX02*
*m*PW hybrid functional	*AB98, PT02, and LZ03*
MPW1K	*LF00*
MPW*X*	*PT02*
CM2/AM1	*LZ98*
CM2/PM3	*LZ98*
CM2/HF/MIDI!	*LZ98*
CM2/HF/MIDI!6D	*LZ98*
CM2/HF/6-31G*	*LZ98*
CM2/BPW91/MIDI!	*LZ98*
CM2/BPW91/MIDI!6D	*LZ98*
CM2/B3LYP/MIDI!	*LZ98*
CM2/BPW91/6-31G*	*LZ98*
CM2/HF/6-31+G*	*LZ98*
CM2/HF/cc-pVDZ	*LX99*
CM2/BPW91/DZVP	*LZ98*
CM3/AM1	*TC03*
CM3/PM3	*TC03*
CM3/HF/MIDI!	*KC05a*
CM3/HF/MIDI!6D	*WT02*
CM3/HF/6-31G*	*WT02*
CM3/MPW*X*/MIDI!	*WT02*
CM3/MPW*X*/MIDI!6D	*WT02*
CM3/MPW*X*/6-31G*	*WT02*
CM3/MPW*X*/6-31+G*	*WT02*
CM3/MPW*X*/6-31+G**	*WT02*
CM3/BLYP/6-31G*	*TC03*
CM3/B3LYP/MIDI!6D	*TC03*
CM3/B3LYP/6-31G*	*TC03*
CM3/B3LYP/6-31+G*	*TC03*
CM3.1/HF/MIDI!	*KC05a*
CM4/DFT/MIDI!6D	*KC05b*
CM4/DFT/6-31G*	*KC05b*
CM4/DFT/6-31+G*	*KC05b*
CM4/DFT/6-31+G**	*KC05b*
SM5.42/HF/MIDI!	*LZ99*
SM5.42/HF/MIDI!6D	*LH98*
SM5.42/HF/6-31G*	*LZ99*

SM5.42/BPW91/MIDI!	*LZ99*
SM5.42/BPW91/MIDI!6D	*ZL98*
SM5.42/B3LYP/MIDI!	*LZ99*
SM5.42/BPW91/6-31G*	*ZL98*
SM5.42/HF/6-31+G*	*LZ99*
SM5.42/HF/cc-pVDZ	*LZ99*
SM5.42/BPW91/DZVP	*ZL98*
SM5.43/HF/6-31G*	*TC04*
SM5.43/B3LYP/6-31G*	*TC04*
SM5.43/MPW*X*/MIDI!	*TC05*
SM5.43/MPW*X*/MIDI!6D	*TC05*
SM5.43/MPW*X*/6-31G*	*TC05*
SM5.43/MPW*X*/6-31+G*	*TC05*
SM5.43/MPW*X*/6-31+G**	*TC05*
SM6/DFT/MIDI!6D	*KC05b*
SM6/DFT/6-31G*	*KC05b*
SM6/DFT/6-31+G*	*KC05b*
SM6/DFT/6-31+G**	*KC05b*
SM6T/DFT/MIDI!6D	*AC06*
SM6T/DFT/6-31G*	*AC06*
SM6T/DFT/6-31+G*	*AC06*
SM6T/DFT/6-31+G**	*AC06*
SM5.42, SM5.43, and SM6 liquid-phase optimizations	*ZL99* and *CR99*
GHO-AIHF	*PG04*
CRK	*MK97*, *LY04*, and *HT08*

SCRF calculations using the generalized Born method

A self-consistent-reaction-field method based on the generalized Born approximation and either Löwdin charges, RLPA charges, CM2 charges, or CM3 charges have been implemented in the program. One can use the SCRF orbitals to calculate the polarization and distortion energies, solvated atomic partial charges, and to investigate solvation effects on atomic charge distributions, multipole moments, and other one-electron properties. In addition, the SCRF calculation is one of the two components required for the calculation of the free energy of solvation.

Solvent-accessible surface area and solvation free energy in aqueous and organic solutions

The calculation of free energies of solvation by SM*x* models requires that one add the so-called CDS term to the polarization energy and the distortion energy. In particular, the standard-state free energy of solvation is given by eq. (1) where G_{CDS} is the specific contribution from short-range effects associated primarily with the first solvation shell. Calculation of G_{CDS} is based on the Analytic Surface Area (ASA) algorithm. Both SM5.42 and SM5.43 can be used to calculate solvation free energies in aqueous solution and organic-solvent solutions. Currently, SM6 is parametrized for aqueous solvent only.

GHO-AIHF QM/MM Calculations

The GHO-AIHF model in combined QM/MM calculations has been tested for a series of closed-shell and open-shell small molecules and ions with various functional groups close to the QM/MM boundary. The rotation barrier around the central C-C bond in *n*-butane has been studied by using GHO-AIHF. The proton affinities of small alcohols, amines, thiols, and acids computed by GHO-AIHF showed that the method is also reliable for energetics. In those tests, various basis sets were used for the QM part, namely: STO-3G, 6-31G(d), 6-31+G(d), 6-31+G(d,p), 6-31++G(d,p), and MIDI!.

As compared to a projected basis scheme and a scheme based on neglect of diatomic differential overlap involving auxiliary orbitals, using hybrid orbitals based on global Löwdin orthogonalized atomic orbitals is more robust. It has been shown that only the non-orthogonality of atoms near the boundary is important, and needs to be removed by the explicit orthogonalization scheme in GHO; therefore only a local orthogonalization is necessary. Considering the localization of the boundary treatment, this method is more theoretically promising. Therefore a local Löwdin orthogonalization algorithm has also been implemented. Instead of doing a Löwdin orthogonalization over the entire QM system, only orbitals on: GHO boundary atoms, QM frontier atom A, and QM atoms directly bonded to A (these atoms are also called geminal atoms) are orthogonalized to each other before the hybridization. By using this local Löwdin orthogonalization method, the mixing of tails from other QM atoms far from the boundary is eliminated and the perturbation introduced to the QM subsystem is minimized.

Although the unparametrized GHO-AIHF method gives reasonable optimized geometries and charges, one can obtain even better results by scaling the integrals involving the boundary orbitals. Such a parametrized version of GHO-AIHF (based on local Löwdin orthogonalization) is available for the MIDI! basis set, in which the scaling factors are obtained from a small training set containing propane, propanol, propanoic acid, *n*-butane, and 1-butene .

Usage

This section highlights important *GAMESS* input options and describes the input options in the namelists $CM2, $GMSOL, $CM2SRP, and $NDDOSRP. Namelists $CM2 and $GMSOL are used to define the specifics of calculations involving CM2, CM3, CM4, SM5.42, SM5.43, SM6, and SM6T. Note that $GMSOL and $CM2 define the same input, and only one of them should be defined in a given input file. If they are both defined in the same input file, the input parameters defined in the second namelist will be ignored. Namelist $CM2SRP is used for reading the external parameters for CM2-SRP methods, and namelist $NDDOSRP is used for reading the external parameters for NDDO-SRP methods.

Note: For GHO-AIHF calculations carried by the *CHARMM*/*GAMESSPLUS* combination package, and the CHARMM input file controls the GHO options. A detailed description is available in documentation for the *CGPLUS* package (http://comp.chem.umn.edu/cgplus, see "*CGPLUS-v1.0 User Manual*")

Notes on *GAMESS* Input

Users of *GAMESSPLUS* should consult at least the first two sections of the *GAMESS Users Manual* that is distributed with the *GAMESS* package. These sections outline the specifics for all aspects of *GAMESS* input.

Most *GAMESS* input is entered in a pseudo-namelist format; where a namelist is a particular group of keywords. The title of this group, the namelist name, is prefixed by a '$' that must begin in the second column of an input file (e.g. $SCF). A namelist accepts as arguments a number of possible keywords that are entered after the namelist name in a free format style that can span over multiple lines. Each keyword accepts either character or numeric values. A namelist is terminated with a $END. Only input between a namelist name and its corresponding $END will be read from a *GAMESS* input file. *GAMESSPLUS* introduces five new namelists to *GAMESS*: $CM2, $GMSOL ($CM2 and $GMSOL define the same input options), $CM2SRP, $NDDOSRP, and $EEQM; these namelists are discussed at length in separate subsections below.

Below is a brief summary of *GAMESS* namelists and their keywords that are most important to *GAMESSPLUS* users. In this section, the namelist name is followed by a brief description, and below the namelist name is a set pertinent keywords and their options:

Namelist $CONTRL			group of global settings for the calculation
SCFTYP	=	RHF	restricted Hartree-Fock calculation (default)
	=	UHF	unrestricted Hartree-Fock calculation
RUNTYP	=	ENERGY	SCF evaluation (default)
	=	GRADIENT	gradient evaluation
	=	HESSIAN	Hessian and possibly vibrational frequency evaluation
	=	OPTIMIZE	geometry optimization to minima
	=	SADPOINT	geometry optimization to saddle points
	=	EEQM	EEQM calculation with a site-site representation of the QM/MM electrostatic interaction
MAXIT	=	*N*	*N* is the maximum number of SCF cycles (default = 30)
ICHARG	=	*N*	*N* is the molecular charge (default = 0)
MULT	=	*N*	*N* is the multiplicity of the electronic state (default = 1)
COORD	=	UNIQUE	input symmetry-unique Cartesian coordinates (default)
	=	CART	input all Cartesian coordinates
	=	ZMT	input *Gaussian*-style *Z*-matrix internal coordinates
ISPHER	=	–1	use Cartesian basis functions (e.g. 6D, 10F; default)
	=	1	use spherical harmonic basis functions (e.g. 5D, 7F)

Note: Symmetry should not be used with *GAMESSPLUS* calculations. The best way to input Cartesian coordinates is to set COORD = UNIQUE (prevents coordinate rotation) and to set the symmetry group of the molecule to C1 in $DATA (see below).

Namelist $SYSTEM		information for controlling the computer's operation
MEMORY=	*N*	*N* is the maximum memory the job can use in words (default = 1 000 000)

Namelist $DFT		density functional theory (DFT) input
DFTTYP=	*N*	*N* is the density functional or hybrid density functional. A full list of the density functionals available in *GAMESSPLUS* that are suggested for use with CM4 and SM6 are given in the section entitled "*Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS.*
HFE=	*N*	Defines the fraction of Hartree-Fock exchange *N* to be used when the MPW*X* functional is used. This keyword must be specified when MPW*X* is used. For CM3 and SM5.43 calculations that use MPW*X* (ICMD = 315 – 319), the HFE = *N* must also be specified in the $GMSOL or $CM2 namelist. For CM4 and SM6 (ICMD = 416 – 419), HFE = *N* must always be specified in the $GMSOL or $CM2 Namelist.
METHOD=	*M*	*M* is GRID to request grid-based DFT calculation (default)

<u>Note:</u> **All CM2, CM3, CM4, SM5.42, SM5.43, and SM6 parameterizations are based on the B3LYP method that uses version III of the VWN correlation functional, so these types of calculations must be carried out using 'DFTTYP=B3LYP3', not 'DFTTYP=B3LYP5'. In addition, the charge and solvation models that are based on DFT were parameterized using grid-based DFT, so 'METHOD=GRID' (the default for METHOD) should always be used.**

Namelist $BASIS			input of available standard basis sets
GBASIS	=	N31	Pople's N-31G basis sets
	=	AM1	AM1 model Hamiltonian
	=	PM3	PM3 model Hamiltonian
NGAUSS	=	*N*	the 'N' of N31 *(*e.g. *N* = 6 for 6-31G)
NDFUNC	=	*N*	*N* is the number of polarizing *d* subshells on heavy atoms (*N* = 1 for 6-31G*; default = 0)
DIFFSP	=	.TRUE.	adds a diffuse *sp* shell to the basis set (default = .FALSE.)

<u>Note:</u> The *d* subshells have 5 functions if ISPHER=1, and they have 6 functions if ISPHER= -1, where ISPHER is defined in the $CONTRL data group. Thus, when using the MIDI! basis set, ISPHER=1 should be specified in the $CONTRL namelist. When using the MIDI!6D basis set, ISPHER=-1 should be specified in the $CONTRL namelist.

Examples:

6-31G* : `$BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END`

6-31+G* : `$BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 DIFFSP=.TRUE. $END`

Namelist $DATA input molecule information

Input consists of the following information:

line 1: title line

line 2: symmetry group (always enter `C1` for *GAMESSPLUS* calculations)

line 3+: molecular coordinates; for Cartesian input, each line consists of the following: atom label, the atom's nuclear charge (which should be a floating-point value, i.e., 1.0 for H), and the atom's *x*, *y*, and *z*-coordinate. If a general basis set is required, then the basis set for each atom follows the atom's coordinates. See examples in the *Input and Output Examples* section and in the test suite.

Namelist $STATPT input information for geometry optimization

NSTEP	=	*N*	*N* is the maximum number of steps (default = 20)
HESS	=	GUESS	guess positive-definite Hessian (default for RUNTYP = OPTIMIZE)
	=	READ	read Hessian from $HESS namelist (default for RUNTYP=SADPOINT)
	=	CALC	calculate the Hessian
HESSEND=		.TRUE.	calculate the Hessian at the end of a successful optimization (default = .FALSE.)

Namelist $FORCE controls Hessian and vibrational frequency evaluation

METHOD	=	ANALYTIC	analytical second derivative (default where available)
	=	NUMERIC	numeric second derivative based on analytical first derivatives (required for liquid-phase Hessian evaluations)
NVIB	=	1	forward-difference for numeric Hessian evaluation(default)
	=	2	central-difference for numeric Hessian evaluation
VIBSIZ	=	*N*	*N* is the displacement size (in Bohr; default = 0.01)
VIBANL	=	.TRUE.	print vibrational analysis (default = .TRUE. for RUNTYP = HESSIAN, = .FALSE. otherwise)

Namelists $GMSOL and $CM2

Namelists $GMSOL and $CM2 control most of the options for the solvation and charge models. These two namelists define the same input, and only one of them should be defined in a given input file. If they are both defined in the same input file, the parameters defined in the second namelist will be ignored. Below are descriptions of all input variables defined in $GMSOL and $CM2.

ISCRF	Controls the type of calculation to be performed:
ISCRF=0	Invokes a gas-phase calculation of Löwdin, RLPA, CM2, CM3, or CM4 charges (default)
ISCRF=1	Invokes a gas- and liquid-phase solution calculation of Löwdin, RLPA, CM2, CM3, or CM4 charges and the calculation of solvation free energy using SCF Scheme I (see SCF Schemes). This method is only available for methods that use nondiffuse basis functions.
ISCRF=2	Invokes a gas- and liquid-phase solution calculation of Löwdin or RLPA charges and CM2, CM3, or CM4 charges and the calculation of solvation free energy using SCF Scheme II (see SCF Schemes). This is the default method for methods that use diffuse basis sets.

Note that when a calculation of RLPA charges is requested, Löwdin charges are also calculated. However, for liquid-phase calculations that use RLPA charges, the Löwdin charges correspond to the converged SCRF achieved with RLPA charges, not with Löwdin charges. Also note that RLPA charges should only be calculated by methods that use the 6-31+G* or 6-31+G** basis sets, while Löwdin charges should be used for all other methods.

ICMD, ICDS Select the set of coefficients to use to evaluate the CM2, CM3, or CM4 charges and the atomic surface tension parameters, respectively. For CM2 and CM3 and SM5.42 and SM5.43 solvation models based on these charges, coefficients have been optimized for specific wave functions. For CM4 and SM6 solvation models based on these charges, coefficients have been optimized for basis sets, and can be used with any good density functional method. A list of density functional methods that are supported by GAMESSPLUS and that are recommended for use with CM4 and SM6 are listed in the section entitled "*Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS*".

ICDS	ICMD	Wave function which corresponds to set of coefficients
0	0	All coefficients equal zero, calculate Löwdin charges
1	1	CM2/HF/MIDI! (default for ICMD)
2	2	CM2/HF/MIDI!6D
3	3	CM2/HF/6-31G*
4	4	CM2/BPW91/MIDI!
5	5	CM2/BPW91/MIDI!6D
6	6	CM2/B3LYP/MIDI!
7	7	CM2/BPW91/6-31G*
8	8	CM2/HF/6-31+G*
9	9	CM2/HF/cc-pVDZ
11	11	CM2/AM1
12	12	CM2/PM3
0	301	CM3/HF/MIDI!
0	302	CM3/HF/MIDI!6D
303	303	CM3/HF/6-31G*
0	311	CM3/AM1
0	312	CM3/PM3
313	313	CM3/B3LYP/6-31G*
0	314	CM3/B3LYP/6-31+G*
315	315	CM3/MPW*X*/MIDI!
316	316	CM3/MPW*X*/MIDI!6D
317	317	CM3/MPW*X*/6-31G*
318	318	CM3/MPW*X*/6-31+G*, uses RLPA charges
319	319	CM3/MPW*X*/6-31+G**, uses RLPA charges
0	320	CM3/B3LYP/MIDI!6D
0	321	CM3/BLYP/6-31G*
0	322	CM3.1/HF/MIDI!
416	416	CM4/DFT/MIDI!6D
417	417	CM4/DFT/6-31G*
418	418	CM4/DFT/6-31+G*, uses RLPA charges
419	419	CM4/DFT/6-31+G**, uses RLPA charges
0	300	Calculate RLPA charges

In the above table, *X* corresponds to the percentage of Hartree-Fock exchange used in the mPW hybrid density functional (see the HFE keyword for more details). The B3LYP method in the table above corresponds to the one requested by the ‘DFTTYP=B3LYP3’ keyword in the $DFT data group; see the section entitled *Notes on GAMESS input* above. For SM6, DFT refers to any good density functional. A list of recommended density functionals that are supported in *GAMESSPLUS* is given in the section entitled “*Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS*”. In cases where the default value of ICDS = 0, a solvation model does not exist for that particular set of charges. For ICMD = 300, all restricted and unrestricted HF, DFT, and hybrid DFT methods are available, but only the 6-31+G* or the 6-31+G** basis set should be used. Note that CM3 for the BLYP and B3LYP methods and CM3.1 for the HF/MIDI! method in the above table uses a different mapping function for compounds that contain N and O than CM3 for MPW*X* and HF does. For more information, see “Parameterization of Charge Model 3 For AM1, PM3, BLYP, and B3LYP” by Thompson, J. D.;

Cramer, C. J.; Truhlar, D. G. *J. Comput. Chem*., **2003**, *24,* 1291 and "Accurate Partial Atomic Charges for High-Energy Molecules with the MIDI! Basis Set" by Kelly, C. P.; Cramer, C. J; Truhlar, D. G. *Theor. Chem. Acc.* **2005**, *113*, 133.

HFE Defines the fraction of Hartree-Fock (HF) exchange to be used when ICMD = 315 to ICMD = 319 or ICMD = 416 to ICMD = 419. For example, if a calculation of CM3 charges by the mPW1PW91/MIDI! method were to be carried out, then ICMD would be set to 315, and HFE would be set to 0.25. The HFE keyword only defines the fraction of HF exchange for a given calculation in the charge and solvation model portion of *GAMESSPLUS*, not in the entire electronic structure part of the *GAMESSPLUS* program (thus, when using the MPW*X* functional, the HFE keyword must also be specified in the $DFT namelist). This keyword is only active when ICMD = 315 – 319 or ICMD = 416 – 419, and in these instances HFE must be explicitly specified by the user in both the $DFT and the $CM2 or $GMSPLUS namelists. If ICMD = 315 – 319 or ICMD = 416 – 419 and HFE is not defined by the user, *GAMESSPLUS* will terminate.

IRADII Determines the atomic radii that will be used during the SCRF calculation:

IRADII=1 Use Bondi's values for the van der Waal's radii to build the molecular cavity, (2.0 Å is used in cases where the atomic radii has not been defined by Bondi). For evaluating the SASA, this is always the default set of radii. For other values of IRADII, these radii are used in cases where the radius of a given atom is not defined for a particular model (see the table below).

IRADII=2 Use the atomic radii optimized for predicting solvation free energies with the SM5.42 solvation model to build the molecular cavity to be used in the SCRF calculation. This is the default for ICMD = 1 to ICMD = 10.

IRADII=3 Use the atomic radii optimized for predicting solvation free energies with the SM5.43 solvation model to build the molecular cavity to be used in the SCRF calculation. This is the default for ICMD = 301 to ICMD = 322.

IRADII=4 Use the atomic radii radii optimized for predicting solvation free energies with the SM6 solvation model to build the molecular cavity to be used in the SCRF calculation. This is the default for ICMD = 0, ICMD = 300, and ICMD = 416 to ICMD = 419.

The four sets of radii described above are listed below. In cases where the atomic radius of a given atom is not defined, the default value is Bondi's radii. For elements whose atomic radii are not listed in Bondi's paper, a value of 2.0 Å is used for the radius. Thus, Generalized Born calculations may be carried out for a molecule containing any element on the periodic table.

Element[a]	IRADII = 1[b]	IRADII = 2[c,d]	IRADII = 3[c,e]	IRADII = 4[c,f]
H	1.20	0.91	0.79	1.02
Li	1.82	1.32	—	—
C	1.70	1.78	1.81	1.57
N	1.55	1.92	1.66	1.61
O	1.52	1.60	1.63	1.52
F	1.47	1.50	1.58	1.47
Ne	1.54	—	—	—
Na	2.27	—	—	—
Mg	1.73	—	—	—
Al	2.50	—	—	—
Si	2.10	—	—	—
P	1.80	2.27	2.01	1.80
S	1.80	1.98	2.22	2.12
Cl	1.75	2.13	2.28	2.02
Ar	1.88	—	—	—
K	2.75	—	—	—
Ni	1.63	—	—	—
Cu	1.40	—	—	—
Zn	1.39	—	—	—
Ga	2.40	—	—	—
Ge	2.10	—	—	—
As	1.85	—	—	—
Se	1.90	—	—	—
Br	1.85	2.31	2.38	2.60
Kr	2.02	—	—	—
Pd	1.63	—	—	—
Ag	1.72	—	—	—
Cd	1.58	—	—	—
In	2.50	—	—	—
Sn	2.20	—	—	—
Sb	2.10	—	—	—
Te	2.06	—	—	—
I	1.98	2.66	—	—
Xe	2.60	—	—	—

[a]For atoms not listed in this table, a radius of 2.0 Å is used. [b]Bondi's values for the van der Waals radius (Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441). [c]In cases where the atomic radius is not defined, Bondi's values are used. In cases where the atomic radius has not been defined by Bondi, 2.0 Å is used for the atomic radius. [d]These radii have been optimized for use with the SM5.42 model. [e]These radii have been optimized for use with the SM5.43 model. [f]These radii have been optimized for use with the SM6 model.

ISTS Determines the type of surface tension functionals:

ISTS=5 Use SM5 functional forms to evaluate G_{CDS} . This is the default for all values of ICMD between 0 and 322.

ISTS=6 Use SM6 functional forms to evaluate G_{CDS} . This is the default for all values of ICMD between 416-419.

SolvRd Value of the solvent radius (in angstroms) for the calculation of the solvent-accessible surface areas of the atoms of the solute (The default is 0 Å for SM5.42 and 0.40 Å for SM5.43 and SM6). For predicting solvation free energies with SM5.42, SolvRd should be set to 0 Å. For predicting solvation free energies with SM5.43 or SM6, the default value of 0.40 Å should be used.

IAQU Determines the solvent type:

IAQU=0 Organic solvent (additional solvent data must be input, see below)

IAQU=1 Aqueous solvent (default, no additional solvent data is required)

Solvent Descriptors: If IAQU=0, then the solvent properties are specified by a series of solvent property (descriptor) values:

Dielec dielectric constant, ε, of solvent (default is the value for water, 78.3)

SolN index of refraction at optical frequencies at 293 K, n_{20}^{D} (default is 1.0)

SolA Abraham's hydrogen bond acidity, $\Sigma\alpha_2^H$ (default is 0.0)

SolB Abraham's hydrogen bond basicity, $\Sigma\beta_2^H$ (default is 0.0)

SolG $\gamma = \gamma_m / \gamma^{\mathrm{o}}$ (default is 0.0), where γ_m is the macroscopic surface tension at air/solvent interface at 298 K, and γ^{o} is 1 cal·mol^{-1}·Å^{-2} (note that 1 dyne/cm = 1.43932 cal·mol^{-1}·Å^{-2})

SolC aromaticity, ϕ : the fraction of non-hydrogenic solvent atoms that are aromatic carbon atoms (default is 0.0)

SolH electronegative halogenicity, ψ : the fraction of non-hydrogenic solvent atoms that are F, Cl or Br (default is 0.0)

For a desired solvent, these values can be derived from experiment or from interpolation or extrapolation of data available for other solvents. Solvent parameters for common organic solvents are tabulated in the *Minnesota Solvent Descriptor Database*. The latest version of this database is available at: http://comp/chem.umn.edu/solvation.

Solvent Temperature: If IAQU=1, then the solvent temperature may be specified using one of the following keywords:

SOLK Specifies that the free energy of solvation for the solute be computed at a single temperature. The current model is only applicable to H, C and O containing compounds for the range of temperatures in which water is liquid (273 to 373 K). (Note the model uses a temperature-dependent dielectric constant, the equation for it was obtained from the *CRC Handbook of Chemistry and Physics 76th edition*, ed. Lide, D. R., 1995, CRC Press, New York.) The default value is SOLK = 298.

READK Causes *GamessPlus* to read in a list of temperatures at which to compute free energies of solvation from a file labeled KELT. The default value is READK=FALSE. KELT must be located in the scratch directory.

Example KELT file:

```
3
273
298
373
```

The first line in the input file must list the number of temperatures at which the free energy of solvation is to be computed. Then each succeeding line should list only one temperature.

AVGK This option is similar to READK however instead of computing the electronic energy of the solute in solution at each temperature; it computes the electronic energy of the solute at the average temperature and scales the electronic energy of the solute using the following factor

$$\frac{\left(1-\frac{1}{\varepsilon_T}\right)}{\left(1-\frac{1}{\varepsilon_{Avg}}\right)}$$

where ε_{Avg} is the dielectric constant of water at the average temperature and ε_T is the dielectric constant of water at a given temperature, *T*. This corresponds to assuming that the charge on the solute does not change significantly as a function of temperature. This approach can compute a large number of temperatures with nearly the same computation time as that of a single temperature. This approach does deviate mildly from the individual computations performed at each temperature; however, the deviation is negligable. The default value is AVGK=FALSE.

IGEOM Determines whether the inputted structure is already optimized in the gas-phase.

IGEOM=0 Perform a geometry optimization in the gas-phase and use the optimized gas-phase structure as the initial guess for the liquid-phase optimization (default)

IGEOM=1 Use the inputted structure as the initial point for the liquid-phase geometry optimization (i.e., do not perform a gas-phase optimization on the inputted structure).

IGAS Determines which gas-phase energy is used in the evaluation of ΔG_S^o.

IGAS=0 Use the gas-phase energy of either the inputted geometry (IGEOM = 1) or of the optimized gas-phase structure (IGEOM = 0; IGAS = 0 is the default). Once this gas-phase energy is obtained, the gas-phase SCF is skipped, i.e., the liquid-phase SCRF begins at first SCF cycle.

IGAS=1	Gas-phase energy is updated for every new geometry.
IGAS=2	Gas-phase energy is read in using the **ETGAS** keyword.
IGAS=3	Gas-phase energy is read in using the **ETGAS** keyword, and gas-phase SCF is always skipped; i.e., the liquid phase SCRF begins at first SCF cycle. This option is not available when ISCRF=2.

ETGAS Accepts an inputted value for the gas-phase energy (in hartrees).

The ETGAS keyword must be defined for IGAS = 2 or 3; the calculation will stop otherwise. Note that if IGAS = 0 or 1, an inputted value of ETGAS will be ignored. Theoretically, the gas-phase SCF energy used in the evaluation of the free energy of solvation is that of the gas-phase equilibrium geometry. Inputting an energy for a non-equilibrium structure or using the IGAS=1 option will lead to theoretically questionable results. This option is not available when ISCRF=2.

ICREAD	Controls the reading of atomic charges from an external file
ICREAD=0	Do not read in atomic charges (default).
ICREAD=1	Read in atomic charges from file "CM2CHG" (case-sensitive).

ICSAVE	Controls the saving of the evaluated CM2, CM3, or CM4 charges to an external file.
ICSAVE=0	Do not save the CM2 or CM3 charges (default).
ICSAVE=1	Save the evaluated CM2 or CM3 charges in file "CM2CHG" (case-sensitive).

ISREAD	Controls the reading of surface tension parameters from an external file
ISREAD=0	Use the surface tension coefficients provided in the code (default).
ISREAD=1	Read and use the surface tension coefficients provided in a file called "CDS_Param" (case-sensitive), and use the surface tension coefficients provided in the code for any surface tension coefficient that is not provided by the user in the file "CDS_Param".
ISREAD=2	Read and use the surface tension coefficients provided in a file called "CDS_Param", and set all other surface tension coefficients to zero.

Note that the CM2CHG and CDS_Param files must be located in the directory defined to be the scratch directory for the calculation (i.e., the directory defined by $SCR in the rungms script, see the section entitled *Notes on Running GAMESSPLUS* below for a description of the rungms script).

User-provided surface tension coefficients are entered into the input file in the following fashion: the first line of file contains the total number of surface tension coefficients that the user has provided in the file, followed by the value of *ICoeff*, where:

***ICoeff* = 1:** use the surface tension coefficients provided in the code for any surface tension coefficients that are not provided by the user in this input file

***ICoeff* = 2:** set all surface tension coefficients that have not been provided to zero

All coefficient types have been assigned a four digit coefficient label (see the table below); the user provides a coefficient label followed by its corresponding value (in cal/Å^2) on each line following the first line until all data is entered.

Note that only one parameter convention can be entered: either a set of solvent descriptors for water or a set of solvent descriptor-dependent coefficients can be entered, but not a mix of both (i.e., coefficient labels must be either all less than 2000 or all greater than 2001).

one-, two-, and three-atom surface tension coefficients:

$\overline{\sigma}_k^A$	$\ddot{\varpi}_i^{(n)}$	$\ddot{\varpi}_i^{(\alpha)}$	$\ddot{\varpi}_i^{(\beta)}$	coefficient type
1000 + *i*, *i* = 1 – 100	2000 + *i*, *i* = 1 – 100	3000 + *i*, *i* = 1 – 100	4000 + *i*, *i* = 1 – 100	atomic coefficient for atoms with nuclear charge *i*
1100 + *j*, *j* = 1 – 100	2100 + *j*, *j* = 1 – 100	3100 + *j*, *j* = 1 – 100	4100 + *j*, *j* = 1 – 100	H-X, where *j* is the nuclear charge on atom X
1201	2201	3201	4201	C-C(1)
1202	2202	3202	4202	C-C(2)
1203	2203	3203	4203	O-C
1204	2204	3204	4204	O-O
1205	2205	3205	4205	N-C(1)
1206	2206	3206	4206	O-N
1207	2207	3207	4207	S-S
1210	2210	3210	4210	C-N
1211	2211	3211	4211	N-C(2)
1212	2212	3212	4212	H-N(2)
1213	2213	3213	4213	H-O(2)
1214	2214	3214	4214	O-P
1215	2215	3215	4215	S-P
1216	2216	3216	4216	N-C(3)

molecular surface tension coefficients:

5001:	$\tilde{\sigma}^{(\gamma)}$
5002:	$\tilde{\sigma}^{(\beta^2)}$
5003:	$\tilde{\sigma}^{(\phi^2)}$
5004:	$\tilde{\sigma}^{(\psi^2)}$

Example:

The following input defines the values of the following four coefficients: $\tilde{\sigma}_{\mathrm{H}}^{(n)}$ = 40.80, $\tilde{\sigma}_{\mathrm{C}}^{(\alpha)}$ = 24.51, $\tilde{\sigma}_{\mathrm{H,O}}^{(\beta)}$ = -346.42, and $\tilde{\sigma}^{(\gamma)}$ = 0.3136. All other coefficients are set to zero (*ICoeff* = 2)

```
4 2
2001    40.80D0
3006    24.51D0
4108  -346.42D0
5001   0.3136D0
```

Namelist $CM2SRP

The current version of *GAMESSPLUS* provides an option for entering external parameters for the CM2 charge model, and this is called CM2-SRP. This option is required when the electronic wave function used for a calculation does not have a matched CM2, CM3, or CM4 parameter set (for example, HF/STO-3G), or when it is desired to 'adjust' the parameters in the charge model to obtain a better estimation of the dipole moment. To carry out a CM2-SRP, CM3-SRP, or CM4-SRP calculation, the namelist $CM2SRP is required. In this namelist, there is only one variable, NAME; this variable is used to indicate the name of a CM2-SRP file, which has a maximum of 8 characters. Note the CM2-SRP file must be located in the scratch directory used throughout a given calculation. The scratch directory used for a particular calculation is defined with the $SCR variable in the rungms script (see the section entitled *Notes on Running GAMESSPLUS* below).

In order to describe the syntax for the input file defined in $CM2SRP, it will be useful to describe the CM2 charge mapping scheme. The CM2, CM3, or CM4 charge, q_k, on an atom *k* is given by

$$q_k = q_k^0 + \sum_k B_{kk'}\left(D_{kk'} + B_{kk'}C_{kk'}\right) \tag{17}$$

where q_k^0 is the Löwdin or Redistributed Löwdin charge on atom *k*, $B_{kk'}$ is the Mayer bond order between atoms *k* and *k*´, and $D_{kk'}$ and $C_{kk'}$ are the CM2 parameters. It is the $D_{kk'}$ and $C_{kk'}$ that may be defined in the CM2-SRP file. Thus the syntax of a CM2-SRP file is:

<Variable type> <Atomic pair no.> <Value>

<Variable type> indicates if the parameter is a *C* parameter or a *D* parameter (acceptable input is C or D), according to equation (17) above, and <Atomic pair no.> corresponds to a particular pair

of atoms *k* and k′ in equation (17). Allowed values of <Atomic pair no.> and the corresponding pair of atoms defined for the value of <Atomic pair no.> are given below:

<Atomic pair no.>	Description
1	H–C
2	H–N
3	H–O
4	H–Si
5	H–P
6	H–S
7	C–N
8	C–O
9	C–F
10	C–Si
11	C–P
12	C–S
13	C–Cl
14	C–Br
15	C–I
16	O–P
17	F–P
18	N–O
19	O–S
20	P–S

Example input:

```
C          1          –0.0200
C          3           0.0149
D          8          –0.0874
D          8           0.0215
```

Namelist $NDDOSRP

Although the AM1 and PM3 methods are parameterized using broad sets of stable molecules, these methods sometimes do not perform equally well for molecules outside the AM1 or PM3 training set, for radical species, or for transition states. Furthermore, sometimes one does not want the parameters that give the best results on the average (i.e., the standard parameters); rather one wants nonstandard parameters that perform the best for a specific reaction or for a limited range of systems. In order to 'fix' the energetic information for a chemical dynamics calculation on a specific reaction or a set of calculations on a specific range of systems, the specific reaction parameter (NDDO-SRP) method can be used. To carry out an NDDO-SRP calculation, the namelist $NDDOSRP is required. In this namelist, there is only one variable, NAME, which is used to indicate the name (with maximum of 8 characters) of an NDDO-SRP file. Note that the NDDO-SRP file must be located in the scratch directory defined for the given calculation. The scratch directory used for a particular calculation is defined with the $SCR variable in the rungms script (see the section entitled *Notes on Running GAMESSPLUS* below).

The syntax of the NDDO-SRP file is a series of lines, each of which having one of two possible forms. Form A is

<Variable type> <Atomic no.> <Value>

where <Variable type> indicates the type of the parameter, and <Atomic no.> indicates the atomic number. Form A may be used for changes in the one-electron, one-center energies (USS or UPP), in the monatomic parameter for one-electron resonance integrals (BETAS or BETAP), and in orbital exponents (ZS or ZP). For example:

USS	6	–49.850
UPP	6	–40.337
USS	8	–99.181
UPP	8	–80.762
BETAS	6	–16.912
BETAP	6	–9.190
BETAS	8	–28.998
BETAP	8	–29.249

Form B is used for modifying the two-center resonance integrals in the NDDO-SRP method without using the arithmetic mean prescription (see last section of *Executive summary*). In form B, each line has the syntax:

BET*xy* <atom which *x* orbital taken from> <atom which *y* orbital taken from> <value>

where BET*xy* is BETSS, BETSP, or BETPP (i.e., *x* and *y* denote orbital types; only S and P type are available).
For example:

BETSP 1 8 -17.711

Note that BETSP 6 8 is different than BETSP 8 6.

GAMESSPLUS keywords Required for Running Standard SMx Calculations

Shown in the below table are the *GAMESSPLUS* keywords required for running standard SM5.42, SM5.43, and SM6 calculations. Note that the keywords entered in the $GMSOL or $CM2 namelist are used by the solvation and charge model portion of *GAMESSPLUS*. Thus, keywords used by the rest of the *GAMESSPLUS* program, such as the level of theory, basis set, scf and geometry optimization options, etc., must also be included in the route section of the input file. For convenience, the ISPHR keyword, which is specified in the $CONTRL namelist, is also included in the table below.

Method	ISPHR[a]	ISTS	ICMD	ICDS	IRADII	SolvRd	HFE
SM5.42/HF/MIDI! [b]	1	5	1	1	2	0.00	N/A
SM5.42/HF/MIDI!6D	-1	5	2	2	2	0.00	N/A
SM5.42/HF/6-31G*		5	3	3	2	0.00	N/A
SM5.42/HF/6-31+G*		5	8	8	2	0.00	N/A
SM5.42/HF/cc-pVDZ [b]		5	9	9	2	0.00	N/A
SM5.42/BPW91/MIDI! [b]	1	5	4	4	2	0.00	N/A
SM5.42/BPW91/MIDI!6D	-1	5	5	5	2	0.00	N/A
SM5.42/BPW91/6-31G*		5	7	7	2	0.00	N/A
SM5.42/BPW91/DZVP		5	10	10	2	0.00	N/A
SM5.42/B3LYP/MIDI! [b]	1	5	6	6	2	0.00	N/A
SM5.43/HF/6-31G*		5	303	303	3	0.40	N/A
SM5.43/B3LYP/6-31G*		5	313	313	3	0.40	N/A
SM5.43/MPW*X*/MIDI! [b]	1	5	315	315	3	0.40	*X*
SM5.43/MPW*X*/MIDI!6D	-1	5	316	316	3	0.40	*X*
SM5.43/MPW*X*/6-31G*		5	317	317	3	0.40	*X*
SM5.43/MPW*X*/6-31+G*		5	318	318	3	0.40	*X*
SM5.43/MPW*X*/6-31+G**		5	319	319	3	0.40	*X*
SM6/DFT/MIDI!6D	-1	6	416	416	4	0.40	*c*
SM6/DFT/6-31G*		6	417	417	4	0.40	*c*
SM6/DFT/6-31+G*		6	418	418	4	0.40	*c*
SM6/DFT/6-31+G**		6	419	419	4	0.40	*c*

In the above table, *X* corresponds to the percentage of Hartree-Fock exchange used in the mPW hybrid density functional (see the HFE keyword for more details). For SM6, DFT refers to any good density functional. A list of recommended density functionals that are supported in *GAMESSPLUS* is given in the section entitled "*Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS*". Also given in this section are the HFE values required by all of the recommended density functionals.

[a] This keyword is specified in the $CONTRL namelist.

[b] Analytic gradients not available for this method.

[c] When using any of the CM4 or SM6 methods, the HFE keyword must always be specified (HFE = 0.00 should be entered when using pure DFT functionals). For more details, see the section entitled "*Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS*".

Namelist $EEQM

Namelist $EEQM controls EEQM calculations with a site-site representation of the QM/MM electrostatic interaction. It is required that RUNTYP=EEQM in namelist $CONTRL to carry out an EEQM calculation. Currently *GAMESSPLUS* doesn't support restricted open-shell Hartree Fock and restricted open-shell Kohn-Sham calculations when the EEQM option is selected.

EETYP Determines the type of EEQM calculation, $V^{\text{EEQM}}\left(\mathbf{R},\boldsymbol{\Phi}\right)=\left\langle\Psi\middle|\hat{H}_0+\hat{\mathbf{Q}}^{\text{T}}\boldsymbol{\Phi}\middle|\Psi\right\rangle$:

EETYP=ENERGY Calculate energy V^{EEQM} and partial charges $\mathbf{Q}$ in the presence of a given electrostatic potential distribution. (default)

EETYP=GRADIENT Calculate energy V^{EEQM} and partial charges $\mathbf{Q}$ plus the gradient $\dfrac{\partial V^{\text{EEQM}}}{\partial \mathbf{R}}$ in the presence of a given electrostatic potential distribution.

EETYP=HESSIAN Calculate energy V^{EEQM} and partial charges $\mathbf{Q}$ plus the gradient $\dfrac{\partial V^{\text{EEQM}}}{\partial \mathbf{R}}$ plus the Hessian $\dfrac{\partial^2 V^{\text{EEQM}}}{\partial \mathbf{R}^2}$ in the presence of a given electrostatic potential distribution.

EETYP=DQDR Calculate energy V^{EEQM} and partial charges $\mathbf{Q}$ plus the derivatives of the charges with respect to coordinates $\dfrac{\partial \mathbf{Q}}{\partial \mathbf{R}}$ in the presence of a given electrostatic potential distribution. See eq. (22).

EETYP=DQDPHI Calculate energy V^{EEQM} and partial charges $\mathbf{Q}$ plus the derivatives of the charges with respect to electrostatic potentials $\dfrac{\partial \mathbf{Q}}{\partial \boldsymbol{\Phi}}$ in the presence of a given electrostatic potential distribution. See eq. (23).

PHI Electrostatic potential distribution, that is, array of electrostatic potentials Φ_k on the atoms k. The units used for PHI are controlled by IUPHI. The default is $\boldsymbol{\Phi}=\mathbf{0}$, which corresponds to a gas-phase calculation.

IUPHI Determines the units of the electrostatic potentials PHI:

IUPHI=0 PHI is in a.u.

IUPHI=1 PHI is in eV. (default)

ICMD Select the set of coefficients to use to evaluate the CM2, CM3, or CM4 charges. This option is the same as ICMD in namelists $GMSOL and $CM2. Two additional options are allowed:

ICMD=998 Use the set of coefficients read in CCMSRP and DCMSRP. RLPA is not used.

ICMD=999 Use the set of coefficients read in CCMSRP, DCMSRP, and ZCMSRP. RLPA is used.

HFE Defines the fraction of Hartree-Fock (HF) exchange to be used when ICMD = 315 to ICMD = 319 or ICMD = 416 to ICMD = 419. This option is the same as HFE in namelists $GMSOL and $CM2.

ZCMSRP Determines the parameters used in the RLPA charge calculation when ICMD=999. The RLPA charge is given by

$$Q_a^0\left(\mathrm{RLPA}\right) = Q_a^0\left(\mathrm{LPA}\right) + Z_a Y_a \sum_{b\neq a} \exp\left(-\alpha_a R_{ab}^2\right) - \sum_{b\neq a} Z_b Y_b \exp\left(-\alpha_a R_{ab}^2\right),$$

where Z_a is a empirical parameter, Y_a is the Löwdin population that is associated with the diffuse basis functions on atom a, α_a is the diffuse orbital exponent on atom a, and R_{ab} is the distance between atom a and b. ZCMSRP(*i*) determines Z_a for atomic number *i*. For example, if one uses $Z_a = 0.11$ for a Cl atom, one sets ZCMSRP(17)=0.11. The default is that all the ZCMSRP are zero.

DCMSRP, CCMSRP Determines the parameters used in the CM2/CM3/CM4 charge calculation when ICMD=998 and ICMD=999. These are the same as C and D in namelist $CM2SRP. The CM2/CM3/CM4 charge is given by

$$Q_a = Q_a^0 + \sum_{b\neq a} B_{ab}\left(D_{ab} + C_{ab} B_{ab}\right),$$

where Q_a^0 is the partial atomic charge from either LPA or RLPA, B_{ab} is the Mayer bond order between atom a and b, and D_{ab} and C_{ab} are empirical parameters. DCMSRP(*i*) and CCMSRP(*i*) determine D_{ab} and C_{ab} for atomic pair *i*. The list of the atomic pairs is the same as that in namelist $CM2SRP. For example, if one uses $D_{ab} = -0.02$ for a H and C pair and $D_{ab} = 0.11$ for a C and Cl pair, one sets DCMSRP(1)=−0.02 and DCMSRP(13)=0.11. The default is that all the DCMSRP and CCMSRP are zero.

DR Defines the displacement size (in bohr) of the coordinate used in computing $\frac{\partial \mathbf{Q}}{\partial \mathbf{R}}$ (EETYP=DQDR) by numerical differentiation of the charges. Default=0.01

DPHI Defines the displacement size (in a.u.) of the electrostatic potential used in computing $\frac{\partial \mathbf{Q}}{\partial \mathbf{\Phi}}$ (EETYP=DQDPHI) by numerical differentiation of the charges. Default=0.01

Input and Output Examples

GAMESSPLUS is transparent if one performs standard *GAMESS* calculations, that is, the program may be used in exactly the same way that one uses *GAMESS*. However, if one would like to use the extra capabilities provided by *GAMESSPLUS*, then, at the very least, either the $GMSOL or the $CM2 namelists are required. To use the CM2-SRP method, the $CM2SRP namelist is required, and to use the NDDO-SRP capability, the $NDDOSRP namelist is required. Below is a set of input examples for a variety of typical *GAMESSPLUS* calculations:

Example 1: Gas phase CM2 dipole moment calculation for water using HF/MIDI!, where MIDI! is inputted as a general basis

```
 $CONTRL SCFTYP=RHF RUNTYP=ENERGY COORD=UNIQUE ISPHER=1 $END
 $GMSOL ISCRF=0 ICMD=1 $END
 $DATA
water gas-phase CM2 dipole moment
C1
O   8.0      0.000000    0.000000    0.000000
S   3
    1    281.8665800        0.0690600000
    2     42.4160000        0.3931590000
    3      9.0956200        0.6656690000
S   2
    1     11.4660300       -0.0808200000
    2      0.8878600        0.5820900000
S   1
    1      0.2788000        1.0000000000
P   2
    1      8.0472400        0.1242710000
    2      1.6684200        0.4765940000
P   1
    1      0.3725100        1.0000000000
D   1
    1      0.8000000        1.0000000000
H1  1.0    0.967300    0.000000    0.000000
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000

H1  1.0   -0.210300    0.944200    0.000000
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000
 $END
```

Example 2: SM5.42R/HF/6-31+G* calculation of a water molecule in methanol solution

```
 $CONTRL SCFTYP=RHF RUNTYP=ENERGY COORD=UNIQUE $END
 $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 POLAR=POPLE DIFFSP=.TRUE. $END
 $CM2 ISCRF=1 IGAS=0 ICMD=8 IAQU=0 SolN=1.3288 SolA=0.43
     SolB=0.47 SolG=31.77 Dielec=32.613 $END
 $DATA
SM5.42R/HF/6-31+G* calculation of water in methanol
C1
O    8.0     0.000000    0.000000    0.000000
H1   1.0     0.967300    0.000000    0.000000
H1   1.0    -0.210300    0.944200    0.000000
 $END
```

Example 3: SM5.42/HF/AM1 geometry optimization of water in aqueous solution

```
 $CONTRL SCFTYP=RHF RUNTYP=OPTIMIZE COORD=UNIQUE $END
 $BASIS GBASIS=AM1 $END
 $STATPT NSTEP=50 $END
 $GMSOL ISCRF=1 ICMD=11 IAQU=1 $END
 $DATA
optimization of water in water
C1
O    8.0     0.000000    0.000000    0.000000
H1   1.0     0.967300    0.000000    0.000000
H1   1.0    -0.210300    0.944200    0.000000
 $END
```

Example 4: Numerical frequency evaluation using central differences and a step size of 0.0005 bohr for water in liquid acetone using SM5.42/HF/PM3

```
 $CONTRL SCFTYP=RHF RUNTYP=HESSIAN COORD=UNIQUE $END
 $FORCE METHOD=NUMERIC NVIB=2 VIBSIZ=0.0005 VIBANL=.TRUE. $END
 $CM2 ISCRF=1 ICMD=12 IAQU=0 Dielec=20.493 SolN=1.3588 SolA=0.04
      SolB=0.49 SolG=33.77 SolC=0.000 SolH=0.000 $END
 $DATA
frequency calculation of water in acetone
C1
O    8.0     0.000000    0.000000    0.000000
H1   1.0     0.967300    0.000000    0.000000
H1   1.0    -0.210300    0.944200    0.000000
 $END
```

Example 5: Transition state optimization for the reaction of Cl- and CH_3Br using *Z*-matrix input and SM5.42/HF/AM1, with numerical Hessian evaluation before and after the optimization using forward differences and a step-size of 0.001 bohr.

```
 $CONTRL SCFTYP=RHF RUNTYP=SADPOINT COORD=ZMT ICHARG=-1 $END
 $BASIS GBASIS=AM1 $END
 $STATPT NSTEP=50 HESS=CALC HSSEND=.TRUE. $END
 $FORCE METHOD=NUMERIC NVIB=1 VIBSIZ=0.001 $END
 $GMSOL ISCRF=1 ICMD=11 IAQU=1 $END
 $DATA
SN2 rxn TS: Cl- + CH3Br
C1
C1
H2   C1    B1
H3   C1    B2    H2    A1
H4   C1    B3    H2    A2    H3    D1
Cl   C1    B4    H2    A3    H3    D2
Br   C1    B5    H2    A4    H3    D3

B1   =     1.09846481
B2   =     1.09846481
B3   =     1.09846481
B4   =     2.03270616
B5   =     2.35911154
A1   =   119.44411379
A2   =   119.44411379
A3   =    94.29818687
A4   =    85.70181313
D1   =  -165.20732678
D2   =    97.39633661
D3   =   -82.60366339
 $END
```

Note: The user cannot enter both a *Z*-matrix and a general basis set in $DATA, i.e., for *Z*-matrix input the user is limited to using only *GAMESS*-supported basis sets that can be entered using the $BASIS namelist.

Example 6: EEQM calculation of charge response kernel with respect to electrostatic potential for water in the presence of an electrostatic potential using MPW1K/6-31G(d).

```
$CONTRL SCFTYP=RHF RUNTYP=EEQM $END
 $EEQM EETYP=DQDPHI ICMD=417 HFE=0.428
  PHI(1)= 1.2873,-0.0734,-0.0734
 $END
 $DFT DFTTYP=MPWX HFE=0.428 NRAD=80 NTHE=16 NPHI=32 $END
 $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
 $DATA
water, MPW1K/6-31G(d)
C1 0
 O           8.0    .0000000000    .0000000000   -.0931046833
 H           1.0   -.7633156080    .0000000000    .5148742072
 H           1.0    .7633156080    .0000000000    .5148742072
 $END
```

Output example 1

Below is an example of typical *GAMESSPLUS* output (for acetamide in chloroform using SM5.42/HF/MIDI!). The output consists of (1) gas-phase and liquid-phase Löwdin and CM2 charges and dipole moments, (2) the atomic and term contributions to the G_P and G_{CDS} terms, and (3) a list of relevant energies. The breakdown of the G_{CDS} term only contains the terms relevant to the inputted molecular structure (i.e., the contribution of the iodine CDS parameter is not outputted for solvated acetamide):

```
         -------------------
          RHF SCF CALCULATION
          -------------------

     NUCLEAR ENERGY =       120.7176594569
     MAXIT =   30     NPUNCH=    2
     EXTRAP=T  DAMP=F  SHIFT=F  RSTRCT=F  DIIS=F  DEM=F  SOSCF=T
     DENSITY CONV=  1.00E-05
     SOSCF WILL OPTIMIZE     640 ORBITAL ROTATIONS, SOGTOL=   0.250
     MEMORY REQUIRED FOR RHF STEP=     56920 WORDS.

 ITER EX DEM  TOTAL ENERGY      E CHANGE  DENSITY CHANGE     ORB. GRAD
   1  0  0  -206.137672838  -206.137672838   0.433942041   0.000000000
          ---------------START SECOND ORDER SCF---------------
   2  1  0  -206.667475470    -0.529802632   0.143753708   0.080707676
   3  2  0  -206.701484404    -0.034008934   0.080332362   0.072062843
   4  3  0  -206.741804366    -0.040319962   0.020454116   0.005919665
   5  4  0  -206.742463230    -0.000658865   0.005185967   0.001776303
   6  5  0  -206.742520665    -0.000057435   0.001931901   0.000609944
   7  6  0  -206.742526207    -0.000005542   0.000394123   0.000218359
   8  7  0  -206.742526834    -0.000000627   0.000187438   0.000050111
   9  8  0  -206.742526898    -0.000000064   0.000029765   0.000012969
  10  9  0  -206.742526902    -0.000000004   0.000009812   0.000003702
  11 10  0  -206.742526903     0.000000000   0.000002039   0.000000898

----------------------------------------------------------
Charge Model 2 (CM2) and 3 (CM3) and
Solvation Model SM5.42 GAMESSPLUS version 4.8 at 298 K
----------------------------------------------------------

 Gas-phase
------------------------------------------------
 Center     Atomic         CM2          Lowdin
 Number     Number         Charge       Charge
------------------------------------------------
    1          8           -0.418       -0.382
    2          6            0.538        0.475
    3          6           -0.253       -0.331
    4          7           -0.717       -0.521
    5          1            0.092        0.118
    6          1            0.083        0.110
    7          1            0.083        0.110
    8          1            0.295        0.211
    9          1            0.295        0.211
------------------------------------------------

 Gas-phase dipole moment (Debye)
------------------------------------------------
              X         Y         Z      Total
 CM2        1.438    -3.869    0.000    4.127
 Lowdin     1.356    -3.340    0.000    3.605
------------------------------------------------


 BREAKDOWN OF G_P AND G_CDS (KCAL/MOL)
------------------------------------------------
 Center     Atomic         Atomic contribution
 Number     Number          G_P          G_CDS
-----------------------------------------------
    1          8          -3.070          0.261
    2          6           0.193         -0.035
    3          6           1.176          0.343
    4          7           1.288         -0.300
    5          1          -0.322         -0.282
    6          1          -0.569         -0.281
    7          1          -0.569         -0.281
    8          1          -1.632         -0.160
    9          1          -1.632         -0.160

       TOTAL              -5.137         -0.895
-----------------------------------------------

  12 11  0  -206.742526903     0.000000000   0.000127049   0.010403370
  13 12  0  -206.742526867     0.000000036   0.018181743   0.010394019
  14 13  0  -206.741887416     0.000639452   0.008988029   0.002303959
  15 14  0  -206.741478988     0.000408428   0.001538903   0.000483566
  16 15  0  -206.741438208     0.000040780   0.000568649   0.000168850
  17 16  0  -206.741452458    -0.000014250   0.000135792   0.000030765
  18 17  0  -206.741455761    -0.000003303   0.000045146   0.000013727
  19 18  0  -206.741455604     0.000000157   0.000012170   0.000004398
  20 19  0  -206.741455279     0.000000325   0.000002647   0.000000839
  21 20  0  -206.741455250     0.000000029   0.000000920   0.000000240
  22 21  0  -206.741455258    -0.000000009   0.000000281   0.000000059

 Solution
-----------------------------------------------
 Center     Atomic          CM2          Lowdin
 Number     Number         Charge        Charge
-----------------------------------------------
    1          8          -0.460         -0.423
    2          6           0.544          0.480
    3          6          -0.251         -0.329
    4          7          -0.724         -0.531
    5          1           0.089          0.115
    6          1           0.098          0.123
    7          1           0.098          0.123
    8          1           0.303          0.221
    9          1           0.303          0.221
-----------------------------------------------

 Solution dipole moment (Debye)
-----------------------------------------------
              X        Y         Z       Total
 CM2        1.576   -4.394     0.000     4.668
 Lowdin     1.486   -3.857     0.000     4.134
-----------------------------------------------
```

```
Breakdown of G_P and G_CDS (kcal/mol)
------------------------------------------------
 Center     Atomic         Atomic contribution
 Number     Number          G_P          G_CDS
------------------------------------------------
    1          8           -3.909         0.261
    2          6            0.326        -0.035
    3          6            1.316         0.343
    4          7            1.406        -0.300
    5          1           -0.329        -0.282
    6          1           -0.777        -0.281
    7          1           -0.777        -0.281
    8          1           -1.824        -0.160
    9          1           -1.824        -0.160

       TOTAL               -6.391        -0.895
------------------------------------------------


Single-subscript surface-tension-coefficient contribution to G_CDS:
--------------------------------------------------------------------------------

SS Area = Small-Sphere (Surface) Area
LS Area = Large-Sphere (Surface) Area = sum of all SS Areas (SM5.42 model)

  Element     SS Area      Sigma         Total
     Z                       k
            (Ang**2)  (cal/Ang**2)  (kcal/mol)
------------------------------------------------

     H          42.404      58.501       2.481
     C          18.637      90.501       1.687
     N          10.027      -2.536      -0.025
     O          17.861     -60.537      -1.081

Subtotal:       88.929                   3.061
------------------------------------------------

Double-subscript surface-tension-coefficient contribution to G_CDS:
--------------------------------------------------------------------------------

  Element      COT *       Sigma         Total
  Z  Z'       SS Area       k,k'
            (Ang**2)  (cal/Ang**2)  (kcal/mol)
------------------------------------------------
  H  C          16.431    -149.434      -2.455
  H  N          12.589    -120.411      -1.516
  H  N(2)        0.000    -275.560       0.000
  H  O           0.000     -66.749       0.000
  H  O(2)        0.000     282.601       0.000
  C  C          11.462    -106.505      -1.221
  C  C(2)        0.000      -1.142       0.000
  O  C          11.755     102.487       1.205
  C  N           3.423     -88.294      -0.302
  N  C          12.301     -28.639      -0.352
  N  C(2)        4.736       0.000       0.000
  N  C(3)        0.000      85.944       0.000
  O  N           0.000     105.726       0.000

Subtotal:                               -4.642
------------------------------------------------
```

```
Contributions to G_CDS that depend on the LS Area but do not on atomic number:
--------------------------------------------------------------------------------

Solvent Descriptors (SD) used to describe the solvent:
---------------------------------------------------
SOLN = index of refraction
SOLA = Abraham's H-bond acidity
SOLB = Abraham's H-bond basicity
SOLG = macroscopic surface tension at air/solvent interface
SOLC = aromaticity: fraction of non-H atoms that are aromatic C
SOLH = electronegative halogenicity: fraction of non-H atoms that are F, Cl, Br

Note: In the evaluation of the LS contribution, SOLB, SOLC and SolH are squared,
      while SOLG is unsquared. SOLN, SOLA and SOLB are unsquared in the
      evaluation of the SS Contribution.

 Solvent       SD          LS          LS           SS        Total
Descriptor    Value       Sigma    Contribution Contribution
   (SD)               (cal/Ang**2)  (kcal/mol)  (cal/Ang**2) (kcal/mol)
------------------------------------------------------------------------
   SOLN       1.446                                -1.383      -1.383
   SOLA       0.150                                -0.186      -0.186
   SOLB       0.020       0.002        0.000       -0.012      -0.012
   SOLG      38.390      12.642        1.124                    1.124
   SOLC       0.000       0.000        0.000                    0.000
   SOLH       0.750      -4.928       -0.438                   -0.438

Subtotal:                 7.717        0.686       -1.581      -0.895
------------------------------------------------------------------------

Summary of SM5.42 solvation calculation by GAMESSPLUS version 4.8 at 298 K
--------------------------------------------------------------------------------

Options: ISCRF = 1, IGAS = 0, ICMD = 1, ICDS = 1, IAQU = 0, DIELEC =  4.71

(0)  E-EN(g) gas-phase RHF      elect-nuc energy         -206.742526903 a.u.
(1)  E-EN(liq) elect-nuc RHF      energy of solute       -206.741455258 a.u.
(2)  G-P(liq) polarization free energy of solvation              -6.391 kcal/mol
(3)  G-ENP(liq) elect-nuc-pol free energy of system      -206.751640660 a.u.
(4)  G-CDS(liq) cavity-dispersion-solvent structure              -0.895 kcal/mol
(5)  G-P-CDS(liq) = G-P(liq) + G-CDS(liq) = (2) + (4)            -7.286 kcal/mol
(6)  G-S(liq) free energy of system = (1) + (5)          -206.753066622 a.u.
(7)  DeltaE-EN(liq) elect-nuc reorganization
     energy of solute molecule (7) = (1) - (0)                    0.672 kcal/mol
(8)  DeltaG-ENP(liq) elect-nuc-pol free energy
     of solvation (8) = (3) - (0)                                -5.719 kcal/mol
(9)  DeltaG-S(liq) free energy of  solvation
     (9) = (6) - (0)                                             -6.614 kcal/mol

------------------------------------------------------------------------------

          -----------------
          DENSITY CONVERGED
          -----------------
     TIME TO FORM FOCK OPERATORS=      12.3 SECONDS (       0.6 SEC/ITER)
     TIME TO SOLVE SCF EQUATIONS=       5.2 SECONDS (       0.2 SEC/ITER)

 FINAL ENERGY IS     -206.7414552585 AFTER  22 ITERATIONS
```

Output example 2: EEQM calculation of charge response kernel with respect to electrostatic potential for water in the presence of an electrostatic potential using MPW1K/6-31G(d). This is the output of input example 6.

```
          -------------
          $EEQM OPTIONS
          -------------
     EETYP  =DQDPHI       IUPHI  =       1
     ICMD   =     417     HFE    =    .428
     DPHI   =1.00E-02     DR     =1.00E-02

          ------------------------
          ELECTROSTATIC POTENTIALS
          ------------------------

          ATOM          PHI(A.U.)      PHI(eV)

       1  O             .047307483   1.287300000
       2  H            -.002697405   -.073400000
       3  H            -.002697405   -.073400000

                          .
                          .
                          .

          --------------------------
              R-MPWX SCF CALCULATION
          --------------------------

     NUCLEAR ENERGY =         9.0229744912

DFT IS SWITCHED OFF, PERFORMING PURE SCF UNTIL SWOFF THRESHOLD IS REACHED.

 ITER EX DEM     TOTAL ENERGY         E CHANGE  DENSITY CHANGE    DIIS ERROR
   1  0  0      -75.7885765352   -75.7885765352     .245832714    .601034454
          * * *   INITIATING DIIS PROCEDURE   * * *
   2  1  0      -76.0309122171     -.2423356819     .070642105    .091199987
   3  2  0      -76.0427974735     -.0118852564     .026571081    .028311564
   4  3  0      -76.0440819683     -.0012844948     .005431369    .005878622
   5  4  0      -76.0441223369     -.0000403686     .002219695    .003258219
 CONVERGED TO SWOFF, SO DFT CALCULATION IS NOW SWITCHED ON.
          * * *   INITIATING DIIS PROCEDURE   * * *
   6  5  0      -76.4143579281     -.3702355912     .047897931    .055345989
   7  6  0      -76.4153865814     -.0010286534     .011904807    .013572266
   8  7  0      -76.4154042610     -.0000176796     .005464921    .012224870
   9  8  0      -76.4155128987     -.0001086377     .000263181    .000260798
 DFT CODE IS SWITCHING BACK TO THE FINER GRID
          * * *   INITIATING DIIS PROCEDURE   * * *
  10  9  0      -76.4157657975     -.0002528988     .000416415    .000396330
  11 10  0      -76.4157660010     -.0000002036     .000054055    .000056104
  12 11  0      -76.4157660045     -.0000000034     .000024524    .000044739
  13 12  0      -76.4157660058     -.0000000013     .000004586    .000008927
  14 13  0      -76.4157660059     -.0000000001     .000000384    .000000688

          -----------------
          DENSITY CONVERGED
          -----------------

                          .
                          .
                          .
     --------------------------
     LOWDIN POPULATION ANALYSIS
     --------------------------

----- GROSS POPULATIONS ON ATOMS -----

        SITE        Z          ELEC          GROSS

   1  O            8.00     -8.787794      -.787794
   2  H            1.00      -.606103       .393897
   3  H            1.00      -.606103       .393897

----- DIPOLE MOMENT CALCULATED FROM LOWDIN PARTIAL CHARGES -----

DIPOLE MOMENT     .905107 A.U.     2.300570 Debye
          DMX     .000001           .000002
          DMY     .000000           .000000
          DMZ     .905107          2.300570


     -----------------------
     CM4 POPULATION ANALYSIS
     -----------------------

----- GROSS POPULATIONS ON ATOMS -----

        SITE        Z          ELEC          GROSS

   1  O            8.00     -8.727499      -.727499
   2  H            1.00      -.636251       .363749
   3  H            1.00      -.636250       .363750

----- DIPOLE MOMENT CALCULATED FROM CM4 PARTIAL CHARGES -----

DIPOLE MOMENT     .835834 A.U.     2.124494 Debye
          DMX     .000001           .000002
          DMY     .000000           .000000
          DMZ     .835834          2.124494

     -----------
     EEQM ENERGY
     -----------

V(QM)            =      -76.3793874795 hartree
        V(QM/MM) =        -.0363785264
V(QM) + V(QM/MM) =      -76.4157660059
```

```
    ------------------------------------------------------------
    CHARGE RESPONSE KERNEL (DQ/DPHI) CALCULATION -- STEP  1 OF  6
    ------------------------------------------------------------

    ELECTROSTATIC POTENTIAL ON ATOM  1 IS CHANGED BY  1.00E-02 A.U.

         ATOM          PHI(A.U.)      PHI(eV)

       1  O             .057307483   1.559413400
       2  H            -.002697405   -.073400000
       3  H            -.002697405   -.073400000


         --------------------------
             R-MPWX SCF CALCULATION
         --------------------------

ITER EX DEM     TOTAL ENERGY        E CHANGE  DENSITY CHANGE    DIIS ERROR
         * * *   INITIATING DIIS PROCEDURE   * * *
  1  0  0      -76.4227877518   -76.4227877518    .006265950    .005596441
  2  1  0      -76.4228502008     -.0000624491    .003080053    .005147627
  3  2  0      -76.4228636696     -.0000134688    .000732871    .001538861
  4  3  0      -76.4228654385     -.0000017688    .000065095    .000072755
DFT CODE IS SWITCHING BACK TO THE FINER GRID
         * * *   INITIATING DIIS PROCEDURE   * * *
  5  4  0      -76.4231207234     -.0002552849    .000450913    .000402486
  6  5  0      -76.4231209372     -.0000002138    .000079777    .000084401
  7  6  0      -76.4231209413     -.0000000041    .000039160    .000083245
  8  7  0      -76.4231209462     -.0000000049    .000002635    .000005880
  9  8  0      -76.4231209462      .0000000000    .000000611    .000000816

         -----------------
         DENSITY CONVERGED
         -----------------

                        .
                        .
                        .

              -----------------------
              CM4 POPULATION ANALYSIS
              -----------------------

         ----- GROSS POPULATIONS ON ATOMS -----

                SITE        Z          ELEC          GROSS

            1  O           8.00     -8.743471      -.743471
            2  H           1.00      -.628265       .371735
            3  H           1.00      -.628264       .371736

         ----- DIPOLE MOMENT CALCULATED FROM CM4 PARTIAL CHARGES -----

         DIPOLE MOMENT     .854184 A.U.     2.171135 Debye
                  DMX     .000001           .000001
                  DMY     .000000           .000000
                  DMZ     .854184          2.171135

              -----------
              EEQM ENERGY
              -----------

         V(QM)             =      -76.3785090551 hartree
               V(QM/MM) =        -.0446118911
         V(QM) + V(QM/MM) =      -76.4231209462
```

```
     -------------------------------------------------------------
     CHARGE RESPONSE KERNEL (DQ/DPHI) CALCULATION -- STEP  2 OF  6
     -------------------------------------------------------------

     ELECTROSTATIC POTENTIAL ON ATOM  1 IS CHANGED BY -1.00E-02 A.U.

          ATOM          PHI(A.U.)      PHI(eV)

       1  O             .037307483   1.015186600
       2  H           -.002697405    -.073400000
       3  H           -.002697405    -.073400000

          --------------------------
              R-MPWX SCF CALCULATION
          --------------------------
ITER EX DEM     TOTAL ENERGY        E CHANGE  DENSITY CHANGE    DIIS ERROR
          * * *   INITIATING DIIS PROCEDURE   * * *
   1  0  0      -76.4079958030   -76.4079958030    .012967634    .010408669
   2  1  0      -76.4082612494     -.0002654464    .006194544    .010158410
   3  2  0      -76.4083125495     -.0000513002    .001567227    .003288009
   4  3  0      -76.4083205513     -.0000080018    .000141059    .000163889
 DFT CODE IS SWITCHING BACK TO THE FINER GRID
          * * *   INITIATING DIIS PROCEDURE   * * *
   5  4  0      -76.4085707745     -.0002502231    .000400689    .000407063
   6  5  0      -76.4085709931     -.0000002186    .000066801    .000069391
   7  6  0      -76.4085709970     -.0000000038    .000033901    .000071098
   8  7  0      -76.4085710006     -.0000000036    .000002724    .000006190
   9  8  0      -76.4085710006      .0000000000    .000000615    .000000814

          -----------------
          DENSITY CONVERGED
          -----------------
                         .
                         .
                         .
               -----------------------
               CM4 POPULATION ANALYSIS
               -----------------------

          ----- GROSS POPULATIONS ON ATOMS -----

                  SITE        Z          ELEC          GROSS

             1  O            8.00     -8.711485      -.711485
             2  H            1.00      -.644258       .355742
             3  H            1.00      -.644257       .355743

          ----- DIPOLE MOMENT CALCULATED FROM CM4 PARTIAL CHARGES -----

          DIPOLE MOMENT     .817434 A.U.     2.077727 Debye
                    DMX     .000001           .000002
                    DMY     .000000           .000000
                    DMZ     .817434          2.077727
                         .
                         .
                         .
          --------------------------------
          CHARGE RESPONSE KERNEL (DQ/DPHI)
          --------------------------------

      Q : PHI      O             H             H
     O            -1.599308       .799655       .799654
     H              .799655      -.872016       .072361
     H              .799654       .072361      -.872015
```

Notes on Running *GAMESSPLUS*

It is recommended that *GAMESS* and/or *GAMESSPLUS* be run using a modified version of the script called "rungms" that is provided in the top-level directory of the *GAMESS* distribution. This file performs several required steps for each calculation:

1. Sets the scratch (working) directory.
2. Initializes some system environmental variables.
3. Launches the *GAMESS* / *GAMESSPLUS* executable using the executable *ddikick.x*, which accepts many arguments, including the *GAMESS* / *GAMESSPLUS* executable name. Every job, including serial jobs, is treated as a parallel job.
4. Defines the scratch directory location in which to store scratch files used during the calculation. Note that external files used in a *GAMESSPLUS* calculation (i.e., the CM2CHG file, the CDS_Param file, or the files defined by the $CM2SRP and the $NDDOSRP namelists) must be located in the scratch directory defined in the rungms script.
5. Deletes all files that are no longer required after the calculation is finished (it does not delete external files, i.e., CM2CHG, CDS_Param, or files defined by the $CM2SRP and the $NDDOSRP namelists).

The user must make a few modifications to the *rungms* script before any calculation is performed:

1. the "set TARGET=" line should be modified as described in the preamble in the file
2. on the following line, the variable SCR must be set to equal to the desired scratch directory name
3. replace all occurrences of ~scr with $SCR
4. the instance of "set GMSPATH=" statement that applies to the user's machine must be set equal to the pathname for the user's *GAMESS*/*GAMESSPLUS* executable. If you do not know which machine you are using, change all instances of "set GMSPATH".
5. On the line beginning "setenv ERICFMT", replace the default pathname with the pathname for the user's *GAMESS*/*GAMESSPLUS* executable.
6. If the *GAMESS* executable was named *name.version.x*, and *name* does not equal "gamess", then every instance of "gamess.$VERNO.x" should be replaced with "*name*.$VERNO.x".

Input files must be named $JOB.inp, where $JOB will be the prefix assigned to all files created during the execution of the calculation. The corresponding output file will be called $JOB.out. The *GAMESS Users Manual* describes all of the other files that are created. After the job is finished, the only files not deleted are $JOB.inp, $JOB.out, $JOB.dat, and $JOB.irc. $JOB.dat and $JOB.irc are restart files; they must be manually deleted between the consecutive runs of the same job if these files are not required for a subsequent calculation.

To start running a calculation with input file $JOB, type:

```
sh path/rungms $JOB version > $JOB.out &
```

where `path` is the directory path to the *rungms* script and `sh` can be replaced with any shell command (e.g. `csh`, `bash`), and `version` is the version number in the executable name (if left out, *version* is assumed to equal 00). If the location of the rungms script is defined in the user's path environment variable (the UNIX command ‘which rungms’ returns the absolute path to the rungms script, not ‘command not found’), then *GAMESS/GAMESSPLUS* can be run by executing the command:

```
./rungms $JOB version > $JOB.out &
```

As noted above, all *GAMESS* jobs are run as parallel jobs. For serial jobs, *ddikick.x* spawns two *GAMESS* processes, although only one process actually does anything. The running of *GAMESSPLUS* has not been tested in parallel. It is recommended that users not run parallel *GAMESSPLUS* jobs without running and checking the full test suite first.

For GHO-AIHF calculations, one needs to run the *CHARMM/GAMESSPLUS* combination package, which is compiled as an integrated executable `charmm`. To start running a calculation with the CHARMM input file $JOB.inp, type:

```
path/charmm <$JOB.inp> $JOB.out
```

where `path` is the directory path to the integrated `charmm` executable.

The *GAMESSPLUS* test suite

The *GAMESSPLUS* test suite is located in the /Input and /Output directories of the *GAMESSPLUS* distribution. For SM6, a separate test suite has been put together—the input and output files for this portion of the test suite are located in the /sm6 directory. Input files of test calculations and scripts to run the test suite and analyze the results are located in the /Input directory. Output files corresponding to a successful installation of *GAMESSPLUS* are located in the /Output directory. In the following sections, a discussion of the contents of the test suite, how to run the test suite, and how to verify that *GAMESSPLUS* is installed correctly by analyzing the test suite is presented.

Note: Because the use of the GHO-AIHF functionality of *GAMESSPLUS* is quite different from other calculations, a separate test suite is provided by the *CGPLUS* package for testing GHO-AIHF through the *CHARMM/GAMESSPLUS* combination package (http://comp.chem.umn.edu/cgplus, see the *CGPLUS-v1.0 Users Manual*).

Description of the test suite for CM2, CM3, CM4, SM5.42, and SM5.43.

A comprehensive test suite for CM2, CM3, CM4, SM5.42, and SM5.43 has been put together for *GAMESSPLUS* (SM6 and SM6T were tested in a different way than SM5.42 and SM5.43; the test suites for SM6 and SM6T are described below). This is the relevant part of a more general test suite that is used by all codes that implement these models.

Nine molecules have been selected such that all of the CM2 parameters and all of the SM5.42 atomic surface tension parameters are tested. The nine molecules are nitroethyne, methyl disulfide, fluorochlorobromoiodomethane, water, hydrazine, acetamide, hydrogen cyanide, hydrogen thionophosphate, and 1-(*P*-fluorophosphino)-1-silylmethanesulfonic acid. These calculations are named test1.*x*2*y* – test9.*x*2*y*, where *x* indicates if a gas-phase calculation of the CM2 charges (*x* = 0), an SM5.42 free energy of solvation evaluation in aqueous solution (*x* = 1), or an SM5.42 free energy of solvation evaluation in liquid chloroform (*x* = 2) is carried out, the number 2 indicates that CM2 charges are used in the calculation, and the value of *y* corresponds to the value of ICMD (see the section entitled *Namelists $GMSOL and $CM2* for allowed values of ICMD) used in the calculation. For calculations corresponding to *x* > 0, SCF Scheme I is used except for those calculations involving SM5.42/HF/6-31+G* (ICMD = 8, test*w*.*x*28, *w* = 1 – 8); many SCRF evaluations using ICMD = 8 and ISCRF = 1 diverge (note that ISCRF=1 is no longer available for methods that use diffuse basis functions because of this property), so where appropriate, SCF scheme II is used (ISRCF=2). In addition, when *x* > 0, when the basis set employed uses Cartesian *d* shells, and when ISCRF = 1, analytical gradient evaluations are performed (i.e., for ICMD = 2, 3, 11, and 12). The table below summarizes this portion of the test suite.

Calculation type	test cases
gas-phase Löwdin and CM2 charges	test9.02*y*, *y*= 1, 2, 3, 6, 8, 9, 11, 12
SM5.42 energy evaluation in aqueous and organic solvent	test*w*.*q*2*y*, where *w* = 1 – 8, *q* = 1 and 2, and *y* = 1, 6, 8, and 9
SM5.42 energy and analytical gradient evaluation in aqueous and organic solvent	test*w*.*q*2*y*, where *w* = 1 – 8, *q* = 1 and 2, and *y* = 2, 3, 11, and 12

Eight of these nine molecules are used to test most of the CM3 and CM4 parameters and all of the SM5.43 parameters (CM3 and SM5.43 do not contain parameters for I). Because CM3 and CM4 are also parameterized for Li and for molecules that containing Si-O, Si-F, and Si-Cl bonds, three additional compounds, orthosilic acid, HSi(OH)FCl and $Li_2C_2N_2OFClSH_{13}$ (test10, test11, and test12, respectively), have been added to the test suite. These test calculations are named test1.*qy* – test12.*qy*, where *q* = 0, 1, or 2 and *y* is the value of ICMD used in the calculation. Note that in the portion of the test suite that tests the CM2 parameters, a value of 2 for CM2 was included in the name before the value of ICMD, while an analogous value of 3 for CM3 and 4 for CM4 is already included in the values of ICMD that request a calculation of CM3 or CM4 charges. The table below summarizes this portion of the test suite:

Calculation type	test cases
gas-phase Löwdin and CM3 charges by HF, AM1, PM3, BLYP, and B3LYP	test*w*.0*y*, where *w* = 1, 2, 4 – 8 for *y* = 301 test*w*.0*y*, where *w* = 1, 2, 4 – 12 for *y* = 302, 312, 320, and 321 test*w*.0*y*, where *w* = 1, 2, 4 – 11 for *y* = 311 test*w*.0*y*, where *w* = 9, 10, 11, and 12 for *y* = 303 and 313
gas-phase Löwdin, RLPA, and CM3 charges by B3LYP	test*w*.0*y*, where *w* = 1, 2, 4 –12 and *y* = 314
gas-phase Löwdin and CM3 charges by MPW*X*	test*w*.*x*.0*y*, where *w* = 9 - 12, *x* = 0, 25, 428, 517, 606, and 999, and *y* = 315, 316 and 317
gas-phase RLPA and CM3 charges by MPW*X*	test*w*.*x*.0*y*, where *w* = 9 - 12, *x* = 0, 25, 428, 517, 606, and 999, and *y* = 318 and 319
gas-phase Löwdin and CM3.1 charges by HF/MIDI!	test*w*.0322, where *w* = 1, 4, 5, 6, and 7
SM5.43 energy and analytical gradient evaluation in aqueous and organic solvent by HF/6-31G* and B3LYP/6-31G*	test*w*.*qy*, where *w* = 1, 2, 4 – 8, *q* = 1 or 2, and *y* = 303 and 313
SM5.43 energy evaluation in aqueous and organic solvent by MPW*X*	test*w*.*x*.*qy*, where *w* = 1, 2, 4 – 8, *x* = 0, 25, 428, 517, 606, and 999, *q* = 1 or 2, and *y* = 315, 316, 317, 318, and 319
gas-phase Löwdin and CM4 charges by MPW*X*	test*w*.*x*.0*y*, where *w* = 9 - 12, *x* = 0, 25, 428, 517, 606, and 999, and *y* = 416 and 417
gas-phase RLPA and CM4 charges by MPW*X*	test*w*.*x*.0*y*, where *w* = 9 - 12, *x* = 0, 25, 428, 517, 606, and 999, and *y* = 418 and 419

The remaining test jobs test the remaining keywords available in *GAMESSPLUS*:

Two test cases are performed using SM5.42/HF/MIDI!6D (ISCRF = 1):

- geometry optimization of 2,4-pentadione in acetonitrile solution (test13).
- transition state optimization for the S_N2 reaction of NH_3 and CH_3Cl in aqueous solution (the Menschutkin reaction; test14). Two numerical Hessian calculations are carried out in test14: one before the optimization to generate a good Hessian guess, and one after the optimization to ensure that the stationary-point is a first-order saddle point.

Test case test16a performs a generalized Born electrostatic solvation energy and gradient evaluation using Löwdin partial atomic charges and UHF/6-31G(d,f) for peroxyl radical in aqueous solution. Test case test16b performs a generalized Born calculation of the electrostatic contribution to the free energy of solvation by RHF/MG3. Test case test16a tests the use of Cartesian *f* functions in the basis set and the use of unrestricted wave functions in generalized Born calculations. Test case test16b tests the use of spherical harmonic *f* functions in the basis set.

Test cases test17a, test17b, test17c, and test17d all perform an SM5.42/MIDI!6D calculation of water solute in acetone solvent using ISCRF = 2:

- test17a tests the ICREAD = 1 option
- test17b tests the ICSAVE = 1 option
- test17c tests the ISREAD = 1 option
- test17d tests the ISREAD = 2 options

The following two test cases test the SRP models available in *GAMESSPLUS*:

- gas-phase CM2-SRP/AM1-SRP evaluation for methanol using Form A input (test20a)
- gas-phase AM1-SRP SCF evaluation for methanol using Form B input (test20b)

Test case test21 tests the calculation of the bond order using an unrestricted wave function. In particular, test21 performs an SM5.42/UHF/6-31G* energy calculation of CH_4^+ in water.

Test case test22 calculates the electrostatic contribution to the aqueous free energy of solvation of water using RLPA charges. The corresponding free energy gradient is also calculated in test22. Test23 tests the new CM3 mapping scheme for compounds that contain N and O. In particular, test23 is a calculation of the electrostatic contribution to the aqueous free energy of solvation of pyramidal formamide using the generalized Born method and CM3 charges calculated by B3LYP/6-31G*. The free energy gradient (which is calculated analytically) is also calculated in test23.

Test cases test24 and test25 test the MPW1K functional. Test cases test26 and test27 both test the B3LYP functional. In test26, the version III VWN correlation functional is used (B3LYP3, the version of B3LYP implemented in *Gaussian*); test27 uses the version V VWN correlation functional (B3LYP5, the version of B3LYP implemented in *GAMESS*). Test case test28 tests the MPW1PW91 functional (also called MPW25), and test case test29 tests the MPW*X* functional and the use of the HFE keyword in the $DFT namelist.

The above portion of the *GAMESSPLUS* test suite may be run conveniently with a set of execution scripts, which are located in the *Input* directory. Because the test suite is extensive, the test calculations that test specific capabilities of *GAMESSPLUS* are grouped together in a single execution script. For example, all calculations that test the CM2 and SM5.42 parameters for ICMD = 1 are included in one execution script. The table below summarizes the files to be run and the name of the corresponding execution script:

Contents of execution script	execution script name
test*w*.*q*2*y*, where *w* = 1 – 9, *q* = 0, 1 or 2, and *y* = 1, 2, 3, 8, 9, 11, and 12	batch.2*y*, *y* = 1, 2, 3, 6, 8, 9, 11, and 12
test*w*.*qy*, where *w* = 1 – 12, *q* = 0, 1, or 2, and y = 301, 302, 303, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 416, 417, 418, or 419.	batch.*y*, *y* = 301, 302, 303, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 416, 417, 418, or 419
all test calculations named test13 and higher	batch.misc

Each execution script can be run interactively using the command:

```
sh batch.2y
sh batch.y
sh batch.misc
```

The execution scripts can be easily modified so that they can be submitted to a queue.

Description of the test suite for SM6.

Eight molecules have been selected such that all of the SM6 atomic surface tension parameters are tested. In all cases, calculations were performed using MPW*X* (where *X* is the percentage of Hartree-Fock exchange). The eight molecules are nitroethyne (test1.*a*.1*b*), methyl disulfide (test2.*a*.1*b*), hydrogen peroxide (test3.*a*.1*b*), water (test4.*a*.1*b*), hydrazine (test5.*a*.1*b*), acetamide (test6.*a*.1*b*), hydrogen cyanide (test7.*a*.1*b*), and hydrogen thionophosphate (test8.*a*.1*b*). In the names above, *a* is the value of *X* in MPW*X* (between 0 and 999), and *b* is the value of ICMD (between 416 and 419).

While testing SM6, it became apparent that two important global issues related to performing liquid-phase calculations with the SM*x* models needed to be tested in a systematic way. The first is related to performing geometry optimizations in solution. This issue is discussed in more detail and is tested in the subsection "*Liquid-Phase Geometry Optimizations*". The second issue is related to the difference between using SCF Scheme I and SCF Scheme II (see the section "*SCF Schemes*" for more information on the difference between these two SCF schemes), for calculations involving diffuse basis sets. This issue is discussed in more detail and is tested in the subsection "*SCF Schemes for Diffuse Basis Functions*".

Liquid-Phase Geometry Optimizations

Two important issues that arise when optimizing geometries is the level of accuracy that should be used for both the SCF convergence tolerance (this is also an issue for single-point calculations) and the gradient convergence tolerance. In the first section of the SM6 test suite, we tested whether or not the default values for these two criteria in *GAMESSLUS* are reasonable for performing liquid-phase geometry optimizations. For both single-point calculations and geometry optimizations in *GAMESSPLUS*, by default, the SCF convergence is set to 10^{-5} a.u., the gradient convergence tolerance is set to 10^{-4} a.u./bohr for the maximum value of the gradient, and to 1/3 the value of tolerance for the maximum value of the gradient for the root mean square of the gradient. (For users familiar with the *Gaussian* program, it is informative to note that in *Gaussian*, by default, the SCF convergence for single-point calculations is set to 10^{-3} a.u., the SCF convergence for geometry optimizations is set to 10^{-8} a.u., the gradient convergence tolerance is set to 0.00045 a.u./bohr for the maximum value of the gradient, and to 2/3 the value of tolerance for the maximum value of the gradient for the root mean square of the gradient).

To test the default options in *GAMESSPLUS*, we performed full geometry optimizations in both the gas and liquid phases (using SCF Scheme I for the liquid-phase part of the calculation) for all jobs with the suffix 1416 (SM6/MPW*X*/MIDI!6D calculations) and 1417 (SM6/MPW*X*/6-31G* calculations), resulting in 96 total test jobs in this portion of the SM6 test suite. As an example of the protocol used for this portion of the test suite, for test job test3.428.1416, the geometry of hydrogen peroxide is first optimized in the gas-phase at the MPW1K/MIDI!6D level of theory, and then optimized in the aqueous phase at the SM6/MPW1K/MIDI!6D level of theory (these two calculations are all carried out as part of one job). For all of the test jobs in this part of the SM6 test suite, the aqueous solvation free energy is computed as the difference between the free energy of the equilibrium geometry in the gas-phase and the free energy of the equilibrium geometry in the aqueous phase. The results of this part of the test suite are summarized below.

In one case (test1.999.1417), an equilibrium geometry in the liquid phase could not be located, even when the SCF convergence tolerance and the gradient convergence tolerance were set to unreasonably large values. However, in all the input files except test1.999.1417, stationary points can be located within 50 steps with the default convergence tolerance value, OPTTOL, which is equal to 0.0001 hartree/bohr. Thus, the above results demonstrate that in many cases, the default options in *GAMESSPLUS* are suitable for performing geometry optimizations in solution. However, the user should be aware that in some cases the default options may lead to convergence problems. In these cases, depending on the non-default options specified, the calculated results should be interpreted with some degree of caution.

SCF Schemes for Diffuse Basis Functions

Another important issue that is tested in this section is the difference between using SCF Scheme I and SCF Scheme II for liquid-phase calculations involving diffuse basis functions. This is an important issue, because past experience has shown that in some cases, when SCF Scheme I is used with diffuse basis functions, the liquid-phase SCF calculation is unable to reach convergence (in particular, drastic fluctuations in the Mayer bond orders occur between the gas phase and the liquid phase when using SCF Scheme I, leading to poor SCF convergence). In cases where this occurs, using SCF Scheme II usually resolves this problem. (In contrast, this issue rarely arises when non-diffuse basis sets are used, as demonstrated by the results from above). To test the difference between the results obtained using SCF Scheme I and SCF Scheme II for basis sets containing diffuse basis functions, we carried out single-point calculations using both SCF Scheme I and SCF Scheme II, for all jobs with the suffix 1418 (SM6/MPWX/6-31+G* calculations) and 1419 (SM6/MPWX/6-31+G* calculations). For these calculations, we used gas-phase geometries calculated at the MPW25/MIDI! level of theory as input. For the test jobs described above, the computed solvation free energy is the difference between the free energy in the gas phase, evaluated at the MPW*X*/6-31+G**//MPW25/MIDI! level of theory, and the free energy in the aqueous phase, evaluated at the SM6/MPW*X*/6-31+G**//MPW25/MIDI! level of theory. The results of this test are summarized below.

For test jobs test3.*a*.1*b*, test4.*a*.1*b*, and test5.*a*.1*b*, (where *a* = 0, 25, 428, 517, 606, and 999 and *b* = 418 and 419) using SCF Scheme I and SCF Scheme II both led to converged, liquid-phase wave functions when the default value of 10^{-5} a.u was used for the SCF convergence tolerance. The computed solvation free energies for these test jobs are listed in the table on the following page.

test job	ISCRF=1[a]	ISCRF=2[b]	diff.
test3.0.1418	-8.080	-8.090	0.010
test3.25.1418	-8.220	-8.215	0.005
test3.428.1418	-8.316	-8.299	0.017
test3.517.1418	-8.372	-8.349	0.023
test3.606.1418	-8.431	-8.402	0.029
test3.999.1418	-8.701	-8.645	0.056
test4.0.1418	-10.014	-10.029	0.015
test4.25.1418	-10.092	-10.085	0.007
test4.428.1418	-10.147	-10.124	0.023
test4.517.1418	-10.185	-10.154	0.031
test4.606.1418	-10.227	-10.189	0.038
test4.999.1418	-10.426	-10.355	0.071
test5.0.1418	-5.159	-5.191	0.032
test5.25.1418	-5.178	-5.210	0.032
test5.428.1418	-5.186	-5.216	0.030
test5.517.1418	-5.161	-5.189	0.028
test5.606.1418	-5.161	-5.189	0.028
test5.999.1418	-5.150	-5.172	0.022
test3.0.1419	-8.011	-8.080	0.069
test3.25.1419	-8.161	-8.235	0.074
test3.428.1419	-8.262	-8.340	0.078
test3.517.1419	-8.305	-8.385	0.080
test3.606.1419	-8.355	-8.437	0.082
test3.999.1419	-8.607	-8.696	0.089
test4.0.1419	-10.091	-10.202	0.111
test4.25.1419	-10.195	-10.309	0.114
test4.428.1419	-10.268	-10.384	0.116
test4.517.1419	-10.297	-10.415	0.118
test4.606.1419	-10.336	-10.455	0.119
test4.999.1419	-10.551	-10.676	0.125
test5.0.1419	-5.154	-5.167	0.013
test5.25.1419	-5.192	-5.210	0.018
test5.428.1419	-5.188	-5.209	0.021
test5.517.1419	-5.175	-5.197	0.022
test5.606.1419	-5.179	-5.202	0.023
test5.999.1419	-5.131	-5.159	0.028
Max.			0.125
Ave.			0.050

[a] Aqueous solvation free energy, in kcal/mol, computed using SCF Scheme I. [b] Aqueous solvation free energy, in kcal/mol, computed using SCF Scheme II.

The table on the preceding page shows that for all test jobs where using SCF Scheme I and SCF Scheme II lead to converged, liquid-phase wave functions, the differences between the aqueous solvation free energies computed using SCF Scheme I and those computed using SCF Scheme II are very small in all cases. Even the maximum value of the difference, 0.125 kcal, is significantly smaller than the average error associated with any of our methods (~ 0.5 kcal in solvation free energies for neutral solutes).

For the remaining test jobs not shown in the table above (test1.*a*.1*b*, test2.*a*.1*b,* test6.*a*.1*b*, test7.*a*.1*b,* and test8.*a*.1*b,* where *a* = 0, 25, 428, 517, 606, and 999 and *b* = 418 and 419), using SCF Scheme I led to poor convergence in the SCF calculation. In all cases, when SCF Scheme II was used instead of SCF Scheme I, the liquid-phase wave functions converged quite readily (when the default value of 10^{-5} a.u was used for the SCF convergence tolerance).

Because of the small difference between the aqueous solvation free energies computed using SCF Scheme I and those computed using SCF Scheme II (see table above), and because for the majority of the test cases where diffuse basis functions were used, SCF Scheme I led to poor convergence, SCF Scheme I is no longer available for calculations that use diffuse basis-functions.

The input and output files (except the ones using SCF Scheme I with diffuse basis functions) for all of the calculations described above are located in the /sm6 directory. A listing of the input and output jobs in this directory is given in the table below.

Calculation type	test cases
Gas-phase and aqueous-phase (SCF Scheme I) geometry optimization using SM6/MPW*X*/MIDI!6D and SM6/MPW*X*/6-31G*	test*x*.*a*.1416, where *w* = 1 – 8, and *a* = 0, 25, 428, 417, 606, and 999. test*x*.*a*.1417, where *w* = 1 – 8, and *a* = 0, 25, 428, 417, 606, and 999.
Single-point, aqueous-phase calculation (SCF Scheme II) using SM6/MPW*X*/6-31+G* and SM6/MPW*X*/6-31+G**	test*x*.*a*.1418_2, where *w* = 1 – 8, and *a* = 0, 25, 428, 417, 606, and 999. test*x*.*a*.1419_2, where *w* = 1 – 8, and *a* = 0, 25, 428, 417, 606, and 999.

Description of the test suite for SM6T.

Three molecules were selected to test the SM6T atomic surface tension parameters for H, C and O containing compounds. In all cases, calculations were performed using MPW*X* (where *X* is the percentage of Hartree-Fock exchange). The three molecules are benzene (test1.*a*.*b*), furfural (test2.*a*.*b*), and water (test3.*a*.*b*). In the names below, *a* is the value of *X* in MPW*X* (between 0 and 999), and *b* is the type of temperature calculation. All of the computations were carried out using the 6-31+G** basis set using ISCRF=2 and ICMD=419. For more information on the effect of basis set and choice of SCRF scheme see the section above entitled "Description of the test suite for SM6."

The test suite for SM6T is found in the /sm6t directory. The directory contains tests of the three keywords associated with SM6T, namely the SOLK, the READK, and the AVGK keywords. Additionally it contains a file labeled "KELT" which must be placed in the *GAMESSPLUS* scratch directory. (The user sets the path for the GamessPlus scratch directory by modifying the rungms script, see the section entitled “Notes on Running *GAMESSPLUS*” for further details)

Calculation type	test cases
Single temperature calculations using the SOLK=*k* keyword.	test*x*.*a*.1.*k*, where *x* = 1 – 3, *a* = 0, 25, 428, 417, 606, or 999, and *k* = 273, 298, 348, or 373.
Multiple temperature calculations using individually computed electronic terms at each temperature using the READK keyword.	test*x*.*a*.2, where *x* = 1 – 3, and *a* = 0, 25, 428, 417, 606, or 999.
Multiple temperature calculations using scaled electronic terms at each temperature using the AVGK keyword.	test*x*.*a*.3, where *x* = 1 – 3, and *a* = 0, 25, 428, 417, 606, or 999.

The SOLK keyword was tested at four different temperatures 273 K, 298 K, 348 K, and 373 K. The results indicate that the temperature dependence of the electrostatic contribution is small even for polar compounds such as furfural. (dipole moment = 1.8 D) It also shows that the effect of the level of theory used does not affect the temperature dependence of the electrostatic contribution to the free energy of solvation. Shown below is a table of the values for ΔG_{ENP} computed in the test suite.

	Temperature (K)			
	273	298	348	373
% Hartree-Fock Exchange	bezene (test1)			
0	-1.74	-1.73	-1.73	-1.72
25	-1.82	-1.82	-1.81	-1.81
417	-1.84	-1.84	-1.83	-1.83
428	-1.84	-1.84	-1.83	-1.83
606	-1.90	-1.90	-1.89	-1.89
999	-1.99	-1.98	-1.98	-1.97
	furfural (test2)			
0	-5.90	-5.89	-5.86	-5.85
25	-5.74	-5.73	-5.70	-5.69
417	-5.56	-5.55	-5.53	-5.52
428	-5.59	-5.58	-5.56	-5.54
606	-5.45	-5.44	-5.41	-5.40
999	-5.33	-5.32	-5.30	-5.29
	water (test3)			
0	-5.62	-5.61	-5.59	-5.58
25	-5.70	-5.69	-5.67	-5.66
417	-5.74	-5.73	-5.71	-5.70
428	-5.75	-5.75	-5.73	-5.72
606	-5.81	-5.80	-5.78	-5.77
999	-5.98	-5.97	-5.95	-5.94

The READK keyword was tested at 7 temperatures, 273, 274, 298, 348, 373, 348, and 373 K. The repeated calculations at 343 K and 372 K were to confirm that the computed value for the electronic structure was not significantly affected by the previous temperature. The effect was negligable. The inclusion of 274 K right after 273 K was to confirm whether the default convergence criteria were sufficient, for GAMESSPLUS the default criteria usually are sufficiently tight to permit convergence of the solute wavefunction even when temperatures are closely spaced, which may be an issue because the READK uses the electronic structure of the solute in solution at one temperature as a starting guess for the electronic structure of the molecule at successive temperatures. Additionally the choice of temperatures was chosen to coincide with the temperatures computed using SOLK to confirm the accuracy of the READK option, the results are nearly identical. The table below shows the values for ΔG_{ENP} computed in the test suite using the READK option.

	Temperature (K)						
	273	274	298	348	373	348	373
% Hartree-Fock Exchange				bezene (test1)			
0	-1.74	-1.74	-1.73	-1.73	-1.72	-1.73	-1.72
25	-1.82	-1.82	-1.82	-1.81	-1.81	-1.81	-1.81
417	-1.84	-1.84	-1.84	-1.83	-1.83	-1.83	-1.83
428	-1.84	-1.84	-1.84	-1.83	-1.83	-1.83	-1.83
606	-1.90	-1.90	-1.90	-1.89	-1.89	-1.89	-1.89
999	-1.99	-1.99	-1.98	-1.98	-1.97	-1.98	-1.97
				furfural (test2)			
0	-5.90	-5.90	-5.89	-5.86	-5.85	-5.86	-5.85
25	-5.74	-5.74	-5.73	-5.70	-5.69	-5.70	-5.69
417	-5.56	-5.56	-5.55	-5.53	-5.52	-5.53	-5.52
428	-5.59	-5.59	-5.58	-5.56	-5.54	-5.56	-5.54
606	-5.45	-5.45	-5.44	-5.41	-5.40	-5.41	-5.40
999	-5.33	-5.33	-5.32	-5.30	-5.29	-5.30	-5.29
				water (test3)			
0	-5.62	-5.62	-5.61	-5.59	-5.58	-5.59	-5.58
25	-5.70	-5.70	-5.69	-5.67	-5.66	-5.67	-5.66
417	-5.74	-5.74	-5.73	-5.71	-5.70	-5.71	-5.70
428	-5.75	-5.75	-5.75	-5.73	-5.72	-5.73	-5.72
606	-5.81	-5.81	-5.80	-5.78	-5.77	-5.78	-5.77
999	-5.98	-5.98	-5.97	-5.95	-5.94	-5.95	-5.94

The AVGK keyword was tested at the same 7 temperatures as the READK keyword. Though the AVGK keyword computes the electronic contribution to the free energy of solvation by scaling the value from an average temperature the difference from results computed using READK are negligible The table below shows the values for ΔG_{ENP} computed in the test suite using the AVGK option.

	Temperature (K)						
	273	274	298	348	373	348	373
% Hartree-Fock Exchange				bezene (test1)			
0	-1.74	-1.74	-1.73	-1.73	-1.73	-1.73	-1.73
25	-1.82	-1.82	-1.82	-1.81	-1.81	-1.81	-1.81
417	-1.84	-1.84	-1.84	-1.83	-1.83	-1.83	-1.83
428	-1.84	-1.84	-1.84	-1.83	-1.83	-1.83	-1.83
606	-1.90	-1.90	-1.90	-1.89	-1.89	-1.89	-1.89
999	-1.99	-1.99	-1.98	-1.98	-1.97	-1.98	-1.97
				furfural (test2)			
0	-5.89	-5.89	-5.89	-5.87	-5.85	-5.87	-5.85
25	-5.73	-5.73	-5.72	-5.70	-5.69	-5.70	-5.69
417	-5.56	-5.56	-5.55	-5.53	-5.52	-5.53	-5.52
428	-5.59	-5.58	-5.58	-5.56	-5.55	-5.56	-5.55
606	-5.44	-5.44	-5.43	-5.42	-5.41	-5.42	-5.41
999	-5.33	-5.33	-5.32	-5.30	-5.29	-5.30	-5.29
				water (test3)			
0	-5.62	-5.62	-5.61	-5.59	-5.58	-5.59	-5.58
25	-5.70	-5.70	-5.69	-5.67	-5.66	-5.67	-5.66
417	-5.74	-5.74	-5.73	-5.71	-5.70	-5.71	-5.70
428	-5.75	-5.75	-5.75	-5.73	-5.72	-5.73	-5.72
606	-5.81	-5.81	-5.80	-5.78	-5.77	-5.78	-5.77
999	-5.98	-5.98	-5.97	-5.95	-5.94	-5.95	-5.94

Description of the test suite for EEQM

Three molecules were selected to test the EEQM calculations with a site-site representation of the QM/MM electrostatic interactions. In all cases, calculations were performed using the MPW*X* density functional (where *X* is the percentage of Hartree-Fock exchange) and the 6-31G(d) basis set. The three molecules are water (test1.*a*.*t*), methyl chloride (test2.*a*.*t*), and the transition state of the S_N2 reaction of methyl chloride with chloride anion (test3.*a*.*t*). The test suite for the EEQM calculations are found in the /eeqm directory, and outputs of the test calculations are found in the /eeqm/output directory.

Calculation type	test cases
EEQM energy calculation	test*x*.*a*.*t*, where *x* = 1–3 and *a*= 0, 25, 428, 606, or 999 and *t*=1 (gas phase) or 6 (embedded)
EEQM gradient calculation	test*x*.*a*.*t*, where *x* = 1–3 and *a*= 0, 25, 428, 606, or 999 and *t*=2 (gas phase) or 7 (embedded)
EEQM Hessian calculation	test*x*.*a*.*t*, where *x* = 1–3 and *a*= 0, 25, 428, 606, or 999 and *t*=3 (gas phase) or 8 (embedded)
EEQM CRK calculation with respect to coordinates	test*x*.*a*.*t*, where *x* = 1–3 and *a*= 0, 25, 428, 606, or 999 and *t*=4 (gas phase) or 9 (embedded)

EEQM CRK calculation with respect to electrostatic potential	test*x*.*a*.*t*, where *x* = 1–3 and *a*= 0, 25, 428, 606, or 999 and *t*=5 (gas phase) or 10 (embedded)

Verifying the Installation of *GAMESSPLUS* Using the Test Suite Results

A PERL script, named compare.pl, is provided in the *GAMESSPLUS* distribution (it is located in the *Input* directory). This script compares important output from test suite output files to corresponding output files in the *Output* directory. To use this script after running some or all of the test suite calculations, use the command

```
./compare.pl output-file(s)
```

where *output-file* is a test suite output file. Note that you can select more than one output file to compare, e.g., you can use the command

```
./compare.pl *.out > compare.out
```

or

```
./compare.pl *.out | more
```

to compare all the test calculations that you have completed.

Special Notes on Basis Sets

MIDI! basis set

The MIDI! basis (also called MIDIX) is similar to the 3-21G basis set with the following key differences:

- With the exception of Si, Br, and I, the MIDI! *s* and *p* orbitals have different exponents, whereas the 3-21G valence *s* orbitals are constrained to have the same exponents as the *p* orbitals.
- The bang (!) indicates that there is a *d* shell on every atom except for carbon and hydrogen atoms and a *p* shell on Li. There are no *p*-polarization functions on hydrogen either. Thus bang is part-way between star and no-star. In the MIDI! basis set, these *d* shells (and the *p* shell for Li) have been carefully optimized.

Note that the following convention regarding MIDI! basis sets: the original MIDI! basis set used 5 spherical harmonic components for each *d* shell; and MIDI! by itself denotes using 5-component *d* shells. However, we sometimes use 6 Cartesian components for *d* shells; this is denoted MIDI!6D. For emphasis, sometimes we use MIDI!5D to denote MIDI!.

The MIDI! basis set is not stored internally in *GAMESS*, but it can be used as an external basis set. For convenience, the MIDI! basis set is provided in *GAMESS* format in the file gmsplus-v4.6/Basis/midi-bang.bas for all 12 elements for which it is defined: H, Li, C, N, O, F, Si, P, S. Cl, Br, I.

MIDI! and MIDIX are synonymous, as are MIDI!6D and MIDIX6D.

cc-pVDZ basis set in *Gaussian*

There are two ways to write down the cc-pVDZ basis set. For example, for the hydrogen atom, it can be written as:

<u>**Form 1:**</u>

```
-H 0
S    4 1.00
  0.1301000000D+02  0.1968500000D-01
  0.1962000000D+01  0.1379770000D+00
  0.4446000000D+00  0.4781480000D+00
  0.1220000000D+00  0.5012400000D+00
S    1 1.00
  0.1220000000D+00  0.1000000000D+01
P    1 1.00
  0.7270000000D+00  0.1000000000D+01
```

or as:

<u>**Form 2:**</u>

```
-H 0
S    3 1.00
  0.1301000000D+02  0.1968500000D-01
  0.1962000000D+01  0.1379770000D+00
  0.4446000000D+00  0.4781480000D+00
S    1 1.00
  0.1220000000D+00  0.1000000000D+01
P    1 1.00
  0.7270000000D+00  0.1000000000D+01
```

These two ways of expressing the basis set give identical wave functions and energies because the basis functions span the same space. Furthermore, they lead to identical Mulliken charges. However, the wave function coefficients and the Löwdin charges are different. Our CM2 charge model was parameterized using the **Form 1** expression. Therefore, when you want to use the cc-pVDZ basis set, use the form that is provided in the file ccpVDZ.bas in the *Basis* directory.

Whenever we refer to the cc-pVDZ basis sets in conjunction with CM2 and/or SM5.42, we refer to the cc-pVDZ basis set as extended in "Accurate Dipole Moments from Hartree-Fock Calculations by Means of Class IV Charges," J. Li, J. Xing, C. J. Cramer, and D. G. Truhlar, *J. Chem. Phys.* **1999**, *111*, 885-892. This extended basis set is identical to the cc-pVDZ basis set for all elements except Br and I, where MIDI! is used instead.

6-31G* and 6-31+G* basis sets in CM*x* (*x* = 2 or 3) and SM*x* (*x* = 5.42, 5.43, 6, or 6T)

Whenever we refer to the 6-31G* and 6-31+G* basis sets in conjunction with CM2, SM5.42R, or SM5.42 calculations, we refer to these basis sets as extended in "A New Class IV Charge Model for Extracting Accurate Partial Charges from Wave Functions," J. Li, T. Zhu, C. J. Cramer, and D. G. Truhlar *J. Phys. Chem. A* **1998**, *102*, 1820-1831. The 6-31G* basis set was not available for I, so we used the MIDI!6D instead. The 6-31+G* basis set was also not available for I, so we used the MIDI!6D basis set augmented by diffuse *s* and *p* shells with an exponent of 0.03. For the user's convenience, the 6-31G* and 6-31+G* basis sets are included in the files 6-31GS.bas and 6-31PGS.bas, respectively. These files are located in the gmsplus-v4.6/Basis directory.

SCF Schemes

If ISCRF = 0, only a gas-phase calculation is performed, and this is the default. If ISCRF equals either 1 or 2, then the code does a liquid-phase SCF calculation. Two different SCF schemes have been implemented. The standard scheme, called scheme I, uses the current solution-phase bond order matrix values in the Fock or Kohn-Sham operator at every step of the iteration. Scheme I is chosen by setting ISCRF = 1.

Scheme II uses the gas-phase bond order matrix values at the current geometry to calculate CM2, CM3, or CM4 corrections of Löwdin or RLPA charges in solution. Scheme II and its physical meaning may be "justified" as follows. The modification to the Löwdin or RLPA charges is parameterized to fit to experimental gas-phase dipole moments. So one can assume that the same modification applies to the solvated molecule with the same geometry in the liquid phase as in the gas phase. The change of atomic partial charges due to the solvation effects is reflected by the change of Löwdin or RLPA charges, and all q_k^{M} are constants. This leads to a very simple expression for solvated Fock matrix elements:

$$F_{ij}^{S} = F_{ij}^{g} - \sum_{n} V_{k_n} (\mathbf{S}^{\frac{1}{2}})_{in} (\mathbf{S}^{\frac{1}{2}})_{nj} \tag{18}$$

where V_{k_n} is the reaction field acting on atom k according to Eq. (3), i, j, and n are basis functions, and k_n is the atom on which basis function n is centered.

Usually scheme I and scheme II give very similar results (see the SM6 test suite for examples). In principle (and in practice), scheme II converges to slightly different results, and thus it could be considered to be a different model. However, in our experience the difference is usually small, and scheme II sometimes has better convergence properties, particularly for larger basis sets and basis sets containing diffuse functions.

SCF Scheme III (ISCRF = 3) performs a liquid-phase SCRF based on user-provided atomic charges (ICREAD = 1). These atomic charges remain constant during the SCRF calculation. Analytical gradients have not been implemented for ISCRF = 3.

SM5.42, SM5.43, SM6, and SM6T calculations cannot use restricted open-shell wave functions, but unrestricted wave functions can be used. With the corrected implementation of UHF gradients, beginning with *GAMESSPLUS-v4.2*, SM5.42, SM5.43, SM6, and SM6T are available for both restricted and unrestricted wave functions, but they are still unavailable for restricted open-shell wave functions.

The SCF procedures used by GHO-AIHF algorithm for combining HF and molecular mechanics are different from the SCF scheme described above. The GHO wave function is only optimized over an active space consisting of fully QM basis functions and active hybrid basis functions. Therefore the conventional SCF procedure is modified to prevent the auxiliary basis functions on GHO boundary atom from mixing with all active basis functions. For convenience in terms of the implementation, the total Fock matrix and density matrix are still constructed in the atomic orbital basis in the usual way. The active Fock matrix and density matrix for solving Roothaan's equations in the GHO active space are then obtained by a certain basis transformation followed by dropping the corresponding auxiliary entries. The specific transformation between atomic basis functions and the orthogonalized hybrid basis functions are also dependent on the specific orthogonalization scheme used by the GHO-AIHF method. The detailed recipe of the modified SCF procedure for GHO-AIHF based different orthogonalization procedures can be found in the reference *PG04*.

Density Functional Methods Recommended for use with CM4 and SM6 in GAMESSPLUS.

The CM4 and SM6 model can be used with any density functional, as long as the density functional gives a reasonable electronic distribution for the molecule of interest. Shown in the table below is a list of density functional methods that are recommended for use with CM4 and SM6 in *GAMESSPLUS*. For the hybrid density functional theory methods recommended for use in *GAMESSPLUS*, the percent Hartree-Fock exchange for each functional is also given. Note that SM6T is applicable to all levels of theory to which SM6 may be applied.

Pure DFT functionals recommended for use with CM4 and SM6 in GAMESSPLUS. Note that the keyword HFE=0.00 should be specified in the $CM2 or $MNGSM namelist when these, or any other pure DFT functionals are used in conjunction with the CM4 or SM6 models .

Method	GAMESSPLUS Keyword	Reference(s)
BLYP	DFTTYP=BLYP (HFE=0.0 must be specified in the $CM2 or $GMSOL namelist only).	A. D. Becke, *Phys. Rev. A* **38**, 3098 (1988). C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988). B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chem. Phys. Lett.* **157**, 200 (1989).
G96LYP	DFTTYP=GLYP (HFE=0.0 must be specified in the $CM2 or $GMSOL namelist only).	P. M. W. Gill, *Mol. Phys.* 89, 433 (1996). C. Adamo and V. Barone, *J. Comp. Chem.* **19**, 419 (1998). C. Lee, W. Yang, and R. G. Parr, *Phys. Rev. B* **37**, 785 (1988). B. Miehlich, A. Savin, H. Stoll, and H. Preuss, *Chem. Phys. Lett.* **157**, 200 (1989).
mPWPW91	DFTTYP=MPW*X* (HFE=0.0 must be specified in both the $DFT namelist, and in the $CM2 or $GMSOL namelist).	C. Adamo and V. Barone, *J. Chem. Phys.* **108**, 664 (1998). K. Burke, J. P. Perdew, and Y. Wang, in *Electronic Density Functional Theory: Recent Progress and New Directions*, Ed. J. F. Dobson, G. Vignale, and M. P. Das (Plenum, 1998) J. P. Perdew, in *Electronic Structure of Solids* ' 91, Ed. P. Ziesche and H. Eschrig (Akademie Verlag, Berlin, 1991) 11 J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Phys. Rev. B* **46**, (1992). J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Phys. Rev. B* **48**, (1993). J. P. Perdew, K. Burke, and Y. Wang, *Phys. Rev. B* **54**, 16533 (1996).

Hybrid DFT functionals recommended for use with CM4 and SM6 in GAMESSPLUS. Note that the HFE keyword should be specified in the $CM2 or $GMSOL namelist when these, or any other hybrid DFT functionals are used in conjunction with the CM4 or SM6 models.

Method	Fraction HFE	*GAMESSPLUS* Keyword(s)	Reference(s)
B3LYP (version III VWN correlation functional)	0.200	DFTTYP=B3LYP3 (HFE = 0.200 must be specified in the $CM2 or $GMSOL namelist only)	P. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch, *J. Phys. Chem.* **98**, 11623 (1994).
mPW1PW91	0.250	DFTTYP=MPWX (HFE = 0.250 must be specified in both the $DFT namelist and the $CM2 or $GMSOL namelist)	C. Adamo and V. Barone, *J. Chem. Phys.* **108**, 664 (1998).
MPW1S	0.060	DFTTYP=MPWX (HFE = 0.060 must be specified in both the $DFT namelist and the $CM2 or $GMSOL namelist)	B. J. Lynch, Y. Zhao, and D. G. Truhlar, *J. Phys. Chem. A* **107**, 1384 (2003).
MPW1N	0.406	DFTTYP=MPWX (HFE = 0.406 must be specified in both the $DFT namelist and the $CM2 or $GMSOL namelist)	B. L. Kormos and C. J. Cramer, *J. Phys. Org. Chem.* **15**, 712 (2002).
MPW1K	0.428	DFTTYP=MPWX (HFE = 0.428 must be specified in both the $DFT namelist and the $CM2 or $GMSOL namelist)	B. J. Lynch, P. L. Fast, M. Harris and D. G. Truhlar *J. Phys. Chem. A* **104**, 4811 (2000).
MPW*X*	0.000 – 0.999	DFTTYP=MPWX (HFE = X must be specified in both the $DFT namelist and the $CM2 or $GMSOL namelist)	

Notation for the solvation models

1. Geometry optimized at level X/Y in the gas phase, followed by a single-point SM*x* solvation calculation at level W/Z, where W/Z is one of the choices supported by ICMD:

   SMx/W/Z//X/Y

2. If X/Y is the same as W/Z, then //X/Y may be substituted by //*g*,, where *g* denotes gas-phase:

   SM*x*/W/Z//*g*

   Previously, solvation calculations carried out using gas-phase geometries were denoted by including an "R" suffix after the name of the SMx model. Here, this older notation has been replaced with the notation above.

3. For a liquid-phase geometry optimization the //X/Y is dropped, and this calculation is denoted as follows:

   SMx/W/Z

   Previously, solvation calculations carried out using liquid-phase geometries were denoted by dropping the "R" suffix after the name of the SMx model. Here, we drop this suffix for all solvation calculations and use the notation described above.

Solvent Parameters

Solvent parameters for common organic solvents are tabulated in the *Minnesota Solvent Descriptor Database*. The latest version of this database is available at: http://comp.chem.umn.edu/solvation.

Program Distribution

The *GAMESSPLUS-v2008* program package consists of a compressed tar file called "gmsplusv2008.tar.gz". The top-level directory of this file system is "gmsplus-v2008"; the following is a list of the subdirectories and their contents:

Code Contains the modified *GAMESS* modules, new *GAMESSPLUS* modules, a modified rungms script, and a script to make new patch files.

bassto.src (*GAMESS)*	grd2b.src (*GAMESS)*	mthlib.src (*GAMESS)*
dft.src (*GAMESS)*	grd2c.src (*GAMESS)*	nddosrp.src (*GAMESSPLUS)*
dftexc.src (*GAMESS)*	inputa.src (*GAMESS)*	rhfuhf.src (*GAMESS)*
eeqm.src (*GAMESSPLUS)*	inputb.src (*GAMESS)*	smx.src (*GAMESSPLUS)*
gamess.src (*GAMESS)*	int1.src (*GAMESS)*	statpt.src (*GAMESS)*
gho.src (*GAMESSPLUS)*	int2a.src (*GAMESS)*	symorb.src (*GAMESS)*
ghodum.src (*GAMESSPLUS)*	mpcgrd.src (*GAMESS)*	rungms
grd1.src (*GAMESS)*	mpcint.src (*GAMESS)*	makepatch
grd2a.src (*GAMESS)*	mpcmol.src (*GAMESS)*	

Basis Contains files with basis set information

631GS.bas 631PGS.bas ccpVDZ.bas midi-bang.bas

Input	Test suite input files, SRP files, and submission and analysis scripts
test*w*.*x2y*.inp	where w = 1 – 8, x = 1 and 2, and y = 1, 2, 3, 6, 8, 9, 11, and 12
test9.02*y*.inp	where y = 1, 2, 3, 6, 8, 9, 11, and 12
test*w*.*qy*.inp	where w = 1, 2, 4 – 12, q = 0, 1, or 2, and y = 301, 302, 303, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 416, 417, 418, or 419.
test*x*.inp	where x = 13, 14, 16a – 16b, 17a – 14d, 20a – 20b, test21, test22, test23, test24, test25, test26, test27, test28, and test29
batch.2*x*	where x = 1, 2, 3, 8, 11, and 12 (submission scripts)
batch.0*x*	where x = 301, 302, 303, 311, 312, 313, 314, 320, 321, 322, 315, 316, 317, 318, 319, 416, 417, 418, or 419.
test17c.CDS	CDS_Param file for test13c
test17d.CDS	CDS_Param file for test13d
T20ASRP	AM1-SRP file for test20a
T20ACSRP	CM2-SRP file for test20a
T20BSRP	AM1-SRP file for test20b
compare.pl	script used to verify correct installation of *GAMESSPLUS*

Output	Test suite standard output files
test*w*.*x2y*.log	where w = 1 – 8, x = 1 and 2, and y = 1, 2, 3, 6, 8, 9, 11, and 12
test9.02*y*.log	where y = 1, 2, 3, 6, 8, 9, 11, and 12
test*w*.*qy*.log	where w = 1, 2, 4 – 12, q = 0, 1, or 2, and y = 301, 302, 303, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 416, 417, 418, or 419.
test*x*.log	where x = 13, 14, 16a – 16b, 17a – 14d, 20a – 20b, test21, test22, test23, test24, test25, test26, test27, test28, and test29

sm6	SM6 test suite input and output files
test*w*.*a*.1*b*.inp	where w = 1 – 8, a = 0, 25, 428, 517, 606, and 999, and b = 416 and 417
test*w*.*a*.1*b*.log	where w = 1 – 8, a = 0, 25, 428, 517, 606, and 999, and b = 416 and 417
test*w*.*a*.1*b*_2.inp	where w = 1 – 8, a = 0, 25, 428, 517, 606, and 999, and b = 418 and 419
test*w*.*a*.1*b*_2.log	where w = 1 – 8, a = 0, 25, 428, 517, 606, and 999, and b = 418 and 419

sm6t	SM6T test suite input (.inp) and output (.log) files
test*w*.*a*.1.*b*.inp	where w = 1 – 3, a = 0, 25, 428, 517, 606, and 999, and b = 273, 298, 348, and 373
test*w*.*a*.1.*b*.log	where w = 1 – 3, a = 0, 25, 428, 517, 606, and 999, and b = 273, 298, 348, and 373
test*w*.*a*.2.inp	where w = 1 – 3, a = 0, 25, 428, 517, 606, and 999
test*w*.*a*.2.log	where w = 1 – 3, a = 0, 25, 428, 517, 606, and 999
test*w*.*a*.3.inp	where w = 1 – 3, a = 0, 25, 428, 517, 606, and 999
test*w*.*a*.3.log	where w = 1 – 3, a = 0, 25, 428, 517, 606, and 999

eeqm	EEQM test suite input (.inp) and output (.log) files
test*w*.*a*.*t*.inp	where w = 1 – 3, a = 0, 25, 428, 606, and 999, and t = 1 – 10
test*w*.*a*.*t*.log	where w = 1 – 3, a = 0, 25, 428, 606, and 999, and t = 1 – 10

Patch Patch files for *GAMESS* version March 24, 2007 R6 and execution script to generate new patch files for other versions of *GAMESS*

bassto.src.patch
dft.src.patch
dftexc.src.patch
gamess.src.patch
grd1.src.patch
grd2a.src.patch
grd2b.src.patch
grd2c.src.patch
inputa.src.patch
inputb.src.patch
int1.src.patch
int2a.src.patch
mpcgrd.src.patch
mpcint.src.patch
mpcmol.src.patch
mthlib.src.patch
rhfuhf.src.patch
statpt.src.patch
symorb.src.patch

gmsplus_solubility Files required to run calculations with the solubility utility program

/lib (contains source file solubility.pl)	install.pl	midi-bang.bas
pentane1.inp	pentane2.inp	
pentane2.log	pentane2.log	

gmsplus_soil Files required to run calculations with the soil sorption utility program

/lib (contains source file soil.pl)
test1.inp
test1.log
install.pl
test2.inp
test2.log
midi-bang.bas
test3.inp
test3.log

A Note on *GAMESS* Versions

In updating *GAMESS*, it is important to understand *GAMESS* versions. *GAMESS* versions are primarily determined by the date printed in the box at the top of the output. However, that is not a unique version indicator. Sometimes individual routines are changed without changing that date. A complete (unique) specification of a version of *GAMESS* therefore requires specifying not only the date in the box at the top of the output but also whether or not any routines have dates later than this and, if so, which routines and what dates.

Whereas previous versions of *GAMESSPLUS* were primarily version specific (sometimes a few versions of *GAMESS* could be used with a particular *GAMESSPLUS* implementation), a great deal of effort has gone into making *GAMESSPLUS* version 2008 compatible with a number of presently existing and any future *GAMESS* implementations.

The present version of *GAMESSPLUS* is designed for use with the March 24, 2007 (R6) version of *GAMESS*. For users with this version of *GAMESS*, updating and compiling can be accomplished using the "Standard Method for Updating and Compiling *GAMESSPLUS*". For users with other versions of GAMESS, updating and compiling can usually be accomplished using the "Makepatch Method for Updating and Compiling *GAMESSPLUS*".

Unfortunately, older versions of *GAMESS* are not always available from the authors of *GAMESS*. If users have difficulties due to this version problem, they should contact the *GAMESSPLUS* developers for assistance.

Standard Method for Updating and Compiling *GAMESSPLUS*

Instructions for modifying the March 24, 2007 (R6) version of *GAMESS* to produce a *GAMESSPLUS* code that can carry out CM2, CM3, CM4, SM5.42, SM5.43, SM6, SM6T, NDDO-SRP, CM2-SRP and EEQM calculations are summarized as follows (for all other versions of *GAMESS*, use the Makepatch Method for Updating and Compiling *GAMESSPLUS*):

1. Obtain the March 24, 2007 (R6) version of *GAMESS* from Iowa State University and obtain *GAMESSPLUS-v2008* from the University of Minnesota.

2. Place both files in the same directory (e.g.; *~jsmith/gamessplus*), and gunzip and untar them. The files will untar into the directories *~jsmith/gamessplus/games* and *~jsmith/gamessplus/gmsplus-v2008*. Move into the gmsplus directory and execute *modgms* (*./modgms*). This script might prompt you for a few bits of system information, and then it will compile *GAMESSPLUS*. If *GAMESSPLUS* compiled successfully, the script should end with something like:

```
------------------ done with all compilations -------------------
Fri Dec 10 13:55:36 CST 2004
456.9u 100.3s 10:17 90% 55+53k 0+0io 14356pf+0w
Fri Dec 10 13:55:36 CST 2004
GAMESS will be linked into the executable image gamessplus.2008.x.
chdir object
xlf -o ../gamessplus.2008.x -q64 -Wl,-m -Wl,-bloadmap:../lked.map gamess.o unport.o …
```

```
.
.
. messages (but no errors) from linker
.
unset echo
0.3u 0.3s 0:06 11% 2+92k 0+0io 4522pf+0w
## End of GAMESPLUSS INSTALL ##
```

The executable gamessplus.2008.x should now be in the ~jsmith/gamessplus/gamess directory.

Makepatch Method for Updating and Compiling *GAMESSPLUS*

Instructions for modifying versions of *GAMESS* other than the March 24, 2007 (R6) version to produce a *GAMESSPLUS* code that can carry out CM2, CM3, CM4, SM5.42, SM5.43, SM6, SM6T, NDDO-SRP, CM2-SRP, and EEQM calculations are summarized as follows:

1. Obtain *GAMESS* from Iowa State University and obtain *GAMESSPLUS-v2008* from the University of Minnesota.

2. Place both files in the same directory (e.g.; *~jsmith/gamessplus*), and gunzip and untar them. The files will untar into the directories *~jsmith/gamessplus/gamess* and *~jsmith/gamessplus/gmsplus-v2008*.

3. Move into the gmsplus/Code directory and execute makepatch (*./makepatch*).

4. Move into the gmsplus directory and execute *modgms* (*./modgms*). This script might prompt you for a few bits of system information, and then it will compile GAMESSPLUS. If GAMESSPLUS compiled successfully, the script should end with something like:

```
------------------ done with all compilations --------------------
Fri Dec 10 13:55:36 CST 2004
456.9u 100.3s 10:17 90% 55+53k 0+0io 14356pf+0w
Fri Dec 10 13:55:36 CST 2004
GAMESS will be linked into the executable image gamessplus.2008.x.
chdir object
xlf -o ../gamessplus.2008.x -q64 -Wl,-m -Wl,-bloadmap:../lked.map gamess.o unport.o …
.
.
. messages (but no errors) from linker
.
unset echo
0.3u 0.3s 0:06 11% 2+92k 0+0io 4522pf+0w
## End of GAMESPLUSS INSTALL ##
```

The executable gamessplus.2008.x should now be in the ~jsmith/gamessplus/gamess directory.

Manually Updating and Compiling *GAMESSPLUS*

If there was a problem with either of the installations outlined above, you can make all the modifications by hand, following the next six steps detailed below.

1. In the gamess/source directory of the *GAMESS* distribution, check the first line of each of the *GAMESS* files listed below for the latest modification date:

bassto.src	September 05, 2005
dft.src	August 20, 2007
dftexc.src	August 20, 2007
gamess.src	December 7, 2007
grd1.src	August 20, 2007
grd2a.src	August 20, 2007
grd2b.src	December 22, 2006
grd2c.src	December 22, 2006
inputa.src	December 7, 2007
inputb.src	May 21, 2007
int1.src	December 7, 2007
int2a.src	December 7, 2007
mpcgrd.src	December 12, 2003
mpcint.src	November 6, 2006
mpcmol.src	November 6, 2006
mthlib.src	August 20, 2007
rhfuhf.src	December 7, 2007
statpt.src	August 20, 2007
symorb.src	August 20, 2007

For each of these files in the *GAMESS* distribution whose latest modification date matches the date above, simply replace the whole file in the *GAMESS* distribution with the corresponding file in the gmssplus-v2008/Code directory. If the latest modification date for one of the above files in the *GAMESS* distribution is later than the corresponding date above, then the modifications should be carried out line-by-line. Note that all modifications to the *GAMESS* code start with "CGMSPLUSSTR" and end with "CGMSPLUSEND". Thus, the corresponding file in the gmsplus-v2008/Code directory can be used as a template for modifying the file in the *GAMESS* distribution. If any problems are encountered in this procedure, contact the *GAMESSPLUS* developers (see http://comp.chem.umn.edu/gamessplus/ for contact information).

2. Copy gmsplus-v2008/Code/smx.src, gmsplus-v2008/Code/nddosrp.src, gmsplus-v2008/Code/ghodum.src, and gmsplus-v2008/Code/eeqm.src to the gamess/source directory.

3. Go to the gamess/tools directory and copy file *actvte.code* to file *actvte.f*, then modify file *actvte.f* based on the instructions given there. Compile *actvte.f* to generate an executable file called *actvte.x* (f77 actvte.f –o actvte.x).

4. Go to the gamess/ddi directory and edit the compddi script. In this file, modify the "set TARGET=" line to list the appropriate machine type. Compile the distributed data interface

portion of the *GAMESS* code by typing ./compddi. When this compilation is completed, move the file named ddikick.x to the games s directory.

5. In the gamess directory, edit the three compile script files in the gamess directory: *compall*, *comp,* and *lked*. In all three files, modify the “`set TARGET=`” line to list the appropriate machine type and the “`chdir`” line to list the directory name where *GAMESS* resides. Add the lines:

```
comp smx
comp nddosrp
comp ghodum
comp eeqm
```

after the “`comp zmatrix`” line in *compall*. Finally, add “`smx.o nddosrp.o ghodum.o eeqm.o`” to the list of object files in file *lked* (near the end of the file). On some platforms, the FORTRAN optimization level may have to be lowered for some of the source files. This can be determined by running the full test suite and checking for any discrepancies between results obtained and those provided in the distribution (see the next section). For example, for SGI computers the optimization level has been set to O3 in the compilation script *comp* for most of the source files. For this setting the gradient evaluations in test*w*.*x*2*y* (where *w* = 1 – 8, *x* = 1 or 2, and *y* = 2, 3, 11, and 12) produces erroneous results, and the geometry optimizations in test cases test13 and test14 both fail. Reducing the optimization level for grd1.src to O1 solves this problem. The optimization level is set in the compilation script *comp*. The following is the portion of the script *comp* that sets compile options for SGI machines, modified to lower the optimization level for grd1.src:

```
if (($TARGET == sgi32) || ($TARGET == sgi64)) then
   set OPT='-O3'
   if ($MODULE == grd2c)  set OPT='-O2'
   if ($MODULE == rxncrd) set OPT='-O2'
   if ($MODULE == grd1) set OPT='-O1'
   set ARCH='-r12000'
                         set INTS=' '
   if ($TARGET == sgi64) set INTS='-i8'
   set FLAGS='-64 -mips4 -automatic -G0 -woff 2290 -OPT:Olimit=0'
   set echo
   f77 -c $OPT $ARCH $FLAGS $INTS $MODULE.f
   unset echo
endif
```

6. Type *compall* and *GAMESS* should start compiling. After *GAMESS* compiles, type *lked*. An executable named *gamess.00.x* will be created. Optionally, typing *lked name version*, where *name* is a filename, and *version* is a number string, produces a *GAMESS* executable named name.version.x (e.g. typing *lked gamessplus 01* gives an executable named *gamessplus.01.x*). Note that modifying *name* will require appropriate modifications to the rungms script; modifying *version* requires the user to type *sh rungms version* to run *GAMESS*.

Note: The compilation of *CHARMM* with *GAMESSPLUS* as a combination package with an integrated executable is supported by a utility package called *CGPLUS*. The step-by-step instructions for modifying *GAMESS*, *GAMESSPLUS*, and *CHARMM* to create the

CHARMM/GAMESSPLUS combination package to perform QM/MM calculations at the HF *ab initio* level with the GHO boundary treatment can be found in the chapter "*Compiling CHARMM with GAMESSPLUS*" of the *CGPLUS-v2008 User Manual* (http://comp.chem.umn.edu/cgplus).

Platforms

GAMESSPLUS-v2008 has been tested with the March 24, 2007 (R6) version of *GAMESS* on the following platforms:

- IBM pSeries 690 and pSeries 655 Nodes (Power 4 processors), running AIX version 5.3, compiled with the XL Fortran compiler version 10.1 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'ibm64')
- Netfinitiy Linux cluster running RedHat Enterprise Linux 3, kernel version 2.4.21, compiled with the g77 compiler, version 3.2.3 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'linux32')
- SGI Altix XE 1300, running SUSE Linux Enterprise Server 10 (x86_64), kernel 2.6.16, compiled with the Intel Fortran compiler, version 10.1 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'linux-ia64')

Note that the GHO-AIHF module has been tested with *CGPLUS-v2008* and *CHARMM* version c30a1 on the IBM SP and the IBM Regatta machines.

Previous versions of *GAMESSPLUS* (all parts except the GHO-AIHF module) have been tested on the following additional platforms:

- IBM pSeries 690 and pSeries 655 Nodes (Power 4 processors), running AIX version 5.2, compiled with the XL Fortran compiler version 9.1 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'ibm64')
- Netfinitiy Linux cluster running Red Hat Linux, kernel version 2.4.21, compiled with the g77 compiler, version 3.2.3 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'linux-pc')
- SGI-Altix with Itanium 2 processors, running Red Hat Linux, kernel 2.4.21, compiled with the Intel Fortran compiler, version 8.0 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'linux-ia64')
- IBM SP with WinterHawk+ (Power 3 processors) nodes, running AIX version 5.1, compiled with the XL Fortran compiler version 7.1.12 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'ibm64')
- IBM SP with NightHawk (Power3 processors), running AIX version 5.1, compiled with the XL Fortran compiler version 7.1.12 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'ibm64')
- IBM Regatta (Power 4 processors, that is, pSeries 690 and pSeries 655 nodes), running AIX version 5.1, compiled with the XL Fortran compiler version 7.1.12 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'ibm64')
- Netfinity Linux cluster running RedHat Linux, version 7.2 and kernel version 2.4.9, compiled with the g77 compiler, version 3.2 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to 'linux-pc')

- Sun Blade 2000 with UltraSparc III processors, running Solaris 8, compiled with Forte Developer 7 Fortran version 7.0 compiler (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to ‘sun64’)
- SGI-Altix 3000 with Madison processors, running RedHat Linux, and compiled with the Intel Fortran compiler, version 8.0 (in the *comp*, *compall*, and *lked* compilation scripts, $TARGET was set to ‘linux-ia64’)

Versions previous to version 3.9 have been successfully tested on:

- SGI Origin 3800 with R14000 CPUs, running IRIX 6.5.12f, compiled with MIPSpro compiler version 7.3.1.2m (in the compilation scripts, *comp*, *compall*, and *lked*, the variable $TARGET was set to ‘sgi64’)

GAMESSPLUS Solubility Utility

Executive Summary

The GAMESSPLUS Solubility Utility program is a utility program for GAMESSPLUS that predicts the solubility of a solute in a given solvent. It utilizes the thermodynamic relationship (see Thompson J. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*, 1661) that exists between the solubility, free energy of solvation, and the pure-substance vapor pressure of a solute, which is shown below for a given solute, A, in a liquid solvent B:

$$S = \left(P_{\mathrm{A}}^{\bullet} \big/ P^{\mathrm{o}} \right) \exp\left[-\Delta G_{\mathrm{S}}^{\mathrm{o}} \big/ RT \right] \tag{1}$$

In this equation, S is the solubility of solute A in solvent B, $P_{\mathrm{A}}^{\bullet}$ is the equilibrium vapor pressure of solute A of a pure solution of A, P^{o} is the pressure of an ideal gas for a given standard-state (a 1 molar standard-state at 298 K is used in this calculation for all phases; therefore P^{o} is 24.45 atm), $\Delta G_{\mathrm{S}}^{\mathrm{o}}$ is the standard-state free energy of solvation of solute A in solvent B, R is the universal gas constant, and T is temperature. This relationship is valid on the condition that all phases in question are ideal (i.e., the saturated solution of the solute in a given solvent, the solute in the gas-phase, and the pure solution of the solute). It has been shown to be valid for a diverse set of liquid and solid solute data in water solvent (see Thompson J. D.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **2003**, *119*, 1661).

The GAMESSPLUS Solubility Utility program uses various features in the GAMESSPLUS program to calculate solubilities. In particular, it uses GAMESSPLUS to calculate $\Delta G_{\mathrm{S}}^{\mathrm{o}}$ for a given solute-solvent system using the SM5.42 or SM5.43 continuum solvation model. It also uses GAMESSPLUS to calculate $P^{\bullet}$, which is defined by the free energy of self-solvation of the solute (the free energy of solvation of the solute in a pure solution of itself), using SM5.42 or SM5.43. Instead of using SM5.42 or SM5.43 to predict the free energy of self-solvation, this utility program can also take as input a user-defined value for the pure-solute vapor pressure.

The solubility of a solute can be calculated with SM5.42 for the following restricted and unrestricted wave functions:

HF/MIDI!
HF/MIDI!6D
HF/6-31G*
HF/6-31+G*
B3LYP/6-31G*
HF/cc-pVDZ
HF/AM1
HF/PM3
B3LYP/MIDI!

and with SM5.43 with HF/6-31G* and B3LYP/6-31G* (using either a restricted or unrestricted formalism).

This section and several subsections below of this manual provide a stand-alone introduction to the GAMESSPLUS Solubility Utility program for users who just want to calculate solubilities. Such users should first install GAMESSPLUS according to the instructions given in the sections entitled *Notes on Running GAMESSPLUS* and *Updating and Compiling GAMESS* of this manual. Users of this utility program should give the appropriate references described in the section entitled *GAMESSPLUS Reference*. Then, everything else they need to know is in the following few sections.

The SM5.42 and SM5.43 continuum solvation models

SM5.42 and SM5.43 are universal Solvation Models that use SM5 functional forms for atomic surface tensions (hence the first three characters in the name of the method are SM5), class IV point charges (hence .4 comes next) of the CM2 or CM3 type (hence 2 or 3), and are parameterized for rigid (hence R) gas-phase geometries.

These solvation models provide a way to calculate electronic wave functions in the liquid phase and free energies of solvation. The SM*x* (*x* = 5.0, 5.2, 5.4, 5.42, and 5.43) solvation models are universal, i.e., they are parameterized for water and any organic solvent. With universal models, one can calculate solvation free energies of a solute in two different solvents (e.g., water and 1-octanol) and use the results to calculate the partition coefficient. In this part of the program, the standard-state free energy of solvation $\Delta G_S^o(\mathbf{R})$ is calculated at the gas-phase geometry and is given by two components:

$$\Delta G_S^o(\mathbf{R}) = \Delta G_{EP} + G_{CDS} \tag{1}$$

where

$$\Delta G_{EP} = \Delta E_E + G_P \tag{2}$$

In this equation, ΔG_{EP} is the bulk electrostatic component of the solvation free energy; it is the sum of the polarization energy G_P (representing favorable solute-solvent interactions and the associated solvent rearrangement cost) and the distortion energy ΔE_E (the cost of distorting the solute electronic charge distribution to be self-consistent with the solvent electric polarization). The quantity, G_{CDS} accounts for first-solvation-shell effects. ΔG_{EP} is determined by a self-consistent reaction field (SCRF) calculation, which allows the solvent-induced change in the solute electronic wave function to be optimized variationally.

The G_{CDS} term is not a self-consistent term; it has no effect on the solute electronic wave function. G_{CDS} is given by

$$G_{CDS} = \sum_k A_k \, \sigma_k \tag{3}$$

where A_k is the solvent accessible surface area of atom *k* (this depends on the solute's 3-D geometry and is calculated by the Analytical Surface Area (ASA) algorithm as described in D. A. Liotard, G. D. Hawkins, G. C. Lynch, C. J. Cramer, and D. G. Truhlar *J. Comput. Chem.* **1995**, *16*, 422–440, and as included in recent versions of AMSOL, GAMESSPLUS, HONDOPLUS, OMNISOL, MN-GSM, ZINDO-MN, and DGSOL), and σ_k is the atomic surface tension of atom *k*. The atomic surface tension σ_k is itself a function of the solute's 3-D geometry and a small set of solvent descriptors (discussed in detail below). The linear parameters in the functional forms for the atomic surface tensions are called the surface tension coefficients. The functional forms themselves are the same in all SM*x* (*x* = 5.42 and 5.43) models.

Usage

Input for the GAMESSPLUS Solubility Utility program is similar to the input for a typical GAMESSPLUS calculation, with some additional modifications discussed below. In order to use the utility program, some knowledge of how to run a GAMESS/GAMESSPLUS calculation is required. A brief description of GAMESS/GAMESSPLUS input that is important to run this utility program is given below. For more detailed information for setting up GAMESS calculations than is given here, see the documentation that comes with the GAMESS distribution, particularly the first two sections of the GAMESS user's manual. These sections specify all aspects of a GAMESS input file.

gamess/gamessplus Input

Most GAMESS input is entered in a pseudo-namelist format; a namelist can be thought of as a particular group of keywords. The title of this group, the namelist name, is prefixed by a '$' that must be entered in column 2 of an input file (e.g. $SCF). A namelist accepts as arguments a number of possible keywords that are entered after the namelist name in a free format style that can span over multiple lines. Each keyword accepts either character or numeric values. A namelist is terminated with a $END. Only input between a namelist name and its corresponding $END will be read from a GAMESS input file.

Below is a brief summary of GAMESS namelists and their keywords that are most important to users who want to compute solubilities with the GAMESSPLUS Solubility Utility program. In this section, the namelist name is followed by a brief description, and below the namelist name is a set pertinent keywords and their options:

Namelist $CONTRL			group of global settings for the calculation
SCFTYP	=	RHF	restricted Hartree-Fock calculation (default)
	=	UHF	unrestricted Hartree-Fock calculation
RUNTYP	=	ENERGY	SCF evaluation (default)
MAXIT	=	*N*	*N* is the maximum number of SCF cycles (default = 30)
ICHARG	=	*N*	*N* is the molecular charge (default = 0)
MULT	=	*N*	*N* is the multiplicity of the electronic state (default = 1)
COORD	=	UNIQUE	input symmetry-unique Cartesian coordinates (default)
	=	CART	input all Cartesian coordinates
	=	ZMT	input *Gaussian*-style *Z*-matrix internal coordinates
ISPHER	=	–1	use Cartesian basis functions (e.g. 6D, 10F; default)
	=	1	use spherical harmonic basis functions (e.g. 5D, 7F)

Note: Symmetry should not be used with GAMESSPLUS calculations. The best way to input Cartesian coordinates is to set COORD = UNIQUE (which prevents coordinate rotation) and to set the symmetry group of the molecule to C1 in $DATA (see below).

Namelist $SYSTEM		information for controlling the computer's operation
MEMORY=	*N*	*N* is the maximum memory the job can use in words (default = 1 000 000)

Namelist $DFT		density functional theory (DFT) input
DFTTYP=	*N*	*N* is BLYP, which requests a BLYP calculation, or *N* is B3LYP3, which requests a B3LYP (as it is implemented in *Gaussian* and *HONDOPLUS*, i.e., using version III of the VWN correlation functional) calculation, or *N* is B3LYP5, which requests a B3LYP calculation, but using version V of the VWN correlation functional.
HFE=	*N*	Defines the fraction of Hartree-Fock exchange *N* to be used when the MPW*X* functional is used. This keyword must be specified when MPW*X* is used. For CM3 and SM5.43 calculations that use MPW*X* (ICMD = 315 – 319), the HFE = *N* must also be specified in the $GMSOL or $CM2 namelist.
METHOD=	*M*	*M* is GRID to request grid-based DFT calculation (default)

Note: All SM5.42 and SM5.43 parameterizations that are based on the B3LYP method use version III of the VWN correlation functional, so these types of calculations must be carried out using 'DFTTYP=B3LYP3', not 'DFTTYP=B3LYP5'. In addition, all SM5.42 and SM5.43 methods that are based on the BLYP method and/or the B3LYP (B3LYP3) method were parameterized using grid-based DFT, so 'METHOD=GRID' (the default for METHOD) should always be used.

Namelist $BASIS			input of available standard basis sets
GBASIS	=	N31	Pople's N-31G basis sets
	=	AM1	AM1 model Hamiltonian
	=	PM3	PM3 model Hamiltonian
NGAUSS	=	*N*	the 'N' of N31 *(*e.g. *N* = 6 for 6-31G)
NDFUNC	=	*N*	*N* is the number of polarizing *d* subshells on heavy atoms (*N* = 1 for 6-31G*; default = 0)
DIFFSP	=	.TRUE.	adds a diffuse *sp* shell to the basis set (default = .FALSE.)

Note: The *d* subshells have 5 functions if ISPHER=1, and they have 6 functions if ISPHER=-1, where ISPHER is defined in the $CONTRL data group.

Examples:

```
6-31G* :     $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 $END
6-31+G* :    $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 DIFFSP=.TRUE. $END
```

Namelist $DATA input molecule information

Input consists of the following information:

line 1: title line

line 2: symmetry group (always enter `C1` for GAMESSPLUS calculations)

line 3+: molecular coordinates; for Cartesian input, each line consists of the following: atom label, the atom's nuclear charge (which should be a floating-point value, i.e., 1.0 for H), and the atom's *x*, *y*, and *z*-coordinate. If a general basis set is required, then the basis set for each atom follows the atom's coordinates. See examples in the *Input and Output Examples* section and in the test suite.

Input Specific to the GAMESSPLUS Solubility Utility

The GAMESSPLUS Solubility Utility adds two new namelists to GAMESSPLUS: the $VAPOR namelist and the $DGS namelist. Both namelists define the specifics for a particular type of SM*x* calculation. For a given solute A in a given liquid solvent B, the $DGS namelist provides the specifics for calculating the standard-state free energy of solvation of the solute A in solvent B, ΔG_S^o. The $VAPOR namelist provides the specifics for calculating the standard-state free energy of self-solvation of solute A, which defines the pure-solute vapor pressure of A, $P^\bullet$. The $VAPOR namelist can also supply a user-given pure-solute vapor pressure of solute A in several different units; in this case, a free energy of self-solvation calculation of solute A is not carried out. These two namelists are used to specify the SM5.42 or SM5.43 parameter set, which corresponds to a particular wave function for which the SM5.42 or SM5.43 parameters were optimized, (the available wave functions are shown in the section entitled *Executive Summary* above). They also specify the solvent descriptors for the two solvents used in the calculation (i.e., the solvent descriptors for the pure solution of the solute required for the calculation of the free energy of self-solvation and the solvent descriptors of the solvent used in the calculation of ΔG_S^o). For example, for the solubility calculation of *n*-pentane in methanol, the $VAPOR namelist would be used to specify the solvent descriptors of *n*-pentane, and the $DGS namelist would be used to specify the solvent descriptors of methanol. Below are descriptions of the keywords used in both $VAPOR and $DGS.

ISCRF Controls the type of calculation to be performed:

ISCRF=1 Calculates the free energy of solvation of a solute using the SM5.42 or SM5.43 solvation model and SCF Scheme I (see SCF Schemes). This is the default value for ISCRF. SCF Scheme I is not available for methods that employ diffuse basis functions. This is the default for all other methods.

ISCRF=2 Calculates the free energy of solvation of a solute using the SM5.42 or SM5.43 solvation model and SCF Scheme II (see SCF Schemes)

ICDS Selects the set of coefficients to use for the SM5.42 or SM5.43 solvation model. Coefficients have been optimized for specific wave functions, as listed below:

Value of ICDS	Wave function (available methods)
1	SM5.42/HF/MIDI!
2	SM5.42/HF/MIDI!6D
3	SM5.42/HF/6-31G*
6	SM5.42/B3LYP/MIDI!
8	SM5.42/HF/6-31+G*
9	SM5.42/HF/cc-pVDZ
11	SM5.42/AM1
12	SM5.42/PM3
303	SM5.43/HF/6-31G*
313	SM5.43/B3LYP/6-31G*
315	SM5.43/MPW*X* /MIDI!
316	SM5.43/MPW*X* /MIDI!6D
317	SM5.43/MPW*X* /6-31G*
318	SM5.43/MPW*X* /6-31+G*
319	SM5.43/MPW*X* /6-31+G**
ICDS = 0	All atomic surface tension coefficients equal zero (G_{CDS} = 0)

By default, ICDS is set equal to 1. Note that the B3LYP method in the table above corresponds to the one requested by the 'DFTTYP=B3LYP3' keyword in the $DFT data group; see the section entitled *Notes on GAMESS input* above.

IAQU	Determines the solvent type:
IAQU=0	Organic solvent (additional solvent data must be input, see below)
IAQU=1	Aqueous solvent (default, no additional solvent data is required)

Solvent Descriptors: If IAQU=0, then the solvent properties are specified by a series of solvent property (descriptor) values:

Dielec dielectric constant, ε, of solvent (default is the value for water, 78.3)

SolN index of refraction at optical frequencies at 293 K, n_{20}^{D} (default is 1.0)

SolA Abraham's hydrogen bond acidity, $\Sigma\alpha_2^H$ (default is 0.0)

SolB Abraham's hydrogen bond basicity, $\Sigma\beta_2^H$ (default is 0.0)

SolG $\gamma = \gamma_m / \gamma^{\text{o}}$ (default is 0.0), where γ_m is the macroscopic surface tension at air/solvent interface at 298 K, and γ^{o} is 1 cal·mol^{-1}·Å^{-2} (note that 1 dyne/cm = 1.43932 cal·mol^{-1}·Å^{-2})

SolC aromaticity, ϕ : the fraction of non-hydrogenic solvent atoms that are aromatic carbon atoms (default is 0.0)

SolH electronegative halogenicity, ψ : the fraction of non-hydrogenic solvent atoms that are F, Cl or Br (default is 0.0)

For a desired solvent, these values can be derived from experiment or from interpolation or extrapolation of data available for other solvents. Solvent parameters for common organic solvents are tabulated in the *Minnesota Solvent Descriptor Database*. The latest version of this database is available at: http://comp/chem.umn.edu/solvation. Note that unless IAQU is set to 1, these solvent descriptors are *required* input. The calculation will not run if they are not specified.

HFE Defines the fraction of Hartree-Fock (HF) exchange to be used when ICMD = 315 to ICMD = 319. For example, if a calculation of CM3 charges by the mPW1PW91/MIDI! method were to be carried out, then ICMD would be set to 315, and HFE would be set to 0.25. The HFE keyword only defines the fraction of HF exchange for a given calculation in the charge and solvation model portion of *GAMESSPLUS*, not in the entire electronic structure part of the *GAMESSPLUS* program (thus, when using the MPW*X* functional, the HFE keyword must also be specified in the $DFT namelist). This keyword is only active when ICMD = 315 – 319, and in these instances HFE must be explicitly specified by the user in both the $DFT and the $CM2 or $GMSPLUS namelists. If ICMD = 315 – 319 and HFE is not defined by the user, *GAMESSPLUS* will terminate.

Input Options Specific to the $VAPOR namelist

There are three other options specific to the $VAPOR namelist. These options are given below.

DENSITY Specifies the density of the pure solution of the solute (in units of mol/L)

This keyword is REQUIRED for all solubility calculations.

PRESSSURE Defines a user-specified pure-solute vapor pressure of the solute, $P^{\bullet}$ in units defined by the user (with the UNITS keyword, described below)

This keyword allows the user to enter a value for the pure-solute vapor pressure of the solute, instead of calculating it. When this keyword is given, the keywords, DIELEC, IAQU, SolA, SolB, SolC, SolG, SolH, SolN, and ICDS are not required in $VAPOR. The default units for the vapor pressure are Pascals, however, the pressure can be given in other units specified by the UNITS keyword, see below.

UNITS Specifies the units of the user-supplied vapor pressure,

The allowed values of this keyword are Pa, atm, bar, and torr, for pressure in units of Pascals, atmospheres, bars, and torr, respectively. The default for this keyword is UNITS=Pa.

Test Calculations

Two test calculations are given with the GAMESSPLUS Solubility Utility program. They are named pentane1.inp and pentane2.inp and are located under the directory /gmsplus-v2008/gmsplus_solubility. The first test calculation computes the solubility of *n*-pentane in water using SM5.42/HF/6-31G* to calculate both $\Delta G_{\mathrm{S}}^{\mathrm{o}}$ and $P^{\bullet}$. The second test calculation computes the solubility of *n*-pentane in water using SM5.42/HF/6-31G(d) to calculate $\Delta G_{\mathrm{S}}^{\mathrm{o}}$ and a user-defined pure-solute vapor pressure with the PRESSURE keyword. Below are the input files and output files for these two calculations.

Input

pentane1.inp:

```
 $CONTRL SCFTYP=RHF RUNTYP=GRADIENT COORD=UNIQUE $END
  $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 POLAR=POPLE $END
  $END
  $DATA
  Solubility calculation of n-pentane
C1
C 6.0 -2.551305  .329567  .000000
C 6.0 -1.279490  -.531652  .000000
C 6.0 .000000  .318908  .000000
H 1.0 .000000  .976801  .884631
H 1.0 -1.277493  -1.188738  .883613
H 1.0 -1.277493  -1.188738  -.883613
H 1.0 -2.584647  .975537  -.889230
H 1.0 -2.584647  .975537  .889230
H 1.0 -3.455107  -.294693  .000000
C 6.0 1.279495  -.531646  .000000
H 1.0 1.277498  -1.188730  -.883614
H 1.0 1.277498  -1.188730  .883614
C 6.0 2.551302  .329578  .000000
H 1.0 2.584636  .975547  .889231
H 1.0 2.584636  .975547  -.889231
H 1.0 3.455111  -.294675  .000000
H 1.0 .000000  .976801  -.884631
  $END
  $vapor ICDS=3 ISCRF=1 IAQU=0 SolN=1.357 SolA=0.0
        SolB=0.0 SolC=0.0 SolH=0.0 Dielec=1.837
        SolG=22.2951 DENSITY=8.607067
  $END

  $DGS ICDS=3 ISCRF=1 $END
```

pentane2.inp:

```
$CONTRL SCFTYP=RHF RUNTYP=GRADIENT COORD=UNIQUE $END
 $BASIS GBASIS=N31 NGAUSS=6 NDFUNC=1 POLAR=POPLE $END
 $END
 $DATA
 Solubility calculation of n-pentane
C1
C 6.0 -2.551305  .329567  .000000
C 6.0 -1.279490  -.531652  .000000
C 6.0 .000000  .318908  .000000
H 1.0 .000000  .976801  .884631
H 1.0 -1.277493  -1.188738  .883613
H 1.0 -1.277493  -1.188738  -.883613
H 1.0 -2.584647  .975537  -.889230
H 1.0 -2.584647  .975537  .889230
H 1.0 -3.455107  -.294693  .000000
C 6.0 1.279495  -.531646  .000000
H 1.0 1.277498  -1.188730  -.883614
H 1.0 1.277498  -1.188730  .883614
C 6.0 2.551302  .329578  .000000
H 1.0 2.584636  .975547  .889231
H 1.0 2.584636  .975547  -.889231
H 1.0 3.455111  -.294675  .000000
H 1.0 .000000  .976801  -.884631
 $END
 $vapor Pressure=0.890832272 Units=Bar Density=8.607067
 $END

 $DGS ICDS=3 ISCRF=1 $END
```

Output

pentane1.log:

```
GAMESSPLUS Solubility Utility calculation
based on GAMESSPLUS 2008

Wed Apr 16 21:03:51 2008

      ------ Summary of Solubility Calculation ------

 Standard-state free energy of self-solvation of solute:

 DeltaE-EN(liq) elect-nuc reorganization:         .002 kcal/mol
 G-P(liq) polarization free energy of solvation:  -.082 kcal/mol
 G-CDS(liq) cavity-dispersion-solvent structure:  -3.164 kcal/mol
 Free energy of self-solvation:                   -3.244 kcal/mol

 Standard-state free energy of solvation of solute:

 DeltaE-EN(liq) elect-nuc reorganization:         .012 kcal/mol
 G-P(liq) polarization free energy of solvation:  -.192 kcal/mol
 G-CDS(liq) cavity-dispersion-solvent structure:  1.766 kcal/mol
 Free energy of self-solvation:                   1.586 kcal/mol

 Density of Solute:                               8.607067 mol/L
 Vapor pressure of solute:
                                                  0.88atm
                                              89083.23Pa
                                                668.18torr
                                                  0.89bar

 Solubility of solute:                            0.0025 mol/L
 Logarithm (base10) of solubility of solute:      -2.61
```

pentane2.log

```
GAMESSPLUS Solubility Utility calculation
based on GAMESSPLUS 2008

Wed Apr 16 21:05:56 2008

      ------ Summary of Solubility Calculation ------

 A value of the pure-solute vapor pressure was supplied by the user,
 so no information regarding the free energy of self-solvation is available.

 Standard-state free energy of solvation of solute:

 DeltaE-EN(liq) elect-nuc reorganization:         .012 kcal/mol
 G-P(liq) polarization free energy of solvation:  -.192 kcal/mol
 G-CDS(liq) cavity-dispersion-solvent structure:  1.866 kcal/mol
 Free energy of self-solvation:                   1.687 kcal/mol

 Density of Solute:                               8.607067 mol/L
 Vapor pressure of solute:
                                                  0.88atm
                                              89083.23Pa
                                                668.18torr
                                                  0.89bar

 Solubility of solute:                            0.0025 mol/L
 Logarithm (base10) of solubility of solute:      -2.61
```

Installing the Solubility Utility Program

A working version of GAMESSPLUS must be installed first, see the sections entitled *Notes on Running GAMESSPLUS* and *Updating and Compiling GAMESS* in this manual.

The GAMESSPLUS Solubility Utility program comes with the GAMESSPLUS version 2008 distribution. It is located in the gmsplus-v2008/gmsplus_solubility directory. In this directory, you will find a PERL script named install.pl (which makes the Solubility Utility program), the directory lib, which contains the source file for the Solubility Utility program, two test cases, pentane1.inp and pentane2.inp, and the file midi-bang.bas, which contains the coefficients necessary to run calculations that use the MIDI! basis set.

To install the GAMESSPLUS Solubility Utility program, the following steps should be taken:

1.) Ensure that install.pl is executable by issuing the following command

```
chmod u+x install.pl
```

The script assumes that the PERL interpreter is located under /usr/local/bin. If it is not, then the first line of install.pl, as well as the file solubility.pl (located under /lib) both need to be modified accordingly.

2.) Run the install.pl script by typing

```
./install.pl
```

This script prompts for two pieces of information. The first is the absolute path to the rungms script. (A guess is offered as to the location of the rungms script. It this guess is correct, copy and paste the guess into the prompt. If not, manually enter the absolute path to the rungms script.) Type in the absolute path and press enter. The second piece of information is the version number of the gamessplus executable. Type in the version number (in the format given by the script) and press enter.

Running the Solubility Utility Program

The name of the Solubility Utility program is solubility.pl. It is executed with the following command:

```
path/gmsplus-v2008/gmsplus_solubility/solubility.pl input
```

where *path* is the location of the GAMESSPLUS distribution and *input* is the name of the input file. The solubility utility program creates two input files, vapor.*input*.log and dgs.*input*.log corresponding to the calculation of $P^{\bullet}$ and $\Delta G_{\mathrm{S}}^{\mathrm{o}}$, respectively. (When the PRESSURE keyword is used, a vapor pressure calculation is not required, so there will be no vapor.*input*.log file in this case). The program also creates a file named *input*.log, which contains a summary of the calculations of $P^{\bullet}$, $\Delta G_{\mathrm{S}}^{\mathrm{o}}$, and the solubility of the solute.

GAMESSPLUS Soil Sorption Utility

Executive Summary

The GAMESSPLUS Soil Sorption Utility program is a utility program for calculating soil sorption coefficients. For a given solute, the soil sorption coefficient (K_{OC}) is defined as

$$K_{\mathrm{OC}} = \frac{C_{\mathrm{soil}} / C^{\circ}_{\mathrm{soil}}}{C_{\mathrm{w}} / C^{\circ}_{\mathrm{w}}} \tag{1}$$

where C_{soil} is the concentration of solute per gram of carbon in standard soil, C_{w} is the concentration of solute per volume of aqueous solution, and $C^{\circ}_{\mathrm{soil}}$ and C°_{w} are the standard state concentrations of organic carbon for soil and aqueous solution, respectively. Typically, a standard state of 1 μg of solute/g of organic carbon is used for $C^{\circ}_{\mathrm{soil}}$, and 1 mol/L is used for C°_{w}.

The GAMESSPLUS Soil Sorption Utility program calculates K_{OC} according to

$$K_{\mathrm{OC}} = \rho_{\mathrm{soil}} \left(\Delta G^{\circ}_{\mathrm{w}} - \Delta G^{\circ}_{\mathrm{soil}} \right) \tag{2}$$

where ρ_{soil} is the density of soil (in g/mL), $\Delta G^{\circ}_{\mathrm{w}}$ is the standard state free energy associated with transferring a solute from the gas phase to aqueous solution, and $\Delta G^{\circ}_{\mathrm{soil}}$ is the standard state free energy associated with transferring a solute from the gas phase to soil. In the GAMESSPLUS Soil Sorption Utility program, calculated $\Delta G^{\circ}_{\mathrm{w}}$ values are for a standard state of 1 mol/L in both the gas and aqueous phase, and $\Delta G^{\circ}_{\mathrm{soil}}$ values are for a standard state of 1 mol/L in both the gas phase and in soil.

Using the above relationship between K_{OC}, ρ_{soil}, $\Delta G^{\circ}_{\mathrm{w}}$, and $\Delta G^{\circ}_{\mathrm{soil}}$, the GAMESSPLUS Soil Sorption Utility program can calculate K_{OC} several ways. First, given a value for ρ_{soil}, the GAMESSPLUS Soil Sorption Utility program can calculate $\Delta G^{\circ}_{\mathrm{w}}$ and $\Delta G^{\circ}_{\mathrm{soil}}$, and then use these calculated values in eq 2 to determine K_{OC}. For calculating $\Delta G^{\circ}_{\mathrm{w}}$, the SM5.42 aqueous continuum solvation model is used. For $\Delta G^{\circ}_{\mathrm{soil}}$ the SM5.42 universal continuum solvation model is used, along with a set of solvent descriptors that have been empirically optimized for modeling bulk soil (for a description of the SM5.42 model for soil, as well as a discussion of modeling soil as a homogenous medium, see Winget, P.; Cramer, C. J.; Truhlar, D. G. "Prediction of Soil Sorption Coefficients Using a Universal Solvation Model", *Environ. Sci. Technol.* **2000**, *34*, 4733).

The GAMESSPLUS Soil Sorption Utility program also allows K_{OC} values to be calculated with user-supplied data for $\Delta G^{\circ}_{\mathrm{w}}$. Thus, in cases where an experimental value for $\Delta G^{\circ}_{\mathrm{w}}$ is known, it can be used in eq 2 to calculate K_{OC}. In fact, a $\Delta G_{\mathrm{w}}^{\circ}$ value obtained from any reliable method (e.g. SM5.43 or SM6) can be used in lieu of the $\Delta G_{\mathrm{w}}^{\circ}$ value calculated using SM5.42 in eq 2.

This section and several subsections below of this manual provide a stand-alone introduction to the GAMESSPLUS Soil Sorption Utility program for users who just want to calculate soil sorption coefficients. Such users should first install GAMESSPLUS according to the instructions given in the sections entitled *Notes on Running GAMESSPLUS* and *Updating and Compiling GAMESS* of this manual. Users of this utility program should give the appropriate references described in the section entitled *GAMESSPLUS Reference*. Then, all of the information required to run a standard calculation with the GAMESSPLUS Soil Sorption Utility program is in the following few sections.

The SM5.42 continuum solvation model for aqueous solution and soil

SM5.42 is a continuum solvation model that uses SM5 functional forms for atomic surface tensions (hence the first three characters in the name of the method are SM5), class IV point charges (hence .4 comes next) of the CM2 type (hence 2). SM5.42 provides a way to calculate electronic wave functions in different liquid phases and in soil, and free energies of transfer between the gas phase and various condensed phases. The SM5.42 solvation model is universal, i.e., it is parameterized for water and any medium that can be characterized by a set of solvent descriptors. In the case of bulk soil, a set of solvent descriptors have been empirically optimized using a data base of experimental K_{OC} values. (see Winget, P.; Cramer, C. J.; Truhlar, D. G. *Environ. Sci. Technol.* **2000**, *34*, 4733).

In the GAMESSPLUS Soil Sorption Utility program, standard-state free energies of transfer are calculated using rigid, gas-phase geometries (i.e. no geometry optimization in solution is performed). Our experience has shown that in most cases, there is very little error associated with using gas-phase geometries to calculate free energies of solvation. Thus, one may use as input gas-phase geometries optimized at any reliable level of theory. However, it is important to point out that in some cases the geometry of a given solute might undergo significant relaxation between the gas-phase and solution (or soil). As a result, the transfer free energy calculated using a rigid, gas-phase geometry might differ significantly from the transfer free energy calculated using a relaxed geometry. In cases where significant geometric relaxation is expected to occur upon transfer from the gas-phase to water or soil, a geometry optimized in soil should be used as input for a GAMESSPLUS Soil Sorption calculation (for water, a separate calculation should be carried out with the main GAMESSPLUS program to obtain the value of $\Delta G_{\mathrm{W}}^{\circ}$ for the relaxed geometry; this value should then be used as input for a soil sorption calculation). The instructions for performing geometry optimizations with the main GAMESSPLUS are outlined in the main part of this manual. In the main GAMESSPLUS program, the solvent descriptors for soil must be explicitly defined in the input file (in the GAMESSPLUS Soil Sorption Utility program, the descriptors are automatically defined, depending on the level of theory). For aqueous solution, solvent descriptors for bulk water do not need to be provided in either the GAMESSPLUS Soil Sorption Utility program or the main GAMESSPLUS program. The solvent descriptors for soil are described in the section below.

Solvent descriptors for bulk soil

For nonaqueous solvents, the SM5.42 universal continuum solvent model uses a set of 7 solvent descriptors to characterize the properties of the solvent of interest. These seven solvent descriptors are as follows: ε, the dielectric constant of the solvent, n, refractive index at the wavelength of the Na D line; α, Abraham's hydrogen bond acidity parameter $\Sigma\alpha_2$; β, Abraham's hydrogen bond basicity parameter $\Sigma\beta_2$; γ, macroscopic molecular surface tension in units of cal mol^{-1} Å^{-2} ; ϕ^2, square of the fraction ϕ of nonhydrogenic solvent atoms that are aromatic carbon atoms (aromaticity); and ψ^2, square of the fraction ψ of nonhydrogenic solvent atoms that are F, Cl, or Br (electronegative halogenecity). For bulk soil, ϕ and ψ are both zero. The remaining five solvent descriptors have been empirically optimized against experimental values. Note that the solvent descriptors for soil, which are listed below, should only be used with the levels of theory for which they are optimized (AM1 and HF/MIDI!).

SM5.42 Solvent Descriptors for Bulk Soil

		AM1	HF/MIDI!
Descriptor	GAMESSPLUS Keyword[a]	\$SCF = AM1	\$SCF = ABINITIO
ε	DIELEC	15.0	15.0
n	SOLN	1.541	1.379
α	SOLA	0.36	0.61
β	SOLB	0.34	0.60
γ	SOLG	63.3	46.0
ϕ	SOLC	0.0	0.0
ψ	SOLH	0.0	0.0

[a]Not required input for the GAMESSPLUS Soil Sorption Utility program.

Usage

A description of the input that is required to run this utility program is given below. For more detailed information for setting up GAMESS or other GAMESSPLUS calculations, see the documentation that comes with the GAMESS distribution and the main part of the GAMESSPLUS manual.

Input

All of the input required to run a calculation with the GAMESSPLUS Soil Sorption program is entered in a pseudo-namelist format; a namelist can be thought of as a particular group of keywords. The title of this group, the namelist name, is prefixed by a '$' that must be entered in column 2 of an input file (e.g. $SCF). A namelist accepts as arguments a number of possible keywords that are entered after the namelist name in a free format style that can span over multiple lines. Each keyword accepts either character or numeric values. A namelist is terminated with a $END. Only input between a namelist name and its corresponding $END will be read from a GAMESS input file.

Below is a brief summary the namelists that are used by the GAMESSPLUS Soil Sorption Utility program.

$SOILDENSITY Soil density (in g/mL)

By default, K_{OC} values are calculated using a soil density of 0.11 g/mL in eq 2, which is a representative value of soil density measured for a variety of soil types from Eastern North America (Federer, C. A.; Turcotte, D. E.; Smith, C. T. *Can. J. For. Res.* **1993**, *23*, 1026). It is recommended that this value be used when calculating K_{OC} values.

$SCF Determines the level of electronic structure theory that will be used to calculate transfer free energies.

AM1	Use AM1. This is the default method.
ABINITIO	Use HF/MIDI!.

$DELTAGW User-supplied value for $\Delta G_{\mathrm{W}}^{\circ}$.

If the $DELTAGW namelist is left blank or is not present, $\Delta G_{\mathrm{W}}^{\circ}$ will be calculated using the SM5.42 continuum solvation model, and this value will be used in eq 2 to calculate K_{OC}. Alternatively, a user-supplied value for $\Delta G_{\mathrm{W}}^{\circ}$ can be entered, and this value will be used in eq 2 to calculate K_{OC}. Note that when entering a value for $\Delta G_{\mathrm{W}}^{\circ}$ in the $DELTAGW namelist, it must be in units of kcal/mol. Also note that the GAMESSPLUS Soil Sorption Utility program uses a standard state of 1 mol/L in both the gas and aqueous phase to calculate K_{OC}. Thus, user-supplied values for $\Delta G_{\mathrm{W}}^{\circ}$ should also correspond to the above standard-state.

$DATA Input molecule information

Input consists of the following information:

line 1: title line
line 2: symmetry group (always enter `C1` for GAMESSPLUS calculations)
line 3+: molecular coordinates; for Cartesian input, each line consists of the following: atom label, the atom's nuclear charge (which should be a floating-point value, i.e., 1.0 for H), and the atom's *x*, *y*, and *z*-coordinate. See examples in the *Input and Output Examples* section and in the test suite.

Test Calculations

Three test calculations are given with the GAMESSPLUS Soil Sorption Utility program. They are named test1.inp, test2.inp, and test3.inp, and are located under the directory /gmsplus-v2008/gmsplus_soil. The first test calculation calculates $\Delta G^{\circ}_{\mathrm{w}}$ and $\Delta G^{\circ}_{\mathrm{soil}}$ values at the AM1/SM5.42 level of theory, and then uses these calculated free energies in eq 2 to compute K_{OC}. The second test calculation calculates $\Delta G^{\circ}_{\mathrm{soil}}$ at the AM1/SM5.42 level of theory, and then uses this calculated value along with an experimental value for $\Delta G^{\circ}_{\mathrm{w}}$ in eq 2 to compute K_{OC}. The third test calculation repeats test calculation 1 at the HF/MIDI! level of theory. Note that for test3, the coefficients for the MIDI! basis set are entered in the data section. This basis set is included with the GAMESSPLUS distribution, and can be found in the directory /gmsplus-v2008/gmsplus_soil. Below are the input files and output files for these three calculations.

Input

test1.inp:

```
  $DATA
Koc calculation for benzene, using calculated values for deltagw and deltagsoil
C1
C    6.0          .000000    1.208141     .697522
C    6.0          .000000    1.208141    -.697522
C    6.0          .000000     .000000   -1.395043
C    6.0          .000000   -1.208141    -.697522
C    6.0          .000000   -1.208141     .697522
C    6.0          .000000     .000000    1.395043
H    1.0          .000000    2.160477    1.247356
H    1.0          .000000    2.160477   -1.247356
H    1.0          .000000     .000000   -2.494707
H    1.0          .000000   -2.160477   -1.247356
H    1.0          .000000   -2.160477    1.247356
H    1.0          .000000     .000000    2.494707
 $END
 $SCF AM1 $END
 $SOILDENSITY 0.11 $END
```

test2.inp:

```
$DATA
Koc calculation for benzene, using an experimental value for deltagw
C1
C    6.0          .000000    1.208141     .697522
C    6.0          .000000    1.208141    -.697522
C    6.0          .000000     .000000   -1.395043
C    6.0          .000000   -1.208141    -.697522
C    6.0          .000000   -1.208141     .697522
C    6.0          .000000     .000000    1.395043
H    1.0          .000000    2.160477    1.247356
H    1.0          .000000    2.160477   -1.247356
H    1.0          .000000     .000000   -2.494707
H    1.0          .000000   -2.160477   -1.247356
H    1.0          .000000   -2.160477    1.247356
H    1.0          .000000     .000000    2.494707
 $END
 $SCF AM1 $END
 $SOILDENSITY 0.11 $END
 $DELTAGW -0.87 $END
```

test3.inp:

```
 $SCF ABINITIO $END
 $SOILDENSITY 0.11 $END
 $DATA
Koc calculation for benzene, using calculated values for deltagw and deltagsoil
C1
C    6.0         .000000    1.198348     .691865
S   3
    1    153.1722600        0.0707400000
    2     23.0730300        0.3953800000
    3      4.9232900        0.6633110000
S   2
    1      5.7255700       -0.0813800000
    2      0.4550400        0.5748530000
S   1
    1      0.1470700        1.0000000000
P   2
    1      4.2513100        0.1099310000
    2      0.8632700        0.4627130000
P   1
    1      0.2013500        1.0000000000

C    6.0         .000000    1.198348    -.691865
S   3
    1    153.1722600        0.0707400000
    2     23.0730300        0.3953800000
    3      4.9232900        0.6633110000
S   2
    1      5.7255700       -0.0813800000
    2      0.4550400        0.5748530000
S   1
    1      0.1470700        1.0000000000
P   2
    1      4.2513100        0.1099310000
    2      0.8632700        0.4627130000
P   1
    1      0.2013500        1.0000000000

C    6.0         .000000     .000000   -1.383730
S   3
    1    153.1722600        0.0707400000
    2     23.0730300        0.3953800000
    3      4.9232900        0.6633110000
S   2
    1      5.7255700       -0.0813800000
    2      0.4550400        0.5748530000
S   1
    1      0.1470700        1.0000000000
P   2
    1      4.2513100        0.1099310000
    2      0.8632700        0.4627130000
P   1
    1      0.2013500        1.0000000000

C    6.0         .000000   -1.198348    -.691865
S   3
    1    153.1722600        0.0707400000
    2     23.0730300        0.3953800000
    3      4.9232900        0.6633110000
S   2
    1      5.7255700       -0.0813800000
    2      0.4550400        0.5748530000
S   1
    1      0.1470700        1.0000000000
```

```
P   2
    1      4.2513100        0.1099310000
    2      0.8632700        0.4627130000
P   1
    1      0.2013500        1.0000000000

C    6.0         .000000   -1.198348     .691865
S   3
    1    153.1722600        0.0707400000
    2     23.0730300        0.3953800000
    3      4.9232900        0.6633110000
S   2
    1      5.7255700       -0.0813800000
    2      0.4550400        0.5748530000
S   1
    1      0.1470700        1.0000000000
P   2
    1      4.2513100        0.1099310000
    2      0.8632700        0.4627130000
P   1
    1      0.2013500        1.0000000000

C    6.0         .000000     .000000    1.383730
S   3
    1    153.1722600        0.0707400000
    2     23.0730300        0.3953800000
    3      4.9232900        0.6633110000
S   2
    1      5.7255700       -0.0813800000
    2      0.4550400        0.5748530000
S   1
    1      0.1470700        1.0000000000
P   2
    1      4.2513100        0.1099310000
    2      0.8632700        0.4627130000
P   1
    1      0.2013500        1.0000000000

H    1.0         .000000    2.130455    1.230017
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000

H    1.0         .000000    2.130455   -1.230017
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000

H    1.0         .000000     .000000   -2.460034
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000

H    1.0         .000000   -2.130455   -1.230017
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000
```

```
H    1.0         .000000   -2.130455    1.230017
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000

H    1.0         .000000     .000000    2.460034
S   2
    1      4.5018000        0.0704520000
    2      0.6814440        0.4078260000
S   1
    1      0.1513980        1.0000000000

 $END
```

Output

test1.log:

```
GAMESSPLUS Soil Sorption Utility calculation
based on GAMESSPLUS 2008

Tue Apr  8 19:58:36 2008

KOC CALCULATION FOR BENZENE, USING CALCULATED VALUES FOR DELTAGW AND DELTAGSOIL
C1
C    6.0          .000000    1.208141     .697522
C    6.0          .000000    1.208141    -.697522
C    6.0          .000000     .000000   -1.395043
C    6.0          .000000   -1.208141    -.697522
C    6.0          .000000   -1.208141     .697522
C    6.0          .000000     .000000    1.395043
H    1.0          .000000    2.160477    1.247356
H    1.0          .000000    2.160477   -1.247356
H    1.0          .000000     .000000   -2.494707
H    1.0          .000000   -2.160477   -1.247356
H    1.0          .000000   -2.160477    1.247356
H    1.0          .000000     .000000    2.494707

$SCF         =    AM1
$SOILDENSITY =    0.11  g/mL

All values are for a temperature of 298 K.
Free energies are calculated using a 1M --> 1M standard state
log = log(base10)
-----------------------------------------------------------------
 Free Energy of Transfer (air --> water)       -1.51  kcal/mol
 Free Energy of Transfer (air --> soil)        -4.32  kcal/mol
 Free Energy of Transfer (water --> soil)      -2.81  kcal/mol
 log Koc                                 3.02
-----------------------------------------------------------------

End of output.
```

test2.log

```
GAMESSPLUS Soil Sorption Utility calculation
based on GAMESSPLUS 2008

Tue Apr  8 20:00:14 2008

KOC CALCULATION FOR BENZENE, USING AN EXPERIMENTAL VALUE FOR DELTAGW AND A CALCULATED
VALUE FOR DELTAGSOIL
C1
C    6.0          .000000    1.208141     .697522
C    6.0          .000000    1.208141    -.697522
C    6.0          .000000     .000000   -1.395043
C    6.0          .000000   -1.208141    -.697522
C    6.0          .000000   -1.208141     .697522
C    6.0          .000000     .000000    1.395043
H    1.0          .000000    2.160477    1.247356
H    1.0          .000000    2.160477   -1.247356
H    1.0          .000000     .000000   -2.494707
H    1.0          .000000   -2.160477   -1.247356
H    1.0          .000000   -2.160477    1.247356
H    1.0          .000000     .000000    2.494707

$SCF         =    AM1
$SOILDENSITY =    0.11  g/mL
$DELTAGW     =    -0.87  kcal/mol

All values are for a temperature of 298 K.
Free energies are calculated using a 1M --> 1M standard state
log = log(base10)
-----------------------------------------------------------------
 Free Energy of Transfer (air --> water)       -0.87  kcal/mol
 Free Energy of Transfer (air --> soil)        -4.32  kcal/mol
 Free Energy of Transfer (water --> soil)      -3.45  kcal/mol
 log Koc                                   3.50
-----------------------------------------------------------------

End of output.
```

test3.log

```
GAMESSPLUS Soil Sorption Utility calculation
based on GAMESSPLUS 2008

Tue Apr  8 20:00:43 2008

KOC CALCULATION FOR BENZENE, USING CALCULATED VALUES FOR DELTAGW AND DELTAGSOIL
C1
                                          ...

$SCF         =    ABINITIO
$SOILDENSITY =    0.11  g/mL

All values are for a temperature of 298 K.
Free energies are calculated using a 1M --> 1M standard state
log = log(base10)
-----------------------------------------------------------------
 Free Energy of Transfer (air --> water)       -1.34  kcal/mol
 Free Energy of Transfer (air --> soil)        -3.77  kcal/mol
 Free Energy of Transfer (water --> soil)      -2.43  kcal/mol
 log Koc                                 2.74
-----------------------------------------------------------------

End of output.
```

Installing the GAMESSPLUS Soil Sorption Utility Program

A working version of GAMESSPLUS must be installed first, see the sections entitled *Notes on Running GAMESSPLUS* and *Updating and Compiling GAMESS* in this manual.

The GAMESSPLUS Soil Sorption Utility program comes with the GAMESSPLUS version 2008 distribution. It is located in the gmsplus-v2008/gmsplus_soil directory. In this directory, you will find a PERL script named install.pl (which makes the Soil Sorption Utility program), the directory lib, which contains the source file for the Soil Sorption Utility program, three test cases, test1.inp, test2.inp, and test3.inp, and the file midi-bang.bas, which contains the coefficients necessary to run calculations that use the MIDI! basis set.

To install the GAMESSPLUS Soil Sorption Utility program, the following steps should be taken:

1.) Ensure that install.pl is executable by issuing the following command

```
chmod u+x install.pl
```

The script assumes that the PERL interpreter is located under /usr/local/bin. If it is not, then the first line of install.pl, as well as the file soil.pl (located under /lib) both need to be modified accordingly.

2.) Run the install.pl script by typing

```
./install.pl
```

This script prompts for two pieces of information. The first is the absolute path to the rungms script. (A guess is offered as to the location of the rungms script. It this guess is correct, copy and paste the guess into the prompt. If not, manually enter the absolute path to the rungms script.) Type in the absolute path and press enter. The second piece of information is the version number of the gamessplus executable. Type in the version number (in the format given by the script) and press enter.

Running the Soil Sorption Utility Program

The name of the Solubility Utility program is soil.pl. It is executed with the following command:

path/gmsplus-v2008/gmsplus_soil/soil.pl input

where *path* is the location of the GAMESSPLUS distribution and *input* is the name of the input file. The soil sorption utility program creates two input files, water.*input*.inp and soil.*input*.inp corresponding to the calculation of $\Delta G_{\mathrm{w}}^{\mathrm{o}}$ and $\Delta G_{\mathrm{soil}}^{\mathrm{o}}$, respectively. (When the $DELTAGW keyword is used, a free energy calculation in water is not required, so there will be no water.*input*.inp file in this case). The program also creates a file named *input*.log, which contains a summary of the calculations of $\Delta G_{\mathrm{w}}^{\mathrm{o}}$, $\Delta G_{\mathrm{soil}}^{\mathrm{o}}$, and the logarithm (base 10) of K_{OC}.

GAMESSPLUS Revision History and Version Summaries

Note: First and second level revisions (e.g., 2.0, 2.1) are enhancements. Third level revisions (e.g., 1.1.1, 2.0.1) are bug fixes. Versions 1.0 through 3.1 were called *GAMESOL*. Beginning with version 3.9 the name is changed to *GAMESSPLUS*. In the revision histories, for each version of *GAMESSPLUS* (*GAMESOL*), we list the authors of that version and we also list the version of *GAMESS* on which it was based.

GAMESOL Version 1.0 (December 1997)

Authors: J. Li, G. D. Hawkins, D. A. Liotard, C. J. Cramer, and D. G. Truhlar
GAMESS version: October 31, 1996

- This version added the Löwdin and CM2 charge models, and the SM5.42R solvation model to the Oct. 31, 1996 version of *GAMESS*. Parameters for following methods were included (5D wave functions and DFT models not yet available in *GAMESS*):

CM2/HF/MIDI!	SM5.42R/HF/MIDI!	CM2/BPW91/6-31G*	
CM2/HF/MIDI!6D	SM5.42R/HF/MIDI!6D	CM2/HF/6-31+G*	SM5.42R/HF/6-31+G*
CM2/HF/6-31G*	SM5.42R/HF/6-31G*	CM2/HF/cc-pVDZ	
CM2/BPW91/MIDI!		CM2/BPW91/DZVP	SM5.42R/BPW91/DZVP
CM2/BPW91/MIDI!6D	SM5.42R/BPW91/MIDI!6D	CM2/AM1	
CM2/B3LYP/MIDI!	SM5.42R/B3LYP/MIDI!	CM2/PM3	

GAMESOL Version 1.1 (March 1998)

Authors: J. Li, G. D. Hawkins, D. A. Liotard, C. J. Cramer, and D. G. Truhlar
GAMESS version: October 31, 1996 and January 6, 1998

- This version can be used with either the Oct. 31, 1996 version or the Jan. 6, 1998 version of *GAMESS*. The same CM2 and SM5.42R methods are supported as in version 1.0.

GAMESOL Version 1.1.1 (September 1998)

Authors: J. Li, G. D. Hawkins, D. A. Liotard, C. J. Cramer, and D. G. Truhlar
GAMESS version: October 31, 1996 and January 6, 1998

- This version fixed two bugs in subroutine SM5CDS for the surface tension functional and a bug in subroutine GBSCRF for option ISCRF = 1. A few typos in the text file of solvent properties have also been corrected.

GAMESOL Version 2.0 (September 1998)

Authors: J. Li, T. Zhu, G. D. Hawkins, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS version: January 6, 1998

- This version added analytical gradients and the capability to optimize solute geometries in liquid solutions by the SM5.42/HF solvation model.

GAMESOL Version 2.0.1 (January 1999)

Authors: J. Li, T. Zhu, G. D. Hawkins, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS version: January 6, 1998

- Bugs in the subroutine for the CDS term were fixed. The bugs were for the terms involving CC triple bonds and amides.
- Some errors in parameters for SM5.42R/AM1 and SM5.42R/PM3 models have been corrected.
- The code now uses the correct coulomb radii for SM5.42R/AM1, SM5.42R/PM3, SM5.42/AM1, and SM5.42/PM3 when these methods are invoked. Note that these coulomb radii are different from the SM5.42R/HF and SM5.42/HF coulomb radii for P and S.
- In version 2.0, *GAMESOL* would crash after the gas-phase Hartree-Fock calculation was done during the gas-phase geometry optimization phase of an SM5.42 calculation. The Hessian matrix from the last step of gas-phase geometry optimization was not being stored, and thus there was no previous Hessian available for the SM5.42 calculation.
- Some common blocks have been rearranged so that longer variables come first, i.e., the common blocks have been made byte aligned (floating-point variables before integer variables before logical variables). Although this is not required on most machines, on some machines running the Linux operation system a compilation error will be generated if the variables in a common block are not properly arranged.
- Some debug lines have been removed.

GAMESOL Version 2.1 (January 1999)

Authors: J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS version: January 6, 1998 and May 6, 1998

- This version added SM5.42R/AM1, SM5.42R/PM3, SM5.42/AM1, and SM5.42/PM3 solvation models. The following models can be used in *GAMESOL*:

Rigid solvation model:	Solvation models with analytical gradient:
SM5.42R/MIDI!6D	SM5.42/MIDI!6D
SM5.42R/6-31G*	SM5.42/6-31G*
SM5.42R/6-31+G*	SM5.42/6-31+G*
SM5.42R/AM1	SM5.42/AM1
SM5.42R/PM3	SM5.42/PM3

- Three new subroutines, GPDER1, SMXPUN, and DOGEOM were added. Subroutine GPDER1 calculates the analytical gradient for SM5.42/AM1 and SM5.42/PM3, and is called by subroutine MPCGRD (in *mpcgrd.src*). Subroutine SMXPUN prints out the energy components from solvation calculations in required by *POLYRATE*, and is called by subroutine DOSOLV. Subroutine DOGEOM initiates liquid-phase geometry optimization, and is called from subroutine DISPLC (in *statpt.src*).
- The capability of inputting specific reaction parameters (SRP) for AM1, PM3, and/or CM2 is added. The former capability is called NDDO-SRP, and the latter is called CM2-SRP. A new subroutine RCM2SRP has been added to the *smx.src* and a new module *nddosrp* has been implemented in file *nddosrp.src*.

- Namelist input is now implemented by calling the *GAMESS* subroutine NAMEIO. This enhances the portability of the code, since namelist input is machine dependent.
- ICDS is set equal to ICMD by default. Thus, a user does not need to specify ICDS.
- The *GAMESOL* version number is printed in the summary of solvation calculations.
- *GAMESOL* now works with both the Jan. 6, 1998 and May 6, 1998 versions of *GAMESS*.
- The Users Manual has been re-organized and improved.

GAMESOL Version 2.2 (April 1999)

Authors: J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS versions: January 6, 1998, May 6, 1998, and December 1, 1998

- *GAMESOL* now supports the Mar. 15, 1999 version of *GAMESS* as well as with the Jan. 6, 1998 and May 6, 1998 version of *GAMESS*.
- New NDDO-SRP options BETSS and BETSP are added with a new test run 11.inp as an example.
- The DATA statements in *smx.src* are moved after the declaration statements of the variables according to the standard FORTRAN syntax.

GAMESOL Version 2.2.1 (April 1999)

Authors: J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS versions: January 6, 1998, May 6, 1998, and March 15, 1999

- The name of directory gamesol.2.2.mod/Dec98 is changed to gamesol2.2.1.mod/Mar99 to reflect that the version of *GAMESS* we received from Iowa State had changes later than the date printed in the box.

GAMESOL Version 2.2.2 (May 1999)

Authors: J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS versions: January 6, 1998, May 6, 1998, and March 15, 1999

- The parameter NUMATM has been assigned to be the same as MXATM (which is 500) in the subroutines BORNRD, CALCDS, CALSTN, DAREAL, GBMOD, GPDER, GPDER1, and SM5CDS (in *smx.src*) to allow for solvation calculation on molecules containing more than 100 atoms.

GAMESOL Version 2.2.3 (July 1999)

Authors: J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS versions: January 6, 1998, May 6, 1998, and March 15, 1999

- The parameter NUMATM has been assigned to 100 in the subroutines BORNRD, CALCDS, CALSTN, DAREAL, GBMOD, GPDER, GPDER1, and SM5CDS (in *smx.src*) due to the size of the executable files.

GAMESOL Version 2.2.4 (August 1999)

Authors: J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS versions: January 6, 1998, May 6, 1998, and March 15, 1999

- The subroutines RNDDOSRP and RCM2SRP were modified to allow them to read the NDDO−SRP and CM2SRP parameter files on both the IBM-SP and SGI Origin2000 workstations. Before compiling *nddosrp.src* and *smx.src* on IBM-SP machines that run the AIX operating system, the "*AIX" at the beginning of lines contained in these subroutines must be removed. No source modification is required for the SGI workstations.

GAMESOL Version 2.2.5 (October 2000)

Authors: J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS versions: January 6, 1998, May 6, 1998, and March 15, 1999

- Corrected an error in the analytical derivative of the N-C=O surface tension term in subroutine SM5CDS.
- Moved the evaluation of DIELEC after the IF(ISOL.LT.0) line in subroutine DOSOLV. The previous implementation would crash during gas-phase calculations on some machines.
- Uncommented the following line in subroutine DISPLC (in *statpt.src*):
  ```
      IF(CVGED) GO TO 700                              !jx0700
  ```
  The absence of this line causes incorrect printing of molecular information after geometry optimization has concluded for both gas-phase and liquid-phase geometry optimizations.
- Common block /GBCNTL/ in subroutine SMXPUN was made consistent with all other instances of this common block.
- The absolute free energy of the solute is now passed to the routines that handle geometry optimization. This should improve convergence in some cases.

GAMESOL Version 3.0 (February 2001)

Authors: J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar

GAMESS version: June 11, 2000

- CM2 and Löwdin atomic charge and SM5.42R solvation energies can now be evaluated using wave functions that use spherical harmonic *d* functions. As a consequence, the following new methods are added to *GAMESOL* in this version:

CM2/HF/MIDI!	SM5.42R/HF/MIDI!
CM2/HF/cc-pVDZ	SM5.42R/HF/cc-pVDZ

- SCRF evaluations can now be performed using Löwdin charges. This option is available for restricted or unrestricted HF wave functions that employ Cartesian or spherical harmonic basis functions. Analytical gradients of SCRF energies with respect to the nuclear coordinates are available for restricted or unrestricted HF wave functions that employ Cartesian basis functions up to *f* shell (i.e., 6D/10F).
- The surface tension parameters for SM5.42R/HF/cc-pVDZ have been added.
- Trapezoidal numerical integration has been replaced by the Gauss-Legendre quadrature method of reference *LH95* in subroutine BORNRD. This change in method improves dramatically the accuracy of the analytical gradient, leading to more successful optimizations.
- Subroutine SETSOL was modified to allow for transition state optimizations and for Hessian evaluations in solution (i.e., RUNTYP = SADPOINT and RUNTYP = HESSIAN are now recognized).
- The new $CM2 keyword ISREAD has been added to allow users to input surface tension coefficients in a file called “CDS_Param”.
- *GAMESOL* now produces more informative output consistent with other codes that implement the SM5.42R and/or the SM5.42 solvation models. In addition, the printing of charges and dipole moments has been changed for option ICREAD = 1.
- A check for whether the Coulomb radii for all of the atoms in the input are available in the code has been introduced. If there are any missing radii and a liquid-phase calculation is being performed, then the program stops. If there are any missing radii and only gas-phase CM2 charges are being calculated, then the program continues, but the printing of NOPOL information is suppressed.
- A new, comprehensive test suite has been added.

GAMESOL Version 3.0.1 (August 2002)
Authors: J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, J. D. Thompson, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS version: June 11, 2000

- In previous versions of GAMESOL, the program would terminate in subroutine SHALF if the number of basis functions for the calculation was greater than 300. However, the maximum number of basis functions allowed in a normal GAMESS calculation is 2047. To make GAMESOL program limitations the same as GAMESS program limitations, which consequently makes GAMESOL transparent to normal GAMESS users, the maximum number of basis functions allowed in subroutine SHALF is now 2047. This modification requires that the parameter MXBAS be set to 2047 throughout the file smx.src. In addition, the array that stores the square root of the overlap matrix is dimensioned to 2047*(2047+1)/2.
- The parameter MXATM, which defines the maximum number of atoms in GAMESOL was inconsistently defined in various subroutines used in GAMESOL. The maximum number of atoms allowed in GAMESS is 500, so to make GAMESOL program limitations consistent with GAMESS program limitations, all occurrences of MXATM in GAMESOL-specific code has been set to 500.
- Mayer's bond order formula for unrestricted wave functions was incorrectly implemented in version previous to 3.0.1. GAMESOL now supports energy calculations using unrestricted wave functions, but not analytical gradients.

GAMESOL Version 3.1 (August 2002)
Authors: J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, J. D. Thompson, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS version: June 11, 2000

- CM2, CM3, Löwdin, RLPA, SM5.42R, and SM5.42 calculations can now be carried out for wave functions that use spherical harmonic f functions.

- The CM3 parameter sets, which allow for evaluation of CM3 charges, were added to this version of *GAMESOL*. This provides the capability to evaluate gas-phase CM3 charges, liquid-phase CM3 charges, and the electrostatic contribution to the free energy of solvation using the generalized Born model. The corresponding CM3/GB gradient may also be evaluated and used for geometry optimizations.

- The redistributed Löwdin population analysis (RLPA) method was also added to this version of *GAMESOL*. This new method can be used to evaluate gas-phase RLPA charges and liquid-phase RLPA charges. The RLPA charges can further be used in a calculation of the electrostatic contribution to the free energy of solvation using the generalized Born model and of the corresponding free energy gradient, which can also be used for geometry optimizations.

- The namelist has been broadened to support the new options as well as the previous ones. The namelist is now called $GMSOL (although the old name, $CM2, may also be used if desired).

- To accommodate the new methods, the allowed values of the ICMD keyword have been extended to include 300, 302, and 303.

- The test suite was extended to test all of the CM3 parameter sets and to test the use of RLPA charges.

- For portability issues, common block /DSOLVA/ was removed in this version of GAMESOL. This common block stored the gradients of the effective Born radii, the gradients of the Coulomb integrals, the gradients of the solvent accessible surface areas, and the gradients of G_{CDS}. These quantities are now stored on GAMESS's direct-access file (i.e., the 'DICTNRY' file, records 267 - 270) and GAMESS main work array.

GAMESSPLUS Version 3.9 (April 2003)
Authors: J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, J. D. Thompson, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS version: January 14, 2003 (R3)

- The name of the program, *GAMESOL*, has been changed to the more general name, *GAMESSPLUS*.

- The parameters for CM3/AM1, CM3/PM3, CM3/BLYP/6-31G*, CM3/B3LYP/MIDI!6D, CM3/B3LYP/6-31G*, and CM3/B3LYP/6-31+G* were added. The new mapping scheme for compounds that contain N and O for these CM3 methods was also added. For these methods, this provides the capability to evaluate gas-phase CM3 charges, liquid-phase CM3 charges, and the electrostatic contribution to the free energy of solvation using the generalized Born (GB) model. For all six methods, the corresponding CM3/GB gradient may also be evaluated and used for geometry optimizations. Modifications were made to routine STVDER in the source file grd1.src and to various subroutines in the source file smx.src. The test suite was extended to test all of the CM3 parameters and to test the new charge-mapping scheme for compounds that contain N and O. The extended test suite also tests the B3LYP method implemented in this version of *GAMESSPLUS*.

- The PM3 parameters for Li, which are necessary to carry out CM3/PM3 calculations for compounds containing Li, were added. Modifications were made to subroutine MPCDAT in the source file mpcdat.src. The source file mpcdat.src is now part of the *GAMESSPLUS* distribution. These parameters are tested in the portion of the test suite that tests the new CM3/PM3 parameters

- The B3LYP hybrid density functional theory method that uses version III of the VWN correlation functional (i.e., the version of B3LYP coded in *Gaussian* and *HONDO/S*) was added. This method is requested by using the keyword 'DFTTYP=B3LYP3' in data group $DFT. Minor modifications were made to subroutines INPGDFT, CALCEXC in the source file dftexc.src. In addition, a new subroutine, called VWN3SCF, was created. This routine is located in the source file smx.src. The source file dftexc.src is now part of the *GAMESSPLUS* distribution. Using the keyword B3LYP now causes the program to stop and ask the user to specify B3LYP3 or B3LYP5.

- In the previous version of *GAMESSPLUS* (*GAMESOL*-version 3.1), common block /DSOLVA/, which stored the gradients of the effective Born radii, the gradients of the

Coulomb integrals, the gradients of the solvent accessible surface areas, and the gradients of G_{CDS}, was removed. The data in these arrays were instead stored on GAMESS's direct-access file (i.e., the 'DICTNRY' file, records 267 - 270). Because the gradients of the effective Born radii and the gradients of the solvent accessible surface areas are each only used locally in one specific subroutine, they do not need to be written to disk, so they are no longer stored on the 'DICTNRY' file.

GAMESSPLUS Version 4.0 (September 2003)
Authors: J. D. Xidos, J. Li. T, Zhu, G. D. Hawkins, J. D. Thompson, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, C. J. Cramer, and D. G. Truhlar
GAMESS version: January 14, 2003 (R3)

- A utility program for calculating solubilities with the GAMESSPLUS program has been added. This utility program computes the solubility for a solute in a given solvent by computing the standard-state free energy of solvation of the solute in the solvent using the SM5.42R solvation model and by computing the pure-solute vapor pressure of the solute (which is defined by the free energy of self-solvation of the solute) with SM5.42R. In addition, the user can specify a value for the vapor pressure instead of calculating it with SM5.42R. The Solubility Utility program is a PERL script, called solubility.pl. It has been added to the GAMESSPLUS distribution package. No modifications have been made to the GAMESSPLUS source code.

GAMESSPLUS Version 4.1 (January 2004)
Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS version: July 3, 2003 (R2)

- The generalized hybrid orbital (GHO) module is added into the *GAMESSPLUS* program to combine *ab initio* HF wave functions with molecular mechanics. Four orbital orthogonalization schemes for this purpose were implemented, with energy and analytical gradients available for all four. Using these methods requires building the *CHARMM/GAMESSPLUS* combination program, which is described in the *CGPLUS* user manual.

- Twelve modules (*bassto*, *gamess*, *grd1*, *grd2a*, *grd2b*, *grd2c*, *inputa*, *inputb*, *int1*, *int2a*, *rhfuhf*, *symorb*) of *GAMESS* have been modified to incorporate the GHO functionality in *GAMESSPLUS*. The GHO-related modifications can be located by the string "QLINK" in the corresponding *.src* files for these modules listed above. A new *gho* module is added to *GAMESSPLUS* in a file called *gho.src* to accommodate routines for defining GHO data structure, constructing the basis transformations, and performing gradient calculations.

- A file called *ghodum.src* has been added to compile *GAMESSPLUS* based on *GAMESS* in the absence of *CHARMM*. The *gho* module in *GAMESSPLUS* is written largely following the *CHARMM* programming style which utilizes include statements and conditional compilations through a pre-processor in *CHARMM*. For compiling the *CHARMM/GAMESSPLUS* as an integrated executable, locating the *CHARMM* pre-processor will be properly handled by the utility package *CGPLUS* when the *gho* module is compiled; therefore it is not a problem. However, for a user who is not planning to use GHO-AIHF in *GAMESSPLUS*, the compilation of *GAMESSPLUS* with *CHARMM* is not necessary, and the compilation of *GAMESSPLUS* with *GAMESS* itself is still needed to run *GAMESSPLUS* enhancements to *GAMESS*. To meet this need, the file *ghodum.src* has been created; it contains all necessary dummy routine required to compile *GAMESSPLUS* with *GAMESS*.

- A set of integral scaling factors is included for GHO-AIHF; the parameters were optimized for GHO-AIHF/MIDI! with the local Löwdin orthogonalization treatment.

- The comment lines to mark the *GAMESSPLUS* modifications have been changed from "CGAMESOLSTR" and "CGAMESOLEND" to "CGMSPLUSSTR" and "CGMSPLUSEND" to be consistent with the name of the program.
- In mpcgrd.src, the lines for reading the gradients of the CDS term and the gradients of Coulomb integrals from the direct access file records 269 and 270 are commented. These calls to DAREAD occurred even for gas-phase MOPAC-type calculations. For gas-phase only mopac type calculations, these resulted in an error because these records (269 and 270) were never created. To avoid this error, a separate subroutine GET_GRAD is called instead to handle both the gas-phase case and the solvation case properly.

The following three items are related to the SM5.42/UHF analytical gradients:

- A bug has been fixed to evaluate the density-weighted matrix correctly for SCRF calculations with UHF wave functions. The density force contribution in analytical gradients for UHF wave functions is:

$$\text{Density force} = \sum_{\mu\nu} W_{\mu\nu} \frac{\partial S_{\mu\nu}}{\partial R} \tag{1}$$

where **W** is the total energy-weighted density matrix (UHF expression):

$$W_{\mu\nu} = \sum_{i=1}^{N_\mathrm{A}} \varepsilon_i^\beta n_{\mathrm{occ},i}^\alpha c_{i\mu}^\alpha c_{i\nu}^\alpha + \sum_{i=1}^{N_\mathrm{B}} \varepsilon_i^\beta n_{\mathrm{occ},i}^\beta c_{i\mu}^\beta c_{i\nu}^\beta \tag{2}$$

In *GAMESSPLUS-v4.0*, the UHF energy-weighted matrix is evaluated in an alternative way:

$$W_{\mu\nu} = \begin{cases} \frac{1}{2}(\mathbf{P}^\alpha \mathbf{F}^\alpha \mathbf{P}^\alpha + \mathbf{P}^\beta \mathbf{F}^\beta \mathbf{P}^\beta)_{\mu\mu} & \mu = \nu \\ (\mathbf{P}^\alpha \mathbf{F}^\alpha \mathbf{P}^\alpha + \mathbf{P}^\beta \mathbf{F}^\beta \mathbf{P}^\beta)_{\mu\nu} & \mu \neq \nu \end{cases} \tag{3}$$

where **P** and **F** are the density matrix and the Fock matrix, for the alpha and the beta set, respectively. The expression in Eq. (3) is identical to Eq. (2) except that the diagonal elements of Eq. (3) are one-half of those in Eq. (2). However, they give identical results when the density force in Eq. (1) is evaluated, because all gradients of the diagonal overlap elements vanish (d$S_{\mu\mu}$/dR = 0).

Note that the above equivalence only holds if one uses the density matrix and the Fock matrix consistently. For example, in gas-phase calculations, both **P** and **F** are obtained without any reaction fields included. In solvation calculations, the solvation effect (for GB model, the mutual polarization of the solute and solvent through the *G*p term) is included as a correction term to the Fock matrix during the SCF procedure:

$$F_{\mu\nu} = F_{\mu\nu}^0 + \frac{\partial G_\mathrm{P}}{\partial P_{\mu\nu}} \tag{4}$$

Unfortunately, the Fock matrix in the conventional storage space is still the gas-phase Fock matrix $\mathbf{F}^{(0)}$. When *GAMESSPLUS* uses Eq. (3) for gradient in any SCRF calculations, the density matrix **P** is the solvated one, but the gas-phase Fock matrix $\mathbf{F}^{(0)}$ without the G_p correction will be used. This inconsistency makes the use of Eq. (3) fail to evaluate the energy-weighted density matrix correctly.

This bug only exists for UHF (and ROHF) cases. For RHF cases, *GAMESSPLUS-v4.0* evaluates the **W** matrix using Eq. (2) for one set of orbitals only.

In *GAMESSPLUS-v4.1*, we fix this bug. When SCRF is on, *GAMESSPLUS-v4.1* uses Eq. (2) instead of Eq. (3) to compute **W** matrix. Modifications have been made in module '*grd1.src*'.

- Modified implementation for gradients of the Mayer bond order based on UHF wave functions:

  In *GAMESSPLUS-v4.0*, the Mayer bond order has been correctly implemented as follows, both of which are correct for both RHF and UHF wave functions.

$$\mathbf{B}_{\mathrm{AB}} = \sum_{\lambda\in \mathrm{A}} \sum_{\omega\in \mathrm{B}} \left[(\mathbf{P}\ \mathbf{S})_{\omega\lambda} (\mathbf{P}\ \mathbf{S})_{\lambda\omega} + (\mathbf{P}^{\mathrm{S}}\mathbf{S})_{\omega\lambda} (\mathbf{P}^{\mathrm{S}}\mathbf{S})_{\lambda\omega} \right] \qquad (5)$$

$$= 2\sum_{\lambda\in \mathrm{A}} \sum_{\omega\in \mathrm{B}} \left[(\mathbf{P}^{\alpha}\mathbf{S})_{\omega\lambda} (\mathbf{P}^{\alpha}\mathbf{S})_{\lambda\omega} + (\mathbf{P}^{\beta}\mathbf{S})_{\omega\lambda} (\mathbf{P}^{\beta}\mathbf{S})_{\lambda\omega} \right] \qquad (6)$$

  where $\mathbf{P}^{\alpha}$ and $\mathbf{P}^{\beta}$ are the alpha and beta spin density matrices, respectively, $\mathbf{P} = \mathbf{P}^{\alpha} + \mathbf{P}^{\beta}$ is the spinless density matrix, and $\mathbf{P}^{\mathrm{S}} = \mathbf{P}^{\alpha} - \mathbf{P}^{\beta}$ is the spin density matrix. Note that for RHF wave functions, the spin density matrix $\mathbf{P}^{\mathrm{S}}$ is zero by definition. However, the gradients of the bond order are not implemented consistently with Eq. (5) in *GAMESSPLUS-v4.0*. Therefore, the GB/CM2/UHF (or SM5.42/UHF) gradients based on CM2 charges were incorrect.

  In the *GAMESSPLUS-v4.1*, the form of the $\partial\mathbf{B}/\partial\mathbf{R}$ part of $\partial G_{\mathrm{P}}/\partial\mathbf{R}$ (the energy gradient) changes due to the implementation of Eq. (5) for the Mayer bond order. Necessary modifications have been made in module '*smx.src*'.

- With the bug fix mentioned above and the correct implementation of the Mayer bond order derivatives, the new version of *GAMESSPLUS* gives the correct analytical gradients for both GB/Löwdin/UHF and SM5.42/UHF calculations. Therefore, *GAMESSPLUS-v4.1* now supports analytical gradients using unrestricted wave functions. Note that *GAMESSPLUS-v4.1* is still unable to carry out analytical gradients for SM5.42/ROHF wave functions. If ROHF is specified with SCRF on, the program prints an error message and stops.

GAMESSPLUS Version 4.2 (March 2004)

Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS version: July 3, 2003 (R2)

- The SM5.43 parameter sets for HF/6-31G* and B3LYP/6-31G* were added. The test suite was updated to test these new models. Modifications to subroutines SETSOL, DOSOLV, COULRD, SM5STN, and OSM5 in smx.src were made.

GAMESSPLUS Version 4.3 (July 2004)

Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS version: May 19, 2004 (R3)

- The *GAMESSPLUS* module was implemented into the May 19, 2004 (R3) version of *GAMESS*. This version of *GAMESS* provides PM3 parameters for lithium that were previously provided by the *GAMESSPLUS* module. Thus the *GAMESSPLUS* distribution no longer contains the file mpcdat.src.

GAMESSPLUS Version 4.3.1 (September 2004)
Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS version: May 19, 2004 (R3)

- A bug was fixed in the program. This bug was due to duplicate lines appearing in the file inputb.src. This bug was introduced in version 4.3.

GAMESSPLUS Version 4.4 (December 2004)
Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS version: May 19, 2004 (R3)

- A script to compile GAMESSPLUS was added. The script *modgms* will automatically make all the changes described in the installation section.

GAMESSPLUS Version 4.5 (January 2005)
Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS version: May 19, 2004 (R3)

- The mPW exchange functional was added for restricted and unrestricted calculations.
- The PW91 correlation functional was added for restricted and unrestricted calculations.
- The keywords MPW1K and mPW1PW91 are now options for DFTTYP in the $DFT input section.
- A numerical issue was resolved in the VWN3 and VWN5 functionals so that they now both work for hydrogen atom.

GAMESSPLUS Version 4.6 (February 2005)
Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS versions: November 22, 2004 (R1), May 19, 2004 (R3)

- This version works with two different versions of GAMESS.
- More helpful error messages are produced when the installation fails.

GAMESSPLUS Version 4.7 (August 2005)
Authors: J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS versions: November 22, 2004 (R1) plus others (see makepatch method for compiling)

An alternate way to compile *GAMESSPLUS* is now available. This way, called the "Makepatch Method for Compiling *GAMESSPLUS*" is explained in the manual. A script called "makepatch.pl" is located in the gmsplus4.7/Code directory, and a directory called "Patches" has been added under the gmsplus4.7 directory. This alternate method of compiling *GAMESSPLUS* was designed with making *GAMESSPLUS* compatible with multiple previous versions of *GAMESS* and future versions of *GAMESS* in mind.

The keyword MPWX is now an option for DFTTYP in the $DFT namelist. Note that using this keyword requires specifying the percentage of Hartree-Fock exchange (*X* in MPW*X*), with the new "HFE" keyword. Note that *X* is a percentage and HFE is a fraction. Thus *X*=100*HFE. This keyword is entered in the $DFT namelist for gas-phase calculations, and also in the $CM2 or $GMSOL namelist for CM3, CM4, SM5.43, or SM6 calculations. A test job, test 29, has been added to the test suite that tests the MPWX and HFE keywords.

The keywords IRADII, ISTS, and SOLVRD were added. The IRADII keyword specifies the set of atomic-number-dependent radii that are used to build the molecular cavity. The ISTS keyword determines whether SM5- or SM6-type functionals are used. The SOLVRD keyword specifies the solvent radius that is used for the SASA. The default value has been set to 0.40 Å. Coulomb radii for SM6 were added. In previous versions of the code, a default value of 0.0 Å was assigned to elements for which an intrinsic Coulomb radius had not been optimized. In version 4.7, the default atomic radius for elements for which an intrinsic Coulomb radius has not been optimized is Bondi's value for the van der Waals radius when available, and 2.0 Å for all other atoms.

The atomic radii used in the SCRF calculation (intrinsic Coulomb radii) are now printed out by default.

The CM3 parameter sets for MPW*X*/MIDI! (ICMD = 315), MPW*X*/MIDI!6D (ICMD = 316), MPW*X*/6-31G* (ICMD = 317), MPW*X*/6-31+G* (ICMD = 318), and MPW*X*/6-31+G** (ICMD = 319) were added and tested.

The CM3 and CM3.1 parameter sets for HF/MIDI! were added and tested (ICMD = 301 and ICMD = 322, respectively).

The CM4 parameter sets for DFT/MIDI!6D (ICMD = 416), DFT/6-31G* (ICMD = 417), DFT/6-31+G* (ICMD = 418), and DFT/6-31+G** (ICMD = 419) were added and tested.

The SM5.43 parameter sets for aqueous and organic solvents for MPW*X*/MIDI! (ICDS = 315), MPW*X*/MIDI!6D (ICDS = 316), MPW*X*/6-31G* (ICDS = 317), MPW*X*/6-31+G* (ICDS = 318), and MPW*X*/6-31+G** (ICDS = 319) were added and tested.

The SM6 parameter sets for aqueous solution for DFT/MIDI!6D (ICDS = 416), DFT/6-31G* (ICDS = 417), DFT/6-31+G* (ICDS = 418), and DFT/6-31+G** (ICDS = 419) were added and tested.

The SM6 test suite has been added. This test suite tests the ability of GAMESSPLUS to perform liquid-phase geometry optimizations and liquid-phase, single-point calculations with methods that use diffuse functions.

SCF Scheme I is no longer available for liquid-phase calculations that involve the use of diffuse basis functions, due to convergence problems.

If ISCRF=1 (SCF Scheme I) is specified for a method that uses diffuse functions, the code prints out a message and automatically switches to ISCRF=2 (SCF Scheme II).

A new method for installing the *GAMESSPLUS* Solubility Utility is used in this version of *GAMESSPLUS*. This new method uses the script called install.pl.

The *GAMESSPLUS* Soil Sorption Utility has been added. This program uses the same installation procedure as the Solubility Utility.

GAMESSPLUS Version 4.8 (May 2006)

Authors: A. C. Chamberlin, J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.-Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS versions: November 22, 2004 (R1) plus others (see makepatch method for compiling)

Minor corrections to the code were implemented

The atomic surface tensions used to compute the temperature dependence of the free energy of solvation have been added.

A new keyword (SOLK) to read in the temperature of liquid aqueous solution and to compute free energies of solvation at the given has been added. The keyword only accepts values within the range of 273 to 373 K

A new keyword (READK) to read in temperatures from a file named *KELT* and to compute free energies of solvation for a particular molecule in aqueous solution has been added. Note that this approach used the solution phase electronic structure as an initial guess for successive temperatures, so one may obtain identical electronic energies if the temperatures are very near one another. To avoid this either use very strict convergence criteria, or do not order similar temperatures very near one another. The differences in electronic energies are negligable anyway.

Another keyword (AVGK) which is similar to READK computes the electrostatics by computing the electronic energy of the solute in solution at the average of all the temperatures in the file *KELT* and then uses a scaling factor to compute the electronic energy at a given temperature. Note this option is considerably faster than the READK option.

An additional test suite for SM6T has been added.

GAMESSPLUS Version 2008 (April 2008)

Authors: M. Higashi, A. C. Chamberlin, J. Pu, J. D. Thompson, J. D. Xidos, J. Li, T. Zhu, G. D. Hawkins, Y.- Y. Chuang, P. L. Fast, B. J. Lynch, D. A. Liotard, D. Rinaldi, J. Gao, C. J. Cramer, and D. G. Truhlar
GAMESS versions: March 24, 2007 (R6) plus others (see makepatch method for compiling)

This version works with the March 24, 2007 (R6) version of the *GAMESS* program. Some modules of GAMESS have been modified. This version also works with the *CGPLUS*-v2008 package.

The capability to carry out electrostatically embedded quantum mechanical (EEQM) calculations with a site-site representation of the QM/MM electrostatic interaction has been added. RUNTYP=EEQM in the $CONTRL namelist carries out the EEQM calculation. The new namelist $EEQM is available.

Two new parameters, MXSATM and MXSBAS have been added in smx.src to reduce the memory requirement. MXSATM and MXSBAS control the maximum number of atoms and basis functions in the smx module. Now these parameters have been set at MXSATM=100 and MXSBAS=512.

The default parameter SolvRd in Namelists $GMSOL and $CM2 has been changed. Now SolvRd=0 Å for SM5.42 calculation, and SolvRd=0.40 Å for SM5.43 or SM6 calculations. In *GAMESSPLUS*-v4.7 and v4.8, SM5.42 calculations without specifying SolvRd=0 Å had errors because the default parameter SolvRd was set at 0.40 Å for all the solvation model calculations.

All the test input files have been rechecked. All the test output files have been replaced with those calculated by *GAMESSPLUS*-v2008. In addition, in many cases of the SM6 geometry optimization calculations, better initial guesses were provided for the test run geometries, and the natural coordinate option was turned on for some of the geometry optimizations.