Fourth and Fifth Virial Coefficients of Polarizable Water

Kenneth M. Benjamin,[†] Andrew J. Schultz,[‡] and David A. Kofke*[,‡]

Department of Chemical and Biological Engineering, South Dakota School of Mines and Technology, Rapid City, South Dakota 57701, and Department of Chemical and Biological Engineering, University at Buffalo, The State University of New York, Buffalo, New York 14260-4200

Received: February 10, 2009; Revised Manuscript Received: April 8, 2009

We report values of the virial coefficients B_4 and B_5 for the Gaussian charge polarizable model (GCPM) of water using the overlap-sampling implementation of Mayer sampling molecular simulation. These results supplement values for the lower-order coefficients B_2 and B_3 reported previously, and in the present work, we provide more precise values of these coefficients as well. The precision of all virial coefficients is such that the standard error in the calculated pressure is significantly less than 1% for most temperatures, with the exception of temperatures near the critical, where the error approaches 100% at the critical density, and supercritical, where the uncertainly in B_5 introduces an error of about 5% in the pressure at the critical density. We examine these coefficients in the context of the equation of state and molecular clustering. Comparisons are made to established molecular simulation data, quantum chemical calculations, and experimental data for real water. Over both sub- and supercritical temperatures, the virial series to B_5 is accurate for densities only up to about half the critical density. In this regime, deviation is observed from experimental data for real water, and it is suggested that further development of the model might do well to further improve the relatively good agreement it has with the experimental second virial coefficient of water. The virial coefficients are used to characterize molecular clusters (dimers through pentamers) in GCPM water under supercritical and saturated vapor conditions between 210 and 673 K.

1.0. Introduction

The virial equation of state[1,2] (VEOS) describes the gas-phase pressure−volume−temperature (*PVT*) behavior via the following expression:

$$\frac{P}{\rho kT} \equiv Z = 1 + B_2\rho + B_3\rho^2 + B_4\rho^3 + B_5\rho^4 + B_6\rho^5 + \ldots \tag{1}$$

where P is the pressure, ρ is the number density, k is the Boltzmann constant, T is the absolute temperature, Z is the compressibility factor, and B_i is the ith virial coefficient. These coefficients can be given directly in terms of configurational integrals of a few molecules interacting with a given intermolecular potential.[2,3]

In recent work we presented results for classical virial coefficients of several water models, using the Mayer sampling Monte Carlo method for calculation of cluster integrals.[4−6] We calculated coefficients up to B_6 for a host of pairwise water models,[5] as well as coefficients up to B_3 for the Gaussian charge polarizable model (GCPM),[6] which is a multibody potential proposed by Paricaud et al.[7] Results of the calculations for GCPM showed a remarkable agreement with experimental virial coefficient data for water, even though these data were not used to fit the model. This performance is in stark contrast to that of the pairwise models, which are optimized to describe the condensed-phase behavior and, being nonpolarizable, cannot also describe well the gas-phase behavior. In the present study, we apply the Mayer sampling method to calculate B_4 and B_5 for GCPM, and we revisit our previous calculations to produce new results having better precision.

In the next section, we review the mathematical formalism for calculating virial coefficients for systems with nonadditive intermolecular forces and describe a few new details of our calculation method. In section 3, we present results for the virial coefficients for GCPM water and discuss thermodynamic properties and molecular clustering predicted from various truncated virial equations of state. It should be noted that this is the first calculation of B_4 and B_5 for GCPM water and indeed, to our knowledge, for any multibody model. Finally, we conclude in section 4.

2.0. Formulas and Methods

Johnson and Spurling provide the formalism needed to compute virial coefficients for nonpairwise additive intermolecular forces.[8] The third virial coefficient, B_3, is represented as the sum of additive and nonadditive parts, as in

$$B_3(T) = B_{3,\mathrm{add}} + \Delta B_3 \tag{2}$$

where the pairwise additive component is of the standard form

$$B_{3,\mathrm{add}} = -\frac{1}{3}\left\langle \iint f_{12}f_{13}f_{23}\,\mathrm{d}\mathbf{r}_2\,\mathrm{d}\mathbf{r}_3 \right\rangle_{\Omega_1\Omega_2\Omega_3} \tag{3}$$

in which molecule 1 is fixed at the origin and $\Omega_1\Omega_2\Omega_3$ indicates averaging the integral over all configurations and orientations of the three respective water molecules. In this and all that follows, $f = e - 1$ is the Mayer function with $e = \exp(-\beta u)$ and u is the intermolecular energy. Subscripts on f, e, and u indicate that the quantity is calculated for interactions among the indicated molecules; two subscripts are a pairwise interaction, three are three-molecule interactions, and so on. These

* Corresponding author. Tel.: (716) 645-1173. Fax: (716) 645-3822. E-mail: kofke@buffalo.edu.
† South Dakota School of Mines and Technology.
‡ The State University of New York.

10.1021/jp901232m CCC: $40.75
Published on Web 05/12/2009

types of integrals are commonly referred to as cluster integrals.[2,3] The nonadditive term for B_3 is

$$\Delta B_3 = -\frac{1}{3}\left\langle \iint (e_{123} - e_{12}e_{13}e_{23})\mathrm{d}\mathbf{r}_2\,\mathrm{d}\mathbf{r}_3 \right\rangle_{\Omega_1\Omega_2\Omega_3} \quad (4)$$

Similarly, for B_4 we have[8]

$$B_4(T) = B_{4,\mathrm{add}} + \Delta B_4 \quad (5)$$

with the pairwise additive term

$$B_{4,\mathrm{add}} = -\frac{1}{8}\left\langle \iiint f_{12}f_{23}f_{34}f_{14}[3 + 6f_{24} + f_{24}f_{13}]\mathrm{d}\mathbf{r}_2\,\mathrm{d}\mathbf{r}_3\,\mathrm{d}\mathbf{r}_4 \right\rangle_{\Omega_1\Omega_2\Omega_3\Omega_4} \quad (6)$$

and the nonadditive term

$$\Delta B_4 = -\frac{1}{8}\left\langle \iiint (\Delta^{(a)}_{1234} + \Delta^{(b)}_{1234})\mathrm{d}\mathbf{r}_2\,\mathrm{d}\mathbf{r}_3\,\mathrm{d}\mathbf{r}_4 \right\rangle_{\Omega_1\Omega_2\Omega_3\Omega_4} \quad (7)$$

where

$$\Delta^{(a)}_{1234} = e_{1234} - e_{12}e_{13}e_{14}e_{23}e_{24}e_{34}$$

and

$$\Delta^{(b)}_{1234} = \sum_{(ijk)} \delta^{(b)}_{(ijk)}$$

$$\delta^{(b)}_{(ijk)} = (e_{ijk} - e_{ij}e_{ik}e_{jk})(f_{ij}f_{ik}f_{jk} + 1)$$

where the sum (b) is taken over all unique triplets of the four indices (four terms).

The expression for B_5 can be given in similar form. The pairwise additive term is available from other sources[2,3] and will not be reproduced here; the correction for nonadditivity is derived as was done for B_4,[2,8] with the result

$$\Delta B_5 = -\frac{1}{30}\left\langle \iiiint (\Delta^{(a)}_{12345} + \Delta^{(b)}_{12345} + \Delta^{(c)}_{12345} + \Delta^{(d)}_{12345})\mathrm{d}\mathbf{r}_2\,\mathrm{d}\mathbf{r}_3\,\mathrm{d}\mathbf{r}_4\,\mathrm{d}\mathbf{r}_5 \right\rangle_{\Omega_1\Omega_2\Omega_3\Omega_4\Omega_5} \quad (8)$$

where

$$\Delta^{(a)}_{12345} = e_{12345} - (e_{12} + e_{13} + e_{14} + e_{15} + e_{23} + e_{24} + e_{25} + e_{34} + e_{35} + e_{45})$$

$$\Delta^{(b)}_{12345} = \sum_{(ijkl)m} \delta^{(b)}_{(ijkl)m}$$

$$\delta^{(b)}_{(ijkl)m} = (e_{ijkl} - e_{ij}e_{ik}e_{il}e_{jk}e_{jl}e_{kl})(f_{im} + f_{jm} + f_{km} + f_{lm} + 1)$$

where the sum (b) is taken over all unique ways of forming the set ($ijkl$) (five terms). Also

$$\Delta^{(c)}_{12345} = \sum_{(ij)(kl)m} \delta_{(ij)(kl)m}$$

$$\delta_{(ij)(kl)m} = (e_{ijm} - e_{ij}e_{im}e_{jm})(e_{klm} - e_{kl}e_{km}e_{lm})(e_{ij}e_{im}e_{jm}e_{kl}e_{km}e_{lm})$$

where the sum (c) is taken over all unique ways of forming two pairs (ij) and (km) (15 terms). Finally

$$\Delta^{(d)}_{12345} = \sum_{(ijk)(lm)} \delta^{(d)}_{(ijk)(lm)}$$

$$\delta^{(d)}_{(ijk)(lm)} = (e_{ijk} - e_{ij}e_{ik}e_{jk})[2(e_{il}e_{im} + e_{jl}e_{jm} + e_{kl}e_{km} - e_{lm}) + f_{lm}(e_{il} + e_{im} + e_{jl} + e_{jm} + e_{kl} + e_{km}) + f_{il}f_{jm} + f_{il}f_{km} + f_{im}f_{jl} + f_{im}f_{kl} + f_{jl}f_{km} + f_{jm}f_{kl}]$$

with the sum (d) taken over all ways of partitioning the five into unique sets of two and three (10 terms).

All of our calculations in this study were performed on classical GCPM water.[7] The key features of this potential model are an exponential-6 representation of repulsion−dispersion interactions and partial charges given by spherical Gaussian charge distributions. These features allow the GCPM to successfully describe the microstructure of liquid water, vapor−liquid equilibrium, thermodynamic properties, transport properties, and even water dimer energies. It should be noted that because the GCPM is polarizable, its electrostatics must be solved self-consistently for each configuration. As outlined in the literature,[7] the electric fields on each molecule and induced dipoles of each molecule must be resolved simultaneously. In this work, the self-consistent solution results from direct matrix inversion of the field equations, rather than through iteration.

It is well-known that quantum effects contribute significantly to the virial coefficients of real water, especially at lower temperatures. By performing entirely classical calculations, we are assuming that these effects are incorporated in some "effective" manner in the GCPM potential.

The Mayer-sampling Monte Carlo (MSMC) method has been presented in detail in previous work,[4−6,9] so the basic procedure will not be described again here. We note that the calculation is performed in an infinite volume with no periodic images or other containing structures. Thus, there is no truncation of the potential, and no Ewald sum or other long-range correction needs to be applied in the calculation. We use the overlap-sampling formulation of MSMC in all of the calculations reported here.

One important feature of these simulations, as compared to our original work,[6] is the application of a coarse orientational average with every configuration. By this we mean that every nominal configuration of molecules in the simulation is for every purpose (i.e., generating configurations and collecting averages) treated as an average over multifold configurations. In particular, for each nominal configuration of n water molecules, all 2^n configurations obtained by inverting each of the molecules through its geometric center are generated (the geometric center is the unweighted average of the atom coordinates; other reasonable choices for the inversion center are also acceptable). The Mayer-sampling integrand γ (used for the configuration weights and averages) for the nominal configuration is then defined as the average of the values over all these configurations. Such averaging of rotated molecules can be visualized as in Figure 1. In this figure, we see two individual water molecules after a given Monte Carlo move (a). The integrand, γ, is computed from the "a" configuration. Then molecule 1 is inverted through its center (b) and γ is computed for the "b" configuration. Then molecule 2 is inverted through its center (c) and γ is computed for the "c" configuration. Then molecule 1 is inverted back to its original configuration and γ for this final configuration is computed, and the average of the four configurations (a−d) is taken as the value for that nominal configuration (and its absolute value is used to define the sampling weight π). After this, molecule 2 was inverted back through its center to its original position, thereby restoring the original configuration of the two water molecules. The phase-

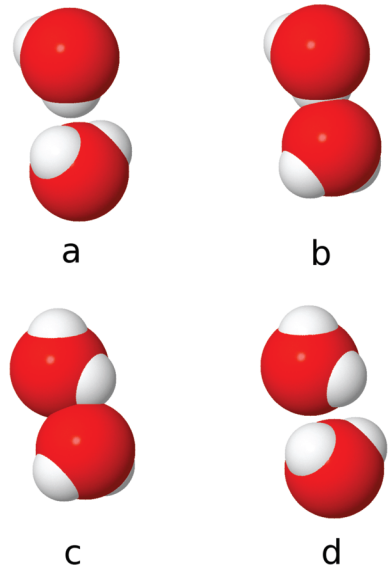

Figure 1. Illustration of rotation−inversion of individual molecules at each Monte Carlo simulation step, to improve averaging and sampling.

space volume associated with each combination is equal, and independent of coordinates, so no special term in the Monte Carlo acceptance criterion is needed in this regard to ensure microscopic reversibility. Further, the sampling weight is divided out in the quantity that we average, so the use of the orientational sum for sampling does not leave any bias in the simulation averages.

This coarse orientational average has the effect of attenuating the potential at large separations, thereby reducing the tendency of the system to sample such separations out of proportion to their importance (i.e., correctly, but inefficiently). The problem is related to the long-range electrostatic features of the potential, and it is significantly exacerbated by the polarizability of the model. It was due to this inherent poor sampling in the absence of orientation averaging that we introduced the small, albeit arbitrary, transition to pairwise additive interactions at 100 Å in our previous work to calculate B_3.[6] With the inclusion of this orientation averaging, we believe these values for B_2 and B_3 are improvements over those published previously, and the newer results for B_4 and B_5 are likewise of higher quality than they would be without it.

Simulations of B_2 and B_3 each sampled 2×10^9 Monte Carlo trials; for B_4 we sampled 6×10^8 configurations, and with B_5, between 1×10^8 and 1.7×10^9 samples were taken (depending on temperature). All of these numbers represent samples of nominal configurations, so if the different configurations generated by the orientation sampling are added, then these numbers increase by a factor of 2^n for coefficient B_n. For each virial coefficient at each specific temperature, an average and standard error of the mean was computed from the respective block averages and is reported with the results.

3.0. Results and Discussion

3.1. Virial Coefficients. Table 1 presents the virial coefficients B_2-B_5 for GCPM water studied in this work. The values display the general conventional behavior of transitioning from negative to positive values as temperature increases. Figure 2 displays these trends by plotting $B_n\rho_c^{n-1}$ versus T/T_c, where n is the order of the coefficient, and ρ_c and T_c are the GCPM critical density (0.3344 g/mL) and temperature (642.21 K), respectively, according to Paricaud et al.[7] The curves are

TABLE 1: Virial Coefficients for GCPM Water[7] As Calculated Using the Mayer-Sampling Method[a]

T (K)	B_2 (L/mol)	B_3 (L/mol)^2	B_4 (L/mol)^3	B_5 (L/mol)^4
210	−12.436(6)	$-1.69(1) \times 10^4$	$-5.8(5) \times 10^8$	
230	−5.979(3)	$-1.65(1) \times 10^3$	$-8.4(5) \times 10^6$	
250	−3.309(2)	−254(1)	$-2.4(1) \times 10^5$	$-5.6(6) \times 10^7$
270	−2.037(1)	−54.0(2)	$-1.15(4) \times 10^4$	$-9.6(8) \times 10^5$
290	−1.3594(6)	−14.68(4)	−980(24)	$-3.0(3) \times 10^4$
298	−1.178(1)	−9.29(3)	−423(9)	$-8.5(6) \times 10^3$
331	−0.7103(3)	−1.795(4)	−19.3(4)	−120(8)
350	−0.5572(3)	−0.807(3)	−4.40(7)	−15(1)
373	−0.4302(3)	−0.344(1)	−0.923(15)	−1.6(1)
400	−0.3307(1)	−0.1410(7)	−0.180(3)	−0.16(1)
420	−0.2778(1)	−0.0773(3)	−0.061(1)	−0.035(4)
438	−0.2405(1)	−0.0469(2)	−0.0233(5)	−0.0067(6)
456	−0.2105(1)	−0.0290(2)	−0.0094(2)	−0.0012(2)
474	−0.1856(1)	−0.01816(7)	−0.0036(1)	0.00018(7)
492	−0.16501(7)	−0.01167(5)	−0.00102(5)	0.00036(4)
510	−0.14770(5)	−0.00741(3)	−0.00011(3)	
528	−0.13287(5)	−0.00474(3)	0.00030(2)	0.00018(1)
546	−0.12010(4)	−0.00300(1)	0.00041(2)	
564	−0.10894(4)	−0.00181(1)	0.00041(2)	$7.0(6) \times 10^{-5}$
582	−0.09928(3)	−0.00104(1)	0.00036(1)	
590	−0.09532(3)	−0.00077(2)	0.00034(1)	
600	−0.09072(2)	−0.000508(7)	0.00032(1)	$2.4(2) \times 10^{-5}$
610	−0.08645(3)	−0.000274(4)	0.000294(7)	
642.21[b]	−0.07426(2)	0.000156(5)	0.000218(5)	$7(1) \times 10^{-6}$
648	−0.07235(2)	0.000221(4)	0.000204(8)	
673	−0.06470(2)	0.000374(4)	0.000157(3)	$1.5(6) \times 10^{-6}$
723	−0.05221(2)	0.000501(4)	$-9.8(4) \times 10^{-5}$	$-2(4) \times 10^{-7}$

[a] Numbers in parentheses represent the confidence limits (standard error of the mean) for the rightmost digits of the value. [b] GCPM critical temperature.[7]

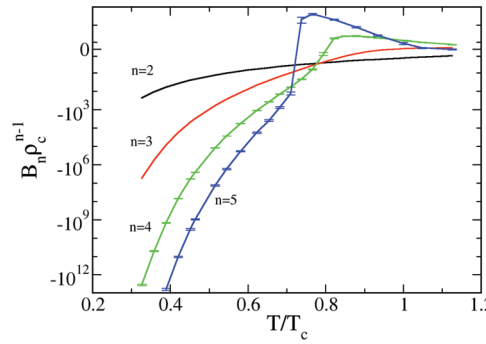

Figure 2. Temperature dependence of virial coefficients B_2-B_5 for GCPM water, plotted as $\sinh^{-1}(B_n\rho_c^{n-1})$ versus T/T_c (where n is the coefficient order and ρ_c and T_c are the critical density and temperature of GCPM water, respectively[7]).

presented on an inverse hyperbolic−sine scale, which is like a log scale when far from zero and is linear in the vicinity of zero; this permits both positive and negative values to be presented together. The virial coefficients of GCPM are as a whole significantly larger in magnitude than virial coefficients of a nonpolar model such as simple Lennard-Jones (LJ). We mean here that the values are larger when put on a corresponding basis and made dimensionless by their respective critical densities. The large and broad peaks that B_4 and B_5 each exhibit at temperatures where they first move into positive values are damped almost to the point of being indiscernible in the LJ model when presented on a critical-density scale. Another way of looking at this is to see the critical density of water as being in some sense anomalously large.

Comparison of the low-order coefficients with experiment was done by us previously,[6] and these more precise values for B_2 and B_3 only further improve the excellent agreement between GCPM water and real water over a wide range of temperatures. With regard to B_4 and B_5, the values we report herein (see Table

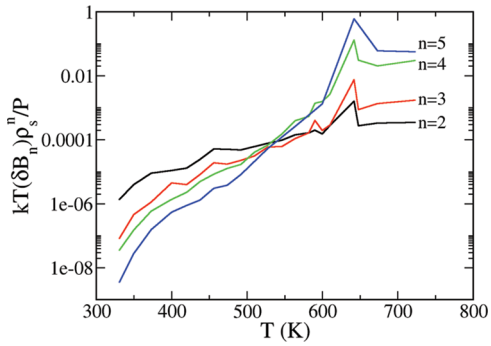

Figure 3. Contribution to the standard error in the pressure from each virial coefficient at a characteristic density, expressed as a fraction of the total pressure.

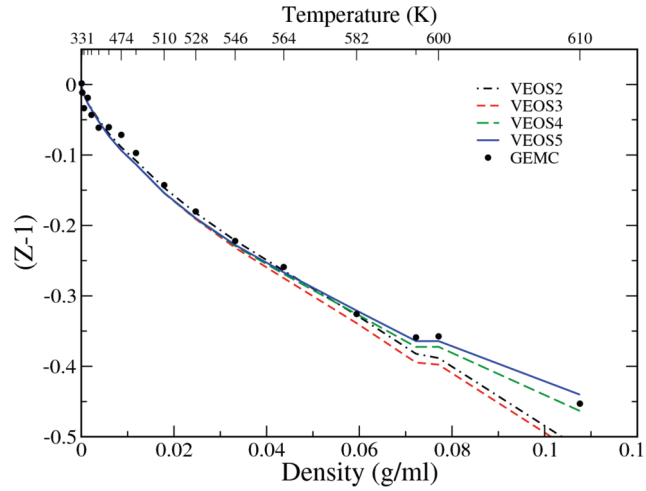

Figure 4. Deviation from ideality along the saturated vapor line of GCPM water. Lines are VEOS predictions at different levels of truncation (according to the legend), and points are molecular simulation data from Paricaud et al.[7] Virial coefficients required at temperatures where direct MSMC calculation was not performed were obtained using an interpolation scheme described elsewhere.[23]

1) are the first reported values for any polarizable water model. As it is too difficult to extract B_5 from *PVT* data, there are no values for real water with which to compare. Hill and MacMillan[10] have reported an analysis of experimental data and give results ostensibly for B_4, along with additional data for B_2 and B_3. Their data for B_2 and B_3 are consistent with values given in other studies and thus are in agreement with ours. Their B_4 data differ markedly from GCPM water—Hill and MacMillan data grow into large positive values with decreasing temperature, whereas the GCPM B_4 takes on even larger negative values; thus, below 528 K our results do not even match the sign of theirs. This could reflect a deficiency in GCPM water, but we do not have sufficient confidence in the Hill−MacMillan B_4 to take this conclusion. First, although we do not have error estimates for the Hill−MacMillan values, we know that stochastic errors in the input data will affect B_4 much more than B_2 (errors are propagated by amounts of order ρ^{-1} and ρ^{-3}, respectively, for the regressed B_2 and B_4, where ρ is a characteristic density). Thus, it is possible that the experimental data are not sufficiently precise to yield a reliable B_4, even though B_2 is given well. Second, it should be noted that Hill and MacMillan compute the coefficients by a straightforward fit of VEOS4 to *PVT* and other data up to saturation conditions. Consequently, the coefficients they regress are not true virial coefficients, which must be determined by examining the *PVT* behavior explicitly in the limit of zero pressure. One can show for example that even if the contribution of B_5 to the pressure is negligible (being of order ρ^5), its neglect can introduce significant errors (order ρ) to the regressed B_4.

The magnitude of the stochastic error in the coefficients we report is best considered in terms of the effect it has on the uncertainty in the pressure when it is added to the virial series, expressed as a fraction of the total pressure. Thus, we consider the quantity $kT(\delta B_n)\rho_s^n/P$, where δB_n is the standard error for coefficient B_n at a given temperature, ρ_s is a characteristic density for this temperature (for subcritical temperatures this is taken as the saturated-vapor density, and for supercritical temperatures this is the critical density), and P is the best available value of the GCPM pressure at T and ρ_s. This fractional error contribution is plotted as a function of temperature in Figure 3. We see that over most temperatures the fractional contribution from each coefficient is less than—often well less than—1%. In the vicinity of the critical temperature, there is a noticeable spike in the relative uncertainty of all coefficients, and in particular, the contribution of the uncertainly in B_5 approaches 100% of the total pressure. Beyond this point the uncertainty contributions drop and level off at about 5% for B_5 down to 0.03% for B_2. This figure suggests that more precision in the total pressure from VEOS could be obtained by carefully reallocating computational effort among the coefficients (such that the incremental improvement in precision for a given increase in computational effort was equal for all n). The balance, however, would vary with the density where VEOS is being applied.

3.2. Thermodynamic Properties. We examine now the *PVT* behavior for saturated water vapor, which represents the largest density of interest for the subcritical vapor. We compare the deviation from ideality, $(Z - 1)$, for GCPM water as determined from various truncated VEOS to values determined from Gibbs ensemble Monte Carlo (GEMC) simulations with GCPM water.[7] In making this comparison, we use the known saturation temperature and vapor density for the model as input to the VEOS and compute the saturation pressure. This, in turn, can be used to compute the quantity $(Z - 1)$. It should be noted that temperature varies in this comparison, increasing with the saturation density.

Figure 4 displays the deviations from ideality as a function of saturated vapor density for the GEMC results as well as virial equations of state truncated at the second through fifth orders (VEOS2, VEOS3, VEOS4, and VEOS5, respectively). The figure shows that all VEOS agree well with each other and with GEMC results out to a density of around 0.06 g/mL. At densities greater than 0.06 g/mL and out to 0.11 g/mL, one finds that VEOS4 and VEOS5 agree with simulation data much better than the simpler VEOS2 and VEOS3 equations. Beyond this density, VEOS4 and VEOS5 both start to trend sharply upward, clearly deviating from the GEMC curve (this is not shown in the figure because the VEOS plots require GEMC data and none are available from the last point shown in the figure to the critical point; the reliance of the VEOS plots on the GEMC data also introduces the "kink" in the VEOS curves at about 0.075 g/mL). We conclude that VEOS5 for GCPM is not converged at saturation densities for temperatures between 600 K and the critical temperature.

In Figure 5 we compare the deviations from ideality of all VEOS for GCPM water with a highly accurate equation of state for real water[11,12] under supercritical conditions. There is noticeable deviation between VEOS and the real-water behavior. This is observable even at low density, where it is clear that the VEOS converged. This discrepancy can only be attributed to some inaccuracy in the GCPM water model. Comparison of

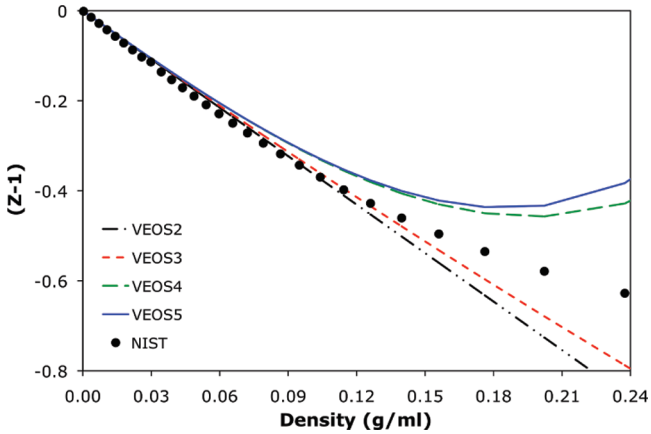

Figure 5. Deviation from ideality for supercritical GCPM water at 673 K. Lines are VEOS predictions at different levels of truncation (according to the legend), and points are results from a highly accurate equation of state for real water.[11,12]

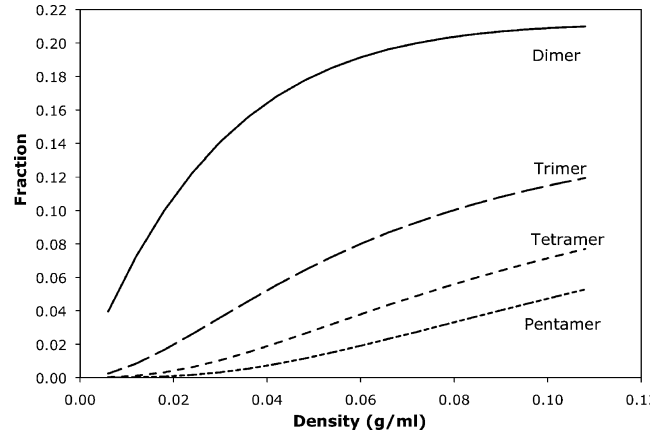

Figure 6. Mole fractions of dimer to pentamer water clusters for supercritical GCPM water at 673 K.

second virial coefficients, and to a lesser extent third virials, to experiment indicates that these coefficients are given accurately by the model; however, this is only in comparison to alternative water models. The GCPM B_2 is slightly greater than experimental values of B_2. This discrepancy is consistent with the deviations that we are observing in Figure 5. The figure further indicates that VEOS3 is converged at this temperature up to a density of about 0.06 g/cm^3. Beyond this point, VEOS4 and VEOS5 split off while remaining close to one another up to 0.18 g/cm^3, at about which differences between them begin to emerge.

3.3. Molecular Clustering. In this section, we examine the phenomenon of molecular clustering as understood through virial coefficients, in a fashion consistent with earlier work.[6,9,13] Briefly, the approach treats the gas as an ideal mixture of noninteracting clusters, which are in "chemical" equilibrium in a manner that is consistent with the density dependence expressed by the virial equation. Equivalently, the clusters may be interpreted in terms of the coefficients of the activity series for the grand partition function, from which the concentration of clusters of a given size follows directly, if given a value of the activity.[2]

In our previous study with GCPM, we could present analysis only of dimer and trimer clusters, because we had only B_2 and B_3 data. We extend those results here to include clusters of up to five molecules. We will keep the discussion brief because most of the qualitative points of interest were stated in our previous work[6] and are not significantly changed by the new results.

3.3.1. Supercritical Conditions. The importance and general features of molecular clustering to chemistry in supercritical water has been discussed and studied previously.[6,9,14−17] Figure 6 shows detailed information about clustering for supercritical GCPM water. At 673 K, there is significant formation of dimers and non-negligible formation of trimers, tetramers, and pentamers under supercritical conditions at densities less than 0.11 g/mL.

3.3.2. Saturated Vapor Conditions. Examination of clustering in the saturated vapor can be performed in reference to the molecular simulation data of Johansson et al.[18] for the SPC/E model for water.[19] Figure 7a,b shows the data for clusters of size n, for n equal to 2 (dimer) through 5 (pentamer). One expects at lower temperatures for the fixed-charged model with a larger dipole moment to exhibit greater clustering, and this is still seen. At higher temperatures the clustering is in remarkable

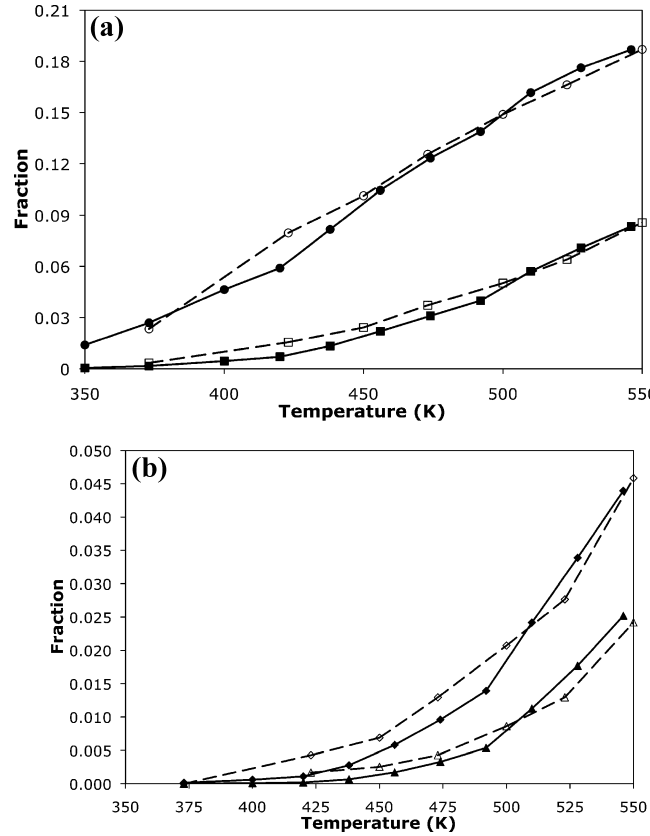

Figure 7. Fraction of water molecules in clusters in the vapor phase along the coexistence curve of water: GEMC simulations with SPC/E water[18] (H-bonded clusters) vs virial-clustering approach with GCPM water. (a) Dimer, (●) virial/GCPM and (○) GEMC/SPCE; trimer, (■) virial/GCPM and (□) GEMC/SPCE. (b) Tetramer, (◆) virial/GCPM and (◇) GEMC/SPCE; pentamer, (▲) virial/GCPM and (△) GEMC/SPCE. Virial coefficients required at temperatures where direct MSMC calculation was not performed were obtained using an interpolation scheme described elsewhere.[23]

agreement, even more so than seen in our previous study considering dimers and trimers. The agreement is no doubt fortuitous, resulting from the cancellation of the effect of increased clustering for SPC/E over GCPM and the use of a more restrictive (chemical) definition of a cluster for the GEMC simulations versus the VEOS calculations. To demonstrate further, in Figure 8 we compare concentrations of tetramers and pentamers derived from the virial coefficients for GCPM and the SPC/E models of water. We see that when calculated on

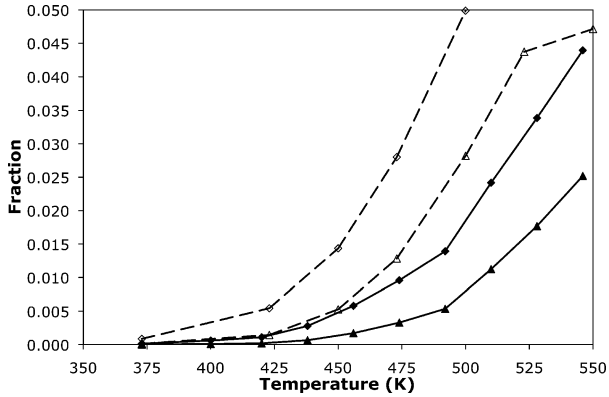

Figure 8. Fractions of water molecules in clusters in the vapor phase along the coexistence curve of water: virial-clustering approach with SPCE water[9] vs virial-clustering approach with GCPM water. Tetramer, (◆) virial/GCPM and (◇) virial/SPCE; pentamer, (▲) virial/GCPM and (△) virial/SPCE. Virial coefficients required at temperatures where direct MSMC calculation was not performed were obtained using an interpolation scheme described elsewhere.[23]

TABLE 2: Comparison of Water Cluster Concentrations in Saturated Vapor at 298 K with $P^{vap}(H_2O) = 0.031$ atm—Computational Quantum Chemistry[20] vs Virial Coefficient Method for GCPM Water

	cluster concentration (clusters/cm^3)	
cluster	from ab initio	from virial coefficient
dimer	9.0×10^{14}	1.1×10^{15}
trimer	2.6×10^{12}	9.0×10^{12}
tetramer	5.8×10^{11}	2.8×10^{11}
pentamer	3.5×10^{10}	7.3×10^{9}

this common basis the clustering in GCPM is significantly less than that of SPC/E.

3.3.3. Ambient Conditions. Table 2 compares the results of Dunn et al.,[20] which are based on ab initio computational chemistry calculations for water molecules in specific energetically favorable arrangements ("chemical clusters"), with cluster concentrations calculated from GCPM virial coefficients for saturated water vapor at 298 K (which includes loosely associated "physical clusters"). The virial-based GCPM concentrations of dimers and trimers are unchanged from what we reported previously and are larger than the ab initio values. On the other hand, the virial-based GCPM tetramers and pentamers are roughly one-half and one-fifth the respective ab initio values. The differences between ab initio and virial tetramers and pentamers, with virial (chemical + physical clusters) being less than ab initio (chemical only), is certainly unexpected and might be best explained by potential limitations or failings of the GCPM model to accurately describe multibody interactions for the gas phase.

4.0. Concluding Remarks

We have demonstrated that virial coefficients up to B_5 for a realistic, nonpairwise additive molecular model can be calculated using the MSMC method. Overall, for GCPM the virial equation of state truncated at B_5 can describe the *PVT* behavior of GCPM water up to the saturation conditions, for temperatures less than about 600 K, or about half the critical density. This behavior extends into the supercritical region, where again we found VEOS4 and VEOS5 to agree only to about half the critical density. The performance contrasts markedly from what we have seen in nonpolar molecules, such as simple Lennard-Jones[4,21] or even alkanes,[22] where VEOS5 provides a good characterization of *PVT* behavior up to the critical density.

The ability of the VEOS of GCPM water to describe real water is a separate matter. Although GCPM exhibits good relative agreement with the experimental values of B_2, the differences produce deviation from real-water *PVT* behavior well before the VEOS itself fails. Further improvements to the GCPM may be found through tuning to obtain even better agreement with the experimental B_2. Our studies indicate that this may significantly improve the ability of the model to describe gas-phase water *PVT* behavior.

We believe that further progress in describing the *PVT* behavior of water to enable application to conditions approaching the critical density will require significantly more coefficients to be included in the series. For water, a more effective approach to the formulation of a thermodynamic model from molecular principles is likely to be found through a different formalism than the standard virial equation of state.

Acknowledgment. Funding for this research was provided by Grants CTS-0414439 and CHE-0626305 from the U.S. National Science Foundation. Support has also been provided by the University at Buffalo School of Engineering and Applied Sciences. Computational support was provided by the University at Buffalo Center for Computational Research. Additional resources were provided by the Open Science Grid, which is supported by the National Science Foundation and the U.S. Department of Energy's Office of Science.

References and Notes

(1) Reed, T. M.; Gubbins, K. E. *Applied Statistical Mechanics*; McGraw-Hill: New York, 1973.

(2) Mason, E. A.; Spurling, T. H. *The Virial Equation of State, The International Encyclopedia of Physical Chemistry and Chemical Physics, Topic 10: The Fluid State, Volume 2*; Pergamon Press Ltd., Oxford, 1969.

(3) McQuarrie, D. A. *Statistical Mechanics*; University Science Books: Sausalito, CA, 2000; Gray, C. G.; Gubbins, K. E. *Theory of Molecular Fluids. Volume 1: Fundamentals*; Oxford University Press: London, 1984.

(4) Singh, J. K.; Kofke, D. A. *Phys. Rev. Lett.* **2004**, *92*, 220601. Erratum: Singh, J. K.; Kofke, D. A. *Phys. Rev. Lett.* **2005**, *94*, 249903.

(5) Benjamin, K. M.; Singh, J. K.; Schultz, A. J.; Kofke, D. A. *J. Phys. Chem. B* **2007**, *111*, 11463.

(6) Benjamin, K. M.; Schultz, A. J.; Kofke, D. A. *J. Phys. Chem. C* **2007**, *111*, 16021.

(7) Paricaud, P.; Predota, M.; Chialvo, A. A.; Cummings, P. T. *J. Chem. Phys.* **2005**, *122*, 244511.

(8) Johnson, C. H. J.; Spurling, T. H. *Aust. J. Chem.* **1974**, *27*, 241.

(9) Benjamin, K. M.; Schultz, A. J.; Kofke, D. A. *Ind. Eng. Chem. Res.* **2006**, *45*, 5566.

(10) Hill, P. G.; MacMillan, R. D. C. *Ind. Eng. Chem. Res.* **1988**, *27*, 874.

(11) Lemmon, E. W. ; McLinden, M. O.; Friend, D. G. Thermophysical Properties of Fluid Systems. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G.; http://webbook.nist.gov/chemistry/fluid.

(12) Wagner, W.; Pruss, A. *J. Phys. Chem. Ref. Data* **2002**, *31*, 387.

(13) Woolley, H. W. *J. Chem. Phys.* **1953**, *21*, 236.

(14) Akiya, N.; Savage, P. E. *Chem. Rev.* **2002**, *102*, 2725.

(15) Mountain, R. D. *J. Chem. Phys.* **1999**, *110*, 2109.

(16) Boero, M.; Terakura, K.; Ikeshoji, T.; Liew, C. C.; Parrinello, M. *J. Chem. Phys.* **2001**, *115*, 2219.

(17) Kalinichev, A. G.; Churakov, S. V. *Fluid Phase Equilib.* **2001**, *271*, 183–184.

(18) Johansson, E.; Bolton, K.; Ahlstrom, P. *J. Chem. Phys.* **2005**, *123*, 024504.

(19) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.

(20) Dunn, M. E.; Pokon, E. K.; Shields, G. C. *J. Am. Chem. Soc.* **2004**, *126*, 2647.

(21) Schultz, A. J.; Kofke, D. A. *Mol. Phys.* **2009**(in preparation).

(22) Schultz, A. J.; Kofke, D. A. *J. Chem. Phys.* **2009**(in preparation).

(23) Schultz, A. J.; Kofke, D. A. *Mol. Phys.* **2009**(in press), DOI: 10.1080/00268970902922633.

JP901232M