JCTC Journal of Chemical Theory and Computation

A Coarse-Grained Model Based on Morse Potential for Water and *n*-Alkanes

See-Wing Chiu,*,† H. Larry Scott,‡ and Eric Jakobsson§

Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana, Illinois 61801, Department of Biological, Chemical and Physical Sciences, Illinois Institute of Technology, Illinois 60616, Department of Molecular and Integrative Physiology, Beckman Institute for Advanced Science and Technology, Department of Biochemistry, Center for Biophysics and Computational Biology, University of Illinois, Urbana, Illinois 61801

Received September 8, 2009

Abstract: In order to extend the time and distance scales of molecular dynamics simulations, it is essential to create accurate coarse-grained force fields, in which each particle contains several atoms. Coarse-grained force fields that utilize the Lennard-Jones potential form for pairwise nonbonded interactions have been shown to suffer from serious inaccuracy, notably with respect to describing the behavior of water. In this paper, we describe a coarse-grained force field for water, in which each particle contains four water molecules, based on the Morse potential form. By molecular dynamics simulations, we show that our force field closely replicates important water properties. We also describe a Morse potential force field for alkanes and a simulation method for alkanes in which individual particles may have variable size, providing flexibility in constructing complex molecules comprised partly or solely of alkane groups. We find that, in addition to being more accurate, the Morse potential also provides the ability to take larger time steps than the Lennard-Jones, because the short distance repulsion potential profile is less steep. We suggest that the Morse potential form should be considered as an alternative for the Lennard-Jones form for coarse-grained molecular dynamics simulations.

1. Introduction

There has been a growing interest in applying coarse-grained (CG) superatom models for molecular dynamics (MD) simulations for a variety of polymers. These CG models consist of superatoms (or beads) that represent groups of atoms, or even several molecules. Coarse graining makes it possible to study larger systems at remarkably longer time scales with acceptable levels of detail. An overall protocol for moving between atomically detailed models and coarse-grained models for polymers by coarse graining and reverse coarse graining was described by Müller-Plathe.[1] The central method is to use a potential of mean force derived from atomistic simulations to parametrize the coarse-grained model, although the parameters can also be tuned by reference to thermodynamic properties. The coarse-grained simulations serve as an "express highway" to an equilibrated state, which is then reverse coarse-grained to atomic detail to reveal a detailed structure of the equilibrated state. A detailed example of the Müller-Plathe approach to coarse graining is given in Reithe et al.[2] An example of the full cycle of coarse-graining followed by simulation (in this case, Monte Carlo), in turn followed by reverse mapping, is given in Spyriouni et al.[3] Wilson et al.[4] provided a review of multiscale simulations as applied to liquid crystals, including macromolecular liquid crystals. A particular feature of simulations of liquid crystals is the ability to use hybrid models in which one component of a molecule (the mesogenic component that induces the ordering characteristic

* Corresponding author e-mail: s-chiu@illinois.edu.

† Beckman Institute for Advanced Science and Technology, University of Illinois.

‡ Illinois Institute of Technology.

§ Center for Biophysics and Computational Biology, University of Illinois.

10.1021/ct900475p
Published on Web 02/17/2010

of liquid crystals) is modeled as being completely rigid, while the rest of the molecule that determines the liquid character is modeled as flexible.[4] Prampolini[5] utilized a variant of this approach in which the properties of the mesogenic component are parametrized by ab initio calculations.

The MARTINI CG force field field[6,7] is especially designed for the simulation of lipids, surfactants, and proteins. This force field is parametrized by extensive calibration of the chemical building blocks of the CG force field against thermodynamical data. Because the MARTINI CG force field is especially designed for a variety of biomolecular simulations (our group's primary application interest), and because it is currently under active development, we take it as a primary reference point for our work in this area. However, we note the existence of coarse-grained simulations of amphiphilic molecules in aqueous environments using other force fields, such as Loison et at,[8] Shelley et al.,[9] and Markvoort et al.[10] A polarizable CG water model has been proposed by Ha-Duong et al.[11]

Another coarse-grained force field of biomolecular interest is the M3B force field,[12] designed to simulate maltooligosaccharides and their interaction with water. A theoretical point of interest is that the nonbonded non-Coulombic interactions in M3B are done with a Morse potential,[13] rather than with a potential of the Lennard-Jones form. In the M3B force field, each water molecule is represented by a single particle interacting with others by a Morse potential. The Morse potential was originally developed as a mathematically efficient way to describe chemical bond formation and dissociation.[13] In ref 12, the applicability of the Morse potential is effectively extended to the making and breaking of hydrogen bonds. Liew and Masuhiro[14] suggest that the application of a Morse-like potential may be an effective general strategy to improve the ability of coarse-grained simulations to properly represent phase changes and coexistence. This point speaks directly to a major problem in the MARTINI force field, which is based on using a Lennard-Jones (LJ) 12−6 interaction potential.[15] The major problem is the fact that its water model freezes at physiological temperatures.[16]

In this paper, we report on the application of a Morse potential to particle−particle interactions in a water model at the MARTINI level of coarse graining, which is four water molecules per particle. We will compare the behavior of this new coarse grained water model to the MARTINI model and to a coarse grained water model due to Shelly et al.,[9] where they have applied a softer 6−4 potential to their CG water model which represents a group of three water molecules. We also develop CG *n*-alkane models based on using the Morse potential.[13] We will refer to the CG water model of Shelley et al.[9] as SSRBK, the acronym of the last names of the authors of the reference. Similarly, we will refer to our models as CSJ.

2. Morse potential

The Morse potential $V_M(r)$, has the form

$$V_M(r) = \varepsilon[e^{\alpha(1-\frac{r}{R_0})} - 2e^{1/2\alpha(1-\frac{r}{R_0})}] \quad (1)$$

In eq 1, R_0 is the distance of the minimum energy ε and α is a parameter that measures the curvature of the potential around R_0. The smaller the value of α, the softer is the potential. In the condensed phase, the density of the system is mainly affected by the value of R_0, the cohesive energy by ε, and the compressibility by α.[12]

In the present work, we use eq 1 as the pair interaction potential function for the CG *n*-alkanes and water (group of four water molecules). The adjustable parameters ε and R_0 are parametrized to fit the experimental densities and heats of vaporization of liquid *n*-alkanes and water via CG simulations. Computational detail is to be described in the Computational Method section. The selection of the values of α for the CG models for *n*-alkanes and water is based on the agreement of the simulated vapor−liquid interfacial tensions with the experimental data.

3. Computational Method

The recent 43A1-S3 atomistically detailed force field for hydrocarbons and lipids[17] was used to carry out MD simulations of the bulk phase for each of the *n*-alkanes from *n*-butane to *n*-heptadecane. The MD trajectories of these hydrocarbon liquids were taken from our previous work.[17] Intramolecular data from these atomistically detailed simulations were mapped according to the CG models as described below and were used in the course of CG parametrization.

GROMACS 4.0.4 modeling software[18] was used for MD simulations performed in this work. User-custom lookup tables for the Morse potential (eq 1) were prepared according to the format as described in the manual of the software,[19] and the LJ parameters in the interaction parameter file for CG atoms were all set to 1.0.

For CG-MD simulations, a time step (Δt) size of 40 fs was generally used unless otherwise specified explicitly. We note that the use of such too large an integration time step in CG-MD may produce energy sinks and induce errors.[16] Because of the softer nature of the repulsive component of the Morse potential function (eq 1), we found that use of a time step (Δt) of 40 fs for CG-MD simulations is permissible in this work. Energy fluctuations, $\Delta E = \langle [E - \langle E \rangle]^2 \rangle^{1/2}$, as a function of Δt were performed to evalulate how large Δt could be used in this work.[16] Details of the test are presented in the Supporting Information. A cutoff of 1.6 nm was used for eq 1 without using switching or shifting the function and for the pair list updating. It should be noted that, when the energy lookup table is used, GROMACS 4.0.4 uses the neighbor search cutoff as the real cutoff for the interaction potential. Neighbor searching is usually done with a larger radius than the cutoff specified for the potential to accommodate for the size of charge groups and diffusion between neighbor list updates. Our CG models carry no partial charges. Each CG interaction site was treated as one charge group. Nonbonding interactions involving the first neighbors in *n*-alkane chains were excluded in all CG-MD calculations, and no bond length constraint such as SHAKE was applied. The nonbonding pair list was updated every 5 time steps. Temperature boundary conditions were set using the Nose-Hoover algorithm.[20] A temperature coupling constant of 0.5 ps was used for $\Delta t = 40$ fs. It was reduced to 0.2 ps when

a smaller time step of 10 fs was used. All simulations were performed at 298 K unless otherwise explicitly specified.

Bulk calculations were performed out for both water and liquid alkanes. For water, calculations were performed on systems of two sizes, 400 water molecules and 3200 water molecules. There was no significant size-dependent effect. For liquid alkane calculations, there were 400 molecules in the butane simulation and 200 molecules in the simulations for the larger alkane molecules.

Calculations were also carried out for systems where the liquid coexisted with a vapor. In those calculations, there were 3200 water molecules, 3200 butane molecules, and 1600 alkane molecules, respectively.

3.1. Enthalpy of Vaporization. For MD simulations of bulk phases of CG water and *n*-alkanes, NPT boundary conditions with isotropic pressure coupling were applied. Pressure boundary conditions were set using the Parrinello−Rahman pressure coupling method.[21] The pressure coupling constant was set to 5 ps for $\Delta t = 0.04$ or 2 ps for $\Delta t = 0.01$ ps. For runs to determine molecular volume and heat of vaporization, each system was simulated for a time length of 20 ns. The enthalpy of vaporization ΔH_{vap} of a liquid was calculated from two simulations for its liquid phase and its gas phase, as previously described:[22]

$$\Delta H_{vap} = V_{Intra}(g) - [V_{Inter}(l) + V_{Intra}(l)] + RT = -V_{Inter}(l) + RT \quad (2a)$$

where V_{Intra} is the intramolecular potential energy per mole of molecules calculated for both the gas (g) and the liquid (l) states and V_{Inter} is the intermolecular energy per mole of molecules. The use of eq 2a is based on the assumption of ideal gas behavior and the assumption that the sum of kinetic and vibrational energies is equal for the gas and liquid states. When eq 2a is applied to calculate ΔH_{vap} for CG water, the term $V_{Inter}(l)$, which represents the inter-CG-water energy per mole of CG water, does not include the intermolecular interaction $V_{S,n}$ among the *n subunits* (water molecules) of each CG site. Theoretically, a correction including $V_{S,n}$ to ΔH_{vap} is required in order to compare directly with the experimental ΔH_{vap}. Thus, for an *n*:1 mapped CG water,

$$\Delta H_{vap} = -\frac{V_{Inter}(l) + V_{S,n}}{n} + RT \quad (2b)$$

One may approximate the binding energy of a water cluster formed by *n* molecules of water in the gas phase, which can be calculated quantum mechanically, as $-V_{S,n}$. The ab initio binding energies of the water trimer (61.9 kJ/mol)[23] and tetramer (128.5 kJ/mol)[24] in their lowest energy configurations have been calculated at a high theoretical level. We use these values as the values for the internal energy in eq 2b in analyzing our computational results.

3.2. Free Energy of Solvation. For simulating a liquid in equilibrium with its vapor, a slab (about 7 nm in thickness) of the equilibrated bulk phase of a CG species under consideration was placed in the center of a simulation box with a vacuum slab thickness of about 3.5 nm on each side. The application of periodic boundary conditions creates a system of alternating liquid and vapor layers, with the liquid layer being 7-nm-thick and the vapor layer being 7-nm-thick. The CG-MD simulation was performed in the NVT ensemble at 298 K. During the simulation, a small number of molecules enter the vapor phase (there are no molecules in the vapor phase at the beginning of the simulation), and ultimately a dynamic steady state is reached in which molecules are evaporating and condensing at approximately the same rate. For water and for the shorter alkanes, the simulations were done for 10 μs, with the last 5 μs trajectories being used for data analysis. For long-chain *n*-alkanes (*n*-tetradecane, *n*-pentadecane, etc.), a time length of 20 μs was used in order to obtain meaningful statistics, because of the smaller number of molecules evaporating. The Gibbs free energy of solvation of a gaseous solute in its own liquid can be calculated from the equilibrium densities of the particles in its vapor phase (ρ_v) and its bulk phase (ρ_l) according to Ben-Naim and Marcus,[25,26]

$$\Delta G_S = -RT \ln\left(\frac{\rho_l}{\rho_v}\right) = -RT \ln\left(\frac{c_l RT}{p_v}\right) \quad (3)$$

where c_l is the molar concentration of the liquid and p_v is the vapor pressure. Since vapor−liquid systems were simulated under the NVT condition, the calculated ΔG_S values are in fact Helmholtz free energies ΔF_S from which $\Delta G_S = \Delta F_S + \Delta(PV)$ can be obtained. Note that the $\Delta(PV)$ correction term applies to the condensed phase before and after solvation of the solute molecules according to Ben-Naim and Marcus' definition of the solvation process[25] and is usually negligible.[19]

Thermodynamic integration (TI) procedure[27] was also carried out under NPT conditions to calculate ΔG_S for alkanes. The free energy of solvation was evaluated over a path that mutated all CG sites in a single CG solute molecule into noninteracting ones via a coupling parameter λ in such a way that $\lambda = 0$ describes the noninteracting phantom site and $\lambda = 1$ describes the fully interacting CG site via the following λ-dependent Morse potential function:

$$V_M(r;\lambda) = \lambda\varepsilon[e^{\lambda\alpha(1-\frac{r}{\lambda R_0})} - 2e^{1/2\lambda\alpha(1-\frac{r}{\lambda R_0})}] = \lambda\varepsilon[e^{\alpha(\lambda-\frac{r}{R_0})} - 2e^{1/2\alpha(\lambda-\frac{r}{R_0})}] \quad (4)$$

Thus, eq 4 is equivalent to eq 1 when $\lambda = 1$, and $V_M(r;\lambda)$ is zero when $\lambda = 0$. In the process of gradual mutation from $\lambda = 0$ to $\lambda = 1$, the internal bonded and nonbonded interactions were kept at their full value so that the integrity of the solute molecule was maintained. The process can be viewed as gradual coupling of a single solute molecule in a vacuum and a box of pure solvent. (Note that, in this terminology, the molecule that is being created by "computational alchemy" is the "solute", and the rest of the molecules in the system comprise the "solvent".) In TI, the integral,

$$\Delta G_S = \int_0^1 \left\langle \frac{\partial H}{\partial \lambda} \right\rangle_\lambda d\lambda \quad (5)$$

is used to evaluate ΔG_S, where *H* is the classical Hamiltonian which depends on configuration variables and momenta. Since there is no λ dependence of the internal energy terms of the solute molecule, it does not contribute directly to

$\partial H(p,q;\lambda)/\partial\lambda$. In addition, there is no mass changed in our coupling scheme; there is also no kinetic-energy contribution to $\partial H(p,q;\lambda)/\partial\lambda$. Hence, eq 5 simply becomes

$$\Delta G_S = \int_0^1 \left\langle \frac{\partial V_M(r;\lambda)}{\partial\lambda} \right\rangle_\lambda d\lambda \quad (6)$$

An analytical derivative of eq 4 was used to evaluate $\langle \partial V_M(r;\lambda)/\partial\lambda \rangle$.

Simulations were started at $\lambda = 1$ from a well equilibrated box (ca. $4 \times 4 \times 4$ nm^3) of about 200 to 400 CG alkanes, depending on the size of the alkane. An alkane (solute) molecule fully interacting with the rest of its own kind of molecules (solvent) was randomly chosen from the initial configuration. The λ interval value was set at 0.05. However, in the region of rapidly changing $\langle \partial V_M(r;\lambda)/\partial\lambda \rangle$, it was reduced to 0.025. Subsequent simulation at the next λ interval value was performed using the last configuration from the simulation at a previous λ value. Each simulation consisted of a 5 ns equilibration period followed by a 15 ns production run. An integration time step of 10 fs was used. The $\langle \partial V_M(r;\lambda)/\partial\lambda \rangle$ value was evaluated for every ps. The integration of eq 6 was performed by the trapezoidal rule.

Since the current GROMACS (version 4.0.4) is not capable of performing the TI procedure based on the Morse potential, the above TI method was carried out manually by using the energy-group-lookup-table option available in the MD software. For each λ value, a potential energy lookup table for the interaction between the target solute and the solvent (eq 4) was set up according to the specification of GROMACS 4.0.4. For solvent−solvent and solute−solute nonbonded interactions, a separate interaction lookup table (eq 1) was made. A separate code was written to calculate $\langle \partial V_M(r;\lambda)/\partial\lambda \rangle$ with the configurations extracted from the simulations trajectories.

3.3. Interfacial Tension. The vapor−liquid interfacial tension γ, a measure of the free energy cost associated with the formation of the interface, was computed from the ensemble average normal P_{ZZ} and lateral P_{XX}/P_{YY} pressure components according to

$$\gamma = \frac{1}{2} L_Z \left(\left\langle P_{ZZ} - \frac{P_{XX} + P_{YY}}{2} \right\rangle \right) \quad (7)$$

The factor 1/2 accounts for the two interfaces present in the chosen setup. In general, a μs of CG simulation was necessary to achieve reliable statistics for the computation of γ. The implementation of eq 7 is contained in a standard Gromacs utility, which we used.

3.4. Coarse-Grained Modeling. *Water Tetramer.* The CG water atom (Figure 1), namely W4, is a single bead which represents a group of four water molecules similar to the MARTINI water model.[6] For comparison, Figure 1 shows the corresponding interaction potentials for the MARTINI model and the SSRBK[9] model. (Note for the SSRBK model that the SSRBK CG particle contains only three waters rather than four, so its potential form is not strictly comparable with the other two.) The CG particles interact through $V_M(r)$, eq 1. The target data for determining the parameters of eq 1 for W4 are the experimental liquid water density and the interfacial tension of air−water or water vapor−water. For a series of MD runs with different α values, we optimized ε and R_0 to have calculated ρ and γ in agreement with the experimental values. The final values of α, ε, and R_0 for W4 so obtained are listed in Table 1. We did not use the experimental ΔH_{vap} of water as a target in the course of parametrization for W4, but very good values for ΔH_{vap} were obtained, as mentioned later in the Results and Discussion section. There, we will examine more closely this and other similar CG water models.

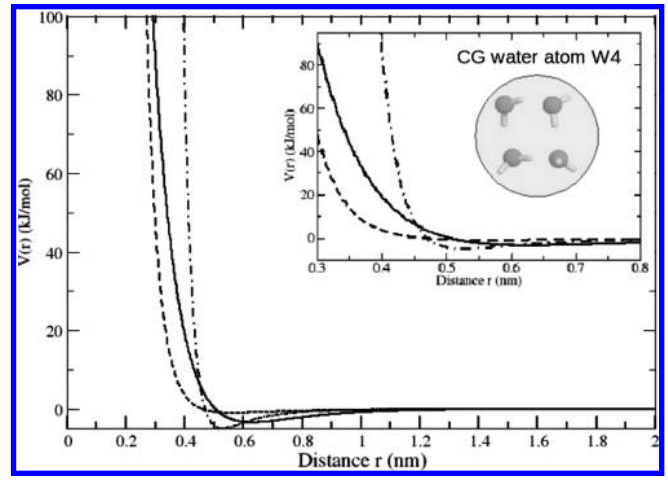

Figure 1. Coarse grained water model and its interaction potential. The CG W4 atom represents a cluster of four water molecules. Solid line, CSJ water (Morse potential); dashed line, SSRBK water[9] (LJ6−4 potential); dot-dashed line, MARTINI water[6] (LJ12−6 potential). It should be noted that the CSJ and MARTINI models are for four-water clusters, while the SSRBK model is for a three-water cluster.

Table 1. Morse Parameters for the CG Water (W4) and the Interaction Sites of 3-Site (C3T, C3M) and 4-Site (C4T, C4M) Mapped Alkanes

interaction site	α	ε (kJ/mol)	R_0 (nm)	atomic mass
W4	7	3.4	0.629	72.062
C3T	12	2.94	0.527	43.089
C3M	12	2.94	0.527	42.081
C4T	12	4.0	0.563	57.116
C4M	12	4.0	0.563	56.108

n-Alkanes. Alkanes with Multiples of 3 or 4 Carbon Atoms. Two CG types of particles for modeling alkanes are employed, 3-site (C3) and 4-site (C4) mapped CG particles. The former is suitable for alkanes with a multiple of three carbon atoms, and the latter is for those with a multiple of four carbon atoms. For example, *n*-hexane is modeled as a linear chain of two C3 CG atoms (Figure 2a), while octane can be represented by a linear chain of two C4 CG atoms, as shown in Figure 2a.

For the 3-site (4-site) CG model, the CG alkanes were constructed from terminal C3T (C4T) and middle C3M (C4M) chain units. These CG units interact through eq 1. In terms of the nonbonding (Morse potential) parameters, the C3T and C4T are not differentiated from C3M and C4M, respectively. However, they are regarded as separate atom types because their masses are slightly different.

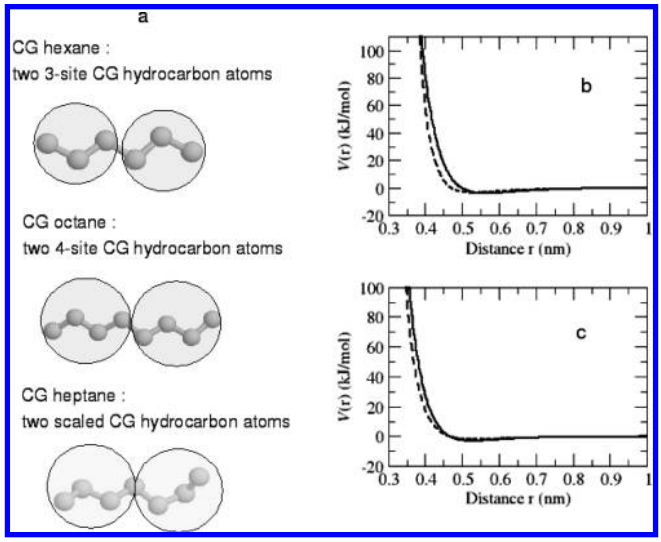

Figure 2. Coarse grained alkane models and their interaction potentials. (a) Hexane is represented by two 3-site CG hydrocarbon atoms. Octane is represented by two 4-site CG hydrocarbon atoms. Heptane is represented by two scaled CG hydrocarbon atoms. (b) Pair interaction potentials for 4:1 mapped CG site. Solid line: Morse potential (this work). Dashed line: LJ 12−6 potential (MARTINI).[6] (c) Pair interaction potentials for 3:1 mapped CG site. Solid line: Morse potential (this work). Dashed line: LJ 9−6 potential (of Nielsen et al.[32]).

Masses for the CG types C3T, C3M, C4T, and C4M are listed in Table 1.

Intramolecularly, the CG alkanes interact through harmonic bond and angular bend potentials,

$$V_{\text{bond}} = \frac{K_{\text{b}}}{2}(r_{\text{b}} - r_{\text{b0}})^2 \tag{8}$$

$$V_{\text{angle}} = \frac{K_{\text{a}}}{2}(\cos\theta - \cos\theta_0)^2 \tag{9}$$

All first neighbors were excluded from nonbonding interaction. The target CG bond lengths (r_{b0}) were calculated from atomistic simulation of dodecane as performed from our previous work.[17] These trajectories were first mapped according to the CG model used. Starting from one end of the alkane chain, n ($n = 3$ for 3-site mapping; $n = 4$ for 4-site mapping) consecutive carbon atoms were counted and their center of mass as the location of the first CG site was calculated. The second CG site was similarly determined by counting further n carbon atoms. Actual masses were used for center of mass calculations. The average distances of two consecutive CG sites were respectively computed from these 3:1 and 4:1 mapped MD trajectories. These are the target bond lengths (r_{b0}) for the CG models and are listed in Table 2. The force field parameters $K_a = 25$ kJ/mol and $\theta_0 = 180°$ (eq 9) were taken from the CG MARTINI force field[6] without further refinement. They are listed in Table 3. It should be noted that the MARTINI n-alkanes and our force field as derived in the main body of this paper do not have torsion potentials. The consequence of not taking torsion potential into account is a uniform dihedral distribution for the CG model, whereas a corresponding atomistic model shows a bias toward lower torsion angles.[29] In further simulations done since the main body of simulations reported in this paper, we find that including torsion angle restraints improves dramatically the agreement of the dihedral angle distributions but has little effect on the thermodynamic properties. Details of the parametrization method and results are presented in the Supporting Information. This extended parameter set includes torsion potential as well as distinction between terminal and nonterminal CG sites in terms of nonbinding interactions. All its nonbonding, bonding, bending, and torsion parameters are parametrized against experimental and atomistic data. In the course of developing this parameter set, we found that setting θ_0 equal to 180° instead of the mean of pseudo-bond angles mapped from atomistic data results in much better correspondence between CG and atomistic bending angle distributions.

Table 2. CG Standard Bond Length r_{b0} and Force Constant K_b Parameters

bond type	r_{b0} (nm)	K_b (kJ mol^{-1} nm^{-2})
C3T, C3M−C3T, C3M	0.36	5000
C4T, C4M−C4T, C4M	0.45	5000

Table 3. CG Standard Bond Angle θ_0 and Force Constant K_a Parameters for 3-Site and 4-Site CG Models

angle type	θ_0 (degree)	K_a (kJ mol^{-1})
C3T, C3M−C3M−C3T, C3M	180	25
C4T, C4M−C4M−C4T, C4M	180	25

CG simulations of liquids hexane and octane were initially performed using the MARTINI bond force constant K_b for alkanes (1250 kJ mol^{-1} nm^{-2}). For different values of α, the parameters ε and R_0 were adjusted so that the simulated ρ and ΔH_{vap} of hexane and octane were in agreement with the experimental values.[30] Their vapor−liquid interfacial tensions were also computed with these initially parametrized sets of $\{\alpha,\varepsilon,R_0\}$. We found that $\{\alpha,\varepsilon,R_0\}$ with $\alpha = 12$ yielded γ values close to the experimental data.[31] In addition, K_b was tuned by matching the radial distribution functions (RDF) for CG hexane and octane to those obtained from mapped atomically detailed MD data. The K_b value so determined, 5000 kJ mol^{-1} nm^{-2}, is good for both 3-site and 4-site CG alkanes. We now have $\alpha = 12$ and $K_b = 5000$ kJ mol^{-1} nm^{-2} set for all CG alkanes.

With the parametrized values of $\alpha = 12$ and $K_b = 5000$ kJ mol^{-1} nm^{-2} and other bonding parameters as listed in Tables 2 and 3, the ε and R_0 were refined by carrying out CG simulations of pentadecane and hexadecane. The target data were their experimental ρ and ΔH_{vap}. The parameters were finally modified slightly so that the CG simulated ρ and ΔH_{vap} were in good agreement with the experiment for both the short (hexane, octane) and long (pentadecane, hexadecane) chain alkanes. Table 1 lists the final values of Morse parameters (α,ε,R_0) for both 3-site and 4-site mapped CG alkanes. Figure 2b shows the nonbonded interaction potentials for the 4-carbon alkane CG interactions in the CSJ

and MARTINI force fields. Figure 2c shows the nonbonded interaction potentials for the 3-carbon alkane CG interactions in the CSJ (this paper) and the NLSK[32] force fields.

Scaling for Alkanes. For chain length with non-multiple-of-three or multiple-of-four carbon atoms, a combination of 3-site and 4-site CG atoms can be introduced. For example, *n*-heptane can be represented by one C3T and one C4T atom. However, this option leads to a nonuniform representation of the chain. In what follows, we introduce a scaling method for the Morse parameters to avoid such unwanted choice.

The number of carbon atoms in an alkane chain, N_C, is divided by three, and the quotient, an integer, gives the number of CG sites, N_{CG}. Each site is then scaled in mass and in center of mass placement by $W_S = N_C/N_{CG}$. In other words, the *extra* carbon atoms (N_C modulo 3) are spread uniformly over the CG chain; i.e., they are evenly shared among the N_{CG} sites. To map the atomistic alkane chain onto a scaled CG representation, W_S carbons are counted starting from one end of the atomistic chain. The center of mass of these atoms determines the location of the first CG site. A reduced mass is given to fractional atoms in the center of mass calculation. The second CG site is then assigned by counting the next W_S carbons and determining their center of mass. If the last counting ends at a fractional site of the atomistic chain, its remainder is used for the current counting. The CG representation of *n*-undecane, for example, has (quotient of 11/3) 3 CG sites. The number of extra carbon atoms is (N_C modulo 3) 2, which are equally shared among the three CG sites. Each CG site has scaling weight of $W_S = N_C/N_{CG} = 11/3$. The first CG site is the center of mass of the first three carbon atoms and 2/3 of the fourth one. The center of mass of the remainder (1/3) of the fourth carbon, its following three carbon atoms, and 1/3 of the eighth carbon atom is then the location of the second CG site. To assign the location of the third CG site, the center of mass of the remainder (2/3) of the eighth carbon atom and the last three carbon atoms is computed.

To obtain the target CG bond lengths (r_{b0}) for *n*-heptane, *n*-decane, *n*-undecane, *n*-tridecane, and *n*-heptadecane, the aforementioned scaling method was used to map the atomically detailed trajectories of these *n*-alkanes from our previous work[17] to their CG representations according to their scaling weights. The time average of the distances between two consecutive CG sites was computed from the atomically detailed simulations for each scaled CG representation, and the result was assigned as the r_{b0} of those CG alkanes whose CG sites have the same scaling weight. Thus, CG *n*-heptane and *n*-tetradecane have the same r_{b0} value. The same bond and bond angle force constants as used in the 3-site/4-site models were applied, and θ_0 was set to 180°. All these parametric values are listed in Table 4.

To calculate R_0 for the scaled CG sites, we assumed that the increase in R_0 from 3-site CG ($W_S = 3$) to 4-site CG ($W_S = 4$) is linear with the increased scaling weight. Hence, the R_0 for a CG site with scaling weight W_s can be simply calculated from

$$R_0(s) = 0.527 + \Delta R_0(W_s - 3) \tag{10}$$

Table 4. Scaled CG Standard Bond Length r_{b0}, Bond Angle θ_0, and Their Force Constants K_b and K_a

alkane	W_S	r_{b0} (nm)	K_b (kJ mol^{-1} nm^{-2})	θ_0 (deg)	K_a (kJ mol^{-1})
heptane, tetradecane	$3^1/_2$	0.400	5000	180	25
decane	$3^1/_3$	0.390	5000	180	25
undecane	$3^2/_3$	0.413	5000	180	25
tridecane	$3^1/_4$	0.371	5000	180	25
heptadecane	$3^2/_5$	0.375	5000	180	25

where the first term is the R_0 value for the 3-site CG model (Table 1), ΔR_0 (0.036 nm) is the difference of the R_0 values between the 4-site and 3-site CG atoms, and the numerical value 3 in the last term is the scaling weight of the 3-site CG atom.

To obtain the ε values of the scaled CG sites, we employed the following combination rule: We compute the geometrical mean value of the parameters ε of two CG sites with known W_S values of w_1 and w_2, respectively. The result is assigned as the ε value for a CG site with a W_S value which is the arithmetic mean of w_1 and w_2:

$$\varepsilon\left(\frac{w_1 + w_2}{2}\right) = (\varepsilon(w_1)\,\varepsilon(w_2))^{1/2} \tag{11}$$

For CG *n*-heptane and *n*-tetradecane with a scaling weight of $3^1/_2$, their $\varepsilon(3^1/_2)$ value was calculated from the already known $\varepsilon(3)$ and $\varepsilon(4)$ values of the 3-site and 4-site CG atoms, respectively (Table 1), according to eq 11. The scaling weight for CG *n*-tridecane is $3^1/_4$. Hence, its $\varepsilon(3^1/_4)$ value was calculated from $\varepsilon(3)$ and $\varepsilon(3\ 1/2)$ using eq 11. The scaling weight for *n*-decane is $3^1/_3$. Its $\varepsilon(3^1/_3)$ value cannot be evaluated in the same way as previously mentioned. Instead, it was approximated by successively computing the value of $\{\varepsilon(w_i)\}$ according to eq 12

$$\varepsilon(w_{i+2}) = (\varepsilon(w_i)\,\varepsilon(w_{i+1}))^{1/2}, w_{i+2} = \frac{w_{i+1} + w_i}{2} \tag{12}$$

For $w_1 = 3\ 1/2$ and $w_2 = 3\ 1/4$, the fisrt two terms of the $\{\varepsilon(w_i)\}$ sequence are the already known $\varepsilon(w_1) = \varepsilon(3\ 1/2)$ and $\varepsilon(w_2) = \varepsilon(3^1/_4)$ values as computed for *n*-heptane and *n*-tridecane, respectively. The third term, $\varepsilon(w_3) = \varepsilon(3^3/_8)$, was then calculated using eq 12. The fourth term $\varepsilon(w_4) = \varepsilon(3^5/_{16})$ was then in turn computed with the known values of $\varepsilon(w_2)$ and $\varepsilon(w_3)$. This procedure was repeated until the term $\varepsilon(w_i)$ with $w_i \approx 3^1/_3$ was obtained. The scaled $\varepsilon(3^2/_3)$value for CG *n*-undecane was computed in the same way using eq 12 with the known $\varepsilon(w_1) = \varepsilon(4)$ and $\varepsilon(w_2) = \varepsilon(3^1/_2)$ values as the first and second terms of $\{\varepsilon(w_i)\}$, respectively. The iterative computation was proceeded until the term $\varepsilon(w_i)$ with $w_i \approx 3^2/_3$ was obtained.

In scaling $\varepsilon(3^2/_5)$ for CG *n*-heptadecane, we first calculated the series (eq 12) through the term $\varepsilon(3^3/_{16})$ starting with the known values of $\varepsilon(w_1) = \varepsilon(3)$ and $\varepsilon(w_2) = \varepsilon(3^1/_4)$. The $\varepsilon(3^{13}/_{32})$ value was then evaluated from $\varepsilon(3^3/_{16})$ and $\varepsilon(3^5/_8)$ according to eq 11. The latter had already been obtained in the course of calculating $\varepsilon(3^2/_3)$. Since $3^{13}/_{32}$ is approximately equal to $3^2/_5$, we hence assigned the so calculated $\varepsilon(3^{13}/_{32})$ as the value of $\varepsilon(3^2/_5)$ for CG *n*-heptadecane.

Table 5. Morse Parameters for Scaled CG Sites[a]

alkane	CG type	W_S	ε (kJ/mol)	R_0 (nm)	α
heptane, tetradecane	CST, CSM	$3^{1/2}$	3.43	0.545	12
decane	CST, CSM	$3^{1}/_{3}$	3.26	0.539	12
undecane	CST, CSM	$3^{2}/_{3}$	3.61	0.551	12
tridecane	CST, CSM	$3^{1}/_{4}$	3.175	0.536	12
heptadecane	CST, CSM	$3^{2}/_{5}$	3.33	0.541	12

[a] Terminal sites are denoted as CST and non terminal sites are designated as CSM. Actual masses for CG sites are used according to their scaling weight W_S.

Table 6. Calculated Density ρ (g/cm^3), Self Diffusion Coefficient D (10^{-9} m^2/s), Heat of Vaporization ΔH_{vap} (kJ/mol), Free Energy of Solvation ΔG_S (kJ/mol), Surface Tension (water-vapor) γ (mN/m), and Isothermal Compressibility κ (10^{-5} bar^{-1}) of CG Water W4 at 298 K for CG Water Models[a]

Water model	ρ	D	ΔH_{vap}	ΔG_S	γ	κ
CSJ[b] W4	0.998	4.3	38.4	−28	71	17
CSJ[b] W4	*0.996*[f]	*4.7*[f]	*38.3*	*−28*[f]	*71*[f]	*26*[f]
MARTINI[c] W	1.005	1.6	30.2	−18	32	9
SSRBK[d] W	0.993	6.6	32.6	−19	71	15
Experiment[e]	0.998	2.3	44.0	−26.5	73	4.5

[a] CG-MD integration time step used was 10 fs. [b] This work. Equation 1 is the pair interaction potential. [c] Ref 6, CG-MD simulations were performed in this work. [d] SSRBK is the acronym of the last names of the authors of ref 9. CG-MD simulations were performed in this work. [e] Refs 30 and 31. The experimental ΔG_S was calculated from the vapor pressure p_v (ref 31) using eq 3. [f] Italic numbers were calculated from simulations performed using an integration time step of 40 fs.

All the scaled Morse parameters obtained in this section are listed in Table 5. They were applied for CG simulations without further parametrization.

4. Results and Discussion

4.1. CG Water. Table 6 lists the calculated density ρ, heat of vaporization ΔH_{vap}, free energy of solvation ΔG_S, and liquid−vapor interfacial tension γ of CSJ water. There, the listed numerical data for MARTINI[6] and SSRBK[9] CG water were also calculated for comparison. The simulation conditions used for these comparisons essentially replicated those used by the workers who developed those force fields except that in order to do the comparison we used a time step of 10 fs, a pairlist update frequency of 5 time steps, the Nose-Hoover temperature coupling method, the and Parrinello−Rahman algorithm for pressure coupling. Both CSJ (this work) and MARTINI[6] CG water models employ 4:1 mapping while the SSRBK[9] model has a single bead to represent three water molecules. The Martini water interacts via a LJ 12−6 potential. The pair potential for the SSRBK water is a softer LJ 6−4. As can be seen from Figure 1, the LJ 12−6 potential for Martini water is much harder than the other two in the short-range region. Using a large integration time step such as 40 fs with a cutoff of 1.2 nm as is done in the MARTINI water model induces errors[16] and produces energy sinks which cause a freezing effect. We found that using a smaller time step of 10 fs with the MARTINI force field does not remove the freezing artifact but delays its occurrence. We hypothesize

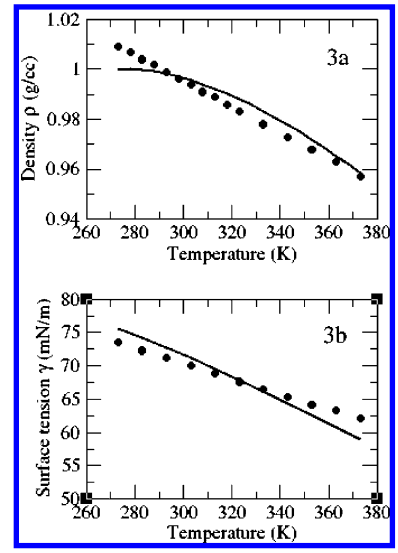

Figure 3. Density ρ and surface tension γ of CG water. (a) Temperature dependent density of CG water: CSJ water (black circle), experiment (solid line). (b) Temperature dependent surface tension of water: CSJ water (black circle), experiment (solid line).

that the fundamental problem lies in the use of the LJ 12−6 potential, which produces a narrow deep well as shown in Figure 1. The SSRBK water model (LJ 6−4) may alleviate this problem but provides only two free parameters, which reduces flexibility in tuning the potential to fit a variety of experimental data. There exists an advantage of using the Morse potential for targeting atomistic or experimental data, in that one can control more freely the interactions for $r \leq R_0$ and $r > R_0$ independently by using different α values for $r \leq R_0$ and $r > R_0$ whenever it is desired. Another 3-parameter candidate for pair interaction is the Buckingham potential[33] (BP) $V_{BP}(r) = b \exp(-r/\rho) - \mu/r^6$ where b, ρ, and μ are constants and r is the interatomic distance. There are several studies in which a Buckingham form (near repulsion is an exponential form) has been used in pairwise interaction potentials.[34−39] A concern in using such a form is that at very short distances, the attractive sixth power term overwhelms the repulsive exponential term and the potential becomes very large and negative, which is nonphysical. It would preclude the use of the Buckingham potential in the early stages of building macromolecular complexes, in which the systematic and automated relief of steric clashes is needed. The Morse potential (eq 1) does not diverge at very small distances and has a finite positive value at $r_{ij} = 0$. We do not find this a problem in our simulations, because with the chosen values of α, the repulsive potential is still large enough to prevent the coalescence of the interacting pair.

Under normal physiological temperatures, both the CSJ and the SSRBK water models are designed to have their intensive properties ρ and γ be consistent with those of water. Both CSJ and SSRBK water models have calculated γ values (71 nN/m at 298K) in good agreement with the measured value, 73 mN/m.[30] Figure 3 shows the temperature dependence of ρ and γ for CSJ water. Both ρ and γ of CSJ water

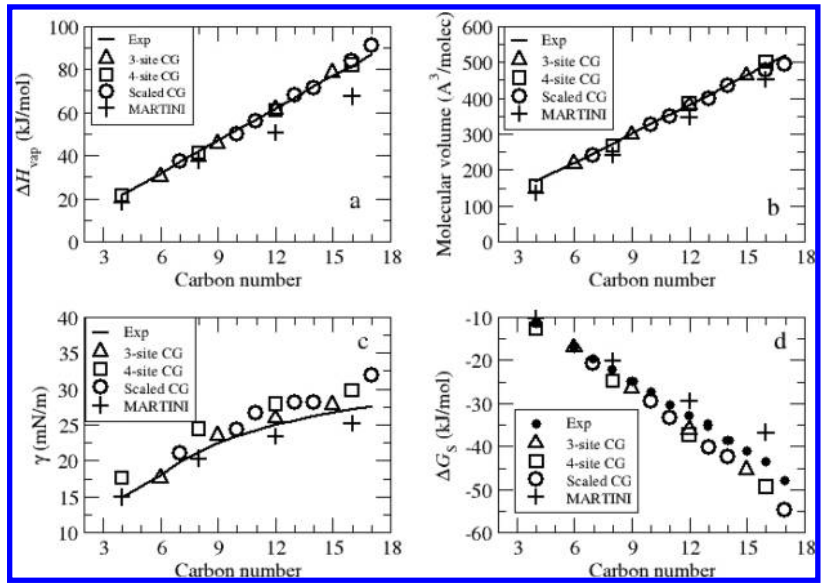

Figure 4. Physical and thermodynamics properties of alkanes at 298 K. (a) Enthalpies of vaporization ΔH_{vap}. (b) Molecular volumes of alkanes. (c) Vapor−liquid interfacial tension γ of alkanes. Data for butane were obtained at 273 K. (d) Gibbs solvation free energies ΔG_S of alkanes.

Table 7. Solvation Free Energy ΔG_S of *n*-Alkanes at 298 K

	ΔG_S (kJ/mol)		
n-alkane	TI (eq 5)	vapor−liquid (eq 3)	experiment[a]
butane	−12.8	−12.6	−11.5
hexane	−17.1	−17.4	−16.9
heptane	−20.8	−21.1	−19.7
octane	−24.9	−25.4	−22.3
nonane	−26.6	−27.3	−24.9
decane	−29.5	−30.1	−27.5
undecane	−33.5	−34.2	−30.5
dodecane	−36.2(−37.3)[b]	−37.2(−36.1)[b]	−33.0
tridecane	−40.3	−42.0	−35.0
tetradecane	−42.3	−43.0	−38.8
pentadecane	−45.5	−43.0	−41.1
hexadecane	−49.4	−50.9	−43.7
heptadecane	−54.9	−52.4	−48.1

[a] Ref 31. The experimental ΔG_S was calculated from the vapor pressure p_v using eq 3. [b] Numbers in parentheses are for the 4-site CG model.

decrease linearly as T increases. The CG water model does not however capture the nonlinear temperature dependence of the experimental data in the temperature range 273−373 K. As can be seen in Figure 3, although the general trends and overall ranges of the properties are similar to the experiment, the forms of the calculated density and surface tension are much more linear than the experimental curves. This is likely a consequence of the loss of some atomistic details that occurs in the coarse graining process. This shows limits of the accuracy which can be expected for coarse-grained calculations as soon as quantities involving derivatives are needed.

The Gibbs excess free energy ΔG_{ex} of water is equivalent to the negative of the solvation free energy of water ΔG_S in its own liquid. The simulated ΔG_S for CSJ water is −28 kJ/mol. All other water models studied in this work give ΔG_S values from −18 to −19 kJ/mol (Table 6). The calculated ΔG_S should correspond to the ΔG_{ex} of a water tetramer (for CSJ and MARTINI W) or trimer (for SSRBK W). Direct comparison of calculated ΔG_S for a CG water representing a cluster of n water molecules to the experimental ΔG_S value (−26.5 kJ/mol calculated from the vapor pressure p_v[31] using eq 3) for water is not strictly appropriate. Baron et al.[40] found that the calculated ΔF_{ex} values for the insertion of an SPC water tetramer into SPC water and into liquid SPC tetramers are, respectively, 22 and 19 kJ/mol. These ΔF_{ex} values compare fairly with the ΔG_{ex} values of the CG water models (CSJ, MARTINI, and SSRBK) suggesting solvation of a CG water bead into its own liquid is reasonably well represented by CG water models.

When a CG water represents more than one water molecule, direct comparison of calculated ΔH_{vap} to the experiment requires a theoretical correction. Presented in Table 6 are ΔH_{vap} values directly calculated from applying eq 2a to MD simulations for the CG water models. Enthalpy of vaporization for water can then be estimated from the simulated ΔH_{vap} for CG water and the binding energy correction term of eq 2b. This has been explained in the Computational Method section. The so calculated ΔH_{vap} values for water are 43.6, 41.5, and 33.2 respectively by using the CG ΔH_{vap} values (Table 6) for CSJ W4, MARTINI W, and SSRBK W. The agreement of these calculated ΔH_{vap} values for water by using eq 2b with the experimental value, 44 kJ/mol,[30] indicates that the CG ΔH_{vap} values (Table 6) are reasonably respresented by the CG water models.

The diffusion coefficients D, calculated from the slopes of the mean square displacements (MSD) in the long time limit using the Einstein relation $\langle \Delta r(t)^2 \rangle = 6Dt$, for the various CG water models (Table 6), range from 1.6×10^{-9} to 6.6×10^{-9} m^2/s. These calculations are all newly done by us for this paper on the various models, so that the computational conditions would be the same (except the cutoff as designed for each water model.) These CG

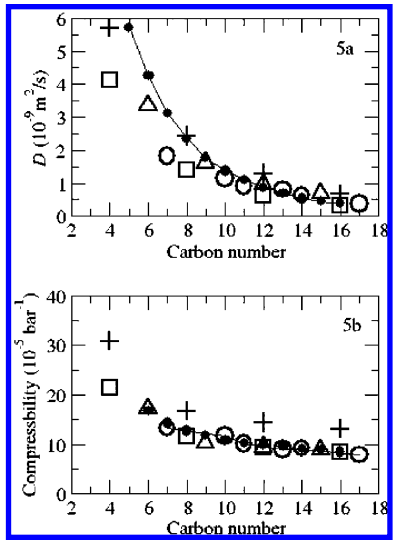

Figure 5. Self-diffusion coefficients D and isothermal compressibilities κ of alkanes at 298 K. (a) Diffusion coefficients were calculated from the slope of the mean square displacement in the long time limit. (b) Isothermal compressibilities were calculated from the relation $\kappa = (\sigma_V^2)/(kT\langle V\rangle)$, where σ_V is the volume (V) fluctuations in an NPT ensemble. Symbols for a and b: triangle, 3-site CG models; square, 4-site CG; circle, scaled CG models; filled circle, experiment.[47]

results are all on the same order of magnitude as the experimental D value (2.3×10^{-9} m^2/s)[41] for water at 298 K. Seeking precise correspondence with the experiment in this case is perhaps not meaningful, since the entity for which the experimental determination is made (single water molecules) does not exist in the CG models. The calculated diffusion coefficients, as listed in Table 6, for CSJ, MARTINI, and SSRBK CG waters were not renormalized by a factor of 4 (for CSJ and MARTINI water models) or 3 (for SSRBK), in contrast to the renormalizations done by Marrink et al.[28] as well as Groot and Rabone.[42] For any single-site CG water model which represents a group of more than one water molecule and is modeled to have the same density as liquid water, the fictitious n water molecules it represents are explicitly and implicitly bound via hydrogen bonding within the volume of the CG particle. In other words, the fictituous water molecules within the CG water model go wherever it goes. Jalabert and Das Sarma[43] have shown the diffusion coefficient of bound particles is the same as that of the center of mass of the particles.

Isothermal compressibilities κ for the water models were calculated from the equation[44] $\kappa = \sigma_V^2/kT\langle V\rangle$, where σ_V is the root-mean-square volume (V) fluctuation in an NPT ensemble. The κ values for CSJ, SSRBK, and MARTINI water (Table 6) are about 4, 3, and 2 times larger than the experimental value[30] (4.5×10^{-5} bar^{-1}) of water, respectively. We note that Marrink et al.[28] reported a smaller value of 6×10^{-5} bar^{-1} for the MARTINI water. Since this trend in the calculated κ relates to the trend in the hardness of the applied interaction potential in the short - range region, we conclude that the softer the potential in the short-range region, the larger the σ_V which in turn results in a larger κ value.

In summary, from the standpoint of coarse graining, it is appropriate to compare CG water properties to experiments for intensive properties (density, surface tension) and thermodynamic properties for equilibrium processes such as free energies of partitioning between two phases. Moreover, direct comparison of CG water results to experiment is also possible for a change-of-state thermodynamic property such as ΔH_{vap} provided a correction is made corresponding to the internal energy of the CG particle.

4.2. Alkanes. *Thermodynamic Properties.* For 3:1 and 4:1 mapped CG alkanes, experimental surface tension was used as a guide to properly choose the value of the Morse α parameter. The nonbonded CG parameters ε and R_0 were fitted against experimental ΔH_{vap} and bulk density. The scaled CG parameters were calculated as described in the Computational Methods section. The simulated

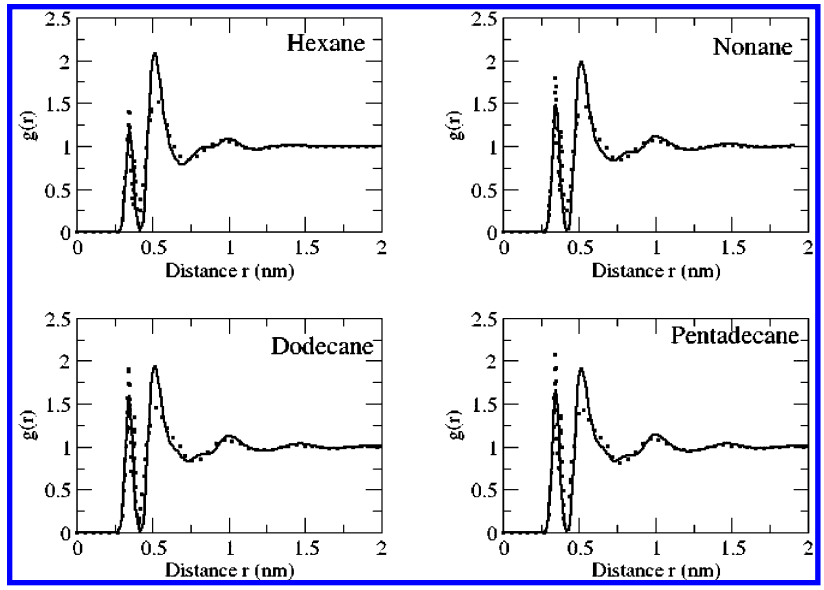

Figure 6. Comparison of radial distribution functions (RDF) for 3-site CG alkanes and corresponding mapped atomistic data. Solid line: CG RDF. Dotted line: mapped atomistic RDF.

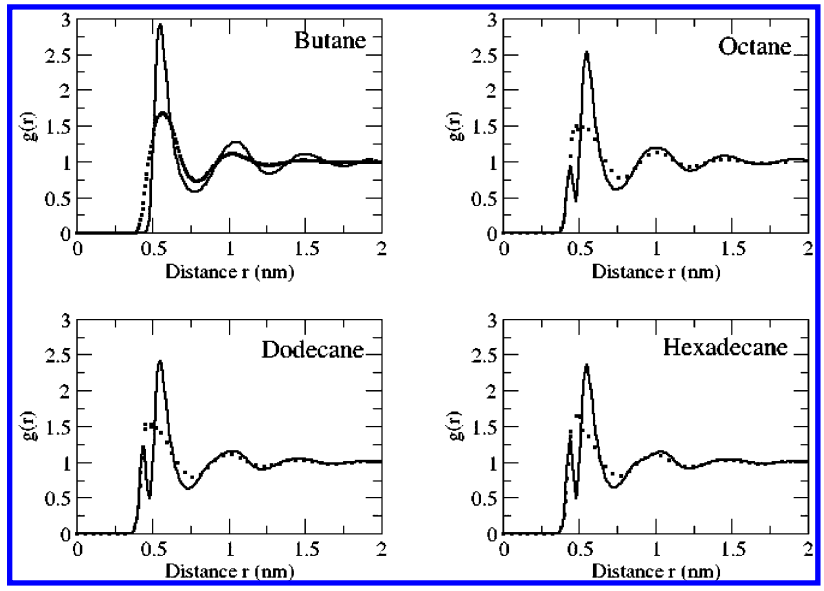

Figure 7. Comparison of radial distribution functions (RDF) for 4-site CG alkanes and corresponding mapped atomistic data. Solid line: CG RDF. Dotted line: mapped atomistic RDF.

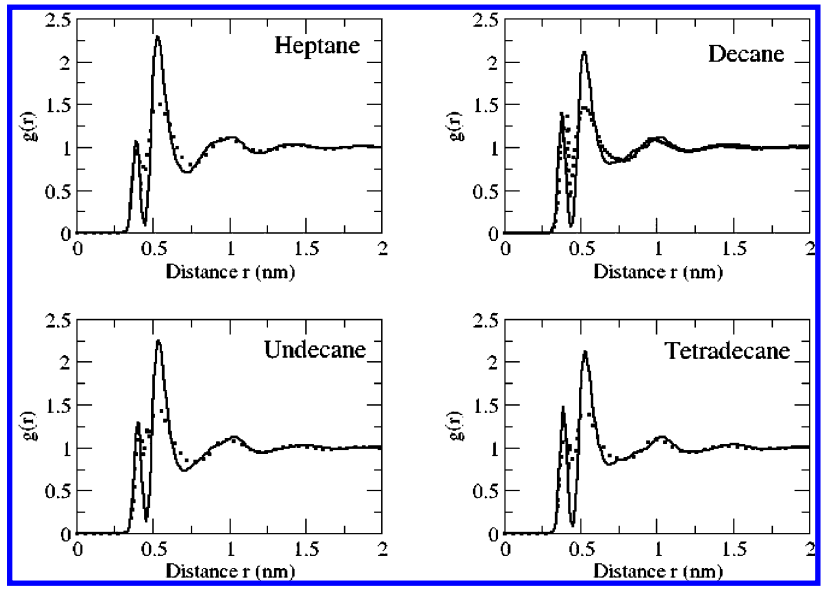

Figure 8. Comparison of radial distribution functions (RDF) for scaled CG alkanes and corresponding mapped atomistic data. Solid line: CG RDF. Dotted line: mapped atomistic RDF.

results are shown in Figure 4. Simulated ΔH_{vap} (Figure 4a) and specific volumes (Figure 4b) of CG alkanes are in good agreement with experimental data.[30] The absolute mean deviation of calculated γ from experimental values is 1.9 mN/m (Figure 4c). The ΔG_S values, as listed in Table 7, calculated by using the TI method (eq 5) agree well with those obtained from simulations of alkane vapor−liquid interfaces (eq 3). Shown in Figure 4d are the calculated ΔG_S for CG alkanes, which were not directly targeted in the parametrization process. They are systematically more negative as compared to the values based on log p measurements[31] (eq 3) for long-chain alkanes. The absolute mean deviation of CG ΔG_S from experimental values is ca. 4 kJ/mol. Although the MARTINI CG force field is less satisfactory in predicting ΔH_{vap} and specific volumes of alkanes, it delivers results of ΔG_S and γ for alkanes in good agreement with experiments.

The results of the pulsed-gradient spin−echo NMR method[45] of measuring self-diffusion establish that the diffusion coefficients D of liquid alkanes decrease nonlinearly as chain length increases,[46] as shown in Figure 5a. It can be seen from the same figure that our CG parameters also reproduce the same trend on this scale. We note that reported D values for MARTINI hexadecane by Baron et al.[40] and Winger et al.[16] are very different. Placing the actual mass on each CG site instead of the uniform value of 72 which is used in the MARTINI force field, we calculated the D values for MARTINI alkanes. The results, in good agreement with the experiment, are also shown in Figure 5a for comparison. Simulated κ values over the whole range of alkanes studied

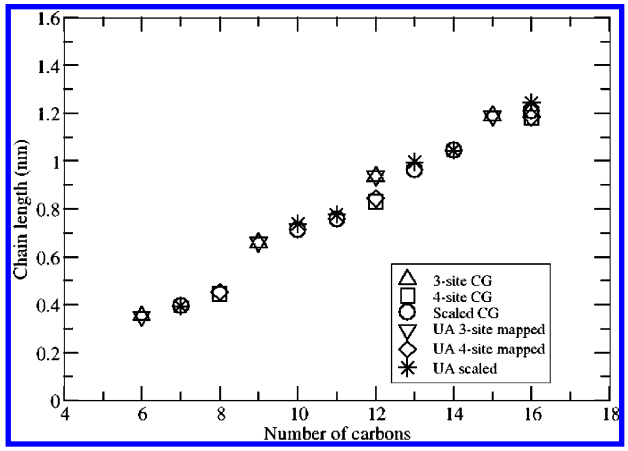

Figure 9. Comparison of mean chain lengths of CG alkanes to atomistic alkanes. Atomistic coordinates for CG mapping were taken from ref 17.

in this work agree well with the experimental results of Díaz Peña and Tardajos[47] as can be seen from Figure 5b.

In order to investigate the effects of taking torsion potentials into account on calculated energetics, CG-MD simulations for bulk CG dodecane, pentadecane (3-site models), and hexadecane (4-site model) were also performed by using the same standard parameter set (from Tables 1−3) with the torsion potential $V_d(\phi) = K_d(1 + \cos(\delta)\cos(m\phi))$ taken into account, where K_d is the dihedral force constant, δ is π, and $m = 1$.The K_d values for the 3-site and 4-site models are 0.75 and 1.2 kJ/mol, respectively. They were parametrized by matching the mean and standard deviation of torsion angles from CG-MD results with those from atomistic data mapped onto the corresponding CG models. We report in the Supporting Information that the calculated physical and thermodynamic properties for CG dodecane, pentadecane, and hexadecane are hardly influenced by the applied torsion potentials.

Structural Properties. The radial distributions (RDFs) for the CG alkanes with comparisons to corresponding mapped atomistic RDFs are shown in Figures 6−8. To compare the CG and atomistic data, the latter were analyzed using the mapping procedure as described in the Computational Method section. In these figures, the locations of the first peaks (except that for butane) correspond to the *ideal* bond lengths of the corresponding CG alkanes, which were obtained from the atomistic data, and the second peaks correspond to the inter-CG site distances. As can be seen in these figures, the structures of CG alkanes reproduce RDFs in good agreement with the atomistic data, to the extent permitted by the coarse graining.

The mean chain length (end-to-end distance) as a function of alkane carbons is plotted in Figure 9. The atomistic data, analyzed using the CG mapping procedure as described previously, are in excellent agreement with the CG data. The chain length distributions for the CG and atomistic data (Figure 10), as exemplified by the 3-site mapped *n*-alkanes (hexane, nonane, doedecane, and pentadecane), also agree well with each other. For a closer look at the agreement, we compared the CG bond angle distributions of the same series of CG alkanes (Figure 11) with those mapped from their corresponding atomistic data.

The consequence of not taking torsion potential into account is a uniform dihedral distribution for the CG model, while the corresponding atomistic model shows a bias toward lower torsion angles.[29] Including torsion potential in CG alkane with four or more consecutive interaction sites does improve dramatically the agreement of dihedral distributions between the CG and the atomistic models. The detailed results are shown in the online Supporting Information. Shown also in the Supporting Information is the comparison of distributions of bending angles with and without torsion potentials applied. It is seen there that taking torsion potential into account has no apparent effect in bending angle distributions.

In summary, inclusion of torsion potentials to the current CG alkane models shows no changed performance of the parameter set except for the excellent correspondence that

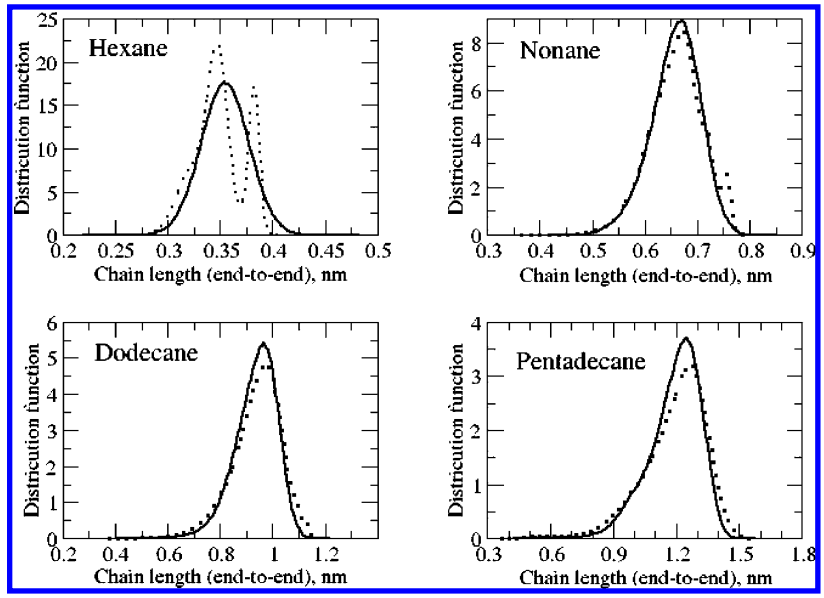

Figure 10. Comparison of chain length distributions of CG alkanes to atomistic alkanes. Solid line: CG results. Dotted line: mapped atomistic results.

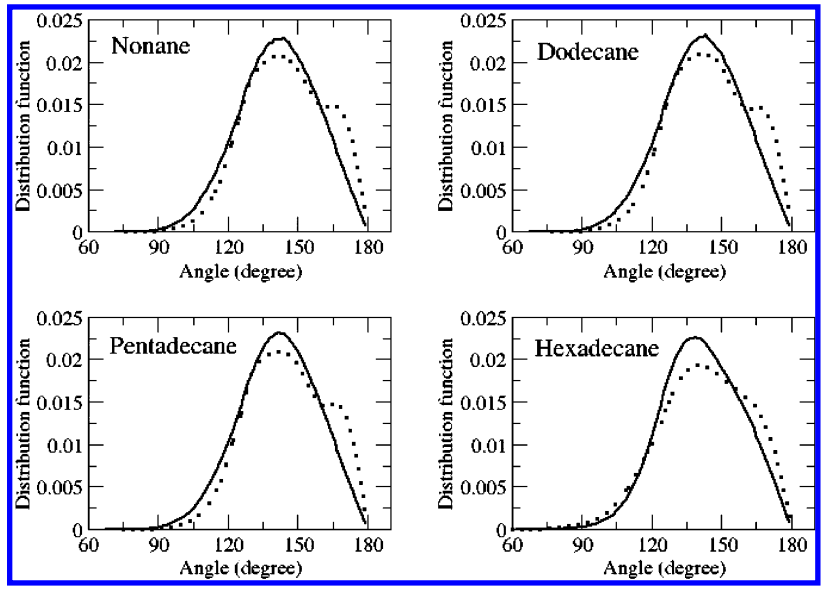

Figure 11. Comparison of bond angle distribution of CG alkanes to atomistic alkanes. Solid line: CG results. Dotted line: mapped atomistic results.

is found between CG and atomistic pseudo dihedral distributions. We note that the addition of a torsion angles makes it necessary to run simulations using a shorter Δt, so that there is a trade-off between the improved accuracy provided by this inclusion, and the capability of running longer and larger simulations without them.

4.3. Integration Time Step Analysis. Details and results for the test to evaluate how large Δt could be used in this work are presented in the Supporting Information. A criterion for a fairly accurate integration of the equations of motion is that ΔE_{tot} should be less than one-fifth of ΔE_{kin} or ΔE_{pot} in the NVE ensemble.[48−50] For both SSRBK and CSJ water models, which have softer repulsive interactions in the short-range region (Figure 1), and for CSJ hexadecane, using Δt of 40 fs is the limit to fulfill the criterion.

5. Conclusion

Coarse grained models for water and alkanes using the Morse potential (eq 1) for pairwise interactions were parametrized using the experimental ΔH_{vap}, bulk density, and surface tension to set the three adjustable parameters in the Morse form. In addition to the three experimental values mentioned above, we find that the CG models also have reasonable values for other properties. The structural properties of CG alkanes are in good agreement with those mapped from atomistic data. The application of an interaction potential softer in the short-range region allows the use of a larger integration time step for CG-MD simulations. We find that a time step of 40 fs can be safely used for the version of the force field in which torsion angle restraints are not included on the alkane chains (as they are not included in the MARTINI force field). However, if greater fidelity is desired for chain structures, a 10 fs time step is required.

Since the Morse potential can be successfully adapted for both highly polar (water) and very nonpolar (alkanes) species, we suggest it as a candidate for the general form of the nonbonded interaction in coarse-grained simulations. An additional benefit is that, due to the softer repulsive force, the Morse potential permits larger time steps than does the L-J form. It should be said the Morse potential is slightly more time-consuming (approximately 20% in our hands) per time step in our implementation (which uses table look up rather than evaluation of exponentials) than the Lennard-Jones form. However, this is more than made up by the larger time step permitted, so that on balance the Morse form is both more efficient and also more faithful in replicating a broad range of experimental data. It may also be that the Morse form calculation can be made more computationally efficient in the future.

Acknowledgment. We gratefully acknowledge support from Grant 5PN2EY016570−06 from the NIH Nanomedicine Roadmap program.

Supporting Information Available: Additional results, discussion, and methodologies inclusing tables and figures. This information is available free of charge via the Internet at http://pubs.acs.org/.

References

(1) Müller-Plathe, F. *ChemPhysChem* **2002**, *3*, 754.

(2) Reith, D.; Pütz, M.; Müller-Plathe, F. *J. Comput. Chem.* **2003**, *24*, 1624.

(3) Spyriouni, T.; Tzoumanekas, C.; Theodorou, D.; Müller-Plathe, F.; Milano, G. *Macromolecules* **2007**, *40*, 3876.

(4) Wilson, M. R.; Stimson, L. M.; Ilnytskyi, J. M.; Hughes, Z. E. Computer Simulations Of Liquid Crystal Polymers And Dendrimers. In *Computer Simulations of Liquid Crystals and Polymers*, Pasini, P., Zannoni, C., Zumer, S., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2005; Vol. 5, pp 7−81.

(5) Prampolini, G. *J. Chem. Theory Comput* **2006**, *2*, 556.

(6) Marrink, Siewert, J.; Risselada, H. J.; Yefimov, S.; Tieleman, D. P.; de Vries, A. H. *J. Phys. Chem. B* **2007**, *111*, 7812.

(7) Monticelli, L.; Kandasamy, Senthil, K.; Periole, X.; Larson, Ronald, G.; Tieleman, D. P.; Marrink, S.-J. *J. Chem. Theory Comput.* **2008**, *4*, 819.

(8) Loison, C.; Mareschal, M. F.; Schmid, F. *J. Chem. Phys.* **2004**, *121*, 1890.

(9) Shelly, J. C.; Shelly, M. Y.; Reeder, R. C.; Bandyopadhyay, S.; Klein, M. L. *J. Phys. Chem B* **2001**, *105*, 4464.

(10) Markvoort, A. J.; Pieterse, K.; Steijaert, M. N. P.; Spijker, P.; Hilbers, P. A. J. *J. Phys. Chem. B* **2005**, *109*, 22649.

(11) Ha-Duong, T.; Basdevant, N.; Borgis, D. *Chem. Phys. Lett.* **2009**, *468*, 79.

(12) Molinero, V.; Goddard, W. A., III. *J. Phys. Chem. B* **2004**, *108*, 1414.

(13) Morse, P. M. *Phys. Rev.* **1929**, *34*, 57.

(14) Liew, C. C.; Masuhiro, M. *Chem. Phys. Lett.* **2003**, *368*, 346.

(15) Lennard-Jones, J. E. *Proc. R. Soc. London, Ser. A* **1925**, *109*, 584.

(16) Winger, M.; Trzesniak, D.; Baron, R.; van Gunsteren, W. F. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1934.

(17) Chiu, S.-W.; Pandit, Sagar, A.; Scott, H. L.; Jakobsson, E. *J. Phys. Chem. B.* **2009**, *113*, 2748.

(18) Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. *J. Chem. Theory Comput.* **2008**, *4*, 435.

(19) van der Spoel, D.; Lindahl, E.; Hess, B.; Kutzner, C.; van Buuren, A. R.; Apol, E.; Meulenhoff, P. J.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; van Drunen, R.; Berendsen, H. J. C. GROMACS User Manual Version 4.0. http://www.gromacs.org/Documentation/Manual (accessed Jan 12, 2010).

(20) Evans, D. J.; Holian, B. I. *J. Chem. Phys.* **1985**, *83*, 4069.

(21) Parinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 182.

(22) Daura, X.; Mark, A. E.; van Gunsteren, W. F. *J. Comput. Chem.* **1998**, *19*, 535.

(23) Flower, J. E.; Shaefer., H. F., III. *J. Am. Chem. Soc.* **1995**, *117*, 446.

(24) Pérez, J. F.; Hadad, C. Z.; Restrepo, A. *Int. J. Quantum Chem.* **2008**, *108*, 1653.

(25) Ben-Naim, A.; Marcus, Y. *J. Chem. Phys.* **1984**, *81*, 2016.

(26) Ben-Naim, A. *J. Solution Chem.* **2001**, *30*, 475.

(27) Wescott, J. T.; Fisher, L. R.; Hanna, S. *J. Chem. Phys.* **2002**, *116*, 2361.

(28) Marrink, S. J.; de Vries, A. H.; Mark, A. E. *J. Phys. Chem. B* **2004**, *108*, 750.

(29) Baron, R.; de Vries, A. H.; Hünenberger, P. H.; van Gunsteren, W. F. *J. Phys. Chem. B.* **2006**, *110*, 8464.

(30) *CRC Handbook of Chemistry and Physics*, 72nd ed.; Lide, D. R., Eds.; CRC Press. Inc.: Boca Raton, FL, 1992.

(31) *Chemical Properties Handbook: Physical, Thermodynamic, Environmental, Transport, Safety, and Health Related Properties for Organic and Inorganic Chemicals*, 1st ed.; Yaws, C. L., Ed.; McGraw-Hill: New York, 1998.

(32) Nielsen, S. O.; Lopez, C. F.; Srinivas, G.; Klein, M. L. *J. Chem. Phys.* **2003**, *119*, 7043.

(33) Buckingham, R. A. *Proc. R. Soc. London, Ser. A* **1938**, *168*, 264.

(34) Engler, E. M.; Andose, J. D.; Schleyer, P. v. R. *J. Am. Chem. Soc.* **1973**, *95*, 8005.

(35) Allinger, N. L. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

(36) Allinger, N. L.; Yuh, Y. H.; Lii, J. H. *J. Am. Chem. Soc.* **1989**, *111*, 8551.

(37) Mayo, S. L.; Olafson, B. D.; Goddard, W. A., III. *J. Phys. Chem.* **1990**, *94*, 8897.

(38) Dillen, J. M. L. *J. Comput. Chem.* **1995**, *16*, 595.

(39) Allinger, N. L.; Chen, K.; Lii, J. H. *J. Comput. Chem.* **1996**, *17*, 642.

(40) Baron, R.; Trzesniak, D.; de Vries, A. H.; Elsener, A.; Siewert, J.; Marrink, S. J.; van Gunsteren, W. F. *ChemPhysChem* **2007**, *8*, 452.

(41) Holz, M.; Heil, S. R.; Sacco, A. *Phys. Chem. Chem. Phys.* **2000**, *2*, 4740.

(42) Groot, R. D.; Rabone, K. L. *Biophys. J.* **2001**, *81*, 725.

(43) Jalabert, R.; Das Sarma, S. *Phys. Rev. A.* **1988**, *37*, 2614.

(44) Herrero, C. P. *J. Phys.: Condens. Matter.* **2008**, *20*, 295230.

(45) Stejskal, E. O.; Tanner, J. E. *J. Chem. Phys.* **1965**, *42*, 288.

(46) von Meerwall, E.; Beckman, S.; Jang, J.; Mattice, W. L. *J. Chem. Phys.* **1998**, *108*, 4299.

(47) Díaz Peña, M.; Tardajos, G. *J. Chem. Thermodyn.* **1978**, *10*, 19.

(48) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311.

(49) Berendsen, H. J. C.; van Gunsteren, W. F. Practical algorithms for dynamic simulations. In *Molecular-Dynamics Simulation of Statistical-Mechanical Systems*; Ciccotti, G., Hoover, W. G., Eds.; North-Holland: Amsterdam, 1986; Course 97, p 43.

(50) Berendsen, H. J. C.; van Gunsteren, W. F. Molecular dynamics simulation: Techniques and approaches. In *Molecular Liquids—Dynamics and Interactions*; Barnes, A. J., Orville-Thomas, W. J., Yarwood, J., Eds.; Reidell: Dordrecht, The Netherlands, 1984; Vol. 47, pp 5−500.

CT900475P