PCCP

Combinations of simulation and integral equation theory

Stefan M. Kast

Institut für Physikalische Chemie, Technische Universität Darmstadt, Petersenstraße 20, 64287 Darmstadt, Germany

Received 9th July 2001, Accepted 9th October 2001
First published as an Advance Article on the web

Much work has been done to accelerate the prediction of structure and thermodynamics, particularly of the chemical potential of condensed phases by means of molecular simulation, requiring large computational power or simplified simulation procedures. This work demonstrates that the combination of simulation with approximate calculations based on integral equation theory provides the necessary information for quantitatively correcting simulation deficiencies. These ideas are illustrated with two examples: the calculation of artefact-free radial distribution functions from truncated potentials and the accurate prediction of excess chemical potentials from a single trajectory with the full potential.

1 Introduction

Modelling increasingly complex systems requires the use of large computational power or alternatively simplified models and simulation procedures in order to yield at least approximate results. The prediction of accurate free energies and chemical potentials by thermodynamic integration or perturbation theory is particularly tedious as this requires to gradually increase the interaction potentials within in a series of Monte Carlo (MC) or molecular dynamics (MD) simulations.[1] Simplifications like the use of continuum models for solvent degrees of freedom[2] are only a partial remedy as this changes the system Hamiltonian. It is therefore of utmost interest to have available a methodology that uses on one hand limited simulation data, generated even with simplified, for instance truncated, potentials, and, on the other hand, a fast corrective device. In this way, the computational speed would be competitive with, for instance, continuum models while simultaneously the original Hamiltonian is retained.

It is the purpose of this work to demonstrate that the use of statistical-mechanical integral equation (IE) theory[3] is a valuable step in this direction. IE theory tries to predict atomic and molecular distribution functions directly from a knowledge of the system Hamiltonian. From these the thermodynamic properties of the system are calculated. Of particular importance in this field are the one-dimensional reference interaction site model (RISM) equations[3,4] (also termed site–site Ornstein–Zernike equations in the literature) that yield the radially averaged site–site distribution functions and, more recently, the 3D-RISM approach.[5,6] For practical purposes, approximations have to be made, notably the introduction of an approximate closure relation like the hypernetted chain (HNC) equation, which lead to well-characterised deficiencies of the theory. For recent reviews, the reader is referred to refs. 7 and 8.

As will be worked out later, we can cope with the systematic drawbacks of both approaches, the simulation and the IE theories, in the sense of complementarity. Advantages of both methodologies, formal ensemble accuracy of the simulation and the computational speed of the IE theories, can be combined to compensate for each other's deficiencies, reaching greater predictive power with limited resources.

Besides the extraction of data for mere comparison of IE theory with simulation, predictive combinations of molecular simulation and integral equation theory have rarely been attempted in the past. The earliest work was done by Verlet[9] (see also ref. 10) for the purpose of extrapolating the radial distribution functions of simple liquids from MD simulations beyond the range determined by the simulation cell. Along these lines Trokhymchuk *et al.*[11] estimated the long range behaviour of liquid water. Their method was later applied to the calculation of small wavenumber structure factors[12] and relative permittivities.[13] With a conceptually much simpler approach, Roux and co-workers[14] used MD simulations of water to extract the solvent susceptibility that was subsequently used for the IE-based prediction of hydration free energies. Kambayashi and Chihara[15] corrected potential truncation artefacts in the spirit of Lado's[16] reference hypernetted chain (RHNC) theory. Polymer (PRISM) IE theories[17,18] extract the intramolecular distribution from single chain MC simulations that are then used for the IE-based prediction of bulk properties. Hirata and co-workers[19,20] coupled the IE treatment of the solvent phase with MC sampling of intramolecular degrees of freedom for biomolecules. Schmidt and Kast[21] recently computed binding constants for crown ether complexes with a combined RISM/MC free energy perturbation approach.

This work utilises perturbation-theoretic approaches to the excess chemical potential in order to address the questions of (1) extrapolation of cutoff simulation data to the full potential and (2) determination of accurate chemical potentials from a single simulation. The distorting effects of applying simple potential truncation, switching or shifting functions as compared to the presumably "exact" Ewald[22] summation for Coulomb potentials, have been known for quite some time.[23–25] They have also been treated on the basis of RISM/HNC theory.[26] Here, a corrective scheme will be described to yield accurate Ewald-limit radial distribution functions for molecular liquids from shifted-force potential simulations. The question of free energy determination from a single trajectory has been addressed in the past on the basis of a Taylor series expansion of the free energy from a single reference state[27] and, more successfully, *via* linear response theory[28,29] by calibrating free parameters with experimental data. In the following, a different method is formulated yielding chemical potentials from radial distribution functions at full potential alone.

DOI: 10.1039/b106075a

This journal is © The Owner Societies 2001

After a description of the basic theory and the models used in the next sections, the results for liquid argon and for water will be presented in section 3. Finally, some conclusions are drawn in the last section.

2 Theory and methods

2.1 Background

The 1D-RISM equation for pure fluids composed of sites α and γ as used in this work is

$$\boldsymbol{h} = \boldsymbol{\omega} * \boldsymbol{c} * \boldsymbol{\omega} + \boldsymbol{\omega} * \boldsymbol{c} * (\rho \boldsymbol{h}), \tag{1}$$

where $\boldsymbol{h} = (h_{\alpha\gamma}(r))$ is the matrix of the total correlation functions (the radial distribution function is $g = h + 1$), $\boldsymbol{c} = (c_{\alpha\gamma}(r))$ are the direct correlation functions, $\boldsymbol{\omega} = (\omega_{\alpha\gamma}(r))$ are the intramolecular correlation functions (for rigid molecules these are normalised Dirac delta functions constraining the site distance), ρ is the site density, r is the distance and the asterisk denotes convolution. This equation must be supplied with a closure relation

$$h_{\alpha\gamma} + 1 = \exp(-\beta u_{\alpha\gamma} + h_{\alpha\gamma} - c_{\alpha\gamma} + B_{\alpha\gamma}) \tag{2}$$

with the pair potential $u_{\alpha\gamma}$ and $\beta = 1/kT$ where k is the Boltzmann constant and T is the absolute temperature; $B_{\alpha\gamma}$ terms are so-called bridge functions. In the HNC approximation B identically vanishes. Eqns. (1) and (2) can be solved simultaneously in an iterative scheme to yield the correlation functions and finally the thermodynamics.

The bridge functions in eqn. (2) have a clear diagrammatic meaning for the atomic and the molecular Ornstein–Zernike equations only[3,30] and were extensively studied by Rosenfeld and Ashcroft.[31] Information about B can be deduced from a variety of sources, for instance from diagrammatic expansion,[30] from a series in powers of the density,[32] from arguments enforcing consistency of various thermodynamic properties[33,34] (see also related work[35]) and from simulation.[15,31,36,37] All these works were concerned with simple liquids only. In all others cases, and this view will be adopted here, eqn. (2) can be interpreted as a formal relation that pragmatically corrects the deficiencies of the HNC theory by the introduction of a nonzero B. In this sense, eqn. (2) and its 3D generalisation have only rarely been used beyond the HNC approximation for studies of multi-site molecular liquids.[14,38,39]

An efficient approach to the extraction of bridge functions from MD simulation data as used in this work is given by constraining the RISM g functions to the values of the simulation by defining[14]

$$B_{\alpha\gamma}(r) = S(r)[\ln(g^{\mathrm{MD}}_{\alpha\gamma}(r)) + \beta u_{\alpha\gamma}(r) - h_{\alpha\gamma}(r) + c_{\alpha\gamma}(r)] \tag{3}$$

where the switching function $S(r)$ is a simple cubic polynomial that goes smoothly from 1 to 0 over a range defined by the simulation cell. Beyond this range the correlation functions are extrapolated within the HNC approximation. We see immediately that in regions where g^{MD} is zero the explicit value of B is numerically ill-defined, *i.e.* particularly for $r \rightarrow 0$. Consequently, much effort is needed to get an exact B near $r = 0$, and the work of Llano-Restrepo and Chapman,[36,37] who essentially applied the umbrella sampling technique, is the only detailed study available in this direction.

The excess chemical potential within the HNC theory is given in closed form[40] by

$$\Delta\mu^{\mathrm{HNC}}[\boldsymbol{h}, \boldsymbol{c}] = \beta^{-1}\rho \sum_{\alpha\gamma} \int \mathrm{d}\boldsymbol{r} \left(\tfrac{1}{2} > h^2_{\alpha\gamma} - c_{\alpha\gamma} - \tfrac{1}{2} h_{\alpha\gamma} c_{\alpha\gamma} \right), \tag{4}$$

where the summation indices run over all site pairs. The square brackets denote the fact that $\Delta\mu$ is a functional of all $h_{\alpha\gamma}(r)$ and $c_{\alpha\gamma}(r)$. The generalisation to nonzero bridge functions is[38]

$$\Delta\mu^{\mathrm{B}}[\boldsymbol{h}, \boldsymbol{c}, \boldsymbol{B}] = \beta^{-1}\rho \sum_{\alpha\gamma} \int \mathrm{d}\boldsymbol{r} \Big(\tfrac{1}{2} h^2_{\alpha\gamma} - c_{\alpha\gamma} - \tfrac{1}{2} h_{\alpha\gamma} c_{\alpha\gamma} + \\ + B_{\alpha\gamma} + \int_0^1 \mathrm{d}\lambda h_{\alpha\gamma}(\lambda) \frac{\partial B_{\alpha\gamma}(\lambda)}{\partial \lambda} \Big). \tag{5}$$

Here, the "charging" or coupling parameter λ "switches on" the potential and varies from 0 to 1. It is obvious for the application of this "direct" formula that $B(r)$ must be known precisely over the full range of distances. This rules out its use when correlation functions from simple molecular simulations are used. If a bridge functions model is applied that consists of a dependence on $t = h - c$, *i.e.* $B = B(t)$ so that the coupling parameters integration in eqn. (5) can be done analytically (see for instance ref. 34), we must furthermore make sure that the essential physics is captured correctly for all values of λ. Hence, utilising eqn. (5) for the prediction of free energies is impractical.

2.2 Perturbation theory

As an alternative to the direct formula one can use perturbation-theoretic approaches. Kovalenko and Hirata[39] derived a first order approximation by expanding the excess chemical potential into a functional Taylor series in terms of Mayer functions of B around the HNC solution:

$$\Delta\mu \approx \Delta\mu^{\mathrm{HNC}}[\boldsymbol{h}^{\mathrm{HNC}}, \boldsymbol{c}^{\mathrm{HNC}}] \\ + \beta^{-1}\rho \sum_{\alpha\gamma} \int \mathrm{d}\boldsymbol{r} g^{\mathrm{HNC}}_{\alpha\gamma}(r)[\exp(B_{\alpha\gamma}(r)) - 1]. \tag{6}$$

Here, an explicit bridge function is needed. For eqn. (6) to be useful we must therefore expect that B is well known at least within the range of nonzero g^{HNC}. This could pose a problem in the case of polar systems where it is known that g^{HNC} tends to have its first major peak at a notably smaller distance as compared to simulation.[14] Application of eqn. (6) under those circumstances can lead to drastic errors.

An interesting feature of the perturbation expansion mentioned above is that it can be used to correct radial distribution functions from simulations with a specified cutoff distance, $r_{\mathrm{c},1}$, to another one, $r_{\mathrm{c},2}$. Let us supply each function in eqn. (6) with a subscript indicating calculations with a specific cutoff distance. By assuming that the bridge functions are short-ranged (for which there is some evidence at least for simple liquids[31]) and that the dominant distorting effect in the case of shifted-force potentials are located around the cutoff distance,[24,26] the second term (and, in fact, higher terms as well) will be almost independent of a given r_{c}. Upon identifying the left-hand side of eqn. (6) with the hypothetical exact chemical potential $\Delta\mu^{\mathrm{MD}}$ (considering further perturbation terms) and subtracting two instances of (6) representing $r_{\mathrm{c},1}$ and $r_{\mathrm{c},2}$, we end up with

$$\Delta\mu^{\mathrm{MD}}(r_{c,2}) - \Delta\mu^{\mathrm{MD}}(r_{\mathrm{c},1}) >\approx \Delta\mu^{\mathrm{HNC}}(r_{\mathrm{c},2}) - \Delta\mu^{\mathrm{HNC}}(r_{\mathrm{c},1}). \tag{7}$$

After some algebra and using the reversible work theorem, this leads to

$$g^{\mathrm{MD}}(r_{\mathrm{c},2}/g^{\mathrm{MD}}(r_{\mathrm{c},1}) \approx g^{>\mathrm{HNC}}(r_{\mathrm{c},2})/g^{\mathrm{HNC}}(r_{\mathrm{c},1}) \tag{8}$$

which represents a simple recipe for extrapolating from $g^{\mathrm{MD}}(r_{\mathrm{c},1})$ to an unknown $g^{\mathrm{MD}}(r_{\mathrm{c},2})$ by readily accessible computations of $g^{\mathrm{HNC}}(r_{\mathrm{c},1})$ and $g^{\mathrm{HNC}}(r_{\mathrm{c},2})$ from IE theory.

For the purpose of free energy predictions an alternative form of perturbation theory could be useful. If we assume that the exact bridge function term represents a small correction to the original potential, *i.e.* $-kTB_{\alpha\gamma} \ll u_{\alpha\gamma}$, we might as well include the bridge potential in an effective potential $u_{\mathrm{eff},\alpha\gamma} = u_{\alpha\gamma} - kTB_{\alpha\gamma}$ and use it in the sense of the HNC approximation.

Repeating the steps described by Singer and Chandler[40] leads to the chemical potential estimate

$$\Delta\mu \approx \Delta\mu^{\mathrm{HNC}}[\boldsymbol{h}^{\mathrm{MD}}, > \boldsymbol{c}^{\mathrm{MD}}] \qquad (9)$$

where the superscript "HNC" means the functional given in eqn. (4). In practice, one would start with the simulated distribution functions g^{MD} and use the bridge function expression (3) for a number of switch ranges S. The correct chemical potential is then estimated by averaging over the results from eqns. (6) or (9). In what follows, results from eqn. (6) will be denoted $\Delta\mu^{\mathrm{I}}$, from eqn. (9) $\Delta\mu^{\mathrm{II}}$.

2.3 Models and numerical methods

The argon model consisted of a Lennard-Jones potential with parameters taken from ref. 41. For water, the rigid 3-site TIP3P model was used.[42] Within the RISM theory small Lennard-Jones parameters were applied to the hydrogen sites to avoid numerical singularities.[39] The CHARMM shifted-force potential[24] was used as an example for potential truncation and subsequent correction for both Lennard-Jones and Coulomb interactions.

The 1D-RISM solutions (eqns. (1)–(3)) were obtained on a logarithmic grid of 512 points ranging from 5.98×10^{-3} Å to a maximum distance of 164.02 Å. The correlation functions were iterated with a variant of the "modified direct inversion of iterative subspace" method developed by Hirata and coworkers,[43] supplied with the nonlinear fast Fourier transformation technique.[44] For the HNC theory, convergent solutions were obtained after roughly 30–40 iterations as defined by a tolerance of $\max[\Delta(\boldsymbol{h} - \boldsymbol{c})r] < 10^{-5}$ between two successive iteration steps. In the case of the bridge function closure (3), the minimal achievable tolerance was 2×10^{-3} as determined probably by the level of statistical noise of the simulated g functions. Correspondingly, the number of iterations increases to around 300.

Reference radial distribution functions from MD simulations were provided by Schilling[45] who also used the models described above. These simulations were performed in the canonical (NVT)[46,47] and in the isothermal-isobaric (NpT)[48] ensembles applying either the CHARMM shifted-force potential or Ewald summation.[22] More details about the simulations will be given in an extensive study of the truncation artefact correction for various cutoff distances and truncation schemes to be published elsewhere.[49]

The models and conditions used within the RISM calculations were chosen in correspondence to the simulation settings. In the case of free energy NpT simulations, the resulting average density was used for the IE solutions. The MD distribution functions were recorded with a spacing of 0.05 Å. The sampling points of the simulation were transformed to those of the RISM solutions by means of cubic spline interpolation.

3 Results and discussion

3.1 Cutoff corrections

To check the validity of eqn. (8) we first look at the results in a case that usually shows very limited cutoff effects: liquid argon. Fig. 1 shows the results of dividing the radial distribution functions from 12 Å and 6 Å cutoff calculations at $T = 87$ K and $\rho = 0.02104$ Å^{-3} in the HNC approximation and from MD simulations. The simulation data was gathered from a 500 ps NVT simulation of a cubic box consisting of 1040 argon atoms using 2000 configurations.[45] Assuming statistical independence, the absolute error of the distribution function reaches a maximum of 0.004 at the onset of the first peak and is on average 0.001–0.002 throughout the rest.

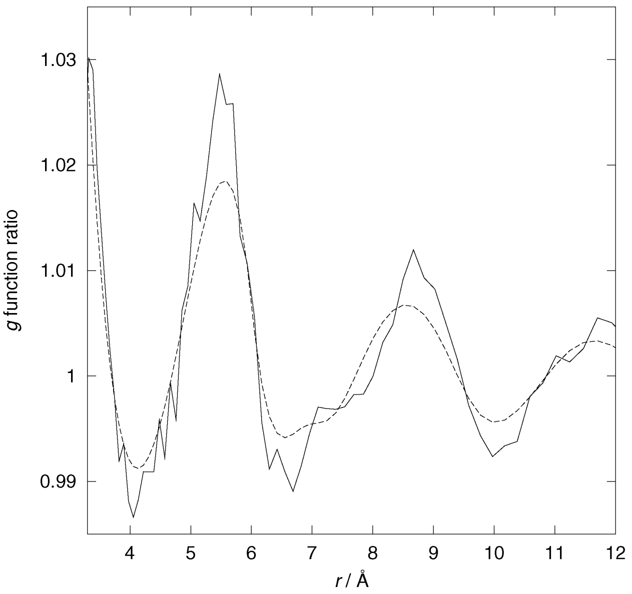

Fig. 1 Ratios of radial distribution functions of liquid argon (12 Å/6 Å cutoff distance); simulation (—) and HNC result (- - -).

As can be seen from the range of the ordinate, in this case very fine effects are resolved, so we might conclude that eqn. (8) is excellently verified. We note a systematic overshooting of the simulation oscillations as compared to the IE results which can be traced back to the size of the simulation cell. Preliminary results (not shown here) indicate that the correspondence of the curves is closer if a 5000 atom box is used.

For polar molecular liquids we can try to extrapolate shifted-force potential results to Ewald summation data,[22] *i.e.* to the limit $r_c \to \infty$. As an example, Figs. 2 and 3 compare H–H, O–H and O–O radial distribution functions of water from different sources. No extrapolation beyond the simulation cell dimensions was attempted here. The cutoff distance was chosen as 8 Å, RISM/HNC solutions and cutoff simulations (NVT) were done at $T = 300$ K and $\rho = 0.03341$ Å^{-3}. Ewald reference data comes from a NpT simulation at 1 bar, yielding an average density of 0.03291 Å^{-3}. Radial distribution functions were generated from a 200 ps trajectory using 2000 configurations. All simulations were performed with 1000 water molecules.[45] The statistical accuracy of the distribution functions is roughly the same as in the argon case described above.

As is obvious from Fig. 2 where only data from shifted-force calculations is depicted, the RISM/HNC solutions show large discrepancies from the MD data in in the region of the first peaks. On the other hand the long range part and particularly the well-known artefacts around the cutoff distance[26] are reproduced quite well.

Fig. 3 shows again the original shifted-force MD data, but also the Ewald reference functions as well as the corrected ones. The corrected results, even for such a small cutoff distance, come very close to the Ewald simulation data. Particularly the massive cutoff artefacts vanish completely. We can therefore safely use very small cutoff distances in the simulation and nevertheless get a clear picture of the solvent structure in the Ewald limit at considerably lower computational cost. For the correction according to eqn. (8), in Fig. 3 neither the dielectrically consistent RISM equations[50] nor the simple charge-scaling procedure[51] were used to adjust the relative permittivity in the IE solutions at full potential. Although the effect of using these methods is barely visible on the level of the g functions, their use is not warranted anyway: The former can be formulated as a bridge function correction to the original HNC equations, the latter changes the Hamiltonian. This would lead

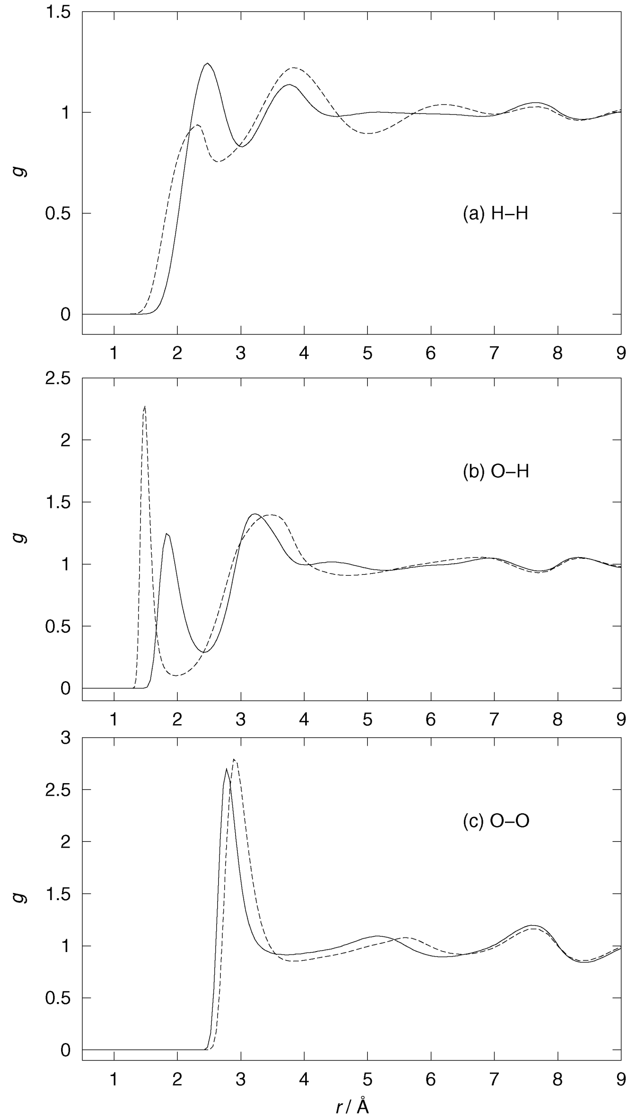

Fig. 2 Radial pair distribution functions of TIP3P water, (a) hydrogen–hydrogen, (b) oxygen–hydrogen, (c) oxygen–oxygen; 8 Å CHARMM cutoff simulation (—), corresponding HNC result (– – –).

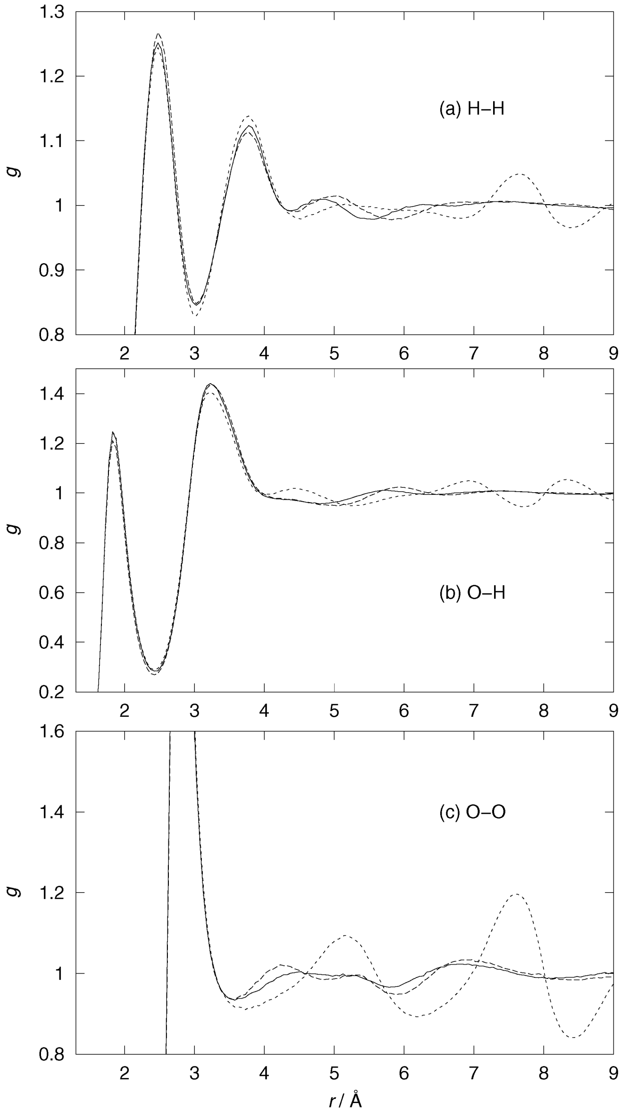

Fig. 3 Radial pair distribution functions of TIP3P water, (a) hydrogen–hydrogen, (b) oxygen–hydrogen, (c) oxygen–oxygen; Ewald simulation (—), corrected 8 Å CHARMM cutoff simulation (– – –), original cutoff simulation (- - - -).

to an element of inconsistency since the modification is applied to the computations without cutoff only.

After Kambayashi and Chihara's study[15] on cutoff correction for simple liquids of spherical particles this is the first time that a combined RISM/MD correction of cutoff artefacts has been done for molecular liquids. The assumptions, based on which eqn. (8) is approximately valid, pose a lower limit on the cutoff distance to allow for successful correction: If the range of the bridge function interferes with regions of notable cutoff artefacts the approximation will fail. This aspect together with a detailed analysis of other observables will be studied elsewhere.[49]

3.2 Determination of excess chemical potentials

For argon with $r_c = 10$ Å and water with $r_c = 12$ Å the validity of the perturbative relations (6) and (9) was checked by predicting the excess chemical potential from the radial distribution functions at full coupling only. The reference data was taken from Schilling[45] who conducted highly accurate thermodynamic integration simulations[1] yielding a statistical error of less than 1 kJ mol^{-1}. The argon g function was derived from a maximum number of 2000 configurations of a 500 ps NpT simulation of 1040 particles at $T = 87$ K and $p = 1$ bar with an average density of 0.02022 Å^{-3}.[45] The radial distribution functions for water were computed from a maximum of 1000 configurations of a 100 ps NpT simulation of 1000 particles at a pressure of 1 bar and $T = 300$ K, yielding on average a density of 0.03305 Å^{-3}.[45] The spacing of the g functions was again 0.05 Å, yielding a similar statistical accuracy of the distribution functions as in the cutoff correction study. The switching function in eqn. (3) was successively applied over a range of 1 Å starting with 13.5, 14.5, 15.5, 16.5 Å for argon and 11, 12, 13, 14 Å for water, respectively. $\Delta\mu^{\mathrm{I}}$ and $\Delta\mu^{\mathrm{II}}$ were finally averaged over these calculations. The results are summarized in Table 1. The standard deviations of the averaging process are also given there.

Table 1 Excess chemical potentials (kJ mol^{-1}) of liquid argon and TIP3P water from different approaches; the error of the MD simulation result is less than 1 kJ mol^{-1}

Substance	$\Delta\mu^{\mathrm{HNC}}$/ kJ mol^{-1}	$\Delta\mu^{\mathrm{MD}}$/ kJ mol^{-1}	$\Delta\mu^{\mathrm{I}}$/kJ mol^{-1}	$\Delta\mu^{\mathrm{II}}$/ kJ mol^{-1}
Ar	−1.56	−3.71	−3.35 ± 0.87	+2.89 ± 0.84
H_2O	−29.10	−26.63	+143.29 ± 40.39	−27.39 ± 0.71

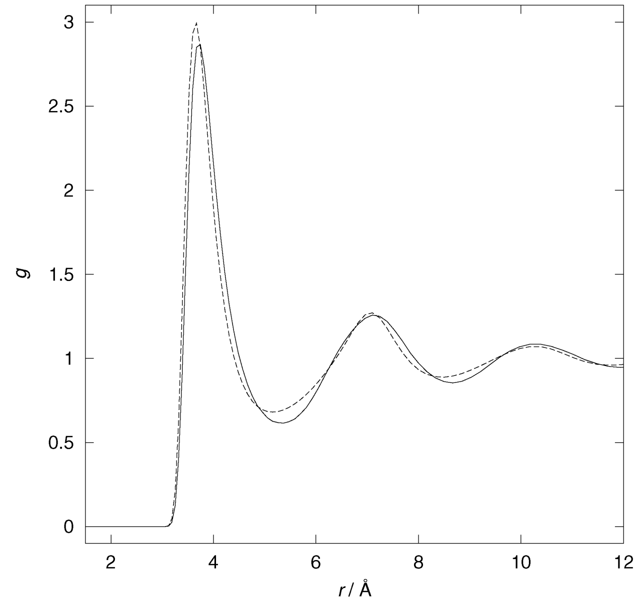

Fig. 4 Radial distribution functions of liquid argon (10 Å cutoff distance), simulation (—) and HNC result (– – –).

$\Delta\mu^{\mathrm{I}}$ is very close to the reference value in the case of argon, while $\Delta\mu^{\mathrm{II}}$ agrees very well with the simulated chemical potential for water. Both numbers are within the statistical error margin of the simulation. $\Delta\mu^{\mathrm{I}}$ for water and $\Delta\mu^{\mathrm{II}}$ for argon deviate substantially. The reason for the failure of $\Delta\mu^{\mathrm{I}}$ for water is related to fact that g^{HNC} is different from 0 in a range where the bridge function is ill-defined.

Looking at Fig. 4 reveals the already good quality of the HNC solution as compared to the simulation result for liquid argon, as has been known for quite some time.[3] This is particularly remarkable since the *g* functions for polar molecular systems like water are considerably off (*cf.* Fig. 2). On the other hand, concerning the magnitude of the perturbing bridge potential in comparison with the physical interaction, a different situation would be expected. Fig. 5 shows this data for argon and for the O–H interaction of water. In fact, the bridge correction for water is small while that for argon it is quite large. The liquid noble gas structure is dominated by repulsive packing effects which explains why a strong perturbation is needed to achieve even small effects. Fig. 5 gives us a clue as to when the perturbation formulae (6) or (9) are applicable and explains the results given in Table 1. This analysis is necessary for reliable predictions.

It is worth looking at the performance of the perturbation theory if only a fraction of the original simulation data is used for the determination of radial distribution functions. This is illustrated for both argon and water in Table 2: The result for the full data set is given in the first column while columns 2 and 3 contain the outcome when every other configuration is used; columns 4 and 5 were yielded by utilising the first and the second half of the configurations, respectively.

Table 2 Perturbation theory excess chemical potentials (kJ mol^{-1}) of liquid argon and TIP3P water from various portions of simulation data

Method	1,2,3...	1,3,5...	2,4,6...	1st half	2nd half
$\Delta\mu^{\mathrm{I}}(\mathrm{Ar})$/kJ mol^{-1}	−3.35	−4.53	−2.12	+0.88	−9.98
$\Delta\mu^{\mathrm{II}}(\mathrm{Ar})$/kJ mol^{-1}	+2.89	+4.15	+1.71	−1.11	+9.91
$\Delta\mu^{\mathrm{I}}(\mathrm{H_2O})$/kJ mol^{-1}	+143.29	+995.01	+415.95	+319.73	+95.80
$\Delta\mu^{\mathrm{II}}(\mathrm{H_2O})$/kJ mol^{-1}	−27.39	−25.00	−26.30	−32.51	−4.39

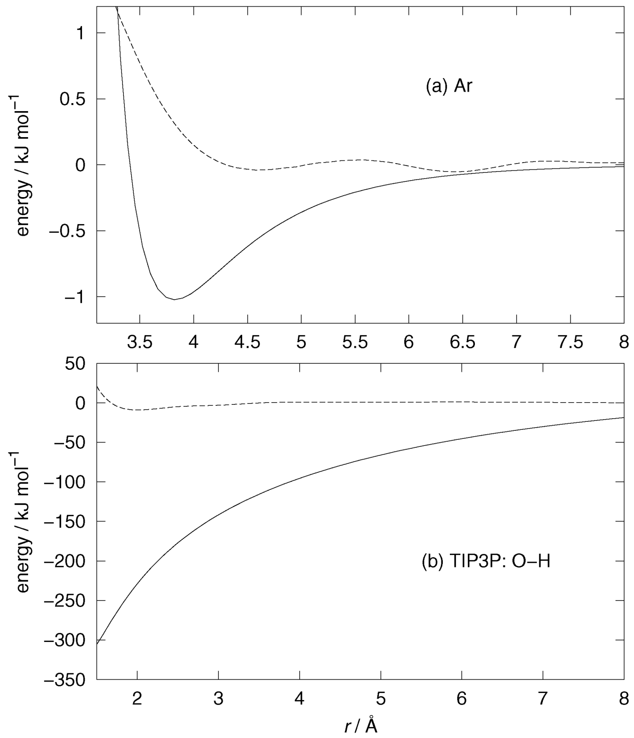

Fig. 5 Pair interaction energy (—) and bridge potential $-kTB$ (– – –) for argon (a) and oxygen–hydrogen of TIP3P water (b).

For water, $\Delta\mu^1$ scatters with no obvious numerical trend while in the argon case $\Delta\mu^{\mathrm{II}}$ appears to be extrapolated systematically in the wrong direction. For both argon ($\Delta\mu^{\mathrm{I}}$) and water ($\Delta\mu^{\mathrm{II}}$), the estimates under the influence of the reduced statistical quality from using only every other configuration are still acceptable. Using less simulation time is, however, deleterious in both cases studied: The average chemical potentials do not overlap with the simulated ones within error bars. This is related to the fact that the volume correlation time of the *NpT* simulation is comparable to the length of the shorter trajectories; the generated ensemble is therefore not representative for the average density of the full trajectory. Using every other configuration instead reduces only the statistical accuracy but avoids those systematic deviations. Analogously to checking the magnitude of the bridge perturbation, this analysis is essential for reliable free energy predictions.

4 Conclusions

It was the aim of this work to demonstrate the complementary character of molecular simulation and integral equation theory. The results clearly suggest that limited simulation data can be used in conjunction with the RISM equations, thus accelerating the computations significantly. The methodology was applied to the extrapolation of radial distribution functions from simulations with the CHARMM shifted-force potential to the Ewald limit, as well as the prediction of excess chemical potentials from the simulated radial distribution functions at full potential only.

We might expect that the prediction of Ewald-limit thermodynamic data from rather short simulations and drastically truncated potentials is within reach, particularly if a potential is used that directly leads to less cutoff artefacts compared to traditional schemes, for instance with the damped Coulomb potential.[52,53] The perturbation-theoretic free energy determination will benefit from better bridge functions that need to be developed or from the 3D-RISM/HNC approaches[5,6] for which a closed-form expression for the chemical potential is known. The loss of orientational information upon using the 1D-RISM equations is responsible for most of their deficiencies.[39] This is also indicated by the work of Krienke and co-workers[54] who thoroughly compared the performance of the RISM/HNC equations with results from MC simulations and the molecular Ornstein–Zernike equation. Extrapolation of the chemical potential deduced from truncated potentials to the Ewald limit can subsequently be done *via* eqn. (7), avoiding known problems with Ewald summation free energy simulations.[55]

An important next step will be the treatment of single solute molecules in a liquid environment. We will probably need efficient smoothing strategies in order to use the simulated distribution functions that suffer from more pronounced statistical noise as compared to pure solvent phases.

Acknowledgements

The author wishes to thank Bernd Schilling for providing his simulation results. Stimulating discussions with him as well as with Friedemann Schmidt and Jürgen Brickmann are also gratefully acknowledged.

References

1 P. Kollman, *Chem. Rev.*, 1993, **93**, 2395.
2 C. J. Cramer and D. G. Truhlar, *Chem. Rev.*, 1999, **99**, 2161.
3 J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, Academic Press, London, 2nd edn., 1991.
4 D. Chandler and H. C. Andersen, *J. Chem. Phys.*, 1972, **57**, 1930.
5 D. Beglov and B. Roux, *J. Phys. Chem. B*, 1997, **101**, 7821.
6 A. Kovalenko and F. Hirata, *Chem. Phys. Lett.*, 1998, **290**, 237.
7 F. Hirata, *Bull. Chem. Soc. Jpn.*, 1998, **71**, 1483.
8 B. Roux and T. Simonson, *Biophys. Chem.*, 1999, **78**, 1.
9 L. Verlet, *Phys. Rev.*, 1968, **163**, 201.
10 E. Lomba, M. Alvarez, G. Stell and J. A. Anta, *J. Chem. Phys.*, 1992, **97**, 4349.
11 A. D. Trokhymchuk, M. F. Holovko, E. Spohr and K. Heinzinger, *Mol. Phys.*, 1992, **77**, 903.
12 A. D. Trokhymchuk, M. F. Holovko and K. Heinzinger, *Z. Naturforsch. Teil A*, 1995, **50**, 18.
13 A. D. Trokhymchuk, M. F. Holovko and K. Heinzinger, *J. Chem. Phys.*, 1993, **99**, 2964.
14 Q. Du, D. Beglov and B. Roux, *J. Phys. Chem. B*, 2000, **104**, 796.
15 S. Kambayashi and J. Chihara, *Phys. Rev. E*, 1994, **50**, 1317.
16 F. Lado, *Phys. Rev. A*, 1973, **8**, 2548.
17 P. G. Khalatur and A. R. Khokhlov, *Mol. Phys.*, 1998, **93**, 555.
18 J. Malenkevitz, K. S. Schweizer and J. G. Curro, *Macromolecules*, 1993, **26**, 6190.
19 M. Kinoshita, Y. Okamoto and F. Hirata, *J. Am. Chem. Soc.*, 1998, **120**, 1855.
20 M. Kinoshita, Y. Okamoto and F. Hirata, *J. Chem. Phys.*, 1999, **110**, 4090.
21 K. F. Schmidt and S. M. Kast, *J. Phys. Chem. B*, submitted.
22 P. P. Ewald, *Ann. Phys. (Leipzig)*, 1921, **64**, 253.
23 P. E. Smith and B. M. Pettitt, *J. Chem. Phys.*, 1991, **95**, 8439.
24 P. J. Steinbach and B. R. Brooks, *J. Comput. Chem.*, 1994, **15**, 667.
25 H. Dufner, S. M. Kast, M. Schlenkrich and J. Brickmann, *J. Comput. Chem.*, 1997, **18**, 660.
26 C. L. Brooks III, *J. Chem. Phys.*, 1987, **86**, 5156.
27 P. E. Smith and W. F. van Gunsteren, *J. Chem. Phys.*, 1994, **100**, 577.
28 H. A. Carlson and W. L. Jorgensen, *J. Phys. Chem.*, 1995, **99**, 10 667.
29 J. Åqvist and T. Hansson, *J. Phys. Chem.*, 1996, **100**, 9512.
30 S. Ichimaru, *Rev. Mod. Phys.*, 1982, **54**, 1017.
31 Y. Rosenfeld and N. W. Ashcroft, *Phys. Rev. A*, 1979, **20**, 1208.
32 Y. Zhou and G. Stell, *J. Chem. Phys.*, 1990, **92**, 5533.
33 A. Vompe and G. A. Martynov, *J. Chem. Phys.*, 1994, **100**, 5249.
34 L. L. Lee, *J. Chem. Phys.*, 1997, **107**, 7360.
35 J. M. Bomont and J. L. Bretonnet, *J. Chem. Phys.*, 2001, **114**, 4141.
36 M. Llano-Restrepo and W. G. Chapman, *J. Chem. Phys.*, 1992, **97**, 2046.
37 M. Llano-Restrepo and W. G. Chapman, *J. Chem. Phys.*, 1994, **100**, 5139.
38 L. Lue and D. Blankschtein, *J. Phys. Chem.*, 1992, **96**, 8582.
39 A. Kovalenko and F. Hirata, *J. Chem. Phys.*, 2000, **113**, 2793.
40 S. J. Singer and D. Chandler, *Mol. Phys.*, 1985, **55**, 621.
41 J. P. Hansen and M. L. Klein, *Phys. Rev. B*, 1976, **13**, 878.
42 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.
43 A. Kovalenko, S. Ten-no and F. Hirata, *J. Comput. Chem.*, 1999, **20**, 928.
44 P. J. Rossky and H. L. Friedman, *J. Chem. Phys.*, 1980, **72**, 5694.
45 B. Schilling, personal communications.
46 S. M. Kast and J. Brickmann, *J. Chem. Phys.*, 1996, **104**, 3732.
47 S. M. Kast, manuscript in preparation.
48 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684.
49 S. M. Kast, B. Schilling and K. F. Schmidt, manuscript in preparation.
50 J. Perkyns and B. M. Pettitt, *J. Chem. Phys.*, 1992, **97**, 7656.
51 P. J. Rossky, B. M. Pettitt and G. Stell, *Mol. Phys.*, 1983, **50**, 1263.
52 D. Wolf, P. Keblinski, S. R. Phillpot and J. Eggebrecht, *J. Chem. Phys.*, 1999, **110**, 8254.
53 D. Zahn, B. Schilling and S. M. Kast, *J. Chem. Phys.*, submitted.
54 J. Richardi, P. H. Fries, R. Fischer, S. Rast and H. Krienke, *Mol. Phys.*, 1998, **93**, 925.
55 P. H. Hünenberger and J. A. McCammon, *J. Chem. Phys.*, 1999, **110**, 1856.