



HAL
open science

Accurate prediction of hydration free energies and solvation structures using molecular density functional theory with a simple bridge functional

Daniel Borgis, Sohvi Luukkonen, Luc Belloni, Guillaume Jeanmairet

► To cite this version:

Daniel Borgis, Sohvi Luukkonen, Luc Belloni, Guillaume Jeanmairet. Accurate prediction of hydration free energies and solvation structures using molecular density functional theory with a simple bridge functional. *The Journal of Chemical Physics*, 2021, 155 (2), pp.024117. 10.1063/5.0057506 . hal-03454131

HAL Id: hal-03454131

<https://hal.science/hal-03454131>

Submitted on 29 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Accurate prediction of hydration free energies and solvation structures using molecular density functional theory with a simple bridge functional

Daniel Borgis,^{1,2, a)} Sohvi Luukkonen,¹ Luc Belloni,³ and Guillaume Jeanmairat^{4,5, b)}

¹*Maison de la Simulation, USR 3441 CNRS-CEA-Université Paris-Saclay, 91191 Gif-sur-Yvette, France*

²*PASTEUR, Département de chimie, École normale supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France*

³*Université Paris-Saclay, CEA, CNRS, NIMBE, 91191, Gif-sur-Yvette, France*

⁴*Sorbonne Université, CNRS, Physico-Chimie des Électrolytes et Nanosystèmes Interfaciaux, PHENIX, F-75005 Paris, France.*

⁵*Réseau sur le Stockage Electrochimique de l'Énergie (RS2E), FR CNRS 3459, 80039 Amiens Cedex, France*

This paper assesses the capability of molecular density functional theory to predict efficiently and accurately the hydration free-energies of molecular solutes and the surrounding microscopic water structure. A wide range of solutes were investigated including hydrophobes, water as a solute, and the FreeSolv database containing 642 drug-like molecules having a variety of shapes and sizes. The usual second-order approximation of the theory is corrected by a third-order, angular-independent bridge functional. The overall functional is parameter-free in the sense that the only inputs are bulk water properties, independent of the solutes considered. These inputs are the direct correlation function, compressibility, liquid-gas surface tension, and excess chemical potential of the solvent. Compared to molecular simulations with the same force field and the same fixed solute geometries, the present theory is shown to describe accurately the solvation free-energy and structure of both hydrophobic and hydrophilic solutes. Overall, the method yields a precision of order $0.5 k_B T$ for the hydration free energies of the FreeSolv database, with a computer speed-up of 3 orders of magnitude. The theory remains to be improved for a better description of the H-bonding structure and the hydration free-energy of charged solutes.

I. INTRODUCTION

Solvation free energies (SFE), and especially hydration free energies (HFE), are fundamental thermodynamic quantities. They are at the heart of a broad range of interesting physical properties such as solubilities, partition coefficients between immiscible solvents, activity coefficients or binding free-energies in solutions.

Computing SFE is difficult as it requires the sampling of all the possible thermodynamic states that can be visited during the transformation from the initial state (usually the neat solvent) to the final one (solute in solvent). Multiple approaches for solvation free energy calculation have been developed along the years. These start from physically-simple but fast implicit solvent approaches where the solvent is treated as a dielectric continuum, plus a surface area correction for cavitation free-energies. Implicit solvent models involve methods based on the Poisson-Boltzmann equation¹, the simpler Generalized Born Approximation², or more sophisticated free-energy functionals coupling hydrophobicity, dispersion, and electrostatics³. They lack information on the microscopic solvent structure that can be indeed an important factor of the solvation process.

Rigorous simulation-based free-energy methods have been developed to take into account the solvent molecules explicitly. They range from simple exponential averaging introduced 60 years ago by Zwanzig⁴ to more sophisticated methods employing non-physical intermediate states⁵⁻⁷, and requiring multiple ergodic molecular dynamics (MD)

or Monte-Carlo (MC) simulations to be run, possibly in parallel, for a single free energy estimate. These methods define today's standard in terms of free energy predictions.

Other approaches accounting explicitly for the molecular nature of the solvent have roots in the liquid state theories. As an intermediate method, let us draw attention on the energy-representation method, in which an energy-based integral equation theory makes it possible to infer the solvation free energy from the sampling of the solute-solvent interaction energy distribution within only two simulations, with and without solute^{8,9}. Turning to 'simulation-free' methods, the so-called morphometric approach to solvation elaborates on scaled-particle theory¹⁰ and fundamental measure theory¹¹ to provide fast estimations of the solvation free-energy of complex molecules¹²⁻¹⁴. Field theory or classical density functional theory (cDFT) with a simplified dipolar solvent model leads to a local¹⁵ or non-local Poisson-Boltzmann-Langevin formalisms^{16,17}, that incorporate solvent and ions sizes. For more realistic molecular models, such as point-charge models for water, integral equation or classical density functional theory approaches have been developed. A first class of those relies on a reference interaction site model (RISM) that is solved through integral equations, as in 3D-RISM¹⁸⁻²⁵, or through a classical DFT (cDFT) formalism using a site-based functional^{26,27}. These methods are having large success and are gaining momentum because of their good balance between precision, simplicity and speed. They have been applied to SFE predictions²⁷⁻³⁵ as well as to a number of structural biology problems³⁶⁻⁴¹. Nevertheless, they rely on site-site correlations only, not on a full molecular description, and hence they require some compromises with phenomenological corrections.

The molecular integral equation theory⁴²⁻⁴⁴ and molec-

^{a)}Electronic mail: daniel.borgis@ens.fr

^{b)}Electronic mail: guillaume.jeanmairat@sorbonne-universite.fr

ular density functional theory (MDFT)⁴⁵⁻⁴⁸ are the only methods based on liquid state theories that keep the full molecular picture by solving the (angular-dependent) molecular Ornstein-Zernike equation (MOZ)^{49,50}. They are known to be diagrammatically consistent and free of some patches that are necessary in the RISM approach, for example to get the dielectric constant correct. Since Ding et al.⁵¹, MDFT can be solved efficiently at the hyper-netted chain (HNC) level of approximation. In a cDFT language, the HNC approximation can be understood as a second-order Taylor expansion of the (unknown) excess functional around a reference homogeneous solvent at density ρ_b . It can thus be considered as exact up to second order in $\Delta\rho = \rho - \rho_b$, and can be used as a sound starting point to be systematically improved. This can be done by developing so-called bridge functionals that bring terms of order $\Delta\rho^3$ and more⁵²⁻⁵⁸.

This paper is the follow up of two recent contributions. The first one was dedicated to MDFT in its 'crudest' HNC approximation⁵⁹. While many of the solvation features are satisfactorily reproduced, some limitations of the HNC approximations were also identified. In particular, the cavity energies are dramatically overestimated. The scope of the second paper was to correct the SFE for hydrophobic solutes starting from an angular independent HNC functional. Introducing a simple weighted-density approximation (WDA) bridge functional beyond HNC allows to properly reproduce the SFE of hydrophobes of various sizes, from microscopic ones to the nanoscale.⁶⁰ Here, we extend this WDA approach to hydrophobic and hydrophilic solvation. We aim to correct the angular-dependent HNC by adding an angular-independent WDA bridge functional, similar to that in Ref.⁶⁰. This bridge accounts in an effective manner for orders in $\Delta\rho$ higher than quadratic that are neglected in HNC.

The outline of the paper is as follows. In section II, we recall briefly the MDFT framework in the HNC approximation and beyond, and we introduce our WDA bridge functional and its parametrisation procedure. In section III, we apply the new functional to hydrophobic solutes and hydrophobic pairs. In section IV, this is extended to the specific case of water in water, and then to molecular solutes of arbitrary 3D-shape and charge distribution, spanning the whole range from hydrophobic to hydrophilic chemical species. Conclusions and perspectives are drawn in section V.

II. MOLECULAR DENSITY FUNCTIONAL THEORY IN THE HNC APPROXIMATION AND BEYOND

A. General expression of the functional

We recall briefly the well-established results of molecular density functional theory: the free energy of solvation of a molecule can be written as

$$\Delta G_{\text{solv}} = \min(\mathcal{F}_{\text{id}}[\rho] + \mathcal{F}_{\text{ext}}[\rho] + \mathcal{F}_{\text{exc}}[\rho]), \quad (1)$$

where \mathcal{F}_{id} , \mathcal{F}_{ext} , and \mathcal{F}_{exc} are the ideal, external and excess functionals depending on the solvent position and orienta-

tion density field, $\rho(\mathbf{r}, \omega)$. The notation ω stands for the three Euler angles that describe the orientation of a rigid body in three dimensions. The whole MDFT procedure can be summarized as how to find the spatial and angular density that minimizes the sum of the three contributions. By virtue of the classical DFT theorems derived by Evans⁶¹, the density minimizing the functional is also the thermodynamic equilibrium density of the solvent in the external potential created by the solute, $\rho_{eq}(\mathbf{r}, \omega)$.

The first term of the right hand side of eq. 1 is the ideal contribution that corresponds to the entropy of the non-interacting liquid

$$\mathcal{F}_{\text{id}}[\rho] = k_B T \int d\mathbf{r} d\omega \rho(\mathbf{r}, \omega) \log\left(\frac{\rho(\mathbf{r}, \omega)}{\rho_b}\right) - \Delta\rho(\mathbf{r}, \omega), \quad (2)$$

where k_B is the Boltzmann constant, T is the temperature and $\rho_b \equiv n_b/8\pi^2$ is the homogeneous bulk solvent density with n_b the number density. The external contribution is the direct cost of the interaction of the solute with the solvent density, the interaction potential $v_{\text{ext}}(\mathbf{r}, \omega)$ being typically given by classical force fields,

$$\mathcal{F}_{\text{ext}}[\rho] = \int d\mathbf{r} d\omega v_{\text{ext}}(\mathbf{r}, \omega) \rho(\mathbf{r}, \omega). \quad (3)$$

The last term of eq. 1 is the excess contribution due to solvent-solvent interaction. As usual in liquid state theories, it is known analytically as an infinite diagrammatic resummation of virial diagrams, but is not tractable numerically. It can be approximated by a Taylor expansion around the homogeneous bulk solvent density ρ_b . Truncation at second order in $\Delta\rho(\mathbf{r}, \omega) = \rho(\mathbf{r}, \omega) - \rho_b$ yields the so-called hyper-netted chain (HNC) approximation and higher orders are gently put into the so-called bridge functional, \mathcal{F}_B :

$$\begin{aligned} \mathcal{F}_{\text{exc}}[\rho] &= \mathcal{F}_{\text{HNC}}[\rho] + \mathcal{F}_B[\rho] \\ &= -\frac{k_B T}{2} \int d\mathbf{r} d\omega d\mathbf{r}' d\omega' \Delta\rho(\mathbf{r}, \omega) \\ &\quad \times c^{(2)}(\mathbf{r} - \mathbf{r}', \omega, \omega') \Delta\rho(\mathbf{r}', \omega') + \mathcal{F}_B[\rho], \quad (4) \end{aligned}$$

where $c^{(2)}$ is the two-body, direct correlation function of the bulk solvent, which depends upon the relative position of two solvent molecules and their respective orientations. It is an input of the theory that needs to be calculated once per solvent and thermodynamic condition. We use the ones obtained by Belloni for either the transferable intermolecular potential 3-points (TIP3P⁶²) or the extended simple point charge (SPC/E⁶³) models of water, using a mixture of MC calculations and integral equations to invert the simulation data and make the asymptotic behaviours well-defined.⁶⁴ Even in its simplest HNC form, the efficient calculation of the excess term is a numerical challenge because of the spatial and angular convolution. For this we use expansions of both the density and the direct correlation function onto generalized spherical harmonics; all technicalities are described in Ref.⁵¹. $\mathcal{F}_B[\rho]$ involves by definition terms of order $\Delta\rho^3$ and higher.

B. Form and parametrization of the bridge functional

Following Ref.⁶⁰, the purpose of this paper is to explore the simplest, conceivable approximations for the bridge functional, which amount to neglect the angular dependence and suppose that \mathcal{F}_B only depends on the number density $n(\mathbf{r}) = \int d\omega \rho(\mathbf{r}, \omega)$, that is to say $\mathcal{F}_B[\rho] = \mathcal{F}_B[n]$.

Various bridge functionals of this sort have been proposed in the past, built in particular from the fundamental measure theory of hard-sphere fluids^{65–68}, and dedicated to simple atomic fluids⁶⁹ or extended to water^{26,27,52}. Inspired by our previous attempts^{53–55}, and the fact that weighted density approximations (WDA) are from its early foundations^{70–72} at the heart of classical DFT, we propose here the simplest possible WDA

$$\mathcal{F}_B[n] = k_B T n_b \int d\mathbf{r} f_B(\Delta \bar{n}(\mathbf{r})/n_b), \quad (5)$$

where $\bar{n}(\mathbf{r})$ is a weighted density using a Gaussian weighting function

$$\bar{n}(\mathbf{r}) = \int d\mathbf{r}' w(|\mathbf{r} - \mathbf{r}'|) n(\mathbf{r}') \quad (6)$$

$$w(r) = (2\pi\sigma_w^2)^{-3/2} \exp(-r^2/2\sigma_w^2). \quad (7)$$

In contrast to Ref.⁶⁰, we propose here a function f_B developed at the lowest cubic order in Δn , *i.e.*, with $x = \bar{n}/n_b$ and $\Delta x = (\bar{n} - n_b)/n_b$

$$f_B(x) = a \Delta x^3. \quad (8)$$

The parameter a is fixed by the liquid-gas coexisting condition, *i.e.*, for an homogeneous fluid of density n , the equality of the free energy of the gas, $n = 0$, and of the liquid at its bulk density, n_b . Expressing the free energy functional in eqs 2-4 for an homogeneous fluid of density $\rho = n/8\pi^2$ in a volume V yields for the free energy per molecule

$$\beta\mathcal{F}(n)/N = f(x) = x \log x - x + 1 - \frac{1}{2} n_b \hat{c}_s(0) \Delta x^2 + a \Delta x^3, \quad (9)$$

with $N = V n_b$. $\hat{c}_s(0) = \frac{1}{(8\pi^2)^2} \int d\mathbf{r} \int d\omega d\omega' c^{(2)}(\mathbf{r}, \omega, \omega')$ is interpreted as the Fourier transform of the spherically averaged direct correlation function taken at $\mathbf{q} = 0$. The liquid-gas coexistence condition yields $f(0) = f(1) = 0$, thus

$$a = 1 - \frac{1}{2} n_b \hat{c}_s(0) = \frac{1}{2} \left(1 + \frac{1}{n_b k_B T \chi_T} \right). \quad (10)$$

In the second equality, we have used the known relation between $\hat{c}_s(0)$ and the isothermal compressibility of the fluid, namely $n_b k_B T \chi_T = 1/(1 - n_b \hat{c}_s(0))$ ⁷³. The parameter a is thus unambiguously fixed by the bulk-fluid isothermal compressibility χ_T , namely $n_b k_B T \chi_T = 0.063$ for SPC/E water and 0.0775 for TIP3P. The dimensionless free energy curve $f(x)$ as function of reduced density $x = n/n_b$ corresponding to SPC/E water is presented in Fig. 1 together with its HNC approximation. It has indeed the form of a double-well with equal minima at $x = 0$ and $x = 1$, in contrast to the HNC version, with only a single well at bulk density.

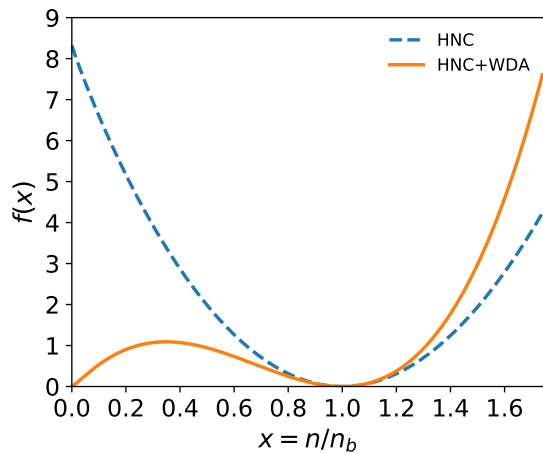


Figure 1. Dimensionless free energy per molecule of bulk SPC/E water, $f(x) = \beta\mathcal{F}(n)/N$ as a function of reduced density $x = n/n_b$, with and without the WDA bridge term.

The only parameter remaining to be fixed is the coarse-graining length σ_w and this can be done by requiring another bulk thermodynamic property to be fulfilled, namely the liquid-vapour surface tension γ . A practical way to evaluate γ is to study the solvation of a hard sphere of increasing radius R since the solvation free energy per surface area, $\Delta G_{\text{solv}}/4\pi R^2$, is known to yield asymptotically γ at large R -values when the pressure, and hence the volumetric term PV , are vanishing. This is the strategy adopted in the following section. Note that in view of previous WDA approaches, one would expect that σ_w corresponds roughly to a coarse-graining over the size of a water molecule, with an equivalent hard sphere diameter $\sigma_{hs} \simeq 2.8$ Å, as inferred from the oxygen-oxygen distribution function. Equating Gaussian and hard sphere volumes $(2\pi)^{3/2}\sigma_w^3 = \pi\sigma_{hs}^3/6$ leads to $\sigma_w \simeq 0.9$ Å.

The present approximation of the overall functional will be termed as HNC+WDA, with all angular correlations considered at HNC level, and a third-order bridge functional taken in an angular-independent, WDA form.

C. Technical issues

The theory developed above and corresponding algorithms are implemented into an in-house high performance Fortran code. With the above choice of bridge functional, the difficult part remains the computation of the HNC excess free energy in eq. 4, as well as its gradients with respect to density, that are non-local in both space and orientations. All technical details on how to handle this part efficiently were presented in Ref.⁵¹. It requires a generalised spherical harmonic expansion of the angular densities to an order m_{max} . The chosen value for m_{max} fixes the number of discretised orientations to be considered on each spatial grid point. For water, we take in general $m_{\text{max}} = 3$, corresponding to 84 orientations per grid point when the C2v symmetry is properly accounted for. This gives sufficient accuracy compared to higher order expansions, *e.g.*

$m_{\max} = 5$ corresponding to 330 orientations. Unless stated, the MDFT calculations presented below were performed with $m_{\max} = 3$. As a rule of thumb, going from $m_{\max} = 3$ to $m_{\max} = 5$ gives slightly better structural results for highly symmetrical solutes, such as spherical ones, but less so for arbitrary 3D-shapes or for the value of the solvation free energy. The evaluation of the bridge free energy, eqs 5-8, and of its gradients can be easily handled through 3D-spatial convolutions and adds only a small overhead to the overall computer process.

For the purpose of comparison with 'exact' simulation data, whenever those were not available from the literature, we have generated our own reference data for the same rigid solute geometry and same force-field parameters as in MDFT. To this end, we have used an original hybrid-4th-dimension Monte-Carlo method (H4D-MC) and the associated code developed by one of us^{74,75}. The method was shown to be more efficient than standard thermodynamics integration methods and more suited to massive parallelisation. Each hydration free energy calculation takes a few tens of cpu.h to obtain a well controlled accuracy of 0.4 kJ/mol. In contrast, a typical MDFT minimisation takes at most minutes for $m_{\max} = 3$ and a spatial grid of size 72^3 .

III. HYDROPHOBIC SOLVATION

A. Hard-sphere solute

We begin by examining the paradigmatic problem of the solvation of a perfectly hydrophobic hard sphere, which is at the basis of the standard scaled particle theory^{10,76} or more recent advances in the theory of hydrophobicity and hydrophobic interactions^{77,78}. We focus first on the solvation free-energy in SPC/E water, for which we can borrow the reference simulation results of Hummer *et al.*⁷⁹ and Huang *et al.*^{77,78}. In Fig. 2-a, we plot the hydration free energy of a hard sphere of increasing radius R computed by MDFT for different values of the coarse-graining parameter σ_w and show that $\sigma_w = 0.91 \text{ \AA}$ does fit perfectly the Hummer *et al.* results, in perfect agreement with the rough estimation given above. In Fig. 3 we plot the corresponding solvation free energy per surface area, $\Delta G_{\text{solv}}/4\pi R^2$, compared to the MC results of Huang *et al.*, including their extrapolation beyond the highest simulated value, $R = 14 \text{ \AA}$. Since the plateau value of $\Delta G_{\text{solv}}/4\pi R^2$ is associated to the liquid-vapor surface tension γ it can be stated that the WDA bridge functional in eq. 5 is parameter-free, since fixed by the values of the bulk liquid compressibility χ_T and liquid-vapor surface tension γ . In Fig. 2-b and Fig. 3 the WDA and simulation results are compared to the bare HNC results and to those corrected by the so-called improved PC correction (PC+) that we have ourselves advocated to improve the HNC results for realistic molecules^{80,81}, and that have been further uptaken by the RISM community³¹. It simply corrects the overestimated pressure in HNC by adding a term proportional to the partial molar volume of

the solute,

$$\Delta G_{\text{solv}}^{\text{PC+}} = \Delta G_{\text{HNC}} - P_{\text{HNC}}^{\text{exc}} \Delta V \quad (11)$$

where ΔG_{HNC} is the SFE predicted by the HNC functional of eqs 2-4, ΔV is the equilibrium partial molar volume of the solute

$$\Delta V = \int d\mathbf{r} \frac{n_{\text{eq}}(\mathbf{r}) - n_{\text{b}}}{n_{\text{b}}}, \quad (12)$$

and $P_{\text{HNC}}^{\text{exc}} = -\frac{1}{2}k_B T n_{\text{b}} \hat{c}_s(0)$ is the bulk excess pressure yielded by the HNC approximation. As pointed out in Ref.⁸², the PC+ correction modifies the SFE without changing the minimization process and it violates the DFT ansatz in the sense that the density minimizing eq. 11 is no longer the one obtained by minimizing eq. 1. As seen in Fig. 2-b, PC+ underestimates the SFE for small HS radii. As for the free energy per surface area in Fig. 3, the approximation cannot capture the correct asymptotic physical behaviour with a plateau and it elicits an overestimation at the nanoscale. Note that the straight pressure correction (PC) introduced in Ref.^{80,81}, that amounts to considering the total HNC pressure in eq. 11 instead of only the excess one, does yield the correct plateau behaviour but with an underestimated value of γ . It leads furthermore to a larger underestimation of ΔG than PC+ at the sub-nanoscale.⁵⁹

As a further test of the WDA bridge, we plot in the top panels of Fig. 4 the solute-water radial distribution function (rdf), $g(r)$, obtained for different values of R with either the HNC or HNC+WDA functionals. In HNC, the first-peak rises continuously to reach a high plateau around 5.5. With WDA a maximum around 2 occurs at $R = 4 \text{ \AA}$ and the peak height further decreases continuously as the sphere grows. This is indeed the correct behaviour when compared to the Monte-Carlo structures provided by Huang *et al.* in the bottom panels. At first sight, the agreement is seen excellent. This is confirmed in Fig. 5 by plotting the value of the maximum of $g(r)$ and the contact value $g(R)$ as function of R . The values inferred from the MC simulations do fit very well into the MDFT curves. HNC+WDA does predict correctly a liquid depletion, characterized by $g(R) < 1$, for values of $R \geq 8 \text{ \AA}$. This depletion will eventually lead to a dewetting transition that is predicted by MDFT in the $R \simeq 50 - 100 \text{ \AA}$ range –thus largely out of the simulation range. Indeed the HNC behaviour appears physically incorrect, except at very small R values below 2 \AA .

B. Lennard-Jones solute

We now depart from perfect hydrophobicity with the addition of solute-solvent attraction. Firstly, back to energetics, we examine the case of a Lennard-Jones spherical particle of increasing size. We rely on the molecular-dynamics results of Fujita *et al.*⁸³ obtained for a Lennard-Jones sphere of diameter σ increasing from $\sigma = 4 \text{ \AA}$ to $\sigma = 18 \text{ \AA}$ in SPC/E water. The well-depth parameter ϵ is taken constant at 0.5 kJ/mol. As seen in Fig. 6, the HNC+WDA results are quite good over the whole range of σ , with only a slight underestimation for the highest diameter considered. The HNC approximation is indeed out of range.

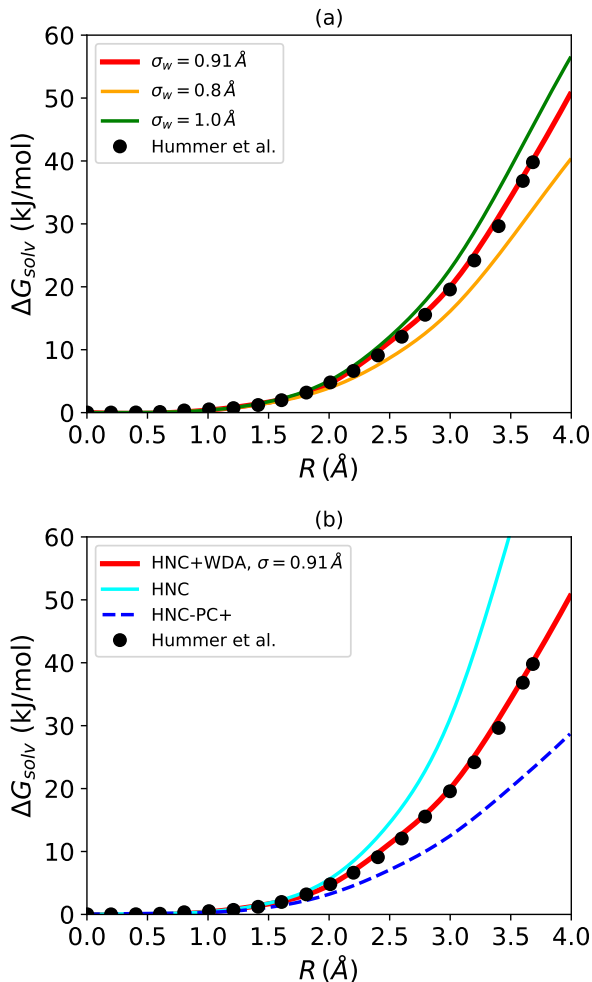


Figure 2. Hydration free energy of a hard sphere as function of HS radius R . The MDFT results are compared to the Monte-Carlo results of Hummer *et al.*⁷⁹: (a) For the HNC+WDA functional and different values of the coarse-graining parameter σ_w in eqs 5-7. (b) According to the different approximations described in the text.

In agreement with our previous studies for Lennard-Jones particles or assembly of Lennard-Jones particles, the HNC-PC+ a-posteriori correction largely improves it, although still underestimating the simulation results for small σ 's and doomed to overestimate them at the nanoscale. Similar conclusions concerning HNC or its close Kovalenko-Hirata (KH) variant were also reached by Fujita *et al.* using 3D-RISM. They never found such agreement as the one that we obtain with HNC+WDA with any of the numerous versions of 3D-RISM that they tried.

As far as structure is concerned, we compare in the top panels of Fig. 7 the radial distribution functions obtained either by MDFT or MD simulations for two microscopic spherical solutes of different size, namely a one-site united-atom representation of methane ($\sigma = 3.73 \text{ \AA}$, $\epsilon = 1.23 \text{ kJ/mol}$) and of neopentane, $\text{C}(\text{CH}_3)_4$ ($\sigma = 6.15 \text{ \AA}$, $\epsilon = 3.49 \text{ kJ/mol}$), immersed in SPC/E water. The HNC approximation tends to systematically overestimate the first and second peaks and slightly displace them to lower values. This

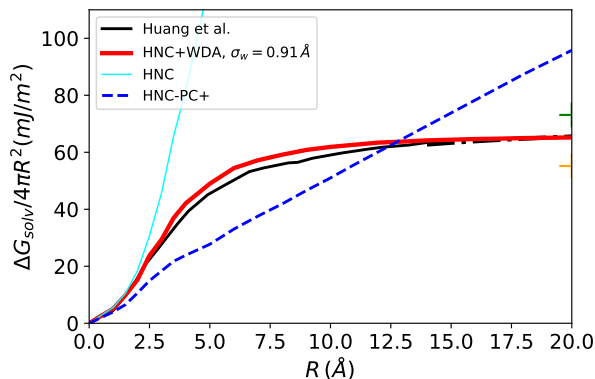


Figure 3. MDFT hydration free energy per surface area as function of HS radius compared to the Monte-Carlo results of Huang *et al.*⁷⁷ according to the different approximation described in the text. The green and orange indicators appearing on the right y-axis give the plateau values that would have been obtained for $\sigma_w = 0.8 \text{ \AA}$ or $\sigma_w = 1.0 \text{ \AA}$, respectively. See the corresponding curves in Fig. 2-a.

is nicely corrected by the WDA bridge, with no visual difference between MDFT and simulation for methane and a little one for neopentane. In the bottom panels, we consider a 5-site model of CH_4 described by the GAFF force field, thus including partial atomic charges, in TIP3P water. The deficiencies of HNC for both the C-O_w and H-O_w rdf's, especially in the first peak, are corrected too, with only a small difference appearing in the first small peak/shoulder in the H-O_w rdf. Here the calculations were performed with a value $m_{\text{max}} = 5$, which gives slightly better converged results than $m_{\text{max}} = 3$.

C. Hydrophobic pairs

We now turn to so-called hydrophobic interactions, namely the effective interaction between two hydrophobic entities in water, which is known to play a key role in many phenomena, for example in biomolecular recognition. We leave the discussion of the simplest paradigmatic case, the potential of mean force (PMF) between two methanes, to the Supplementary Material. Rather, we borrow from Fujita *et al.*⁸³ a slightly more complex test system that they have explored using both 3D-RISM and simulation⁸³. It consists in a pair of coronene molecules which are separated by a varying distance R while keeping the plane of each molecule orthogonal to the R -direction. In Fig. 8 is plotted the solvent-induced PMF, computed for a center-center distance R as $\Delta G_{\text{solv}}(R) - \Delta G_{\text{solv}}(\infty)$. It is seen that HNC+WDA reproduces correctly the overall behaviour of the simulation curve. HNC is improved by the pressure correction but the agreement remains far below that of HNC+WDA. In particular, the solvent-induced free-energy barrier for association, i.e the barrier encountered when going from the solvent-separated to the contact pair, is essentially correct for HNC+WDA whereas it is overestimated by $\sim 90 \text{ kJ/mol}$ by HNC and $\sim 20 \text{ kJ/mol}$ by HNC-PC+.

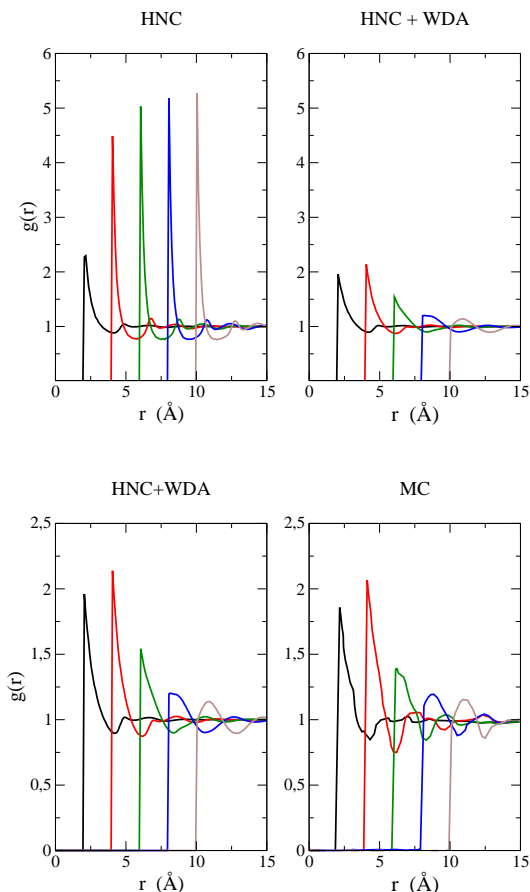


Figure 4. Radial solute-water pair distribution function for values of $R = 2, 4, 6, 8, 10$ Å (black, red, green, blue and brown lines, respectively). The top panels compare the HNC and HNC+WDA approximations whereas the bottom ones confronts HNC+WDA to the Monte-Carlo results of Huang *et al.*^{77,78}. The top-right and bottom-left panels are redundant but presented at different scales for comparison.

D. Hydrophobic molecules dataset

As in Ref.⁶⁰, we devote this section to hydrophobic solvation by testing the predictive power of the present HNC-WDA functional for the hydration free energies of an extended dataset of hydrophobic solutes. To this end, we consider the calculation of the non-electrostatic contribution to the hydration free-energy of drug-like molecules, a quantity that is systematically computed in free-energy perturbation methods (growing slowly the Lennard-Jones potential before introducing the charges), and that is a main source of error when using MDFT-HNC. We refer to Mobley’s FreeSolv database, containing 642 drug-like molecules for which the experimental solvation free-energies, and more importantly here, the solvation free-energies computed by simulation and decomposed into their non-electrostatic and electrostatic contributions, are available^{84,85}. We have generated ourselves our own reference data with the same force-field parameters and starting solute geometries as in Mobley’s FEP calculations, but

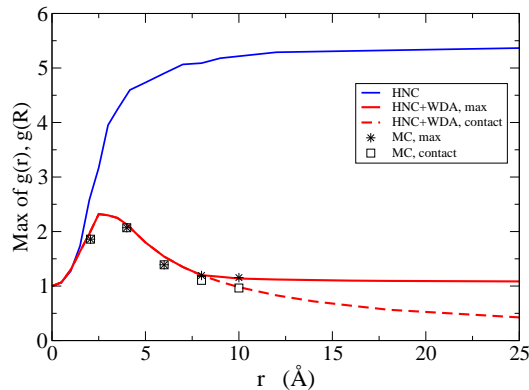


Figure 5. Maximum value of the hard-sphere/water-oxygen pair distribution function $g(r)$, and its value at contact $g(R)$, as a function of HS radius R . The corresponding MC results of Huang *et al.*^{77,78} for radii up to 10 Å are given by the star and square symbols, respectively.

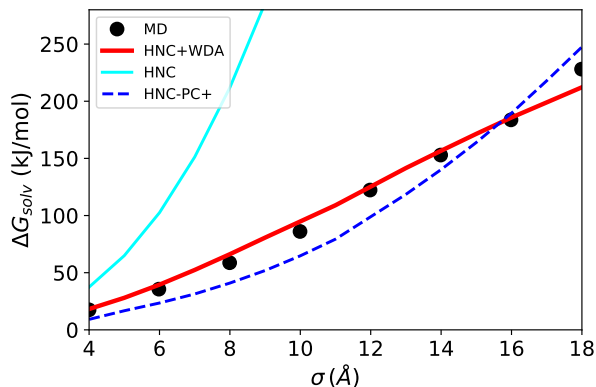


Figure 6. Hydration free energy of a spherical Lennard-Jones particle of increasing diameter σ . The dots represent the MD simulation results of Fujita *et al.*⁸³ whereas the lines correspond to different MDFT approximations explained in the text.

keeping the solute geometries fixed in order to compare directly with MDFT. This was done using the H4D-MC method and the associated code^{74,75}. We are thus testing our approach on a dataset of virtual hydrophobic molecular solutes with a large variety of shapes and sizes. In Fig. 9 we compare the MC results obtained for the non-electrostatic contribution to the solvation free-energy (all partial charges put to zero) to the same quantities obtained with MDFT in its present HNC+WDA version. In Ref.⁶⁰ this was done using a fully angular-independent functional, *i.e.*, using an angular-independent HNC reference inferred from the water structure factor, complemented by a WDA bridge functional similar to that defined in eqs 5-8. Scanning the whole ‘hydrophobic’ Mobley database we found in that study a mean absolute error (MAE) of $0.35 k_B T$ between the MDFT and Monte-Carlo hydration free energies. As it appears in Fig. 9, this error is now decreased to $0.25 k_B T$ when using the full angular-dependent HNC ref-

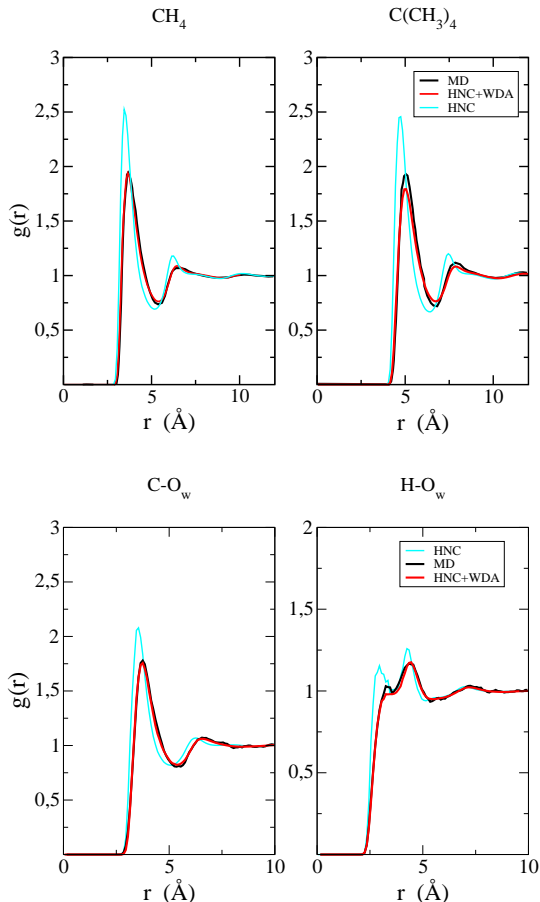


Figure 7. Water structure around simple hydrophobic solutes. Top panels: a one-site unified-atom representation of methane and neopentane in SPC/E water. Bottom panels: a five-site model of methane in TIP3P water, with the C-O_w and H-O_w site-site pair distribution functions represented.

erence. The mean slope is very close to unity. This is much better than that was obtained previously using the HNC approximation and the best pressure correction, equivalent to HNC-PC+⁵⁹. That yielded a mean slope of 0.55 and a MAE of $0.7 k_B T$; see the top panel of Fig. 9. Note that an isolated, utmost outlier appears around $+20 \text{ kJ/mol}$ in the HNC+WDA results; it concerns the ethion molecule, composed of 41 atoms, the largest molecule in the database, that is curiously better described by HNC-PC+.

We conclude here our application to hydrophobic solvation, although many other features remain to be studied within our MDFT framework: the entropy-enthalpy contributions to the solvation free energy and the temperature effects, the cross-over between the so-called enthalpic and entropic solvation regimes, the effective interaction between nanoscale hydrophobic objects.... We reserve this to future work. For now, we have already in hand many clues indicating that our HNC+WDA version of MDFT is able to capture the main physical features of hydrophobic solvation at a qualitative and even quantitative level. How does this extend to hydrophilic solvation? This is the matter of the next section.

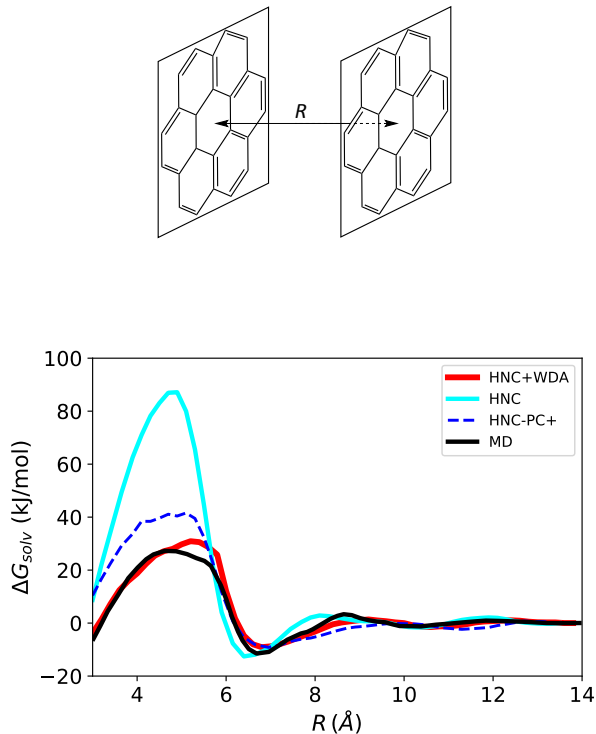


Figure 8. Top: Representation of two interacting coronene molecules distant by R and kept perpendicular to the R -direction. Bottom: Solvent-induced PMF as function of the center-to-center distance R . The MD results are taken from Ref.⁸³.

IV. HYDROPHOBIC TO HYDROPHILIC SOLVATION

A. Water as a solute

As the paradigm for a hydrophilic solute, we begin with the water molecule itself, *i.e.*, we consider a water molecule immersed in the same water model. Immediately appears a limitation of the HNC+WDA functional developed above: the computed solvation free energy turns out to be -16.2 and -15.8 kJ/mol for SPC/E and TIP3P, respectively, whereas our reference Monte-Carlo values are -29.5 and -23.3 kJ/mol . There is a clear underestimation of the electrostatic contribution to the free energy. To correct that deficiency, we propose the simplest conceivable patch correcting the theory, which is to slightly reinforce the solvent electrostatic potential through an additional bridge term that has the form

$$\mathcal{F}_B^{\text{elec}}[\rho(\mathbf{r}, \omega)] = \int d\mathbf{r} d\mathbf{r}' \frac{\rho_{\text{solute}}(\mathbf{r}) \lambda \rho_w(\mathbf{r}')}{4\pi\epsilon_0 |\mathbf{r} - \mathbf{r}'|}, \quad (13)$$

where $\rho_{\text{solute}}(\mathbf{r})$ denotes the solute charge density and $\rho_w(\mathbf{r})$ the solvent charge density defined by

$$\rho_w(\mathbf{r}) = \int d\mathbf{r}' d\omega \sigma(\mathbf{r} - \mathbf{r}', \omega) \rho(\mathbf{r}', \omega) \quad (14)$$

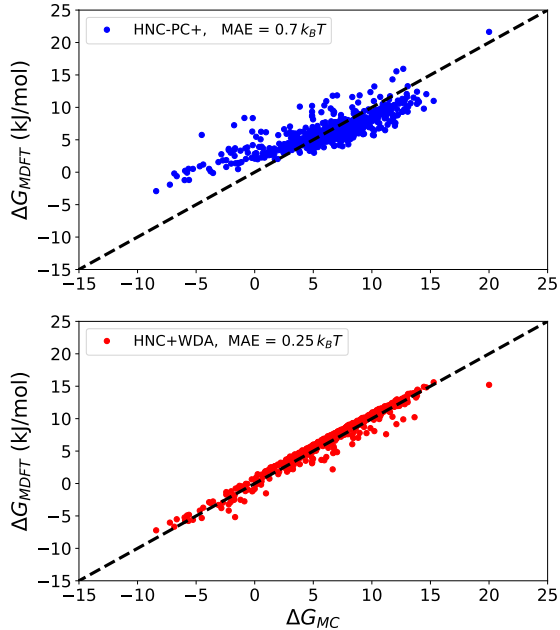


Figure 9. Correlation between the MDFT and Monte-Carlo solvation free energies for FreeSolv dataset of 642 drug-like organic molecules with all partial charges put to zero. Top: MDFT in the HNC-PC+ approximation (results are taken from Ref.⁵⁹). Bottom: Using the HNC+WDA functional developed in this work. The mean absolute errors (MAE) are indicated.

$\sigma(\mathbf{r}, \omega)$ being the charge density of a single water molecule located at origin with orientation ω . The additional bridge term appears thus as a functional of the full angular density $\rho(\mathbf{r}, \omega)$. The parameter λ is fixed by requiring the correct solvation free energy for water in water, *i.e.*, the correct excess chemical potential of the bulk solvent. This condition yields $\lambda = 0.18$ for SPC/E, and 0.14 for TIP3P. The new term amounts to slightly renormalize (increase) the solvent charges seen by the solute while keeping the water-water interactions unchanged. From now on, the modified functional including this new electrostatic term on top of WDA bridge of eqs 5-8 will be referred to as HNC+WDA*. How does it affect the liquid structure? It can be seen in Fig. 10 that the effect of the new electrostatic term is not spectacular and does not solve completely the known limitations of HNC⁵⁹: a first peak of $g_{OO}(r)$ that is too broad, a second one that is misplaced, a first peak of $g_{HO}(r)$ that is underestimated. It can be seen, however, that HNC+WDA* slightly improves both the HNC and HNC+WDA results, with for example a first peak of $g_{OO}(r)$ that is narrower and better located.

B. Test sets of organic molecule

At this point, it has to be checked how the HNC+WDA* functional, somewhat phenomenological but parametrized

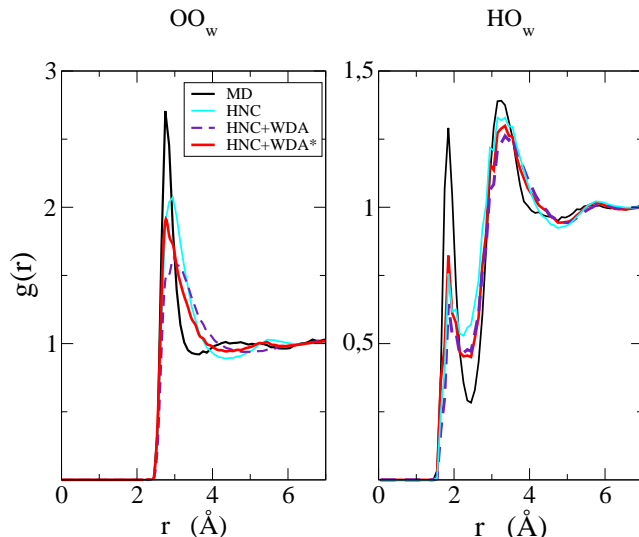


Figure 10. OO_w and HO_w radial distribution functions for a TIP3P molecule in TIP3P water (subscript w). The different approximations of MDFT described in the text are compared to simulation results.

on the sole properties of water, extends to the solubility prediction of other hydrophilic molecules. To make a first quick assessment, we have constructed a reduced dataset of 19 test organic molecules, including the standard dataset constituted by 14 amino-acid side-chain analogs, quinoline (see below), plus 4 special molecules picked in the Mobley’s dataset (the two molecules having the smallest, negative ‘hydrophobic’ SFE, and the two having the largest positive ones. This includes the largest molecule in the dataset, ethion). See the SM for a more detailed description of the test set. Five of them, including part of the amino-acid analogs, turn out to be hydrophobic while eleven solutes can be referred to as hydrophilic with markedly negative SFE’s. Compared to reference simulation results obtained with the H4D-MC method^{74,75}, it is seen in Fig. 11 that the trend of the SFE’s when going from hydrophobic to largely hydrophilic solvation is reproduced quantitatively, with a mean absolute error of 0.75 $k_B T$ over this rather reduced dataset. The highest discrepancy is observed for molecule 19, namely Ethion, that was already pointed out as an outlier for the hydrophobic contribution to the SFE in Fig. 9.

As for the solvent microscopic structure, it seems illusory to test visually tens of molecules. In Fig. 12 we have selected as a typical example the quinoline molecule, already studied in Ref.⁵⁹. Looking at the different solute-site/water radial distribution functions, we can claim that the agreement between MDFT and simulation is quite good, and that HNC+WDA* does improve the structure with respect to straight HNC –the HNC+WDA results (not shown) are rather close to the HNC+WDA* ones. As can be anticipated, the largest discrepancy appears for the nitrogen atom, wearing a strong negative charge, for which the H-bond structuring effect is better accounted for compared to

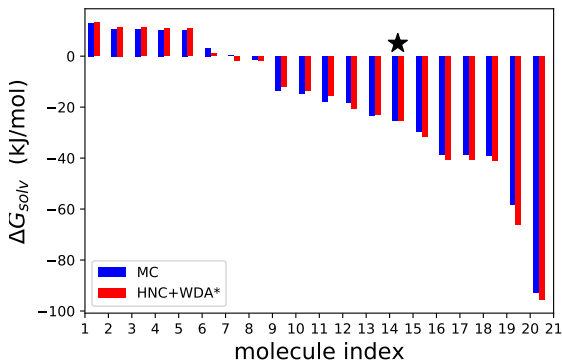


Figure 11. Hydration free energies of a reduced test set of 19 organic molecules in addition to water (denoted by the black star). The MDFT results, obtained with a HNC+WDA* functional adjusted on the sole properties of water, are compared to simulation results obtained with the H4D-MC method^{74,75}. The molecules are numbered from 1 to 20 and are sorted by decreasing solvation free energies (see the Supplementary Material).

HNC, but appears still underestimated.

We conclude this section devoted to neutral molecules by now spanning the FreeSolv database, *i.e.*, the whole 642 drug-like molecules wearing their full partial charges^{84,85}. The results concern in fact 635 molecules, since 7 out of the 642 (1 % of the database) lead to convergence problems. In Fig. 13, we show the correlation between the HNC+WDA* hydration free energies and those that we have obtained through the H4D-Monte-Carlo method for the same rigid molecule geometries. We find a slope very close to 1, and a mean absolute error of $0.45 k_B T$ (~ 1 kJ/mol) over the whole database, thus far below 4 kJ/mol that is considered as chemical accuracy for such thermochemical quantities.

C. Charged species

We have shown in the previous sections that with the appropriate bridges MDFT is able to quantitatively predict the solvation properties of hydrophobic and hydrophilic neutral molecules, we now turn to ionic species. In Fig. 14, we display the water structure obtained with MDFT-HNC and using the WDA* bridge for a typical monovalent cation, K^+ , and a typical monovalent anion, Cl^- . There we plot either the reduced number density, equivalent to the rdf, $g(r) = n(r)/n_b$, or the reduced polarisation density, defined here as $P(\mathbf{r}) = |\int d\omega \hat{\omega} \rho(\mathbf{r}, \omega)/n_b|$ where, for every orientation ω , $\hat{\omega}$ represent the unitary vector along the dipole direction. Again, compared to MC results, HNC+WDA* does slightly improve the HNC predictions, in particular increasing the first peak height which is a little bit too small in HNC. Overall the structure around the cation is very satisfactorily reproduced, be it in terms of number density or polarisation density. This is less so for the anion, for which a shift of the second peak occurs, reminiscent to that observed for water. Overall, the structural differences between HNC+WDA*, HNC+WDA (not represented here)

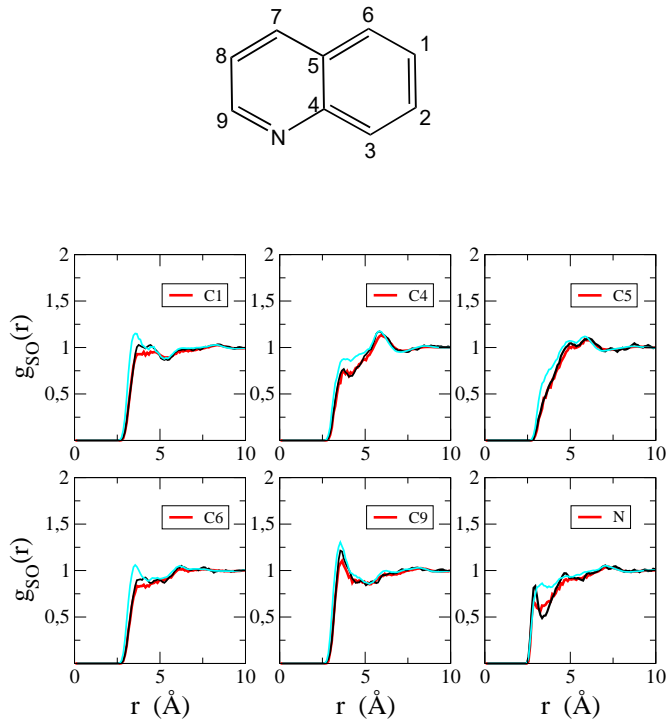


Figure 12. Top: Picture of the quinoline molecule with its carbon atoms numbered. Bottom: site-water oxygen radial distribution functions for a selection of different atomic sites of quinoline. The color code is the same as in Fig. 10: the simulation curves are in black; the MDFT ones within the HNC or HNC+WDA* are in cyan and red, respectively.

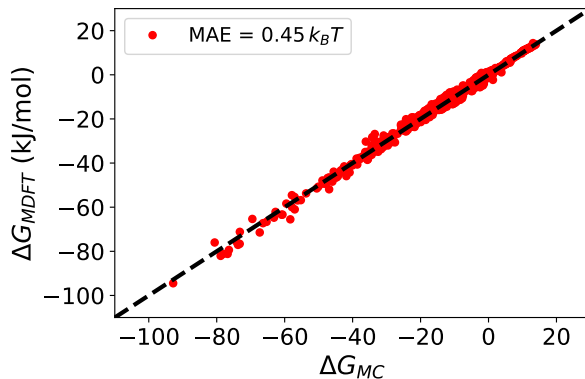


Figure 13. Correlation between the MDFT and Monte-Carlo hydration free energies for the FreeSolv dataset of drug-like organic molecules. The mean absolute error is indicated.

and HNC are relatively small. This is not the case for the solvation free energies, as can be seen in Table I. There we report our previous results obtained with HNC and its a-posteriori PC+ correction (which has here a minor influence). Here we find that HNC+WDA underestimate in absolute value the SFE's of both the cation and the anion, and more so for the anion. On the other hand HNC+WDA* accounting for the electrostatic bridge correction of eq. 13 largely overestimates them. The λ -factor selected to yield

	MC	HNC	HNC+WDA	HNC+WDA*	HNC-PC+
Cl ⁻	-360	-317	-328	-403	-317
K ⁺	-315	-293	-294	-384	-304

Table I. Absolute hydration free energies in kJ/mol obtained for different approximations of MDFT and compared to Monte-Carlo results.

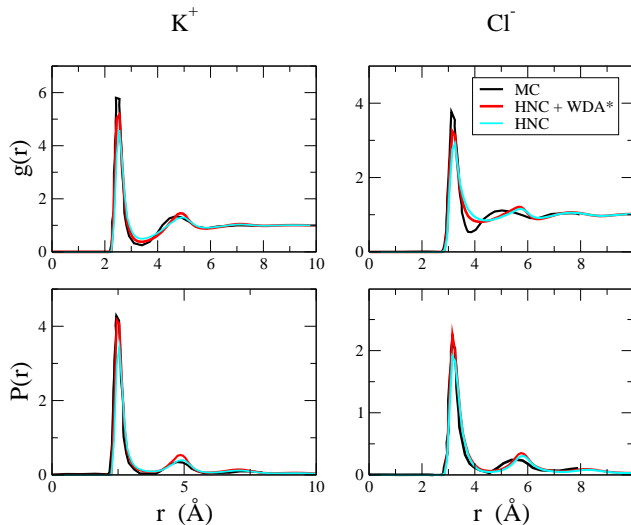


Figure 14. Water structure around a cation and an anion computed within different MDFT approximations and compared to MC simulation results. The top panels display the ion-water radial distribution function, $g(r) = n(r)/n_b$, whereas the bottom ones display the reduced polarization density, as defined in the text.

the correct chemical potential of water, and further applied to arbitrary neutral molecules, obviously does not extend to ionic species. If any λ -factor to be defined, it should take a smaller value for charged species, and furthermore a different one for negative or positive charges. Despite the fair amount of successes that we have reported so far for both hydrophobic and hydrophilic solvation, we reach here the caveats and the limits of the simple bridge functional approach developed conscientiously in the paper.

V. CONCLUSIONS AND PERSPECTIVE

In a previous work we have explored the capability of molecular density functional theory at its lowest, second-order, theoretical level, namely in the HNC approximation, to predict the hydration free-energies of molecular solutes and the surrounding microscopic solvent structure.⁵⁹ In the present work, following Ref.⁶⁰, we have proposed a simple, third-order, angular-independent bridge functionals, complementing the fully angular-dependent HNC functional, that was shown to be extremely powerful to describe the solvation of hydrophobic entities, avoiding the need of any a-posteriori corrections. This lead to the further defini-

tion of a simple angular-dependent correction that extends the predictability of the theory to hydrophobic solvation. Among many positive features described in the paper, managing to reproduce the simulation solvation free-energies for the whole FreeSolv database with an accuracy of half a $k_B T$, below what can be considered as chemical accuracy, moreover with a functional adjusted only on bulk liquid properties and not on any part of the database to be studied, is a success. We thus consider that, with the present version, the molecular density functional theory of water has reached a sufficient degree of maturation and predictability to be diffused and applied to relevant chemical problems where water plays a key role, for example hydration effects in the structure and function of biomolecules, molecular recognition, or in QM/MM approaches to describe explicit water around quantum mechanical solutes⁸⁶⁻⁸⁸. Let us also recall that, even in its crudest HNC approximation, MDFT and associated liquid-state theories are known to work much better for non-associated liquids^{42,46}, *i.e.*, many of the non-polar or polar aprotic solvents that are used in chemistry. As often, despite the extremely simple chemical structure of its components, and their simple point charge representation in models such as SPCE/E or TIP3P, the description of liquid water is the problem.

Indeed, the work presented in this paper is not the end of the story of MDFT for water. It does not make it possible to describe ionic species consistently. Furthermore, the simple angular-dependent, additional bridge functional proposed in eq. 13, although solving part of the problem for hydrophilic solvation, remains ad-hoc and marginally improves the description of the H-bonding water structure. The ultimate bridge functional for water should be sought from the start as an angular-dependent functional, as evidenced by a previous study for neat water⁶⁴. A positive indication emerging from this study, in particular the form of the WDA functional that was chosen in eqs 5-8, is that a systematic third-order theory, thus involving a reasonable approximation of the three-body direct correlation function $c^{(3)}(1, 2, 3)$ in addition to the two-body one contained in the HNC functional, might prove sufficient. This is a direction for future improvements.

Acknowledgement

This work was supported by the Agence Nationale de la Recherche, projet ANR BRIDGE AAP CE29.

Availability of data

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

- ¹B. ROUX and T. SIMONSON, *Biophys. Chem.* **78**, 1 (1999).
- ²J. MONGAN, C. SIMMERLING, J. A. MCCAMMON, D. A. CASE, and A. ONUFRIEV, *J. Chem. Theory Comput.* **3**, 156 (2007).
- ³J. DZUBIELLA, J. M. J. SWANSON, and J. A. MCCAMMON, *Phys. Rev. Lett.* **96**, 1 (2006).
- ⁴R. W. ZWANZIG, *J. Chem. Phys.* **22**, 1420 (1954).
- ⁵C. H. BENNETT, *J. Phys. Chem. Lett.* **22**, 245 (1976).
- ⁶S. KUMAR, J. M. ROSENBERG, D. BOUZIDA, R. H. SWENDSEN, and P. A. KOLLMAN, *J. Comp. Chem.* **13**, 1011 (1992).
- ⁷M. R. SHIRTS and J. D. CHODERA, *J. Chem. Phys.* **129**, 124105 (2008).
- ⁸N. MATUBAYASI and M. NAKAHARA, *J. Chem. Phys.* **117**, 3605 (2002).
- ⁹N. MATUBAYASI and M. NAKAHARA, *J. Chem. Phys.* **119**, 9686 (2003).
- ¹⁰H. S. ASHBAUGH and L. R. PRATT, *Rev. Mod. Phys.* **78**, 159 (2006).
- ¹¹R. ROTH, *J. Phys. Condens. Matter* **22** (2010).
- ¹²R. ROTH, Y. HARANO, and M. KINOSHITA, *Phys. Rev. Lett.* **97**, 1 (2006).
- ¹³S. CHIBA, Y. HARANO, R. ROTH, M. KINOSHITA, and M. SAKURAI, *J. Comput. Chem.* **33**, 550 (2012).
- ¹⁴Y. HARANO, R. ROTH, and S. CHIBA, *J. Comput. Chem.* **34**, 1969 (2013).
- ¹⁵P. KOEHL and M. DELARUE, *J. Chem. Phys.* **132**, 064101 (2010).
- ¹⁶R. RAMIREZ, R. GEBAUER, M. MARESCHAL, and D. BORGIS, *Physical Review E* **66**, 031206 (2002).
- ¹⁷R. RAMIREZ and D. BORGIS, *J. Phys. Chem. B* **109**, 6754 (2005).
- ¹⁸F. HIRATA and P. J. ROSSKY, *Chem. Phys. Lett.* **83**, 329 (1981).
- ¹⁹D. BEGLOV and B. ROUX, *J. Phys. Chem. B* **101**, 7821 (1997).
- ²⁰A. KOVALENKO and F. HIRATA, *Chem. Phys. Lett.* **290**, 237 (1998).
- ²¹A. KOVALENKO and F. HIRATA, *J. Chem. Phys.* **113**, 2793 (2000).
- ²²A. KOVALENKO and F. HIRATA, *J. Phys. Chem. B* **103**, 7942 (1999).
- ²³Y. MARUYAMA, N. YOSHIDA, H. TADANO, D. TAKAHASHI, M. SATO, and F. HIRATA, *J. Comput. Chem.* **35**, 1347 (2014).
- ²⁴F. HOFFGAARD, J. HEIL, and S. M. KAST, *J. Chem. Theory Comput.* **9**, 4718 (2013).
- ²⁵J. HEIL and S. M. KAST, *J. Chem. Phys.* **142** (2015).
- ²⁶Y. LIU, S. ZHAO, and J. WU, *J. Chem. Theory Comput.* **9**, 1896 (2013).
- ²⁷Y. LIU, J. FU, and J. WU, *J. Phys. Chem. Lett.* **4**, 3687 (2013).
- ²⁸G. N. CHUEV, M. V. FEDOROV, and J. CRAIN, *Chem. Phys. Lett.* **448**, 198 (2007).
- ²⁹D. S. PALMER, A. I. FROLOV, E. L. RATKOVA, and M. V. FEDOROV, *Molecular Pharmaceutics* **8**, 1423 (2011).
- ³⁰J. F. TRUCHON, B. M. PETTITT, and P. LABUTE, *J. Chem. Theory Comput.* **10**, 934 (2014).
- ³¹M. MISIN, D. S. PALMER, and M. V. FEDOROV, *J. Phys. Chem. B* **120**, 5724 (2016).
- ³²T. LUCHKO, N. BLINOV, G. C. LIMON, K. P. JOYCE, and A. KOVALENKO, *J. Comput. Aided. Mol. Des.* **30**, 1115 (2016).
- ³³J. JOHNSON, D. A. CASE, T. YAMAZAKI, and T. LUCHKO, *J. Phys. Condens. Matter* **28**, 344002 (2016).
- ³⁴D. ROY and A. KOVALENKO, *J. Phys. Chem. A* **123**, 4087 (2019).
- ³⁵N. TIELKER, D. TOMAZIC, L. EBERLEIN, S. GÜSSREGEN, and S. M. KAST, *J. Comput. Aided. Mol. Des.* **34**, 453 (2020).
- ³⁶N. YOSHIDA, T. IMAI, S. PHONGPHANPHANEE, A. KOVALENKO, and F. HIRATA, *J. Phys. Chem. B* **113**, 873 (2009).
- ³⁷T. LUCHKO, S. GUSAROV, D. R. ROE, C. SIMMERLING, A. DAVID, J. TUSZYNSKI, and A. KOVALENKO, *J. Chem. Theory Comput.* **6**, 607 (2011).
- ³⁸M. C. STUMPE, N. BLINOV, D. WISHART, A. KOVALENKO, and V. S. PANDE, *J. Phys. Chem. B* **115**, 319 (2011).
- ³⁹J. JOHNSON, D. A. CASE, T. YAMAZAKI, S. GUSAROV, A. KOVALENKO, and T. LUCHKO, *J. Phys. Condens. Matter* **28** (2016).
- ⁴⁰C. NGUYEN, T. YAMAZAKI, A. KOVALENKO, D. A. CASE, M. K. GILSON, T. KURTZMAN, and T. LUCHKO, *PLoS One* **14**, 1 (2019).
- ⁴¹G. M. GIAMBAŞU, D. A. CASE, and D. M. YORK, *J. Am. Chem. Soc.* (2019).
- ⁴²J. RICHARDI, C. MILLOT, and P. H. FRIES, *J. Chem. Phys.* **110**, 1138 (1999).
- ⁴³M. LOMBARDEO, C. MARTÍN, S. JORGE, F. LADO, and E. LOMBA, *J. Chem. Phys.* **110**, 1148 (1999).
- ⁴⁴L. BELLONI and I. CHIKINA, *Mol. Phys.* **112**, 1246 (2014).
- ⁴⁵L. GENDRE, R. RAMIREZ, and D. BORGIS, *Chem. Phys. Lett.* **474**, 366 (2009).
- ⁴⁶D. BORGIS, L. GENDRE, and R. RAMIREZ, *J. Phys. Chem. B* **116**, 2504 (2012).
- ⁴⁷S. ZHAO, R. RAMIREZ, R. VUILLEUMIER, and D. BORGIS, *J. Chem. Phys.* **134**, 194102 (2011).
- ⁴⁸G. JEANMAIRET, M. LEVESQUE, R. VUILLEUMIER, and D. BORGIS, *J. Phys. Chem. Lett.* **4**, 619 (2013).
- ⁴⁹L. BLUM and A. J. TORRUELLA, *J. Chem. Phys.* **56**, 303 (1972).
- ⁵⁰L. BLUM, *J. Chem. Phys.* **57**, 1862 (1972).
- ⁵¹L. DING, M. LEVESQUE, D. BORGIS, and L. BELLONI, *J. Chem. Phys.* **147**, 094107 (2017).
- ⁵²M. LEVESQUE, R. VUILLEUMIER, and D. BORGIS, *J. Chem. Phys.* **137**, 034115 (2012).
- ⁵³G. JEANMAIRET, M. LEVESQUE, and D. BORGIS, *J. Chem. Phys.* **139**, 154101 (2013).
- ⁵⁴G. JEANMAIRET, M. LEVESQUE, V. SERGHEVSKIY, and D. BORGIS, *J. Chem. Phys.* **142**, 154112 (2015).
- ⁵⁵C. GAGEAT, L. BELLONI, D. BORGIS, and M. LEVESQUE, (2017), arXiv: 1709.10139.
- ⁵⁶S. ZHAO, Z. JIN, and J. WU, *J. Phys. Chem. B* **115**, 6971 (2011).
- ⁵⁷S. ZHAO, Z. JIN, and J. WU, *J. Phys. Chem. B* **115**, 15445 (2011).
- ⁵⁸Y. LIU, S. ZHAO, and J. WU, *J. Chem. Theory Comput.* **9**, 1896 (2013).
- ⁵⁹S. LUUKKONEN, M. LEVESQUE, L. BELLONI, and D. BORGIS, *J. Chem. Phys.* **152**, 064110 (2020).
- ⁶⁰D. BORGIS, S. LUUKKONEN, L. BELLONI, and G. JEANMAIRET, *J. Phys. Chem. B* **124**, 6885 (2020).
- ⁶¹R. EVANS, *Advances in Physics* **28**, 143 (1979).
- ⁶²W. L. JORGENSEN, J. CHANDRASEKHAR, J. D. MADURA, and M. L. KLEIN, *J. Chem. Phys.* **79**, 926 (1983).
- ⁶³H. J. C. BERENDSEN, J. R. GRIGERA, and T. P. STRAATSMA, *J. Phys. Chem.* **91**, 6269 (1987).
- ⁶⁴L. BELLONI, *J. Chem. Phys.* **147**, 164121 (2017).
- ⁶⁵Y. ROSENFELD, *Phys. Rev. Lett.* **63**, 980 (1989).
- ⁶⁶E. KIERLIK and M. L. ROSINBERG, *Phys. Rev. A* **42**, 3382 (1990).
- ⁶⁷R. ROTH, R. EVANS, A. LANG, and G. KAHL, *J. Phys. Condens. Matter* **14**, 12063 (2002).
- ⁶⁸Y.-X. YU and J. WU, *J. Chem. Phys.* **117**, 10156 (2002).
- ⁶⁹Y. ROSENFELD, *J. Chem. Phys.* **98**, 8126 (1993).
- ⁷⁰S. NORDHOLM, M. JOHNSON, and B. C. FREASIER, *Australian J. Chem.* **33**, 2139 (1980).
- ⁷¹P. TARAZONA, *Phys. Rev. A* **31**, 2672 (1985).
- ⁷²W. A. CURTIN and N. W. ASHCROFT, *Phys. Rev. A* **32**, 2909 (1985).
- ⁷³J.-P. HANSEN and I. R. McDONALD, *Theory of Simple Liquids: With Applications to Soft Matter*, Academic Press, Amsterdam, 4th edition, 2013.
- ⁷⁴L. BELLONI, *J. Chem. Phys.* **151**, 021101 (2019).
- ⁷⁵S. LUUKKONEN, *Hydration of drug-like molecules with molecular density functional theory and the hybrid-4th-dimension Monte-Carlo approach*, PhD thesis, Université Paris-Saclay, 2020.
- ⁷⁶F. M. FLORIS, M. SELMI, A. TANI, and J. TOMASI, *J. Chem. Phys.* **107** (1997).
- ⁷⁷D. M. HUANG, P. L. GEISSLER, and D. CHANDLER, *J. Phys. Chem. B* **105**, 6704 (2001).
- ⁷⁸D. M. HUANG and D. CHANDLER, *J. Phys. Chem. B* **106**, 2047 (2002).
- ⁷⁹G. HUMMER, S. GARDE, A. E. GARCÍA, A. POHORILLE, and L. R. PRATT, *Proc. Nat. Acad. Sci. USA* **93**, 8951 (1996).
- ⁸⁰V. P. SERGHEVSKIY, G. JEANMAIRET, M. LEVESQUE, and D. BORGIS, *J. Phys. Chem. Lett.* **5**, 1935 (2014).
- ⁸¹V. SERGHEVSKIY, G. JEANMAIRET, M. LEVESQUE, and D. BORGIS, *J. Chem. Phys.* **143**, 184116 (2015).
- ⁸²T. Y. HSU and G. JEANMAIRET, *J. Chem. Phys.* **154** (2021).
- ⁸³T. FUJITA and T. YAMAMOTO, *J. Chem. Phys.* **147**, 014110 (2017).
- ⁸⁴D. L. MOBLEY, C. I. BAYLY, M. D. COOPER, M. R. SHIRTS, and K. A. DILL, *J. Chem. Theory Comput.* **5**, 350 (2009).

- ⁸⁵G. D. R. MATOS, D. Y. KYU, H. H. LOEFFLER, J. D. C., M. R. SHIRTS, and D. L. MOBLEY, *J. Chem. Eng. Data* **62**, 1559 (2017).
- ⁸⁶G. JEANMAIRET, M. LEVESQUE, and D. BORGIS, *J. Chem. Theory Comput.* **16**, 7123 (2020).
- ⁸⁷T. KLOSS, J. HEIL, and S. M. KAST, *J. Phys. Chem. B* **112**, 4337 (2008).
- ⁸⁸H. SATO and S. SAKAKI, *J. Phys. Chem. A* **108**, 1629 (2004).