JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

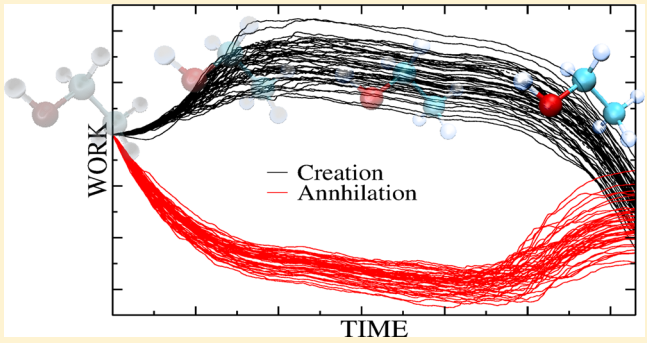
Fast Switching Alchemical Transformations in Molecular Dynamics Simulations

Piero Procacci* and Chiara Cardelli

Dipartimento di Chimica, Università di Firenze, Via della Lastruccia 3, Sesto Fiorentino, Florence I-50019, Italy

Supporting Information

ABSTRACT: We present an efficient and rigorous implementation of the fast switching alchemical transformation for systems where electrostatic interactions are treated using the smooth particle mesh Ewald method. Free energies are computed using bidirectional nonequilibrium alchemical trajectories by applying the Crooks fluctuation theorem and the Bennett acceptance ratio to the collection of the final alchemical works. The technique is used for the evaluation of the 1-octanol/water partition coefficients for some selected organic molecules. Fast switching alchemical tranformations appear to be competitive, both in accuracy and in efficiency, with respect to the traditional methods based on independent equilibrium simulations of intermediate states.



INTRODUCTION

Alchemical transformations,[1−6] whereby a chemical species is transformed into another via a pathway of nonphysical (alchemical) states, play a prominent role in modern computational chemistry. These methods provide in fact a computational viable tool to access to the free energy differences between well-defined chemical states. Many physical processes, such as ligand binding in complex biological system or transfer of a molecule from one phase to an other, can be equivalently expressed as a composition of alchemical transformations and hence for each of them the associated free energy change can be determined. In the context of computational drug discovery for example, modern approaches in ligand binding affinities evaluation proceed by constructing ad hoc thermodynamic cycles involving the reversible alchemical decoupling of the ligand with its surroundings.[7−9]

Alchemical transformations are generally implemented using equilibrium methods on selected alchemical intermediate states between the end target states whose free energy difference is at stake. These methods include all common free energy methods, that is, free energy perturbation schemes (FEP),[10] thermodynamic integration[11] and the multiple Bennett acceptance ratio (MBAR).[12−15] The equilibrium approach with MBAR has nowadays become a standard in alchemical simulations after its implementation in the popular GROMACS package by Shirts and co-workers[16,17] and in the NAMD program by Chipot and co-workers.[4,18]

The technique was originally invented by Jorgensen and Ravimohan[1] and essentially is based on the set up of independent equilibrium simulations of several nonphysical intermediate states defined with respect to a λ coordinate appropriately coupled to the potential energy function and such that for the values $\lambda = 1$ and $\lambda = 0$ one recovers the two target end states for which the free energy difference has to be determined. The λ coordinate involves a linearly turning off of the electrostatic and van der Waals interactions of the alchemical species with the remainder of the system, that is, the nonalchemical "solvent". In this first stage, the system configurations of each intermediate alchemical state are saved and, in the following step, the saved data are postprocessed using the MBAR analysis. The postprocessing step requires the calculation (usually with single point (SP) energy determination) of the distribution of the potential energy differences between contiguous intermediate states, i.e. $\delta(E - (V_{\lambda_{k+1}}(x(\lambda_k)) - V_{\lambda_k}(x(\lambda_k))))$ and $\delta(E - (V_{\lambda_k}(x(\lambda_k)) - V_{\lambda_{k-1}}(x(\lambda_k))))$. From these potential energy distributions, the free energy differences between intermediate states, ΔG_k, can be evaluated using MBAR, thus recovering the free energy difference between the target end states as $\Delta G = \sum_k \Delta G_k$. In the FEP scheme as well as in the MBAR approach, the accuracy in the determination of free energy differences crucially depends on the overlap of the potential energy distribution of the contiguous intermediate states. It follows then that the number of intermediate states needed for determining the free energy difference between the target end states depends on the underlying (and unknown) free energy function with respect to the alchemical λ coordinate. A "smooth" free energy function requires less intermediate states for a reliable free energy determination.

Special Issue: Free Energy Calculations: Three Decades of Adventure in Chemistry and Biophysics

Received: February 18, 2014
Published: April 8, 2014

With this respect, in order to avoid numerical instabilities and the insurgence of secondary minima, the intermediate states in alchemical transformations involving molecules or part of molecules are defined so that the charging/discharging and van der Waals coupling/decoupling processes of the alchemical species are done sequentially and not simultaneously, producing on the overall a smoother behavior of the free energy as a function of the λ coordinate. In general—for example, in the calculation of the water solvation free energy of simple organic molecules—from few to few tens of intermediate states are needed.

The method is indeed reliable but quite cumbersome. In the present paper, we follow a different route, the so-called "fast switching" or "fast-growth" method,[17,19,20] in its bidirectional variant based essentially on growth and annihilation using nonequilibrium steered molecular dynamics simulation and exploitation of the Crooks theorem.[21,22] In this method, rather than performing long (in the nanoseconds) equilibrium simulations at *fixed* λ_k states (where k labels the alchemical state), fast nonequilibrium simulations are produced connecting the target end states by switching on and off *continuously* the λ alchemical parameter between 0 and 1. In each of these fast nonequilibrium simulations, where the interactions of the alchemical system with the solvent (but not the interaction within itself) are turned off or on by means of a continuous λ variation, the work W done for the switching on/off process is computed producing the distribution $P_{on}(W)$, $P_{off}(-W)$. These latter work distributions are related to the free energy difference between the target end states, ΔG, via the Crooks fluctuation theorem[21,22] (CFT)

$$\frac{P_{on}(W)}{P_{off}(-W))} = \exp[\beta(W - \Delta G)] \tag{1}$$

We are not the first to use the fast switching approach for bidirectional processes combined with the Crooks theorem. The technique has been discussed recently[23] by Gapsys et al. proposing a modification of the Beutler et al. soft-core potential[3] for the van der Waals interactions implemented in the GROMACS program. Such modification was claimed to prevent the insurgence of singularities and numerical instabilities produced by conventional approaches. However, the main source of numerical instabilities in the study of Gapsys et al. stems from the *simultaneous* modification the van der Waals and electrostatics interactions, which is known to produce a less smooth variation of the underlying free energy function.[17] A simple and effective way to eliminate the instabilities is that of switching on (or off) electrostatic and van der Waals interactions sequentially. Here, we show that such an approach, along with the use of the Beutler/GROMACS soft-core potential, contrary to what is claimed in ref. 23, is fully practical and computationally efficient within CFT. Also in ref 23, though the fast switching trajectories are standardly produced using the efficient PME-Ewald[24] to rigorously account for the long-range electrostatic energy and forces, the electrostatic contribution to the work, which is of course still an energy with a nonabsolutely convergent r^{-1} behavior, is incoherently computed using a simple spherical cutoff, thus introducing a systematic error that critically depends on the total charge within the cutoff sphere.[25] Besides, such a nonrigorous approach prevents the optimized use of efficient multiple time step schemes within the PME framework for electrostatic interactions, with a significant impact on the overall efficiency of the algorithm. In the present paper, the alchemical work is computed numerically to $\delta\lambda^2$ precision with a rigorous and coherent treatment of the electrostatic interaction with PME. We show that in "fast" solvents like water as well as in more viscous media such as 1-octanol, a few hundred of the independent simulations for the creation and annihilation process of a few tens of picoseconds can produce very accurate solvation free energies. These independent, nonequilibrium alchemical simulations can be run on parallel platforms with no communication overhead. Therefore, although the overall simulation time is comparable to that of the MBAR approach with fixed intermediate states (tens to hundreds of nanoseconds), the wall-clock time is *at least* 1 order of magnitude smaller. The paper is organized as follows. In section 2 ,we revise the theory of continuous alchemical transformations focusing on the calculation of the alchemical work when electrostatic interactions are rigorously accounted for using PME. In section 3, we describe in detail the methods and the implementation of the fast switching technique. In section 4, we show some results concerning the calculation of the partition coefficients of some organic molecules. Concluding remarks and perspectives are discussed in section 5.

THEORY

In the following, we shall describe in details the theory of continuous alchemical transformations, with focus on the issues and technicalities regarding the implementation in molecular dynamics codes using the Ewald method. As we will see, running a simulation using standard implementation of the Ewald methods of a system where atomic charges are varying, implies the insurgence of nontrivial terms in the energy and forces that must be considered for producing correct trajectories. In a nutshell, Ewald resummations consist of adding and subtracting to the atomic point charge a spherical Gaussian charge distribution bearing the same charge, so that the electrostatic potential is split in a fast dying term (the Erfc term) due to the sum of the point charge and the neutralizing charge distribution and evaluated in the direct lattice, and in a slowly decaying term (the Erf term) due to the added Gaussian spherical distributions[26] evaluated in the reciprocal lattice. Thanks to this trick, the conditionally convergent electrostatic energy sum is split in two absolutely convergent series. In standard implementations of the Ewald resummation technique, as we will see later on, the electrostatic potential at the atomic position $\mathbf{r}_i$ is actually not available with mixing of the interactions between alchemical and nonalchemical species in the so-called Ewald reciprocal lattice contribution (i.e., the Erf part). The smooth particle mesh Ewald method[24] (SPME) makes no exception, with the additional complication that the atomic point charges (including the alchemical charges) are now smeared over nearby grid points to produce a regularly gridded charge distribution, to be evaluated using fast Fourier transform (FFT). Due to the extraordinary efficiency, SPME is still an unrivaled methodology for the evaluation of electrostatic interactions in complex systems.[27] Moreover, SPME can be straightforwardly incorporated in fast multiple time step schemes producing extremely efficient algorithms for systems of biological interest, for example.[28] For these reasons, it is highly desirable to devise rigorous and efficient approaches to account for alchemical effects in a system treated with PME.

Production of the MD Trajectory with an Externally Driven Alchemical Process. In a system of N particles subject to a continuous alchemical transformation, only the

nonbonded potential energy function is modified because of the presence of alchemical species. The full nonbonded energy of the system is given by

$$V(\mathbf{r}_1, ..., \mathbf{r}_N, \lambda, \eta) = \sum_i \sum_{j>i} [1 - \lambda_{ij}(t)] \frac{Q_i Q_j}{r_{ij}} \mathrm{erfc}(\alpha r_{ij}) - \frac{\alpha}{\pi^{1/2}} \sum_i [1 - \lambda_i(t)]^2 Q_i^2 + \frac{1}{2\pi V} \sum_{\mathbf{m}\neq 0} \frac{e^{-\pi^2 \mathbf{m}^2/\alpha^2}}{\mathbf{m}^2} \times \left[\sum_i \sum_{j\neq i} [1 - \lambda_{ij}(t)] Q_i Q_j e^{-i2\pi \mathbf{m}\cdot\mathbf{r}_{ij}} + \sum_i [1 - \lambda_i(t)]^2 Q_i^2 \right] + \sum_i \sum_{j>i} 4\varepsilon_{ij} [1 - \eta_{ij}(t)] \times \left(\frac{1}{[\gamma \eta_{ij}(t) + (r_{ij}/\sigma_{ij})^6]^2} - \frac{1}{\gamma \eta_{ij}(t) + (r_{ij}/\sigma_{ij})^6} \right) \quad (2)$$

where V the unit cell volume, $\mathbf{m}$ is a reciprocal lattice vector, and α is the Ewald convergence parameter related to the width of the Gaussian spherical charge distribution. The first term in the nonbonded energy eq 2 is limited to the zero-cell and corresponds to the electrostatic interactions in the direct lattice; the second term refers to the self-interactions of the Gaussian charge distributions and the third term corresponds to the interactions between Gaussian distributions in the zero cell as well as in the infinite direct lattice, reformulated as an absolutely convergent summation in the reciprocal lattice. The last term in eq 2, finally, corresponds to the modified atom−atom van der Waals interaction introduced in ref 16 incorporating a soft-core parametrization, where the infinity in the Lennard-Jones interaction is smoothed to zero as a function of the η_i. The parameter γ is a positive constant (usually set[29] to 0.5) that controls the smoothing to zero of the derivatives of the Lennard-Jones function as r tends to zero.[30] In the present general formulation, according to eq 2, all atoms of the systems, whether alchemical or not, are characterized by an additional, time dependent and externally driven "coordinate", the $\lambda_i(t)$ parameter controlling the charging/discharging of the system and the $\eta_i(t)$ parameter for switching on or off the atom−atom Lennard-Jones potential. The time dependence of the $\eta_i(t)$ and $\lambda_i(t)$ atomic factors is externally imposed using an appropriately selected time protocol. The nonbonded potential energy of eq 2 coincides with the standard potential energy of a system with no alchemical species when all the alchemical atomic factors, $\lambda_i(t)$ and $\eta_i(t)$, referring to electrostatic and van der Waals interactions, are constant and equal to zero. At the other extreme, when $\lambda_i(t) = \eta_i(t) = 1$, the alchemical species disappears according to the "mixing" rules for $\lambda_{ij}(t)$ and $\eta_{ij}(t)$ factors specified in Table 1. These rules are such that the modified alchemical potential is enforced only when one of the two interacting atoms is alchemical while atom−atom interactions within a given alchemical species are accounted for with the standard potential or simply set to zero when they do refer to atoms on different alchemical species. In general, the time protocol for the λ_i and η_i van der Waals and electrostatic atomic parameters may differ from each other and for different alchemical species. A simple and sufficiently flexible scheme[31] would be that, for example, of allowing only two sets of alchemical species, that is, the species to be annihilated and the species to be created, hence defining two different time protocols for the λ_i and two more for the η_i atomic parameters. Such a scheme allows, for example, the determination of the energy difference when one group in a molecule is replaced by an other group in a single alchemical simulation.

Table 1. Atom−Atom Combination Rules for Alchemical and Nonalchemical Species[a]

i	j	$\lambda_{ij}(t)$	$\eta_{ij}(t)$
alchemical	solvent	$\lambda_i(t)$	$\eta_i(t)$
solvent	alchemical	$\lambda_j(t)$	$\eta_j(t)$
solvent	solvent	0	0
alchemical G	alchemical G	0	0
alchemical A	alchemical A	0	0
alchemical A	alchemical G	1	1
alchemical G	alchemical A	1	1

[a]The alchemical systems may contains three species: (i) the alchemical growing subsystems, (ii) the alchemical annhiliating subsystems, and (iii) the nonalchemical solvent. The $\lambda_i(t)$ and $\eta_i(t)$ atomic factors within each of this species are all identical and equal to $\lambda_{G/A/S}(t)$, $\eta_{G/A/S}(t)$, where the index G,A,S label the growing, annhilating, and solvent species, respectively.

As remarked by others,[29] it is convenient in, for example, an alchemical creation to switch on first the van der Waals parameters changing η for the alchemical atoms from one to zero and then charge the system varying λ from one to zero. Although for soft-core Lennard-Jones term and the direct lattice electrostatic term the combination rules described in Table 1 can be straightforwardly implemented at a very limited computational cost in a standardly written force routine, the same rules cannot be directly applied to the reciprocal lattice part. In common implementation of the Ewald method, for computational convenience, the reciprocal lattice space double sum is rewritten in terms of a squared charge weighted structure factors as

$$V_{\mathrm{rl}} = \frac{1}{2\pi V} \sum_{\mathbf{m}\neq 0} \frac{\exp(-\pi^2 |\mathbf{m}|^2/\alpha^2)}{|\mathbf{m}|^2} S(\mathbf{m}) S(-\mathbf{m}) \quad (3)$$

In a system subject to a continuous alchemical transformation, the charge weighted structure factor becomes a function of the atomic factors $\lambda_i(t)$

$$S(\mathbf{m}, \lambda) = \sum_i^N (1 - \lambda_i(t)) Q_i \exp(-2\pi i \mathbf{m}\cdot\mathbf{r}_i) \quad (4)$$

In the PME method, the sum of eq 4 is done via FFT by smearing the atomic charges on a regular grid in the direct lattice.[24,28] In this approach, all charge−charge interactions between alchemical and nonalchemical species are almost inextricably mixed in the PME Ewald reciprocal lattice contribution and the application of the rules reported in Table 1 requires an extra effort indeed, an effort that has apparently deterred many[23,26,32] from using the full Ewald method for computing the work done during continuous alchemical transformations. To this end and with no loss of generality, it is convenient to classify the system in an alchemical "solute" and in a nonalchemical "solvent", with only the former being externally driven. We then label with $q(t)$ and Q the time-dependent alchemical charges and the full time-invariant atomic charges of the solute, respectively, and with Q_S the charges on the solvent. The alchemical $q(t)$ and full Q solute charges are related by $q(t) = (1 - \lambda(t))Q$. When

dx.doi.org/10.1021/ct500142c | J. Chem. Theory Comput. 2014, 10, 2813−2823

evaluating the reciprocal lattice energy via eq 3, the situation for the charge−charge electrostatic interactions is represented in Table 2. In the direct lattice, the rules reported in Table 2 can

Table 2. Charge−Charge Interactions in Alchemical Transformations Using the Ewald Summation[a]

direct lattice (Erfc)			reciprocal lattice (Erf)		
QQ/r	$(q(t)Q_s)/r$	Q_sQ_s/r	$(q(t)q(t))/r$	$q(t)Q_s/r$	Q_sQ_s/r
only interactions ≥1−4			all interactions		

[a]The atomic charges labeled $q(t)$, Q, and Q_s refer to the alchemical charge, to the full (time-invariant) solute charge, and to the solvent (nonalchemical) charge, respectively.

be implemented straightforwardly by excluding in the double atomic summation of eq 2 all the so-called 1−2 and 1−3 contacts. These atom−atom contacts involve directly bonded atoms of atoms bound to a common atom for which no electrostatic charge−charge contribution should be evaluated. In the reciprocal lattice, however, because of the structure of eq 3, all intrasolute interactions are implicitly of the kind $q_i(t)$ $q_j(t)\mathrm{erf}(\alpha r)/r_{ij}$ and 1−2 and 1−3 pairs are automatically considered in the sum of eq 3. The latter terms may be standardly removed in the zero cell by subtracting from the energy the quantity

$$V_{\mathrm{intra}} = \sum_{ij-\mathrm{excl.}} q_i q_j \frac{\mathrm{erf}(\alpha r_{ij})}{r_{ij}} \tag{5}$$

Regarding the 1−4 interactions, these are fully included in the reciprocal lattice sum, whereas in popular force fields, only a portion of them are considered via the so-called fudge factor f. What must be subtracted in this case is the complementary interaction $q_i(t)q_j(t)(1-f)\mathrm{erf}(\alpha r)/\mathrm{r_{ij}}$.

It should be stressed here that, when the reciprocal lattice sum is computed using eq 3, the zero cell Erf contribution of the 1−2, 1−3, and 1−4$(1-f)$ interactions must be removed whether the two involved charges are alchemical or not. So, alchemically driven simulations imply no changes on the subtraction of these peculiar self-interactions with respect to a normally implemented program with no alchemical changes. The routines that implement eq 5 must be called, therefore, using the atomic charges $q_i = (1-\lambda_i(t))Q_i$ whether alchemical or not (i.e., whether λ_i is different from zero or not). With the same spirit, the self-interaction in the zero cell, that is, the term $-(\alpha/\pi^{1/2})\sum_i[1-\lambda_i(t)]^2Q_i^2$, must be computed using the same charges.

We have seen in Table 2 that in the direct lattice the intrasolute nonbonded electrostatic interactions are computed using the full time invariant solute charges Q_i as alchemical changes affect only solute−solvent interaction energies. To recover the bare Coulomb potential for intrasolute interaction in a system subject to an alchemical transformation one must then subtract, as done for the 1−2, 1−3, and 1−4$(1-f)$ pairs, the Erf $q(t)q(t)$ contribution, and *add* a QQ Erf term to the total energy of the system, producing the alchemical correction to the electrostatic energy

$$V_{\mathrm{alch}} = \sum_{ij>14} Q_iQ_j[1-(1-\lambda_i(t))(1-\lambda_j(t))]\frac{\mathrm{erf}(\alpha r_{ij})}{r_{ij}} \tag{6}$$

where the summation is extended to all nonbonded intrasolute interactions. It should be stressed that the energy of eq 6 is a nontrivial additive term that *must* be included in simulations of continuous alchemical transformations. Such a term stems from the time-dependent alchemical charges $q(t)$ and is due to the peculiar implementation of the Ewald method. Although overlooked in recent implementation of fast switching alchemical transformations,[23] V_{alch} is indeed a large contribution (3−4 kcal mol^{-1} per solute atom) and its neglect may lead to severe errors in the electrostatic energies and to incorrect MD trajectories.

We can finally rewrite down the total energy of a system subject to an alchemical transformation as

$$\begin{aligned} V(\mathbf{r}_1, ..., \mathbf{r}_N, \lambda, \eta) = {} & \sum_i \sum_{j>i} [1-\lambda_{ij}(t)]\frac{Q_iQ_j}{r_{ij}}\mathrm{erfc}(\alpha r_{ij}) \\ & + V_{\mathrm{rl}} - V_{\mathrm{intra}} - \frac{\alpha}{\pi^{1/2}}\sum_i [1-\lambda_i(t)]^2Q_i^2 + V_{\mathrm{alch}} \\ & + \sum_i \sum_{j>i} 4\varepsilon_{ij}[1-\eta_{ij}(t)]\left(\frac{1}{[\alpha\eta_{ij}(t)+(r_{ij}/\sigma ij)^6]^2}\right. \\ & \left. - \frac{1}{\alpha\eta_{ij}(t)+(r_{ij}/\sigma ij)^6}\right) \end{aligned} \tag{7}$$

where V_{rl}, V_{intra}, V_{alch} are defined in eqs 3, 5, and 6, respectively. All terms in eq 7, except for the self-term $(\alpha/\pi^{1/2})\sum_i[1-\lambda_i(t)]^2Q_i^2$, contribute to the atomic forces that can be standardly computed by taking the derivatives of the energy eq 7 with respect to the atomic position $\mathbf{r}_i$ producing the correct trajectories for alchemically driven systems under periodic boundary conditions and treated with the Ewald sum. In Figure 1, we report the time record of the intrasolute electrostatic

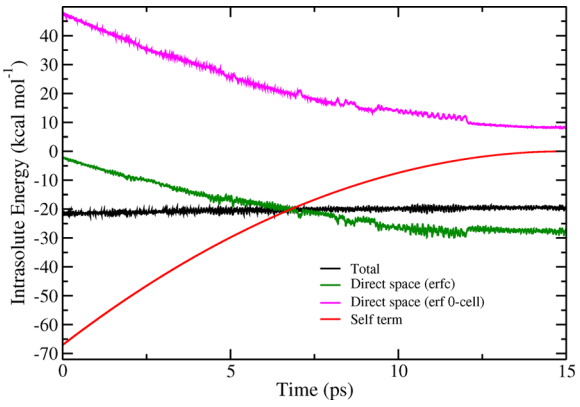


Figure 1. Time record for the intrasolute energy arising form electrostatic interactions during the alchemical discharging of ethanol in water at T = 300 K and P = 1 atm. The simulation went on for 15 ps. The red curve is due the self-term, $-(\alpha/\pi^{1/2})\sum_i[1-\lambda_i(t)]^2Q_i^2$. The green curve is due to the direct lattice contribution. The magenta curve includes the terms $-V_{\mathrm{intra}}$ (eq 5) and V_{alch} (eq 6).

energy during the discharging of a molecule of ethanol in water in standard conditions. In spite of the huge changes in the contributing energy terms (k_BT is only 0.59612 kcal mol^{-1}), the total intrasolute energy remains approximately constant during the transformation, modulated by the intramolecular motion, exactly as it should. The changes in the solute self-term, $-(\alpha/$

dx.doi.org/10.1021/ct500142c | J. Chem. Theory Comput. 2014, 10, 2813−2823

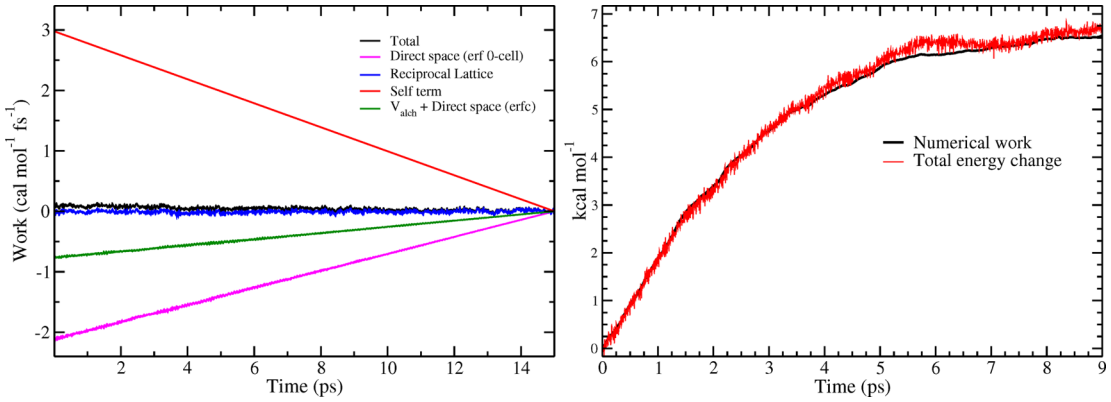


Figure 2. Left: Time record for the intrasolute reciprocal lattice contributions to the differential work (eq 10) arising from electrostatic interactions during the alchemical discharging of ethanol in water at T = 300 K and P = 1 atm. The simulation went on for 15 ps. The red curve is due the self-term, $-(\alpha/\pi^{1/2})\sum_i[1-\lambda_i(t)]^2Q_i^2$. The green curve is due to the direct lattice contribution and to V_{alch}. The magenta curve includes the terms $-V_{\text{intra}}$ (eq 5). The blue curve is due to the full reciprocal lattice PME term, eq 3. Right: Total energy change (red line) and numerical work (black line) computed using eq 10 for the discharging of ethanol in water in an alchemical trajectory lasting for 9 ps.

$\pi^{1/2})\sum_i[1-\lambda_i(t)]^2Q_i$, compensate at all time steps the variation of the direct lattice and of Erf intrasolute corrections. This balance does occur provided that *all* terms in the energy of eq 7 are accounted for, including the intrasolute alchemical Erf correction V_{alch} of eq 6.

In a multiple time scheme, the individual contributions to the nonbonded forces evolve in time with disparate time scales and, hence, must be partitioned in appropriately defined "integration shell" as described in detail in ref 33. Therefore, in condensed phases, the direct lattice term is integrated in the fast short-ranged nonbonded shell, whereas the reciprocal lattice summations (including the Erf intramolecular correction terms in V_{intra}) are usually assigned, with an appropriate choice[28,33] of the Gaussian parameter α, to the intermediate nonbonded shell. The Lennard-Jones term, finally, is split among the short-range, intermediate-range, and long-range integration shells. The potential subdivision for condensed phases is basically unaffected by the implementation of alchemical transformation, except for the intrasolute self-term V_{alch} and for the now time-dependent self-term, $(\alpha/\pi^{1/2})\sum_i[1-\lambda_i(t)]^2Q_i^2$. [Note: This last term does not contribute to the atomic forces but only to the alchemical work and is constant for all nonalchemical species. The work done by an alchemical species through this term is simply given by $W = \pm(\alpha/\pi^{1/2})\sum_i Q_i^2$, depending whether the alchemical species has been charged or discharged.] The latter can be safely included in the intermediate shell, whereas the former, evaluated in the direct lattice, must be integrated in the short-range shell. The $\lambda_i(t)$ and $\eta_i(t)$ factors, finally, must be updated according to the predefined time protocol before the force computation of the fast short-range nonbonded shell. More details on the implementation of alchemical transformations with multiple time step schemes are given in the Supporting Information.

Calculation of the Alchemical Work. The work done on the system by the driven alchemical coordinates during a simulation of length τ can be written as

$$W = \int_0^\tau \frac{\partial H(x,\lambda,\eta)}{\partial\lambda}\dot{\lambda}\,\mathrm{d}t + \int_0^\tau \frac{\partial H(x,\lambda,\eta)}{\partial\eta}\dot{\eta}\,\mathrm{d}t \tag{8}$$

In an NVT or NPT extended Lagrangian simulation with an ongoing alchemical process, the alchemical work, eq 8, could be computed simply by monitoring the changes in the *total* energy of the systems, and that includes the real potential and kinetic energy of system and the potential and kinetic energies of the barostat and the thermostats. This energy, if no velocity scaling is implemented (i.e., no heat is artificially transferred to or absorbed from the *extended* system), is a constant of the motion and, hence, any variation of it must correspond to the work done on the system.[34] Alternatively the work can be computed by analytically evaluating the λ and η derivatives of the nonbonded energy eq 7. Both of these methods have counterindications. Small drifts in the total energy adds up in the work as a extra dissipation term[35,36] that may reduce the accuracy in the free energy determination via the Crooks theorem, especially when using multiple time step schemes. The method based on the derivatives, if alchemical species are annihilated and created within the same process, requires the constant tagging of the *two* creation and annihilation works, as the increments $\delta\lambda_{G/A}$ or $\delta\eta_{G/A}$ have opposite signs for creation (G species) and annihilation process (A species). Besides, although all direct lattice Erfc and reciprocal lattice Erf corrections terms pose no difficulties in λ derivation with a moderate extra cost of the force routines, the analytic derivation of reciprocal lattice energy V_{rl}, eq 3, with respect to λ implies the calculation of three gridded charge arrays, that is, one for the whole system and two more for the discharging and for the charging alchemical solutes

$$\frac{\partial V_{\text{rl}}}{\partial\lambda_i} = -\frac{1}{2\pi V}\sum_{\mathbf{m}\neq 0}^{\infty}\frac{\exp(-\pi^2|\mathbf{m}|^2/\alpha^2)}{|\mathbf{m}|^2}(S(\mathbf{m})S_{G/A\ \text{slt}}(-\mathbf{m}) + S(-\mathbf{m})S_{G/A\ \text{slt}}(\mathbf{m})) \tag{9}$$

where with the notation $S_{G/A\ \text{slt}}(\mathbf{m})$ we refer to the gridded charge arrays obtained for charging alchemical species ($1 \le \lambda \le 0$) and for the discharging ($0 \le \lambda \le 1$) if they are both present. The work can also by computed numerically observing that the differential work due to a $\delta\lambda$ or $\delta\eta$ increment of the alchemical factors is given by

$$dw = \frac{1}{2}(E(\lambda+\delta\lambda,x) - E(\lambda-\delta\lambda,x) + E(\eta+\delta\eta,x) - E(\eta-\delta\eta,x)) \tag{10}$$

which is correct to order $o(\delta\lambda^2)$ and $o(\delta\eta^2)$. Equation 10 requires just one extra calculation of the energy within the direct space force loop using the λ_i values at the previous step

with no need for tagging annihilating and creating species. For computing the work arising from the reciprocal lattice sum, eq 3, the gridded charge array must be computed at every step of the intermediate-range shell using the current charges and those at the previous step with a very limited computational cost. *Both* of these arrays must then undergo FFT. As for the direct lattice, for the reciprocal lattice term, there is no need for tagging creating or annihilating species. The different means to access the alchemical work can be used as a powerful check to test the coherency of the trajectories and of the computed numerical work, eq 10. The alchemical work indirectly evaluated monitoring the changes of total energy of the system must follow closely the profile of the numerical work computed using eq 10. Such a test is reported in Figure 2 (right) for the discharging of ethanol in water.

In a multiple time step scheme, the alchemical work must be computed exactly as the energy is computed, hence evaluating more often the contributions arising from the fast shells with respect to the terms evolving more slowly. In the scheme reported in the Supporting Information, we succinctly describe the implementation of the alchemical process and the associated work calculation in a molecular dynamics code, highlighting the parts of the code that must be modified because of the presence of alchemical species with respect to a normal MD code. In Figure 2, we report the behavior of the various contributions to the intrasolute differential work computed during the transformation. In the reciprocal lattice term (blue curve), the intrasolute and solute−solvent contributions are mixed. Hence, the integrated total differential work (black curve) is, expectedly, slightly positive because of a loss of long-range electrostatic energy because of ethanol discharging. Again, paralleling the situation seen for the intrasolute energy, the work due to the self-term approximately cancels the end Erfc intrasolute contributions.

We conclude this section with some comments on the time protocol that drives the alchemical transformation. In our implementation, the charges and the Lennard-Jones potential can be switched on and off independently, by setting up different time protocols for η_i and λ_i alchemical coordinates. Such an approach is much more flexible and powerful than that based on the definition of a single alchemical parameter, implying the simultaneous variation of Lennard-Jones and electrostatic interactions. As remarked in ref 23, if the η_i and λ_i factors are varied coherently (i.e., only *one* type of alchemical coordinate Λ_i is defined), catastrophic numerical instabilities may arise, especially in complex solutes with competing conformational structures. According to ref 23, one way to circumvent this problem is to switch electrostatic and Lennard-Jones interactions separately as we do here. Contrary to what is stated in ref 23, we show that such an approach has no significant impact on the simulation time and poses no difficulties whatsoever for the equilibrium simulations of the target end states.

Concerning this latter issue, for the evaluation of solvation free energy via alchemical transformations, the target end states are (i) the decoupled solute (in the gas phase) and the pure solvent (in the liquid state) and (ii) the solution. For the decoupled state (i), in principle, two independent standard simulations are needed, one for the isolated solute and the other for pure solvent. However, the decoupled state can be sampled in one single simulation using the nonbonded energy of eq 7, by setting the alchemical solute λ_i and η_i factors all equal to one. In fact, according to eq 7 and to the rules of Table 1, when the alchemical solute λ_i and η_i terms are all equal to one, the solute is not felt by any means by the solvent and evolves in time independently, subject only to the intramolecular interactions with no contribution from the solute lattice images. The intrasolute electrostatic energy, in particular, has no contribution from the reciprocal lattice sum as the λ_i referring to the solute are all equal to 1 in eq 3. It has, indeed, a direct lattice contribution for nonbonded intrasolute evaluated in the zero cell according to the rules specified in Table 1 *plus* the alchemic correction term that simply corresponds (with all solute λ_i set to 1) to the complementary Erf part, thus recovering the bare intrasolute Coulomb energy. At the other extreme end of the alchemical transformation ($\lambda_i = 0$, $\eta_i = 0$), according to eq 7, the solute is fully charged, interacting normally with the solvent and with the solute images via the term eq 3. We now come to the issue of the efficiency of a code with distinct Lennard-Jones and charge alchemical parameters. Of course, also in this case simultaneous switching of λ_i and η_i remains perfectly possible. To avoid numerical instabilities at the early stage of the creation process or at the end of the annihilation, it is sufficient in the first case to slightly delay the charge switching and, in the last case, to anticipate the discharging process. In Figure 3, we report the work computed

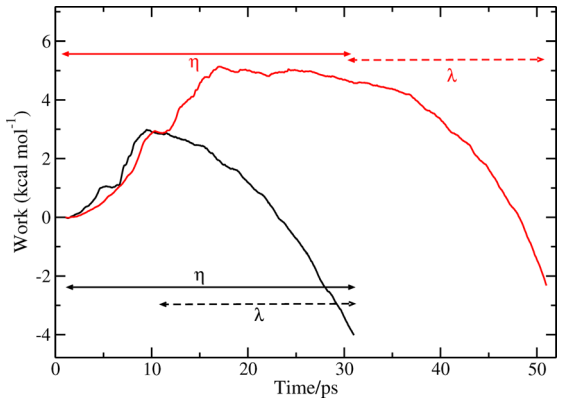


Figure 3. Alchemical work produced in the creation of ethanol in water T = 300 K and P = 1 atm using two different time protocols represented by the black and red horizontal lines.

in the alchemical creation of ethanol in water conducted with two different time protocol. In the red nonequilibrium trajectory, the Lennard-Jones η_i parameters for ethanol are prudently brought from 1 to 0 in 30 ps, and in the next 20 ps the solute is charged. In the black trajectories lasting for 30 ps, in the first 10 ps, the η_i coordinates alone are brought from 1 to 0.5 and then, in the last 20 ps, they are brought to zero (fully switched on ethanol) *together* with the charging process that is started at 10 ps. As one can see, both trajectories are regular with no instabilities, yielding negative and comparable works with limited dissipation with respect to the reversible work ($\simeq$ −4 kcal mol^{-1}; see the next section) in spite of short duration of the nonequilibrium alchemical transformations. We must stress here, that in the fast switching nonequilibrium method with determination of the free energy difference between end states via the CFT, once the equilibrium configurations of the starting end states have been prepared, the simulation time per trajectory indeed does correspond to the wall-clock time if the independent nonequilibrium trajectories are performed in parallel. For the creation of ethanol in water, the CPU time

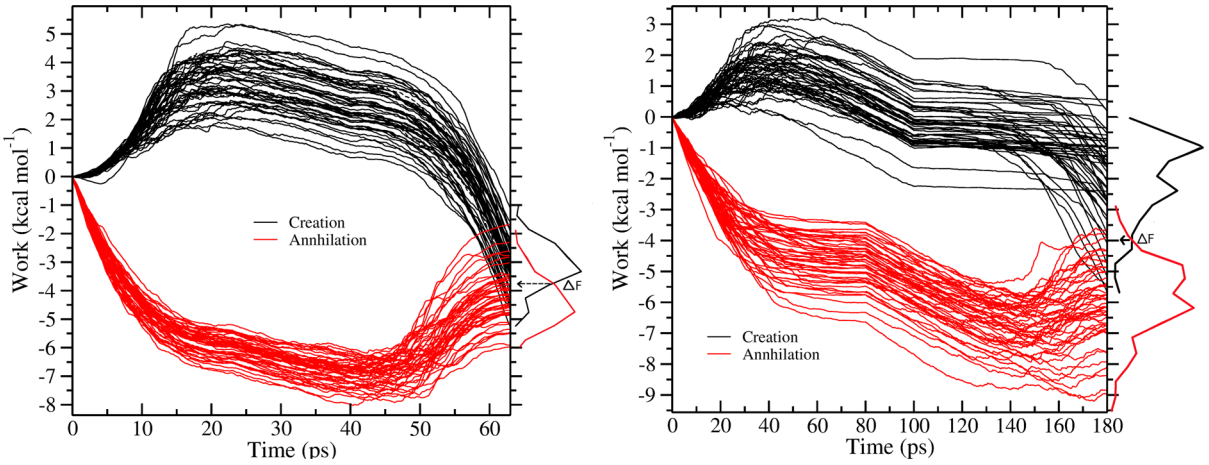


Figure 4. Computed work as a function of the simulation time in nonequilibrium alchemical trajectories for ethanol in water (left) and in 1-octanol (right). The black and red curves refer to the *forward* process (decoupled state to solution) and to the *reverse* process with inverted time schedule (solution to decoupled state). On the right the corresponding work distributions are shown.

amounts to few minutes on a low-end desktop computer for both time protocols.

METHODS

We apply the fast switching/CFT method to the determination of the free energy of solvation of some small organic molecules in water and in octanol, that is, ethanol, benzene, benzamide, and pentane. Water is described by the TIP3P model[37], whereas the standard GAFF/AMBER parameters[38] have been used for octanol as well as for ethanol, benzene, benzamide, and pentane. Atomic charges were computed on the optimized structures using the electrostatic potential fit (ESP) scheme by Merz−Singh−Kollman[39] at the B3-LYP/6-31G(d) level of theory. For the polar molecules ethanol and benzamide, two set of ESP charges were computed, one for the simulation in octanol and the other for the simulation in water, using the polarizable continuum model with the integral equation formalism variant (IEFPCM).[40] The simulations in water and octanol included 520 and 59 solvent molecules, respectively, and were run in the isothermal isobaric ensemble. The constancy of temperature (298 K) was enforced using a Nosé−Hoover thermostat,[41] whereas the external pressure of 0.1 MPa was imposed using a modification of the Parrinello−Rahman Lagrangian, allowing for isotropic volume changes.[42] The electrostatic interaction were treated using the smooth particle mesh Ewald method,[24] with an α Gaussian parameter set to 0.43 Å^{-1}, 24 grid points in each direction (approximately 1 Å of grid spacing in all cases) and a fourth order B-spline interpolation for the evaluation of the gridded charge array. The cutoff for Lennard-Jones interactions was set to 15 Å. The equations of motion were integrated using a multiple time step r-RESPA algorithm[43] with a potential subdivision specifically tuned for biomolecular systems.[28,42] The starting end states for the fast alchemical transformations were prepared by running equilibrium simulations (lasting about 1 ns) with the nonbonded energy function of eq 7 with the λ_i and η_i parameter set to 1 (decoupled gas-phase solute and pure solvent) and set to zero (the solution). A total of 256 full restarting equilibrium configurations for both states were saved, each 5 ps. Starting from these initial points, forward (decoupled state to solution) and reverse (solution to decoupled state) fast switching alchemical processes were conducted in parallel on 256 cores. The creation time protocol for the nonequilibrium process is the same for all solutes with the annhilitation process being done with an inverted time schedule as prescribed by the CFT. In water, the Lennard-Jones η_i factors of the solute are fully switched on in the first 40 ps, whereas the solute is charged in the following 23 ps for a total of 63 ps simulation time. In octanol, the switching on of the solute Lennard-Jones potential lasts for 100 ps, and the charging process is performed for the following 80 ps, for a total of 180 ps simulation time. As discussed when commenting Figure 3, the adopted time protocol for creation and annihilation is a rather prudent one with no simultaneous switching of the Lennard-Jones and electrostatic interactions. With this respect, the CPU time for running serially 63 ps alchemical simulation in water using the ORAC code[27] on a low end desktop workstation is in the order of 15 min, whereas for octanol, it is less than 1 h. The adoption of the protocol with delayed charging (black curve in Figure 3) would have spared just few minutes. We remark that these are indeed the wall-clock times needed to produce the nonequilibrium works in parallel. The $n_{F/R}$ final acquired forward and reverse works, $W_i(F/R)$, are then postprocessed by feeding them to a simple code that implements the Bennett acceptance ratio[5,12]

$$\sum_{i=1}^{n_F} \frac{1}{1+\frac{n_F}{n_R}e^{\beta(W_i(F)-\Delta G)}} - \sum_{i=1}^{n_R} \frac{1}{1+\frac{n_R}{n_F}e^{\beta(W_i(R)+\Delta G)}} = 0 \tag{11}$$

for instantly recovering the free energy change ΔG between the end states.

RESULTS

In Figure 4, we show the work done as a function of the simulation time for few tens of nonequilibrium fast switching trajectories for ethanol in water (left) and in octanol in the forward process (black lines) and in the reverse process (red lines). For the latter, *minus* the work is reported in order to produce the reverse $P_R(-W)$ distribution satisfying the CFT, eq 1. The associated $P_F(W)$ and $P_R(-W)$ final work distributions, shown on the opposite side in the y axis, cross each other at ΔG as indicated in the Figure. According to the CFT,[5,20,21,34] the distance between the ΔG crossing point on the y-axis and the mean work of the two $P_F(W)$ and $P_R(-W)$ distribution corresponds approximately to the mean dissipation

in the forward and reverse process. The mean dissipative work $W_d = \langle W \rangle - \Delta G$ decreases to zero with increasing simulation time, reducing to zero in the limit of quasistatic (equilibrium) alchemical transformations. Although in a given solvent, the forward ($W_d^F = \langle W \rangle_F - \Delta G$) and reverse ($W_d^R = \Delta G - \langle W \rangle_R$) dissipation works are similar to each other, implicating an approximately symmetric behavior of the dissipation in the annihilation and creation process, W_d appears markedly different when comparing water and octanol. During the fast switching in water (lasting 63 ps), the mean dissipation (either in the forward or in the reverse process) is rather small (less than 1 kcal mol^{-1}). In octanol, in spite of a much longer simulation time (180 ps), the mean dissipated work is at least twice as much. Similar results are consistently found for the other polar solute, that is, benzamide (see Figures 5 and 6).

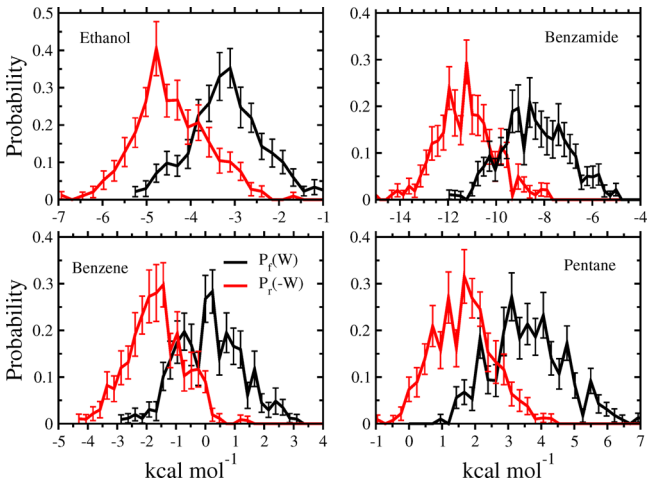


Figure 5. Forward ($P_F(W)$) and reverse ($P_R(-W)$) final work distributions calculated from 256 fast switching alchemical creation and annihilation of some selected organic molecules in water.

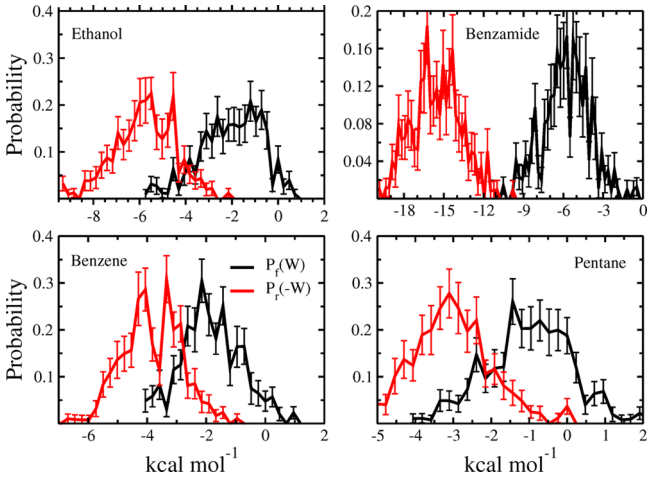


Figure 6. Forward ($P_F(W)$) and reverse ($P_R(-W)$) final work distributions calculated from 256 fast switching alchemical creation and annihilation of some selected organic molecules in 1-octanol.

The reason for such behavior lies in the very nature of the two solvents in standard conditions. Water is a "fast" solvent made of simple molecules with only translational and rotational degrees of freedom. As the solute molecule is growing in the solvent, the water molecules can rapidly reorganize around it in energetically favorable configurations, especially during the charging process (i.e., when the λ_i are varied). Octanol is almost ten times more viscous than water and has complex conformational motions with gauche−trans equilibria along the aliphatic chain. Because of such viscosity, as the dipole of octanol is mostly contributed to by the hydroxyl moiety, the establishment of solute−solvent H-bonds is a much rarer event in octanol with respect to water. In the equal lasting nonequilibrium creation of polar solute in water and octanol, energetically unfavorable configurations are sampled with a much larger probability in the latter solvent, thus producing more dissipative trajectories. In Figures 5 and 6, we report the forward and reverse (with a minus sign) work distributions for all analyzed solutes in water and in octanol. The distributions have been computed with a resolution of 0.25 kcal mol^{-1}, and the error bars were evaluated by block bootstrapping using 50 forward and reverse sets, each containing 128 works randomly sampled out of the 256 set of works. The standard error on $P(W)$ is W dependent and is at a maximum in correspondence of the mean work (i.e., at the maximum of the distribution) and can be as large as 20−30% (see, e.g., benzamide in octanol, Figure 6). Nonetheless, as noticed in previous CFT applications,[5,34] these large uncertainties on the work distributions are not propagated to the final free energy value computed by applying the BAR method, eq 11. As a matter of fact, the BAR-computed free energy from nonequilibrium works produced using the fast switching method turns out to be rather precise as shown in Table 3. The errors reported in Table 3 (evaluated by block bootstrapping) are comparable to those obtained with the MBAR equilibrium method (reported, for example, in refs 16 and 31). In Table 3, we summarize the results obtained for the free energy of solvation of ethanol, benzamide, benzene, and pentane in water and octanol. In general, the calculated values compares very satisfactorily with the experimental data. We must stress here that the experimental data concerning the partition coefficients and solvation free energies are not always coherent as they have been taken from different sources (i.e., ref 44 and ref 45 for the solvation free energies in water and octanol and ref 46 for the partition coefficients). So, for example, the recommended partition coefficient for pentane is reported[46] to be 4.4, whereas the partition coefficient derived from the experimental solvation energies in water[44] and in octanol[45] of pentane is 3.5, in much closer agreement with the calculated value from the fast switching simulation.

In the fast switching approach, the behavior of the free energy as a function of the alchemical coordinates can be straightforwardly reconstructed using the collection of bidirectional realizations. In Figure 7, we report the potential of mean force (PMF) along the alchemical coordinates η for the switching on of the Lennard-Jones potential and λ for the charging process for all solutes in water and 1-octanol calculated from 256 nonequilibrium trajectories. The PMFs have been evaluated using a simple and effective estimator based on the combination of bidirectional exponential averages.[47] In the η switching on portion of the alchemical transformation (i.e., for the 0−40 ps and 0−100 ps time intervals for the forward process in water and 1-octanol, respectively), the behavior of the PMF as a function of the alchemical coordinates reflects the process of the formation of a cavity in the bulk solvent to accommodate the solute and the generation of the uncharged solute inside the cavity. The following process (40 to 63 ps in water, 100 to 180 ps in octanol), that is, the charging of the solute in the solvent

dx.doi.org/10.1021/ct500142c | J. Chem. Theory Comput. 2014, 10, 2813−2823

Table 3. Computed and Experimental Solvation Free Energies (kcal mol^{-1}) and 1-Octanol/Water Partition Coefficients

	water		octanol		log($P_{o/w}$)	
	calculated	experimental[44]	calculated	experimental[45]	calculated	experimental[46]
ethanol	−3.80 ± 0.04	−5.0	−4.00 ± 0.04	−4.4	0.1	−0.3
benzamide	−9.78 ± 0.07	−11.0	−10.28 ± 0.15	−11.9	0.4	0.7
benzene	−0.79 ± 0.04	−0.9	−2.82 ± 0.04	−3.7	1.5	2.2
pentane	2.56 ± 0.05	2.3	−2.03 ± 0.05	−2.4	3.4	4.4

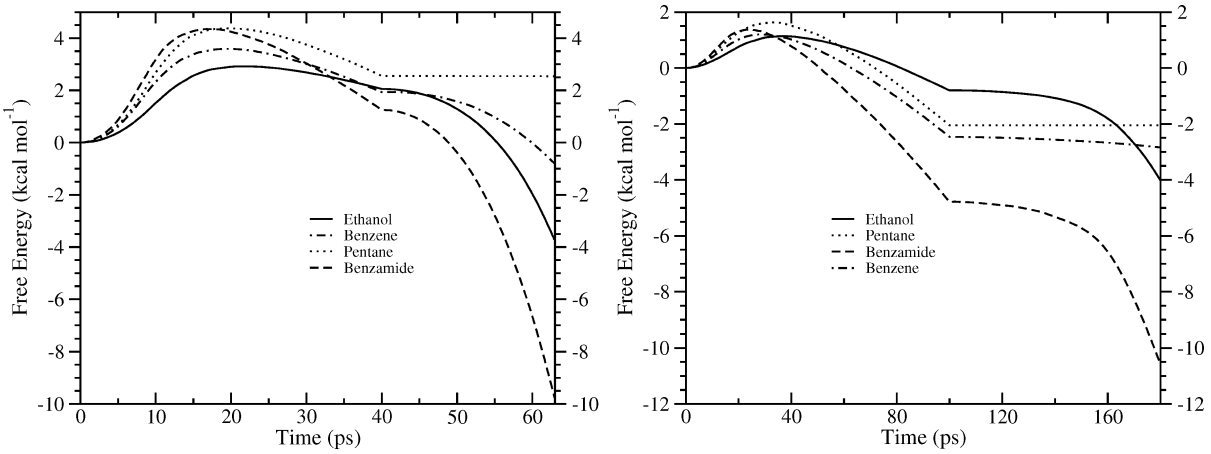


Figure 7. Potential of mean force for the alchemical creation of some organic molecules in water (left) and 1-octanol. (right).

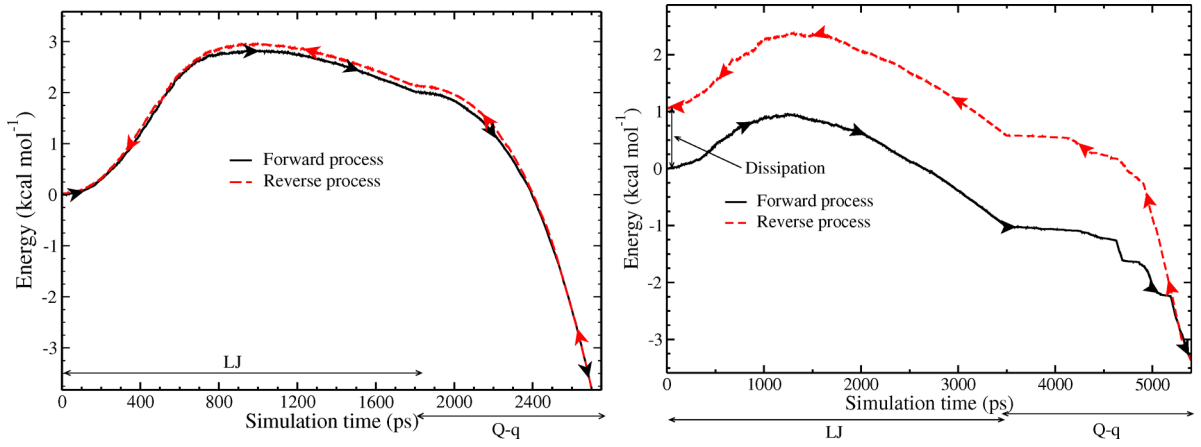


Figure 8. Work obtained during the creation (black line) and time-reversal annihilation (red line) of ethanol in water and 1-octanol. The simulations in water and 1-octanol lasted for 3 and 5.4 ns, respectively.

corresponding to the variation of the λ coordinate in eq 7, adds up the so-called electrostatic contribution to the solvation free energy.

The adoption of the time protocol whereby in the forward process Lennard-Jones and electrostatic interactions are switched on sequentially, provides an opportunity to compare the MD results obtained with the fast switching method to the results obtained with the continuum models.[48] In these approaches, the solvation free energy is partitioned into various additive contributions, namely dispersion−repulsion and electrostatic terms and the so-called cavitation free energy.[49] For example the sum of the cavitation energy and of the dispersion−repulsion terms for ethanol in water is reported[50] to be 2.2 kcal mol^{-1} to be compared with the value of 2.1 kcal mol^{-1} obtained from the PMF evaluated with the fast switching atomistic simulation. In the same study, the electrostatic contribution to the solvation free energy evaluated via dielectric continuum for ethanol in water is −6.4 kcal mol^{-1} and −5.8 in the simulation. For a nonpolar solute such as pentane, only the cavitation and dispersive−repulsive terms contribute to the free energy and these are 1.6 and 2.5 kcal mol^{-1} for the continuum approach and the simulation, respectively.

The PMFs in water and 1-octanol shown in Figure 7 all display a common pattern with an increase in the change of the η coordinate up to a maximum at about η = 0.5 (i.e., 20 ps in water and 50 ps in 1-octanol) followed by a decrease up to completion of the turning on of the Lennard-Jones interactions (occurring at η = 0). On the basis of this observation, it is tempting to identify the cavitation energy with the values of PMF at $\eta \simeq 0.5$, although one should, before, evaluate the effect of the γ parameter used in the definition of the Lennard-Jones soft-core potential on the PMF profiles in the η range.

The partition coefficient between the 1-octanol and the water phases is routinely used in the pharmaceutical practice as a

dx.doi.org/10.1021/ct500142c | J. Chem. Theory Comput. 2014, 10, 2813−2823

paradigm to assess the distribution of drugs in hydrophobic and hydrophilic compartments such as lipid bilayers or blood serum. It is, therefore, worthwhile to further analyze the behavior of these two solvents, water and 1-octanol, with respect to the dissipative phenomena observed during the alchemical switching on or off of the four solutes. As outlined previously, if the simulation is performed at infinitely slow speed, then the continuous alchemical transformation crosses a series of equilibrium alchemical states becoming equivalent to a thermodynamic integration. For a given solute, what "infinitely small speed " actually means, depends on the characteristics of the solvent. The fast pulling experiments performed in this study have shown that for simple organic solutes, the alchemical dissipative work (i.e., the difference between the mean nonequilibrium alchemical work and the free energy of solvation) is in general smaller in water than in 1-octanol, especially for polar solutes. In water in particular, the alchemical dissipation is surprisingly small in a switching time as short as 63 ps, rarely exceeding 2 kcal mol $^{-1}$ in all analyzed cases (see Figure 5). For a comparison, the dissipative work in the driven unfolding of helical deca-alanine *in vacuo* exceeds the 10 kcal mol^{-1} for nonequilibrium simulation lasting at least twice as much, and reduces to about 1 kJ mol^{-1} for simulation times in the range 5−10 ns.[51] Based on these data, we expect that in water, a few nanoseconds of simulation in the alchemical creation/annhilation of a polar solute such as ethanol should de facto correspond to a quasireversible regime. As a matter of fact, the profiles reported in Figure 8 (left), referring to a thermodynamic cycle for the switching on/off of ethanol in water as obtained from forward and reverse alchemical simulations lasting 3 ns, show that the dissipation is negligible for such a time duration (i.e., the alchemical process is in essence reversible), happily recovering the smooth PMF profile reported in Figure 7 like in thermodynamic integration.

The situation is totally different in octanol, as shown in Figure 8 (right). The switching on of ethanol in an alchemical experiment lasting for 5.4 ns (i.e., nearly the double with respect to the simulation performed in water) produces a very irregular profile with several bumps in the charging part of the process. These bumps are very likely due to the transient interactions of the oxidril moieties of the 1-octanol solvent molecules with the same group on the solute. For the reverse process, contrarily to what observed in water, the alchemical work stays consistently above the forward work with a strong hysteresis and producing a dissipation of about 1 kcal mol^{-1} at the end of the thermodynamic cycle, that is, only the half with respect to the mean dissipation observed in fast pulling trajectories (see Figure 6). In spite the longer simulation time, the process in this case clearly appears to be still in a highly nonreversible regime. The PMF for ethanol in 1-octanol, computed using fast switching trajectories and reported in Figure 7, has indeed a smooth behavior consistently lying in between the forward and reverse works evaluated in the 5.4 ns simulations. The benefit of a fast switching approach in this case is evident. A total of 256 trajectories lasting only 180 ps and running in parallel are able to produce an essentially exact PMF when nanosecond simulations of the same systems are far from emulating the performances of thermodynamic integration.

We finally conclude this section with some comments on the computational aspects of the fast switching method. For the case of ethanol in water, we have seen that 128 nonequilibrium alchemical trajectories in the forward and reverse direction lasting 63 ps can produce solvation free energies with numerical uncertainty of the same order of (if not less than) that obtained in the equilibrium MBAR approach as implemented in popular MD packages.[31] Although the total simulation time for the fast switching and equilibrium methods are approximately comparable (a few tens of nanoseconds), it is certainly not so for the wall-clock time. The parallelization of the fast switching approach is indeed straightforward as no communication overhead whatsoever must be paid running *one* fast serial trajectory per core. In the case of the equilibrium technique, the same amount parallel threads must be distributed in the simulation of the few tens of intermediate states, implying communication overhead and suboptimal speed-up ratio for the parallel simulation of one single intermediate state. Besides, the postprocessing stage in the equilibrium method requires, for each λ state, a single point energy calculation done using the parameters of contiguous λ systems for all sampled configuration points. In the fast switching method, the postprocessing stage is limited to the application of eq 11 to the collection of forward and reversible *final* nonequilibrium works.

■ CONCLUSIONS AND PERSPECTIVES

We have discussed a rigorous implementation of continuous alchemical transformations for molecular systems where the electrostatic interaction are treated using the smooth particle mesh Ewald method. This latter approach, by proceeding with the construction of a gridded charge array for subsequent FFT, prevents the conventional application of alchemical schemes based on the knowledge of the local electrostatic potential. We have shown that the correct alchemical trajectory, whereby the growing/annihilating solute interacts via the alchemical charges only with the nonalchemical solvent, can be recovered by judiciously subtracting the unwanted intrasolute Erf modulated charge−charge interactions in the direct lattice automatically included in the SPME reciprocal lattice contribution.

The technique has been used for implementing fast switching alchemical transformations to evaluate the solvation energies of some selected small organic molecules in water and 1-octanol. Bidirectional nonequilibrium trajectories, starting from the initial thermodynamic end states (i.e., (i) the decoupled solute in the gas-phase and the pure liquid and (ii) the solution), have been set up in a massively parallel environment. The solvation free energies have been computed by applying Crooks fluctuation theorem and the Bennett acceptance ratio using the forward and reverse work distribution obtained in the bidirectional experiments. The results compares very satisfactorily with the experimental data. The simulations have highlighted the different natures of the two solvents, water and 1-octanol, with respect to dissipation phenomena during the growth or annihilation process on the solutes. The reversible regime for continuous alchemical is reached for times that are much longer in 1-octanol than in water.

The fast switching approach for evaluating free energy difference can be straightforwardly parallelized with no communication overhead and is competitve, both in efficiency and in precision, with the conventional approach based on equilibrium simulations of intermediate states implemented in popular MD packages. Given the short simulation time needed per nonequilibrium trajectories for recovering accurate estimates of the free energy differences between end states both in water and in 1-octanol, the technology can provide in a massively parallel environment an effective and automated

mean for a rapid access to the partition coefficient of molecules in the lipid bilayer and in the water phase, in order to reliably predict in silico the lypophilicity of lead compounds for effective drug delivery.

The algorithm is made available to the public in the current release of the ORAC code at the Internet address http:/www.chim.unifi.it/orac

■ ASSOCIATED CONTENT

Supporting Information

Pseudocode for implementing alchemical transformations in MD drivers with PME and multiple time step schemes. This material is available free of charge via the Internet at http://pubs.acs.org/.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: procacci@unifi.it.

Funding

CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development, and national and European research programs.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff; see http://www.cresco.enea.it for information. We thank Riccardo Chelli for carefully proofreading the manuscript and for precious suggestions.

■ REFERENCES

(1) Jorgensen, W.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050−3054.
(2) Jorgensen, W.; Thomas, L. *J. Chem. Theory Comput.* **2008**, *4*, 869−876.
(3) Beutler, T.; Mark, A.; van Schaik, R.; Gerber, P.; van Gunsteren, W. *Chem. Phys. Lett.* **1994**, *222*, 5229−539.
(4) Dixit, S.; Chipot, C. *J. Phys. Chem. A* **2001**, *105*, 9795−9799.
(5) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.
(6) Chodera, J.; Mobley, D.; Shirts, M.; Dixon, R.; Branson, K.; Pande, V. *Curr. Opin Struct. Biol.* **2011**, *21*, 150−160.
(7) Woo, H.-J.; Roux, B. *Proc. Nat. Acad. Sci. U. S. A.* **2005**, *102*, 6825−6830.
(8) Deng, Y.; Roux, B. *J. Phys. Chem. B* **2009**, *113*, 2234−2246.
(9) Gumbart, J. C.; Roux, B.; Chipot, C. *J. Chem. Theory Comput.* **2013**, *9*, 974−802.
(10) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420−1426.
(11) Kirkwood, J. G. *J. Chem. Phys.* **1935**, *3*, 300−313.
(12) Bennett, C. H. *J. Comput. Phys.* **1976**, *22*, 245−268.
(13) Shirts, M. R.; Chodera, J. D. *J. Chem. Phys.* **2008**, *129*, 124105.
(14) Chelli, R. *J. Chem. Theory Comput.* **2010**, *6*, 1935−1950.
(15) Procacci, P. *J. Chem. Phys.* **2013**, *139*, 124105.
(16) Shirts, M.; Pande, V. *J. Chem. Phys.* **2005**, 134508.
(17) Shirts, M.; Mobley, D.; Chodera, J. *Annu. Rep. Comput. Chem.* **2007**, *3*, 41−59.
(18) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, L.; Schulten, K. *J. Comput. Chem.* **2005**, *26*, 1781−1802.
(19) Hendrix, D.; Jarzynski, C. *J. Chem. Phys.* **2001**, *114*, 5961−74.
(20) Hummer, G. *J. Chem. Phys.* **2001**, *114*, 7330−7337.
(21) Crooks, G. E. *J. Stat. Phys.* **1998**, *90*, 1481−1487.
(22) Crooks, G. E. *Phys. Rev. E* **2000**, *61*, 2361−2366.
(23) Gapsys, V.; Seeliger, D.; de Groot, B. *J. Chem. Theory Comput.* **2012**, *8*, 2373−2382.
(24) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *101*, 8577−8593.
(25) Wolf, D.; Keblinski, P.; Phillpot, S.; Eggebrecht, J. *J. Chem. Phys.* **1999**, *110*, 8254−8282.
(26) Anwar, J.; Heyes, D. *J. Chem. Phys.* **2005**, *122*, 224117.
(27) Marsili, S.; Signorini, G. F.; Chelli, R.; Marchi, M.; Procacci, P. *J. Comput. Chem.* **2010**, *31*, 1106−1116.
(28) Procacci, P.; Paci, E.; Darden, T.; Marchi, M. *J. Comput. Chem.* **1997**, *18*, 1848−1862.
(29) Shirts, M.; Pitera, J.; Swope, W.; Pande, V. *J. Chem. Phys.* **2003**, *119*, 5740−5761.
(30) In the formulation of eq 2, we have implicitly assumed the so-called "tin-foil" boundary conditions: the Ewald sphere is immersed in a perfectly conducting medium and, hence, the dipole term on the surface of the Ewald sphere is zero. See deLeeuw, S. W.; Perram, J. W.; Smith, E. R. *Proc. R. Soc. London, Ser. A* **1980**, 373:27.
(31) See, for example, the GROMACS manual and the tutorial for alchemical calculations: *Hands-on tutorial Solvation free energy of ethanol* available at http://www.gromacs.org (accessed 04/04/2014). For NAMD, see the tutorial: *In silico alchemy: A tutorial for alchemical free-energy perturbation calculations with NAMD* available at http://www.ks.uiuc.edu (accessed 04/04/2014).
(32) Hummer, G.; Pratt, L.; Garcia, A. *J. Chem. Phys.* **1995**, *99*, 14188−14194.
(33) Procacci, P.; Marchi, M. *J. Phys. Chem.* **1996**, *100*, 10464−10469.
(34) Procacci, P.; Marsili, S.; Barducci, A.; Signorini, G. F.; Chelli, R. *J. Chem. Phys.* **2006**, *125*, 164101.
(35) Lechner, W.; Oberhofer, H.; Dellago, C.; Geissler, P. L. *J. Chem. Phys.* **2006**, *124*, 044113.
(36) Sivak, D. A.; Chodera, J. D.; Crooks, G. E. *Phys. Rev. X* **2013**, *3*, 011007.
(37) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926−935.
(38) Wang, J.; Wolf, R.; Caldwell, J.; Kollman, P.; Case, D. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(39) Singh, U.; Kollman, P. *J. Comput. Chem.* **1984**, *5*, 129−145.
(40) Tomasi, J.; Mennucci, B.; Cancés, E. *J. Mol. Struct.: THEOCHEM* **1999**, *464*, 211−226.
(41) Nosé, S. *Mol. Phys.* **1984**, *52*, 255−268.
(42) Marchi, M.; Procacci, P. *J. Chem. Phys.* **1998**, *109*, 5194−5202.
(43) Tuckerman, M.; Berne, B. J. *J. Chem. Phys.* **1992**, *97*, 1990−2001.
(44) Abraham, M. H.; Whiting, G. S. *J. Chem. Soc., Perkin Trans.* **1990**, *2*, 291−300.
(45) Li, J.; Zhu, T.; Hawkins, G. D.; Winget, P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G.; Abraham, M. H.; Whiting, G. S. *Theor. Chem. Acc.* **1999**, *103*, 9−63.
(46) Sangster, J. *J. Phys. Chem. Ref. Data* **1989**, *18*, 1111−1227.
(47) Chelli, R.; Procacci, P. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1152−1158.
(48) Cramer, C.; Truhlar, D. *Chem. Rev.* **1999**, *99*, 2161−2200.
(49) Pierotti, R. *Chem. Rev.* **1976**, *76*, 717−726.
(50) Goncalves, P. B.; Stassen, H. *Pure Appl. Chem.* **2004**, *76*, 231−240.
(51) Procacci, P.; Marsili, S. *Chem. Phys.* **2010**, *375*, 8−15.