PCCP

Accepted Manuscript

This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the author guidelines.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard Terms & Conditions and the ethical guidelines, outlined in our author and reviewer resource centre, still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/pccp

Journal Name

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Solvation Free Energies for Periodic Surfaces: Comparison of Implicit and Explicit Solvation Models†

Stephan N. Steinmann,*[a] Philippe Sautet,[a,b] and Carine Michel[a]

Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

The evaluation of solvation energies is a great challenge. We focus here on an organic molecule chemisorbed at a metal-liquid interface, as a prototypical system, essential in tribology, electrochemistry and heterogeneous catalysis. We compare an established implicit solvation scheme with a strategy based on molecular mechanics (MM) free energy perturbation (FEP) seeded by QM computations. First, we benchmark the approaches against experimental hydration energies of standard (organic) molecules and find acceptable errors in the order of 0.06 eV (1.3 kcal/mol). Then, the impact of various parameters on the solvation energy of an adsorbate have been assessed on a typical system of interest, levulinic acid adsorbed at a Ru(0001)/water interface. We identify the need for dipole corrections or symmetric slabs when including solvation effects on metallic surfaces. The MM-FEP scheme is revealed to be as reliable as the implicit solvent for water. In the case of levulinic acid, both PCM and MM-FEP agree that the bulk solvation effect is not sufficient to change the adsorption mode from bidentate to mono-dentate, despite the fact that the COOH group is desolvated in the bidentate case. MM-FEP has the great advantage of being more easily generalized to other solvents and to be further improved which will be particularly useful to study solvent and (counter-)ion effects on interfacial reactions.

1 Introduction

The adsorption/desorption processes of organic molecules and ions at solid-liquid interfaces are key in heterogeneous catalysis, electrochemistry and tribology. The most extreme example is probably the electrochemical double layer where the electrolyte plays a major role in the competition between the adsorption of reactants and desorption of products. Another illustration is, as it has been shown particularly for biomass-related heterogeneous catalysis, that liquid water can modify significantly the catalytic activity of metallic supported catalysts compared to the more traditional gas- phase conditions.[1–3] However, first principles modeling of reactions under such conditions is still a great challenge, especially when the adsorbate is chemisorbed, forming covalent bonds with the catalyst.

For around a decade, several attempts have been made to include the influence of water on (electro-)catalytic processes based on micro-solvation at the first principles based density functional theory (DFT) level. In a pioneering study of the de-protonation of acetic acid at water/Pd(111) interface, Neurock and co-workers included several water molecules in their surface model.[4] This approach was successfully used to assess the impact of water molecules on catalytic steps by several groups. The explicit solvent included varies from one or few selected water molecules[5,6] to ice-like water[7–10] or the optimization of several snapshots extracted from a short *ab initio* molecular dynamics simulations.[1,11]

These schemes focus on surface reaction energies, i.e., they allow a reasonably balanced description of the reaction energy of reactions of the type A* $\longrightarrow$ B* where the asterisk indicates surface adsorbed species. However, the adsorption of reactants from solution and desorption of products back into solution is not easily accessible in schemes based on microsolvation.

Another strategy has been recently proposed by Heyden and coworkers.[12] In their implicit solvation model for surfaces (iSMS), reaction energies are computed with DFT applying periodic boundary conditions and solvation effects are estimated separately on large metallic clusters using a polarizable continuum model (PCM) for the solvent description.[12] PCM, which accounts

[a] *Univ Lyon, Ecole Normale Supérieure de Lyon, CNRS Université Lyon 1, Laboratoire de Chimie UMR 5182, 46 allée d'Italie, F-69364, LYON, France; E-mail: stephan.steinmann@ens-lyon.fr*

[b] *Department of Chemical and Biomolecular engineering, University of California, Los Angeles, Los Angeles, CA 90095, United States*

† Electronic Supplementary Information (ESI) available: It contains the full list of molecules and their solvation free energies as computed herein. Data comparing FEP computations based on DFT with and without the dipole correction and the influence of the geometry on the adsorption energy is presented as well. The effect of changing the energy of adsorption of H_2O on Ru(0001) on solvation energies is discussed. See DOI: 10.1039/b000000x/

‡ Computational resources generously provided by the mesocenter PSMN. This work was granted access to the HPC resources of CINES and IDRIS under the allocation 2014-080609 made by GENCI.

for the electrostatic interaction between the solute and the solvent, is well-established in molecular chemistry[13] but less common for metal/liquid interfaces[14,15] since it has only recently become available to the public for periodic boundary conditions.[16] PCM can not describe any direct solvent effects and might be inaccurate for specific hydrogen bonds,[17] which makes the combination of micro-solvation with PCM particularly attractive.[2,18–21] In aprotic solvents, the non-electrostatic solvent effects can become dominating, which is particularly challenging for PCM.[22] In subsequent work, Heyden turned the iSMS approach into an explicit solvent scheme, dubbed eSMS.[23] Borrowing the idea of Yang and co-workers to optimize geometries on an approximate QM/MM free energy surface,[24] the cluster is no longer immersed in a PCM, but surrounded by explicit solvent, treated by molecular mechanics (MM). However, this is still in the perspective of surface reaction energies, rather than the assessment of adsorption/desorption events.

An explicit solvent molecular dynamics simulation coupled with Free Energy Perturbation, FEP,[25] is a "logical" improvement along these lines, as it describes the physics of hydrogen bonds, hydrophobic and cavitation effects and is equally well adapted for "isolated" molecules in solution and periodic surfaces. FEP is popular in bio-molecular simulations to compute free energy differences in order to obtain binding or solvation free energies of drugs.[26] Furthermore, enzyme catalysis[27] and solvation effects for organic reactions have been extensively studied within the framework of QM/MM-FEP, where reactions are described in the QM subsystem, while solvation is treated by MM.[28] The size (both in lateral and out of plane direction) of a solid-liquid interface limits the usefulness of "brute force" ab initio molecular dynamics. Furthermore, the reliable description of the evolution of the solvation shell requires extensive sampling that is clearly out of reach at a first principles level. Hence, we herein compare two more approximate methods to include the bulk solvent effect on any adsorption or surface reaction energy.

Our present development computes solvation energies in an ONIOM like[29] approximation avoiding any cluster computations and the associated problems of how to choose the size of these clusters or how to correct for the limited cluster size. More importantly, the periodicity allows us, in principle, to assess coverage effects, which is beyond the scope of eSMS. The system consists of two parts: the quantum mechanical subsystem (QM) and the solvent, which is treated by molecular mechanics (MM). The QM subsystem describes the interaction of the molecule with the surface or, in the case of molecules, describe relative energies of different conformations in vacuum. The interaction between the QM and the MM subsystem is evaluated at the MM level. We call this scheme, which is fully detailed in the next section, MM-FEP. We reserve the acronym QM/MM-FEP for a (self-consistent) electrostatic embedding approach, which is beyond the scope of the current paper. An implicit solvent is compared to the MM-FEP scheme. In particular, a polarized continuum model (PCM) has been implemented in the periodic boundary DFT code VASP under the name VASPsol by the group of Hennig.[16] The implicit solvent also allows to incorporate bulk solvent effects on the QM geometry and electronic structure, which should improve the quality of the MM-FEP estimates.

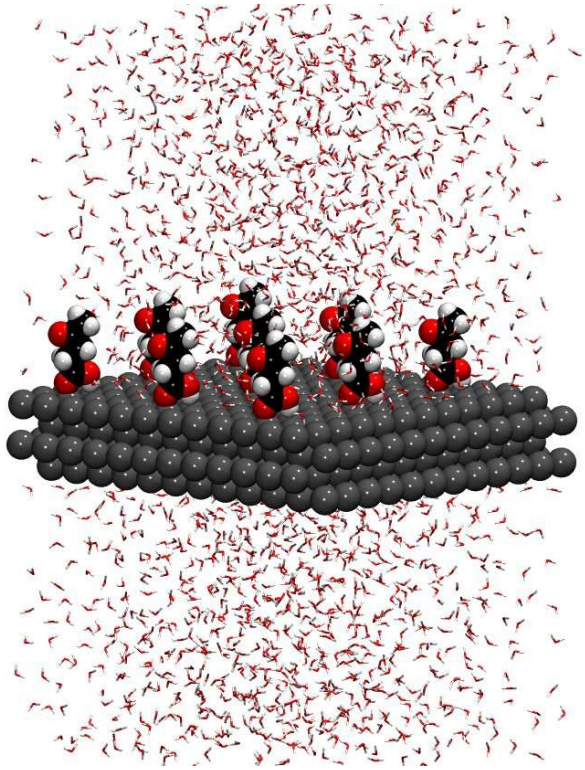

Fig. 1 Model of levulinic acid chemisorbed at the water/Ru(0001) interface. The solute (levulinic acid and the Ru atoms) is represented by van der Waals spheres and the water solvent molecules with lines. The depicted system corresponds to the unit cell for MM computations. The solute is kept in its PBE-dDsC chemisorption geometry; water is described with the TIP3P force field; the water/solute interaction is provided by a mixed TIP3P — QM-UFF description: the electrostatic interactions are TIP3P – QM and the Lennard-Jones interactions are TIP3P–UFF.

After setting the stage by describing in detail the approximations involved in the assessed schemes, we first benchmark the implicit solvent (PCM) and our MM-FEP for hydration energies of small molecules. Then we assess the influence of solvation on the adsorption energy of an organic molecule on a metallic surface. We have chosen the adsorption of a bi-functional molecule, levulinic acid (LA), at the water/Ru(0001) interface as a typical test case (see Fig. 1). Levulinic acid is an essential platform molecule in cellulosic biomass valorization that can be converted to molecules of interest using Ru supported catalysts in water such as γ-valerolactone.[30] Two chemical functions can interact with either the solvent or the metal catalyst, namely a ketone and a carboxylic acid and the preferred adsorption mode might be impacted by the presence of water.

2 Theory

2.1 Decomposition of Energies in Solvent

In the context of continuum solvation models, it is customary to divide the influence of the solvent into two effects:[18–20] direct participation in the reaction (or cooperative co-adsorption) and indirect influences. Hence, the Gibbs free energy of solvation

($\Delta_s G$) is divided into two terms

$$\Delta_s G = \Delta_s G_{direct} + \Delta_s G_{indirect} \quad (1)$$

where we only introduce the approximation that solvent molecules can be classified as "strongly" or "weakly" bound. This distinction is not always obvious. However, a well defined scheme has been put forward that minimizes the introduced error by assessing the differential solvation energy induced by including each additional explicit solvent molecule.[18]

The direct participation occurs, for instance, in proton-shuttling mechanisms and requires explicit water molecules to be simulated on equal footing with the reactant, i.e., at the explicit QM level. Care has to been taken in order to obtain meaningful results, first because of the number of solvent molecules needs to be chosen carefully[18] and second because the same standard state needs to be applied to the explicit and implicit solvent contributions.[20]

$\Delta_s G_{indirect}$ accounts for the remaining "indirect", bulk solvation effects in the context of solvent embedded QM methods, this term can be further decomposed:

$$\Delta_s G_{indirect} = \Delta_s E_{pol} + \Delta_s G_{inter} + \Delta_s G_{sps} \quad (2)$$

where $\Delta_s E_{pol}$ accounts for the polarization of the electronic wave function in the presence of a solvent. $\Delta_s G_{inter}$, on the other hand, represents the interaction energy between the (polarized) solute and the solvent and includes the solvent re-organization energy. Finally, $\Delta_s G_{sps}$ accounts for non-trivial entropic terms due to translation, rotation and vibrations of the solute that differ between the gas- and solution-phase.

2.2 Polarizable Continuum Model

PCM refers here to the VASPsol implementation, where

$$\Delta_s G^{PCM}_{inter} = -\int dr \varepsilon(\vec{r}) \frac{|\nabla\phi(\vec{r})|^2}{8\pi} + \tau \int dr |\nabla S(\vec{r})| \quad (3)$$

The first term is a generalized Poisson equation for the electrostatic interaction and the second term accounts for cavitation energy. $\varepsilon(\vec{r})$ represents the relative permittivity which depends on the electron density and $\phi(\vec{r})$ is the total electrostatic potential. τ is the surface tension and $S(\vec{r})$ the cavity shape function, which is given by

$$S(r) = \frac{1}{2}\mathrm{erfc}\left\{\frac{\ln(\rho(r)/\rho_c)}{\sigma\sqrt{2}}\right\} \quad (4)$$

where ρ is the electron density and ρ_c is the isodensity value around which the cavity is created. σ modulates the diffuseness of the cavity around ρ_c, i.e., how fast "inside" switches to "outside" as a function of the electron density. Due to the derivative of the cavity shape function, the cavitation energy is at the origin of significant numerical instabilities, especially when uniform grids are used which is the case for VASP. Since the cavitation energy usually only gives a small contribution (in the order of 0.05 eV), it is often neglected.[16,31] The model contains three empirical parameters (ρ_c, σ, τ) that have been fitted to reproduce the reference data,[32] comparing the computed $\Delta_s G_{indirect} = \Delta_s E_{pol} + \Delta_s G_{inter}$ with the experimental solvation Gibbs free energy. However, the last term in Eq. 2, i.e., the the modification of the phase-space accessible to the solute when immersed in a solvant, is unaccessible within PCM.

2.3 Free Energy Perturbation

The aim of the presented MM-FEP scheme is to describe the same physics as the PCM but in a way that allows to systematically improve the scheme. Furthermore, the MM-FEP scheme can be applied to a wider range of systems with a similar expected accuracy, in particular also to (counter-)ion effects on interfacial reactions such as electrocatalysis and corrosion, which are tricky to describe by a PCM. At the end of this section we also discuss extensions for the present FEP that will improve its predictive power.

Free energy perturbation is a well established method and the interested reader is encouraged to consult the excellent reviews on the topic.[26–28] However, for sake of completeness, we give the most important equations in the following. The basic idea of the free energy perturbation methods is to obtain the free energy difference between state A and state B as a Boltzmann-weighted average between two potential energy functions U_0 and $U_0 + \Delta U$, evaluated for a representative ensemble of system configurations. In practice, the size of the perturbation is reduced by doing the transformation step-wise, i.e., the interval from 0 to 1 is divided into n "windows" and the difference between state A and B becomes a function of the coupling parameter λ:

$$U(\lambda) = U_0 + \Delta U(\lambda) \quad (5)$$

The coupling parameter runs from 0 (initial/reference state) to 1 (final state). Since molecules interacting with the environment at $\lambda = 0$ and non-interacting at $\lambda = 1$ fully contribute to the initial state but not to the final state, we refer to them as "disappearing", while "appearing" molecules have the opposite characteristics. Typically, one assesses a free energy difference of solvation between two closely related molecules by gradually replacing the interactions of molecule A with the solvent by the interactions of molecule B with the solvent.

The free energy difference between the two states is computed as[25]

$$\Delta G = -k_B T \sum_i \ln \left\langle \exp\left[-\frac{\Delta U(\lambda_{i+1}) - \Delta U(\lambda_i)}{k_B T}\right]\right\rangle_{\lambda_i} \quad (6)$$

where k_B is Boltzmann's constant and the brackets indicate thermodynamic averages obtained from molecular dynamics or Monte Carlo simulations. In order to improve statistical accuracy, the simple-overlap sampling formula allows to combine two trajectories, one in the "forward" (fw) direction ($0 \to 1$) and on in the "backward" (bw) direction ($1 \to 0$). The corresponding free energy estimate is given by:

$$\Delta G = -k_B T \sum_i \ln \frac{\left\langle \exp\left[-\frac{U(\lambda_{i_{fw}+1}) - U(\lambda_{i_{fw}})}{2k_B T}\right]\right\rangle_{\lambda_{i_{fw}}}}{\left\langle \exp\left[-\frac{U(\lambda_{i_{bw}}) - U(\lambda_{i_{bw}-1})}{2k_B T}\right]\right\rangle_{\lambda_{i_{bw}}}} \quad (7)$$

where i goes from 1 to the total number of windows (N), i_{fw} it

the ith window of the forward and i_{bw} is the $(N-i+1)$th window of the backward run.

The key question for our simulations is the definition of $U(\lambda)$. To enable an efficient sampling of solvent phase-space, we describe the solvent by the classical TIP3P water model.[33] In all our simulations, this corresponds to U_0 of Eq. 5. The solute is described through the combination of QM derived point charges[34] and UFF Lennard-Jones parameters,[35] which are, to the best of our knowledge, the only ones available for most elements in the periodic table. Hence, the interaction between the solute and the solvent is provided by a mixed TIP3P — QM-UFF description: the electrostatic interactions are TIP3P — QM and the Lennard-Jones interactions are TIP3P — UFF. Following the study of Jorgensen and co-workers, the point charges for the QM regions are derived from slightly modified CM5 charges,[34,36] which initially have been developed to yield accurate gas-phase dipole moments and are used in the SM12 implicit solvation model.[37] Please note that the Lennard-Jones parameters for water are similar for TIP3P and UFF: the O – O equilibrium distance is virtually the same. The well depth of the O – O interaction in TIP3P includes the vdW interaction with the hydrogens, while UFF assigns individual contributions that sum up to about the same interaction energy. For the hydration energy of water we obtained similar values for a TIP3P water molecule and our UFF-CM5 water molecule: 0.26 and 0.24 eV in comparison to 0.27 eV, its experimental value.

A certain degree of "(self-)consistency" between the computation of the solute and the solvation free energy is desirable in order to capture the subtle coupling between $\Delta_s E_{pol}$ and $\Delta_s G_{inter}$. However, only the dominant term ($\Delta_s G_{inter}$) can be approximated by MM-FEP. On the other hand, the polarization component $\Delta_s E_{pol}$ is naturally included in the self-consistent implementations of implicit solvent models. Hence, the two approaches can be combined, to benefit from the strengths of each method.

In contrast to the situation for organic molecules,[38] to the best of our knowledge no reliable, extremely fast semi-empirical electronic structure method is available for transition metal surfaces. Furthermore, in contrast to attempts to characterize the structural properties of interfacial water by classical molecular mechanics simulations,[39] low-cost force-field methods are generally not available for reactive adsorption events on metal surfaces, which are the main systems of interest herein. Therefore, the QM region is kept frozen during all our MM computations, very similar to what has been done for reactions in solution.[33]

Although our MM-FEP scheme relies on several approximations, it has the same merit as the PCM: it can be applied to all kinds of systems and, as we demonstrate hereafter, it provides a similar accuracy as the implicit solvent. The advantage is, however, that MM-FEP can be systematically improved and can "easily" account for all the relevant physics, especially when other solvents than water pure are involved. In particular, MM-FEP naturally accounts for the size of solvent molecules and ions (rather than an infinitesimal small point charges), which need to be included for more realistic simulations of processes at electrode surfaces.[40,41] The improvements we are envisioning are, on the one hand, self-consistency between the QM and the MM computations (especially the polarization of the solute) and on the other hand an improved MM description, which ultimately would allow to include the changes in accessible phase-space of the solute, but for which reasonably accurate force fields are required.

2.4 Impact of the Solvation on Adsorption Energies

Our target in this work is to assess the impact of the water solvent on the adsorption reaction of a given substrate (here levulinic acid, named LA in the following) on a surface (here Ru(0001)). This process can be described by the following reaction:

$$\mathrm{LA}^{\mathrm{solv}} + \mathrm{Ru(0001)}^{\mathrm{solv}} \longrightarrow \mathrm{LA\ @\ Ru(0001)}^{\mathrm{solv}}$$

Since the free energy of adsorption in vacuum is much more accessible than the one in solvent, it is customary to write the Gibbs free energy of adsorption in solution $\Delta_a G^{solv}$ as the following sum

$$\Delta_a G^{solv} = \Delta_a G^{vac} + \Delta_s \Delta_a G \tag{8}$$

where $\Delta_a G^{vac}$ is the Gibbs free energy of adsorption in vacuum and $\Delta_s \Delta_a G$ corresponds to the variation of the solvation Gibbs free energy ($\Delta_s G$) along the adsorption process (Δ_a), i.e., the solvation energy of the surface with the adsorbed molecule, minus the solvation energy of the separated molecule and surface (see Fig. 2). Note that we use the index "s" and "a" to symbol for a reaction energy (Δ) for the "solvation" and "adsorption" process, respectively.

Since we have no experimental benchmark data for the solvation effect on adsorption energies, we do not include any terms that affect PCM and MM-FEP to the same extent such as estimates for the modified available phase space ($\Delta_s G_{sps}$). This implies that thermal effects (e.g., entropy of adsorption) are neglected in the present approach, but could be included at least approximately by "standard" statistical mechanics formulas. More rigorous treatments than the ideal gas, harmonic oscillators exist, but necessitate significantly higher computational resources.[42,43] In short, in our study, we consider the approximation $\Delta_a G^{vac} \approx \Delta_a E^{vac}$

In addition, as discussed in section 2.1, $\Delta_s G$ can be decomposed in direct and indirect terms, leading to

$$\Delta_s \Delta_a G = \Delta_s \Delta_a G_{direct} + \Delta_s \Delta_a G_{indirect} \tag{9}$$

We do not expect any significant direct solvent effects for the adsorption of levulinic acid on Ru(0001) and therefore do not include any explicit water molecules in the QM subsystem. In other words, $\Delta_s \Delta_a G_{direct}$ is neglected. Hence, the Gibbs free energy of adsorption in solution considered in this article reads

$$\Delta_a G^{solv} = \Delta_a G^{vac} + \Delta_s \Delta_a G_{indirect} = \Delta_a E^{vac} + \Delta_s \Delta_a E_{pol} + \Delta_s \Delta_a G_{inter} \tag{10}$$

The first term can be easily computed based on periodic DFT computations and the strategies of refinements are well-known. In particular, we are using a dispersion corrected GGA functional, which has been shown to provide robust results.[44] The second term, $\Delta_s \Delta_a E_{pol}$ will be shown to be a rather small contribution in the PCM. As detailed above, it is not re-evaluated within the FEP scheme. However, it could be included in a computationally significantly more expensive QM/MM-FEP scheme, which is beyond

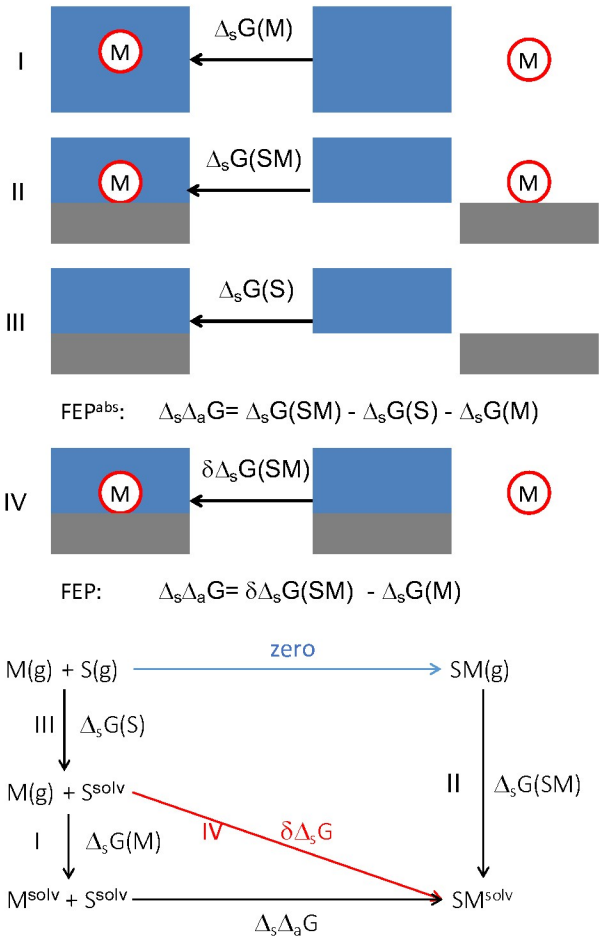

Fig. 2 Schematic of the two possibilities to assess $\Delta_s\Delta_a$G, evaluated by FEP, which only accounts for the variation of the solvation free energy upon adsorption of the molecule on the surface. (top): "M" and "S" refer to the molecule (red circle) and surface (gray rectangle), respectively, while the solvent is depicted by the blue rectangle. Steps II-I-III define the scheme dubbed FEP^{abs}, where the solvation energy for each species needs to be computed separately. Alternatively, steps IV-I directly give the solvation energy contribution to adsorption. (bottom): Corresponding thermodynamic cycles, with roman letters indicating the corresponding processes depicted in the top scheme. The FEP contribution to the adsorption process in vacuum evaluates to strictly zero (horizontal blue line in the thermodynamic cycle) and its value is evaluated at the DFT level. The rectangle represents the "standard" cycle (FEP^{abs}), while the triangle avoids the absolute solvation free energies for metallic slabs by assessing process IV, in red.

the scope of the present paper, but parallels the approach that has been published during the revision of this paper.[45] Herein, we focus on different approximations for the dominant $\Delta_s\Delta_a G_{inter}$ term, which describes the intermolecular solvent–solute contribution to the free energy of solvation, including the solvent re-organization energy.

Evaluating the solvation energy contribution to the adsorption energy, $\Delta_s\Delta_a G_{inter}$, within the PCM framework is straight forward:

$$\Delta_s\Delta_a G = \Delta_s G(SM) - \Delta_s G(S) - \Delta_s G(M) \qquad (11)$$

where S, M and SM represent the individual surface, molecule and the surface with the adsorbed molecule, respectively. The same scheme can be applied to the FEP computations, where we call it FEP^{abs} (see processes I-III in Fig. 2). However, the absolute solvation free energies for extended systems $(\Delta_s G(SM), \Delta_s G(S))$ are difficult to converge. Hence, we have devised an alternative, avoiding these two terms: when the initial state is defined as metal surface in the solvent and the molecule in the gas-phase and the final state is the molecule adsorbed on the solvated surface, we can directly assess $\delta\Delta_s$G=Δ_sG(S) - Δ_sG(SM) (see process IV in Fig. 2 and the corresponding triangular thermodynamic cycle). This perturbation is, obviously, much smaller than completely desolvating an entire metal surface and hence converges rather easily.

3 Methods and Computational Details

The presented computational strategy consists of three distinct phases: i) DFT computations, ii) setting up the molecular mechanics computations, iii) running the free energy perturbation computations. These steps are described in the following subsections.

3.1 DFT Computations

All DFT computations are performed with a locally modified version of VASP 5.3.5,[46,47] patched to include the implicit solvent module VASPsol[16] and including the recently implemented[44] dispersion correction dDsC[48] that has become available to the public in version VASP 5.4.1. The electronic structure is described at the PBE level,[49] with an energy cut-off of at least 400 eV for the plane-wave basis set. The electron–ion interactions are described by the PAW formalism.[50,51]

The metal surface is modeled as a p(4×4) Ru(0001) slab with a lattice constant of 2.70 Å and a thickness of 4 layers, with the two lowest ones fixed in the bulk position. One levulinic acid molecule is chemisorbed on this slab, which corresponds to a low coverage situation. Such a coverage appears reasonable considering the modest chemisorption energy, but determination of optimal coverage is beyond the scope of this paper. The vacuum layer completes the z-direction to reach 40.5 Å. Dipole corrections in z-direction are included in some computations during the SCF procedure, which is indicated by a subscript "dip" when appropriate. Molecules are computed in a cubic box of a length of 20 Å. All computations are performed at the Gamma point. Integrating the Brillouin zone more accurately or using thicker, symmetric slabs changes the DFT contribution (relative energy changes of ~ 0.1

eV), but leaves the solvation energy unchanged compared to the dipole corrected results (see ESI†).

All geometries are optimized to reach a gradient smaller than 0.02 eV/Å with wave functions converged to $1{\cdot}10^{-6}$ eV. In most of the computations the precision setting of VASP is set to "normal". For testing purposes, we have also performed computations with more stringent accuracy settings, particularly, the FFT grids are improved from 54×54×98 to 90×90×160 grid points for the surface model by using "accurate" precision settings, and a plane-wave cut-off of 600 eV is used. The automatic optimization of the real-space projection operators is used. If not stated otherwise, the default implicit solvent settings are used, except that the cavitation energy is not included, i.e., the cavity surface tension τ is set to zero to increase the numerical robustness. A 2nd order Methfessel-Paxton smearing is applied to the metallic surfaces (width of 0.2 eV). Reported charges are based on a Hirshfeld analysis,[52] which is a by-product of the dDsC dispersion correction.

3.2 Setting up the MM Force Field

The CM5 charges, which have been developed for the SMD implicit solvent model,[36] are evaluated based on the Hirshfeld charges given by VASP, transformed by the cm5pack utility available from the University of Minnesota,[53] and modified to take into account the recommended scaling factor of 1.27 for explicit solvent computations.[34] This scaling value is similar to the one suggested by Grimme for his QM derived force field.[54] The transformation from Hirshfeld to CM5 charges depends on the geometry. However, cm5pack is not a periodic code. Therefore, the DFT cell is first replicated in all 3 directions in order to minimize finite size effects and then only the results of the central unit are exploited, resulting in charges for all atoms in the QM region.

For the MM computations of surface bound species, the relatively small DFT unit-cell has been replicated to a 3×3 super super-cell, corresponding to a formally p(12×12). This allows for a better description of the bulk solvent region due to less "enforced" periodicity, increases the statistical averaging and avoids technical issues with MM programs that are intended for large unit cells. The Lennard-Jones (LJ) parameters are taken from UFF.[35] The initial solvent distribution is obtained from the predefined TIP3P[33] box with about 35 Å of water surrounding the system.‡ This generates an orthorhombic box that does not account for the periodicity of the surface, since the QM system was treated like a molecule, resulting in water surrounding the entire system. From the thus generated orthorhombic box with in plane angles of 90 degrees, the "original", hexagonal (rhombus of length 32.4 Å with 60/120 degrees in-plane angles) cell is cut out by removing all the solvent which is not on top of the metal surface. This restores the correct periodicity of the entire system. Typically, there are about 1900 water molecules surrounding the metal surface. The final unit cell for one of the adsorption modes is represented in Fig. 1.

‡ The system setup is performed using tleap of the AmberTools.

3.3 Free Energy Perturbation

All MM computations are performed with NAMD 2.9[55]. In addition to being highly efficient, NAMD allows us to fix the QM region in its starting position during the entire FEP computation and to still apply a barostat in the out-of-plane direction. In other words, the QM region is kept frozen during the evaluation of the solvation free energy, a strategy well known from applications to molecular systems.[56] This also avoids the necessity of periodic bonded interactions, which would be necessary otherwise in order to reliably describe a metal surface.

The electrostatics are evaluated according to the Particle Mesh Ewald summation, with a 4th order interpolation scheme and a grid spacing of 1 Å. The temperature (300 K) is controlled by a Langevin thermostat with a damping coefficient of 1 ps^{-1}. The Langevin barostat for keeping the pressure constant to 1 bar is used with a piston period of 100 fs and a decay of 150 fs. The LJ terms are cut off at 11 Å, after having been switched to zero, starting from 9 Å. The water geometry is kept fixed and a time step of 2 fs is applied. To fully exploit the efficiency of NAMD and its multiple-time step algorithm,[57] the full electrostatics are evaluated every second time step, while the LJ interactions are evaluated every time step. Tests have shown that evaluating the full electrostatics every time step barely affects the results (see ESI†).

Prior to the FEP simulation, the system is minimized for 2000 steps, then heated from 100 K to 300 K within 10 ps and then equilibrated during 200 ps.

The FEP computations are performed with a variable number of equally spaced windows (at least 10), each of the length of 200 ps, of which the first 100 ps are used to equilibrate. Data presented in the ESI† for windows of 500 ps with 250 ps equilibration each, yield the same results, demonstrating a well converged setup. Note, that this is especially true for the surfaces, since the "primitive" system is repeated 9 times, yielding statistics that are roughly equivalent to 0.9 ns per window. To avoid the "end-point catastrophe" (diverging interactions) in the FEP computations, the default settings are applied, i.e., a soft-core potential is used for the LJ potential.[58,59] The electrostatic interactions are scaled twice as fast than the LJ parameters (e.g., "incoming" particles have partial charges starting from $\lambda = 0.5$).[60] Furthermore, the internal interactions of the solute do not contribute to the MM-FEP free energy estimates, which is achieved through "decoupling".[61] Note, that these "internal" energy differences are fully accounted for in the QM computations and therefore included in the final energy.

The free energy difference is obtained from joining the forward and backward run through the simple overlap sampling,[62] with outputs written every 50 fs.

The error estimates were obtained through the ParseFEP toolkit.[63] Since these errors were found to be below 0.01 eV for all species, we do not report them herein. For hydration energies of individual molecules the errors are, however, given in the ESI†.

4 Results and Discussion

We will discuss two distinct problems: First, the computation of free energy of solvation $\Delta_s G$ for small molecules, for which

we have compared the PCM solvation energies to results of our straight forward MM-FEP scheme. Second, we present the adsorption Gibbs energies $\Delta_a G^{solv}$ of levulinic acid on Ru(0001) and the impact of the inclusion of the solvation free energies contributions $\Delta_s\Delta_a G$, first as obtained from PCM and then FEP. We discuss the numerical precision and the sensitivity of these quantities to the geometry and to the polarization of the QM system by an implicit solvent. We conclude by discussing some shortcomings of the presented method.

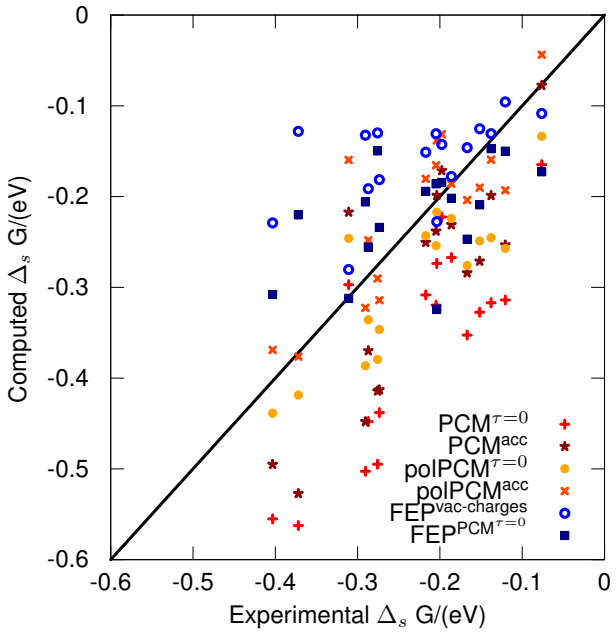

Fig. 3 Computed solvation free energies $\Delta_s G$ of small molecules compared to experimental data (see ESI†for their distributions in form of a histogram around the average error). Two settings are tested for the PCM: PCM$^{\tau=0}$ refers to the use of standard parameters, except that the cavity surface tension τ is set to zero to improve numerical stability and PCM^{acc} uses the default parameters for the PCM, but with increased numerical precision (600 eV plane-wave basis set and more accurate FFT grids). polPCM includes the polarization contribution to the solvation energy, i.e., $\Delta_s E_{pol}^{PCM}$. FEP computations are either based on vacuum charges (FEP$^{vac\text{-}charges}$) or on an electronic structure surrounded by an implicit solvent (PCM$^{\tau=0}$).

4.1 Molecules

In Fig. 3, we compare the hydration energies $\Delta_s G$ obtained with six methods to experiment for 17 molecules. The set of molecules (see ESI† for detailed information) includes typical polar organic molecules such as alcohols, amines and carboxylic acids.

Two different PCM setups are tested: PCM$^{\tau=0}$ (to which we will also simply refer as "PCM", as it is the stable, default variant as implemented in VASPsol [16]) uses default parameters, except that the cavitation energy is neglected ($\tau = 0$), which is needed in some applications for sake of numerical stability.[16,31] As clearly seen, PCM$^{\tau=0}$ overestimates the solvation free energy rather dramatically (mean absolute deviation, MAD of 0.14 eV (> 3 kcal/mol)).§ The cavitation energy is, by definition, positive and its inclusion reduces the solvent affinity. Indeed, including this contribution in PCM^{acc} yields somewhat better results (MAD=0.08 eV). Nevertheless, both schemes do not compare well with experiment. However, as discussed above, the parametrization of the PCM was done taking into account the energetic cost of polarizing the electronic wave function. Hence, a fair comparison assesses rather polPCM ($\Delta_s E_{pol}^{PCM} + \Delta_s G_{inter}^{PCM}$). Indeed, the MAD drops from 0.14 to 0.07 eV for PCM$^{\tau=0}$ and becomes, with 0.04 eV, excellent for polPCM^{acc}. Jorgensen and co-workers reported a MAD of around 0.04 eV for their FEP computations that are based on the sophisticated OPLS-AA force field to describe the Lennard-Jones interactions and the dynamics of the molecules.[34] Let us, nevertheless emphasize that the physics captured by the two schemes (PCM and Jorgensen's FEP) is somewhat different, as the PCM does not account for any dynamic properties (e.g., conformational changes), while the FEP is missing $\Delta_s E_{pol}$. Hence, we conclude that both schemes are accurate due to some error cancellation based on parameter fitting. This error cancellation implies that functional groups or types of molecules (e.g., transition metal complexes) require some benchmarking prior to being confident that predictions are in line with experiments. Very similar conclusions have been drawn from an earlier comparison.[64]

Alternatively, the hydration energies are computed by our MM-FEP approach. As a first approach we use vacuum charges to evaluate the FEP free energy of solvation. The MAD of 0.08 eV indicates that the scheme is reasonable, but not very accurate. Using atomic charges which have been evaluated from a DFT calculation that includes an implicit solvent (PCM$^{\tau=0}$), the performance is improved, yielding a MAD of 0.06 eV. This performance is remarkable considering that the LJ parameters for the FEP computations have not been "tuned" for this application and that in general the LJ parameters need to depend on the chemical environment of the atom.[65] Furthermore, this accuracy is sufficient for most applications we target, i.e., adsorption and reactions on metal surfaces, where other sources of errors (density functional approximation, thermal contributions, surface coverage and direct solvent effects) give much larger uncertainties.

In conclusion, our approach of keeping the DFT geometry fixed during the FEP computation and the application of UFF Lennard-Jones parameters together with atomic charges derived from PCM computations yields reasonably accurate results. Our molecule of interest (levulinic acid) has a solvation energy of about -0.7 eV and -0.5 eV at the PCM$^{\tau=0}$ and FEP level, respectively.

4.2 LA@Ru(0001)

We now turn to our system of interest, the adsorption of levulinic acid (LA) at a liquid water/Ru(0001) interface. For such a sys-

§ The histogram of the errors around the mean error is shown in the ESI†, to represent the scattering of the data. For all the methods, about half of the errors are located within ±25 meV of the mean error, with the remaining errors being quite similarly distributed. Hence, the trend of the MAD is a reasonable descriptor to compare these different methods.

	$\Delta_a E^{vac}$	$\mu^{vac,dip}$	$\Delta_a E^{PCM}$	$\Delta_s\Delta_a G^{PCM}$	μ^{PCM}	$\Delta_a E^{PCM,dip}$	$\Delta_s\Delta_a G^{PCM,dip}$	$\mu^{PCM,dip}$
ket*-COOH*	-1.70	-1.08	-1.11	-0.42	-4.84	-1.60	0.24	-1.55
COOH*	-1.12	-0.58	-1.05	0.19	-2.22	-1.13	0.31	-0.69
ket*	-1.06	-0.77	-0.76	-0.25	-3.36	-0.97	0.04	-1.05
ket*-iso	-0.92	-0.32	-0.42	-0.54	-2.66	-0.64	-0.26	-0.83

Table 1 Components of the PCM$^{\tau=0}$ adsorption free energies in eV: $\Delta_a G^{PCM} = \Delta_a E^{vac} + \Delta_s\Delta_a E^{PCM}_{pol} + \Delta_s\Delta_a G^{PCM}_{inter} = \Delta_a E^{PCM} + \Delta_s\Delta_a G^{PCM}_{inter}$ The "dip" superscript refers to the use of the dipole correction during the SCF procedure. Note that the energy cost to polarize the wave function in the PCM is responsible for $\Delta_a E^{PCM} > \Delta_a E^{vac}$. Hence, even a negative $\Delta_s\Delta_a G^{PCM}_{inter}$ can lead to more weakly adsorbed molecule in implicit solvent than in vacuum. For ease of comparisons, $\Delta_a E^{vac}$ is given as well.

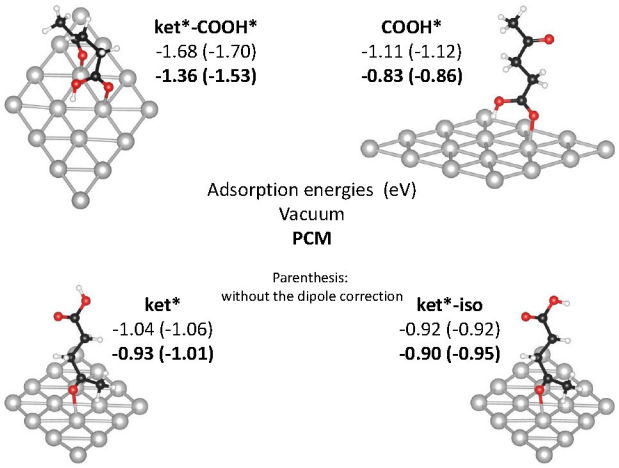

Fig. 4 Representation of the four adsorption modes studied of levulinic acid and the adsorption energies computed by different schemes; numbers in parenthesis refer to the absence of the dipole correction. The first row gives the adsorption energies in vacuum (Δ_aE^{vac}). The second row (bold) gives the adsorption energy when including an implicit solvent (Δ_aG$^{\mathrm{PCM}^{\tau=0}}$), which naturally includes the polarization contribution $\Delta_s\Delta_a E_{pol}$.

tem, reference experimental data is, unfortunately, not available. In this part, we assess the impact of numerical settings on the results we obtain with the PCM and MM-FEP approaches and analyze their consistency. We focus on four adsorption modes (see Fig. 4 for their representations). In the most stable mode, LA forms a loop, so that the two chemical functions interact with the metallic surface (labeled **ket*-COOH***): the ketone is adsorbed on top of a Ru atom through its O atom and the carboxyl group is chemisorbed on a neighboring site in a similar configuration. In the other configurations we considered, the levulinic acid is oriented perpendicular to the surface. In **COOH***, the carboxylic acid is chemisorbed while the ketone is not interacting with the metallic surface. In **ket*** and **ket*-iso**, it is the carbonyl that is chemisorbed, and they differ only by the orientation of the OH bond of the carboxylic acid.

4.2.1 PCM

Fig. 4 depicts the adsorption modes of levulinic acid on Ru(0001) and the associated adsorption energies $\Delta_a G$ as computed by DFT in vacuum and in implicit solvent (PCM$^{\tau=0}$). The components of the adsorption energy ($\Delta_a E^{vac} + \Delta_s\Delta_a E^{PCM}_{pol}$, $\Delta_s\Delta_a G^{PCM}_{inter}$) are reported in Table 1.

Comparing the adsorption energies in vacuum with and without the dipole correction (numbers without/within parenthesis on the first lines in Fig. 4), we could conclude that this correction is not necessary. However, in implicit solvent, we find a significant difference upon including the dipole correction. Inspecting the solvation free energy (Table 1), we find that $\Delta_s\Delta_a G^{PCM}$ (**ket*-COOH***) changes by more than 0.5 eV due to the use of the dipole correction. Computing the dipole moment of the system reveals that it is 1.1 eÅ in vacuum and 1.6 eÅ in PCM when using the dipole correction, but in the absence of this correction, the dipole moment increases to 1.5 and 4.8 eÅ in vacuum and PCM, respectively. The latter value clearly indicates a spurious, macroscopic dipole moment. This statement is confirmed by both, increasing the number of metallic layers (which increases the spurious dipole moment) and the lack of a reasonable convergence as a function of the void separating two slabs, even at box sizes around 100 Å. From these tests (documented in the ESI†) we conclude that the origin of the problem is not a dipole-dipole interaction between cells, but the polarizability of the metallic slab. The artificial polarization occurs because the "bottom" of the slab is supposed to mimic a bulk material that cannot be polarized at all, but the solvent is on both sides of the slab and thus favors a polarization across the metallic slab. The negative $\Delta_s\Delta_a G^{PCM}$ nicely illustrate how the PCM solvates this macroscopic dipole, leading to heavily biased estimates of solvent effects. In order to verify the reliability of the dipole correction for the chosen box size, we have performed simulations with symmetric slabs (see ESI†). The use of symmetric slabs precludes the artificial polarization of the "bottom" slab. These computations have evidenced that the dipole correction is, despite its magnitude, highly accurate. Hence, we conclude that the dipole correction (or the use of symmetric slabs) is absolutely necessary when dealing with metallic (and thus highly polarizable) slabs on which polar molecules are adsorbed, surrounded by a continuum model. In addition, tests (see ESI†) have revealed that the geometry (optimized in vacuum or implicit solvent) does not influence these results, suggesting that re-optimization in solvent would not be necessary.

To gain chemical insights, we consider now the dipole corrected results. Based on the adsorption energies in vacuum, we find that **ket*-COOH*** is by far the most and **ket*-iso** the least stable adsorption mode. Accounting implicitly for the solvent overall weakens the adsorption free energies Δ_aG compared to vacuum, since the isolated molecule is more stabilized by the solvent than when adsorbed on the surface. This reduction strongly depends on the adsorption mode and is particularly small in the **ket*** structures. Therefore, the implicit solvent reduces the range of relative energies between different adsorption modes from 0.76

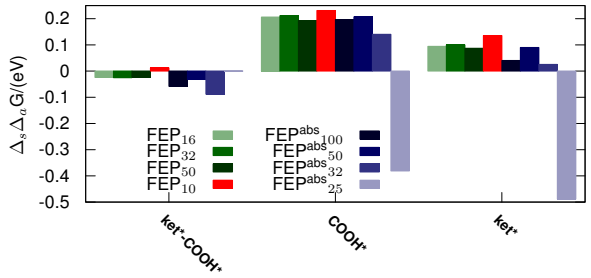

Fig. 5 Computed solvation free energy contributions $\Delta_s\Delta_a G^{FEP}$ to the adsorption of levulinic acid in three distinct adsorption modes according to different FEP setups: The index gives the number of "windows" used, while the superscript "abs" refers to the computation of the free energy of adsorption by two separate computations: one for the naked metal surface and one for the surface with the molecule adsorbed. In the absence of this superscript, the FEP is computed between surface with the molecule adsorbed as the end state and the naked metal surface as the initial state.

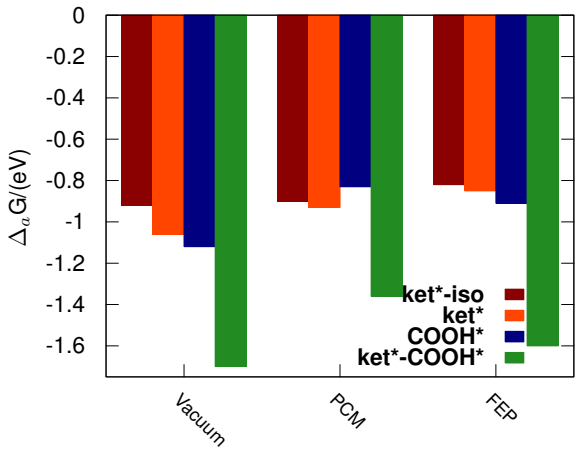

Fig. 6 Comparison of the adsorption free energy in vacuum, an implicit solvent ($PCM^{\tau=0}$) and in explicit solvent (FEP).

to 0.53 eV. This can easily be rationalized considering that desolvating COOH implies a significant destabilization compared to adsorption modes where the acid is still solvated (**ket*** and **ket*-iso**). Furthermore, the **ket*-iso** conformation is particularly stabilized by the solvent, because the O−H proton is more exposed to the solvent and therefore more strongly solvated (see Table 1). We also observe that the bi-adsorbed structure **ket*-COOH*** is slightly better solvated than **COOH***. At first sight, this may seem surprising since the ketone function is desolvated during the **ket*-COOH*** adsorption and not for **COOH***. However, the dipoles generated by the two oxygenated functions are parallel in the loop-like adsorption mode **ket*-COOH*** and not in the straight ones. Since the metallic surface is highly polarizable, this leads to an enhanced charge transfer between the surface (negative) and the molecule (positive) in **ket*-COOH*** (0.28 e) compared with **COOH*** (0.14 e). This is also reflected in the overall (out-of-plane) dipole moment of 1.6 and 0.7 e/Å, respectively. In a nutshell, the larger surface dipole generated in **ket*-COOH*** compensates the desolvation of the ketone function.

4.2.2 FEP

As outlined in the Theory section, we apply two different approaches to assess the solvation energy contribution to adsorption $\Delta_s\Delta_a$G: The straight forward scheme (FEP^{abs}), which assesses the solvation energy of each system individually and the scheme simply dubbed "FEP", which takes a short-cut by determining the change in solvation energy of the surface when adding a molecule $\delta\Delta_s G$. Of course, the two approaches are expected to converge to the same $\Delta_s\Delta_a$G result, especially taking into account error compensation when annihilating the interactions with the entire metal slab. Note, however, that following an ONIOM-like approach[29], in both cases FEP is only used to assess the difference in adsorption energy between vacuum and solution (second term in Eq. 10) and that the major energetic contribution is described at the DFT level (first term in Eq. 10).

Let us now turn to the solvation free energy computed by FEP. Our first concern is precision, before we can assess the accuracy of the proposed scheme. In order to address this point, we systematically increase the number of windows from 10 to 50 for the perturbation where only the adsorbate (dis)appears (denoted FEP) and from 25 to 100 for the absolute free energy of solvation (denoted FEP^{abs}), i.e., the slab and the adsorbate disappear and the solvation energy of the empty slab is assessed independently by a similar computation. We would like to use 10 windows (FEP_{10}), which worked well for the isolated molecules. Therefore, in Fig. 5 we put FEP_{10} in between FEP_{50} and $FEP^{abs}{}_{100}$, to compare it to the best converged results. As expected, $\Delta_s\Delta_a G^{FEP}$ converges very slowly with respect to the number of windows when the perturbation is large (i.e., the entire system is solvated, FEP^{abs}), while with 10 windows we already get results accurate within 0.05 eV when the perturbation is only the adsorbate. This demonstrates the efficiency of our scheme.

4.2.3 Comparison of PCM and FEP for Surfaces

After having discussed the PCM and FEP results individually, we now compare the two approaches, both in term of overall results and the contribution of the solvation energy. The comparison of the adsorption energies in vacuum, $PCM^{\tau=0}$ and FEP are presented in Fig. 6 The major result is that the most stable conformation is **ket*-COOH***, independently on the environment, which rationalizes the observation that vacuum conditions (with micro-solvation) are often enough to gain valuable insight from DFT. The implicit solvent affords one of the "expected" qualitative results, i.e. that **ket*** is more stable than **COOH***, since the carboxylic acid does not need to be completely desolvated. The FEP adsorption free energies, on the other hand, are closer to the trends in vacuum than to the PCM. These differences are small and might depend significantly on the adsorption strength of water on the metal surface, which is inaccurately described by both methods.

While the total adsorption energies are the ultimately relevant quantities, the comparison of $\Delta_s\Delta_a G_{inter}$, i.e., the modification of the adsorption energy due to the solvation energy allows a more detailed comparison of the two methods. The comparison between the PCM and FEP results for $\Delta_s\Delta_a G_{inter}$ is given in Fig. 7,

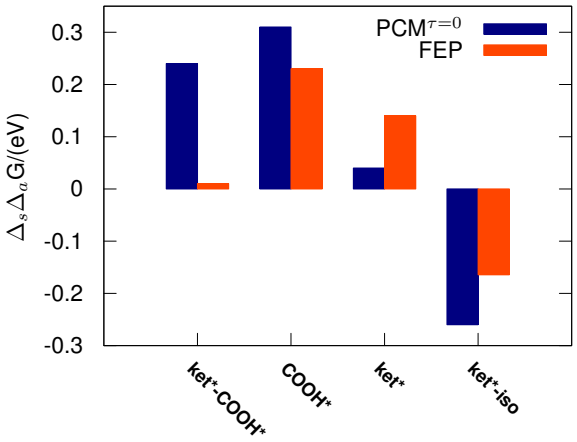

Fig. 7 Comparison of the solvation free energy contribution to adsorption $\Delta_s\Delta_a G$ as computed by the PCM$^{\tau=0}$ and FEP approaches.

i.e., we compare the balance of solvation energy between the adsorbed state on the one hand and the isolated molecule and a clean surface on the other hand (see Fig. 2) . The positive sign of most of these contributions indicates that accounting for the solvation free energy weakens the adsorption free energy, i.e., that the solvent interacts less strongly with the adsorbed state than with the isolated molecule and the clean surface. This is rather intuitive, keeping in mind that the fraction of the molecule interacting directly with the surface is inaccessible for the solvent and assuming that chemisorption does not induce a large polarization of the solvent accessible surface area. As discussed above, this assumption is not valid in the absence of the dipole correction, leading to spurious results. There is only one exception to this "expected" behavior, which is **ket*-iso**, where the proton is better exposed to the solvent than in the other orientations and in solution: the lowest energy solution conformation is derived from the **COOH*** adsorption mode. PCM and FEP give a negative contribution to the adsorption free energy of **ket*-iso**, which results in a relative stabilization of this adsorption mode compared to the vacuum results.¶

Probably the most important message of Fig. 7 is that both approaches give the same sign for the solvation energy contribution for each adsorption mode. This gives credence that the dominating physics captured by the two methods is, as expected, similar. However, there is one rather significant difference: MM-FEP solvates **ket*-COOH*** better than **COOH*** to a larger extent than PCM. This increased relative solvation can be rationalized as follows: in the **COOH*** adsorption mode, a hydrophobic zone close to the metal surface is created (see ESI† for average water distributions). This reduces the access of the solvent to the (polarized) metal surface as measured by the radial distribution function between Ru and the water molecules, which is lower for **COOH*** than for **ket*-COOH***, thereby weakening the solvation energy. For **ket*-COOH***, on the other hand, the molecule adsorbs rather flatly on the surface, which exposes roughly half of the molecular surface to the solvent. Furthermore, this flat adsorption allows the water molecules to better solvate the large surface dipole (*vide supra*). As discussed in the ESI†, the creation of a hydrophobic zone in the "upright standing" conformations could be overestimated in our current model, mostly because water adsorbs too weakly on Ru(0001) with the default UFF LJ parameters (physisorption of -0.03 eV compared to the chemisorption of -0.7 eV according to DFT). However, even when increasing the interaction between water and Ru, the stability ordering of the four conformations remains the same although the difference between **ket*-COOH*** and **COOH*** decreases from 0.7 eV to 0.4 eV (see ESI†). The failure of the UFF Lennard-Jones parameters to capture the chemisorption of water on Ru(0001) cannot come as a surprise, since these parameters are only able to account for weakly bound water molecules (i.e., the "indirect" contributions). Here, the "direct" solvent contributions would mainly describe the chemisorption of water molecules. However, both the FEP and the PCM model neglect this chemisorption of water molecules, providing a fair comparison between the two approaches. In other words, a more complete model would need to consider chemisorbed water molecules on the Ru surface in the QM subsystem in order to capture the "direct" solvent effect.

Our work highlights two points that warrant further studies: first, the impact of the strong adsorption of water on the Ru(0001) surface might change the solvation free energies significantly, given that this effect is basically absent in PCM and the present MM-FEP approach. Second, the solvation free energy might strongly depend on the coverage. MM-FEP is ideally suited for elucidating the coverage effect, as it takes the finite size and chemical shape of the solvent fully into account. Furthermore, the entropy of adsorption is neglected in the present approach, which might introduce a slight imbalance in disfavor of the more flexible adsorption modes **ket*** and **COOH*** compared to the roughly immobile, bi-dentate **ket*-COOH***. These questions can be addressed by straightforward improvements of the current MM-FEP scheme along two axes: on the one hand, improved force fields for the interface would allow for approximations of the adsorption entropy and on the other hand (self-)consistent coupling between QM and MM, with eventually a full QM resampling[66,67] would improve the energetics.

5 Conclusion

Solvent effects for adsorption free energies of organic molecules on (metal) surfaces are a great challenge to assess both experimentally and theoretically. Our target is to evaluate the solvation free energy contribution to chemisorption of organic molecules chemisorbed at the liquid-metal interface. To this end, we describe the chemisorption on a periodic surface by DFT. To assess the solvation free energy, we have compared two schemes: implicit solvation and molecular mechanics based free energy perturbation (MM-FEP). Our results demonstrate the benefit of polarizing the QM subsystem by an implicit solvent prior to MM-FEP, but also the need for dipole corrections or symmetric slabs in order to avoid artificial polarization of the metallic slab. We

¶ The same observation applies to isolated levulinic acid, but the effect is not strong enough to change the preferred conformation: the energy difference is reduced from 0.18 to 0.05 eV when going from vacuum to PCM$^{\tau=0}$.

have shown that the MM-FEP relative solvation energies converge at low to modest computational cost to an acceptable precision, given the approximations involved in the overall scheme. The method is validated for bulk solutions on a set of 17 standard "organic" molecules, delivering accurate results (MAD=0.06 eV). In the test case of bifunctional molecule levulinic acid chemisorption at the Ru-water interface, the comparison of MM-FEP with an implicit solvent model shows that trends are often similar. Most notably, the bulk solvation effect is not strong enough to desorb the carboxylic acid from the surface, despite its desolvation being necessary upon chemisorption. The demonstration that even such a simplistic MM-FEP approach yields valuable results is encouraging, since, in contrast to PCM, the MM-FEP approach can be systematically improved and is well adapted to assess coverage effects and to model solvation effects in ionic liquids and electrolytes as well as solvents at high temperature and pressure.

References

1 B. S. Akpa, C. D'Agostino, L. F. Gladden, K. Hindle, H. Manyar, J. McGregor, R. Li, M. Neurock, N. Sinha, E. H. Stitt, D. Weber, J. A. Zeitler and D. W. Rooney, *J. Catal.*, 2012, **289**, 30.

2 C. Michel, J. Zaffran, A. M. Ruppert, J. Matras-Michalska, M. Jedrzejczyk, J. Grams and P. Sautet, *Chem. Comm.*, 2014, **50**, 12450.

3 C. Michel and P. Gallezot, *Acs Catalysis*, 2015, **5**, 4130.

4 S. K. Desai, V. Pallassana and M. Neurock, *J. Phys. Chem. B*, 2001, **105**, 9171.

5 C. Michel, F. Auneau, F. Delbecq and P. Sautet, *ACS Catal.*, 2011, **1**, 1430.

6 L. Arnadottir, E. M. Stuve and H. Jonsson, *Chem. Phys. Lett.*, 2012, **541**, 32.

7 D. Cao, G. Q. Lu, A. Wieckowski, S. A. Wasileski and M. Neurock, *J. Phys. Chem. B*, 2005, **109**, 11622.

8 S. A. Wasileski and M. J. Janik, *Phys. Chem. Chem. Phys.*, 2008, **10**, 3613.

9 B. N. Zope, D. D. Hibbitts, M. Neurock and R. J. Davis, *Science*, 2010, **330**, 74.

10 D. D. Hibbitts and M. Neurock, *J. Catal.*, 2013, **299**, 261.

11 J. Liu, X.-M. Cao and P. Hu, *Phys. Chem. Chem. Phys.*, 2014, **16**, 4176.

12 M. Faheem, S. Suthirakun and A. Heyden, *J. Phys. Chem. C*, 2012, **116**, 22458.

13 C. J. Cramer and D. G. Truhlar, *Chem. Rev.*, 1999, **99**, 2161.

14 S. A. Petrosyan, A. A. Rigos and T. A. Arias, *J. Phys. Chem. B*, 2005, **109**, 15436.

15 O. Andreussi, I. Dabo and N. Marzari, *J. Chem. Phys.*, 2012, **136**, 064102.

16 K. Mathew, R. Sundararaman, K. Letchworth-Weaver, T. A. Arias and R. G. Hennig, *J. Chem. Phys.*, 2014, **140**, 084106.

17 G. Konig and S. Boresch, *J. Phys. Chem. B*, 2009, **113**, 8967.

18 J. R. Pliego and J. M. Riveros, *J. Phys. Chem. A*, 2001, **105**, 72417247.

19 D. Asthagiri, L. R. Pratt and H. S. Ashbaugh, *The Journal of Chemical Physics*, 2003, **119**, 27022708.

20 V. S. Bryantsev, M. S. Diallo and W. A. Goddard III, *J. Phys. Chem. B*, 2008, **112**, 97099719.

21 H.-F. Wang and Z.-P. Liu, *J. Phys. Chem. C*, 2009, **113**, 17502.

22 B. Mennucci, J. Tomasi, R. Cammi, J. R. Cheeseman, M. J. Frisch, F. J. Devlin, S. Gabriel and P. J. Stephens, *J. Phys. Chem. A*, 2002, **106**, 6102.

23 M. Faheem and A. Heyden, *J. Chem. Theory Comput.*, 2014, **10**, 3354.

24 H. Hu, Z. Lu and W. Yang, *J. Chem. Theory Comput.*, 2007, **3**, 390.

25 R. W. Zwanzig, *J. Chem. Phys.*, 1954, **22**, 1420.

26 P. Kollman, *Chem. Rev.*, 1993, **93**, 2395.

27 J. Aqvist and A. Warshel, *Chem. Rev.*, 1993, **93**, 2523.

28 O. Acevedo and W. L. Jorgensen, *Acc. Chem. Res.*, 2010, **43**, 142.

29 M. Svensson, S. Humbel, R. D. J. Froese, T. Matsubara, S. Sieber and K. Morokuma, *J. Phys. Chem.*, 1996, **100**, 19357.

30 W. R. H. Wright and R. Palkovits, *ChemSusChem*, 2012, **5**, 1657.

31 S. N. Steinmann, C. Michel, R. Schwiedernoch and P. Sautet, *Phys. Chem. Chem. Phys.*, 2015, **17**, 13949.

32 D. Gunceler, K. Letchworth-Weaver, R. Sundararaman, K. A. Schwarz and T. A. Arias, *Modelling and Simulation in Materials Science and Engineering*, 2013, **21**, 074005.

33 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926.

34 J. Z. Vilseck, J. Tirado-Rives and W. L. Jorgensen, *J. Chem. Theory Comput.*, 2014, **10**, 2802.

35 A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, *J. Am. Chem. Soc.*, 1992, **114**, 10024.

36 A. V. Marenich, S. V. Jerome, C. J. Cramer and D. G. Truhlar, *J. Chem. Theory Comput.*, 2012, **8**, 527.

37 A. V. Marenich, C. J. Cramer and D. G. Truhlar, *J. Chem. Theory Comput.*, 2013, **9**, 609.

38 M. P. Repasky, C. R. W. Guimaraes, J. Chandrasekhar, J. Tirado-Rives and W. L. Jorgensen, *J. Am. Chem. Soc.*, 2003, **125**, 6663.

39 D. T. Limmer, A. P. Willard, P. Madden and D. Chandler, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 4200.

40 S. N. Steinmann and P. Sautet, *J. Phys. Chem. C*, 2016, **120**, 5619.

41 S. Ringe, H. Oberhofer, C. Hille, S. Matera and K. Reuter, *J. Chem. Theory Comput.*, 2016, **12**, 40524066.

42 T. P. Straatsma and J. A. McCammon, *J. Chem. Phys.*, 1989, **90**, 3300.

43 M. Strajbl, Y. Y. Sham, J. Villa, Z.-T. Chu and A. Warshel, *J. Phys. Chem. B*, 2000, **104**, 4578.

44 S. Gautier, S. N. Steinmann, C. Michel, P. Fleurat-Lessard and P. Sautet, *Phys. Chem. Chem. Phys.*, 2015, **17**, 28921.

45 H.-K. Lim, H. Lee and H. Kim, *J. Chem. Theory Comput.*, 2016, **Just Accepted**, 10.1021/acs.jctc.6b00469.

46 G. Kresse and J. Hafner, *Phys. Rev. B*, 1993, **47**, 558.

47 G. Kresse and J. Furthmuller, *Phys. Rev. B*, 1996, **54**, 11169.

48 S. N. Steinmann and C. Corminboeuf, *J. Chem. Theory Comput.*, 2011, **7**, 3567.

49 J. P. Perdew, K. Burke and M. Ernzerhof, *Phys. Rev. Lett.*, 1996, **77**, 3865.

50 P. E. Blochl, *Phys. Rev. B*, 1994, **50**, 17953.

51 G. Kresse and D. Joubert, *Phys. Rev. B*, 1999, **59**, 1758.

52 F. L. Hirshfeld, *Theor. Chem. Acc.*, 1977, **44**, 129.

53 *Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. CM5PAC, version 2011; University of Minnesota: Minneapolis, MN, 2011.*

54 S. Grimme, *J. Chem. Theory Comput.*, 2014, **10**, 4497.

55 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem.*, 2005, **26**, 1781.

56 W. L. Jorgensen and J. K. Buckner, *J. Phys. Chem.*, 1986, **90**, 4651.

57 M. Tuckerman, B. J. Berne and G. J. Martyna, *J. Chem. Phys.*, 1992, **97**, 1990.

58 T. C. Beutler, A. E. Mark, R. C. van Schaik, P. R. Gerber and W. F. van Gunsteren, *Chem. Phys. Lett.*, 1994, **222**, 529.

59 M. Zacharias, T. P. Straatsma and J. A. McCammon, *J. Chem. Phys.*, 1994, **100**, 9025.

60 A. Pohorille, C. Jarzynski and C. Chipot, *J. Phys. Chem. B*, 2010, **114**, 10235.

61 S. Boresch and M. Karplus, *J. Phys. Chem. A*, 1999, **103**, 103.

62 N. Lu, D. A. Kofke and T. B. Woolf, *J. Comput. Chem.*, 2004, **25**, 28.

63 P. Liu, F. Dehez, W. Cai and C. Chipot, *J. Chem. Theory Comput.*, 2012, **8**, 2606.

64 M. Orozco, W. L. Jorgensen and F. J. Luque, *J. Comput. Chem.*, 1993, **14**, 14981503.

65 D. Casanova, S. Gusarov, A. Kovalenko and T. Ziegler, *J. Chem. Theory Comput.*, 2007, **3**, 458476.

66 M. Retegan, M. Martins-Costa and M. F. Ruiz-Lopez, *J. Chem. Phys.*, 2010, **133**, 064103.

67 G. Konig, P. S. Hudson, S. Boresch and H. L. Woodcock, *J. Chem. Theory Comput.*, 2014, **10**, 1406.