

Cite this: DOI: 10.1039/xxxxxxxxxx

From In Silica To In Silico: Retention Thermodynamics At Solid-Liquid Interfaces[†]

Krystel El Hage,^a Raymond Bemish^b and Markus Meuwly^{*a}Received Date
Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

The dynamics of solvated molecules at the solid/liquid interface is essential for a molecular-level understanding for the solution thermodynamics in reversed phase liquid chromatography (RPLC). The heterogeneous nature of the systems and the competing intermolecular interactions makes solute retention in RPLC a surprisingly challenging problem which benefits greatly from modelling at atomistic resolution. However, the quality of the underlying computational model needs to be sufficiently accurate to provide a realistic description of the energetics and dynamics of systems, especially for solution-phase simulations. Here, the retention thermodynamics and the retention mechanism of a range of benzene-derivatives in C18 stationary-phase chains in contact with water/methanol mixtures is studied using point charge (PC) and multipole (MTP) electrostatic models. The results demonstrate that free energy simulations with a faithful MTP representation of the computational model provide quantitative and molecular level insight into the thermodynamics of adsorption/desorption in chromatographic systems while a conventional PC representation fails in doing so. This provides a rational basis to develop more quantitative and validated models for the optimization of separation systems.

1 Introduction

Solute retention in reversed-phase liquid chromatography is an important process dominated by intermolecular interactions and has received considerable attention during the past decade.^{1–8} Due to the molecular complexity of the system and because QSAR (quantitative structure activity relationship)-like interpretations of experimental data remained largely inconclusive, it is of major importance to improve the understanding of solute retention using computational modelling so that the composition, functionalization and other physico-chemical properties of specific columns can be quantitatively predicted.^{9–11} Extensions of molecularly-resolved pictures to the mesoscale for liquid chromatography are particularly relevant as they have the potential to reduce development times and potentially improved selectivity of such columns.

Despite the seemingly simple chemical composition of such systems, the atomistic understanding underlying the separation process remains elusive since the system is highly dynamical, heterogeneous and disordered^{1–4}. Use of computational and experimental spectroscopic techniques such as IR^{12–14}, Raman¹⁵ and NMR^{16–18} has provided important insights concerning the water/methanol mixture and proved to be useful in order to characterize these highly heterogeneous systems at the molecular level^{1,7,19–23}. The process is molecularly driven and requires bottom up strategies because decomposition schemes such as QSAR are not sufficiently detailed.

The separation mechanism involves specific and non-specific interactions between the solute and its environment which includes the solvent mixture, the support surface and the functionalized alkyl chains chemically bonded to the support surface.^{24–29} Challenges further increase when separating structurally or functionally similar compounds (for example, isomers) where their corresponding peaks elute close to one another.³⁰ Progress has been made in understanding surface

^a Department of Chemistry, University of Basel, Klingelbergstrasse 80, CH-4056 Basel, Switzerland

^b Air Force Research Laboratory, Space Vehicles Directorate, Kirtland AFB, New Mexico 87117, USA

*To whom correspondence should be addressed. Email: m.meuwly@unibas.ch

density, chain length, temperature, and solvent composition effects on retention.^{31–34} However, quantitative studies are rare⁸ and thus no quantitative retention model is available that would allow the chromatographic parameters, such as capacity factor, resolution and separation factor, for individual analytes separated under given conditions to be accurately predicted. Given the cost of industrial-scale experimentation, distinct composition parameters and operation scenarios: choice of solvent structure and mixtures, alkyl chains length and types and silica layer properties can now be computationally tested for separation viability. Such has been the case of *in silico* drug discovery.^{35–37}

A typical chromatographic system consists of a solid silica substrate tethered with hydrophobic alkyl chains forming the stationary phase, a solvent mixture and analyte molecules (see Figure 1). The alkyl chains in such systems are highly flexible which modulates the intermolecular interactions at the stationary/mobile interface which significantly influences retention and chemical selectivity.³⁹ The nature of the solvent mixture and alkyl chains is decisive for the separation process. Commonly used mixtures are water (H₂O), together with an organic co-solvent such as acetonitrile (ACN), methanol (MeOH) or tetra-hydro furan (THF). The change in the solvent composition mainly modulates the hydrophobic nature of the environment which affects the elution time and improves the resolution of the analyte.³⁹ Experimental studies showed that separation of compounds depends decisively on the hydrophobicity of the studied compounds.^{40–42} Contrary to normal phase liquid chromatography (NPLC), the stationary phase in RPLC is less polar than the mobile phase, and so the retention time increases as the polarity of the mobile phase decreases. Thus, non-polar analytes are more strongly retained than polar ones.⁴³

One of the eminent questions for RPLC is whether, at a molecular level, the retention process is dominated by partitioning or by adsorption or whether it is a combination of the two. Adsorption chromatography is based on the retention of solute by surface adsorption, whereas partition chromatography is based on the adsorption of solute by bulk stationary phase usually held in place by an inert scaffold of solid particles.⁵ The retention mechanism based on hydrophobic interactions depends on the surface coverage of the stationary phase (C₁₈ in the present work). In materials with dense coverage $\geq 3.8 \mu\text{mol}/\text{m}^2$, partition predominates, whereas in materials with less dense coverage $\leq 3.8 \mu\text{mol}/\text{m}^2$, adsorption is of increasing importance. Because higher density coverage shows a homogeneous coverage of the supported surface and the unblocked silanols are shielded by the C18 ligands in materials with a high surface load.

Recent work has shown that atomistic simulations with ac-

curate force fields are suitable to quantitatively describe the thermodynamics of complex systems.^{8,44,45} They allow to explicitly determine the hydration free energy^{44,45} and the free energy of desorption from the surface⁸ which can be directly compared with the experimentally measured partition coefficients k .^{4,39} However, these force fields can be based on point charges (PC) or multipoles (MTP) depending on the molecules considered.⁴⁵ For the specific case of benzene derivative solutes, first a PC force field was used to model the entire system and then a combined PC/MTP force field, where only the solutes (here benzene derivatives) are treated with multipoles (in order to take into account the anisotropic electron distribution of the aromatic cycle and the moieties attached to it) while the chromatographic system is treated with PCs, and the resulting quantitative atomistic simulations were compared to experiment. This strategy was previously found to provide the accuracy of MTP simulations but at a lower computational cost.⁴⁴

The present work discusses the use, merit and shortcomings of point charge- and multipole-based force fields for molecular simulations of solute molecules in heterogeneous systems at the solid/liquid interface. First, the parametrizations of both types of models are described. Next, the free energy of adsorption for two different solvent compositions are determined and specific cases are discussed. Finally, conclusions are presented.

2 Theoretical and Computational methods

2.1 The System

The model silica support consists of two 8.75 Å thick segments of the (101) face of a crystalline quartz lattice with dimensions 36 × 41 Å. This results in two -OH capped surfaces with a vicinal silanol density of 3.1 μmol/m². A model for a chromatographic column was then generated by covalently tethering C₁₈ alkyl chains to the silanol oxygen atoms of the quartz surface and orthogonal to the bulk quartz. The C₁₈ chains were randomly distributed over the surface silanols to give a surface coverage of 2.65 μmol/m² which is a typical value in experiments.³⁹ An equilibrated 80 Å thick solvent box (water/methanol) was superimposed on this arrangement which leads to a unit cell with dimensions of 36 × 41 × 97 Å³, see Figure 1.

The water model in the present work is the transferable intermolecular potential three point (TIP3P)⁴⁸ model. For the methanol (MeOH) molecule, parameters are taken from previous work³ with an OH-equilibrium bond length of 0.96 Å.⁴⁹ The CT3-OH1 bond, the HA-CT3-OH1 and HA-OH1-CT3 angles, and the dihedrals⁴⁹ use conventional CHARMM parameters. Two solvent compositions were considered: for a 20:80 MeOH/water mixture 307 MeOH molecules and 2966 H₂O molecules are in the system,

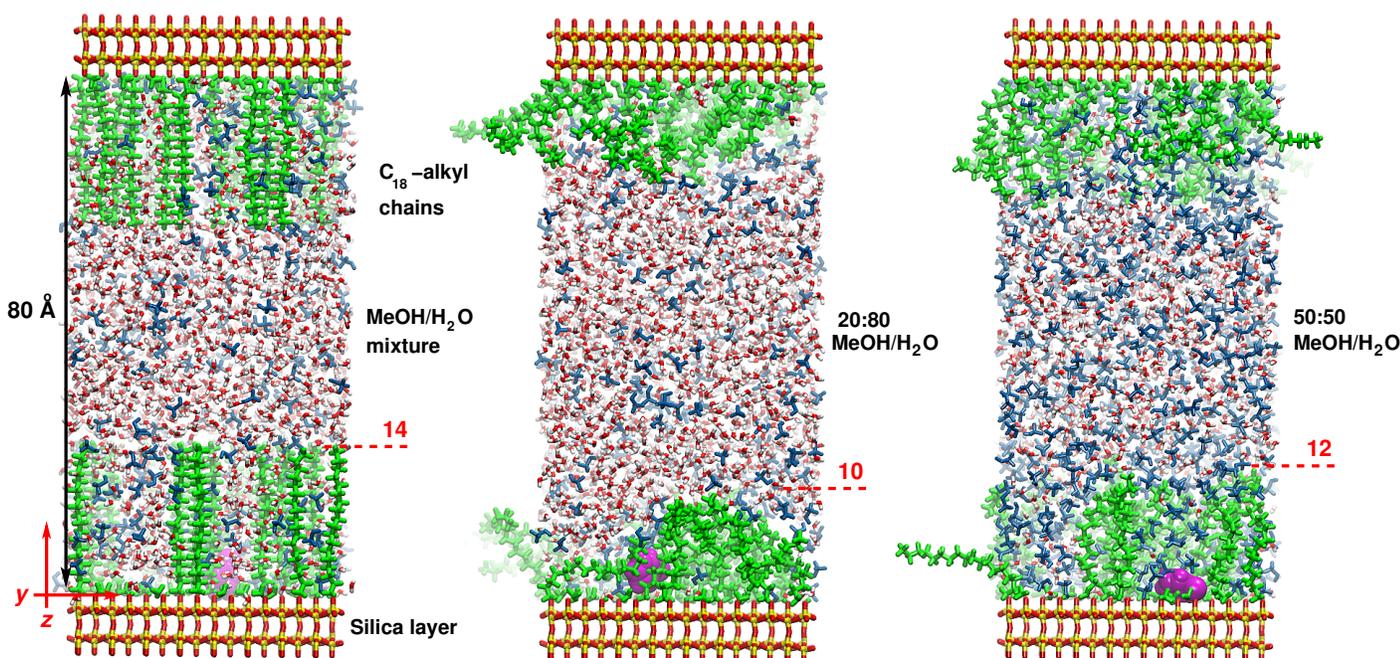


Fig. 1 Schematic representation of a MeOH/H₂O chromatographic system together with the analyte (magenta, vdW representation). The silica layer is in licorice (silica in yellow, oxygen in red), the C₁₈-alkyl chains (green) grafted onto it, H₂O (red), MeOH (blue), and PhCl (magenta). These snapshots are taken from different umbrella windows with surface-to-center of mass (COM) of PhCl distance of 3.0 Å; they illustrate the intercalation of the analyte in the highly dynamical C₁₈-functionalized silica column. The length of the alkyl chains is 14 Å, but the typical extent is found to be 10 Å and 70 Å for 20:80 and 12 Å and 68 Å 50:50 MeOH/H₂O solvent mixtures, respectively.³⁸

whereas a 50:50 MeOH/water mixture corresponds to 743 MeOH and 1362 H₂O molecules.

2.2 Force Field Parametrization

Two parametrizations for the solutes are considered in the following. One of them is based on a point charge (PC) representation of the electrostatics as is used in typical conventional force fields, including CHARMM⁵⁰, AMBER⁵¹, GROMOS⁵² and OPLS⁵³, although exceptions to this have been explored.⁵⁴ Another possibility is to assign atom-centered multipoles (MTPs) to all or selected atoms. Here MTPs up to quadrupoles on non-hydrogen atoms and monopoles on hydrogens are used which has been shown to provide accurate representations of the electrostatic potential^{55–57} and in particular to correctly capture the electron anisotropy around substituted, halogenated phenyls.^{44,45,58,59} The parametrization of PhH and PhCl was that from previous work.⁵⁹ For PhOH and PhCH₃, the parametrization procedure, for PC and MTP, followed the same protocol.⁴⁵ The starting topology and the bonded parameters are those of CGenFF^{60,61} and only the electrostatic and van der Waals parameters were optimized. Briefly, first the point charges are optimized to represent the molecular ESP. In a next step these PCs are frozen and the atomic dipole and quadrupole moments are optimized

to further improve ESP compared to the reference *ab initio* calculated electrostatic potential (ESP) at the MP2/aug-cc-pVDZ level of theory.^{62,63} Next, the van der Waals parameters, starting from values of the CGenFF force field⁶⁴, are scaled to reproduce thermodynamic properties: density ρ , heat of vaporization ΔH_{vap} (determined from a simulation of the pure liquid according to $\Delta H_{\text{vap}}(T) = \langle E_{\text{gas}}(T) \rangle - \langle E_{\text{liq}}(T) \rangle + RT$ where $\langle E_{\text{gas}} \rangle$ and $\langle E_{\text{liq}} \rangle$ are the ensemble average of potential energies of one molecule in the gas and liquid phases, T is the temperature and R is the gas constant) and hydration free energy ΔG_{hyd} (determined using thermodynamic integration from simulations of one solute molecule in a $25 \times 25 \times 25 \text{ \AA}^3$ box of TIP3P⁴⁸ water molecules).⁴⁵

The best-fit thermodynamic data for the four compounds based on PC and MTP models for the electrostatics are summarized in Table 1. It is important to point out that the van der Waals parameters for one compound differ depending on whether a PC or a MTP electrostatic model was used. A previously introduced score S which weights hydration free energies more heavily than heats of formation or pure liquid density is used to compare the different models on an equal footing.⁴⁵ The ΔG_{hyd} is more heavily weighted because in the present study we are primarily concerned with this quantity.

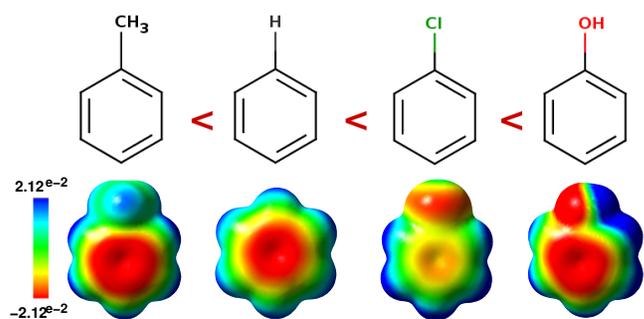


Fig. 2 The solutes used in this study, along with the Molecular Electrostatic Potential (MEP) at the $10^{-3} e a_0^{-3}$ isodensity surface.⁴⁶ The MEP is that from which the PCs and MTPs were derived by fitting to the ESP. Phenyl derivatives PhR are arranged according to increasing relative polarity: PhCH₃ 0.099 < PhH 0.111 < PhCl 0.188 < PhOH 0.680. The values of relative polarity are normalized from measurements of solvent shifts of absorption spectra.⁴⁷

PC	ρ	ΔH_{vap}	ΔG_{hyd}	S
PhH	0.86 (0.88)	7.53 (7.89)	-0.77 (-0.86)	0.43
PhCl	1.11 (1.11)	9.68 (9.79)	-0.66 (-1.12)	1.09
PhOH	1.07 (1.07)	14.01 (13.81)	-6.49 (-6.62)	0.20
PhCH ₃	0.9 (0.86)	8.50 (n/a)	-0.81 (-0.90)	0.04
MTP	ρ	ΔH_{vap}	ΔG_{hyd}	S
PhH	0.90 (0.88)	7.88 (7.89)	-0.89 (-0.86)	0.01
PhCl	1.14 (1.11)	10.13 (9.79)	-1.11 (-1.12)	0.35
PhOH	1.08 (1.07)	12.47 (13.81)	-6.47 (-6.62)	0.09
PhCH ₃	1.03 (0.86)	8.84 (n/a)	-0.86 (-0.90)	0.08

Table 1 Thermodynamic data calculated from optimized parameters and compared with the experimental values^{65–67} given in parenthesis using both PC and MTP electrostatics. Missing experimental data is represented by the symbol n/a (not available). The density is reported in g/cm³, the heat of vaporization and the free energy of hydration are reported in kcal/mol at 298 K. The score $S = \sum_{i=1}^3 w_i (\text{Obs}_i - \text{Calc}_i)^2$ is that from Ref.⁴⁵ with $w_\rho = 1$, $w_{\Delta H} = 3$ and $w_{\Delta G} = 5$ to differently weight the three observables and put most weight on the hydration free energy.

2.3 Molecular Dynamics Simulations

For all molecular dynamics (MD) simulations of the chromatographic system the same protocol was used using the CHARMM program⁶⁸. After solvation, the systems were subjected to 2000 steps of steepest descent minimization to relieve strain. The system was then heated to $T = 300$ K and equilibrated at constant volume and temperature for 500 ps. Next, MD simulations were carried out with a time step of 1 fs. SHAKE⁶⁹ was used to constrain all bonds involving hydrogen atoms. All simulations were carried out with periodic boundary conditions. For PC-PC interactions, particle mesh Ewald (PME)⁷⁰ was used with a grid

spacing of 1 Å and a relative tolerance of 10^{-6} . Nonbonded interactions were truncated at a distance of 14 Å on an atom-by-atom basis, using a shift function for the electrostatic interactions and a switch algorithm for the van der Waals interactions.⁷¹ For higher MTP interactions a power-law dependent switching was employed.⁵⁹

The atom positions of the bulk quartz surface with the exception of the exposed hydrogen atoms of the silanol groups were held fixed during the simulation (1664 atoms). Initially, the system without analyte molecules was heated to $T = 300$ K and equilibrated at this temperature for 500 ps, followed by 5 ns of production simulations. Next, the analyte molecules were introduced in the stationary phase at 3 Å above the silica surface and overlapping solvent molecules were removed. Then, the system was again minimized, heated and equilibrated using the same protocol.

Umbrella sampling simulations⁷² were used to evaluate the free energy of binding along the z -direction $G(z)$ of the analyte away from the surface (which is the xy -plane). The reaction coordinate was the distance of the center of mass of the analyte above the surface and the probe molecule was moved away from the surface along the z -direction towards the middle of the mobile phase, see Figures 3 and 4. The sampling consisted of 28 evenly spaced windows, separated by 1 Å for optimal overlap of the histograms. For every window, the simulation included 500 ps of equilibrium simulation and data accumulation took place for the next 2 ns. This gives 56 ns of sampling for a given system. The overall free energy profile was constructed from a weighted histogram analysis (WHAM).^{73,74}

Convergence of these free energy simulations was assessed by running independent simulations with 2 ns of equilibration and 5 ns of accumulation for PhCl, using both PC and MTP representations, in both solvent compositions 20:80 and 50:50. The results show that convergence within an US window is obtained after ≈ 1 ns of accumulation time after which the free energy contribution for the particular window fluctuates around its average.⁷⁵

3 Results

3.1 Thermodynamics of Intercalation

For all analytes, experimental data for the retention time and partition coefficient for a C₁₈ column in different MeOH/H₂O compositions (20:80 and 50:50) and for different C₁₈ coverage is available.³⁹ In order to make direct contact with this data, free energy simulations using umbrella sampling (US) were carried out. Typical snapshots from such simulations

are shown in Figure 1. During US, the analytes' height above the surface z is constrained to within a window of 1 Å but it is free to rotate. It is evident that the C₁₈ chains are not fully extended but in a partially collapsed state as was already found in previous simulations^{3,7,20,21}. The typical average extent is found to be 10 Å and 12 Å above the solid support for 20:80 and 50:50 MeOH/H₂O solvent mixtures, respectively.³⁸

The computed adsorption free energy can then be compared in two different ways with experiments: first by calculating the change in free energy of the same compound in two different solvent compositions ($\Delta\Delta G_{r_1 \rightarrow r_2}$, where r_1 and r_2 are two different MeOH/H₂O ratios r) and, secondly, by comparing the difference in free energy of two different compounds in the same solvent ($\Delta\Delta G_{c_1 \rightarrow c_2}$, where c_1 and c_2 are two different compounds c). The free energy difference between the solvated and adsorbed states is experimentally estimated from the measured partition coefficient K by the relationship:

$$\Delta G = -RT \ln K \quad (1)$$

where R is the gas constant, and T is the temperature. K is the molar concentration of analyte in the stationary phase divided by the molar concentration of the analyte, which corresponds to the equilibrium constant K_e . Thus, for a given solute X in two different solvent compositions (here 20:80 and 50:50 MeOH/H₂O) the change in free energy is

$$\Delta\Delta G_{r_1 \rightarrow r_2} = -RT \ln \left(\frac{k'_{50:50}}{k'_{20:80}} \right) \quad (2)$$

and the difference in adsorption free energy for two compounds c_1 and c_2 is

$$\Delta\Delta G_{c_1 \rightarrow c_2} = -RT \ln \left(\frac{k'_{c_2}}{k'_{c_1}} \right) \quad (3)$$

The variable k' is the capacity factor which is directly proportional to the retention time t_r , which, in turn, is a measure for the free energy difference for the solute in the mobile phase versus the stationary phase, respectively. Hence, k' can be used instead of K in equation 1.⁷⁶ The capacity factors for all solutes considered in the present work are those from Figure 4C in Ref.³⁹ for a C₁₈ system density of 2.60 μmol/m². These experimental values can be directly compared with the change in free energy from free energy simulations, see Figures 3 and 4. As the free energy $\Delta G(z)$ depends on the position above the surface, $\Delta G(z)$ was averaged over a finite range $3 \leq z \leq 14$ Å within the alkyl chains, see Figure 3. The experimental and computed free energy values using both PC and MTP representations depending on the solvent composition and compared to PhCl as a reference are reported in Tables 2 and 3, respectively.

$\langle \Delta\Delta G \rangle_{20:80 \rightarrow 50:50}$ (kcal/mol)	Calc. PC	Calc. MTP	Exp.
PhCl	1.56	0.83	0.82
PhCH ₃	0.50	0.96	0.91
PhH	2.02	0.77	0.86
PhOH	1.71	0.59	0.55

Table 2 Comparison of computed and experimental³⁹ relative free energy differences in 20:80 and 50:50 MeOH/H₂O solvent mixtures using both PC and MTP electrostatics.

The results in Table 2 indicate that all analytes are more stable in 20:80 solvent composition compared to 50:50, which also agrees with experiment.³⁹ However, the PC model fails to reproduce the experimental desorption free energy for all four compounds, while the MTP model reproduces experiments to within ≈ 0.1 kcal/mol. For PhCl, the relative free energy for the two different solvent compositions is $\Delta\Delta G_{20:80 \rightarrow 50:50} = 0.82$ kcal/mol from experiment, compared with an average value of 0.83 kcal/mol from free energy simulations using a MTP model. The same simulations based on a PC model yield a difference of 1.56 kcal/mol which considerably overestimates stabilization in the 20:80 column. Similar results are found for PhCH₃, PhH and PhOH all of which confirm that PC models are less suitable for quantitative studies than MTP-based models. The correlation coefficient between experimentally measured and MTP-computed values for $\langle \Delta\Delta G \rangle_{20:80 \rightarrow 50:50}$ is 0.92 and the slope is 0.88. For PC-based simulations no meaningful correlation between computed and experimental values can be found.

Alternatively, changes in stabilization free energies relative to a (arbitrary) reference (here PhCl) can be considered, too, see Table 3. For PhH, the comparison with experiment yields $\langle \Delta\Delta G \rangle_{\text{PhCl} \rightarrow \text{PhH}} = 0.60$ and 0.61 for 20:80 and 50:50 MeOH/H₂O solvent composition respectively, which compares with 0.57 and 0.51 kcal/mol from the MTP simulations. Simulations with the PC model yield -0.33 and 0.13 kcal/mol, respectively. For PhOH, the results for $\langle \Delta\Delta G \rangle_{\text{PhCl} \rightarrow \text{PhOH}}$ from MTP simulations are within less than 0.2 kcal/mol of the experiments whereas those from simulations with PC differ by 0.4 kcal/mol (for 20:80) or agree very closely (for 50:50). Finally, for PhCH₃ results from MTP simulations agree well with experiments whereas those from PC simulations differ by 0.4 and 0.8 kcal/mol depending on the mixture.

The free energy profiles for PhH, PhCl, PhOH and PhCH₃ in MeOH/H₂O mixtures of 20:80 (left panel) and 50:50 (right panel) using MTP and PC electrostatics are illustrated in Figures 3 and 4, respectively. The reference point $z = 0$ are the O-atoms of the hydroxylated silica surface (see inset in Figure 3 (left

$\langle \Delta \Delta G \rangle_{\text{PhCl} \rightarrow \text{c}_2}$ (kcal/mol)	20:80			50:50		
	Calc. PC	Calc. MTP	Exp.	Calc. PC	Calc. MTP	Exp.
PhH	-0.33	0.57	0.60	0.13	0.51	0.61
PhCH ₃	0.40	0.05	0.06	-0.66	0.08	0.20
PhOH	1.03	1.25	1.43	1.17	1.01	1.16

Table 3 Comparison of the computed and experimentally measured³⁹ free energy difference $\langle \Delta \Delta G \rangle_{\text{PhCl} \rightarrow \text{c}_2}$ (kcal/mol) in 20:80 and 50:50 MeOH/H₂O solvent using both PC and MTP electrostatics with PhCl as the reference.

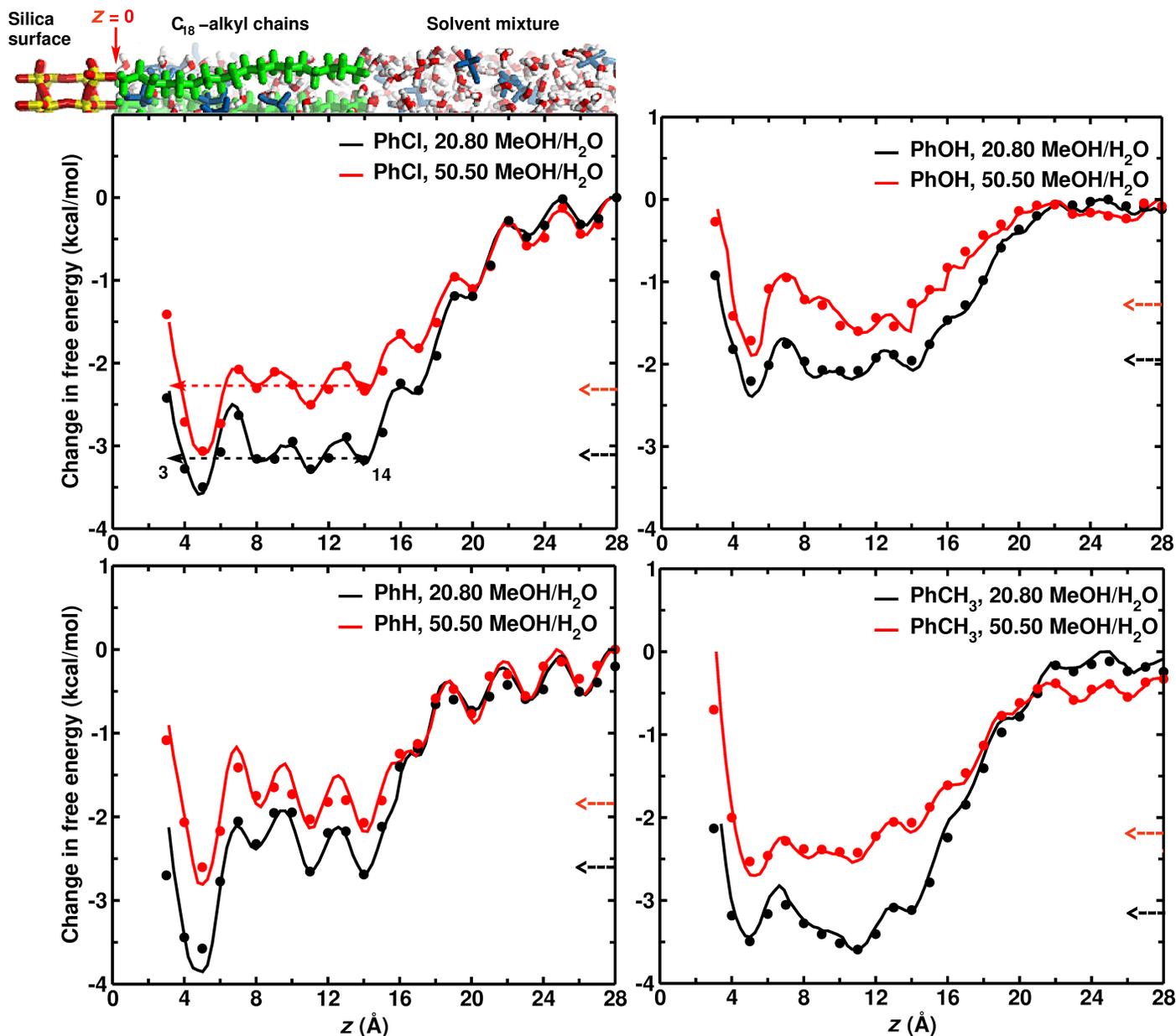


Fig. 3 Comparison of free energy difference for PhH, PhCl, PhOH and PhCH₃ in MeOH/H₂O mixtures of 20:80 (black) and 50:50 (red) calculated with MTP electrostatics. The reference point $z = 0$ are the O-atoms of the hydroxylated silica surface. The inset illustrates the chromatographic column. The region over which the free energy value was averaged is marked with a double-headed red arrow. The dashed arrows along the y-axes indicate the average free energy value for the respective profile. The points are raw data from the 28 bins from umbrella sampling.

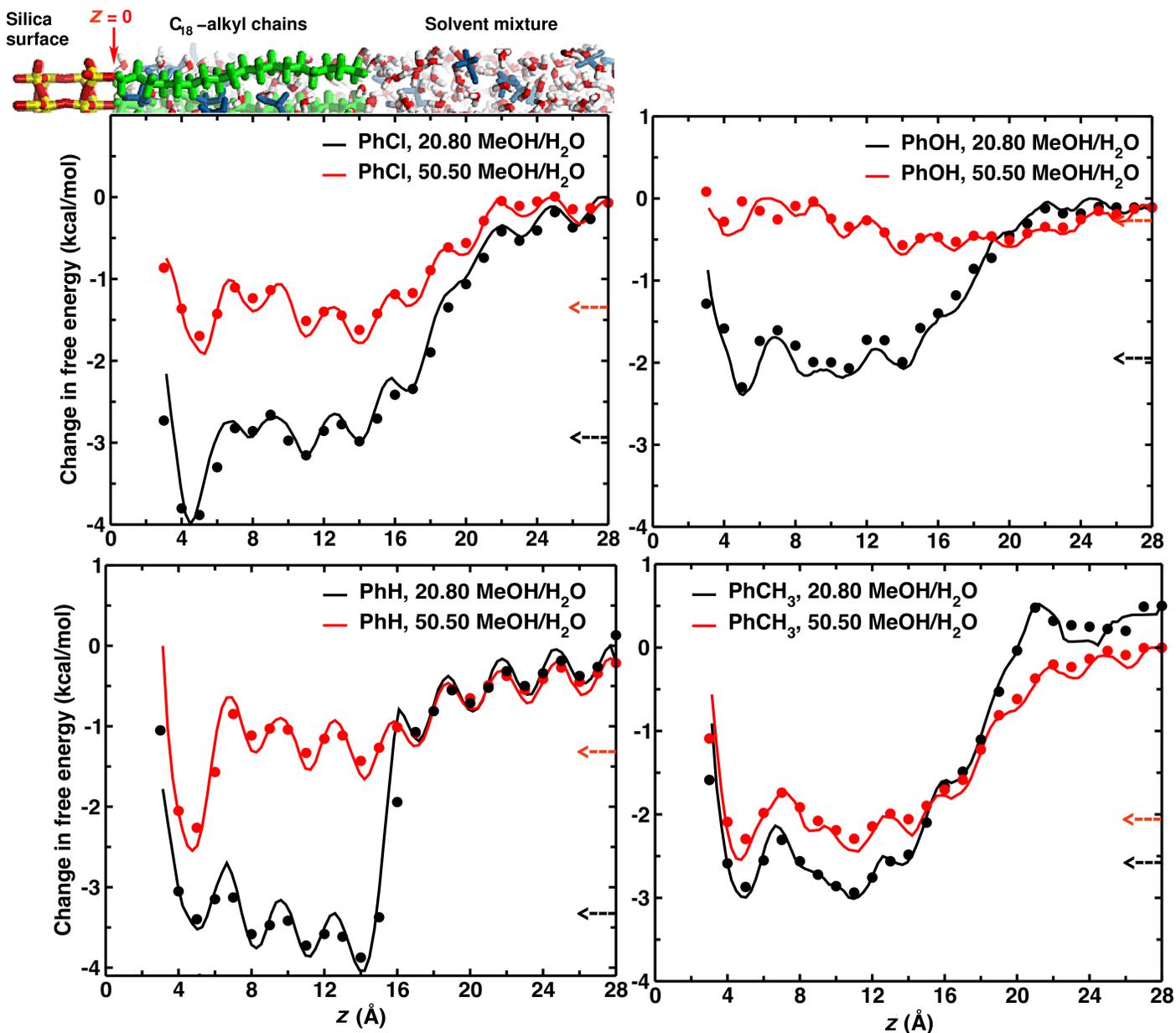


Fig. 4 Comparison of free energy difference for PhH, PhCl, PhOH and PhCH₃ in MeOH/H₂O mixtures of 20:80 (black) and 50:50 (red) calculated with PC electrostatics. The reference point $z = 0$ are the O-atoms of the hydroxylated silica surface. The inset illustrates the chromatographic column. The dashed arrows along the y -axes indicate the average free energy value for the respective profile. The points are raw data from the 28 bins from umbrella sampling.

panel)). Figures 3 and 4 show that the free energy profiles exhibit an attractive region (up to $z \approx 14$ Å, which is the average height of the stationary phase) after which the free energy increases by 2 to 3 kcal/mol (depending on the compound, the electrostatic model and the solvent composition) upon transfer from the hydrophobic stationary phase into the polar mobile phase. This transition region extends from $14 \leq z \leq 18$ Å.

It is also noted that the free energy curves exhibit undulations, up to 0.5 kcal/mol in magnitude, along the z -coordinate. Analysis of the average number of solvent molecules (MeOH/H₂O) $\langle N_{\text{solvent}} \rangle$ in a 6 Å shell around the center of mass of the solute from the US simulations at $z = 7$ (local maximum) and $z = 8$ (local minimum) for PhCl in the 50:50 MeOH/H₂O mixture, finds 3.6 and 2.0 solvent molecules on average, respectively. This indicates that increased solvation of the solute in the stationary phase leads to a reduced retention free energy. This molecular level interpretation shows how different structuring of the mobile phase (here MeOH/water) between the solute and the alkyl chains in a chromatographic system affects the retention free energy of the solute (see Figure 8 B).

The present simulations find minima along the free energy curves for $5 < z < 10$ Å above the surface which is consistent with the interpretation of previous experiments that report an increase in the interaction energy for $z < 18$ Å and a minimum in the range between 5 and 10 Å above the stationary phase.³² Considering the minimum at $z \approx 14$ Å along the free energy curves before moving into the solvent shows the following order of stability: PhCl > PhCH₃ > PhH > PhOH, for both solvent compositions when using the MTP model (Figure 3), while using the PC model the order of stability is PhCH₃ > PhH > PhCl > PhOH (Figure 4). The observed order of stability using the MTP electrostatics correlates with the retention factors k' observed experimentally for these probes ($k'_{\text{PhCl}} > k'_{\text{PhCH}_3} > k'_{\text{PhH}} > k'_{\text{PhOH}}$)³⁹ while the order of stability using the PC model does not correlate with the retention factors but rather follows the order of polarity of the molecules PhCH₃ < PhH < PhCl < PhOH.

The difference between the free energy profiles using PC and MTP models originates from the electronic anisotropy around the phenyl ring (see Figure 2). This anisotropy is not accounted for when using a PC representation for the electrostatics⁴⁴ and affects the intercalation of the solutes within the alkyl-chains (stacking of the phenyl ring between the alkyl-chains, which represents the majority of the interactions) and thus desorption from the stationary phase. This shows that all phenyl-like solutes need to be represented with an MTP force field in order to accurately reproduce the experimental observables.

Another notable difference in the free energy curves is their behaviour for PhOH in the 50:50 MeOH/H₂O mixture. Using a MTP representation there is an adsorption stabilization of -1.5 kcal/mol at around $z = 5$ Å whereas simulations with a PC model yield only a broad and unspecific minimum less than -0.5 kcal/mol deep which is indicative of partitioning.

3.2 Dynamics of Intercalation

The unbiased dynamics of the solute is also of interest. For this, the dynamics of the solute depending on its initial position with respect to the surface was investigated. Figures 5 and 6 show the evolution of the distance between the center of mass of the analyte (X) and the surface, using MTP and PC models in the simulations, respectively, for 20:80 (left panels) and 50:50 MeOH/H₂O mixtures (right panels) for different initial positions of the solute (at 3, 9, 15, 21, and 27 Å above the surface) for 5 ns simulations. Independent of the chemical modification on the solute the probe molecules prefer to sample regions close to the surface - i.e. the stationary phase and the alkyl chains - and not the solvent mixture. This is consistent with the free energy simulations and with previous observations,^{8,10,19} and was linked to the fact that the interface is rough and that the alcohol in water/alcohol solvent mixtures wets the surface but does not alter the stationary phase structure.⁷⁷ The surface of the stationary phase as defined by the average position of the CH₂ group extending farthest away from the silica support is indicated by dashed lines at 10 Å and 70 Å for 20:80 MeOH/H₂O and 12 Å and 68 Å for 50:50 MeOH/H₂O solvent mixtures, respectively.³⁸

Depending on solvent composition, the probability for the analyte to sample the surface differs slightly. For 20:80 mixtures the interaction is typically stronger, i.e. the probe molecules sample regions closer to the surface whereas for 50:50 mixtures it is more loosely bound. This confirms the results from Table 2 and 3. A typical example is the free dynamics of PhCH₃ which remains localized at one surface for a 20:80 mixture whereas for 50:50 it interacts with either side of the stationary phase because specific solvent-surface interactions are preferred over hydrophobic solute-surface interactions, and so the solute migrates faster to the mobile phase.³⁹

Comparing the different probe molecules in 20:80 solvent mixture suggests that PhCl intercalates almost permanently, i.e. interacts most strongly for the time scale considered here (an aggregate of 140 ns) when using MTP electrostatics. Conversely, PhOH as the other extreme exchanges readily between the two interfaces with MTP electrostatics (Figure 5). This is even more pronounced for a 50:50 solvent mixture. However, when

comparing the different solutes in 20:80 and 50:50 solvent mixtures using PC electrostatics (Figure 6) it is observed that PhCH₃ intercalates almost permanently over the time scale considered, followed by PhH, while PhCl samples regions in the solvent mixtures that are close to the alkyl chain/solvent interface. As for PhOH it migrates to the solvent mixture but does not exchange between the two interfaces (Figure 6).

3.3 PhCl Anchors to Silanol Oxygens of the Surface

As mentioned earlier, with increasing polarity of the solute, its stability in the stationary phase typically decreases. This is, however, not the case for PhCl as found in experiments³⁹ and when using MTP electrostatics in simulations⁸. Using a PC representation in the simulations leads to ranking in the stability from simulations which follows the order of polarity. Hence, for PhCl it is meaningful to further analyse the molecular origin for the increased interaction strength with the stationary phase which can be captured with MTP electrostatic but not with a simpler PC representation.

PhCl, where Cl is covalently bonded to the phenyl ring, does not follow this trend because the halogen atom exhibits a sigma-hole, that results in a positive electrostatic potential along the C-Cl bond. Such a charge distribution (“janus-like”) can interact with negatively charged sites (here, the silanol oxygens) and the negative electrostatic potential on the flanks can interact with positively charged sites (here, the silanol hydrogens).⁷⁸ Thus PhCl uses the sigma-hole feature of the halogen as a double anchor to stick to the silica surface, which then reinforces its stability in the stationary phase.⁸ On the contrary, a simple point charge model for the electrostatic potential, which assigns a partial charge to each atom, is unable to reproduce the thermodynamics of halogenated compounds, as it fails to describe the positive lobe of the sigma-hole, and the quadrupolar electrostatics around aromatic rings.^{44,45,79,80}

To substantiate the difference between a PC and a MTP representation, Figure 7 compares structural and thermodynamic properties for PhCl in 50:50 MeOH/H₂O solvent composition. First, US simulations show that with PC the solute less favourably interacts with the stationary phase compared to simulations with MTPs by 0.95 kcal/mol. Moreover, the free energy profile with PC electrostatics (red) resembles that of PhH in the same 50:50 MeOH/H₂O solvent mixture (Figure 4), suggesting that a PC model for PhCl is barely capable of distinguishing between PhH and PhCl. The time evolution of the PhCl-surface distance along 5 ns starting from $z = 3$ as initial position in 50:50 MeOH/H₂O solvent mixture (Figure 7B) shows that PhCl simulated with PC electrostatics (green) compared to MTP (black) migrates more

rapidly into the mobile phase.

In addition, the surface-to-PhCl center of mass distance time series obtained from PC simulations (Figure 7B) resemble more those for PhH and PhCH₃ (Figure 6, sampling of the interfacial region) than the time series obtained from MTP simulations of PhCl (Figure 7B, interaction with the surface). This, again, is indicative that a PC model has difficulties to distinguish between PhCl, PhH and PhCH₃. Moreover, the corresponding distance probability distribution $P(z)$ averaged over 140 ns of MD (inset in Figure 7B) mirrors the MTP-PC difference in the free energy obtained with US. Third, Figure 7C reports the cumulative number of water (red), methanol (blue) and silica (green) oxygen atoms within 4 Å of Cl during 5 ns of MD simulation starting from $z = 3$ using both MTP (left panel) and PC (right panel) simulations. Compared to PC sampling, only the simulations with MTP find the Cl atom interacting with the surface-silanol oxygens. In addition, when detaching from the surface, the solute engages in interactions (see Figure 8B) with the mobile phase oxygens (water or methanol oxygens). This rationalizes the functional role of the solvent even for regions deep within the stationary phase (solvent structuring Figure 8B). Finally, Figure 7D shows the projection of the C-Cl bond onto the xy -plane for different z -ranges from MTP (grey and black segments) and PC simulations (light and dark red segments). MTP simulations show how the C-Cl bond points towards the silica surface for $z < 5$ Å, and for larger values of z , the Cl atom interacts with the water and methanol oxygens and thus the molecule is oriented towards the mobile phase. On the contrary, for $z < 5$ Å PC simulations the C-Cl bond points towards the mobile phase.

Figure 8 shows the PhCl-sampled conformations using MTP electrostatics. At short distances from the stationary phase, PhCl samples two well-defined conformations corresponding to a halogen bond between Cl and the O of the silanol and to a hydrogen bond between Cl and the hydroxyl H of the silanol, respectively (Figure 8A). Further away from the surface ($z > 6$ Å) the C-Cl bond samples a wider range of conformations and the solvent molecules tend to structure between the analyte and the silica surface (Figure 8B). In summary, the positive cone of the halogen in PhCl acts as an anchor to engage in interactions with the silica surface (Figure 8A), which then reinforces its stability in the stationary phase (Table 3), and successively cheats its way towards the top.

4 Conclusion

Retention in HPLC simultaneously depends on various types of intermolecular interactions between the solute and the stationary phase, the solute and the mobile phase, and the stationary and

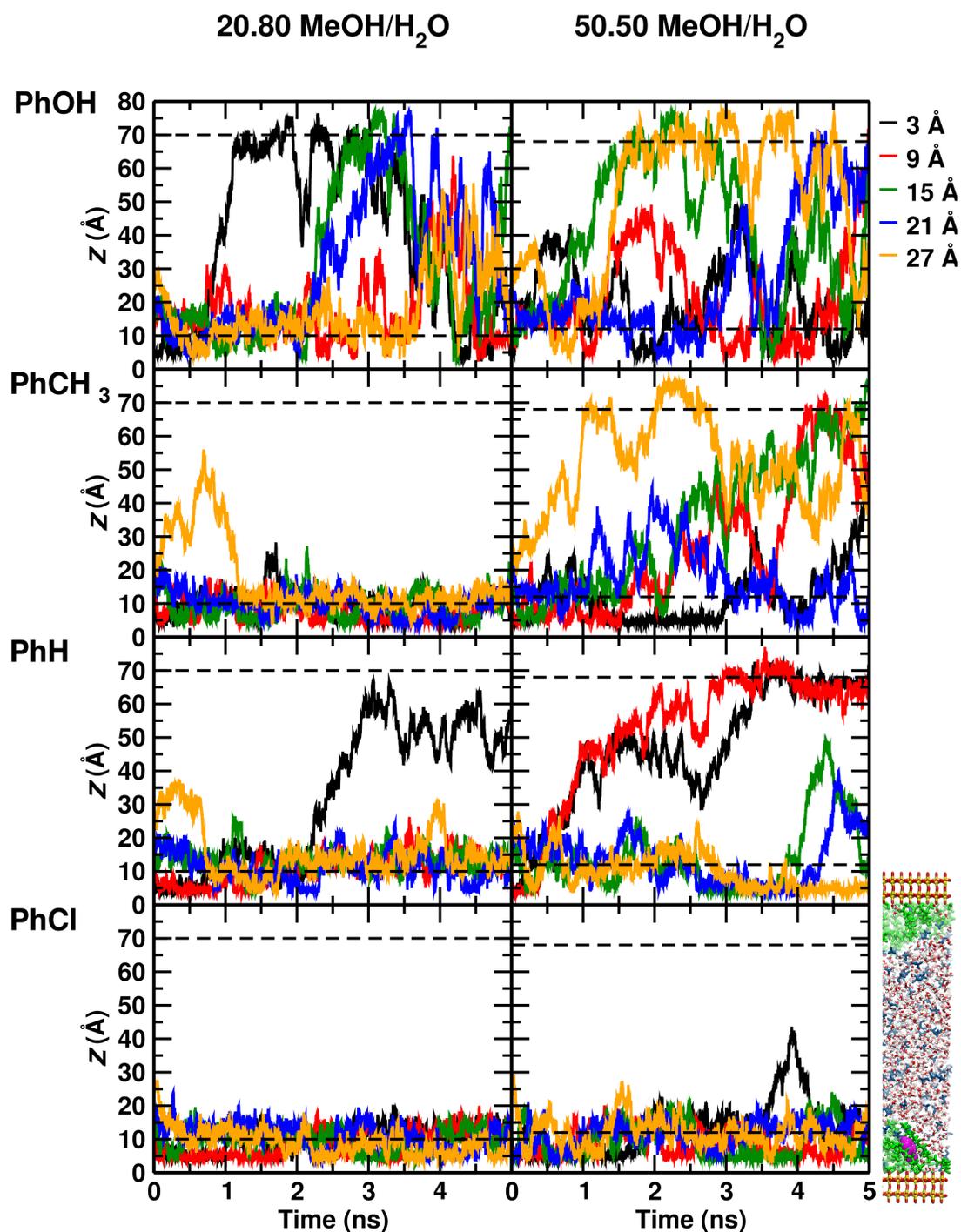


Fig. 5 Evolution of the Surface-to-center of mass (COM) of X distance along 5 ns for different initial position: 3, 9, 15, 21, and 27 Å and in both MeOH/H₂O solvent mixtures, 20:80 (left panels) and 50:50 (right panels), when using MTP electrostatics. The origin is at the position of the oxygen atoms of the silanol group on the surface. The inset at the bottom right of the figure illustrates a side view (*yz*-plane) of the starting structure after equilibration. The average extension of the alkyl chains are indicated by horizontal dashed lines (10 Å and 70 Å for 20:80 and 12 Å and 68 Å for 50:50 MeOH/H₂O solvent mixtures, respectively).³⁸

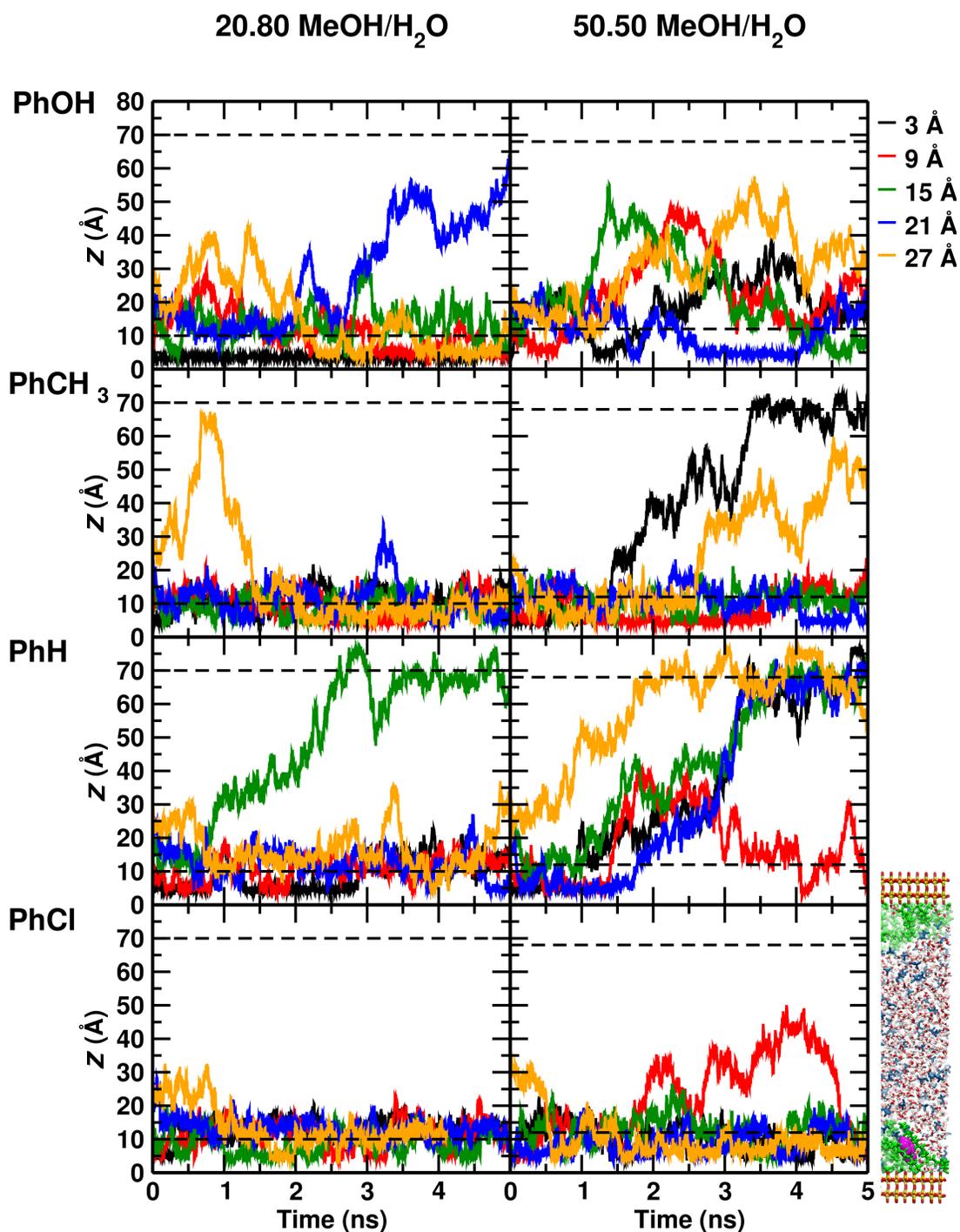


Fig. 6 Evolution of the Surface-to-center of mass (COM) of X distance along 5 ns for different initial position: 3, 9, 15, 21, and 27 Å and in both MeOH/H₂O solvent mixtures, 20:80 (left panels) and 50:50 (right panels), when using PC electrostatics. The origin is at the position of the oxygen atoms of the silanol group on the surface. The inset at the bottom right of the figure illustrates a side view (*yz*-plane) of the starting structure after equilibration. The average extension of the alkyl chains are indicated by horizontal dashed lines (10 Å and 70 Å for 20:80 and 12 Å and 68 Å for 50:50 MeOH/H₂O solvent mixtures, respectively).³⁸

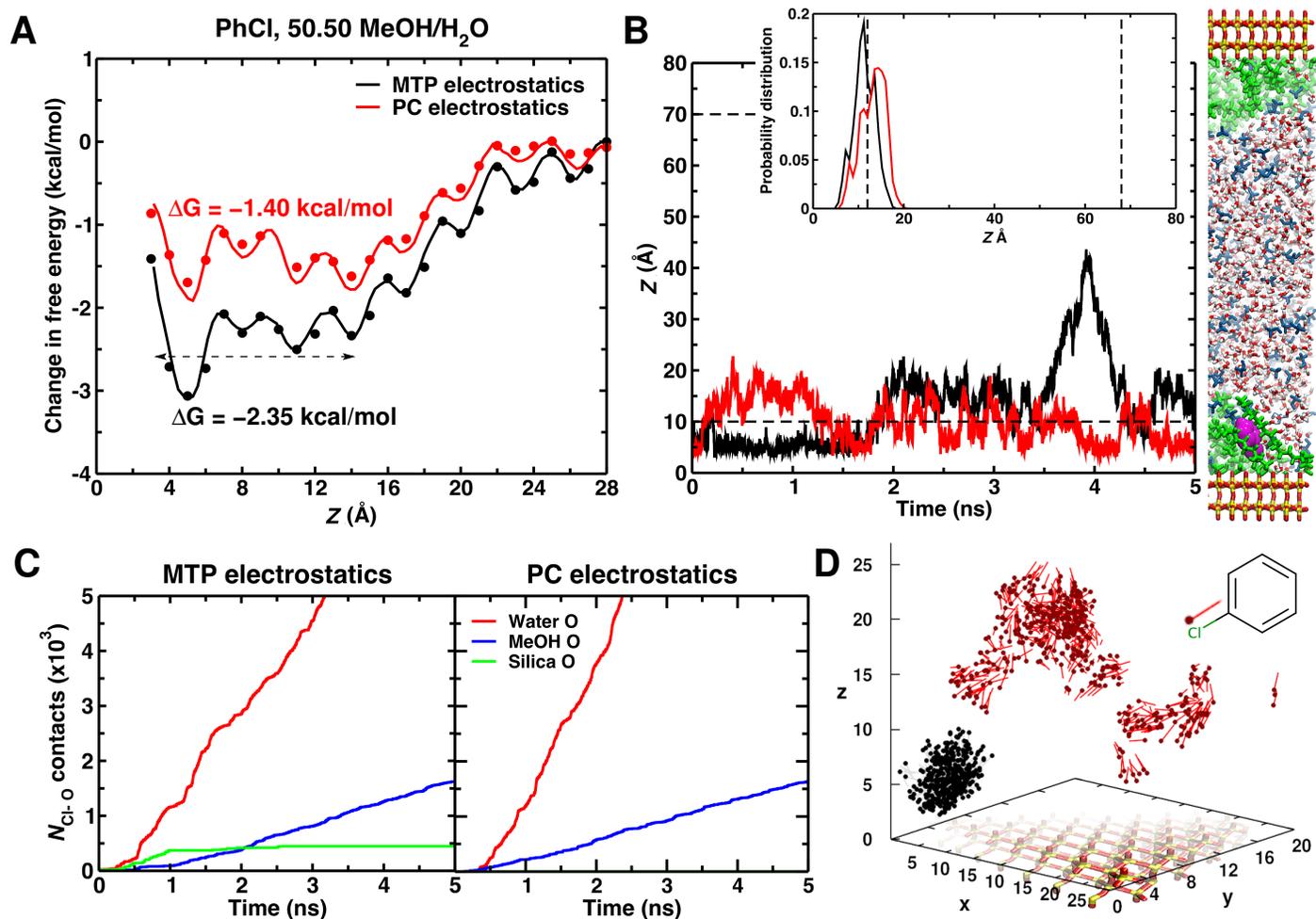


Fig. 7 Comparison of structural and thermodynamic properties for PhCl in 50:50 MeOH/H₂O solvent composition between MTP and PC simulations. (A) Free energy profiles with MTP (black) and PC (red) electrostatics. The stabilization free energies were averaged over z -ranging from 3 to 14 (marked by a black dashed arrow) and are given next to them. The points are raw data from the 28 umbrellas. (B) Evolution of the surface-to-PhCl center of mass distance $z(t)$ along 5 ns starting from $z = 3$ as initial position. The origin is at the position of the oxygen atoms of the silanol group on the surface, see sketch to the right. The inset shows the distance probability distribution function $P(z)$ averaged over 28 independent MD runs of 5 ns each (140 ns in total). The average extension of the alkyl chains (dashed lines) are at 10 Å and 70 Å.³⁸ (C) Cumulative number of contacts between Cl and O of water (red), methanol (blue) and silica (green). This was done by counting the oxygen atoms within 4 Å of Cl during 5 ns of MD simulation (starting from $z = 3$) using MTP (left) and PC (right) electrostatic representation of PhCl in 50:50 MeOH/H₂O solvent mixture. (D) Projection of the C-Cl bond orientation onto the xy -plane for different z -ranges, during the 1 ns of MD simulation starting from $z = 3$. The gray and red segments represent the C-Cl bond of PhCl from MTP and PC runs, respectively. The bold black and dark red dots highlight the Cl atom, to emphasize the orientation of the analyte relative to the stationary phase.

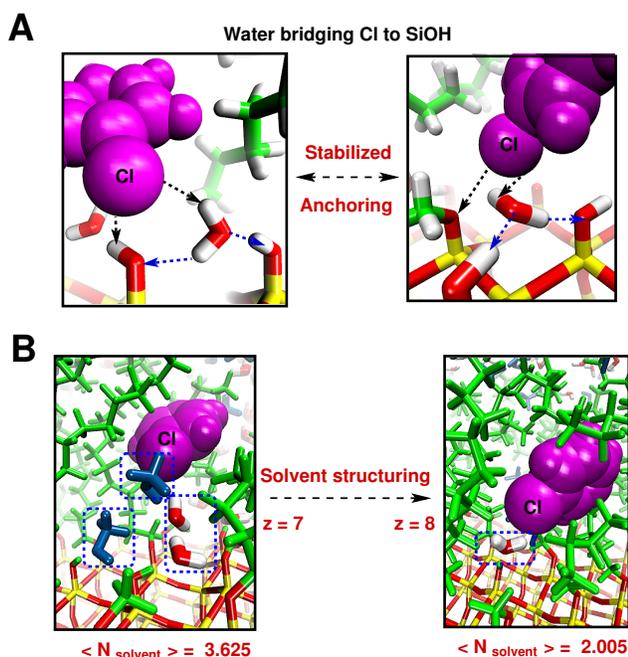


Fig. 8 Structural dynamical properties of PhCl intercalation in 50:50 MeOH/H₂O solvent composition. (A) Stabilized Cl anchoring by water bridges formation between Cl and the SiOH, whether Cl interacts with the silica hydrogen (left) or the silica oxygen (right). (B) Snapshots from two subsequent umbrella windows, at $z = 7$ and 8 Å, illustrating solvent molecules (MeOH/H₂O). The average number of solvent molecules $\langle N \rangle$ in a 6 Å shell from the center of mass of the solute calculated from the US trajectories at $z = 7$ and 8 Å is also indicated. Intercalating solvent molecules affecting the calculated free energies are highlighted by dashed squares.

the mobile phases. The present work demonstrates that the combination of accurate MTP force fields and free energy simulations provides the necessary accuracy required for quantitative interpretations of experiments, contrary to entirely PC-based models. It is shown that MD simulations lead to a more detailed microscopic understanding of the solute transfer process, and these studies show that experimentally relevant quantities in chromatography can be accurately computed with current computer capabilities. The results rationalize previous difficulties^{24,25,30} encountered in studying retention thermodynamics, explore the effects of a larger number of chromatographic parameters, and extend MD sampling techniques from methane^{9,11} to chemically more meaningful and challenging solutes.

Improved characterization of the molecular-level details of retention in chromatographic systems offers a number of future possibilities. For one, this will assist in predicting retention/separation times from computation and to eventually develop more effective, rapid scoring techniques based on molecular-level information. Secondly, this information can be

used for column design. In the past, a universal solute/solvent retention index system of RPLC has been developed and tested with a library of compounds and mobile phases.⁸¹ It has been found by examining RPLC retention data by principal component analysis (PCA)⁸² and target transformation factor analysis⁸³ that the data obtained from three different columns share a common factor space and that three vectors are sufficient to describe this retention data. The resulting eigenvector matrix associated with analyte compounds from singular value decomposition is found to be characteristic of the retention behavior of the compounds and independent of the mobile phases and reversed-phase columns used for the measurements.⁸¹ This opens up the possibility to combine results from quantitative MD simulations - which can generate large amounts of meaningful data with molecular resolution - with machine learning techniques. Another future avenue is to extract the enthalpic and entropic contributions for libraries of compounds in chromatographic systems of different chemical composition.

Acknowledgments

This work was supported by the Swiss National Science Foundation through grants 200021-7117810 and the NCCR MUST (to MM).

References

- 1 A. Fouqueau, M. Meuwly and R. J. Bemish, *J. Phys. Chem. B*, 2007, **111**, 10208–10216.
- 2 M. Orzechowski and M. Meuwly, *J. Phys. Chem. B*, 2010, **114**, 12203–12212.
- 3 P. K. Gupta and M. Meuwly, *J. Phys. Chem. B*, 2012, **116**, 10951–10959.
- 4 P. K. Gupta and M. Meuwly, *Chem. Phys. Chem.*, 2016, **17**, 2938–44.
- 5 R. K. Lindsey, J. L. Rafferty, B. L. Eggimann, J. I. Siepmann and M. R. Schure, *J. Chromatogr. A*, 2013, **1287**, 60 – 82.
- 6 J. L. Rafferty, J. I. Siepmann and M. R. Schure, *J. Chromatogr. A*, 2011, **1218**, 2203–2213.
- 7 J. L. Rafferty, J. I. Siepmann and M. R. Schure, *J. Chromatogr. A*, 2012, **1223**, 24–34.
- 8 K. El Hage, P. K. Gupta, R. Bemish and M. Meuwly, *The Journal of Physical Chemistry Letters*, 2017, **8**, 4600–4607.
- 9 S. J. Klante and T. L. Beck, *J. Chem. Phys.*, 1996, **100**, 5931–5934.
- 10 T. L. Beck and S. J. Klante, in *Unified Chromatography*, American Chemical Society, Washington, DC, 1999, ch. 5, pp. 67–81.
- 11 J. T. Slusher and R. D. Mountain, *J. Phys. Chem. B*, 1999, **103**, 1354–1362.

- 12 G. Eaton, A. S. Pena-Nunez and M. C. R. Symons, *J. Chem. Soc. Faraday Trans.*, 1988, **84**, 2181–2193.
- 13 D. S. Venables and C. A. Schmuttenmaer, *J. Chem. Phys.*, 1988, **108**, 4935–4944.
- 14 D. S. Venables and C. A. Schmuttenmaer, *J. Chem. Phys.*, 2000, **113**, 11222–11236.
- 15 K. L. Rowlen and J. M. Harris, *Anal. Chem.*, 1991, **63**, 964–969.
- 16 E. V. Goldammer and H. G. Hertz, *J. Phys. Chem.*, 1970, **74**, 3734–3755.
- 17 H. Leiter, K. J. Patil and H. G. Hertz, *J. Solution. Chem.*, 1983, **12**, 503–517.
- 18 E. D. Dawson and S. L. Wallen, *J. Am. Chem. Soc.*, 2002, **124**, 14210–14220.
- 19 R. D. Mountain, *J. Phys. Chem. A*, 1999, **103**, 10741–10748.
- 20 K. A. Lippa and L. C. Sander, *J. Chromatogr. A*, 2006, **1128**, 79–89.
- 21 J. L. Rafferty, J. I. Siepmann and M. R. Schure, *J. Chromatogr. A*, 2008, **1204**, 11–19.
- 22 J. L. Rafferty, J. I. Siepmann and M. R. Schure, *J. Chromatogr. A*, 2009, **1216**, 2320–2331.
- 23 A. Braun, J. Fouqueau, M. Meuwly and R. J. Bemish, *Phys. Chem. Chem. Phys.*, 2008, **10**, 4765–4777.
- 24 J. G. Dorsey and K. A. Dill, *Chem Rev*, 1989, **89**, 331–346.
- 25 J. G. Dorsey and W. T. Cooper, *Anal. Chem.*, 1994, **66**, 857A–867A.
- 26 F. Gritti, G. Félix, M. F. Achard and F. Hardouin, *J. Chromatogr. A*, 2001, **922**, 51–61.
- 27 J. W. Wolcott, R. G. and Dolan, L. R. Snyder, S. R. Bakalyar, M. A. Arnold and J. A. Nichols, *J. Chromatogr. A*, 2000, **869**, 211–230.
- 28 K. Krupczyńska, B. Buszewski and P. Jandera, *Anal. Chem.*, 2004, **76**, 227A–234A.
- 29 B. Buszewski and S. Noga, *Anal. Bioanal. Chem.*, 2012, **402**, 231–247.
- 30 J. Dorsey, W. Cooper, J. Wheeler, H. Barth and J. Foley, *Anal. Chem.*, 1994, **66**, A857–A867.
- 31 K. B. Sentell and J. G. Dorsey, *J. Chromatogr. A*, 1989, **461**, 193–207.
- 32 G. E. Berendsen, K. A. Pikaart and L. d. Galan, *J. Liq. Chromatogr.*, 1980, **3**, 1437–1464.
- 33 L. A. Cole and J. G. Dorsey, *Anal. Chem.*, 1992, **64**, 1317–1323.
- 34 A. Alvarez-Zepeda, B. N. Barman and D. E. Martire, *Anal. Chem.*, 1992, **64**, 1978–1984.
- 35 S. Myers and B. A., *Nature Biotechnol.*, 2001, **19**, 727–730.
- 36 S. Ekins, J. Mestres and B. Testa, *Br. J. of Pharmacol.*, 2007, **152**, 9–20.
- 37 C. Taft, V. Da Silva and C. Da Silva, *J. Pharm. Sci.*, 2008, **97**, 1089–1098.
- 38 J. Braun, A. Fouqueau, R. J. Bemish and M. Meuwly, *Phys. Chem. Chem. Phys.*, 2008, **10**, 4765–4777.
- 39 B. Buszewski, Z. Suprynowicz, P. Staszczuk, K. Albert, B. Pfeleiderer and E. Bayer, *J. Chromatogr. A*, 1990, **499**, 305–316.
- 40 L. C. Sander, K. A. Lippa and S. A. Wise, *Anal. Bioanal. Chem.*, 2005, **382**, 646–668.
- 41 M. L. Miller, R. W. Linton, S. G. Bush and J. W. Jorgenson, *Anal. Chem.*, 1984, **56**, 2204–2210.
- 42 S. A. Wise and L. C. Sander, *J. High Resol. Chromatogr.*, 1985, **8**, 248–255.
- 43 P. Jandera, *Liquid Chromatography-Normal Phase*, Encyclopedia of analytical science, Elsevier, Oxford, 2nd edn, pp. 142–152.
- 44 K. El Hage, T. Bereau, S. Jakobsen and M. Meuwly, *J. Chem. Theory. Comput.*, 2016, **12**, 3008–3019.
- 45 F. Hédin, K. El Hage and M. Meuwly, *J. Chem. Inf. Model.*, 2016, **56**, 1479–1489.
- 46 R. Dennington, T. Keith and J. Millam, *GaussView Version 5*, Semichem Inc. Shawnee Mission KS 2009.
- 47 C. Reichardt, *Solvents and Solvent Effects in Organic Chemistry*, Wiley-VCH Publishers, 3rd edn, 2003.
- 48 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and C. L. Klein, *J. Chem. Phys.*, 1983, **79**, 926–935.
- 49 A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. Dunbrack Jr., J. Evanseck, M. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, III, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.
- 50 B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *Journal of Computational Chemistry*, **30**, 1545–1614.
- 51 D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *Journal of Computational Chemistry*, **26**, 1668–1688.
- 52 C. Oostenbrink, A. Villa, A. E. Mark and W. F. V. Gunsteren, *Journal of Computational Chemistry*, **25**, 1656–1676.
- 53 G. A. Kaminski, R. A. Friesner, J. Tirado-Rives and W. L. Jorgensen, *The Journal of Physical Chemistry B*, 2001, **105**, 6474–6487.
- 54 O. T. Unke, M. Devereux and M. Meuwly, *J. Chem. Phys.*,

- 2017, **147**, 161712.
- 55 C. Kramer, P. Gedeck and M. Meuwly, *J. Comput. Chem.*, 2012, **33**, 1673–1688.
- 56 C. Kramer, P. Gedeck and M. Meuwly, *J. Chem. Theory. Comput.*, 2013, **9**, 1499–1511.
- 57 C. Kramer, T. Bereau, A. Spinn, K. R. Liedl, P. Gedeck and M. Meuwly, *J. Chem. Inf. Model.*, 2013, **53**, 3410–3417.
- 58 W. L. Jorgensen and P. Schyman, *J. Chem. Theory. Comput.*, 2012, **8**, 3895–3901.
- 59 T. Bereau, C. Kramer and M. Meuwly, *J. Chem. Theory. Comput.*, 2013, **9**, 5450–5459.
- 60 K. Vanommeslaeghe and A. D. MacKerell, Jr., *J. Chem. Inf. Model.*, 2012, **52**, 3144–3154.
- 61 K. Vanommeslaeghe, E. P. Raman and A. D. MacKerell, Jr., *J. Chem. Inf. Model.*, 2012, **52**, 3155–3168.
- 62 T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007–1023.
- 63 D. E. Woon and T. H. Dunning, *J. Chem. Phys.*, 1993, **98**, 1358–1371.
- 64 K. Vanommeslaeghe and A. D. MacKerell, *J. Chem. Inf. Model.*, 2012, **52**, 3144–3154.
- 65 S. Kim, P. A. Thiessen, E. E. Bolton, J. Chen, G. Fu, A. Gindulyte, L. Han, J. He, S. He, B. A. Shoemaker, J. Wang, B. Yu, J. Zhang and S. H. Bryant, *Nucl. Acids Res.*, 2016, **44**, D1202–D1213.
- 66 D. L. Mobley, *eScholarship*, 2013.
- 67 D. L. Mobley and J. P. Guthrie, *J Comput Aided Mol Des*, 2014, **28**, 711–720.
- 68 B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan and M. Karplus, *J. Comput. Chem.*, 1983, **4**, 187–217.
- 69 J. P. Ryckaert, G. Ciccotti and H. J. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.
- 70 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
- 71 B. R. Brooks, et al, *J. Comput. Chem.*, 2009, **30**, 1545–1614.
- 72 G. M. Torrie and J. P. Valleau, *J. Comput. Phys.*, 1977, **23**, 187–199.
- 73 S. Kumar, J. M. Rosenberg, D. Bouzida, R. H. Swendsen and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 1011–1021.
- 74 B. Roux, *Comput. Phys. Commun.*, 1995, **91**, 275–282.
- 75 K. El Hage, P. Mondal and M. Meuwly, *Molecular Simulation*, 2018, **0**, 1–18.
- 76 C. P. R. Walter J. Weber Jr, Yu-Ping Chin, *Wat. Res.*, 1986, **20**, 1433–1442.
- 77 M. E. Montgomery, M. A. Green and M. J. Wirth, *Anal. Chem.*, 1992, **64**, 1170–1175.
- 78 K. El Hage, J.-P. Piquemal, Z. Hobaika, R. G. Maroun and N. Gresh, *J. Comput. Chem.*, 2015, **36**, 210–221.
- 79 K. El Hage, J.-P. Piquemal, Z. Hobaika, R. G. Maroun and N. Gresh, *J. Comput. Chem.*, 2013, **34**, 1125–1135.
- 80 K. El Hage, V. Pandeyarajan, N. B. Phillips, B. J. Smith, J. G. Menting, J. Whittaker, M. C. Lawrence, M. Meuwly and M. A. Weiss, *J. Biol. Chem.*, 2016.
- 81 C. E. Reese, L. Huang, S.-H. Hsu, S. Tripathi and C. H. Lochmüller, *J. Chromatogr. Sci.*, 1996, **34**, 101–110.
- 82 H. Hotelling, *Journal of Educational Psychology*, 1933, **24**, 417–441.
- 83 E. Malinowski, *Factor Analysis in Chemistry.*, Wiley, New York, NY, 1991.