

Testing the Limitations of MD-based Local Electric Fields Using the Vibrational Stark Effect in Solution: Penicillin G as a Test Case

Jacek Kozuch,^{a†*} Samuel H. Schneider,^{a†} Chu Zheng,^a Zhe Ji,^a Richard T. Bradshaw,^b and Steven G. Boxer^{a*}

^aDepartment of Chemistry, Stanford University, Stanford, CA 94305-5012, USA

^bDepartment of Chemistry, King’s College London, Britannia House, 7 Trinity Street, London, SE1 1DB, UK
Email: jkozuch@stanford.edu; sboxer@stanford.edu

[†]Present addresses:

J.K. - Experimental Molecular Biophysics, Department of Physics, Freie Universität Berlin, 14195 Berlin, Germany

S.H.S. - Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94158, USA

Abstract: Non-covalent interactions underlie nearly all molecular processes in the condensed phase from solvation to catalysis. Their quantification within a physically consistent framework remains challenging. Experimental vibrational Stark effect (VSE)-based solvatochromism can be combined with molecular dynamics (MD) simulations to quantify the electrostatic forces in solute-solvent interactions for small rigid molecules and, by extension, when these solutes bind in enzyme active sites. While generalizing this approach towards more complex (bio)molecules, such as the conformationally flexible and charged penicillin G (PenG), we were surprised to observe inconsistencies in MD-based electric fields. Combining synthesis, VSE spectroscopy, and computational methods, we provide an intimate view on the origins of these discrepancies. We observe that the electric fields are correlated to conformation-dependent effects of the flexible PenG side-chain, including both local solvation structure and solute conformational sampling in MD. Additionally, we identified that MD-based electric fields are consistently overestimated in 3-point water models in the vicinity of charged groups; this cannot be entirely ameliorated using polarizable force fields (AMOEBA) or advanced water models. This work demonstrates the value of the VSE as a direct method for experiment-guided refinements of MD force fields and establishes a general reductionist approach to calibrating vibrational probes for complex (bio)molecules.

1. Introduction

Electrostatic interactions are one of the major physical forces that govern the strength of non-covalent interactions in condensed phase systems ranging from simple solvents to complex materials and biological assemblies. As such, traditional experimental observables, e.g. solvation energies, binding constants, enzymatic rates, and ion channel conductance, among many others, are often described as having a significant electrostatic origin.^{1–3} These experimental observables arise from complex, dynamic convolutions of many local interactions between multiple functional groups or atoms. This is particularly true for biomolecules, such as metabolites, peptides, proteins, and DNA, whose intrinsic chemical complexity has evolved towards highly specific biological function and interactions.

There have been extensive efforts into understanding how these experimental observables correlate with the properties of solvation, lipid or protein environments, for instance, using dyes molecules as reporters.^{4–7} However, it is rarely possible to deconvolute such effects precisely into a set of relevant local atomic interactions. This issue impedes the use of direct physics-based interpretations to describe the underlying interatomic forces and quantify their contributions towards biological function. As a consequence, many

biophysical studies have utilized theoretical approaches in order to interpret experimental observables in terms of a molecular picture with dynamic information and atomistic resolution.⁸ For example, molecular dynamics (MD) has been utilized extensively to structurally rationalize the physical and chemical origins of many biological processes.^{9–12} Despite its ability to simulate structure and dynamics of condensed phase systems within a consistent classical mechanics approach, the reliability of MD simulations to recapitulate local physical forces is intrinsically limited by the benchmarking methods used to obtain MD force field parameters.¹³ Generally, force fields are carefully benchmarked and parameterized using either *ab initio* quantum chemistry simulations in the gas phase, average bulk properties (e.g. hydration energies, densities, and dielectric constants), and/or direct structural observables such as radial distribution functions of atoms in bulk media.^{14–17} However, there is no *a priori* guarantee that such careful parameterizations, in particular those based on bulk observables, will lead to accurate descriptors of the underlying *local* interactions that are crucial for many *in silico* applications such as the rational design of novel drugs, ligand docking, and *de novo* design of proteins and enzymes.

The vibrational Stark effect (VSE) is an approach that

enables quantification of the *local* electrostatic force, i.e. the electric field, underlying non-covalent interactions^{18–21} in various settings including solvents,^{22,23} electrode interfaces,^{24,25} and biological systems.^{26–28} The linear VSE describes the shift of the vibrational frequency ($\bar{\nu}_{obs}$; units of cm^{-1}) of a vibrational probe, arising from the interaction of its difference dipole or Stark tuning rate ($|\Delta\vec{\mu}_{probe}|$; units of $\text{cm}^{-1}/(\text{MV}/\text{cm})$), with the electric field exerted by the environment ($|\vec{E}_{env}|$; units of MV/cm) according to:

$$\bar{\nu}_{obs} = \bar{\nu}_0 - |\Delta\vec{\mu}_{probe}| \cdot |\vec{E}_{env}| \quad (1)$$

where $\bar{\nu}_0$ is the vibrational frequency in the absence of an electric field, i.e. in vacuum. Given the small size of vibrational Stark probes, such as the C=O or C≡N stretch, the VSE can be used to evaluate the projection of the electric field from the environment at precise and functionally relevant locations, such as specific chemical bonds undergoing a chemical transformation. This approach has been recently demonstrated to experimentally quantify the electrostatic contributions to catalytic rates of enzymes such as ketosteroid isomerase,^{29,30} 4-chlorobenzoyl-CoA dehalogenase and serine proteases,³¹ providing direct evidence for the role of large active site electric fields along reactive bonds to achieve significant rate accelerations, referred to as electrostatic catalysis.¹⁹

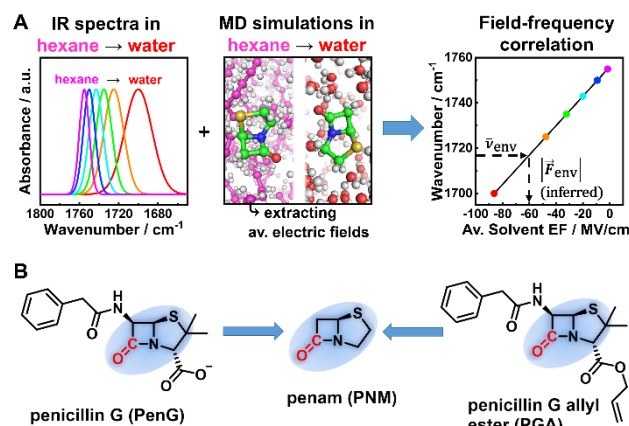


Figure 1. (A) Schematic illustration of the solvatochromic approach to obtain electric field-frequency calibrations utilizing the VSE. IR spectra of a VSE probe-containing solute are measured in solvents ranging from hexanes to water to obtain the vibrational frequencies of the probe’s IR absorption. MD simulations are then employed to reproduce the solvation environment and extract the average electric fields exerted by the solvent on the VSE probe. The average value of the simulated electric fields and experimental vibrational peak frequencies are combined in a field-frequency correlation plot and then modelled by the linear VSE equation (eq. 1) indicated by the black regression line through the data points. The electric fields sensed by the same VSE probe in, for instance, an environment of an enzyme active site, can be inferred from its vibrational frequency via the correlation, as indicated by the dashed arrows. (B) Structures of penicillin G (left), and penicillin G allyl ester (right), with the rigid β -lactam core containing the C=O vibrational probe (highlighted in red) and modeled with penam (center). The common penam core is highlighted in blue.

All applications of the VSE rely on a proper quantitative calibration of the vibrational probe in terms of the electric

field exerted on the bond due to the environment, e.g., the solvent. A particularly convenient method is the combination of experimental vibrational solvatochromism and MD simulations. This method provides an absolute field-frequency calibration over an electric field range of up to 100 MV/cm, relevant for investigations in complex environments, such as enzyme active sites.^{23,28,32–34} In this approach (schematically illustrated in Figure 1A), the experimental vibrational frequency of the VSE probe ($\bar{\nu}_{obs}$) in various solvents is correlated with the simulated average electric fields projected onto the probe bond axis and modeled using the linear VSE equation (Eq. 1). This field-frequency calibration provides (a) the magnitude of the Stark tuning rate as the *solvatochromic slope*, describing the sensitivity of the probe to the environment’s electric field, and (b) the zero-field vibrational frequency $\bar{\nu}_0$, i.e. the frequency in vacuum ($|\vec{E}_{env}| = 0$; the y-intercept). Implicit to this approach is the assumption that the MD simulations can accurately recapitulate the experimental solvation environment of the probe, enabling the field-frequency calibration to be used in more complex environments to quantify electrostatic effects. The correspondence between vibrational solvatochromism and MD simulations is supported by experimental results from vibrational Stark spectroscopy (VSS),^{23,31,35} which allows independent determination of the Stark tuning rate in a defined externally applied electric field³⁶ and supports the application of the linear VSE (Eq. 1). Further support is obtained by experiments demonstrating a solvent-independent anharmonicity²² as well as simulations at higher levels of theory using polarizable force fields, DFT, and QM/MM simulations.^{33,37,38}

In all prior work the model systems used for calibration were small, conformationally rigid molecules where the vibrational probe was either the sole or predominant polar functional group. Therefore, in order to advance the understanding of electrostatic interactions in biological systems using vibrational probes, it is important to expand the scope of this approach to include more complex (bio)molecules that are conformationally flexible, charged, or contain multiple polar functional groups. An expansion of this approach requires further testing, refinement, and benchmarking of the reliability of MD force fields for the determination of local electric fields. The VSE can augment this benchmarking approach by providing direct and quantitative information on the nature and strength of noncovalent molecular interactions in condensed phase systems with applications to complex biomolecules as demonstrated below.

Among the many complex and biomedically relevant molecules, β -lactam antibiotics are a prime example where conformational flexibility, charge, and multiple functional groups are integral to their overall function in antibiotic action and resistance.³⁹ This led to the development of many classes of β -lactam antibiotics, such as penicillins, cephalosporins, monobactams, etc.,⁴⁰ as well as various (bio)synthetic approaches to provide access to a steadily growing number of derivatives that can overcome resistance (at least temporarily).^{41,42} Penicillin G (PenG), which was the first β -lactam antibiotic used clinically, contains a rigid bicyclic β -lactam core structure (i.e. the penam structure) and a bulky,

flexible phenylacetamido (PAA) side chain as well as a negatively charged carboxylate group (Figure 1B). The antibiotic function and the mechanism of antibiotic resistance of PenG and all other β -lactam antibiotics are based on the reactivity of the β -lactam carbonyl group (highlighted in red color in Figure 1B) towards nucleophilic attack.⁴³ As such, determination of the parameters of the VSE-based field-frequency calibration of the β -lactam C=O would facilitate the analysis of the electrostatic contributions exerted by proteins that are responsible for antibiotic action and resistance, e.g. transpeptidases and β -lactamases.^{44,45}

In the course of extending the concept of electrostatic catalysis to TEM β -lactamases and their catalytically relevant interactions with PenG (and other β -lactams, such as cephalosporins), we encountered conflicting results in the solvatochromic calibration. As described in detail below, the solvatochromic response of PenG's β -lactam C=O group evaluated using fixed charge (FC) and polarizable (POL) MD simulations suggests an anomalously low Stark tuning rate in contrast to previously benchmarked carbonyl compounds. As the role of MD simulations in the solvatochromic calibration is to model the experimental solvation electrostatics, this observation suggests possible inaccuracies or oversimplifications in the methods and/or force fields used for determining the MD-derived solvation electric fields for PenG (and other solutes of similar structural complexity). In order to understand this observation, we synthesized the β -lactam penam core (PNM), i.e., the smallest rigid functional moiety of PenG, as well as the PenG allyl ester (PGA), where the charged carboxylate is chemically masked (Figure 1B). Given that the core structure containing the β -lactam carbonyl is the same in these three molecules, we would expect that the β -lactam C=O groups of all three molecules exhibit similar Stark tuning rates. Therefore, using this set of molecules, their vibrational Stark spectra obtained in a defined external field, and their experimental VSE solvatochromism, we aimed to understand the effects of increasing chemical complexity, i.e., charge, conformational degrees of freedom, and multiple functional groups, on solvation and the corresponding electric fields as determined from MD simulations.

2. Materials and Methods

All organic solvents were obtained from Sigma Aldrich (anhydrous and at >99 % purity); D₂O (99 % atom D, Cambridge Isotope). Penicillin G sodium salt was obtained from Sigma Aldrich at highest grade available.

2.1 Synthesis

The syntheses of penicillin G allyl ester, PGA (allyl (2R,5R,6R)-3,3dimethyl-7-oxo-6-(2-phenylacetamido)-4-thia-1-azabicyclo[3.2.0]heptane-2-carboxylate), and penam, PNM (4-thia-azabicyclo[3.2.0]heptan-7-one), were performed according to the previously described procedure by Jeffrey *et al.*,⁴⁶ and by Chiba *et al.* and Cambie *et al.*, respectively,^{47,48} and are described in detail in the Supporting Information (SI Section 1).

2.2 Vibrational Solvatochromism

Infrared spectra were recorded using a Bruker Vertex 70 equipped with a liquid nitrogen-cooled mercury cadmium telluride (MCT) detector and under constant purging of the

sample chamber with dry air. As described previously,³² 20 – 30 μ L of the solute solutions were added into a demountable cell composed of two CaF₂ optical windows (19.05 mm diameter, 3 mm thickness, Lambda Research Optics, Inc.), which were separated by a suitable combination of 25, 50 or 100 μ m thick Teflon spacers. The concentration was ca. 10 mM. In cases where solubility became a limiting factor, a saturated solution was prepared and diluted by a factor of at least two before injection into the cell (to exclude solute aggregation, spectra were compared to those of even more diluted solutions). Each recorded FTIR transmission spectrum consisted of 256 scans, which, depending on the solute concentration, was repeated, and averaged to reduce the noise. Spectra were recorded in a spectral window of 4000 to 1000 cm^{-1} with 1 cm^{-1} spectral resolution and an aperture of 2 to 3.5 mm. Solvent spectra were recorded at similar conditions and used as reference to obtain the absorbance spectra according to the Lambert-Beer law. Peak positions were determined using a pseudo-Voigt fitting in OPUS 6.5.

2.3 Vibrational Stark Spectroscopy (VSS)

PenG, PNM and PGA were dissolved in glass forming solvents as specified for each experiment at concentrations of 100 mM. Each solution was injected in a sample-cell consisting of two offset CaF₂ optical windows (1 mm thickness, 12.7 mm diameter, FOCtek Photonics), which were coated with a 4.5 nm Ni layer and separated by Teflon spacers of a thickness of approximately 26 μ m. The filled sample cell is immediately immersed into liquid N₂ in a custom-built cryostat.⁴⁹ Vibrational Stark spectra were recorded on the Bruker Vertex 70 described above, with voltage ranging from 1.5 to 3.6 kV applied using a Trek 10/10 high-voltage power amplifier (corresponding to applied fields of 0.5–1.4 MV/cm taking into account the spacer thickness), 1 cm^{-1} resolution, 64 scans of field-on/off transmission spectra each. Stark spectra at different fields were compared to ensure that intensities scale linearly with the square of the field strength as expected for an isotropic sample.⁵⁰ The Stark tuning rates were determined from the second derivative contribution of the Stark spectra, which was obtained from a numerical fit of the zeroth, first, and second derivative contributions of the best-fit Voigt profile of the experimental low temperature absorbance spectrum, as described previously.⁵⁰ The resulting Stark tuning rates are reported as $|\Delta\tilde{\mu}|f$, where the local field factor f denotes the difference between applied electric field and the actual electric field experienced by the probe,⁵⁰ as discussed in detail with the results.

2.4. Fixed Charge Molecular Dynamics Simulations

Fixed charge (FC) MD simulations were performed using the general AMBER force field (GAFF)^{51,52} in GROMACS 2018,⁵³ as described previously.³² Briefly, parameters of the solutes were obtained using AmberTools18 (AM1-BCC)^{54,55} based on optimized structures from DFT simulations in Gaussian⁵⁶ (b3lyp/6-311++g(2d,2p)). The solute was placed in the center of a cubic solvent box (edges were at a distance of 2.0 nm from solutes and periodic boundary conditions were used; organic solvent parameters were obtained from virutalchemistry.org;^{13,57} water models were used as implemented in GROMACS or taken from literature as described in the main text) and, in the case of

penicillin G, one solvent molecule was replaced by a sodium ion (ff99SB-ildn). The system was energy-minimized (steepest descent until forces $< 1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$) and equilibrated for 100 ps as an NPT ensemble at 300 K and 1 bar (2 fs steps; Bussi thermostat with time constant of 0.1 ps; Berendsen barostat; PME method, cutoff distances set to 1.0 nm; constraints using the LINCS algorithm). MD production runs were carried out over 2 ns using the stochastic dynamics integrator implemented in GROMACS (SD with inverse friction constant of 2 ps; Parrinello-Rahman barostat; otherwise, same conditions as in the NPT equilibration). Full precision coordinates and forces at the C and O atoms of the C=O bond of interest were extracted every 100 steps. The same trajectory was rerun (i.e., atom coordinates were kept as in the MD production run using keyword “rerun”) with all partial charges of the solvents neutralized, to obtain forces absent of intermolecular electrostatic interactions. The difference between the forces from both trajectories yielded electrostatic forces due to the solvent, which were divided by the respective partial charge to yield the solvent electric fields \vec{F}_C and \vec{F}_O at the C and O, respectively.³²

Solute side-chain electric fields were extracted by subtracting the total intramolecular forces (from the rerun trajectory with discharged solvent) and the residual non-electrostatic forces from an additional rerun trajectory with discharged solvent and discharged solute, except for the C=O bond of interest. It should be noted that describing side-chain MD electric fields on a meaningful absolute scale is not straightforward due to the simplified assumption of atomic point charges and force field-specific cutoffs; the latter is set in GAFF such that electrostatic forces are only considered for 1-4 and more distant bonded interactions, with force field specific scaling factors. In the simplest attempt to correct for parameter-specific contributions, the side-chain fields were obtained by subtracting the total intramolecular field of PNM from the total intramolecular electric fields of PenG and PGA for a given solvent. Using this approach, we self-consistently remove the contribution of the shared rigid penam core (as indicated in Figure 1B), whereas all other electrostatic forces due to the side-chain groups and their degrees of freedom are accounted for in the average side-chain electric field. A detailed scheme summarizing each step of how electric fields are obtained using MD is shown in Scheme S1 and Table S1.

2.5 Polarizable Molecular Dynamics Simulations

Polarizable (POL) MD simulations were performed using the AMOEBA09 force field⁵⁸ in Tinker 8.7,⁵⁹ as described previously.³³ Water models were used as elaborated in Section 3. Parameters for tetrahydrofuran and hexane were taken from our previous work³³ and dibutyl ether was parameterized using Poltype with multipoles fitted to electrostatic potentials from MP2/aug-cc-pvtz calculations.⁶⁰ All other solvents were used as implemented in the AMOEBA09 force field. Solutes were parameterized using Poltype, with polarization groups assigned between conformationally rigid groups⁶¹ (the penam unit was one polarization group in each case) and with electrostatic potential fits at the MP2/6-311g++(2d,2p) level. For PenG and PGA, the fit was performed simultaneously to 4 and 6 individual

conformations, respectively. Initial torsions of PGA from Poltype were refined based on the phenylacetamido parameters of PenG and parameters obtained for allyl acetate.

Solutes were placed in the center of a cubic solvent box with edges at a distance of 20 Å from solutes (periodic boundary conditions were used) and one solvent molecule was replaced by a sodium ion for PenG (AMOEBA09). Each system was first energy-minimized (steepest descent until forces $< 0.5 \text{ kcal mol}^{-1} \text{ Å}^{-1}$), and then underwent an NVT-equilibration for 50 ps at 300 K and an NPT-equilibration for 100 ps at 300 K and 1 bar, both with 1 fs steps. Minimization, NVT and NPT steps were performed with a convergence threshold for induced dipoles of 10^{-2} D , a 9.0 Å van der Waals cutoff, a 7.0 Å real space electrostatics cutoff, as well as the particle mesh Ewald method and analytical van der Waals correction. In the equilibration steps, the velocity-Verlet integrator was used with the Andersen thermostat, and the Berendsen barostat was added to the NPT step with a pressure coupling time constant of 1 ps. MD production runs were performed over 500 ps with 1 fs steps, an increased van der Waals cutoff of 12.0 Å, and lowered induced dipole threshold to 10^{-5} D . The Beeman integrator was used together with the Bussi thermostat and the Monte-Carlo barostat (inverse friction constant 1 ps; ‘molecular’ volume-scaling). The induced dipoles were written into a separate file using the keyword ‘save-induced’. Next, all solvent molecules were removed from the trajectory and the resulting intramolecular induced dipoles were extracted using the ANALYZE option in Tinker. As outlined previously and along a similar line of reasoning as above, the difference between extracted induced and intramolecularly (solute side-chain)-induced dipoles was taken and divided by the polarizability parameter of the corresponding atom (multiplying by 299.79 to convert from D Å^{-3} to MV/cm) in order to obtain the solvent-exerted electric fields \vec{F}_C and \vec{F}_O at the C and O, respectively.³³

Solute side-chain electric fields were obtained by redefining the β -lactam C=O group as a separate polarization group and re-calculating the intramolecular induced dipole moments. Similar to the approach described for FC MD, these electric fields were offset to the value of PNM, to isolate the contributions due to the side-chains (see Scheme S1 and Table S1 for a summary of each step).

2.6 Electric Field on Projected C=O Bond

Electric fields exerted on the C=O bond in both FC and POL simulations were determined as the average along the C=O axis (using the unit vector of the CO bond \hat{u}_{CO})³²

$$F = |\vec{F}_{\text{env}}| = \frac{1}{2} (\vec{F}_C \cdot \hat{u}_{CO} + \vec{F}_O \cdot \hat{u}_{CO}). \quad (2)$$

Note that in this work the shorter notation F will be used to unclutter figures and captions but is interchangeable with the more complete notation $|\vec{F}_{\text{env}}|$. Finally, the obtained fields are averaged over the entire trajectory for the correlation with the vibrational frequencies from solvatochromic experiments. While this approach does not account for effects due to electric field gradients, we show in Table S2 that gradients are consistent across the three molecules.

3. Results and Discussion

3.1. Inconsistencies in MD-Solvatochromic Calibration of Penicillin G

Solvatochromic Calibration of PenG. For solvatochromic calibration, IR spectra are acquired in a set of solvents, typically ranging from hexanes to water, so that a large span of solvent electric fields is sampled, typically over a range of ca. 50 to 100 MV/cm for carbonyl groups. This is illustrated in Figure 2A using the C=O stretch of acetophenone (ACP), which is an extensively benchmarked carbonyl probe.^{32,33} As shown in previous studies, evaluating the C=O peak position in terms of the average solvent electric field using both fixed charge (FC; tip3p used for water) as well as polarizable MD simulations using AMOEBA09 (POL; amoeba03 water as implemented in the AMOEBA09 parameters) yields linear field-frequency plots in accordance with the linear VSE.³³ Note that all field-frequency plots in the following figures span the same ranges on both axes to aid comparisons; in this way the solvatochromic slopes, representing the apparent Stark tuning rates, can be compared visually. For ACP, both levels of theory result in similar Stark tuning rates of the C=O stretch of ca. $0.4 \text{ cm}^{-1}/(\text{MV}/\text{cm})$. The offset of ca. 5 cm^{-1} (corresponding to differences of the zero-field frequency) results from the explicit consideration or absence of polarization effects in POL or FC calculations, respectively, and was discussed in detail in earlier work.³³

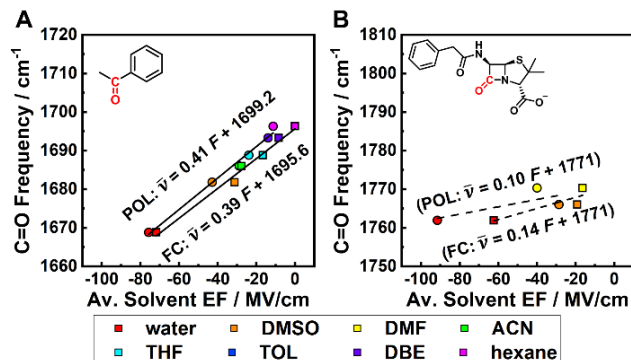


Figure 2. Comparison of the MD-solvatochromic field-frequency calibration of the C=O shown in red for (A) acetophenone and (B) PenG. Observed carbonyl peak frequencies are plotted on the vertical axis and average calculated values of the electric field projected on the C=O probe on the horizontal axis. Squares and circles represent average solvent electric fields obtained from fixed charge (FC) and polarizable (POL) MD simulations, respectively, and black lines are linear fits to the data. Due to the limited number of data points (limited by solubility), the linear fits of PenG in (B) are only indicated as dashed lines. For ACP, the linear fit yields $\bar{\nu} = (0.41 \pm 0.02) F + (1699.2 \pm 0.9)$ (POL, with $R^2 = 0.99$) and $\bar{\nu} = (0.39 \pm 0.02) F + (1695.6 \pm 0.7)$ (FC, with $R^2 = 0.98$); for PenG, the linear regression results in $\bar{\nu} = (0.10 \pm 0.08) F + (1771 \pm 5)$ (POL, with $R^2 = 0.58$) and $\bar{\nu} = (0.14 \pm 0.05) F + (1771 \pm 3)$ (FC, with $R^2 = 0.86$).

In contrast to small model compounds like ACP, a common problem faced when working with biomolecules is their limited solubility in a subset of solvents. PenG is a typical example due to the negatively charged carboxylate, so IR spectra of the sodium salt can only be acquired in D₂O, DMSO and DMF (Figure 2B, see Figure S1 for IR spectra). While one can protonate the carboxylate to increase

solubility, this can lead to potential inconsistencies in interpretation based on eq. 1 due to shifted values of $\bar{\nu}_0$ between protonated and deprotonated PenG, which cannot be corrected for in a straightforward manner. Using the solvatochromism of PenG in D₂O, DMSO, and DMF, it appears that the C=O frequency is relatively invariant across these different solvents ($\Delta\bar{\nu} \sim 8 \text{ cm}^{-1}$). This is in contrast to many previously calibrated carbonyls, such as ACP (Fig. 2A), where the solvatochromic shift is considerably larger ($\Delta\bar{\nu} \sim 20 - 30 \text{ cm}^{-1}$) across the same set of solvents.³¹ This weak solvatochromic shift is peculiar when compared to the calculated average electric fields from FC and POL MD simulations (Figure 2B), which span ca. 60 MV/cm between water and DMSO or DMF, a similar range as obtained between water and hexanes for ACP. Thus, based on the correlation of fields and frequencies (Figure 2B), the resulting apparent Stark tuning rate, as derived from the solvatochromic slope, of the β -lactam C=O of PenG would be ca. $0.1 - 0.14 \text{ cm}^{-1}/(\text{MV}/\text{cm})$, which is anomalously low relative to other carbonyl model compounds benchmarked with this approach.³¹

Solvatochromic Calibration and Vibrational Stark Spectroscopy of PenG Derivatives. In order to investigate the possible origins of the anomalously low Stark tuning rate in the MD-assisted vibrational solvatochromism analysis of PenG, we first tested whether the VSE of the β -lactam C=O of the bicyclic penam core of PenG is governed by similar intrinsic Stark tuning rates, independent of additional chemical complexity, charge, flexibility, etc. We synthesized PNM, the smallest rigid functional moiety of PenG, and the neutral PGA (Figure 1B).⁴⁶⁻⁴⁸ We performed VSS, in which each of the molecules is isotropically immobilized in a frozen glass matrix (at 77 K) and exposed to a defined external electric field, in order to quantify the experimental Stark tuning rate.³⁶ Figures 3A – C show the 77K IR absorbance and Stark spectra in the C=O spectral range, the latter represented as difference spectra in the presence and absence of the externally applied electric field. The Stark spectrum of the β -lactam C=O is dominated by the characteristic second derivative contribution of the absorption spectra, from which the Stark tuning rate can be extracted as $|\Delta\bar{\nu}|f$, that is, convoluted with the local field factor f .⁶² This factor describes the difference between the externally applied and the locally sensed electric field ($f = \vec{F}_{local}/\vec{F}_{ext}$) and was observed to have a value of $ca. 2.0 \pm 0.5$, independent of the employed solvent³¹ (the local field factor is described in more detail in the supporting material of ref. 18 and 63). The Stark tuning rates of the β -lactam C=O stretch of all three molecules are observed to be *identical* within experimental and fitting error in a narrow range of $1.40 - 1.44 \text{ cm}^{-1}/(\text{MV}/\text{cm})$, and in line with observed values for other carbonyl groups in prior work.³¹ This accordance of the Stark tuning rates directly confirms our hypothesis that the *intrinsic* Stark tuning rates of PenG, PGA and PNM are similar due to the shared, rigid penam moiety. As such, we might expect that solvatochromic calibrations of the three molecules should result in similar solvatochromic slopes.

We then compared the solvatochromic calibrations of the three molecules to determine the origins of the anomalous field-frequency calibration of PenG. The solubility of PNM permits solvatochromic experiments over the entire solvent polarity range from hexanes to water. The corresponding solvatochromic calibration found using FC and POL MD fields (Figure 3E), both spanning a range of ca. 70 MV/cm, yields consistent solvatochromic slopes of 0.78 and 0.69 $\text{cm}^{-1}/(\text{MV}/\text{cm})$, respectively. These values lie in the expected range for aliphatic amide carbonyls,^{31,33} and, importantly, are in line with the Stark tuning rates of $|\Delta\vec{\mu}|f \approx 1.4$

$\text{cm}^{-1}/(\text{MV}/\text{cm})$ as determined using VSS ($f \approx 2$, Fig. 3A-C). The esterification of PenG, yielding PGA, also improves the solubility across the entire solvent range and allows the corresponding analysis (Figure 3F) providing solvatochromic slopes of 0.46 and 0.39 $\text{cm}^{-1}/(\text{MV}/\text{cm})$ with FC and POL MD solvent fields, respectively. While both values are still consistent with other carbonyl probes, they are lower by a factor of ca. 1.8 relative to PNM.

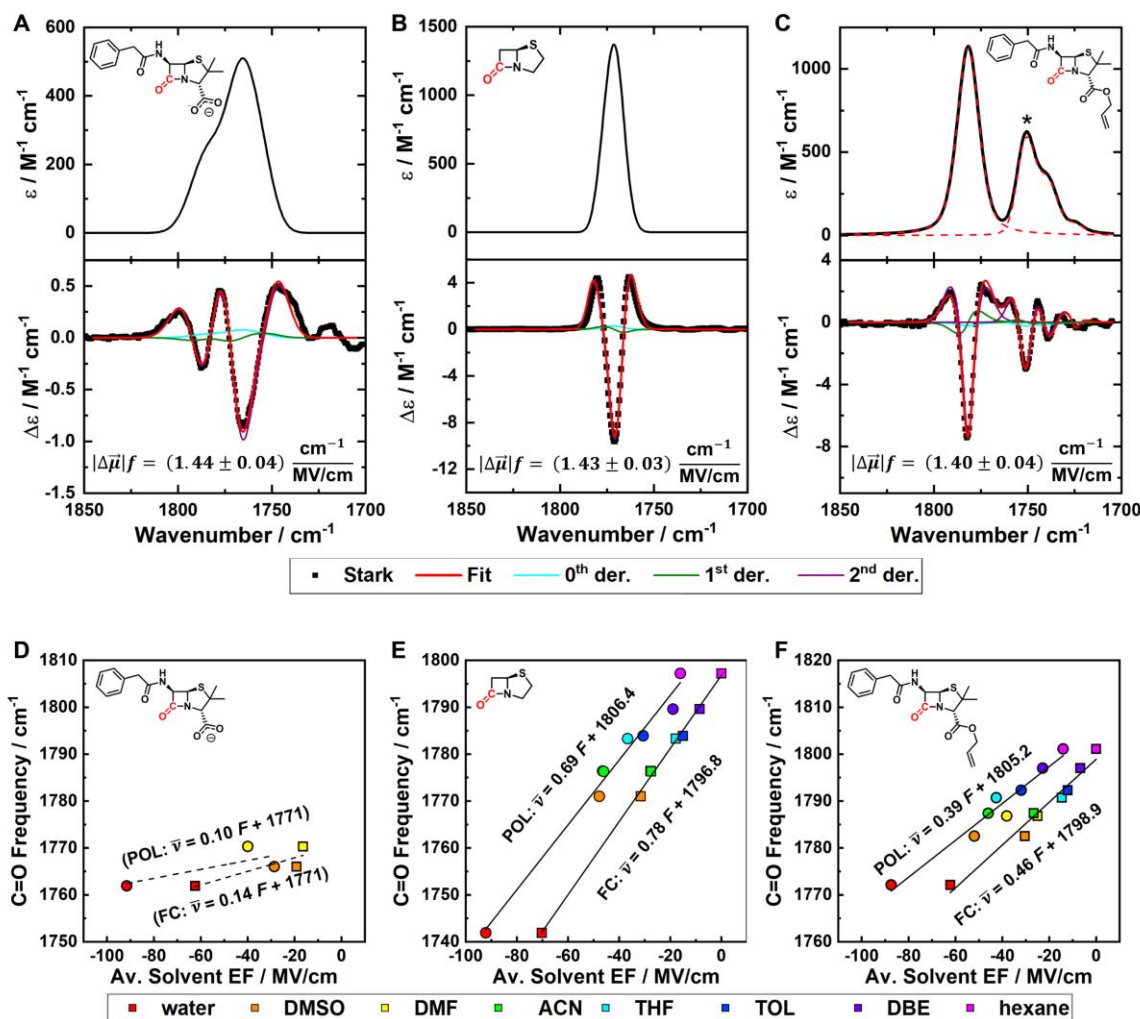


Figure 3. Vibrational Stark spectra and MD-solvatochromic field-frequency calibrations of β -lactam C=O modes of (A & D) PenG, (B & E) PNM, and (C & F) PGA (structures are shown as insets). Top panel: Infrared absorbance and vibrational Stark spectra obtained at 77 K in 1:1 glycerol/D₂O for PenG (A), and in 2MeTHF for PNM and PGA (B and C) are shown normalized to the absolute concentration of the solute, i.e. as extinction coefficient ϵ or its difference $\Delta\epsilon$ in an applied field of 1 MV/cm. Both features visible in the absorbance spectrum of PenG (A), which can be tentatively assigned to two distinct conformations only visible at 77 K, were modelled with similar Stark tuning rates. For PGA (C), the features marked with the asterisk (*) are assigned to the ester C=O mode. The extracted values of $|\Delta\vec{\mu}|f$ are shown below the corresponding Stark spectra. Note that IR bands are typically narrower in 2MeTHF than in glycerol/D₂O due to greater inhomogeneous broadening in the latter. Bottom panel: MD-Solvatochromic field-frequency calibrations via correlation of IR peak frequencies in various solvents with MD-calculated average electric fields. Squares and circles represent average solvent fields obtained from FC and POL MD simulations, black lines are linear fits to the data. For PenG (D) fitting results are provided in the caption of Figure 1B and only indicated as dashed lines due to the limited number of data points; for PNM (E), the linear fit yields $\bar{\nu} = (0.69 \pm 0.04) F + (1806.4 \pm 1.9)$ (POL, with $R^2 = 0.98$) and $\bar{\nu} = (0.78 \pm 0.02) F + (1796.8 \pm 0.6)$ (FC, with $R^2 = 0.99$); for PGA (F), the linear fit results in $\bar{\nu} = (0.39 \pm 0.04) F + (1805.2 \pm 1.7)$ (POL, with $R^2 = 0.95$) and $\bar{\nu} = (0.46 \pm 0.04) F + (1798.9 \pm 1.0)$ (FC, with $R^2 = 0.97$).

Based on the consistency of the experimental intrinsic Stark tuning rates from VSS for the three molecules, it is striking that the solvatochromic trend of PenG yields an aberrant, ca. 6-fold smaller solvatochromic slope relative to the structurally similar β -lactam C=O of PNM (compare Figure 3D and E), and this can now be traced as due to the PAA and/or carboxylate groups present in PenG. Furthermore, it is surprising that the field-frequency trends of PenG and PGA produce such significantly different solvatochromic slopes given that the only difference is the esterification of the carboxylate group, which should have minimal effect on the C=O's electric field response. Together with the identical experimental Stark tuning rates from VSS for all three β -lactams, the apparent discrepancies in the field-frequency calibrations suggest that the solvation environment from FC and POL MD simulations are unable to recapitulate the experimental vibrational solvatochromism. Both levels of MD simulations overestimate the local averaged solvent electric field at the β -lactam C=O in the presence of additional functional groups leading to solvatochromic slopes inconsistent with the experimental Stark tuning rates. We note that slight differences of the electrostatic parameters of the three solutes' β -lactam C=O do not explain these discrepancies (Table S6 and S7). As such, we analyze and discuss below the contributions of side-groups, i.e., the PAA and carboxylate, to this observed effect.

3.2 Origins of Inconsistencies in MD-derived Electric Fields

Conformational Sampling and Side-Chain Electric Fields in MD Simulations. One potential factor contributing to the discrepancy between the solvatochromic slopes and the Stark tuning rates is the conformational heterogeneity of more complex solutes like PenG and PGA, which was not considered explicitly in the above solvatochromic analysis. In particular, reorientation of the PAA amide group, which is proximate to the β -lactam C=O, would not only modulate *intermolecular* (solvent/ β -lactam C=O) interactions, via rotamer-dependent solvation, but also direct interactions with the polar side chain itself (side chain/ β -lactam C=O), which would together contribute to the entire flexible, dynamic electrostatic environment (note: PGA also contains the flexible allyl ester group, but its conformational effect is negligible since it is oriented away from the β -lactam due to the angled bicyclic structure; see structure in Figure 4A).

In order to test the role of PAA side group rotamers on both electrostatic contributions, we analyzed the MD simulations in terms of the dihedral angle around the C-N bond connecting the β -lactam and the PAA group, i.e. the C-C-N-C torsion highlighted in Figure 4A. Notably, we observe that across all solvents two main conformations are sampled by both PenG and PGA, denoted 'conformation 1' or 'conformation 2', with the amide N-H bond directed either towards the β -lactam carbonyl or the S atom, respectively (Figure 4A, B, C). These conformations are approximately equally prevalent using the FC force field, whereas AMOEBA09 parameters show a higher preference for conf. 2 (note: some POL MD simulations only sampled one conformation; in these cases, structures for the alternative conformations were obtained *via* restrained MD trajectories as described in SI

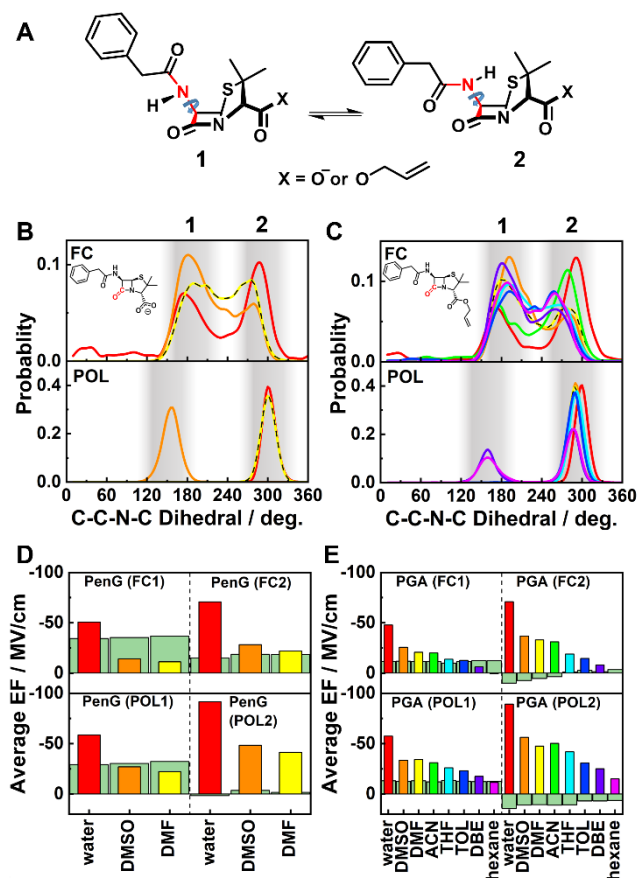


Figure 4. Average electric fields due to solvent and side-chain orientations from fixed-charge (FC) and polarizable (POL) MD associated with distinct conformations of PenG and PGA. (A): Reorientation around the C-C-N-C bond (dihedral is shown in red and with a blue arrow indicating the specific rotating bond) leads to different alignments of the strong dipole of the amide group with respect to β -lactam C=O. (B,C): Evaluating this distribution of this dihedral angles over the MD trajectories for PenG (A) and PGA (B) in all solvents (solvent color code similar to previous figures; the yellow lines for DMF are highlighted using a black dashed line) reveals two major conformations, labelled 1 and 2, where the amide is aligned approximately parallel with the β -lactam C=O (1) or is directed towards the sulfur (2), as shown in A. (D,E): Average solvent (rainbow colors) and electric fields due to side-chain orientation (green) exerted on the β -lactam C=O in FC and POL MD simulations when conformation 1 or 2 is adopted. Note that in PGA's conf. 1 the columns of the fields due to side-chain orientation are located behind the columns of the solvent fields.

section 4). Extracting the conformation-dependent average solvent electric fields we observe that consistently higher solvent fields, by a factor of ca. 1.6, are experienced by the β -lactam C=O in conf. 2 relative to conf. 1 (Figure 4D, E; rainbow-colored columns). The origin of this dihedral-dependent solvent electric field can, at least in part, be identified as a perturbation of the second solvation sphere around the β -lactam O atom, which shows a clear conformation-dependence, as demonstrated in the water radial distribution functions for PenG and PGA (see Figure S7 and S8).

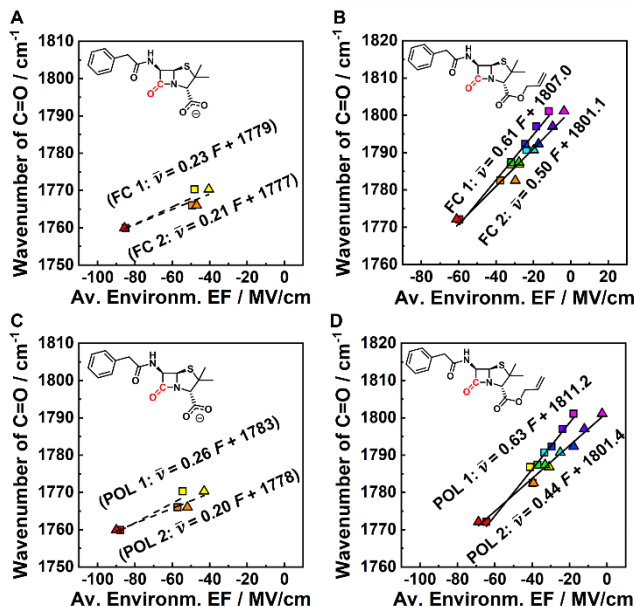


Figure 5. Conformation-dependent solvatochromic VSE analysis of PenG and PGA using average environmental electric fields obtained from FC (A,B) and POL simulations (C,D). Average environmental electric fields are obtained as a sum of average solvent and side-chain electric fields due to side groups of PenG and PGA. Based on the two conformations observed in MD trajectories separate VSE evaluations are performed labelled FC 1 and FC 2, along with POL 1 and POL 2, referring to the corresponding conformations depicted in Figure 3 (solvent colors follow the same pattern as in Figure 2; data points for conf. 1 and 2 are shown as squares and triangles, respectively; the corresponding picture without the contribution of the side-chain fields is shown in Figure S4). The linear regressions provide the following results in (A): $\bar{\nu} = (0.23 \pm 0.07) F + (1779 \pm 3)$ (PenG FC 1, with $R^2 = 0.92$) and $\bar{\nu} = (0.21 \pm 0.07) F + (1777 \pm 4)$ (PenG FC 2, with $R^2 = 0.87$); in (B): $\bar{\nu} = (0.61 \pm 0.04) F + (1807.0 \pm 0.9)$ (PGA and FC 1, with $R^2 = 0.98$) and $\bar{\nu} = (0.50 \pm 0.03) F + (1801.1 \pm 0.6)$ (PGA and FC 2, with $R^2 = 0.98$); in (C): $\bar{\nu} = (0.26 \pm 0.10) F + (1783 \pm 2)$ (PenG and POL 1, with $R^2 = 0.88$) and $\bar{\nu} = (0.20 \pm 0.05) F + (1778 \pm 1)$ (PenG and POL 2, with $R^2 = 0.94$); in (D): $\bar{\nu} = (0.63 \pm 0.05) F + (1811.2 \pm 0.8)$ (PGA and POL 1, with $R^2 = 0.96$) and $\bar{\nu} = (0.44 \pm 0.02) F + (1801.4 \pm 0.4)$ (PGA and POL 2, with $R^2 = 0.98$).

One would expect that these conformational changes of PenG’s and PGA’s side-chains influence the β -lactam C=O stretch predominantly via through-space electrostatics, as the β -lactam C=O and the PAA sidechain are separated exclusively by single bonds (see structure in Figure 4A). This assertion is supported by a comparison of DFT-based β -lactam C=O stretch vibrational frequencies and side-chain FC MD electric fields, both as a function of the C-C-N-C dihedral angle in vacuum (Figure S5). As such, we have determined the average side-chain electric fields from the FC and POL MD trajectories specific to conf. 1 and conf. 2 in all solvents (Figures 4D & E). For PGA’s conf. 1 (Figure 4E), we observe constant side-chain electric fields of ca. -12 MV/cm for FC and POL parameters, indicating that the net electrostatic effect of PGA’s side groups contributes a stabilizing electric field exerted onto the C=O dipole. This stabilization is due to the orientation of PAA’s amide group in conf. 1, in which the N-H dipole is aligned roughly

antiparallel with respect to the C=O’s dipole ($\sim 20^\circ$). In contrast, for conf. 2, the corresponding field due to the side-chain shows a pronounced solvent dependence with rather positive fields ranging from ca. +10 to -5 MV/cm or ca. +15 to +5 MV/cm for FC and POL MD, respectively, due to a slight solvent dependent variation of the dihedral angle. The electric fields of positive sign in this conformation create a destabilizing field at the β -lactam C=O dipole due to the unfavorable dipolar alignment with the PAA’s amide C=O ($\sim 30 - 50^\circ$). For PenG (Figure 4D), we observe that all fields from the side-chain are negative in FC MD, except for conf. 2 in POL MD, where their contribution is negligible. These side-chain fields show a constant difference of ca. -25 or -15 MV/cm in FC or POL MD simulations, respectively, in comparison to PGA. Given that the PAA group assumes similar rotamers in both molecules (Figure 4B & C), we can conclude that this difference originates from the presence of the negative charge at the carboxylate in PenG, which is neutralized in PGA by the esterification.

Using these side-chain contributions, we updated the previous solvatochromic calibrations shown in Figure 3 by combining the average solvent fields with the average side-chain fields to obtain the total averaged environmental electric fields for conf. 1 and conf. 2, where “environment” refers to electrostatic forces exerted upon the β -lactam C=O from both the solvent and the solute’s side groups (Figure 5). Using both FC and POL MD, the resulting slopes of the field-frequency calibration increase for both PenG and PGA in conf. 1 and 2, although to a larger extent for conf. 1. Based on the key observation that the intrinsic Stark tuning rate of the β -lactam C=O determined using VSS is identical for the three molecules (Figure 3A – C), we can compare these values to PNM’s slopes of 0.79 and 0.68 $\text{cm}^{-1}/(\text{MV}/\text{cm})$ for FC and POL MD, respectively. For PenG, despite the addition of side-chain electrostatic effects and conformational deconvolution, a substantial discrepancy of a factor of 3 – 4 relative to PNM’s solvatochromic slope remains, indicating further complications arising from the negatively charged carboxylate. In contrast, for neutral PGA, which showed a discrepancy of the solvatochromic slopes by a factor of ca. 1.8 in comparison to PNM (Figure 3), selective consideration of conf. 1 and addition of side-chain electrostatic contributions in the analysis reduced this discrepancy to factors of 1.3 and 1.1 for FC MD and POL MD, whereas PGA’s conf. 2 results in larger discrepancies of 1.6. In light of these conformational differences, we note that unrestricted FC and POL MD simulations led to markedly different conformational samplings, i.e., either a mixture of both conformations or predominantly conf. 2, respectively. This demonstrates the sensitivity of MD-predicted electric fields to complications with *ab initio* quantum chemistry-derived MD parameters of solutes or potential incompatibilities with solvent parameters, leading to these discrepancies.

Overestimated MD Electric Fields in Water in the Presence of Charged Functional Groups. In order to further unravel the origin(s) of the breakdown in the field-frequency calibration of PenG relative to PNM, we performed a comparison with the *expected* field-frequency correlation of PenG based on the Stark tuning rate from VSS and the

experimental gas phase frequency.⁶⁴ This comparison is based on the direct observation using VSS that the β -lactam C=O Stark tuning rates of PNM, PGA and PenG are the same (Figure 3A-C) and thus should show similar slopes in solvatochromic calibrations. We can readily determine an expected field-frequency trend for PenG using eq. 1 by (a) equating PenG’s field-frequency slope to the one obtained in the solvatochromic calibration of PNM (Figure 3E) and (b) using the experimental gas phase β -lactam C=O frequency of the PenG anion of 1778 cm^{-1} as the zero-solvent field frequency; the latter was determined in IR multiphoton dissociation measurements.⁶⁴ Furthermore, to account for side-chain electrostatic contributions, we extracted the average side-chain electric fields for the PenG anion from *in vacuo* MD simulations for both conformations, in analogy to Figure 4. This approach allows us to be consistent within the concept of the average environmental electric fields as used in the previous section and Figure 5.

Figure 6 compares the FC- and POL-based data points with the expected trend so we can visualize which data points are associated with the largest deviation from the predicted environmental electric fields. Beginning with aprotic organic solvents DMF and DMSO in Figure 6A and B, we see that the data points of PenG fall on the expected trends of conf. 1 and conf. 2 for FC and POL MD results. Here, PenG in conf. 1 in DMF (POL) shows the highest deviation from the expected line by ca. -8 MV/cm. However, comparing this deviation to the scatter observed for previous solvatochromic calibrations of other carbonyl probes,³¹ it still fits within the framework of the solvatochromic analysis. Therefore, we can conclude that PenG’s experimental C=O frequencies in the gas phase, DMF, and DMSO are consistent with the Stark tuning rate of PNM, as expected given the shared penam-core and the VSS results. In contrast, the data points in water in both FC and POL MD are consistently displaced from the expected trend by -25 or -30 MV/cm for conf. 1 and 2, respectively. This observation indicates that both water models used in FC and POL simulations (tip3p and amoeba03 waters, respectively) overestimate the local environmental field exerted on the β -lactam C=O, pointing to inaccuracies of the solvation structure and/or dynamics in the proximity of the charged carboxylate of PenG.

The overestimation of electrostatic interactions of the charged PenG with water can be traced to issues experienced in simulations of the thermodynamics of ion hydration, which are found to be highly sensitive to force fields and parameterizations (see ref. 65 and citations therein). Interestingly, these overestimated fields in water are observed to the same extent for POL MD-based force fields, even though one might expect that the inclusion of polarization effects is an obvious missing factor required to describe water molecules in the proximity of charged groups. Since the accurate modelling of H-bond interactions and dynamics is undoubtedly a significant challenge in MD methods,^{66–71} we sought to test more elaborate and updated water models that are available for FC MD and POL MD simulations to assess whether further refinement and improvements of force fields may be needed in order to accurately recapitulate the water dynamics and solvation environment.

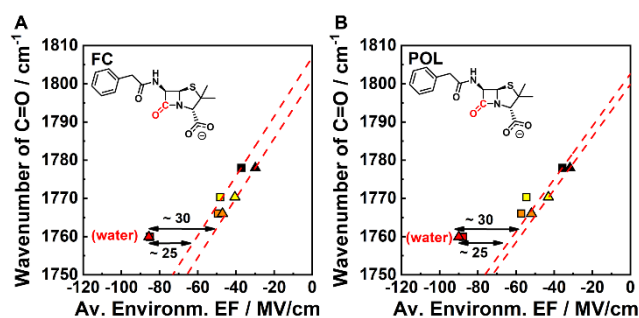


Figure 6. Comparison of the expected solvatochromic field-frequency correlation of PenG with the data obtained using (A) FC and (B) POL-based average environmental electric fields, i.e. the sum of solvent and side-chain electric field contributions. Data points correlating solvatochromic frequencies of PenG and MD-based environmental electric fields are shown as colored squares and triangles for conf. 1 and 2, respectively. The experimental gas phase frequency of PenG is shown as black squares and triangles. The expected correlations for PenG’s conf. 1 (top) and conf. 2 (bottom) are shown as red dashed lines. Discrepancies between MD and expected fields are indicated using the horizontal double-headed arrows and the associated ΔF values.

Electric Fields using Different Water Models. Water models of increasing complexity have been introduced into FC force fields⁶⁵ to more accurately model the properties of water molecules relative to the most commonly used 3-point tip3p model,^{72,73} which has been employed in previous solvatochromic calibrations and vibrational frequency maps.^{18,74} More elaborate models include the 4-point tip4p,⁷⁵ the 5-point tip5p⁷⁶ or 7-point tip7p⁷⁷ models that attempt to describe a more realistic charge distribution and geometry of the water molecule. Within AMOEBA, the flexible 3-point amoeba03 water model⁷⁸ has been updated to improve symmetric and asymmetric stretching frequencies (here referred to as amoeba03*), was re-parameterized using the Force-Balance algorithm as amoeba14,⁷⁹ or was simplified to improve the cost of simulations with (inexpensive) iamoeba.⁸⁰ We sought to evaluate whether alternative FC and AMOEBA water models or their later re-parameterized versions (i.e. tip4/5pew,^{81,82} tip4p2005,⁸³ tip5p2018,⁸⁴ tip3/4pfb⁸⁵) can alleviate the discrepancy between the VSE analyses of PenG, PNM and PGA by testing their average electric fields in individual MD trajectories.

Comparing the absolute magnitudes of electric fields exerted on the β -lactam C=O using the FC water models one finds a general trend in solvent electric fields of $|F(\text{tip5p})| < |F(\text{tip7p})| < |F(\text{tip3p})| < |F(\text{tip4p})|$ spanning up to ca. 30 MV/cm across each of the β -lactams and their corresponding conformations (Figure 7A, Figure S6). This trend is in line with the partial charge on the H atoms of the water models (see Table S5) as well as the solvation structure and dynamics in MD simulations (i.e., H-bond distance and fluctuations, see Figure S7). Among the tested AMOEBA water models (Figure 7B), solvent fields span an even larger range projected on the C=O of up to 60 MV/cm with $|F(\text{iamoeba})| < |F(\text{amoeba03})| \approx |F(\text{amoeba03*})| < |F(\text{amoeba14})|$. In both FC and POL, side-chain electric fields do not change the observed trends of the solvent fields as they remain constant across all tipNp and all AMOEBA water models, with the

exception of iamoeba, which shows slightly more positive side-chain fields; likely owing to the lack of stabilizing ‘mutual’ (induced dipole – induced dipole) polarization.

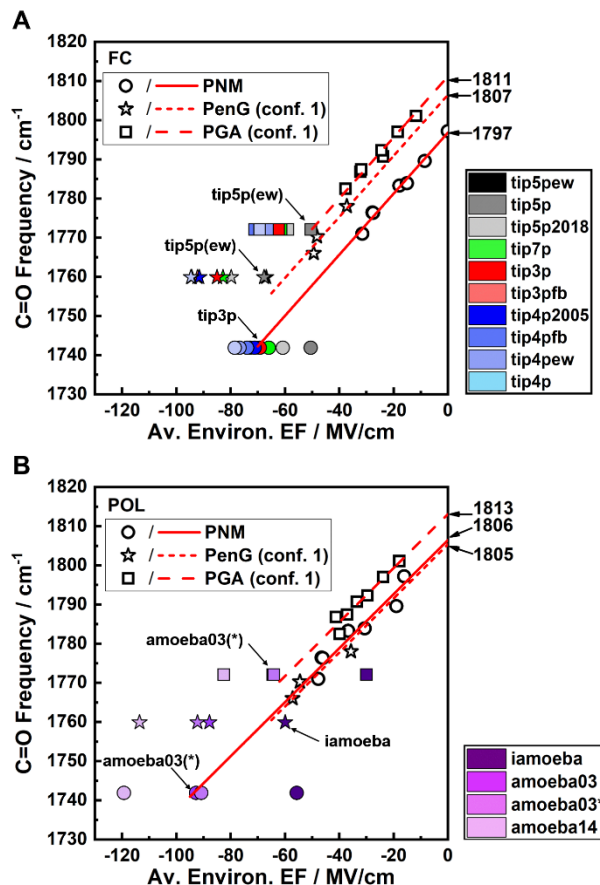


Figure 7. Effect of different water models on the solvatochromic calibrations of PNM, PenG, and PGA using average environmental electric fields from FC (A) and POL MD (B). Data points of organic solvents and gas phase are displayed as hollow circles, stars, or squares for PNM, PenG and PGA, respectively. The same but colored symbols are used for data points with different water models (see legends for color code). Full, short dashed and long dashed lines (for PNM, PenG and PGA, respectively) represent fits with the solvatochromic slope of PNM of 0.78 or 0.69 cm⁻¹/(MV/cm) obtained from FC (A) or POL (B) in Figure 3. Best matching data points of water models are highlighted with model name. Zero-environmental field frequencies are display on the right corner of each figure.

In Figure 7 we added these water models into the solvatochromic calibrations of PNM, PenG and PGA to evaluate which of the models is most consistent with the intrinsic Stark tuning rate of the β -lactam C=O stretch; based on the improved agreement of solvatochromic slopes (see above), we only show conf. 1 for PenG and PGA. We performed a fit to the data points in organic solvent (and gas phase for PenG) using eq. 1 and PNM’s solvatochromic slopes of 0.78 and 0.69 cm⁻¹/(MV/cm) in FC and POL MD, respectively. By extrapolating the fits towards the vibrational frequency in water, we observe that among the considerably varying water models tip3p (and tip4p2005), and amoeba03 (and amoeba03*) provide the best agreement with the expected linear trends of PNM, the simplest of the three molecules.

This is in line with previous work on small rigid solutes using the VSE^{31–33} and can be understood as both 3-point models lie in the middle of their respective trends comparing different water models. As such, they can be regarded as describing an averaged picture of the different geometrical and electrostatic aspects of the set of tested water models. In contrast, for PenG and PGA (FC), tip5p and tip5pew provide a better agreement, alleviating the discrepancies discussed above. This result suggests that the geometry and charge distribution of tip5p(ew) more appropriately describes the β -lactam C=O’s local solvation environment in the presence of the multiple functional groups of PGA as well as the charged carboxylate of PenG. In POL MD, amoeba03(*) provides a good match for PGA, pointing towards the importance of polarization for interactions with multiple functional groups (when comparing to the corresponding 3-point tip3p in FC MD). In turn, for PenG (POL MD) iamoeba appears to alleviate the discrepancy shown in Figure 6; however, we stress that this change of electric fields hardly reflects improved solvation electrostatics as it originates from the exclusion of mutual polarization by induced dipoles in the “inexpensive” iamoeba.

The large variations in the water electric fields, despite careful parameterization of each water force field against bulk experimental observables (e.g., hydration energies, densities, dielectric constants etc.), suggests that caution must be used when applying these models to quantify local electric fields around polar and charged functional groups. It is difficult to pinpoint whether improvements of geometric or electrostatic properties are necessary for a more reliable description of hydration electrostatics around the polar and charged functional groups of PenG (see SI section 9 for brief comparison of MD and DFT-based electric field estimations). As such, it will be interesting to test in future studies whether the current improvements of polarizable MD models such as AMOEBA+,⁸⁶ AMOEBA+cf,⁸⁷ aniso-AMOEBA,⁸⁸ or the MB-pol many body potential⁸⁹ will provide more accurate descriptions of solvation electrostatics for solutes of diverse structures and charge distributions, or if *ab initio* MD methods must be employed. In addition to such advances, we highlight the necessity of benchmarking force fields against molecular observables, such as the VSE, to account for the highly heterogeneous behavior of water when interacting with a charged functional group, which is often undervalued when focusing on bulk averaged properties. In this context, the VSE is particularly useful as a direct and, most importantly, molecular method for further experiment-guided refinements (e.g., allowing us to probe electrostatics experimentally for isolated ions, and not only ion pairs) to test force fields using the electric field associated with non-covalent interactions.

3.3 Conclusions and Implications for MD-based VSE Solvatochromic Calibrations

Experimental VSE-based determinations of electric fields in complex environments, such as enzyme active sites, rely on accurate MD simulations in solvents as calibration. This and previous work has demonstrated that MD force fields accurately describe the solvatochromic behavior of small and conformationally rigid molecules containing C=O

and C≡N groups as vibrational probes within the framework of the linear VSE.^{23,31–33} However, the conformationally flexible and charged PenG highlights discrepancies for structurally more complex (bio)molecules; that is different PenG derivatives yield inconsistent solvatochromic slopes. We showed that explicit consideration of the electric fields due to the PAA side-chain consistently reduces the discrepancies due conformational flexibility. However, the remaining discrepancy due to the charged carboxylate in water was neither improved using polarizable AMOEBA force fields or more advanced water models. As such, this work demonstrates the particular value of the VSE as a direct method for future experiment-guided refinements of MD force fields providing quantitative insight into the local electrostatics of non-covalent interactions in various condensed phase environments ranging from solvents to enzyme active sites. It may be that QM methods are required to fully capture these effects, though with a loss of physical intuition.

Based on these results we highlight the importance of utilizing MD-based solvent electric fields of small, conformationally rigid molecules, such as PNM, to verify the solvatochromic calibrations of more complex derivatives like PenG or PGA. Moreover, we argue that the bicyclic penam-core of PNM can be used as an appropriate proxy for quantifying the influence of environmental electrostatics on the β -lactam C=O bond in general. In other words: the response of PenG's and PGA's β -lactam C=O to the cumulative electrostatic effect due to solvent and orientation-dependent side-chain conformations is well captured by PNM's solvent response. Such a strategy is supported by the relatively good agreement, within the limitations and accuracy of the presented approach, between the zero-field frequencies of PNM, PGA and PenG after correction for side-chain electrostatics (Figure 7; compare to Figure 3). Here, a better correspondence is observed for POL relative to FC MD, which is in line with previous results showing that inclusion of polarization in MD simulations improves the prediction of gas phase frequencies extrapolated from VSE solvatochromism.³³ This reductionist strategy enables explicit quantification of the effects of both non-covalent interactions from the solvent and through space effects of the flexible side-chains self-consistently for a given class of solutes sharing a shared probe-containing core. These results establish methods for extending the VSE to complex (bio)molecules to quantify environment-dependent electrostatic forces, e.g., upon binding of PenG to its targets the penicillin-binding proteins or β -lactamases.⁹⁰

Notes

Structure and parameter files for GROMACS and TINKER of PenG, PGA, PNM and ACP (GAFF and AMOEBA) as well as dibutyl ether (AMOEBA) are deposited under <https://github.com/jacekkozuch/MD-parameters/releases/tag/v0.1-preprint>.

Acknowledgment

We thank Jay Ponder (Washington University) for support with Tinker and AMOEBA simulations, Robert Lesniak from the Du Bois Lab (Stanford University) for assistance with chemical synthesis, and Yasmin Binti Shamsudin, Jared Weaver and Jacob Kirsh

for helpful discussions. J.K. acknowledges the Deutsche Forschungsgemeinschaft for a research fellowship (KO5464/1). This work was supported in part by NIH Grant GM118044.

Abbreviations

VSE – vibrational Stark effect, VSS – vibrational Stark spectroscopy, FC – fixed charge; POL – polarizable; solutes: ACP – acetophenone, PenG – penicillin G, PGA – penicillin G allyl ester, PNM – penam; PAA – phenylacetamide; solvents for solvatochromism: DMSO – dimethyl sulfoxide, DMF – dimethylformamide, ACN – acetonitrile, THF – tetrahydrofuran, TOL – toluene, DBE – dibutyl ether, HEX – hexane(s); VSS solvent: 2MeTHF – 2-methyl tetrahydrofuran.

References

- (1) Zhou, H.-X. X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. *Chem. Rev.* **2018**, *118* (4), 1691–1741.
- (2) Warshel, A.; Sharma, P. K.; Kato, M.; Xiang, Y.; Liu, H.; Olsson, M. H. M. Electrostatic Basis for Enzyme Catalysis. *Chem. Rev.* **2006**, *106* (8), 3210–3235.
- (3) Roux, B.; Berneche, S.; Im, W. Ion Channels, Permeation, and Electrostatics: Insight into the Function of KcsA. *Biochemistry* **2000**, *39* (44), 13295–13306.
- (4) Reichardt, C. Solvatochromic Dyes as Solvent Polarity Indicators. *Chem. Rev.* **1994**, *94* (8), 2319–2358.
- (5) Catalán, J.; Del Valle, J. C.; Díaz, C.; Palomar, J.; De Paz, J. L. G.; Kasha, M. Solvatochromism of Fluorophores with an Intramolecular Hydrogen Bond and Their Use as Probes in Biomolecular Cavity Sites. *Int. J. Quantum Chem.* **1999**, *72* (4), 421–438.
- (6) Clarke, R. J. The Dipole Potential of Phospholipid Membranes and Methods for Its Detection. *Adv. Colloid Interfac.* **2001**, *89–90*, 263–281.
- (7) Loving, G. S.; Sainlos, M.; Imperiali, B. Monitoring Protein Interactions and Dynamics with Solvatochromic Fluorophores. *Trends Biotechnol.* **2010**, *28* (2), 73–83.
- (8) Dong, F.; Olsen, B.; Baker, N. A. Computational Methods for Biomolecular Electrostatics. *Methods Cell Biol.* **2008**, *84* (07), 843–870.
- (9) Karplus, M.; Kuriyan, J. Molecular Dynamics and Protein Function. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (19), 6679–6685.
- (10) Adcock, S. A.; McCammon, J. A. Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. *Chem. Rev.* **2006**, *106* (5), 1589–1615.
- (11) Welborn, V. V.; Ruiz Pestana, L.; Head-Gordon, T. Computational Optimization of Electric Fields for Better Catalysis Design. *Nat. Catal.* **2018**, *1* (9), 649–655.
- (12) Welborn, V. V.; Head-Gordon, T. Fluctuations of Electric Fields in the Active Site of the Enzyme Ketosteroid Isomerase. *J. Am. Chem. Soc.* **2019**, *141* (32), 12487–12492.
- (13) Caleman, C.; Van Maaren, P. J.; Hong, M.; Hub, J. S.; Costa, L. T.; Van Der Spoel, D. Force Field Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities, Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and Dielectric Constant. *J. Chem. Theory Comput.* **2012**, *8* (1), 61–74.
- (14) Zhang, H.; Jiang, Y.; Cui, Z.; Yin, C. Force Field Benchmark of Amino Acids. 2. Partition Coefficients between Water and Organic Solvents. *J. Chem. Inf. Model.* **2018**, *58* (8), 1669–1681.
- (15) Horton, J. T.; Allen, A. E. A.; Dodda, L. S.; Cole, D. J. QUBESKit: Automating the Derivation of Force Field Parameters from Quantum Mechanics. *J. Chem. Inf. Model.* **2019**, *59* (4), 1366–1381.
- (16) Sherrill, C. D.; Merz, K. M. Quantum Mechanical Methods for Quantifying and Analyzing Non-Covalent Interactions and for Force-Field Development. In *Many-Body Effects and Electrostatics in Biomolecules*; Cui, Q., Meuwly, M., Ren, P., Eds.; Pan Stanford Publishing: New York, 2016; pp 65–120.
- (17) Guillot, B. A Reappraisal of What We Have Learnt during Three Decades of Computer Simulations on Water. *J. Mol. Evol.* **2002**, *101*, 219–260.
- (18) Fried, S. D.; Boxer, S. G. Measuring Electric Fields and Noncovalent Interactions Using the Vibrational Stark Effect. *Acc. Chem. Res.*

- 2015**, 48, 998–1006.
- (19) Fried, S. D.; Boxer, S. G. Electric Fields and Enzyme Catalysis. *Annu. Rev. Biochem.* **2017**, 86 (1), 387–415.
 - (20) Völler, J.-S.; Biava, H.; Hildebrandt, P.; Budisa, N. An Expanded Genetic Code for Probing the Role of Electrostatics in Enzyme Catalysis by Vibrational Stark Spectroscopy. *Biochim. Biophys. Acta* **2017**, 1861 (11), 3053–3059.
 - (21) Slocum, J. D.; Webb, L. J. Measuring Electric Fields in Biological Matter Using the Vibrational Stark Effect of Nitrile Probes. *Annu. Rev. Phys. Chem.* **2018**, 69 (1), 253–271.
 - (22) Schneider, S. H.; Kratochvil, H. T.; Zanni, M. T.; Boxer, S. G. Solvent-Independent Anharmonicity for Carbonyl Oscillators. *J. Phys. Chem. B* **2017**, 121 (10), 2331–2338.
 - (23) Deb, P.; Haldar, T.; Kashid, S. M.; Banerjee, S.; Chakrabarty, S.; Bagchi, S. Correlating Nitrile IR Frequencies to Local Electrostatics Quantifies Noncovalent Interactions of Peptides and Proteins. *J. Phys. Chem. B* **2016**, 120 (17), 4034–4046.
 - (24) Staffa, J. K. J. K.; Lorenz, L.; Stolarski, M.; Murgida, D. H. D. H.; Zebger, I.; Utesch, T.; Kozuch, J.; Hildebrandt, P. Determination of the Local Electric Field at Au/SAM Interfaces Using the Vibrational Stark Effect. *J. Phys. Chem. C* **2017**, 121 (40), 22274–22285.
 - (25) Patrow, J. G.; Sorenson, S. A.; Dawlaty, J. M. Direct Spectroscopic Measurement of Interfacial Electric Fields near an Electrode under Polarizing or Current-Carrying Conditions. *J. Phys. Chem. C* **2017**, 121 (21), 11585–11592.
 - (26) Hu, W.; Webb, L. J. Direct Measurement of the Membrane Dipole Field in Bicyclic Using Vibrational Stark Effect Spectroscopy. *J. Phys. Chem. Lett.* **2011**, 2 (15), 1925–1930.
 - (27) Bagchi, S.; Fried, S. D.; Boxer, S. G. A Solvatochromic Model Calibrates Nitriles' Vibrational Frequencies to Electrostatic Fields. *J. Am. Chem. Soc.* **2012**, 134 (25), 10373–10376.
 - (28) Biava, H.; Schreiber, T.; Katz, S.; Völler, J.-S. S.; Stolarski, M.; Schulz, C.; Michael, N.; Budisa, N.; Kozuch, J.; Utesch, T.; et al. Long-Range Modulations of Electric Fields in Proteins. *J. Phys. Chem. B* **2018**, 122 (35), 8330–8342.
 - (29) Fried, S. D.; Bagchi, S.; Boxer, S. G. Extreme Electric Fields Power Catalysis in the Active Site of Ketosteroid Isomerase. *Science*. **2014**, 346 (6216), 1510–1514.
 - (30) Wu, Y.; Boxer, S. G. A Critical Test of the Electrostatic Contribution to Catalysis with Noncanonical Amino Acids in Ketosteroid Isomerase. *J. Am. Chem. Soc.* **2016**, 138, 11890–11895.
 - (31) Schneider, S. H.; Boxer, S. G. Vibrational Stark Effects for Diverse Carbonyl Probes Applied to the Re-Interpretation of IR and Raman Data in Terms of Electric Fields at Enzyme Active Sites. *J. Phys. Chem. B* **2016**, 120 (36), 9672–9684.
 - (32) Fried, S. D.; Bagchi, S.; Boxer, S. G. Measuring Electrostatic Fields in Both Hydrogen Bonding and Non-Hydrogen Bonding Environments Using Carbonyl Vibrational Probes. *J. Am. Chem. Soc.* **2013**, 135 (30), 11181–11192.
 - (33) Fried, S. D.; Wang, L.-P. P.; Boxer, S. G.; Ren, P.; Pande, V. S. Calculations of the Electric Fields in Liquid Solutions. *J. Phys. Chem. B* **2013**, 117 (50), 16236–16248.
 - (34) Błasiak, B.; Ritchie, A. W.; Webb, L. J.; Cho, M. Vibrational Solvatochromism of Nitrile Infrared Probes: Beyond the Vibrational Stark Dipole Approach. *Phys. Chem. Chem. Phys.* **2016**, 18 (27), 18094–18111.
 - (35) Fried, S. D.; Boxer, S. G. Thermodynamic Framework for Identifying Free Energy Inventories of Enzyme Catalytic Cycles. *Proc. Natl. Acad. Sci. USA* **2013**, 110 (30), 12271–12276.
 - (36) Andrews, S. S.; Boxer, S. G. Vibrational Stark Effects of Nitriles. I. Methods and Experimental Results. *J. Phys. Chem. A* **2000**, 104 (51), 11853–11863.
 - (37) List, N. H.; Beerepoot, M. T. P.; Olsen, J. M. H.; Gao, B.; Ruud, K.; Jensen, H. J. A.; Kongsted, J. Molecular Quantum Mechanical Gradients within the Polarizable Embedding Approach—Application to the Internal Vibrational Stark Shift of Acetophenone. *J. Chem. Phys.* **2015**, 142 (3), 034119.
 - (38) Choi, J. H.; Cho, M. Vibrational Solvatochromism and Electrochromism of Infrared Probe Molecules Containing C=O, C=N, C=O, or C-F Vibrational Chromophore. *J. Chem. Phys.* **2011**, 134 (15), 154513.
 - (39) Essack, S. Y. The Development of B-Lactam Antibiotics in Response to the Evolution of β -Lactamases. *Pharm. Res.* **2001**, 18 (10), 1391–1399.
 - (40) Bush, K.; Maclellan, M. J. New β -Lactam Antibiotics and β -Lactamase Inhibitors. *Expert Opin. Ther. Pat.* **2010**, 20 (10), 1277–1293.
 - (41) Kamath, A.; Ojima, I. Advances in the Chemistry of B-Lactam and Its Medicinal Applications. *Tetrahedron* **2012**, 68 (52), 10640–10664.
 - (42) Guzmán-Chávez, F.; Zwahlen, R. D.; Bovenberg, R. A. L.; Driessen, A. J. M. Engineering of the Filamentous Fungus *Penicillium Chrysogenum* Cell Factory for Natural Products. *Front. Microbiol.* **2018**, 9, 2768.
 - (43) Drawz, S. M.; Bonomo, R. A. Three Decades of B-Lactamase Inhibitors. *Clin. Microbiol. Rev.* **2010**, 23 (1), 160–201.
 - (44) Wilke, M. S.; Lovering, A. L.; Strynadka, N. C. J. β -Lactam Antibiotic Resistance: A Current Structural Perspective. *Curr. Opin. Microbiol.* **2005**, 8 (5), 525–533.
 - (45) Pratt, R. F. β -Lactamases: Why and How. *J. Med. Chem.* **2016**, 59 (18), 8207–8220.
 - (46) Jeffrey, P. D.; McCombie, S. W. Homogeneous, Palladium(0)-Catalyzed Exchange Deprotection of Allylic Esters, Carbonates, and Carbamates. *J. Org. Chem.* **1982**, 47, 587–590.
 - (47) Chiba, T.; Sakaki, J.; Takahashi, T.; Aoki, K.; Kamiyama, A.; Kaneko, C.; Sato, M. Studies on Amino Acid Derivatives. Part 7. General Method for the Synthesis of Penam and Cepham and Their Substituted Derivatives. *J. Chem. Soc. Perkin Trans. 1* **1987**, No. 0, 1845.
 - (48) Cambie, R.; Clark, G.; Jones, T.; Rutledge, P.; Strange, G.; Woodgate, P. Vic-Iodo Thiocyanates and Iodo Isothiocyanates. IX. A Synthesis of Penam and Other Polycyclic B-Lactams. *Aust. J. Chem.* **1985**, 38 (5), 745.
 - (49) Andrews, S. S.; Boxer, S. G. A Liquid Nitrogen Immersion Cryostat for Optical Measurements. *Rev. Sci. Instrum.* **2000**, 71 (9), 3567–3569.
 - (50) Bubblitz, G. U.; Boxer, S. G. Stark Spectroscopy: Applications in Chemistry, Biology, and Materials Science. *Annu. Rev. Phys. Chem.* **1997**, 48, 213–242.
 - (51) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and Testing of a General Amber Force Field. *J. Comp. Chem.* **2004**, 25 (9), 1157–1174.
 - (52) Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A. Automatic Atom Type and Bond Type Perception in Molecular Mechanical Calculations. *J. Mol. Graph. Model.* **2006**, 25, 247–260.
 - (53) James, M.; Murtola, T.; Schulz, R.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, 1–2, 19–25.
 - (54) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved Side-Chain Torsion Potentials for the Amber Ff99SB Protein Force Field. *Proteins* **2010**, 78 (8), 1950–1958.
 - (55) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, Efficient Generation of High-Quality Atomic Charges. AM1-BCC Model: II. Parameterization and Validation. *J. Comp. Chem.* **2002**, 23 (16), 1623–1641.
 - (56) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. Gaussian 09. Gaussian, Inc.; Wallingford CT 2013.
 - (57) van der Spoel, D.; van Maaren, P. J.; Coleman, C. GROMACS Molecule & Liquid Database. *Bioinformatics* **2012**, 28 (5), 752–753.
 - (58) Ponder, J. W.; Wu, C.; Pande, V. S.; Chodera, J. D.; Schnieders, M. J.; Haque, I.; Mobley, D. L.; Lambrecht, D. S.; Distasio, R. A.; Head-gordon, M.; et al. Current Status of the AMOEBA Polarizable Force Field. *J. Phys. Chem. B* **2010**, 114, 2549–2564.
 - (59) Schnieders, M. J.; Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagarde, L.; Piquemal, J.; Ren, P.; Ponder, J. W. Tinker 8: Software Tools for Molecular Design. *J. Chem. Theory Comput.* **2018**, 14, 5273–5289.
 - (60) Wu, J. C.; Chattree, G.; Ren, P. Automation of AMOEBA Polarizable Force Field Parameterization for Small Molecules. *Theor. Chem. Acc.* **2012**, 131, 1138.
 - (61) Bradshaw, R. T.; Essex, J. W. Evaluating Parameterization Protocols for Hydration Free Energy Calculations with the AMOEBA Polarizable Force Field. *J. Chem. Theory Comput.* **2016**, 12, 3871–3883.
 - (62) Boxer, S. G. Stark Realities. *J. Phys. Chem. B* **2009**, 113, 2972–2983.
 - (63) Lin, C. Y.; Romei, M. G.; Oltrogge, L. M.; Mathews, I. I.; Boxer, S.

- G. Unified Model for Photophysical and Electro-Optical Properties of Green Fluorescent Proteins. *J. Am. Chem. Soc.* **2019**, *141* (38), 15250–15265.
- (64) Dunbar, R. C.; Oomens, J.; Orlova, G.; Bohme, D. K. IRMPD Spectroscopic Investigation of Gas-Phase Complexes of Deprotonated Penicillin G with Ba²⁺, Zn²⁺ and Cd²⁺. *Int. J. Mass. Spectr.* **2011**, *308*, 330–337.
- (65) Ren, P.; Chun, J.; Thomas, D. G.; Schnieders, M. J.; Marucho, M.; Zhang, J.; Baker, N. A. Biomolecular Electrostatics and Solvation: A Computational Perspective. *Quart. Rev. Biophys.* **2012**, *45* (4), 427–491.
- (66) Auer, B.; Kumar, R.; Schmidt, J. R.; Skinner, J. L. Hydrogen Bonding and Raman, IR, and 2D-IR Spectroscopy of Dilute HOD in Liquid D₂O. *Proc. Natl. Acad. Sci. USA* **2007**, *104* (36), 14215–14220.
- (67) Nicodemus, R. A.; Corcelli, S. A.; Skinner, J. L.; Tokmakoff, A. Collective Hydrogen Bond Reorganization in Water Studied with Temperature-Dependent Ultrafast Infrared Spectroscopy. *J. Phys. Chem. B* **2011**, *115*, 5604–5616.
- (68) Liu, J.; He, X.; Zhang, J. Z.; Qi, L. Hydrogen-Bond Structure Dynamics in Bulk Water: Insights from Ab Initio Simulations with Coupled Cluster Theory. *Chem. Sci.* **2018**, *9* (8), 2015–2386.
- (69) Asbury, J. B.; Steinel, T.; Kwak, K.; Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L.; Fayer, M. D. Dynamics of Water Probed with Vibrational Echo Correlation Spectroscopy. *J. Chem. Phys.* **2004**, *121* (24), 12431–12446.
- (70) Li, X.; Walker, B.; Michaelides, A. Quantum Nature of the Hydrogen Bond. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6369–6373.
- (71) Ceriotti, M.; Fang, W.; Kusalik, P. G.; Mckenzie, R. H.; Michaelides, A.; Morales, M. A.; Markland, T. E. Nuclear Quantum Effects in Water and Aqueous Systems: Experiment, Theory, and Current Challenges. *Chem. Rev.* **2016**, *116*, 7529–7550.
- (72) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926–935.
- (73) Dick, T. J.; Madura, J. D. A Review of the TIP4P, TIP4P-Ew, TIP5P, and TIP5P-E Water Models. *Annu. Rep. Comp. Chem.* **2005**, *1*, 59–74.
- (74) Schmidt, J. R.; Corcelli, S. A.; Skinner, J. L. Ultrafast Vibrational Spectroscopy of Water and Aqueous N-Methylacetamide: Comparison of Different Electronic Structure/Molecular Dynamics Approaches. *J. Chem. Phys.* **2004**, *121* (18), 8887–8896.
- (75) Jorgensen, W. L.; Madura, J. D. Temperature and Size Dependence for Monte Carlo Simulations of TIP4P Water. *Mol. Phys.* **1985**, *56*, 1381–1392.
- (76) Mahoney, M. W.; Jorgensen, W. L. A Five-Site Model for Liquid Water and the Reproduction of the Density Anomaly by Rigid, Nonpolarizable Potential Functions. *J. Chem. Phys.* **2000**, *112*, 8910.
- (77) Zhao, C. L.; Zhao, D. X.; Bei, C. C.; Meng, X. N.; Li, S.; Yang, Z. Z. Seven-Site Effective Pair Potential for Simulating Liquid Water. *J. Phys. Chem. B* **2019**, *123* (21), 4594–4603.
- (78) Ren, P.; Ponder, J. W. Polarizable Atomic Multipole Water Model for Molecular Mechanics Simulation. *J. Phys. Chem. B* **2003**, *107* (24), 5933–5947.
- (79) Laury, M. L.; Wang, L. P.; Pande, V. S.; Head-Gordon, T.; Ponder, J. W. Revised Parameters for the AMOEBA Polarizable Atomic Multipole Water Model. *J. Phys. Chem. B* **2015**, *119* (29), 9423–9437.
- (80) Wang, L. P.; Head-Gordon, T.; Ponder, J. W.; Ren, P.; Chodera, J. D.; Eastman, P. K.; Martinez, T. J.; Pande, V. S. Systematic Improvement of a Classical Molecular Model of Water. *J. Phys. Chem. B* **2013**, *117* (34), 9956–9972.
- (81) Horn, H. W.; Swope, W. C.; Pitera, J. W.; Madura, J. D.; Dick, T. J.; Hura, G. L.; Head-Gordon, T. Development of an Improved Four-Site Water Model for Biomolecular Simulations: TIP4P-Ew. *J. Chem. Phys.* **2004**, *120* (20), 9665–9678.
- (82) Rick, S. W. A Reoptimization of the Five-Site Water Potential (TIP5P) for Use with Ewald Sums. *J. Chem. Phys.* **2004**, *120* (13), 6085–6093.
- (83) Abascal, J. L. F.; Vega, C. A General Purpose Model for the Condensed Phases of Water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123* (23).
- (84) Khalak, Y.; Baumeier, B.; Karttunen, M. Improved General-Purpose Five-Point Model for Water: TIP5P/2018. *J. Chem. Phys.* **2018**, *149* (22).
- (85) Wang, L.-P.; Martinez, T. J.; Pande, V. S. Building Force Fields: An Automatic, Systematic, and Reproducible Approach. *J. Phys. Chem. Lett.* **2014**, *5*, 1885–1891.
- (86) Liu, C.; Piquemal, J. P.; Ren, P. AMOEBA+ Classical Potential for Modeling Molecular Interactions. *J. Chem. Theory Comput.* **2019**, *15* (7), 4122–4139.
- (87) Liu, C.; Piquemal, J. P.; Ren, P. Implementation of Geometry-Dependent Charge Flux into the Polarizable AMOEBA+ Potential. *J. Phys. Chem. Lett.* **2020**, *11* (2), 419–426.
- (88) Das, A. K.; Demerdash, O. N.; Head-Gordon, T. Improvements to the AMOEBA Force Field by Introducing Anisotropic Atomic Polarizability of the Water Molecule. *J. Chem. Theory Comput.* **2018**, *14* (12), 6722–6733.
- (89) Lambros, E.; Paesani, F. How Good Are Polarizable and Flexible Models for Water: Insights from a Many-Body Perspective. *J. Chem. Phys.* **2020**, *153*, 60901.
- (90) Kozuch, J.; Schneider, S. H.; Boxer, S. G. Biosynthetic Incorporation of Site-Specific Isotopes in β -Lactam Antibiotics Enables Biophysical Studies. *ACS Chem. Biol.* **2020**, *15* (5), 1148–1153.

Supporting Information for

Testing the Limitations of MD-based Electric Fields Using Vibrational Stark Effect in Solution: Penicillin G as a Test Case

Jacek Kozuch, Samuel H. Schneider, Chu Zheng, Zhe Ji, Richard T. Bradshaw, and Steven G. Boxer

Table of Contents

1. Synthesis of Penicillin G Allyl Ester and Penam	p. S2
2. Overview of determination of electric fields in this work	p. S4
3. IR Spectra of Penicillin G, Penicillin G Allyl Ester and Penam	p. S6
4. Conformational analysis of POL MD electric fields for PenG and PGA	p. S8
5. Correlation of DFT C=O frequencies with side-chain MD electric fields for PenG and PGA in vacuum	p. S9
6. Side-Chain Electric Fields on the β -Lactam C=O	p. S11
7. Water models	p. S12
8. Comparison of Electrostatic MD parameters for the Penam cores	p. S16
9. Comparison of Electric Fields from FC force fields, POL force fields and DFT	p. S16

1. Synthesis of Penicillin G Allyl Ester and Penam

For the chemical synthesis the following compounds were used: triethyl amine (Sigma Aldrich, for synthesis), allyl bromide (Sigma Aldrich, 99%), sodium iodide (Sigma Aldrich, 99.5%), sodium sulfite (Sigma Aldrich, 98%), cysteamine hydrochloride (Sigma Aldrich, 98%), ethyl propiolate (Sigma Aldrich, 99%), 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (Sigma Aldrich, 98%), hexanes (> 98.5%), ethyl acetate (> 99.5%), diethyl ether (> 99%), dichloromethane (> 99.8%).

The synthesis of penicillin G allyl ester, PGA (allyl (2R,5R,6R)-3,3dimethyl-7-oxo-6-(2-phenylacetamido)-4-thia-1-azabicyclo[3.2.0]heptane-2-carboxylate), was performed according to the previously described procedure by Jeffrey and McCombie.¹

Penicillin G allyl ester: 2.025 g of penicillin G sodium salt (5.68 mmol) was dissolved in 20 mL of dry DMF and 424 mg of sodium iodide (0.5 eq, 2.83 mmol) and 750 μ L of allyl bromide (1.05 g, 1.5 eq, 8.67 mmol) were added under stirring. After 18 h, 20 mL water was added together with 420 mg of sodium sulfite and the aqueous phase was extracted 4 times with 20 mL (each) of 5:1 ethyl ether/dichloromethane. The organic phase was washed 4 times with 10 mL water, dried over sodium sulfate and the solvent was removed at vacuum. The product was purified using silica chromatography (2:1 hexane/ethyl acetate) to yield 1.974 g of a colorless highly viscous oil (93 % yield). ¹H-NMR (300Hz) in CDCl₃: 1.44 (s, 3H), 1.45 (s, 3H), 3.64 (s, 2H), 4.39 (s, 1H), 4.64 (dt, J = 6.0 Hz, 1.3 Hz, 2H), 5.35 (m, 2H), 5.51 (d, J = 4.1 Hz, 1H), 5.65 (dd, J = 9.0 Hz, 4.2 Hz, 1H), 5.9 (ddt, J = 16.6 Hz, 10.2 Hz, 6.1 Hz, 1H), 6.06 (broad d, J = 9.1 Hz, 1H), 7.31 (m, 5H). MS in ESI+ mode: 375.2 [M + H]⁺, 393.2 [M + H₂O + H]⁺, 397.3 [M + Na]⁺, 413.3 [M + K]⁺. R_f = ~ 0.3 (hexane/ethyl acetate 2:1).

The synthesis of penam, PNM (4-thia-azabicyclo[3.2.0]heptan-7-one), was performed according to procedures by Chiba et al.² and Cambie et al.³.

Ethyl thiazolidin-2-ylacetate: 1.55 g cysteamine (0.02 mol) was suspended in 10 mL dry ethanol and cooled using an ice bath, into which 2 g of ethyl propiolate (1 eq) in 10 mL dry ethanol was added dropwise. After 10 h of stirring at RT under Ar, the solvent was removed at reduced pressure. The product was isolated using silica column chromatography with hexane/ethyl acetate 3:1 as eluent to give 2.1 g of a colorless, low viscous oil (60 %). ¹H-NMR (300Hz) in CDCl₃: 1.29 (t, J = 7.1 Hz, 3H), 2.55 (broad s, 1H), 2.78 (dd, J = 15.8 Hz, 6.2 Hz, 1H), 2.87 (dd, J = 15.8 Hz, 6.5 Hz, 1H), 2.94 – 2.99 (m, 2H), 3.12 (ddd, J = 12.3 Hz, 7.5 Hz, 6.1 Hz, 1H), 3.42 (dddd, J = 12.4 Hz, 6.0 Hz, 4.8 Hz, 0.5 Hz, 1H), 4.19 (q, J = 7.1 Hz, 2H), 4.88 (t, J = 6.3 Hz, 1H). MS in ESI+ mode: 115.98 [M + H – C₂H₅S]⁺, 176.07 [M + H]⁺. R_f = 0.15 (hexane/ethyl-acetate 3:1).

Thiazolidin-2-ylacetic acid hydrochloride: 1.75 g of ethyl thiazolidon-2-ylacetate was stirred in 50 mL concentrated hydrochloric acid for 1 h at RT. 17 mL water was added and the mixture continued stirring for 15 h, after which the water was removed at reduced pressure. The residue was taken up in a minimal amount of ethanol (~120 mL) and recrystallized by adding ethyl ether (~250 mL) to yield 1.015 g of the colorless product as the hydrochloride (56 %). ¹H-NMR (300Hz) in DMSO: 2.90 (dd, J = 17.2 Hz, 8.0 Hz, 1H), 3.21 – 3.05 (m, 3 H), 3.43 (dt, J = 11.6 Hz, 6.7 Hz, 1H), 3.56 (dt, J = 12.1 Hz, 6.2 Hz, 1H), 4.90 (dd, J = 8.0 Hz, 6.1 Hz, 1H), 10.90 (broad s, 1H). MS in ESI+ mode: 130.97 [M - H₂O]⁺, 148.01 [M + H]⁺.

4-Thia-1-azobicyclo[3.2.0]heptan-7-one: 92 mg of thiazolidin-2-ylacetic acid hydrochloride were suspended in 200 mL dry DCM under Ar atmosphere. 249 mg 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (2.6 eq) was added and the solution was stirred for 23 h at RT under Ar. 70 μ L triethylamine (1 eq) were added and the solution was stirred for 2 h. The mixture was washed with water and brine, dried over sodium sulfate, and the solvent was removed at reduced

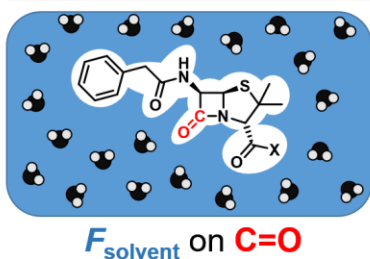
pressure. The residue was absorbed on 500 mg silica and the product was isolated using silica column chromatography (hexane/ethyl-acetate 9:1, using KMnO_4 for staining) at a yield of 22 mg (34 %). ^1H -NMR (400 Hz) in DMSO: 2.83 (m, 2H), 3.19 (m, 2H), 3.51 (dd, $J = 4.0$ Hz, 0.8 Hz, 1H), 3.93 (m, 1H), 4.94 (dd, $J = 4.0$ Hz, 2.4 Hz, 1H). MS in ES+ mode: 88.01 $[\text{M} - \text{C}_2\text{H}_2\text{O} + \text{H}]^+$, 129.98 $[\text{M} + \text{H}]^+$, 148.02 $[\text{M} + \text{H}_2\text{O} + \text{H}]^+$. $R_f = 0.20$ (hexane/ethyl-acetate 9:1). Note: 4-Thia-1-azobicyclo[3.2.0]heptan-7-one decomposes rapidly in chloroform yielding a compound with a carbonyl stretch at $1700 - 1750\text{ cm}^{-1}$ depending on the solvent.

2. Overview of determination of electric fields in this work

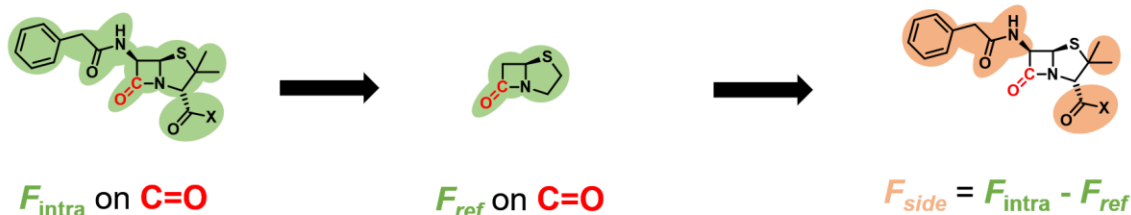
Table S1. Overview for the determination of solvent, side-chain and total environmental fields in this work; see corresponding Fig. S1.

For small, rigid solutes (e.g. PNM) – one conformation	
Fixed-charge MD solvatochromism: <ol style="list-style-type: none"> Vibrational frequencies: <ul style="list-style-type: none"> Determine peak positions from FTIR spectra of solute in solvents Average solvent electric fields: <ul style="list-style-type: none"> Run fixed-charge MD trajectories of solutes with similar solvents Determine electric field F on C=O bond from electrostatic forces f due to solvent and charge q on C and O – in GROMACS as: $f_{\text{solvent}} = f_{\text{total}} - f_{\text{solute}} \text{ with } F = f/q$ Plot vibrational frequencies and average electric fields: <ul style="list-style-type: none"> Fit using eq. 1 	Polarizable (AMOEBA) MD solvatochromism: <ol style="list-style-type: none"> Vibrational frequencies: <ul style="list-style-type: none"> Determine peak positions from FTIR spectra of solute in solvents Average solvent electric fields: <ul style="list-style-type: none"> Run AMOEBA MD trajectories of solutes with similar solvents Determine electric field F on C=O bond from induced dipoles u due to solvent and polarizabilities a on C and O – in Tinker as: $u_{\text{solvent}} = u_{\text{total}} - u_{\text{solute}} \text{ with } F = u/a$ Plot vibrational frequencies and average electric fields: <ul style="list-style-type: none"> Fit using eq. 1
For conformationally flexible solutes (e.g. PenG, PGA)	
Fixed-charge MD solvatochromism (separately for different conformations, if required): <ol style="list-style-type: none"> Vibrational frequencies: <ul style="list-style-type: none"> Determine peak positions from FTIR spectra of solute in solvents Average solvent electric fields: <ul style="list-style-type: none"> Run fixed-charge MD trajectories of solutes with similar solvents Determine electric field F on C=O bond from electrostatic forces f due to solvent and charge q on C and O – in GROMACS as: $f_{\text{solvent}} = f_{\text{total}} - f_{\text{solute}} \text{ with } F = f/q$ Average side-chain electric fields: <ul style="list-style-type: none"> Determine intramolecular electric fields on C=O from MD trajectories above (i.e. solute fields on C=O): F_{intra} Determine intramolecular electric field on C=O of rigid reference molecule from MD trajectory in vacuum (i.e. PNM in the present work): F_{ref} Difference yields conformation-dependent electric fields due to side-chains: $F_{\text{side}} = F_{\text{intra}} - F_{\text{ref}}$ Average environmental electric fields: <ul style="list-style-type: none"> Add average solvent and side-chain electric fields: $F_{\text{env}} = F_{\text{solvent}} - F_{\text{side}}$ Plot vibrational frequencies and average electric fields: <ul style="list-style-type: none"> Fit using eq. 1 	Polarizable (AMOEBA) MD solvatochromism (separately for different conformations, if required): <ol style="list-style-type: none"> Vibrational frequencies: <ul style="list-style-type: none"> Determine peak positions from FTIR spectra of solute in solvents Average solvent electric fields: <ul style="list-style-type: none"> Run fixed-charge MD trajectories of solutes with similar solvents Determine electric field F on C=O bond from induced dipoles u due to solvent and polarizabilities a on C and O – in Tinker as: $u_{\text{solvent}} = u_{\text{total}} - u_{\text{solute}} \text{ with } F = u/a$ Average side-chain electric fields: <ul style="list-style-type: none"> Determine intramolecular electric fields on C=O (as one separate polarization) from MD trajectories above (i.e. solute fields on C=O): F_{intra} Determine intramolecular electric field on C=O (again as one separated polarization group) of rigid reference molecule from MD trajectory in vacuum (i.e. PNM in the present work): F_{ref} Difference yields conformation-dependent electric fields due to side-chains: $F_{\text{side}} = F_{\text{intra}} - F_{\text{ref}}$ Average environmental electric fields: <ul style="list-style-type: none"> Add average solvent and side-chain electric fields: $F_{\text{env}} = F_{\text{solvent}} - F_{\text{side}}$ Plot vibrational frequencies and average electric fields: <ul style="list-style-type: none"> Fit using eq. 1

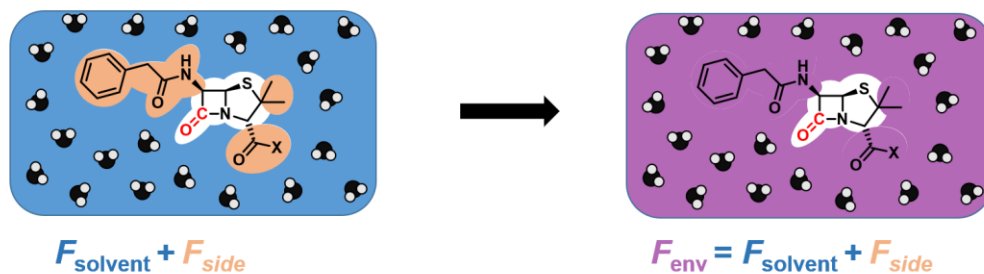
1. Average solvent electric fields:



2. Average side-chain electric fields:



3. Average environmental electric fields:



Scheme S1. Overview of the determination of solvent, side-chain and total environmental fields in this work. (1) The average solvent electric field F_{solvent} on the β -lactam $\text{C}=\text{O}$ bond is extracted from the MD trajectories. (2) From the same trajectories the total average intramolecular field F_{intra} are determined from all atoms of the solute. A separate in vacuo MD trajectory of PNM yields PNM's average intramolecular field F_{ref} . The difference yields the average side-chain electric field F_{side} (the combination of the phenylacetamide and carboxylate/ally ester through-space electrostatic forces). (3) Adding the F_{solvent} and F_{side} results in the average environmental electric field F_{env} due to the dynamic local environment onto the penam core.

3. IR Spectra of Penicillin G, Penicillin G Allyl Ester and Penam

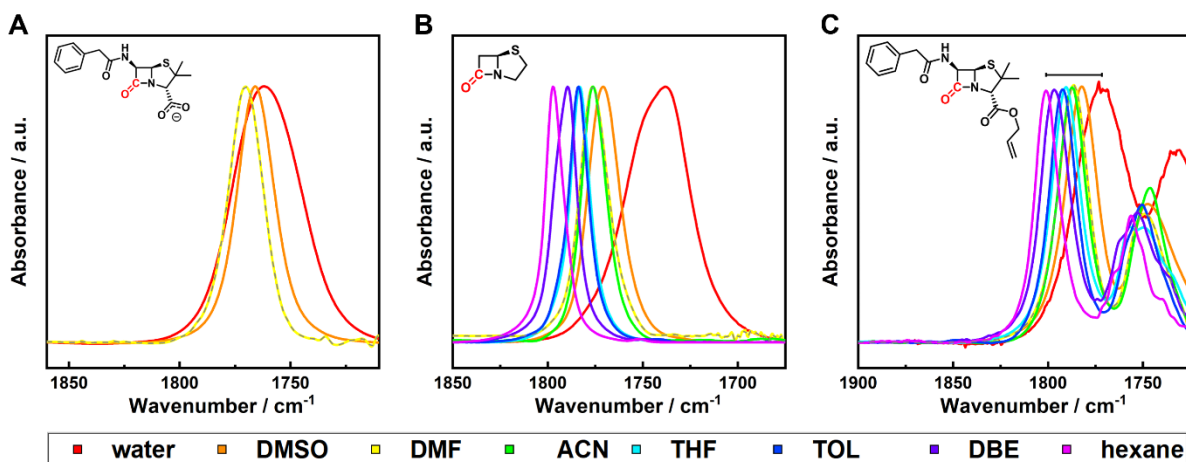


Figure S1. FTIR spectra of (A) PenG, (B) PNM and (C) PGA in the region of the β -lactam C=O mode. In C, the β -lactam C=O mode is labeled with the bar; the IR absorption between ca. 1750 – 1700 cm^{-1} is assigned to the ester mode. All spectra are normalized to equal peak maximum intensity of the β -lactam C=O. Color code is shown below; yellow spectra (DMF) are highlighted using a gray dashed line. Water is D_2O at pD 7.4.

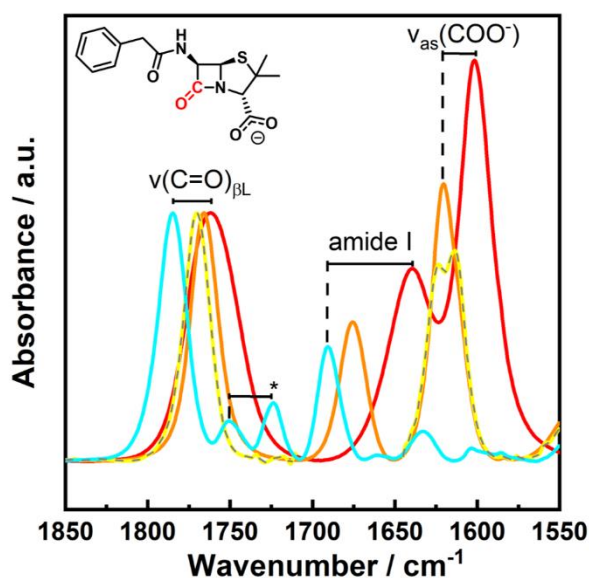


Figure S2. FTIR spectrum from Fig. S1A with the spectrum of PenG in THF added and expanded to 1550 cm^{-1} . Beside the β -lactam C=O (labeled as $\nu(\text{C}=\text{O})_{\beta\text{L}}$), the spectra in water, DMSO and DMF display the amide I and asymmetric carboxylate ($\nu_{\text{as}}(\text{COO}^-)$) modes; all spectra are normalized to similar peak intensities of the β -lactam C=O stretch. The THF spectrum differs in that two peaks at ca. 1720 and 1750 cm^{-1} appear (marked with *), which originate from the protonated carboxylate (potentially in two different orientations/environments) as the $\nu_{\text{as}}(\text{COO}^-)$ disappears. Colors are chosen as in Figure S1.

Table S2. IR peak positions of PenG, PNM, and PGA in the solvents shown in Fig. S1 with the associated average MD electric fields obtained from fixed charge and AMOEBA simulations. In addition the electric field drop, i.e. the difference between the fields at the O and C atoms of the C=O (as discussed in Fried, Wang, Boxer, Ren, Pande, *J. Phys. Chem. B*, 2013, 117, 16236-16248) is shown.

		Fixed Charge MD		Polarizable AMOEBA MD	
Solute/solvent	IR Frequency / cm ⁻¹	Average Electric Field / MV/cm	Average Electric Field Drop across C=O / MV/cm	Average Electric Field / MV/cm	Average Electric Field Drop across C=O / MV/cm
PenG					
Water (tip3p)	1761.9 (1759.9) ^a	-62.4	-31.0	-91.6	-36.6
DMSO	1766.0	-19.2	-6.2	-28.6	-10.3
DMF	1770.3	-16.4	-2.5	-39.9	-6.83
PNM					
Water (tip3p)	1704.0	-70.2	-42.7	-92.1	-51.7
DMSO	1733.7	-31.5	-4.7	-47.7	-9.9
DMF	1737.6	-27.8	-4.4	-46.5	-9.1
ACN	1739.8	-27.5	-6.5	-46.1	-11.8
THF	1740.3	-17.8	-2.7	-36.7	-8.32
TOL	1742.2	-15.1	-3.6	-30.6	-8.8
DBE	1746.3	-8.5	-1.1	-18.9	-5.3
HEX	1750.2	0.0	0.1	-16.1	-5.3
PGA					
Water (tip3p)	1772.1	-62.6	-24.8	-87.4	-29.7
DMSO	1782.5	-30.4	-5.8	-52.0	-10.2
DMF	1786.8	-25.0	-3.8	-38.1	-6.2
ACN	1787.4	-26.6	-3.3	-46.1	-10.4
THF	1790.7	-14.7	-4.2	-42.6	-9.4
TOL	1792.3	-12.2	-1.6	-32.0	-7.5
DBE	1797.0	-6.8	-2.1	-22.8	-9.1
HEX	1801.1	0.0	0.1	-14.1	-5.0

^a This frequency refers to the major, lower frequency component for the β -lactam C=O stretch band obtained in Fig. S3

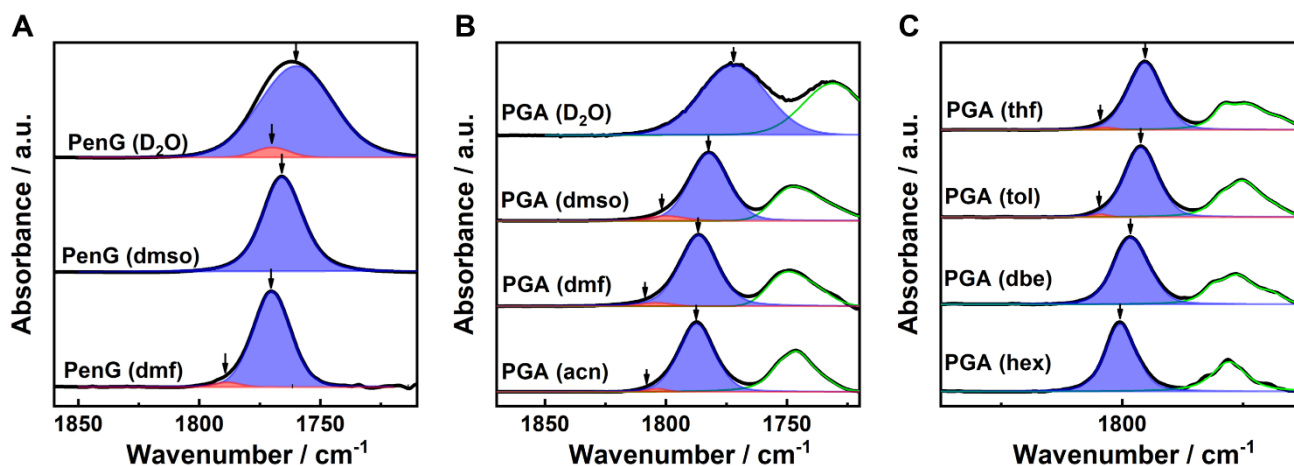


Figure S3. Peak fitting of FTIR spectra of (A) PenG and (B, C) PGA in the region of the β -lactam C=O mode using pseudo-Voigt and Gaussian peak shapes. Components modelling the β -lactam C=O stretch are shown as light red and light blue areas; the ester C=O stretch is indicated by green lines. Arrows highlight the position the components contributing to the β -lactam C=O stretch band with a major lower frequency and minor (or absent) higher frequency contribution, separated by ca. 15 cm^{-1} .

4. Conformational analysis of POL MD electric fields for PenG and PGA

In contrast to FC simulations, only the solute-solvent systems PenG+DMSO, PGA+Hex, PGA+DBE showed sampling of conformation 1 (or a mixture of 1 and 2) and all other systems exclusively yielded conformation 2. To obtain data on the unsampled conformation, the AMOEBA MD run was repeated with the option “RESTRAINT-TORSION 10 8 12 14 1 110 -130” for conformation 1 or “RESTRAINT-TORSION 10 8 12 14 1 -130 -10” for conformation 2. The combined resulting average electric fields are shown in Table S2.

Table S3. Data set used for Fig. 3 in the main text and Fig. S4 below.

Solute/solvent	Fixed Charge MD			Polarizable AMOEBA MD		
	Average Solvent Electric Field / MV/cm	Conf. 1 Average Solvent Electric Field / MV/cm	Conf. 2 Average Solvent Electric Field / MV/cm	Average Electric Field Solvent / MV/cm	Conf. 1 Average Solvent Electric Field / MV/cm	Conf. 2 Average Solvent Electric Field / MV/cm
PenG						
Water (tip3p)	-62.4	-50.6	-70.5	-91.6	-58.5 ^a	-91.6
DMSO	-19.2	-14.1	-28.3	-28.6	-28.6	-44.3 ^a
DMF	-16.4	-11.3	-22.0	-39.9	-22.2 ^a	-39.9
PGA						
Water (tip3p)	-62.2	-47.8	-70.1	-87.4	-57.5 ^a	-87.4
DMSO	-30.4	-25.4	-37.5	-52.0	-33.4 ^a	-52.0
DMF	-25.0	-20.3	-33.2	-38.1	-34.2 ^a	-38.1
ACN	-26.6	-19.6	-32.9	-46.1	-31.1 ^a	-46.1
THF	-14.7	-13.6	-19.2	-42.6	-26.0 ^a	-42.6
TOL	-12.2	-13.6	-15.3	-32.0	-23.0 ^a	-32.0
DBE	-6.8	-5.0	-8.9	-22.8	-17.5	-25.1
HEX	0.0	0.0	0.0	-14.1	-11.6	-15.1

^a Average electric fields for these conformations were determined from restraint simulations as indicated in the text.

Based on the specific conformation dependent solvent electric fields shown in Table S2, we determined separate solvatochromic analyses for both conformations shown in Figure S4. Addition of the intramolecular fields to obtain average environmental fields results in Figure 5 in the main text.

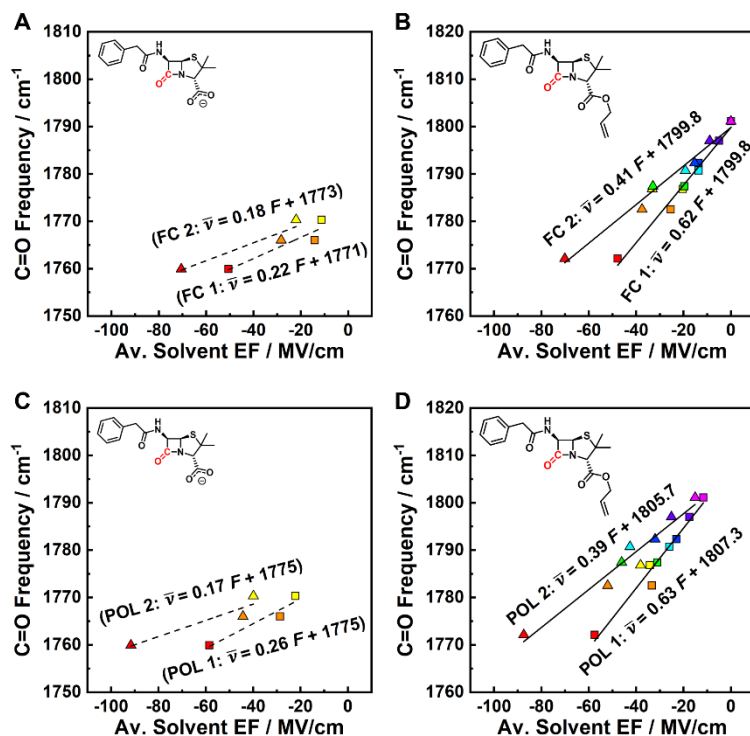


Figure S4. Conformation-dependent solvatochromic VSE analysis of PenG and PGA using electric fields obtained from FC (A,B) and POL MD simulations (C,D). Based on the two conformations observed in MD trajectories separate VSE evaluations are performed labelled with FC 1 and FC 2 as well as POL 1 and POL 2, referring to the corresponding conformation depicted in Figure 4 (solvent colors refer to the same pattern as in Figure 2; data points for conf. 1 and 2 are shown as squares and triangles, respectively). The linear regressions provide the following results in (A): $\bar{\nu} = (0.22 \pm 0.08) F + (1771 \pm 3)$ (PenG FC 1, with $R^2 = 0.88$) and $\bar{\nu} = (0.18 \pm 0.06) F + (1773 \pm 3)$ (PenG FC 2, with $R^2 = 0.91$); in (B): $\bar{\nu} = (0.62 \pm 0.03) F + (1799.8 \pm 0.7)$ (PGA and FC 1, with $R^2 = 0.98$) and $\bar{\nu} = (0.41 \pm 0.02) F + (1799.8 \pm 0.8)$ (PGA and FC 2, with $R^2 = 0.98$); in (C) $\bar{\nu} = (0.26 \pm 0.07) F + (1775 \pm 3)$ (PenG and POL 1, with $R^2 = 0.93$) and $\bar{\nu} = (0.17 \pm 0.06) F + (1775 \pm 4)$ (PenG and POL 2, with $R^2 = 0.88$); in (D): $\bar{\nu} = (0.63 \pm 0.05) F + (1807.3 \pm 1.5)$ (PGA and POL 1, with $R^2 = 0.97$) and $\bar{\nu} = (0.39 \pm 0.04) F + (1805.7 \pm 1.8)$ (PGA and POL 2, with $R^2 = 0.95$).

5. Correlation of DFT C=O frequencies with side-chain MD electric fields for PenG and PGA in vacuum

To analyze if PenG's and PGA's side-chain groups contribute to the shift of the β -lactam C=O stretch via through space electrostatics or through bond effects, we compared gas-phase DFT vibrational frequency calculations with MD simulations of the corresponding side-chain electric fields. First, we optimized the geometries of the PenG anion and PGA using density functional theory (DFT) at the b3lyp/6-311++g(2d,2p) level with the C-C-N-C dihedral restrained between the phenylacetamido group and the penam core and then determined the vibrational frequencies of the β -lactam C=O using normal mode analysis. The frequencies were scaled by a factor of 0.9679 according to M. P. Andersson and P. Uvdal.⁴ Then, we calculated fixed charge MD trajectories in vacuum with restricted dihedrals of the side chains similar to those in the DFT calculation and extracted the averaged intramolecular electric fields exerted on the β -lactam C=O. It should be noted that within the Amber force field intramolecular electrostatic interactions are excluded for 1-2 and 1-3 pairs, and 1-4 pair electrostatics are scaled down by a factor of 0.833; all further distant pairs are considered via direct electrostatics. As such one must be careful when interpreting absolute values of the intramolecular electric fields, which contain considerable unphysical (i.e. forcefield-dependent) offsets due these electrostatic cutoffs. Note that PenG, PGA and PNM contain slightly different atomic partial charges, which thus leads to molecule specific offsets in

absolute intramolecular electric fields (see further discussion below). In the simplest attempt to correct for such offsets, we subtracted from the intramolecular electric fields of PenG and PGA the intramolecular electric field of PNM (17.2 MV/cm); in this way we are self-consistently accounting for the contribution of the penam core in order to obtain side-chain electric fields in the simplest approximation.

Plotting the DFT C=O frequencies and averaged side-chain electric fields against the restrained dihedral angle (Figure S5 A and B), we obtain very similar trends for the PenG anion and PGA. That is, the vibrational frequency and fields follow a similar trend spanning a total of ca. 30 cm⁻¹ and ca. 50 MV/cm, respectively, with minima and maxima at ca. 150 and 330 deg. In Figure S5 C, the DFT C=O frequencies are correlated with the side-chain MD electric fields. Interestingly, both molecules show linear trends with slopes of 0.62 ± 0.05 and 0.70 ± 0.10 cm⁻¹/(MV/cm) for PenG and PGA, respectively, which surprisingly reproduce the Stark tuning rate of the β -lactam C=O within the limitations and accuracy of this approach. This indicates that the observed frequency shifts of the β -lactam C=O due the side groups of PGA and PenG can be treated within the classical electrostatic approximation of MD simulations, i.e. via a through space electric field exerted on the β -lactam C=O.

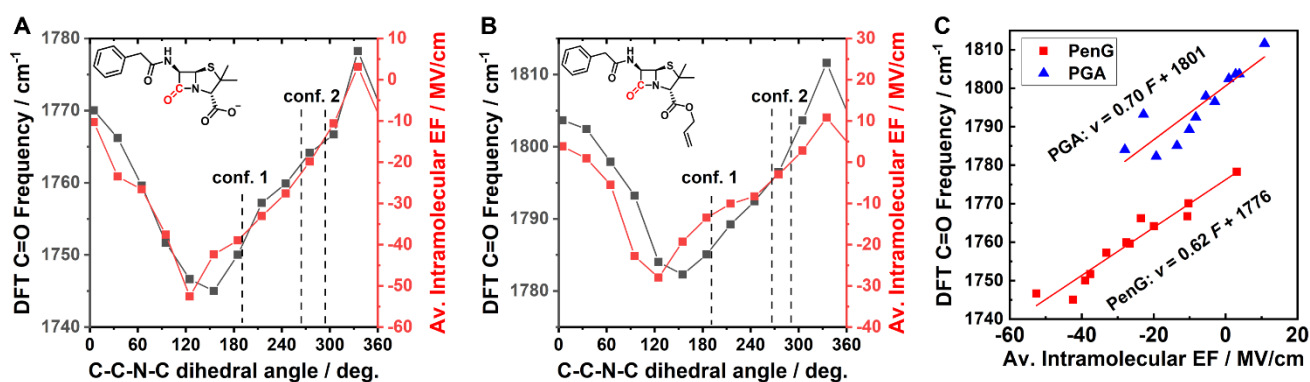


Figure S5. Correlation of DFT-based β -lactam C=O frequencies with MD-derived average intramolecular electric fields (both restrained at specified C-C-N-C dihedrals) for the PenG anion (A) and PGA (B). (A,B): Vibrational frequencies of β -lactam C=O are shown on the left y-axis (scaled by factor of 0.9769, see SI text); average side-chain MD electric fields exerted the β -lactam C=O shown on x-axis. (C): Correlation of DFT C=O frequencies with the average side-chain MD electric fields at corresponding dihedrals results in linear trends with $\bar{\nu} = (0.62 \pm 0.05) F + (1776 \pm 1)$ (PenG with $R^2 = 0.93$) and with $\bar{\nu} = (0.70 \pm 0.10) F + (1801 \pm 1)$ (PGA with $R^2 = 0.80$).

6. Side-Chain Electric Fields on the β -Lactam C=O

Table S4. Average side-chain electric fields of PNM, PenG, PGA in conf. 1 and 2 for FC and POL MD simulations. All side-chain electric fields are reported in reference to PNM's total intramolecular field (17.2 MV/cm), to isolate the contributions of the side groups, i.e. the phenyl acetamido chain, the carboxylate, or the allyl ester group. To obtain a similar approach for POL MD, where the penam core was one polarization group for each of the molecules, we calculated MD vacuum trajectories of PNM, PenG and PGA and recalculated the electric fields with the C=O set as a separate group. This approach did not shift total intramolecular fields of PNM, but shifted those of PGA and PenG by -27.7 and -47.7 MV/cm, which were added to the side-chain fields when the penam core was one polarization group.

Solute/solvent	Fixed Charge MD			Polarizable AMOEBA MD		
	Average Intramol. Electric Field / MV/cm	Conf. 1 Average Intramol. Electric Field / MV/cm	Conf. 2 Average Intramol. Electric Field / MV/cm	Average Intramol. Electric Field / MV/cm	Conf. 1 Average Intramol. Electric Field / MV/cm	Conf. 2 Average Intramol. Electric Field / MV/cm
PNM						
Water (tip3p)	-0.7			0		
DMSO	0.1			0		
DMF	0.0			0		
ACN	-1.4			0		
THF	-1.9			0		
TOL	-0.9			0		
DBE	-1.2			0		
HEX	-0.7			0		
PenG						
Water (tip3p)		-34.4	-15.1		-29.3	1.6
DMSO		-35.3	-18.7		-30.3	-3.5
DMF		-36.9	-18.5		-32.3	-1.6
vacuum		-37.2	-29.8		-35.7	-31.7
PGA						
Water (tip3p)		-11.8	9.9		-13.1	14.5
DMSO		-11.9	7.3		-12.5	11.0
DMF		-11.5	5.3		-13.1	10.5
ACN		-11.8	3.5		-12.1	11.1
THF		-10.0	-0.9		-13.6	10.7
TOL		-11.6	-2.7		-12.7	6.9
DBE		-12.3	-1.5		-12.4	6.9
HEX		-12.2	-3.6		-12.5	6.4

7. Water models

For FC MD (Figure S6, Table S5), the observed trend in water model-dependent electric fields follows the increasing partial charge (in units of e) on the H atoms from ca. 0.24 (tip5p models; 0.39 for tip5p2018) to 0.42 (tip3p models) and 0.52 – 0.56 (tip4p models), which systematically modulates the electric field based on Coulombic interactions. Tip7p deviates from this correlation with a partial charge of 0.58, which is due to the considerably more complex electrostatics with the virtual site between H and O atoms with a charge of ca. -0.46. In addition, the above trend is associated with a shortening of the average heavy-atom H-bond distance to the C=O, which decreases by ca. 0.2 Å (Figure S7). This shortening leads to a closer proximity and less fluctuations of water's H atoms relative to the carbonyl O atom. Interestingly, tip7p combines these solvation properties, adopting an H position similar to that of tip5p, but with a much narrower distribution as observed for tip4p models.

For POL MD, (Figure S6, Table S5), Amoeba03* yields similar solvent electric fields as the original amoeba03 water as only bonded parameters were refined. However, amoeba14 water results in increased electric field magnitudes at the β -lactam C=O due to ca. 0.2 Å-shorter and more homogeneous H bonds than amoeba03 (see Figure S8). Although iamoeba also yields ca. 0.2 Å-shorter H-bonds, there is greater heterogeneity which leads to a decrease in the electric fields as polarization due to induced dipoles is not considered in this direct polarization model (in contrast to mutual polarization used in the other amoeba models, where induced dipoles contribute to polarization).

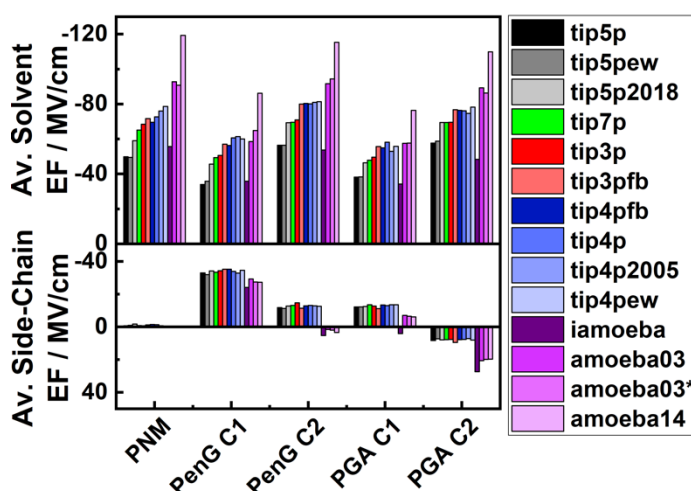


Figure S6. Effect of different water models on the average solvent (top) and side-chain electric fields (bottom) exerted on the β -lactam C=O in FC and POL MD simulations. The water models tip5p (black and gray), tip7p (green), tip3p (red colors) and tip4p (blue colors) for PNM as well as PGA and PenG in both conformations ("C#") 1 and 2 are sorted in accordance with the general trend of increasing electric fields magnitudes as discussed in the text. Amoeba water models are shown in purple colors.

Table S5. Average MD solvent and side-chain electric fields determined in various water models for PNM and PenG's and PGA's separate conformations, as in Table S3.

Solute/solvent	Fixed Charge/Polarizable AMOEBA MD						
	Partial charge on H / e	Average Solvent Electric Field / MV cm ⁻¹	Conf. 1 Average Solvent Electric Field / MV cm ⁻¹	Conf. 2 Average Solvent Electric Field / MV cm ⁻¹	Average Intramol. Electric Field / MV cm ⁻¹	Conf. 1 Average Intramol. Electric Field / MV cm ⁻¹	Conf. 2 Average Intramol. Electric Field / MV cm ⁻¹
PenG							
tip5p-ew	0.241		-33.9	-56.4		-33	-11.8
tip5p	0.241		-35.8	-56.4		-31.9	-11.3
tip5p-2018	0.394137		-45.6	-69.3		-34.2	-12.8
tip7p	0.58014 (-0.45837) ^a		-49.4	-69.5		-33.4	-13.1
tip3p	0.417		-50.6	-70.9		-34.4	-14.8
tip3p-fb	0.42422		-57	-80.0		-35.2	-11.4
tip4p-fb	0.52		-56.3	-80.4		-35.2	-12.9
tip4p	0.52		-60.6	-80.1		-34	-13.2
tip4p-2005	0.5564		-61.4	-81.0		-32.9	-12.9
tip4p-ew	0.52422		-60	-81.3		-34.6	-12.6
iameboa	-		-35.8	-53.7		-24.13	5.2
amoeba03	-		-58.5	-91.6		-29.3	1.5
amoeba03*	-		-64.8	-94.6		-27.42	1.9
amoeba14	-		-86.2	-115.4		-27.31	3.37
PNM							
tip5p-ew	0.241	-49.8			-0.7		
tip5p	0.241	-49.5			-1.0		
tip5p-2018	0.394137	-59.0			-1.8		
tip7p	0.58014 (-0.45837) ^a	-65.1			-0.8		
tip3p	0.417	-68.5			-0.7		
tip3p-fb	0.42422	-71.7			-1.2		
tip4p-fb	0.52	-69.6			-1.5		
tip4p	0.52	-72.6			-1.3		
tip4p-2005	0.5564	-76.0			-0.7		
tip4p-ew	0.52422	-78.6			0.1		
iameboa	-	-55.6			0		
amoeba03	-	-92.7			0		
amoeba03*	-	-90.8			0		
amoeba14	-	-119.3			0		
PGA							
tip5p-ew	0.241		-38.2	-57.6		-12.2	8.3
tip5p	0.241		-38.3	-58.8		-12.2	7.2
tip5p-2018	0.394137		-46.4	-69.4		-12.5	7.9
tip7p	0.58014 (-0.45837) ^a		-47.9	-69.4		-13.5	7.8
tip3p	0.417		-49.6	-69.6		-12.7	7.7
tip3p-fb	0.42422		-55.7	-76.8		-11.2	9.5
tip4p-fb	0.52		-54.9	-76.4		-13.4	7.8
tip4p	0.52		-58.2	-76.1		-13	7.7
tip4p-2005	0.5564		-52.9	-74.7		-13.5	7.1
tip4p-ew	0.52422		-55.8	-78.3		-13.5	8
iameboa	-		-34.2	-48.4		4.12	27.37
amoeba03	-		-57.5	-89.3		-7	20.6
amoeba03*	-		-57.6	-86.3		-6.48	19.8
amoeba14	-		-76.3	-109.9		-6.17	19.73

^a This charge refers to the virtual site located halfway between the O and H atom.

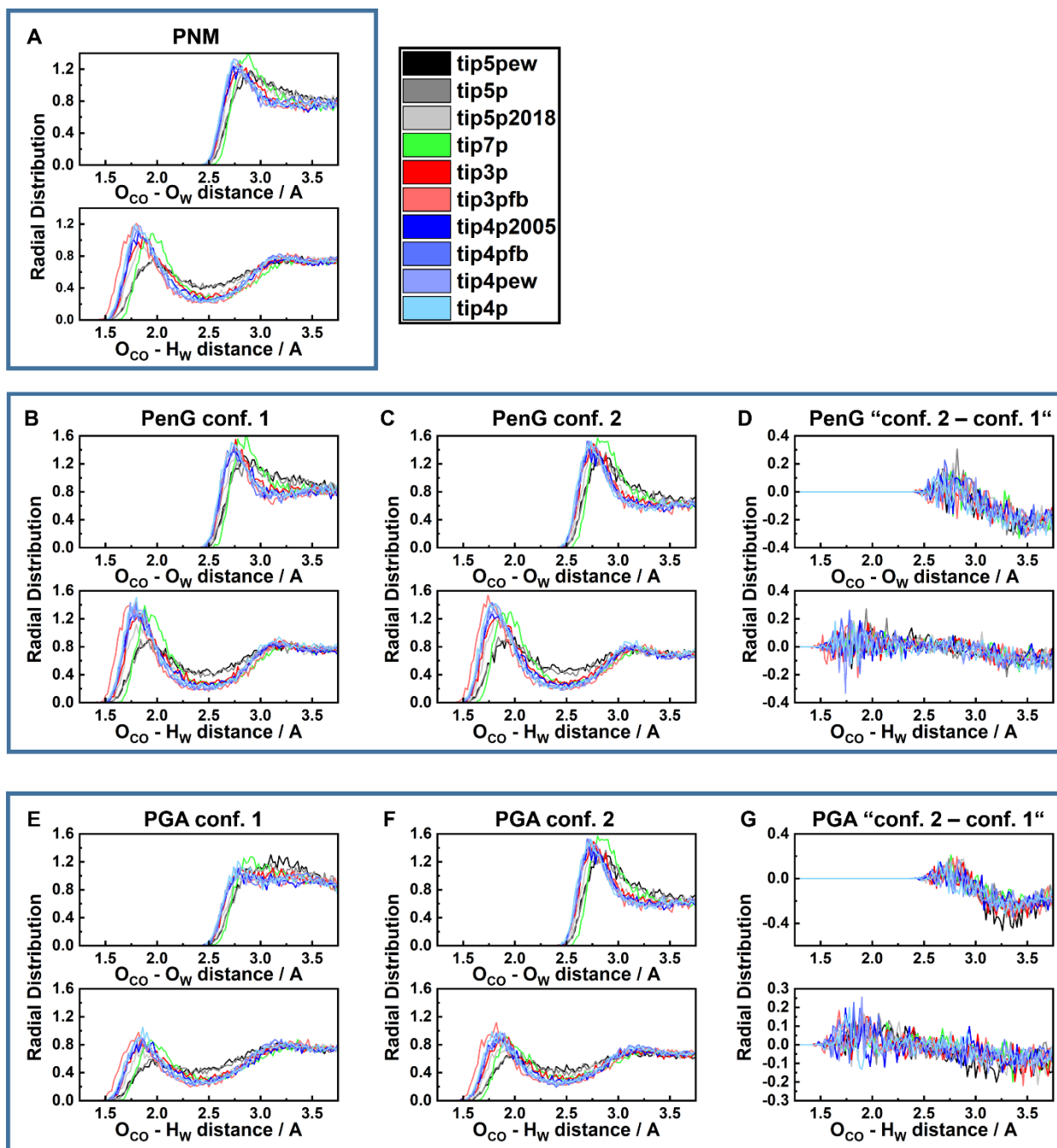


Figure S7. Radial distribution functions of water atom positions for FC MD simulations in reference to the O atom of the β -lactam C=O groups (i.e. $O_{CO} - O_W$ and $O_{CO} - H_W$ distances shown on top and bottom, respectively, for each case) for PNM, PenG and PGA with the water models shown in the legend. (A): Radial distribution functions for PNM. (B,C): Radial distribution functions for PenG’s conformations 1 and 2, as labelled. (D): Difference between the radial distributions in B and C. (E,F): Radial distribution functions for PGA’s conformations 1 and 2, as labelled. (G): Difference between the radial distributions in E and F.

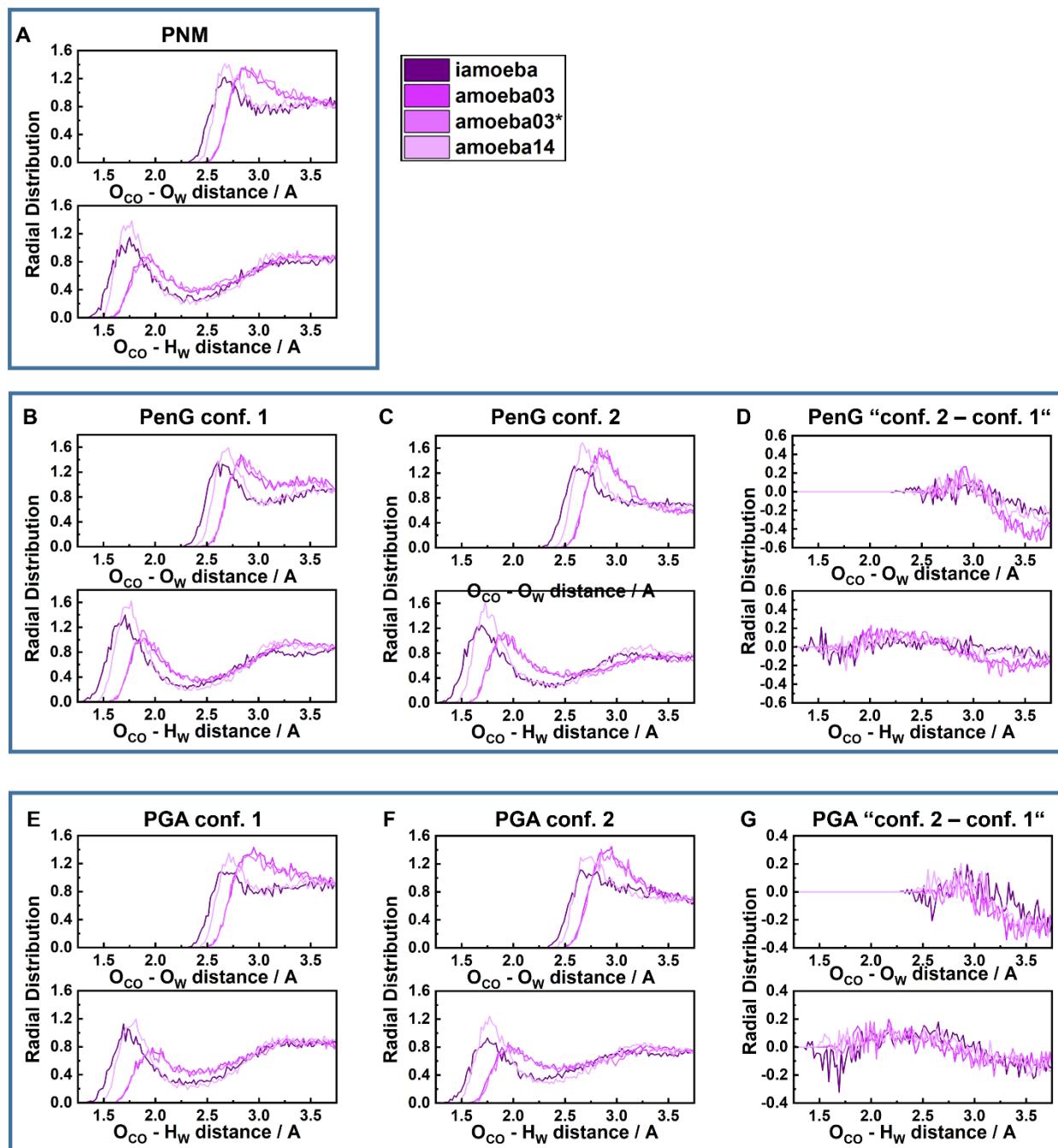


Figure S8. Radial distribution functions of water atom positions for POL MD simulations in reference to the O atom of the β -lactam C=O groups (i.e. $O_{CO} - O_W$ and $O_{CO} - H_W$ distances shown on top and bottom, respectively, for each case) for PNM, PenG and PGA with the water models shown in the legend. (A): Radial distribution functions for PNM. (B,C): Radial distribution functions for PenG's conformations 1 and 2, as labelled. (D): Difference between the radial distributions in B and C. (E,F): Radial distribution functions for PGA's conformations 1 and 2, as labelled. (G): Difference between the radial distributions in E and F.

8. Comparison of Electrostatic MD parameters for the Penam cores

Table S6. Charge parameters for PenG, PGA and PNM used in FC MD.

		PenG	PGA	PNM
C	charge / e	0.650501	0.651501	0.667501
O	charge / e	-0.572501	-0.516501	-0.534501

Table S7. Multipole parameters for PenG, PGA and PNM used in POL MD.

		PenG	PGA	PNM
C	charge / e	0.58689	0.57804	0.54137
	dipole (X) / e bohr	-0.24782	-0.19770	-0.25139
	dipole (Z) / e bohr	-0.04286	-0.09096	-0.27306
	quadrupole (XX) / e bohr ²	0.09078	0.03877	-0.13238
	quadrupole (YY) / e bohr ²	-0.45408	-0.33729	-0.25279
	quadrupole (XZ) / e bohr ²	0.00386	-0.00782	-0.03915
	quadrupole (ZZ) / e bohr ²	0.36330	0.29852	0.38517
O	charge / e	-0.66147	-0.60352	-0.60581
	dipole (X) / e bohr	-0.03378	-0.00785	0.04409
	dipole (Z) / e bohr	-0.17824	-0.16200	-0.17138
	quadrupole (XX) / e bohr ²	-0.18476	-0.26429	-0.20693
	quadrupole (YY) / e bohr ²	0.32031	0.29959	0.39592
	quadrupole (XZ) / e bohr ²	-0.10790	-0.08865	0.19010
	quadrupole (ZZ) / e bohr ²	-0.13555	-0.03530	-0.18899

9. Comparison of Electric Fields from FC force fields, POL force fields and DFT

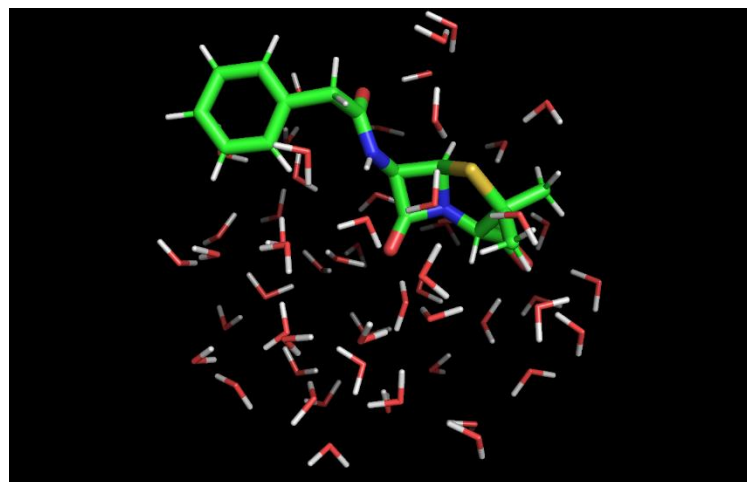


Figure S10. Excerpt structure obtained from FC MD with average solvent field, containing one PenG molecule and 56 water molecules.

In order to test whether the discrepancy of MD electric fields on PenG's β -lactam C=O is due to electrostatic issues of the 3-point water models, we performed a direct comparison of solvent electrostatics based on FC force fields, POL field fields, and DFT (b3lyp/6-311++g**) in the style of a similar comparison in proteins⁵. For this we extracted one frame of the FC MD trajectory of PenG in tip3p water, in which the electric field at the β -lactam C=O was closest to the average. From this frame we extracted an excerpt containing (a) the PenG molecule and (b) a shell of water molecules within 8 Å distance of the C=O's O atom; in total, this excerpt contained 1 PenG molecule and 56 water molecules (Figure S10). Next, we determined the electric fields from this structure using the three methods with the

same atomic positions (i.e. without any further optimization). Maintaining the atom positions allows us to specifically isolate the differences due to the electrostatic properties of the water molecules with tip3p and amoeba03 parameters as well as at the chosen DFT level. We are aware that for a definitive answer extensive sampling would be needed to provide a statistically relevant picture. As such, we see this comparison as a simple test pointing towards potential, more extensive studies. Using this approach, we obtain solvent electric fields of -65 MV/cm, -75 MV/cm and -30 MV/cm from FC force fields, POL field fields, and DFT (b3lyp/6-311++g**). As such, we observe that DFT yields solvent fields smaller by $|\Delta F| = 35 - 45$ MV/cm than the MD water models tip3p and amoeba03. This difference is in line with the overestimated fields in tip3p and amoeba03 water observed in Figure 6 and it is therefore possible that averaged DFT electric field (corresponding to an entire MD trajectory) could compensate for the discrepancy in water.

References:

- (1) Jeffrey, P. D.; McCombie, S. W. Homogeneous, Palladium(0)-Catalyzed Exchange Deprotection of Allylic Esters, Carbonates, and Carbamates. *J. Org. Chem.* **1982**, *47*, 587–590. <https://doi.org/10.1002/anie.201201190>.
- (2) Chiba, T.; Sakaki, J.; Takahashi, T.; Aoki, K.; Kamiyama, A.; Kaneko, C.; Sato, M. Studies on Amino Acid Derivatives. Part 7. General Method for the Synthesis of Penam and Cepham and Their Substituted Derivatives. *J. Chem. Soc. Perkin Trans. 1* **1987**, No. 0, 1845. <https://doi.org/10.1039/p19870001845>.
- (3) Cambie, R.; Clark, G.; Jones, T.; Rutledge, P.; Strange, G.; Woodgate, P. Vic-Iodo Thiocyanates and Iodo Isothiocyanates. IX. A Synthesis of Penam and Other Polycyclic B-Lactams. *Aust. J. Chem.* **1985**, *38* (5), 745. <https://doi.org/10.1071/ch9850745>.
- (4) Andersson, M. P.; Uvdal, P. New Scale Factors for Harmonic Vibrational Frequencies Using the B3LYP Density Functional Method with the Triple- ζ Basis Set 6-311+G(d,P). *J. Phys. Chem. A* **2005**, *109* (12), 2937–2941. <https://doi.org/10.1021/jp045733a>.
- (5) Bradshaw, R. T.; Dziedzic, J.; Skylaris, C. K.; Essex, J. W. The Role of Electrostatics in Enzymes: Do Biomolecular Force Fields Reflect Protein Electric Fields? *J. Chem. Inf. Model.* **2020**, *60* (6), 3131–3144. <https://doi.org/10.1021/acs.jcim.0c00217>.