



# A two-directional vibrational probe reveals different electric field orientations in solution and an enzyme active site

Chu Zheng<sup>1,3</sup>, Yuezhi Mao<sup>1,3</sup>, Jacek Kozuch<sup>2</sup>, Austin O. Atsango<sup>1</sup>, Zhe Ji<sup>1</sup>, Thomas E. Markland<sup>1</sup>✉ and Steven G. Boxer<sup>1</sup>✉

**The catalytic power of an electric field depends on its magnitude and orientation with respect to the reactive chemical species. Understanding and designing new catalysts for electrostatic catalysis thus requires methods to measure the electric field orientation and magnitude at the molecular scale. We demonstrate that electric field orientations can be extracted using a two-directional vibrational probe by exploiting the vibrational Stark effect of both the C=O and C-D stretches of a deuterated aldehyde. Combining spectroscopy with molecular dynamics and electronic structure partitioning methods, we demonstrate that, despite distinct polarities, solvents act similarly in their preference for electrostatically stabilizing large bond dipoles at the expense of destabilizing small ones. In contrast, we find that for an active-site aldehyde inhibitor of liver alcohol dehydrogenase, the electric field orientation deviates markedly from that found in solvents, which provides direct evidence for the fundamental difference between the electrostatic environment of solvents and that of a preorganized enzyme active site.**

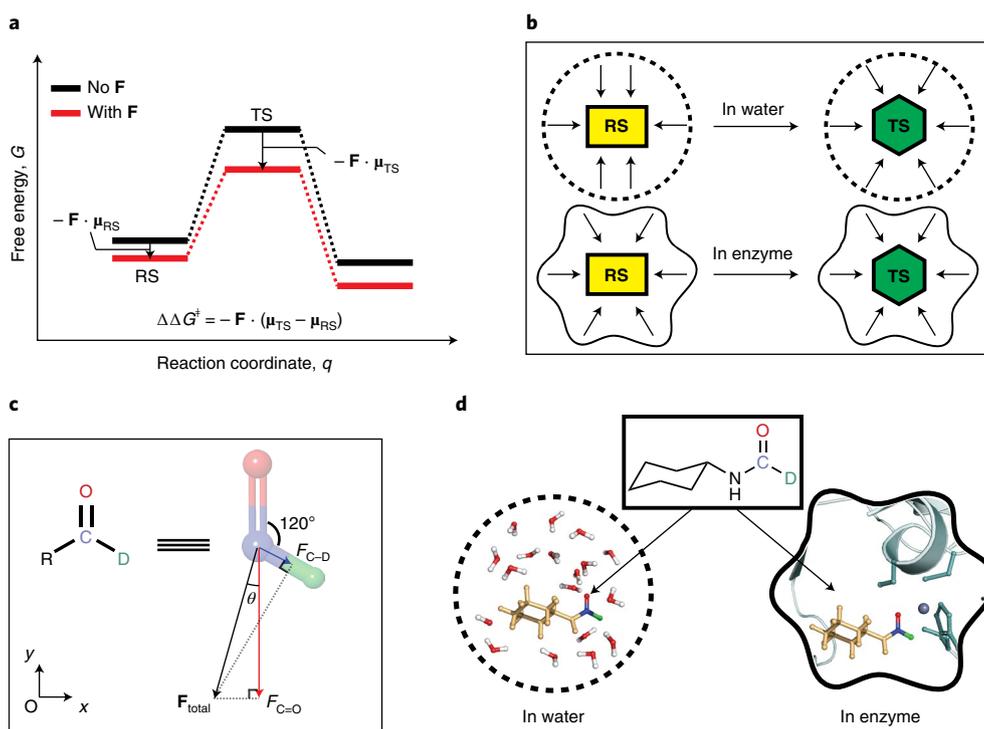
The physical basis for the remarkable proficiency and specificity of enzymes under mild physiological conditions has been widely debated. One essential contribution is electrostatic preorganization<sup>1,2</sup>, the precise positioning of functional groups with charges and dipoles by the protein scaffold that preferentially stabilizes the transition state (TS) over the reactant state (RS) due to the electric field at the active site, which leads to a decrease in the activation free energy barrier (Fig. 1a). This mechanism is directly supported by studies of several model enzymes<sup>2–7</sup>, among which a prime example is ketosteroid isomerase<sup>8,9</sup>, in which a large electric field (about  $-150\text{ MV cm}^{-1}$ ) exerted by functional groups at the active site was measured, which leads to an about  $10^5$ -fold rate of acceleration and accounts for a large fraction of the reduction in the activation free energy compared with that of the same reaction in water. Electric fields can be measured by exploiting the vibrational Stark effect (VSE), in which the infrared or Raman frequency shifts are related to the projection of the field on a particular bond. This is accomplished by using VSE probes calibrated via a combination of vibrational Stark spectroscopy (VSS), vibrational solvatochromism and molecular dynamics (MD) simulations<sup>2,8,10,11</sup>.

The concept of electrostatic preorganization is illustrated in Fig. 1b, which compares a reaction that involves charge reorganization occurring in water to that at the active site of an enzyme. As solvent molecules, such as water, fluctuate around the reactant to stabilize the RS dipole moment on average, solvent configurations that stabilize the TS are rare and thus require substantial reorganization in the course of a reaction. In contrast, an electrostatically preorganized catalyst active site requires minimal adjustments to preferentially stabilize the charge distribution in the TS, and thus minimizes the reorganization energy<sup>1</sup>. Although this viewpoint is supported by computational studies<sup>11,12–19</sup>, there is limited direct experimental evidence for the critical assumption that the electric field orientation experienced by a substrate at enzyme active sites

differs from that in solvents. The challenge to experimentally test this hypothesis lies in the fact that most previous studies only measured the projection of the electric fields along one chemical bond at a time, and even when this is the bond involved in the catalysis, it only provides information on the magnitude of the projected electric field, but not on the orientation of the total field.

In this work, we directly measured the orientation of electric fields in condensed-phase environments using a vibrational probe that bore two chemical bonds whose vibrational frequencies were both markedly shifted by the electric fields that arose from their chemical environments, which allowed us to measure the fields experienced by this probe along two directions (Fig. 1c). Using the deuterated (aldehyde H) form of *N*-cyclohexylformamide (CXF-D; Fig. 1d), a liver alcohol dehydrogenase (LADH) inhibitor<sup>20,21</sup>, we obtained the projections of the electric field that arose from the solvents and LADH active site on both the carbonyl (C=O) and carbon–deuterium (C–D) bonds that are approximately  $120^\circ$  from each other (Fig. 1c). Intriguingly, we found that the vibrational frequencies of C=O ( $\nu_{\text{C=O}}$ ) and C–D ( $\nu_{\text{C–D}}$ ) shifted in opposite directions as the solvent polarity increased, which indicated the coexistence of stabilization (C=O) and destabilization (C–D) of the chemical bonds around a single atom in solution. To elucidate these observations and introduce a field-frequency map for this probe in complex condensed phase environments, such as enzyme active sites, we developed a computational framework that combined MD simulations with recently introduced electronic structure partitioning methods<sup>22,23</sup>, which gives an improved accuracy and reliability in predicting the solvent electric fields compared with those of previously available methods. For a wide range of solvents, we observed a strong correlation between the electric fields experienced by the C=O and C–D bonds, which suggests that different solvents, despite their diverse chemical structures and polarities, exert electric fields of a similar orientation on the aldehyde moiety. By contrast,

<sup>1</sup>Department of Chemistry, Stanford University, Stanford, CA, USA. <sup>2</sup>Experimental Molecular Biophysics, Department of Physics, Freie Universität Berlin, Berlin, Germany. <sup>3</sup>These authors contributed equally: Chu Zheng, Yuezhi Mao. ✉e-mail: [tmarkland@stanford.edu](mailto:tmarkland@stanford.edu); [sboxer@stanford.edu](mailto:sboxer@stanford.edu)



**Fig. 1 | Electrostatic catalysis and the two-directional vibrational probe used in this work.** **a**, The model of electrostatic catalysis in enzymes: the reaction barrier is reduced by preferential stabilization of the TS because of the preorganized electric field,  $F$ , at the active site. **b**, Schematic illustration of the environmental response to the formation of the TS from RS during a chemical reaction that involves charge displacement (from rectangular to hexagonal for illustrative purposes). The top row illustrates the solvent reorganization in water (arrows represent the solvent dipoles), which contrasts with the minimal change in electric field orientation at the active site of enzymes due to electrostatic preorganization in the bottom row. **c**, The concept of probing the orientation of an electric field by using the C=O and C–D stretch vibrations of an aldehyde. The total electric field on the two-dimensional plane of the aldehyde moiety is the vector sum of the fields measured along the two directions. **d**, The structure of the VSE probe CXF-D used in this article to measure the electric field orientations in solvents (left) and at the active site of LADH (right). The crystal structure of the wild-type LADH in complex with NADH and CXF was resolved at 1.43 Å (Protein Data Bank: 7RM6).

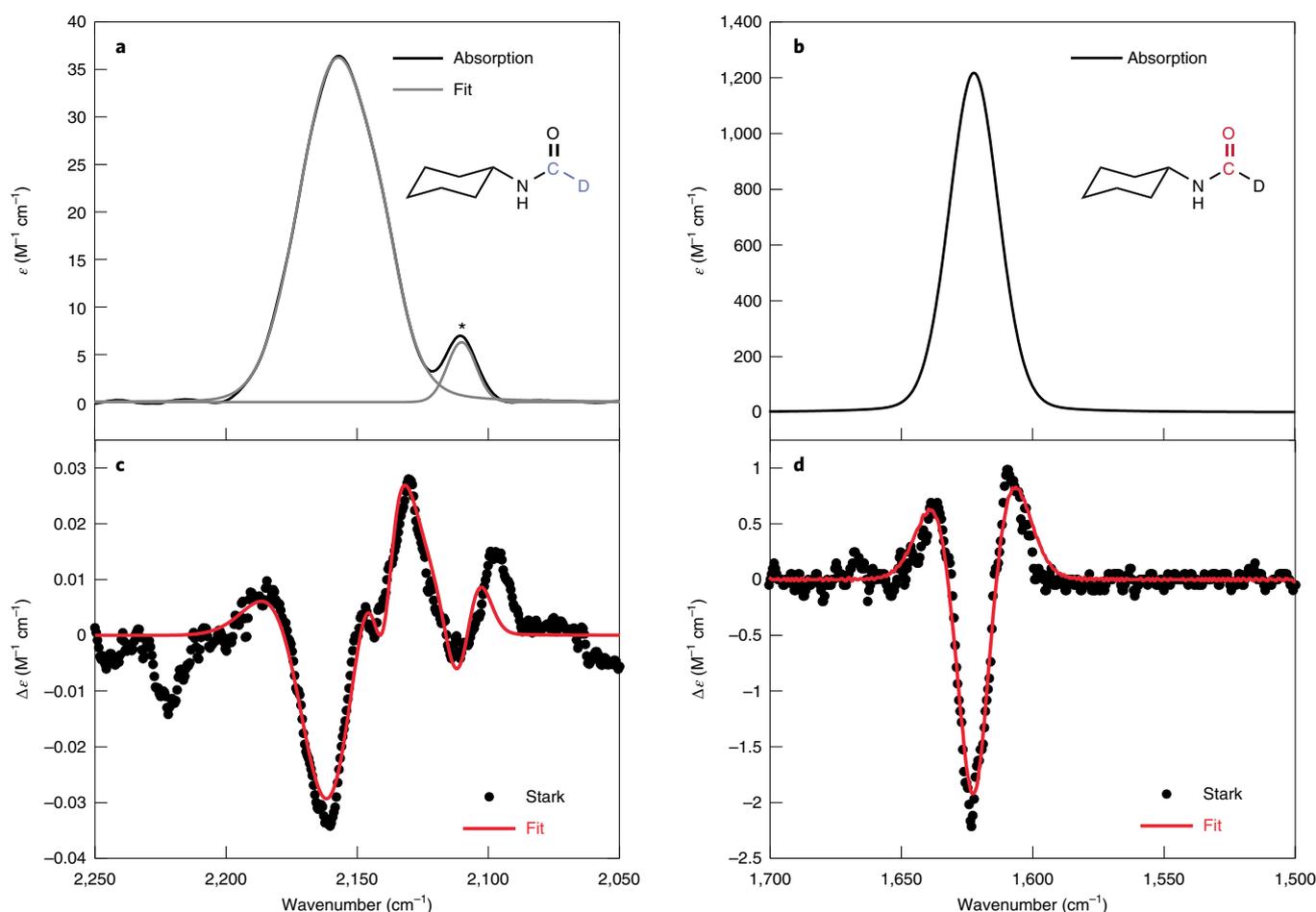
at the active site of LADH, we reveal that the aldehyde moiety, which participates in the enzyme-catalysed reaction, experienced a substantially different electric field orientation. This provides direct evidence for the intrinsic differences between the electrostatic environments of solvents and an enzyme active site experienced by a chemically relevant species.

## Results and discussion

**Experimental validation of two-directional electric field measurements by vibrational Stark spectroscopy.** The infrared absorption spectra for the C–D and C=O stretches in CXF-D are shown in Fig. 2a,b, respectively. Although the absorbance of the C–D stretch is typically ~100 times weaker than that of the C=O (or C≡N) stretch<sup>24</sup>, the single C–D stretch in CXF-D shows a reasonably large extinction coefficient ( $\epsilon_{\max} \approx 3 \text{ M}^{-1} \text{ cm}^{-1}$ , which is about 1/30th that of C=O), which facilitates its application as a VSE probe. To validate the spectroscopic response of C=O and C–D in CXF-D to an external electric field, we used VSS, in which an external electric field of known magnitude ( $\sim 1.0 \text{ MV cm}^{-1}$ ) was applied to the sample in a frozen glass (Fig. 2). The Stark spectrum is dominated by the second-derivative band shape from which the Stark tuning rates,  $|\Delta\mu_{\text{C=O}}|f_{\text{C=O}} = 1.13 \pm 0.01 \text{ cm}^{-1} (\text{MV cm}^{-1})^{-1}$  and  $|\Delta\mu_{\text{C-D}}|f_{\text{C-D}} = 1.32 \pm 0.04 \text{ cm}^{-1} (\text{MV cm}^{-1})^{-1}$ , were obtained (see Supplementary Fig. 1 and Supplementary Table 1 for details of the fitting). The Stark tuning rate we obtained for the aldehyde carbonyl is in good agreement with the reported values for other carbonyl compounds<sup>3,11</sup>, and the values for both C=O and C–D can

be well reproduced ( $1.01$  and  $1.22 \text{ cm}^{-1} (\text{MV}^{-1} \text{ cm}^{-1})$ , respectively) using ‘Stark in silico’ calculations, in which we performed ab initio frequency calculations with an external electric field applied along the bond directions while treating the solvent as a dielectric continuum (Supplementary Method 13, Supplementary Text 3 and Supplementary Table 30). Here  $f_{\text{C=O}}$  and  $f_{\text{C-D}}$  are scalar approximations of the local field factors<sup>25</sup> in the directions of the C=O and C–D bonds, respectively. The local field factor is a tensor that describes the difference between the externally applied electric field and the actual electric field experienced by VSE probes and could be different for the C=O and C–D stretches<sup>10,25</sup> (Supplementary Text 3). Although the sign of the Stark tuning rates cannot be explicitly determined by VSS as the sample is isotropic, for vibrational transitions the Stark tuning rates should be positive in most cases because the excited-state dipole moment is expected to be larger than that of the ground state due to vibrational anharmonicity. The dominance of the second-derivative band shape in the Stark spectra (Fig. 2c,d) indicates that the frequency shifts of both the C=O and C–D stretch modes in response to an external electric field are primarily determined by the linear VSE<sup>26</sup>. Similar VSS results were also found for methyl formate-D, a different compound that also carries a deuterated aldehyde group (Supplementary Fig. 2), which demonstrates the generality of using a deuterated aldehyde to probe electric fields in two directions.

**Opposite frequency shifts of C=O and C–D bonds upon solvation.** Figure 3 shows that  $\nu_{\text{C=O}}$  and  $\nu_{\text{C-D}}$  exhibit frequency shifts in opposite directions upon increasing the solvent polarity, a trend



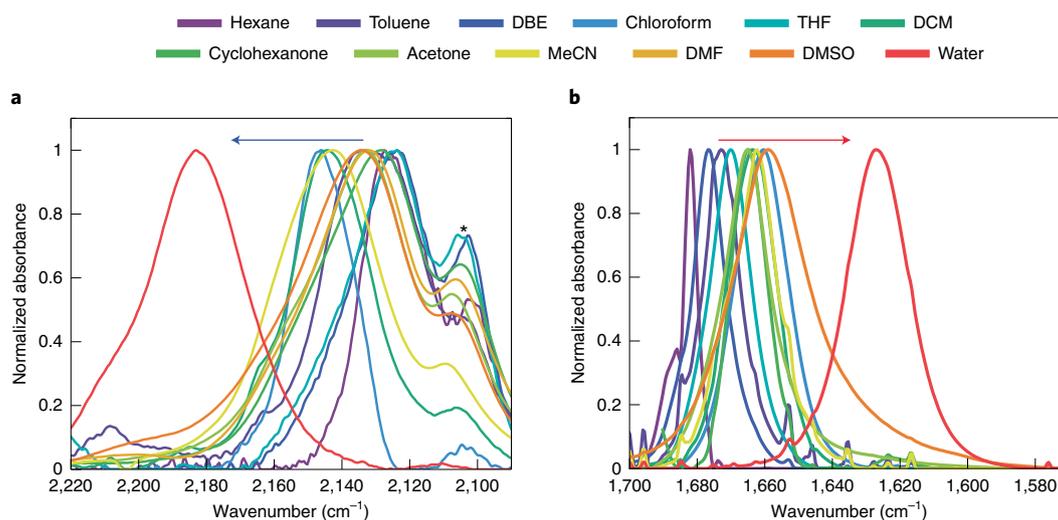
**Fig. 2 | Infrared absorption and Stark spectra of CXF-D at 77 K demonstrating the VSEs of the C–D (left) and C=O (right) bonds. a–d,** Infrared absorption spectra (**a,b**) and the corresponding vibrational Stark spectra scaled to an applied field of 1.0 MV cm<sup>-1</sup> with best fits (red lines) (**c,d**). See Supplementary Fig. 1 and Supplementary Table 1 for the fitting details. The small shoulder peak labelled with an asterisk at around 2,100 cm<sup>-1</sup> in **a** is most likely due to a Fermi resonance (Supplementary Text 5 and Supplementary Table 31). The C–D stretch was measured in 2-methyltetrahydrofuran and the C=O stretch in D<sub>2</sub>O/glycerol-D<sub>3</sub> (v/v 1:1). Different solvents were used because the C=O stretch of CXF-D shows multiple bands in 2-methyltetrahydrofuran, which most likely arise from dimerization due to intermolecular hydrogen bonds.

that is also observed for methyl formate-D (Supplementary Fig. 3). Although it has been widely observed that vibrational probes, such as C=O and C≡N, undergo redshifts with increasing solvent polarity<sup>27</sup>, for this vibrational probe two bonds that arise from the same atom show frequency shifts in opposite directions, where C–D exhibits a blueshift and C=O a redshift. As shown below, this arises because solvent tends to organize around the more polar C=O bond due to attractive electrostatic interactions, resulting in a redshift, which in turn leads to destabilizing interactions sensed by the C–D bond at an about 120° angle from the C=O direction.

**Evaluation of the solvent electric fields on C=O and C–D bonds using molecular simulation.** To unravel the origins of this unusual solvatochromic behaviour, we performed MD simulations and quantum mechanical (QM) electronic structure calculations to quantify the solvent electric fields projected onto the C=O and C–D bonds. Although fixed-charge MD simulations have been successfully used to establish field-frequency maps for C=O in various compounds<sup>3,8,10,11,28</sup>, the map produced for C–D using this method provides a poor correlation between the experimentally observed frequency shifts and the calculated fields (Supplementary Fig. 4). To more accurately quantify the fields experienced by the C–D probe, we therefore adapted two recently developed electronic

structure partitioning schemes<sup>22,23</sup> to obtain more accurate solvent electric fields for configurations obtained from MD sampling. The QM-based approach can significantly improve the accuracy of field calculations because it avoids the use of any parameters assigned by the force field, such as atomic charge or polarizability, but instead obtains the electric field directly from the electrostatic potential created by electrons and nuclei in the system. As shown in Supplementary Tables 6–12, the results obtained from both SPADE<sup>22</sup> (subsystem projected AO decomposition)- and ALMO<sup>23</sup> (absolutely localized molecular orbitals)-based partitioning schemes gave consistent predictions of the solvent electric fields when combined with solvent configurations obtained from either fixed-charge<sup>11</sup> or polarizable force fields<sup>28,29</sup>. This approach thus provides a reliable way to quantify the solvent electric fields projected onto the vibrational probes, which enabled us to uncover the physical origin of the unusual solvatochromic shifts as well as to establish the field-frequency map for CXF-D.

Figure 4a,b shows that for both C–D and C=O, over a wide range of solvent polarities (dielectric constants from 1 to 80), there is a strong linear correlation between the ensemble-averaged solvent electric fields obtained from our simulations and the experimentally measured frequencies. The ensemble-averaged solvent electric field along the C=O bond is negative, which has also been



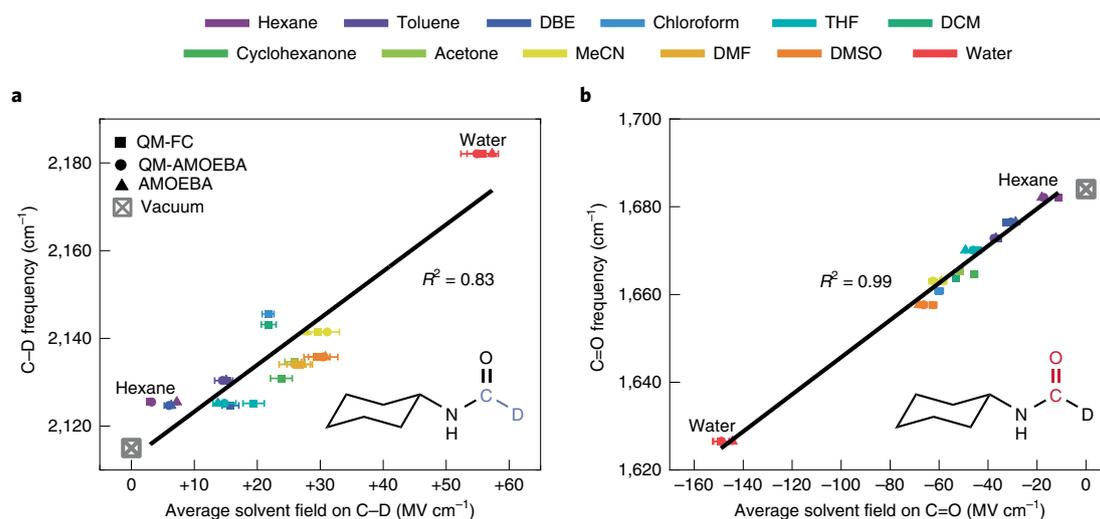
**Fig. 3 | Normalized infrared spectra of CXF-D in a series of solvents with various polarities.** **a, b**, The spectra demonstrate the solvatochromic shifts in the C–D (**a**) and C=O (**b**) frequencies. The concentrations of CXF-D in the measurements for  $\nu_{\text{C-D}}$  and  $\nu_{\text{C=O}}$  were 20 and 1 mM, respectively. The blue and red arrows indicate the direction of the frequency shifts of C–D and C=O, respectively, with the increase in solvent polarity. The small shoulder peak labelled with an asterisk at around 2,100  $\text{cm}^{-1}$  in **a** is most likely due to a Fermi resonance, which is evidenced by its minimal frequency shift with different solvents and yet a more pronounced change in intensity as the C–D stretch band shifts. DBE, dibutyl ether; THF, tetrahydrofuran; DCM, dichloromethane; MeCN, acetonitrile; DMF, dimethylformamide; DMSO, dimethylsulfoxide.

found for many other carbonyl-containing molecules<sup>3,11,28</sup>. The negative electric field along the C=O bond signifies an attractive electrostatic interaction with solvents because the field is aligned with the bond dipole, which lies in the O  $\rightarrow$  C direction. In contrast, the C–D bond in CXF-D has a more subtle dipole moment in which both C and D are positively charged, but C is more so (Supplementary Table 27), which gives a dipole in the D  $\rightarrow$  C direction. The electric field along the C–D bond is positive and thus antiparallel with the bond dipole direction, which indicates that the electrostatic interaction with the solvent destabilizes the C–D covalent bond, consistent with the blueshift observed in its vibrational frequency as the solvent polarity increases. The correlation between the experimental frequency shifts and calculated fields thus provides strong support that the electrostatic interactions are responsible for modulating the vibrational frequencies. We note that the linearity of the correlation for C–D is not as good as that for C=O ( $R^2=0.83$  versus 0.99). This most likely arises due to a Fermi resonance between the C–D stretch and the overtone of C–D bending modes that appears as a shoulder peak at around 2,100  $\text{cm}^{-1}$  (Supplementary Text 5 and Supplementary Table 31), consistent with previous studies on the vibrational spectra of aldehydes<sup>30</sup>. To further validate our computational procedure, we also performed ‘Stark *in silico*’ calculations in vacuum with an external electric field applied along the C=O or C–D direction, from which we observed a strong correlation between the strength of the applied field and the computed C=O and C–D frequencies (Supplementary Fig. 5 and Supplementary Table 14).

The stark contrast in the solvent electric field directions on C=O and C–D manifests how these two bonds, connected to a common atom centre, interact with the solvents. As C=O has a substantially larger bond dipole than C–D (Supplementary Table 27), it is preferentially stabilized by the solvent, via both the local arrangements of the solvent molecules and their electronic polarization (Supplementary Fig. 17). Owing to the  $\sim 120^\circ$  angle between the C=O and C–D bonds, the field induced by the solvent ordering around C=O causes C–D to experience a field that is opposite to its bond dipole direction, which results in its destabilization. To further corroborate this physical picture, we performed simulations of *N*-cyclohexylacetamide (Supplementary Table 15), in which the

deuterium atom in CXF-D is replaced by a methyl group and thus removes any specific interactions that may exist between C–D and the solvent molecules. In this methylated analogue, we observed that the positive field along the C–C(methyl) bond, which is in the same direction as C–D, is largely retained (reduced by  $\sim 20\%$ ) and thus confirms that the positive electric field along C–D is a secondary effect due to the preferential stabilization of C=O. This picture of preferential stabilization of the more polar C=O group provides additional details about the solute–solvent interactions that are not captured by the Onsager reaction field model<sup>31</sup>, one of the most widely used models for solvation based on the concept that globally the molecular dipole of a solute experiences a stabilizing reaction field. Our results show clear evidence that locally, that is, at the level of individual chemical bonds<sup>32</sup>, there are both stabilizing and destabilizing electrostatic interactions experienced by solute molecules in solutions. Given its electrostatic origin, the trend in the opposite frequency shifts of C=O and C–D upon increasing solvent polarity can be qualitatively captured using a simple polarizable continuum model<sup>33</sup> for the solvents (Supplementary Table 16); however, a quantitative description of the solvatochromic effects necessitates considering the specific molecular ordering of the solvent around the probe.

To further elucidate the competition between the C=O and C–D dipoles in organizing the solvent around them, we examined the solvatochromic behaviour of the carbonyl group in acetyl chloride, in which the C–D is replaced by the much more polar carbon–chlorine (C–Cl) bond. In this case, C–Cl can much more effectively compete with C=O for solvent organization, resulting in negative (stabilizing) electric fields along both bonds with the magnitude of that along C=O reduced notably (Supplementary Fig. 12). Remarkably, we found that the C=O of acetyl chloride showed only a 2  $\text{cm}^{-1}$  redshift going from hexane to acetonitrile (acetyl chloride is not stable in water), a much smaller solvent shift than that of CXF-D (19  $\text{cm}^{-1}$ ) or any other carbonyl-containing molecule observed to date (Supplementary Fig. 7 and Supplementary Table 17). Given that the C=O of acetyl chloride has a typical Stark tuning rate as obtained from both VSS experiments and ‘Stark *in silico*’ calculations (Supplementary Fig. 8, Supplementary Table 1 and 18), this extraordinarily small solvation redshift indicates that



**Fig. 4 | Construction of the field-frequency maps for C–D and C=O in the CXF–D probe using experimentally measured frequency shifts and computed electric field projections along these bonds. a, b,** Field-frequency correlations of C–D (**a**) and C=O (**b**) plotted using the calculated solvent electric fields (ensemble average) and the experimentally measured peak frequencies. Electric fields obtained using three methods are shown and used for the least-squares fits, and additional results can be found in Supplementary Tables 6–13. QM-FC refers to the use of density functional theory and the SPADE electronic structure partitioning scheme<sup>22</sup> (Supplementary Methods) to calculate solvent electric fields based on fixed-charge MD trajectories. QM-AMOEBA refers to the use of the same QM method to calculate the solvent electric fields when based on AMOEBA MD trajectories. AMOEBA refers to the use of the native parameters from the AMOEBA force field, such as the atomic permanent multipoles and polarizabilities, to calculate the solvent electric fields. The error bars represent the standard errors of the solvent field distributions. ‘Vacuum’ shows the results of ab initio harmonic frequency calculations of CXF–D in vacuo at the B3LYP/6-31+G(d,p) level (scaled by 0.964<sup>37</sup>), which serves as a cross-validation of the zero-field case and was not included in the fits. The least-squares regression lines based on the results of all three types of electric field calculations and the measured infrared frequencies are  $\nu_{\text{C–D}} = 1.07F_{\text{C–D}}^{\text{solv}} + 2,112.61$  ( $R^2 = 0.83$ ) and  $\nu_{\text{C=O}} = 0.42F_{\text{C=O}}^{\text{solv}} + 1,687.97$  ( $R^2 = 0.99$ ), where  $F_{\text{C–D}}^{\text{solv}}$  and  $F_{\text{C=O}}^{\text{solv}}$  denote the projections of the solvent electric fields along the C–D and C=O directions, respectively.

the C=O of acetyl chloride, unlike that of CXF–D, gains a much lesser degree of electrostatic stabilization as the solvent polarity increases (Supplementary Text 1). The observation that the C–Cl bond suppresses the preferential stabilization of C=O by the solvent in acetyl chloride further underlines the ubiquity of distinct local solvation effects governed by bond dipoles, which may have far-reaching implications in, for example, small-molecule catalyst design.

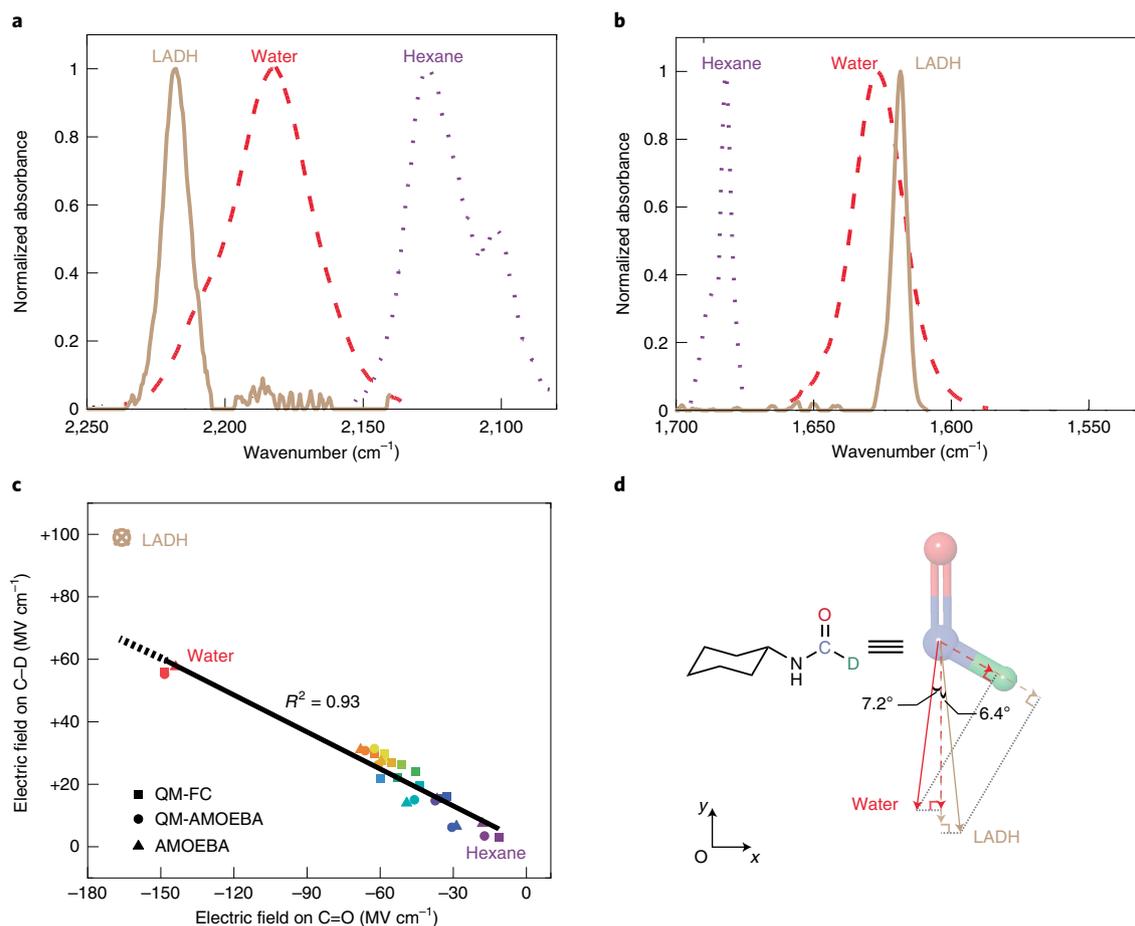
**Probing the active-site electric field of LADH.** LADH catalyses a hydride transfer reaction, which is ubiquitous in many metabolic pathways, for the interconversion between alcohol and aldehyde using NAD<sup>+</sup> or NADH as the redox cofactor<sup>34</sup>. To probe the magnitude and orientation of the electric fields in wild-type LADH, we introduced the CXF–D probe, which serves as an aldehyde-analogue inhibitor, to its active site. As shown in Fig. 5a,b, this leads to red- and blueshifts in the C=O and C–D frequencies, respectively, which are even larger ( $\nu_{\text{C=O}} = 1,618 \text{ cm}^{-1}$  and  $\nu_{\text{C–D}} = 2,218 \text{ cm}^{-1}$ ) than those in water ( $\nu_{\text{C=O}} = 1,627 \text{ cm}^{-1}$  and  $\nu_{\text{C–D}} = 2,182 \text{ cm}^{-1}$ ), the most polar solvent that we investigated (Supplementary Fig. 9 and Supplementary Tables 19 and 20). In addition, the linewidths of both the C=O and C–D spectra in the enzyme were much narrower than those in water, probably due to the smaller range of fluctuations possible in the active-site environment, as has been observed in many enzymes<sup>2</sup>. Based on the field-frequency maps for CXF–D that we obtained using a wide range of solvents (Fig. 4), the measured frequency shifts indicated that the active-site electric field projected on C=O was  $-166 \text{ MV cm}^{-1}$ , whereas that on C–D was  $+99 \text{ MV cm}^{-1}$ , which are both larger in magnitude than the average values observed in any of the solvents. CXF–D’s C=O, similar to that of the substrate in aldehyde reduction reactions, is known to strongly interact with Ser-48 and Zn<sup>2+</sup> at the active site<sup>20,21</sup> (Fig. 1 and Supplementary Fig. 10). Hence these specific interactions are the most likely sources

of the large active-site electric field. In contrast, based on the crystal structure we obtained at 1.43 Å (Supplementary Fig. 10), the C–D in CXF–D does not interact strongly with the nearby residues, which suggests that the main contributions to its large positive electric field are from Ser-48 and Zn<sup>2+</sup>, which preferentially stabilize C=O and destabilize C–D as a secondary effect.

**Comparison of the electric field orientation at the LADH active site to that in the solvents.** As shown in Fig. 5c, there is a strong linear correlation ( $R^2 = 0.93$ ) between the average electric fields projected on C=O and C–D for CXF–D in the solvents. Despite the substantial changes in the electric field magnitudes across the solvents due to their diverse polarities, the ratio between the average solvent electric field projected on C=O ( $F_{\text{C=O}}$ ) and C–D ( $F_{\text{C–D}}$ ) remains essentially unchanged at  $F_{\text{C–D}}/F_{\text{C=O}} = -0.39 \pm 0.02$ . Using this ratio and assuming that the angle between the C=O and C–D groups is fixed at 120° for the *sp*<sup>2</sup> carbon we extracted the orientation of the electric field vector in the aldehyde plane, namely its angle to the bond dipole direction of C=O, labelled  $\theta$  in Fig. 1c. This angle was shown (Supplementary Text 2) to be given by:

$$\theta = \arctan \left[ \frac{1}{\sqrt{3}} \left( 1 + 2 \frac{F_{\text{C–D}}}{F_{\text{C=O}}} \right) \right]. \quad (1)$$

Therefore, the nearly constant value of  $F_{\text{C–D}}/F_{\text{C=O}}$  across all the solvents indicates that the electric field orientations in the aldehyde plane were remarkably similar. This finding is further supported by the results in other polar protic solvents besides water (methanol and ethanol; Supplementary Text 4 and Supplementary Figs. 15 and 16) and implies that, despite their distinct chemical structures and polarities<sup>35</sup>, all these solvents tend to organize in a similar way to electrostatically stabilize the solute (vibrational probe). From the average value of this ratio we obtained an angle  $\theta$  of  $7.2 \pm 1.3^\circ$ , that is,



**Fig. 5 | Comparison of the electric field orientations in solvents and at the LADH active site revealed by two-directional electric field measurements.**

**a,b**, Normalized infrared spectra of C–D (**a**) and C=O (**b**) with CXF–D bound to the active site of wild-type LADH to form a ternary complex with NADH. The spectra of CXF–D in hexane and water are shown for reference. **c**, Correlation between the average electric fields on C–D and C=O in each solvent (colour code as in Figs. 3 and 4) and LADH. The electric field calculations based on QM–FC, QM–AMOEBa and AMOEBa are as defined in Fig. 4. The least-squares fit does not include the point for LADH, and the result is  $F_{C-D} = -0.39F_{C=O} + 1.21$  ( $R^2 = 0.93$ ). **d**, Visualization of the electric fields exerted by water (red arrow) and LADH (light brown arrow) based on the results in **c**. The dashed arrows represent the measured electric fields projected on C=O and C–D, from which the orientations of the total electric field on the two-dimensional plane of the aldehyde moiety were obtained (see equation (1)). The solvents exert electric fields of different magnitudes but nearly the same direction: an angle of  $-7.2 \pm 1.3^\circ$  with respect to the C=O dipole (the error is based on the slope's standard error in the linear fit). In contrast, the active-site electric field of LADH has a larger magnitude and is oriented in a direction that differs from the direction in the solvents by  $-14^\circ$ .

the electric field was almost parallel to the C=O bond dipole and pointed slightly away from C–D (Fig. 5d). This small value of  $\theta$  further supports the physical picture that the solvents preferentially stabilize C=O, and the small deviation from being exactly parallel ( $0^\circ$ ) may be beneficial to stabilize the solute globally by reducing the destabilization of C–D compared with that of a field that is perfectly parallel with the C=O bond dipole.

In contrast, the electric field at the active site of LADH is oriented differently in the aldehyde plane as compared with that in the solvents. Using the  $F_{C=O}$  and  $F_{C-D}$  values for LADH ( $-166$  and  $+99$   $\text{MV cm}^{-1}$ , respectively) and equation (1), the angle  $\theta$  obtained is  $-6.4^\circ$ . This angle corresponds to the electric field vector experienced by the aldehyde group of CXF–D being slightly skewed towards the opposite side of the C=O bond dipole direction, that is, towards C–D, compared with that in the solvents in which it is away from C–D (Fig. 5d). Hence, in the two-dimensional plane of the aldehyde moiety the electric field orientation at the LADH active site is rotated by  $\sim 14^\circ$  relative to that in the solvents, which leads to a field orientation that enhances the positive field projection along the C–D bond and thus further destabilizes it. This provides

experimental evidence that unlike solvents whose organization around the solute is largely directed by the solute bond dipoles to achieve a global electrostatic stabilization, the enzyme active site is preorganized, which tunes the relative stabilization of the aldehyde moiety's C=O and C–D bonds by imposing a distinct electric field on the reactive chemical moiety (the aldehyde group) that is rarely accessible in solvents in terms of both magnitude and orientation.

## Conclusions

In summary, we have introduced an approach based on the VSE to measure both the magnitude and orientation of the electric fields in condensed phase systems. By exploiting a two-directional vibrational probe, CXF–D, we showed that its deuterated aldehyde group can be calibrated using experiments and a computational framework based on MD and electronic structure calculations to report the electric field strengths along both the C=O and C–D directions. Using this probe, we demonstrated the coexistence of stabilizing and destabilizing electrostatic interactions experienced by the individual chemical bonds of a solute in a wide range of solvents, as well as the nearly invariant field orientation across solvents of

distinct chemical structures and polarities. This provides direct experimental insights into how local solvent organization around a solute molecule driven by the difference in bond dipoles affords global electrostatic stabilization. In contrast, we showed that the electric field at the active site of LADH, as measured by the CXF-D probe, possesses a markedly different orientation that provides a greater electrostatic destabilization of the C–D bond than that observed in the solvents. Hence, although the electric fields at an enzyme's active site and in aqueous solutions may be of comparable magnitudes, the orientations of these fields are intrinsically different. This supports the hypothesis that enzymes may provide preorganized electrostatic environments that maximize their catalytic power<sup>1,2</sup>. Finally, the experimental and computational framework that we introduce here to probe electric field orientations using deuterated aldehydes can be widely applied to other systems in complex environments and with complex solutes, including homogeneous catalysts, and thus serves as a prototype for the characterization of electric field orientations to stimulate future investigations on the functional role of field orientation in enzymatic and electrostatic catalysis.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41557-022-00937-w>.

Received: 8 September 2021; Accepted: 25 March 2022;

Published online: 05 May 2022

### References

1. Warshel, A. et al. Electrostatic basis for enzyme catalysis. *Chem. Rev.* **106**, 3210–3235 (2006).
2. Fried, S. D. & Boxer, S. G. Electric fields and enzyme catalysis. *Annu. Rev. Biochem.* **86**, 387–415 (2017).
3. Schneider, S. H. & Boxer, S. G. Vibrational Stark effects of carbonyl probes applied to reinterpret IR and Raman data for enzyme inhibitors in terms of electric fields at the active site. *J. Phys. Chem. B* **120**, 9672–9684 (2016).
4. Carey, P. R. Spectroscopic characterization of distortion in enzyme complexes. *Chem. Rev.* **106**, 3043–3054 (2006).
5. Dong, J. et al. The strength of dehalogenase–substrate hydrogen bonding correlates with the rate of Meisenheimer intermediate formation. *Biochemistry* **42**, 9482–9490 (2003).
6. Tonge, P. J. & Carey, P. R. Length of the acyl carbonyl bond in acyl-serine proteases correlates with reactivity. *Biochemistry* **29**, 10723–10727 (1990).
7. Deng, H. et al. Source of catalysis in the lactate dehydrogenase system. Ground-state interactions in the enzyme–substrate complex. *Biochemistry* **33**, 2297–2305 (1994).
8. Fried, S. D., Bagchi, S. & Boxer, S. G. Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **346**, 1510–1514 (2014).
9. 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.* **138**, 11890–11895 (2016).
10. Fried, S. D. & Boxer, S. G. Measuring electric fields and noncovalent interactions using the vibrational Stark effect. *Acc. Chem. Res.* **48**, 998–1006 (2015).
11. 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.* **135**, 11181–11192 (2013).
12. Wu, Y., Fried, S. D. & Boxer, S. G. A preorganized electric field leads to minimal geometrical reorientation in the catalytic reaction of ketosteroid isomerase. *J. Am. Chem. Soc.* **142**, 9993–9998 (2020).
13. Welborn, V. V. & Head-Gordon, T. Computational design of synthetic enzymes. *Chem. Rev.* **119**, 6613–6630 (2019).
14. Welborn, V. V., Pestana, L. R. & Head-Gordon, T. Computational optimization of electric fields for better catalysis design. *Nat. Catal.* **1**, 649–655 (2018).
15. Fuxreiter, M. & Mones, L. The role of reorganization energy in rational enzyme design. *Curr. Opin. Chem. Biol.* **21**, 34–41 (2014).
16. Kamerlin, S. C., Sharma, P. K., Chu, Z. T. & Warshel, A. Ketosteroid isomerase provides further support for the idea that enzymes work by electrostatic preorganization. *Proc. Natl Acad. Sci. USA* **107**, 4075–4080 (2010).
17. Bim, D. & Alexandrova, A. N. Electrostatic regulation of blue copper sites. *Chem. Sci.* **12**, 11406–11413 (2021).
18. Shaik, S., Mandal, D. & Ramanan, R. Oriented electric fields as future smart reagents in chemistry. *Nat. Chem.* **8**, 1091–1098 (2016).
19. Xu, L., Izgorodina, E. I. & Coote, M. L. Ordered solvents and ionic liquids can be harnessed for electrostatic catalysis. *J. Am. Chem. Soc.* **142**, 12826–12833 (2020).
20. Deng, H., Schindler, J. F., Berst, K. B., Plapp, B. V. & Callender, R. A Raman spectroscopic characterization of bonding in the complex of horse liver alcohol dehydrogenase with NADH and N-cyclohexylformamide. *Biochemistry* **37**, 14267–14278 (1998).
21. Ramaswamy, S., Scholze, M. & Plapp, B. V. Binding of formamides to liver alcohol dehydrogenase. *Biochemistry* **36**, 3522–3527 (1997).
22. Claudino, D. & Mayhall, N. J. Automatic partition of orbital spaces based on singular value decomposition in the context of embedding theories. *J. Chem. Theory Comput.* **15**, 1053–1064 (2019).
23. Khaliullin, R. Z., Head-Gordon, M. & Bell, A. T. An efficient self-consistent field method for large systems of weakly interacting components. *J. Chem. Phys.* **124**, 204105 (2006).
24. Adhikari, R., Zimmermann, J. & Romesberg, F. E. Transparent window vibrational probes for the characterization of proteins with high structural and temporal resolution. *Chem. Rev.* **117**, 1927–1969 (2017).
25. Bublitz, G. U. & Boxer, S. G. Stark spectroscopy: applications in chemistry, biology, and materials science. *Annu. Rev. Phys. Chem.* **48**, 213–242 (1997).
26. Boxer, S. G. Stark realities. *J. Phys. Chem. B* **113**, 2972–2983 (2009).
27. Baiz, C. R. et al. Vibrational spectroscopic map, vibrational spectroscopy, and intermolecular interaction. *Chem. Rev.* **120**, 7152–7218 (2020).
28. Kozuch, J. et al. Testing the limitations of MD-based local electric fields using the vibrational Stark effect in solution: penicillin G as a test case. *J. Phys. Chem. B* **125**, 4415–4427 (2021).
29. Fried, S. D., Wang, L. P., Boxer, S. G., Ren, P. & Pande, V. S. Calculations of the electric fields in liquid solutions. *J. Phys. Chem. B* **117**, 16236–16248 (2013).
30. Eggers, D. F. & Lingren, W. E. C–H vibrations in aldehydes. *Anal. Chem.* **28**, 1328–1329 (1956).
31. Onsager, L. Electric moments of molecules in liquids. *J. Am. Chem. Soc.* **58**, 1486–1493 (1936).
32. Levinson, N. M., Fried, S. D. & Boxer, S. G. Solvent-induced infrared frequency shifts in aromatic nitriles are quantitatively described by the vibrational Stark effect. *J. Phys. Chem. B* **116**, 10470–10476 (2012).
33. Tomasi, J., Mennucci, B. & Cammi, R. Quantum mechanical continuum solvation models. *Chem. Rev.* **105**, 2999–3093 (2005).
34. Sekhar, V. C. & Plapp, B. V. Rate constants for a mechanism including intermediates in the interconversion of ternary complexes by horse liver alcohol dehydrogenase. *Biochemistry* **29**, 4289–4295 (1990).
35. Reichardt, C. & Welton, T. *Solvents and Solvent Effects in Organic Chemistry* (John Wiley & Sons, 2011).
36. Epifanovsky, E. et al. Software for the frontiers of quantum chemistry: an overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **155**, 084801 (2021).
37. Johnson, R. D. III *NIST Computational Chemistry Comparison and Benchmark Database*. NIST Standard Reference Database Number 101, Release 21 August 2020 (NIST, accessed 2 January 2021).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

### Data availability

The gene sequence for the wild-type enzyme in this study has been deposited in GenBank (accession code [OM863576](https://doi.org/10.6084/m9.figshare.19248108)). The X-ray coordinates and structural factors have been deposited in the Protein Data Bank as entry 7RM6 for the wild-type enzyme in complex with NADH and CXF. The electronic structure partitioning methods for DFT calculations of solvent electric fields at atomic positions were implemented in the Q-Chem software package<sup>36</sup>, which is available in the 5.4.2 and later releases. Source data are provided with this paper. All the data that support the finding of this study are available within this article and its Supplementary Information and available on Figshare (<https://doi.org/10.6084/m9.figshare.19248108>).

### Code availability

The code used for processing MD trajectories and calculating electric field projections along bond directions is available at [https://github.com/YuezhiMao/2D\\_VSE\\_probe](https://github.com/YuezhiMao/2D_VSE_probe) and Zenodo (<https://zenodo.org/record/6300009#.Yhw0GBPMKS4>)

### Acknowledgements

We express our gratitude to B. V. Plapp at the University of Iowa for providing detailed guidance on the expression and purification of LADH and much other valuable advice. We thank C.-Y. Lin for his help with X-ray crystallography, S. Schneider, J. Weaver, J. Kirsh and S. D. E. Fried for helpful discussions, T. Carver at the Stanford Nano Shared Facilities for nickel coating the Stark windows, T. McLaughlin from Stanford University Mass Spectrometry (SUMS) for technical support with the mass spectrometry. C.Z. is grateful for the Stanford Center for Molecular Analysis and Design (CMAD) Fellowship. J.K. acknowledges support from the German Research Foundation DFG-Project-ID221545957-SFB1078/B9. This work was supported in part by NIH Grant GM118044 (to S.G.B.) and the National Science Foundation grant no. CHE-1652960 (to T.E.M.). T.E.M. also acknowledges support from the Camille Dreyfus Teacher-Scholar Awards Program. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, is supported by the US Department of Energy, Office of Science, Office of Basic Energy Sciences under contract no. DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office

of Biological and Environmental Research, and by the National Institutes of Health, National Institute of General Medical Sciences (P30GM133894). The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of NIGMS or NIH. This research also used resources of the National Energy Research Scientific Computing Center (NERSC), a US Department of Energy Office of Science User Facility operated under contract no. DE-AC02-05CH11231, and the Sherlock cluster operated by the Stanford Research Computing Center.

### Author contributions

C.Z. and S.G.B. designed the research. C.Z. performed the experiments and fixed-charge MD simulations. Y.M. developed the computational protocol that combined MD simulations and QM calculations to quantify the solvent electric field contributions, implemented the electronic structure partitioning schemes in the Q-Chem software package and performed the QM calculations. J.K. performed the AMOEBA polarizable MD simulations. A.O.A. analysed the data from the MD simulations and wrote the codes to extract the truncated solute-solvent structures for the QM calculations. Z.J. synthesized *N*-[formyl-<sup>2</sup>H]cyclohexylformamide. C.Z., Y.M., T.E.M. and S.G.B. discussed the results and wrote the manuscript. All the authors contributed to improving the manuscript.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41557-022-00937-w>.

**Correspondence and requests for materials** should be addressed to Thomas E. Markland or Steven G. Boxer.

**Peer review information** *Nature Chemistry* thanks Steven Corcelli, Jahan Dawlaty, Feng Gai, Valerie Welborn for their contribution to the peer review of this work.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

## Supplementary information

---

# **A two-directional vibrational probe reveals different electric field orientations in solution and an enzyme active site**

---

In the format provided by the authors and unedited

# Supplementary Information

## A two-directional vibrational probe reveals different electric field orientations in solution and an enzyme active site

Chu Zheng<sup>1,†</sup>, Yuezhi Mao<sup>1,†</sup>, Jacek Kozuch<sup>2</sup>, Austin Atsango<sup>1</sup>, Zhe Ji<sup>1</sup>, Thomas E. Markland<sup>1,\*</sup>,  
and Steven G. Boxer<sup>1,\*</sup>

1. Department of Chemistry, Stanford University, Stanford, CA 94305, USA
2. Experimental Molecular Biophysics, Department of Physics, Freie Universität Berlin, 14195 Berlin, Germany

†These authors contributed equally to this work.

\*Correspondence to: [tmarkland@stanford.edu](mailto:tmarkland@stanford.edu); [sboxer@stanford.edu](mailto:sboxer@stanford.edu)

### Table of Contents

<b>Materials and Methods</b> .....	<b>1</b>
<b>Supplementary Texts</b> .....	<b>19</b>
<b>Supplementary Figures</b> .....	<b>29</b>
<b>Supplementary Tables</b> .....	<b>41</b>
<b>References</b> .....	<b>57</b>

## Materials and Methods

Chemicals that have been used in this study are from the following sources without further purification unless mentioned otherwise: hexane (Sigma-Aldrich, 99%), toluene (Sigma-Aldrich, 99.8%), dibutyl ether (Sigma-Aldrich, 99%), chloroform (Sigma-Aldrich, 99.8%), tetrahydrofuran (Sigma-Aldrich, 99.8%), dichloromethane (Sigma-Aldrich, 99.8%), cyclohexanone (Sigma-Aldrich, 99.8%), acetone (Sigma-Aldrich, 99.8%), acetonitrile (Sigma-Aldrich, 99.8%), dimethyl formamide (Sigma-Aldrich, 99.8%), dimethyl sulfoxide (Sigma-Aldrich, 99.8%), methyl formate-D (Sigma-Aldrich, 99 atom % D), Cyclohexylamine (99%), Sodium

1 Formate-<sup>13</sup>C (99% atom % <sup>13</sup>C), 2-methyl tetrahydrofuran (Sigma-Aldrich, 99%), Glycer(ol-D<sub>3</sub>)  
2 (Sigma-Aldrich, 99 atom % D), deuterium oxide (ACROS, 100.0 atom % D), acetyl chloride (Sigma-  
3 Aldrich, 98%), Formic acid (formyl-D) (Cambridge Isotope Laboratories, 98 atom % D), and 1,1'-  
4 carbonyldiimidazole (Sigma-Aldrich).

### 5 **1. Synthesis of N-[formyl-<sup>13</sup>C]-Cyclohexylformamide (<sup>13</sup>C-CXF)**

6 The synthesis of <sup>13</sup>C-CXF was based on a previously reported method.<sup>1</sup> 1.38 mL  
7 concentrated HCl was added into a round-bottom flask that contained sodium [<sup>13</sup>C]-formate (1.0  
8 g) in toluene (15 mL) and refluxed for 1 hour. 2 mL of cyclohexylamine was added into the flask,  
9 followed by the installation of a Dean-Stark trap. The reaction mixture was refluxed, during which  
10 the Dean-Stark trap pushed the reaction equilibrium toward the product side by removing water.  
11 TLC was used to monitor the reaction using potassium permanganate as the staining reagent,  
12 and the final product was purified by silica chromatography (eluted with ethyl acetate/hexane  
13 3:2) and dried through rotary evaporation and vacuum. The final product remained as a colorless  
14 oil without further drying. <sup>1</sup>H-NMR (400 MHz, CD<sub>3</sub>Cl): δ 8.06 (d, *J* = 192 Hz, 1H), 5.81 (broad s,  
15 0.25H, conformer 1), 5.60 (broad s, 0.75H, conformer 2), 3.82 (m, 0.75H, conformer 2), 3.27 (m,  
16 0.25H, conformer 1), 1.92-1.09 (m, 10H); <sup>13</sup>C-NMR (100 MHz, CD<sub>3</sub>Cl): δ 163.5 (conformer 1), 160.3  
17 (conformer 2), 47.1, 34.6 (conformer 1), 32.9 (conformer 2), 25.4 (conformer 2), 25.0 (conformer  
18 1), 24.7. *m/z* (EI-MS): 129.0 [M+H]<sup>+</sup>.

### 19 **2. Synthesis of N-[formyl-<sup>2</sup>H]-Cyclohexylformamide (CXF-D)**

20 CXF-D was synthesized using an alternative method. To a solution of 1,1'-  
21 carbonyldiimidazole (2 g) in tetrahydrofuran (30 mL) was added a solution of formic acid (formyl-  
22 D, 600 mg) in tetrahydrofuran (10 mL). The mixture was stirred at room temperature for 1 hour

1 under nitrogen before the addition of a solution of cyclohexylamine (1.26 g) in tetrahydrofuran  
2 (10 mL). The mixture was stirred at room temperature for another 3 hours. The resulting mixture  
3 was concentrated and purified by silica chromatography (eluted with  
4 methanol/dichloromethane 1:20, visualized by potassium permanganate). The final product was  
5 dried through rotary evaporation and vacuum as a colorless oil in 41% yield. <sup>1</sup>H-NMR (400 MHz,  
6 CD<sub>3</sub>Cl): δ 8.05 (s, 0.01H), 5.99 (broad s, 0.25H, conformer 1), 5.86 (broad s, 0.75H, conformer 2),  
7 3.82 (m, 0.75H, conformer 2), 3.28 (m, 0.25H, conformer 1), 1.94-1.07 (m, 10H); <sup>13</sup>C-NMR (100  
8 MHz, CD<sub>3</sub>Cl): δ 47.1, 34.6 (conformer 1), 32.9 (conformer 2), 25.4 (conformer 2), 25.0 (conformer  
9 1), 24.7. *m/z* (EI-MS): 129.1 [M+H]<sup>+</sup>.

### 10 **3. Fourier Transform Infrared (FTIR) spectroscopy.**

11 FTIR spectroscopy was performed on a Bruker Vertex 70 spectrometer with a liquid  
12 nitrogen-cooled mercury cadmium telluride (MCT) detector. The sample chamber was constantly  
13 purged with dry air to remove atmospheric water. Data collection and processing were  
14 performed using the spectroscopy software OPUS 5.0.

### 15 **4. Vibrational Stark Spectroscopy (VSS).**

16 The compound of interest was dissolved in a glass-forming solvent, such as 2-methyl  
17 tetrahydrofuran (2MeTHF) or a mixture of D<sub>2</sub>O and glycer(ol-D<sub>3</sub>) (v:v = 1:1). 10 μL sample was  
18 added into the assembled Stark cell, followed by immediately plunging into a custom-built  
19 cryostat that was filled with liquid nitrogen. The Stark cell is assembled using two Calcium  
20 Fluoride (CaF<sub>2</sub>) optical windows (1 mm thick, 12.7 mm diameter, FOCtek Photonics) that were  
21 coated with nickel metal on one side (4.5 nm thick) and separated by Teflon spacers (26 μm thick).  
22 The Stark spectra were recorded by a Bruker Vertex 70 spectrometer at 1 cm<sup>-1</sup> resolution with 64

1 scans in the presence of an applied electric field between 0.5 - 1.7 MV/cm. The external electric  
2 field was applied via a Trek 10/10 high-voltage power amplifier. The linear Stark tuning rates were  
3 obtained as previously described<sup>2,3</sup> from the contribution of the second-derivative component to  
4 the Stark spectrum (see Supplementary Table 1). The experimentally determined Stark tuning  
5 rate is reported as a product of the local field factor  $f$ , which reflects the difference between the  
6 applied electric field and the actual local electric field that is experienced by the vibrational probe,  
7 with the experimentally extracted value of  $|\Delta\mu|$ .<sup>4</sup> More discussion on the local field factor is in  
8 Supplementary text 3.

## 9 **5. Vibrational solvatochromism.**

10 As described previously,<sup>3,5</sup> a demountable IR cell was assembled with two CaF<sub>2</sub> optical  
11 windows (19.05 mm diameter, 3 mm thick, Lambda Research Optics Inc.), which were separated  
12 by two Teflon spacers (25  $\mu\text{m}$  and 50  $\mu\text{m}$  thick). The sample was prepared by dissolving the  
13 compound of interest in various solvents at low concentration which typically ranged from 1 – 20  
14 mM. 20  $\mu\text{L}$  sample was added into the IR cell, which was then tightly capped and wrapped with  
15 parafilm to minimize solvent evaporation and water absorption. The IR spectra were collected in  
16 a wavenumber window that spans 4000  $\text{cm}^{-1}$  – 1000  $\text{cm}^{-1}$  with 256-1024 scans at 1  $\text{cm}^{-1}$  resolution.

## 17 **6. Isotope-edited FTIR spectroscopy for liver alcohol dehydrogenase (LADH).**

18 To measure the C=O frequency of CXF-D bound to LADH, the <sup>13</sup>C isotopically labeled  
19 compound, N-[<sup>13</sup>C-formyl] cyclohexyl formamide (<sup>13</sup>C-CXF), was used as the background (see  
20 Supplementary Fig. 9). LADH was concentrated to ca. 1.6 mM with a total volume of ca. 60  $\mu\text{L}$ ,  
21 and then buffer exchanged into a D<sub>2</sub>O buffer solution (66 mM sodium phosphate, 100mM

1 potassium chloride at pD 8.2). 2  $\mu\text{L}$  NADH solution (50 mM) was added into the protein sample,  
2 which was then split into two portions of ca. 25 $\mu\text{L}$ . 1  $\mu\text{L}$  CXF-D (40 mM) and  $^{13}\text{C}$ -CXF (40 mM)  
3 were added into each portion respectively. Similar to the solvatochromism experiments, 20  $\mu\text{L}$   
4 protein sample was injected into the assembled IR cell, and the IR spectra (4000  $\text{cm}^{-1}$  – 1000  $\text{cm}^{-1}$ )  
5 were collected with 512 scans at 1  $\text{cm}^{-1}$  resolution. Nine spectra were collected from each  
6 sample for data analysis. The IR spectra of CXF-D,  $^{13}\text{C}$ -CXF, and  $\text{D}_2\text{O}$  buffer were also collected for  
7 subtracting the IR signal of the unbound CXF-D from the protein spectra. The procedure for data  
8 processing has been described in detail in previous work.<sup>6,7</sup> Briefly, the raw data in the form of  
9 transmission spectra were converted to absorption spectra. Then difference FTIR spectra were  
10 generated by subtracting the protein spectra with  $^{13}\text{C}$ -CXF bound from that of CXF-D. A scaling  
11 factor was typically used to compensate for the difference of the background signal between two  
12 spectra that resulted from slight differences of the pathlength for each sample.<sup>1,7</sup> In most cases,  
13 the scaling factor involved in this work was very close to unity (usually 0.95-1.05), indicating that  
14 the IR cell assemblies of isotope-labeled samples were highly consistent. The resultant difference  
15 FTIR spectra was minimally baselined, followed by removing the peak of the unbound inhibitor.  
16 All the nine spectra were processed consistently as described above and analyzed using the built-  
17 in programs PeakPick and CurveFit in OPUS 5.0 to obtain the peak position and FWHM. The  
18 results shown in Supplementary Tables 19-20 are the mean values over those nine spectra with  
19 the standard deviations as the corresponding error bar. These nine spectra were combined and  
20 averaged leading to the final spectral results shown in Supplementary Fig. 9.

21 To measure the C-D frequency of CXF-D in LADH, the protein sample was concentrated to  
22 ca. 6 mM because of the weak C-D absorption. The other steps are similar to what was described

1 above except that the buffer solution was made of normal water (H<sub>2</sub>O) instead of heavy water  
2 (D<sub>2</sub>O).

### 3 **7. Plasmid construction and protein sequence**

4 The gene sequence<sup>8</sup> of wild-type LADH was inserted into a recombinant DNA expression  
5 plasmid (pET-15b, Genscript Inc.) along with a strep tag and a cleavage site of *Tobacco etch virus*  
6 (TEV) protease. The gene sequence and protein sequence are shown below.

#### 7 **Gene sequence**

8 atggcgagctggtctcatccgcagtttgaaaaaggcgcggaaaatctgtattttcagggcgcgggtagcacagcaggaaaagtaataaa  
9 atgcaaagcggctgtgctgtgggaggaaaagaaccattttccatcagaggaggtggagggtgcacccccgaaggccatgaagtcgta  
10 taaagatggtggccacaggaatttgcctcagatgaccacgtggttagtggaaacccttgcacacctcttctgtgatcgcaggccatga  
11 ggcagcgggcattgtggagagcattggagaaggcgtcactacagtaagaccaggtgataaagtcatcccactcttactccccagtggtg  
12 aaaatgcagggtttgaagcacctgaaggcaacttctgcttgaaaaatgatctgagcatgcctcggggaacatgcaggatggtacca  
13 gcaggttcacctgcagagggaagcccatccaccacttcttggcaccagcaccttctccagtagaccggtgggacgagatctcagtg  
14 ccaagatcgatgcggcctcaccgctggagaaagtctgtctcattggctgtggattttctactgggtatgggtctgcagcaaggttccaag  
15 gtcaccagggtccacctgtgccgtttggccttggaggagtgggctgtctgttatcatgggctgtaaagcagccggagcggccagg  
16 atcattgggtggacatcaaaaagacaagtttgcaaaggccaagaagtgggtgccactgagtgtgtaaccctcaggactacaaga  
17 aacctatccaggaggtgctgacagaaatgagcaatggaggtgtggattttccttgaagtattggctcgacactatggtgactgc  
18 cttgtcatgctgtcaagaagcatatggtgtgagcgtcattgtgggagtacctcctgattccaaaatctctatgaatcctatggtgctact  
19 gaggtagctacctggaaaggagctattttggcggtttaaagagtaaagattctgtcccaaacttggccgattttatggctaaaaag  
20 ttgactggatccttaatacaccatgtttacctttgaaaaataaatgaaggatttgacctgctctggagagagtatccgtacc  
21 atcctgacgttttga

22

#### 23 **Protein sequence of LADH with a strep tag (GREEN) and TEV cleavage site (YELLOW)<sup>a</sup>**

24 MASWSHPQFEKGAENLYFQ|GAGSTAGKVIKCKAAVLWEEKPFSIEEVEVAPPKAHEVRIKMOVATGICRSD  
25 DHVVSGLTVPLPVIAGHEAAGIVESIGEGVTTVRPGDKVIPLFTPQCCKRVCCKHPEGNFCLKNDLSMPRGT  
26 MQDGTSRFTCRGKPIHHFLGTSTFSQYTVVDEISVAKIDAASPLEKVCLIGCFSTGYGSAVKVAKVTQGSTCA  
27 VFGLGGVGLSVIMGCKAAGAARIIGVDINKDKFAKAKEVGATECVNPQDYKKPIQEVLTMSNGGVDFSFEV  
28 IGRDLMVTALSCQEAAYGVSIVGVPPDSQNLSMNPMLLLSGRTWKGAIFGGFKSKDSVPKLVADFMMAK  
29 FALDPLITHVLPFEKINEGFDLLRSGESIRTLTF

30 <sup>a</sup>"|" indicates the cleavage position.

31

32

33

## 1 **8. Protein expression and purification**

2 Wild-type LADH was expressed using the *Escherichia coli* cell line [BL21(DE3), New  
3 England BioLabs Inc.] in Lysogeny broth (LB) medium (Miller) in the presence of ampicillin (0.1  
4 g/L). Typically, a 5 mL starter culture was prepared overnight and then added to 1-L autoclaved  
5 LB medium (25 g LB in 1 L tap water; tap water is essential to produce good yields, we thank  
6 Professor Plapp for this information), which was incubated in a shaker at 37 °C and 200 rpm.  
7 When the O.D.<sub>600</sub> had reached 0.6-0.8, protein over-expression was induced with 0.2 mM  
8 Isopropyl β-D-1-thiogalactopyranoside (IPTG) and the cell culture was grown overnight (16-18  
9 hours) at 23 °C. Cell pellets were collected by centrifugation at 6000 rpm for 30 minutes at 4 °C  
10 and then resuspended in 100mM Tris-HCl buffer at pH 8.0 containing 150 mM NaCl and 10% (v/v)  
11 glycerol. Cells were lysed using a mechanical homogenizer (Avestin EmulsiFlex-C3), and the  
12 lysates were centrifuged at 15000 rpm for 90 minutes at 4 °C. The resultant supernatants were  
13 further filtered (0.22 μm) and then loaded onto the strep column (StrepTrap HP, Catalog No.  
14 28907547, GE Healthcare). After the strep-tag purification, the protein was treated with TEV  
15 protease (Gene and Cell Technologies Inc.) to cleave the strep tag for 16 hours at 4 °C. The tag-  
16 free protein was then purified by anion exchange column (HiTrap Q HP, Catalog No. 17115401,  
17 GE Healthcare) at pH 9.5 (20 mM piperazine). The identity and purity of protein were  
18 characterized by electrospray ionization mass spectrometry (Waters 2795 HPLC with ZQ single  
19 quadrupole MS in Stanford University Mass Spectrometry facility).

20

21

22

## 1 **9. X-Ray Crystallography of LADH.**

2 The crystallization of LADH is based on a modified hanging-drop method as previously  
3 described.<sup>9</sup> The protein solution for growing crystals contains 25 mg/mL LADH, 4 mM NADH and  
4 20 mM N-cyclohexylformamide. The initial drop consists of 1  $\mu$ L protein sample in 1  $\mu$ L 50mM  
5 Tris-HCl buffer (pH 7.80 at room temperature) from the reservoir solution that contains varying  
6 amounts of PEG 400 (10% -28%). Crystals are grown at 4 °C and normally harvested within a week.  
7 To prepare samples for X-ray crystallography, a single crystal was looped and dipped into the  
8 cryoprotectant (mother liquor with 30% PEG400) before flash cooling in liquid nitrogen. The X-  
9 ray diffraction data were collected at 100 K at the Stanford Synchrotron Radiation Lightsource  
10 (Menlo Park, CA) at beamline BL12-2. The structures were solved using the Molecular  
11 Replacement Module from PHENIX, where the initial model was a published 1.14 Å structure of  
12 wild-type LADH in complex with NAD<sup>+</sup> and 2,3,4,5,6-pentafluorobenzyl alcohol (PDB: 4DWV).  
13 Several rounds of manual and automatic structure refinement were performed using COOT and  
14 PHENIX to ensure the data fitting converges according to some standard metrics. The data  
15 collection and refinement statistics are shown in Supplementary Table 29.

## 16 **10. Fixed-Charge (FC) MD simulations of aldehydes in solvents.**

17 The molecule of interest was constructed in Gaussview 6.0, and the geometry was  
18 optimized by Density Functional Theory (DFT) in Gaussian 16<sup>10</sup> at the B3LYP/6-311++G(2d,2p)  
19 level. The resultant structure was then used for parameterization through the Antechamber  
20 program of AmberTools16<sup>11</sup> based on the GAFF force field. The force field parameters for organic  
21 solvents were taken from Coleman et al.,<sup>12</sup> and water was parameterized using the TIP3P model.  
22 All simulations were performed by GROMACS 2018<sup>13</sup> based on a previously established protocol.<sup>3</sup>

1 A single solute molecule was placed in the center of a cubic solvation box that has a size of 4×4×4  
2 nm<sup>3</sup> filled with solvent molecules. Under periodic boundary conditions, the system was first  
3 energy minimized until the maximum force is less than 1000 kJ/mol/nm, followed by NVT and  
4 NPT equilibration runs (100 ps for each). MD production runs were performed over 1 ns with 2  
5 fs steps, with a van der Waals cutoff of 10 Å (with analytical VdW correction), an electrostatics  
6 cutoff of 10 Å (with particle mesh Ewald method), an SD (leap-frog stochastic dynamics)  
7 integrator, the Berendsen thermostat<sup>14</sup> and the Parrinello-Rahman barostat.<sup>15</sup> The electric field  
8 vectors on the relevant atoms (C, D, O) were obtained via dividing the electrostatic forces acting  
9 on these atoms by the corresponding atomic charges as described in previous work<sup>3,16</sup>. To extract  
10 the solvent contributions to the electric fields, the MD frames obtained were recalculated with  
11 atomic charges on all solvent molecules zeroed out, providing the field contributions from the  
12 solute atoms themselves which were then subtracted from the electric field values obtained in  
13 the original production run. The electric field projections along the C=O and C-D directions are  
14 then evaluated using the following equation:

$$15 \quad F_{CX} = \frac{1}{2} (\vec{F}_C \cdot \hat{r}_{CX} + \vec{F}_X \cdot \hat{r}_{CX})$$

16 where  $F_{CX}$  denotes the field projection along the C=O/C-D bond,  $\vec{F}_C$  and  $\vec{F}_X$  are the electric fields  
17 acting on the C and O/D atoms, respectively, and  $\hat{r}_{CX}$  refers to a unit vector along the direction  
18 of the C=O/C-D bond.

19

20

21

## 1 **11. AMOEBA polarizable MD simulations of aldehydes in solvents.**

2 Polarizable MD simulations were performed using the AMOEBA09 force field in Tinker  
3 8.7.<sup>17</sup> Solute parameters were generated using Poltype as described previously,<sup>16,18</sup> with  
4 polarization groups defined according to conformationally rigid groups and electrostatic  
5 potential fitting based on QM calculations at the MP2/6-311++G(2d,2p) level. Solvent parameters  
6 were taken from previous work,<sup>16,18</sup> and amoeba03 water was used as implemented in the  
7 AMOEBA09 force field. Solute molecules were solvated in a 4.6×4.6×4.6 nm<sup>3</sup> box and the system  
8 was equilibrated using the procedure as reported previously.<sup>16,18</sup> MD production runs were  
9 performed over 1 ns with 1 fs steps, with a van der Waals cutoff of 12 Å (with analytical VdW  
10 correction), an electrostatics cutoff of 7 Å (with particle mesh Ewald method), an induced dipole  
11 convergence threshold set to 10<sup>-5</sup> D, the Beeman integrator,<sup>19</sup> the Bussi thermostat<sup>20</sup> and the  
12 Monte-Carlo barostat (with an inverse friction constant of 1 ps as well as molecular volume  
13 scaling). In order to determine the solvent electric fields acting on solute atoms, the solvent-  
14 specific induced dipoles (i.e., total induced dipoles “minus” intramolecular/solute induced  
15 dipoles) were divided by the atomic polarizability.<sup>16,18</sup> The electric field projections along the C=O  
16 and C-D bond directions were calculated using the same equation as in fixed-charge simulations.

17

18

19

20

21

## 1 **12. New approaches for QM-based solvent electric field calculations.**

2           Obtaining the frequency-field correlation for the CXF-D vibrational probe in solution  
3 requires accurate evaluation of the solvent electric fields along the C=O and C-D bonds. While  
4 calculations based on fixed-charge<sup>3</sup> or polarizable<sup>18</sup> force fields have been successful in  
5 quantifying the solvent electric fields acting on C=O in vibrational probes such as acetophenone,  
6 the accuracy of MM-based approaches in evaluating the electric fields experienced by C-D, which  
7 has a very different bond dipole and surrounding solvent environment than those of C=O, has  
8 never been systematically assessed to the best of our knowledge. As shown in Supplementary  
9 Fig. 4, calculations based on fixed-charge MM yields a poor field-frequency correlation for the C-  
10 D probe. To avoid the potential bias arising from the uncertainties in assigning parameters to the  
11 MM force fields, here we have developed a new computational protocol that employs DFT-based  
12 electronic structure methods to calculate the solvent electric field projections on both the C=O  
13 and C-D bond directions.

14           While in calculations based on MM force fields it is straightforward to partition the  
15 electric fields experienced by the probe into contributions from solvents and those from the  
16 solute atoms themselves, such a partition is not readily available for QM methods since the  
17 electronic wavefunction obtained from a conventional QM calculation is delocalized in the full  
18 system and thus in general non-partitionable. Here we adapted two recently developed  
19 electronic structure partitioning methods to achieve such a partition and thereby quantify the  
20 solvent electric field contributions, which are illustrated in Supplementary Fig. 11. The first  
21 approach, which we refer to as the SPADE (Subsystem Projected AO Decomposition)<sup>21</sup>  
22 partitioning scheme in the present paper, is a recently developed method for partitioning the

1 space spanned by occupied molecular orbitals (MOs) in projection-based embedding theory.<sup>22,23</sup>  
2 It shows better performance than other commonly used approaches such as the method based  
3 on Pipek-Mezey localization and Mulliken population screening.<sup>24</sup> Starting from the fully  
4 converged MOs obtained from a self-consistent field (SCF) calculation for the entire solute-  
5 solvent system, this method transforms the occupied orbitals based on the result of a singular  
6 value decomposition of their subsystem-projected MO coefficient matrix<sup>21</sup>. These transformed  
7 occupied orbitals can then be assigned to the solute and solvent fragments of the system, which  
8 remain orthogonal to each other, and the one-particle electron density matrix constructed from  
9 the occupied orbitals that are assigned to the solvent fragment, together with the solvent nuclei,  
10 will be used to evaluate the electric fields exerted on solute atoms.

11 The second approach, which is based on absolutely localized molecular orbitals<sup>25</sup> and thus  
12 referred to as the ALMO approach throughout this paper, requires a fragment partition of the  
13 system with the solute molecule (e.g. CXF-D) as one fragment and all solvent molecules as  
14 another. The MOs are fragment-designated at the beginning of the calculation (Scheme 2 in  
15 Supplementary Fig. 11), which are variationally optimized subject to the constraint that they are  
16 expanded in the atomic orbital basis functions belonging to the same fragment alone, i.e., the  
17 MO coefficient matrix remains fragment-block-diagonal. The converged ALMOs that belong to  
18 the solvent part of the system are then used to evaluate the solvent electric fields experienced  
19 by the solute vibrational probe. Differing from the SPADE method, the ALMO approach precludes  
20 charge transfer between the solute and solvent regions<sup>26</sup>, whereas the other electrostatic and  
21 non-electrostatic (e.g. Pauli repulsion) interactions between the solute and solvent are  
22 retained.<sup>27</sup> The results in Supplementary Tables 6 and 10 show that these two electronic

1 structure partitioning schemes yield similar results for the electric fields experienced by C=O and  
2 C-D, suggesting that they are both robust schemes for quantifying the solvent electric field  
3 contributions.

4         Due to the high computational demand of QM, we performed electric field calculations  
5 on molecular clusters consisting of a single solute molecule (CXF-D) and solvent molecules, which  
6 were extracted from GAFF (fixed-charge) or AMOEBA (polarizable) MD trajectories. For each  
7 solvent, 100 equally spaced frames were selected from a 1 ns trajectory, except that 500 equally  
8 spaced frames were selected for water due to the particularly large standard deviation associated  
9 with its resultant electric fields, i.e., large inhomogeneous broadening. A truncated solute-  
10 solvent cluster was then extracted from each MD frame, based on the criterion that the center-  
11 of-mass of a solvent molecule included is within a range of 7 Å from any one of the atoms  
12 constituting CXF-D. Note that the atomic coordinates in MD frames were subject to periodic  
13 boundary conditions, which had been taken into account in the generation of solute-solvent  
14 clusters for non-periodic QM calculations. The structures of these clusters were then used for the  
15 QM electric field calculations. We also generated solute-solvent clusters using a larger cutoff  
16 radius (8 Å) for four representative solvents (hexane, toluene, DMSO, and water). The changes  
17 in the obtained mean values of the electric fields experienced by C=O and C-D were all smaller  
18 than 10% relative to the results calculated with the 7 Å cutoff (Supplementary Table 21), with  
19 the largest relative deviation (9.5%) occurring in the electric field of hexane projected along C-D  
20 (which has a particularly small absolute value of ~1.5 MV/cm), indicating that a cutoff value of  
21 7 Å is sufficient for the accurate evaluation of solvent electric fields. With CXF-D as the solute,  
22 such a cluster typically consists of 200—500 atoms.

1 We implemented the two electronic structure partitioning methods introduced above  
2 (SPADE and ALMO) for electric field calculations in a locally modified version of the Q-Chem 5.3  
3 software package.<sup>28</sup> All the solvent electric field calculations were performed at the B3LYP/6-  
4 31+G(d) level of theory,<sup>29,30</sup> with a (75, 302) integration grid (75 radial shells with 302 Lebedev  
5 points in each) for the exchange-correlation functional. The standard and ALMO-based SCF  
6 calculations were both converged to a root-mean-square (RMS) error below  $10^{-8}$  a.u. Since the  
7 solvent electric field is non-uniform in space, the electric field exerted on a chemical bond (which  
8 involves two atoms) cannot be uniquely defined. In this work, we investigated two distinct  
9 schemes to estimate the strength of electric field along a given chemical bond. Denoting the two  
10 bonding atoms as C and X (X = O or D), the first approach (utilized to generate the data in  
11 Supplementary Tables 6-7) utilizes the QM package to calculate the solvent electric fields exerted  
12 on the two atoms, which are two 3-vectors ( $\vec{F}_C$  and  $\vec{F}_X$ ). We then calculate the projection of these  
13 two electric field vectors along the bond direction ( $\hat{r}_{CX}$ ), and the average of these two projections  
14 can be used to estimate the electric field experienced by the CX bond:

$$15 \quad F_{CX} = \frac{1}{2} (\vec{F}_C \cdot \hat{r}_{CX} + \vec{F}_X \cdot \hat{r}_{CX})$$

16 Note that this approach was also employed in the electric field calculations based on fixed-charge  
17 and polarizable force fields. The second approach uses the electrostatic potential (ESP) values on  
18 the C and X atoms ( $\varphi_C$  and  $\varphi_X$ ) obtained from electronic structure calculations (utilized to  
19 generate the data in Supplementary Tables 8-9). The electric field along the CX bond is evaluated  
20 using

$$21 \quad F_{CX} = (\varphi_C - \varphi_X)/r_{CX}$$

1 where  $r_{\text{CX}}$  is the CX bond length. It is easy to show that these two approaches yield the same  
2 results if the electric component increases or drops linearly along the bond direction. The results  
3 in Supplementary Tables 6-9 show that the two approaches introduced above give similar  
4 estimations for the strength of electric field on chemical bonds, making no significant difference  
5 to the frequency-field maps that one would obtain. The excellent agreement between the results  
6 of these two distinct approaches underpinned the physical meaning of the “electric field on bond”  
7 concept that we have been using to discuss the vibrational Stark effect / solvatochromism of the  
8 C=O and C-D stretching modes throughout this paper.

### 9 **13. *Ab initio* vibrational frequency calculations in the gas phase or with implicit solvent.**

10 The harmonic frequencies for the vibrational modes of CXF-D were computed using the  
11 Q-Chem 5.3 software package.<sup>28</sup> At a given level of theory (e.g., B3LYP/6-31+G(d)), we first  
12 optimized the geometry of CXF-D until the maximum component of the gradient was below  $10^{-4}$   
13 a.u. and the energy change relative to that in the previous optimization step was smaller than  $10^{-7}$   
14 a.u. The harmonic frequencies were then computed through analytic second nuclear derivatives  
15 (for DFT-based calculations) or numerical differentiation based on nuclear gradients (for  
16 calculations using resolution-of-identity (RI)-MP2 whose second nuclear derivatives are  
17 unavailable in Q-Chem). A (99, 590) numerical quadrature was employed for the DFT-based  
18 calculations, which corresponds to 99 atom-centered radial shells with 590 Lebedev points in  
19 each. The obtained harmonic frequencies were then uniformly scaled with a multiplicative factor  
20 that is specific to each level of theory, which can be retrieved from the Computational Chemistry  
21 Comparison and Benchmark Database (CCCBDB).<sup>31</sup> For the B3LYP/6-31+G(d) level of theory, the  
22 same scaling factor as that for B3LYP/6-31+G(d,p) (0.964) was adopted (see Fig. 4 and

1 Supplementary Table 16). The same procedure was also applied to calculate the harmonic  
2 frequencies of acetyl chloride in the gas phase (Supplementary Table 22).

3 We also performed DFT-based vibrational frequency calculations (at the B3LYP/6-31+G(d)  
4 level) for CXF-D with a uniform dipolar electric field applied along the direction of either of these  
5 two bonds (Supplementary Fig. 5 and Supplementary Table 14), which are referred to as “Stark  
6 *in silico*” calculations throughout the discussions in this paper. The strength of the electric field  
7 applied along C=O ranged from -160 to +20 MV/cm, and that along C-D from -20 to +80 MV/cm.  
8 In these calculations, we first re-oriented the geometry of CXF-D optimized without the external  
9 field such that the C=O or C-D bond was aligned with the z-axis in the Cartesian frame. The  
10 geometry of CXF-D was then re-optimized at each given magnitude of the external electric field  
11 along the z direction, followed by a harmonic frequency calculation under the same field. We  
12 note that the external electric field applied along either of the bond directions will exert a torque  
13 on the CXF-D molecule, such that the C=O or C-D bond in the optimized geometry will not be  
14 exactly parallel to the applied field which remains along the z direction. Nevertheless, the  
15 deviation of the C=O or C-D bond direction from the z axis was found to be very small in the  
16 optimized geometry, and thus had negligible impacts on the field-frequency maps obtained from  
17 these calculations. A similar set of calculations was also performed for the acetyl chloride  
18 molecule with an external electric field applied along either the C=O or C-Cl directions, with the  
19 field strength ranging from -100 to +20 MV/cm (Supplementary Fig. 13 and Supplementary Table  
20 18).

21 As a computationally less costly way to characterize the solvatochromism associated with  
22 the C=O and C-D stretch frequencies, we also performed harmonic frequency calculations for

1 CXF-D (at the B3LYP/6-31+G(d) level of theory) in solvents described by the conductor-like  
2 polarizable continuum model (C-PCM)<sup>32,33,34</sup> (Supplementary Table 16). In the C-PCM calculations,  
3 the surface of the solute cavity was determined by coalescing each atom's van der Waals (vdW)  
4 radius<sup>35</sup> (scaled by a factor of 1.2). For both hydrogen and heavy atoms, 302 Lebedev grid points  
5 are placed on the surface of scaled vdW spheres. To ensure the smoothness of the potential  
6 energy surface of solute in PCM solvent, the switching/Gaussian (SWIG) method<sup>36,37</sup> as  
7 implemented in Q-Chem was employed in the C-PCM calculations. The other computational  
8 details for geometry optimization and harmonic frequency analysis are the same as those for the  
9 gas-phase calculations described above.

10 To shed light on the difference in the local field factors associated with the C=O and C-D  
11 bonds, we further performed "Stark *in silico*" calculations in PCM to model the local solvation  
12 environment experienced by the vibrational probe within a frozen glass. These calculations  
13 capture the vibrational frequency shifts when a chemical bond experiences both an external  
14 electric field and the field arising from the local solvation environment, and the local field factor  
15 can be estimated by calculating the ratio of the frequency shift calculated in PCM ( $\Delta\nu_{\text{PCM}}$ ) to that  
16 in vacuum ( $\Delta\nu_{\text{vac}}$ ) at a given strength of the external field:  $f = |\Delta\nu_{\text{PCM}}|/|\Delta\nu_{\text{vac}}|$ . The  
17 computational details for the "Stark *in silico*" calculations as well as the use of PCM are the same  
18 as described above. Dielectric constants 6.97 and 62.046 were used for the two glass-forming  
19 solvents, 2-MeTHF and the 1:1 (v/v) mixture of D<sub>2</sub>O/glycer(ol-D<sub>3</sub>), respectively.

#### 20 **14. Calculations of molecular electrostatic properties.**

21 To shed light on the electrostatic interactions between the aldehyde group of CXF-D and  
22 solvents, we obtained the atomic charge populations on the C, D, and O atoms from electronic

1 structure calculations at the B3LYP/6-31+G(d) level of theory. The calculations were performed  
2 at the optimized geometry of CXF-D (with the same QM theory). Multiple charge population  
3 schemes that are available in the Q-Chem were employed, including Mulliken,<sup>38</sup> iterative  
4 Hirshfeld,<sup>39</sup> ChelpG,<sup>40</sup> Merz-Kollman,<sup>41</sup> and the charge model 5 (CM5).<sup>42</sup> The atomic charges  
5 obtained from these schemes (Supplementary Table 27) vary from each other to a large extent,  
6 but the results are qualitatively consistent, that is, there is a strong bond dipole associated with  
7 C=O, and both the carbon and deuterium atoms are positively charged in the C-D bond while the  
8 D atom is less so.

9         At the same level of theory, we generated the electrostatic potential (ESP) map of the  
10 CXF-D molecule. Given the converged electron density of CXF-D obtained from the DFT  
11 calculation, the ESP in the 3-space was calculated on a cubic uniform grid. The ESP map of CXF-D  
12 shown in Supplementary Fig. 6 was evaluated on the 1.2× vdW radii surface using the IQmol  
13 visualization package. The ESP map was also generated for the acetyl chloride molecule in the  
14 same way.

15  
16  
17  
18  
19  
20  
21  
22  
23

# Supplementary Texts

## 1. The vibrational solvatochromism of acetyl chloride.

Since the carbon-chlorine bond (C-Cl) has a larger bond dipole than C-D, we expect that the solvent field's orientation around acetyl chloride should be significantly different from that on CXF-D. In particular, unlike the observed bond destabilization (positive field) on C-D, we expect that the C-Cl bond should experience favorable electrostatic interactions with solvents (negative field). Following this intuition, the C=O of acetyl chloride would consequently experience a smaller negative solvent field than CXF-D's C=O since the electric field orientation would be shifted toward the C-Cl's bond axis. The computational results shown in Supplementary Fig. 12 are in good agreement with our predictions: the C=O of acetyl chloride has a smaller electric field than that in CXF-D (Supplementary Fig. 12), and the C-Cl experiences a negative solvent electric field (Supplementary Fig. 12), which is in opposition to the field experienced by CXF-D's C-D.

The FTIR spectra of acetyl chloride in three solvents (Supplementary Fig. 7; note that acetyl chloride is not stable in water) show its much smaller solvatochromic shifts in  $\bar{\nu}_{\text{CO}}$  than that of CXF-D, which also qualitatively agrees with the *ab initio* frequency calculations using PCM to describe solvent effects (Supplementary Tables 16 and 22). However, it is noteworthy that the observed frequency shift from hexane to acetonitrile is only  $2 \text{ cm}^{-1}$ , which is much smaller than the corresponding shift for CXF-D ( $19 \text{ cm}^{-1}$ ). We used VSS to experimentally calibrate the sensitivity of acetyl chloride's C=O to an applied external electric field (Supplementary Fig. 8), showing that its Stark tuning rate is  $1.80/f \text{ cm}^{-1}/(\text{MV}/\text{cm})$ , where  $f$  denotes the local field factor.  $f$  has been found to be about 2 for carbonyls, so the actual Stark tuning rate for the C=O of acetyl

1 chloride should be around  $0.90 \text{ cm}^{-1}/(\text{MV}/\text{cm})$ , which is in good agreement with the result  
2 obtained from *ab initio* VSE calculations ( $1.07 \text{ cm}^{-1}/(\text{MV}/\text{cm})$ , see Fig. S13). The  $2\text{-cm}^{-1}$  frequency  
3 shift over this range of solvents thus suggests that the solvent electric field experienced by C=O  
4 has changed by only  $\sim 2 \text{ MV}/\text{cm}$  moving from hexane to acetonitrile, a much smaller value than  
5 those obtained from the solvent electric field calculations (Supplementary Tables 23-26), which  
6 are typically larger than  $20 \text{ MV}/\text{cm}$  despite the variations with the computational method  
7 employed. We speculate that the unusually small frequency shift of acetyl chloride's C=O may  
8 arise from through-bond polarization when C-Cl is stabilized by solvents. Another possibility is  
9 that the force field parameterization for the Cl atom in acetyl chloride might be insufficiently  
10 accurate such that in the MD simulations the interaction between C-Cl and the solvents was  
11 underestimated. Nonetheless, the extraordinarily small C=O solvatochromic shift in acetyl  
12 chloride compared with other carbonyl groups probed to date reflects the unusual tug-of-war  
13 between the C=O and C-Cl bonds for solvent stabilization.

14

15

16

17

18

19

20

1 **2. Derivation of Equation 1 in the main text (calculation of the angle between the electric**  
2 **field vector and the C=O dipole direction in the aldehyde plane).**

3 Here we show how one can extract the orientation of the electric field vector in the plane  
4 of the aldehyde moiety from the ratio of electric field projections on C-D ( $F_{C-D}$ ) to that on C=O  
5 ( $F_{C=O}$ ). As shown in Fig. 1c, denoting the angle between the electric field vector in the plane and  
6 the C=O bond dipole direction (from O to C) as  $\theta$  (which is positive when it is on the opposite side  
7 as C-D with regard to the C=O bond dipole), with the assumption that the angle between C=O  
8 and C-D is fixed at  $120^\circ$ , the electric field projections along C=O and C-D are related to the total  
9 electric field in the aldehyde plane through the following equalities:

10 
$$F_{C-D} = |\vec{F}_{\text{tot}}| \cos\left(\theta + \frac{\pi}{3}\right), \quad F_{C=O} = |\vec{F}_{\text{tot}}| \cos(\pi - \theta).$$

11 Therefore, we have

12 
$$\frac{F_{C-D}}{F_{C=O}} = \frac{\cos\left(\theta + \frac{\pi}{3}\right)}{\cos(\pi - \theta)} = \frac{\frac{1}{2}(\cos \theta - \sqrt{3} \sin \theta)}{-\cos \theta} = -\frac{1}{2}(1 - \sqrt{3} \tan \theta).$$

13 Solving for the angle  $\theta$  gives Equation 1 in the main text

14 
$$\theta = \arctan\left[\frac{1}{\sqrt{3}}\left(1 + 2\frac{F_{C-D}}{F_{C=O}}\right)\right].$$

15

16

17

18

19

### 1 3. Discussion on the local field factor.

2 In vibrational Stark spectroscopy (VSS) experiments, an external voltage ( $V$ ) of known  
3 magnitude was applied to a parallel capacitor of known pathlength ( $d$ ) that contains the molecule  
4 of interest. The applied external electric field ( $|\vec{F}_{\text{ext}}|$ ) at the macroscopic level can thus be  
5 obtained by  $|\vec{F}_{\text{ext}}| = V / d$ . However, because the sample is embedded in a frozen dielectric  
6 medium such as 2-MeTHF or  $D_2O$ /glycer( $ol-D_3$ ) mixture at 77 K rather than in vacuum, the actual  
7 electric field ( $|\vec{F}_{\text{local}}|$ ) experienced by a vibrational probe at the microscopic level differs from  
8  $|\vec{F}_{\text{ext}}|$  as a result of the inhomogeneous polarization of the dielectric medium induced by the  
9 external electric field. This difference is taken into account using the local field factor  $f$ ,<sup>4,43</sup> i.e.,  
10  $|\vec{F}_{\text{local}}| = |\vec{F}_{\text{ext}}|f$ . In general the local field factor  $f$  is a tensor and several models are available  
11 for estimating its value<sup>4</sup>, but we approximate it as a scalar. Because of the uncertainty in the  
12 actual value of the local field factor for any particular glass used to measure Stark spectra, we  
13 report the observed value of  $\Delta\mu$  scaled by the local field factor as:  $|\Delta\mu|f$ . Vibrational  
14 solvatochromism provides another means to estimate  $\Delta\mu$  from the slope of the correlation  
15 between experimentally measured frequencies and calculated electric field projections on bonds  
16 (e.g. Fig. 4). While the local field correction does not enter into solvatochromism measurements,  
17 a comparison of the slope of solvatochromism fit with the VSS data gives an estimate for the local  
18 field correction, to the extent that the calculated fields are meaningful. This issue has been  
19 discussed in depth elsewhere.<sup>44,45</sup> For  $C=O$  and  $C\equiv N$   $f \approx 2$  while for vibrational transitions that  
20 involve hydrogen, such as  $O-H(D)$ ,  $N-H$ , and  $S-H$ ,  $f \approx 1$ . Since previously only one vibrational  
21 transition of a molecule was studied each time, only one corresponding local field factor was  
22 obtained from the VSS experiment. However, in this work, the two-directional vibrational probe

1 carrying C=O and C-D around the same carbon atom offers an opportunity to estimate the local  
2 field factors in the directions of two chemical bonds. In particular, by comparing the Stark tuning  
3 rates obtained from the solvatochromism fits (Fig. 4) to those measured from VSS experiments  
4 (Fig. 2), the resulting value of  $f$  in the VSS studies of CXF-D in the direction of C=O ( $f_{C=O}$ ) is  $\sim 2.7$ ,  
5 which falls in the range of estimated values in previous studies,<sup>3,44</sup> while in the direction of C-D  
6 ( $f_{C-D}$ ) the value is  $\sim 1.2$  (Figs. 2 and 4).

7         There are two most likely reasons for the different local field factors for C=O and C-D in  
8 the aldehyde probes reported here. First, the VSS experiments for C=O and C-D were performed  
9 in different solvents at different concentrations, i.e., 1 M CXF-D in 2-MeTHF for C-D, where the  
10 high concentration is required due to the weak absorbance of C-D, and 50 mM CXF-D in  
11 D<sub>2</sub>O/glycer(ol-D<sub>3</sub>) mixture for C=O. Here we used a different solvent for the measurement of C=O  
12 because CXF-D could form dimers through intermolecular hydrogen bonds in 2-MeTHF  
13 complicating the C=O absorption spectrum. Second, the local field factors for C=O and C-D may  
14 be *intrinsically different* due to their distinct local solvation environments. As discussed in the  
15 main text of this paper, the solvents tend to organize around CXF-D to preferentially stabilize C=O  
16 while destabilizing C-D as a secondary effect, leading to a negative (stabilizing) electric field along  
17 C=O and a positive (destabilizing) field along C-D (Fig. 3 and Fig. 4). Since the local field factor  
18 originates from the anisotropic polarization of the dielectric medium imposed by the external  
19 electric field, it is anticipated that the different solvation (electrostatic) environments  
20 experienced by C=O and C-D could lead to different local field factors. That is, the external electric  
21 field is screened by solvents to different extents along C=O and C-D depending on their local  
22 solvation environment. Although the samples of VSS experiments are embedded in a frozen glass

1 environment that is different from the liquid phase used in the solvatochromism experiments  
2 and in the enzyme, it is conceivable that the model of preferential solvation revealed in the liquid  
3 phase also applies to the frozen glass, i.e., the local solvation environments of C=O and C-D are  
4 different in both the liquid phase and the frozen glass phase.

5 To gain additional insights into the different local field factors for C=O and C-D in CXF-D,  
6 we extended the “Stark *in silico*” calculations (DFT frequency calculations with uniform electric  
7 fields applied along the bond directions, see Supplementary Fig. 5) to environments with implicit  
8 solvation, i.e., instead of performing these calculations *in vacuo*, we employed the C-PCM  
9 model<sup>32,33,34</sup> as a first approximation to describe the solvation environment in the frozen glass  
10 formed by 2-MeTHF and the D<sub>2</sub>O/glycer(ol-D<sub>3</sub>) mixture. The field-on calculations were performed  
11 at a field strength of +/- 20MV/cm for C-D/C=O, and the experimental dielectric constants for the  
12 liquid phase of these solvents were employed in the calculations with PCM [ $\epsilon = 6.97$  was used  
13 for 2-MeTHF and 62.0<sup>46</sup> for the 1:1 (v/v) mixture of D<sub>2</sub>O/glycer(ol-D<sub>3</sub>)]. The results are  
14 summarized in Supplementary Table 30. Remarkably, the Stark tuning rates obtained using this  
15 relatively simple theoretical model are in good agreement with the VSS experimental results:  
16 1.01 (sim.) vs. 1.13 (exp.) cm<sup>-1</sup>/(MV/cm) for C=O, and 1.22 (sim.) vs. 1.29 (exp.) cm<sup>-1</sup>/(MV/cm) for  
17 C-D, where “sim.” and “exp.” denote simulated and experimental values, respectively. The  
18 resulting local field factors  $f$  [estimated using the ratio of the Stark tuning rate calculated with  
19 PCM ( $|\Delta\mu|_{\text{PCM}}$ ) to that in the gas phase ( $|\Delta\mu|_{\text{vac}}$ ), i.e.,  $f = |\Delta\mu|_{\text{PCM}} / |\Delta\mu|_{\text{vac}}$ ] are  $\sim 1.5$  for C=O and  
20  $\sim 0.9$  for C-D regardless of the choice of solvent. We note that the local field factors obtained  
21 here are smaller than those estimated by comparing the VSS-measured values to the slopes of  
22 the solvatochromism fits ( $\sim 2.7$  for C=O and  $\sim 1.2$  for C-D), which is likely due to the

1 overestimation<sup>44</sup> in  $|\Delta\mu|_{\text{vac}}$  by the “Stark *in silico*” calculations performed in the gas phase  
2 (Supplementary Fig. 5). Nevertheless, the observed trend that  $f_{\text{CD}} < f_{\text{CO}}$  is well reproduced by  
3 these calculations. The fact that the use of different solvents only causes small changes to the  
4 Stark tuning rates as well as the local field factors obtained suggests that the different local field  
5 factors for C=O and C-D are more likely to stem from the different solvent organization around  
6 these two bonds rather than the change in the solvent dielectric constants.

7

#### 8 **4. Vibrational solvatochromism of CFX-D in alcohols.**

9 Despite strong linear field-frequency and field-field correlations observed in a wide range  
10 of solvents (Figs. 4 and 5), water is quite different from most of the other solvents as it can form  
11 hydrogen-bonding (HB) interactions with carbonyl groups (C=O), leading to a significant redshift  
12 in the C=O frequency. It has been shown that this large redshift is largely due to the vibrational  
13 Stark effect given that the redshifts of carbonyl groups in water follow the same linear field-  
14 frequency correlation with other solvents.<sup>3,5,44</sup> The vibrational Stark spectroscopy<sup>3</sup> and recent 2D  
15 IR studies<sup>47</sup> have shown that the Stark tuning rate and bond anharmonicity of C=O are largely  
16 unaffected upon HB formation, further demonstrating that the HB interaction experienced by  
17 C=O can be quantitatively described within the framework of electrostatic interactions.

18 However, up until now we have not extensively explored polar protic solvents other than  
19 water. This is mostly because: (1) experimentally, the IR spectra in non-aqueous protic solvents,  
20 which usually have weaker HB capabilities than water, may be complicated by the coexistence of  
21 HB and non-HB populations; (2) on the computational side, the electric field calculations for  
22 protic solvents place a premium on both the accuracy of the MD force field and the solvent

1 parameterization for describing HB interactions, which is more subtle than other non-specific  
2 intermolecular interactions. Despite these potential challenges, we have studied methanol and  
3 ethanol here as two other protic solvents besides water and show that, despite additional  
4 complexity due to the heterogeneity of solvation environment, the results in alcohols are in good  
5 agreement with the established field-frequency and field-field correlations.

6         Supplementary Figure 15 shows the IR spectra of CFX-D in methanol (a-b) and ethanol (c-  
7 d). The non-symmetric C=O and C-D IR bands indicate the heterogeneity of solvation  
8 environments in alcohols as mentioned earlier. As a minimum model, we fit these spectra with  
9 two peaks, which correspond to the HB and non-HB populations (Supplementary Table 3).

10         To obtain the ensemble-averaged solvent electric fields for both HB and non-HB  
11 populations, we analyzed MD trajectories and grouped MD frames into HB and non-HB categories  
12 based on a commonly used criterion for an HB.<sup>48</sup> We then calculated solvent electric fields for  
13 both two populations using the SPADE method (Supplementary Methods 12, Supplementary  
14 Table 6). Supplementary Figure 16 shows that both HB and non-HB populations in alcohols form  
15 good linear field-frequency correlations along with other solvents for both C-D and C=O,  
16 indicating that the solvatochromic shifts in alcohols are largely due to the vibrational Stark effect,  
17 which is similar to the case of water. We then studied the field-field correlation and plotted the  
18 results in Supplementary Fig. 16, where both HB and non-HB populations in alcohols fall on the  
19 same line as other solvents, demonstrating that this linear field-field correlation also applies to  
20 protic solvents other than water. This observation further supports our conclusion and shows  
21 that the HB interactions in solutions increase the magnitude of electric fields but have a small

1 effect on the field orientation, demonstrating the fundamental similarity of electric field  
2 orientations experienced by solute in a broad range of solvents.

### 3 **5. Fermi resonance analysis for the C-D vibration of CXF-D.**

4 The Fermi resonance coupling constant ( $W$ ) can be estimated by Eq. 1, where  $\Delta$  is the  
5 frequency separation between the two perturbed vibrations that are in Fermi resonance and  $R$   
6 denotes the ratio of their intensities.<sup>49,50,51</sup>

7

$$8 \quad W = \frac{1}{2} \cdot \Delta \cdot \sqrt{1 - \left(\frac{R - 1}{R + 1}\right)^2}$$

9 (Eq. 1)

10 Based on Eq. 1, we obtain the Fermi coupling coefficients ( $W$ ) of CXF-D in the solvents  
11 investigated in this work as shown in Supplementary Table 31 (note that in water the Fermi  
12 resonance is very weak and thus undetectable). It shows that the coupling constant  $W$  varies in  
13 different solvents ranging from 7.5 to 13  $\text{cm}^{-1}$ . It is worth noting that the perturbed frequency of  
14 the combination/overtone mode ( $\nu_2$ ) that forms Fermi resonance with the C-D stretch mode ( $\nu_1$ )  
15 does not significantly shift in solvents. We think that is likely due to an interplay between the  
16 intrinsic sensitivity of that vibrational mode to solvent polarity and the coupling strength with  
17 the C-D stretching mode in different solvents. We also note that the Fermi resonance feature has  
18 vanished in LADH, which is most likely because the significant blueshift of the C-D band in the  
19 LADH active site makes its energy ( $\nu_1$ ) and that of the other combination/overtone mode ( $\nu_2$ )  
20 substantially different, and therefore eliminates the Fermi resonance between them. This can

1 already be seen in the solvatochromism data (Fig. 3) as the peak intensity of the Fermi-resonance  
2 band progressively decreases and eventually vanished as the C-D band shifts to the blue going  
3 from hexane to water. Combination/overtone modes in the protein environment are also  
4 expected to be quite different from those in simple solvents.

5

6

7

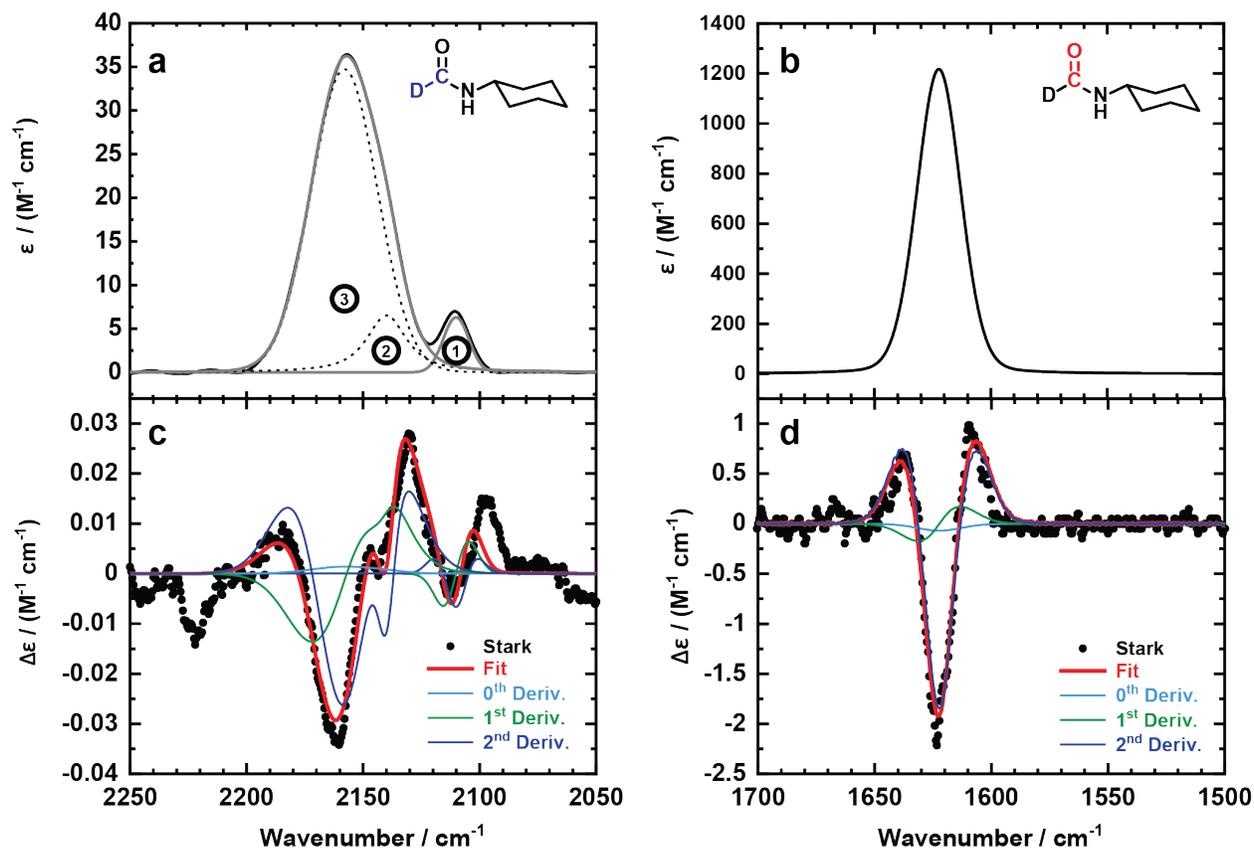
8

9

10

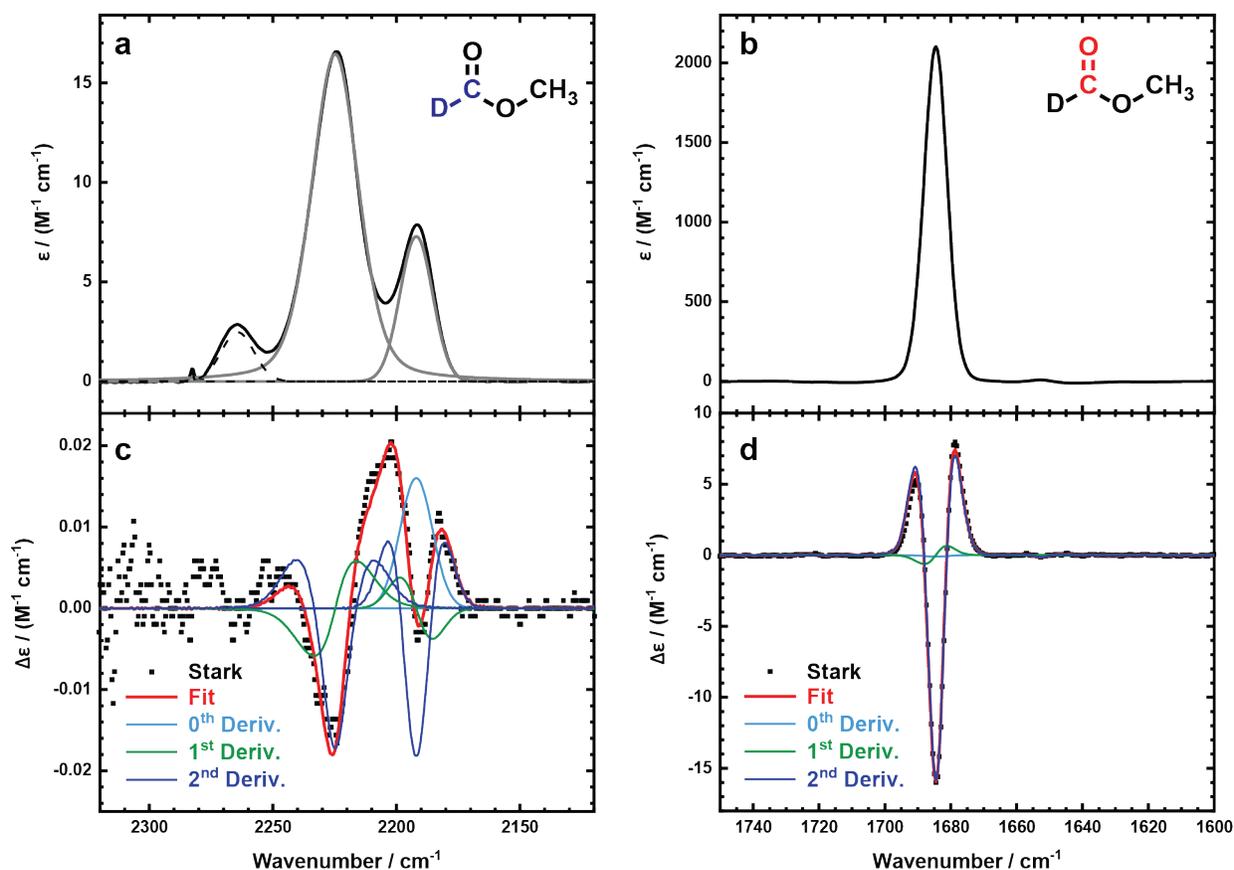
11

# 1 Supplementary Figures



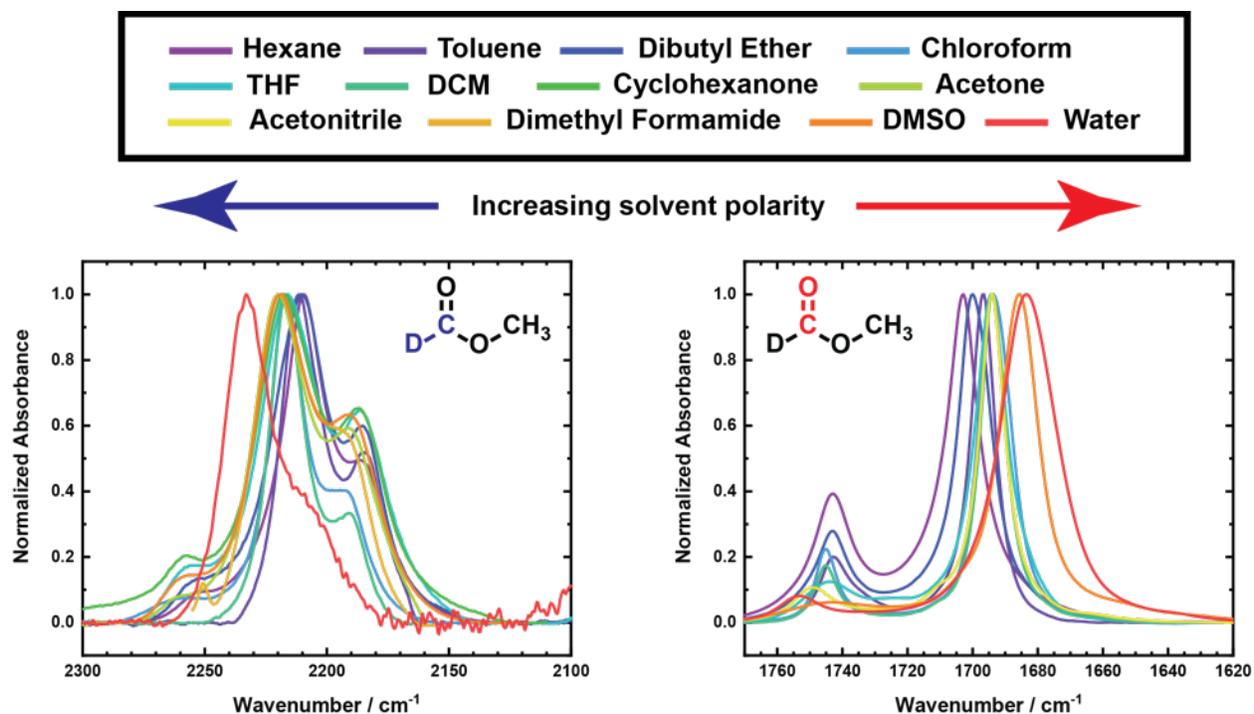
2  
 3 **Fig. 1. Vibrational Stark spectroscopy of CXF-D measured at 77 K with the derivative**  
 4 **contributions.** The upper panels (a and b) are infrared absorption spectra, and the lower panels  
 5 (c and d) are the corresponding vibrational Stark spectra, which are scaled to an applied electric  
 6 field of 1.0 MV/cm, with best fits (red) and derivative contributions (see Supplementary Table 1  
 7 for full details of fitting). The C-D spectra were measured using 1 M CXF-D in 2-methyl  
 8 tetrahydrofuran, and the absorption spectrum is fitted to a sum of three Voigt profile  
 9 components (dash lines) labeled as 1, 2, and 3 in (a). Peak 1 is believed to result from Fermi  
 10 resonance with the overtone of the C-D bending mode (see the section of Results and discussion).  
 11 We note that the Stark spectrum of Peak 1 in (c) around 2090  $\text{cm}^{-1}$  is not fully captured by the  
 12 “best fit”. However, the poor fit for that part of the spectrum has a negligible effect on the value  
 13 of the Stark tuning rate of C-D ( $|\Delta\vec{\mu}_{\text{C-D}}|f_{\text{C-D}}$ ) that is extracted. Even if we fit the Stark spectrum  
 14 with only the derivative contributions of the main C-D band (i.e., without Peak 1 at all), the  
 15 obtained  $|\Delta\vec{\mu}_{\text{C-D}}|f_{\text{C-D}}$  remains essentially the same compared to the fitting result that includes the  
 16 shoulder peak [differing by  $\sim 0.01 \text{ cm}^{-1}/(\text{MV}/\text{cm})$ ]. Peaks 2 and 3 are both considered to  
 17 correspond to the C-D stretch mode, and there are two peaks likely due to the dimerization of  
 18 CXF-D under the required condition of high concentration. This is evidenced by the fact that the  
 19 peaks 2 and 3 have the same Stark tuning rate and can be grouped together (gray line in a) to fit  
 20 the Stark spectrum (c), while peak 1 has very different fitting coefficients from that of peaks 2  
 21 and 3. The C=O spectra were measured using 50 mM CXF-D in  $\text{D}_2\text{O}/\text{glycer}(\text{ol-D}_3)$  ( $v/v = 1/1$ ).

1



2

3 **Fig. 2. Vibrational Stark spectroscopy of methyl formate-D in 2-methyl tetrahydrofuran at 77 K**  
 4 **with the derivative contributions.** The upper panels (a and b) are infrared absorption spectra,  
 5 and the lower panels (c and d) are the corresponding vibrational Stark spectra, which are scaled  
 6 to an applied electric field of 1.0 MV/cm, with best fits (red, see Supplementary Table 1 for full  
 7 details of fitting). For C-D, the concentration of methyl formate-D was 4 M because the extinction  
 8 of the C-D bond is very small; for C=O, the concentration of methyl formate-D was 0.1 M.



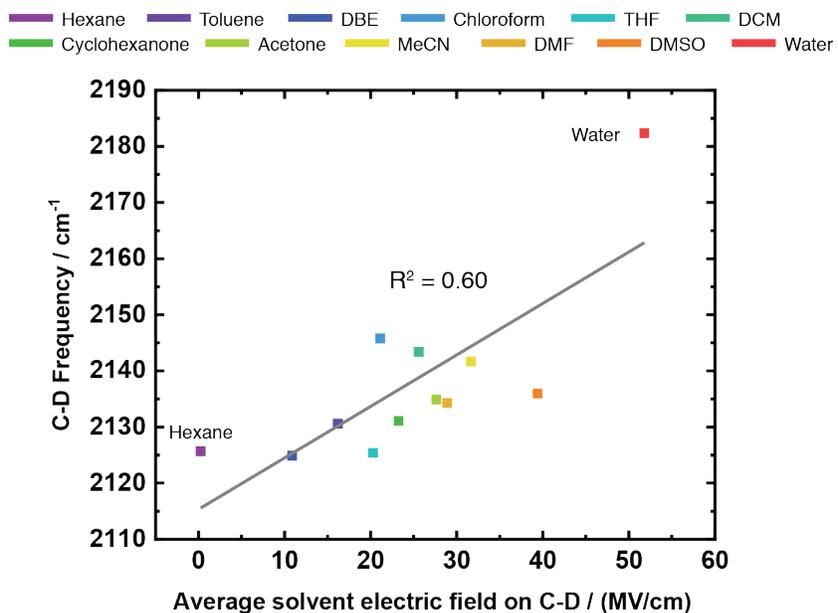
1

2 **Fig. 3. Vibrational solvatochromism of methyl formate-D.** The FTIR spectra were recorded when  
 3 20 mM methyl formate-D was dissolved in a series of organic solvents with various polarities.  
 4  $\bar{\nu}_{C-D}$  blue-shifts while  $\bar{\nu}_{C=O}$  red-shifts as the solvent polarity increases. The shoulder peaks in the  
 5 C-D spectrum at 2180 cm<sup>-1</sup> and in the C=O spectrum at 1740 cm<sup>-1</sup> are likely due to Fermi  
 6 resonances; those peaks had minimal frequency shifts yet notable changes in intensity in  
 7 different solvents.

8

9

10



1

2 **Fig. 4. The correlation between experimentally measured C-D frequencies of CXF-D in solvents**  
 3 **and the average solvent electric fields obtained from fixed-charge MD simulations based on**  
 4 **the electrostatic force divided by the atomic charge<sup>3</sup>. The least-square fit is  $\bar{\nu}_{C-D} = 0.92F_{C-D}^{solv} +$**   
 5  **$2115.3 (R^2 = 0.60)$ , where  $F_{C-D}^{solv}$  is the solvent electric field projection along C-D.**

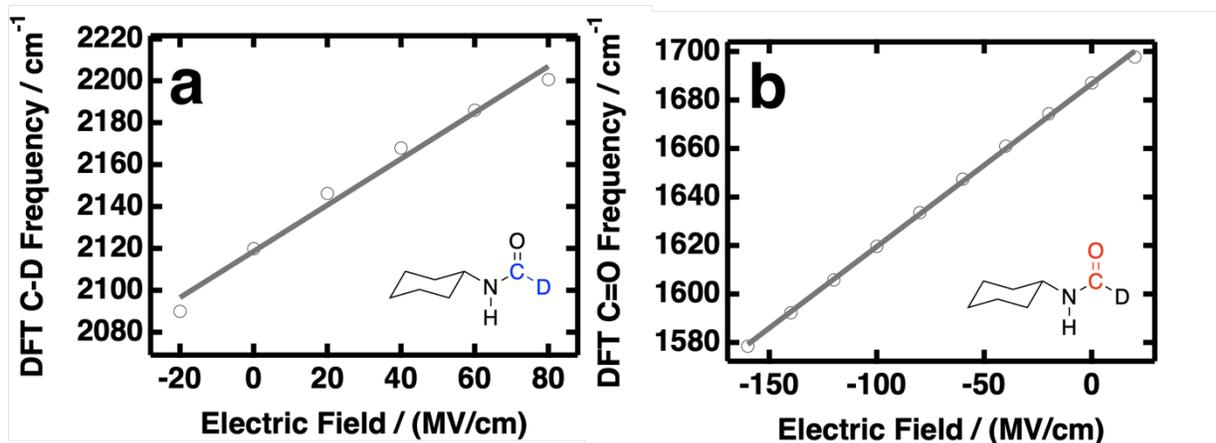
6

7

8

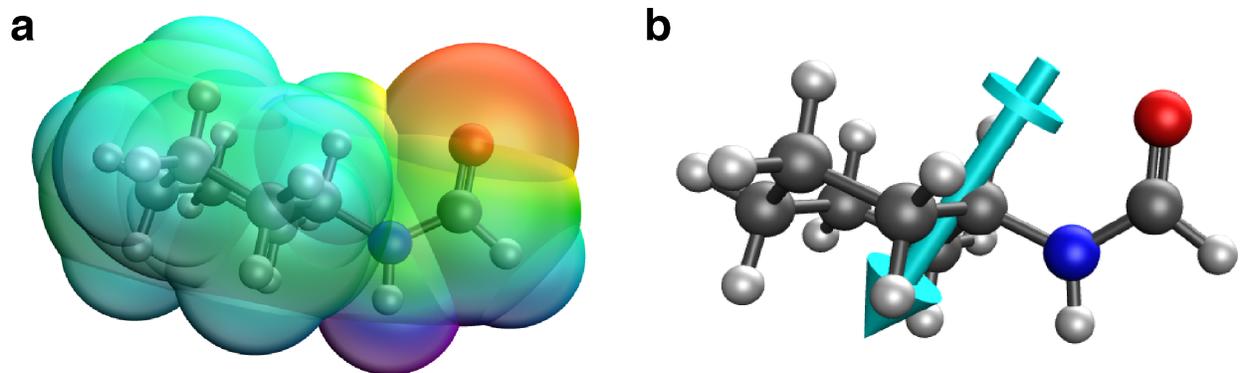
9

10

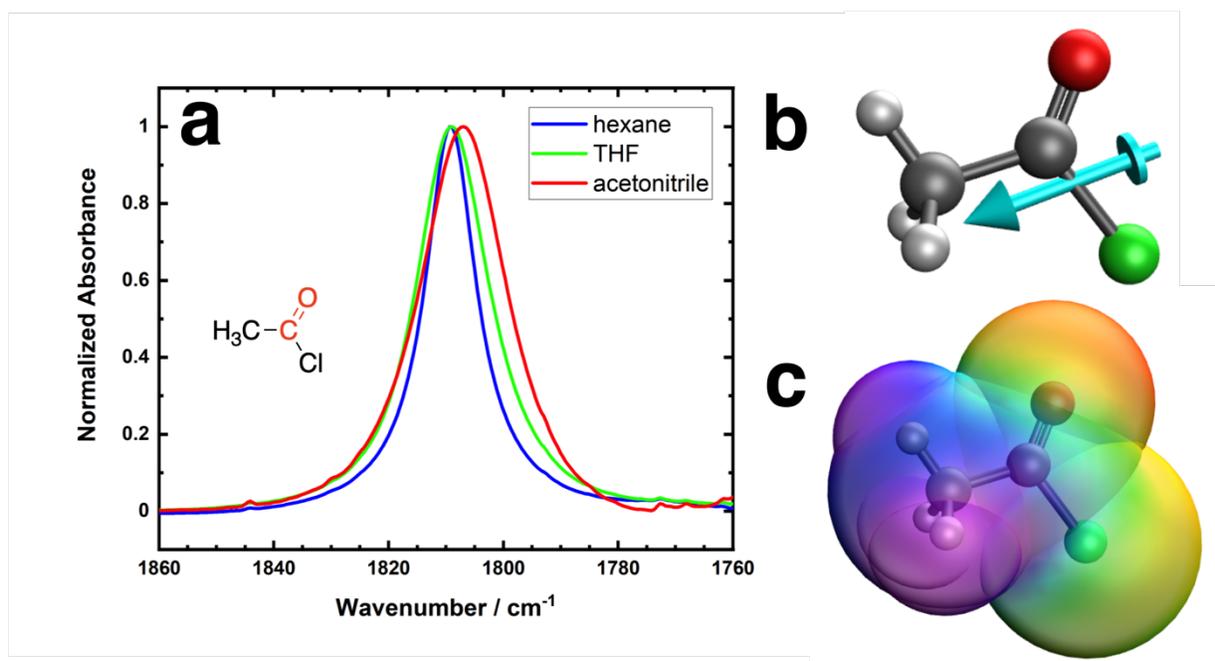


1  
 2 **Fig. 5. Vibrational Stark effect of C-F-D's C-D (a) and C=O (b) stretch modes revealed using gas-**  
 3 **phase "Stark *in silico*" calculations.** A uniform external electric field was applied along the bond  
 4 axis of either C-D or C=O, and *ab initio* frequency calculations were performed at the B3LYP/6-  
 5 31+G(d) level (scaled by 0.964<sup>31</sup>) to determine the vibrational frequencies as a function of the  
 6 strength of the external electric field, whose original data are shown in Supplementary Table 18.  
 7 (a) The C-D stretch frequency of C-F-D blue-shifts as the magnitude of positive electric field along  
 8 C-D increases, consistent with the results of vibrational solvatochromism (Fig. 3). A linear fitting  
 9 for the field-frequency correlation of C-D gives  $\bar{\nu}_{C-D} = 1.10F_{C-D} + 2118.70$  ( $R^2 = 0.98$ ). (b)  
 10 The C=O stretch frequency of C-F-D red-shifts as the magnitude of negative electric field along  
 11 C=O increases. A linear fitting for the field-frequency correlation yields  $\bar{\nu}_{C=O} = 0.67F_{C=O} +$   
 12  $1686.80$  ( $R^2 = 0.9993$ ). It is noteworthy that the Stark tuning rate of C-D [ $1.10 \text{ cm}^{-1}/(\text{MV}/\text{cm})$ ]  
 13 obtained purely *in silico* here is larger than that of C=O [ $0.67 \text{ cm}^{-1}/(\text{MV}/\text{cm})$ ]. This is in good  
 14 agreement with the Stark tuning rates obtained from solvatochromism fits ( $1.07$  and  $0.42 \text{ cm}^{-1}/(\text{MV}/\text{cm})$   
 15  $^1/(\text{MV}/\text{cm})$  for C-D and C=O, respectively, see Fig. 4). We also note that previous studies have  
 16 shown that DFT methods tend to overestimate the Stark tuning rates (by as much as 50%)<sup>52</sup>,  
 17 which may explain the larger Stark tuning rates obtained in "Stark *in silico*" calculations than  
 18 those from solvatochromism fits.

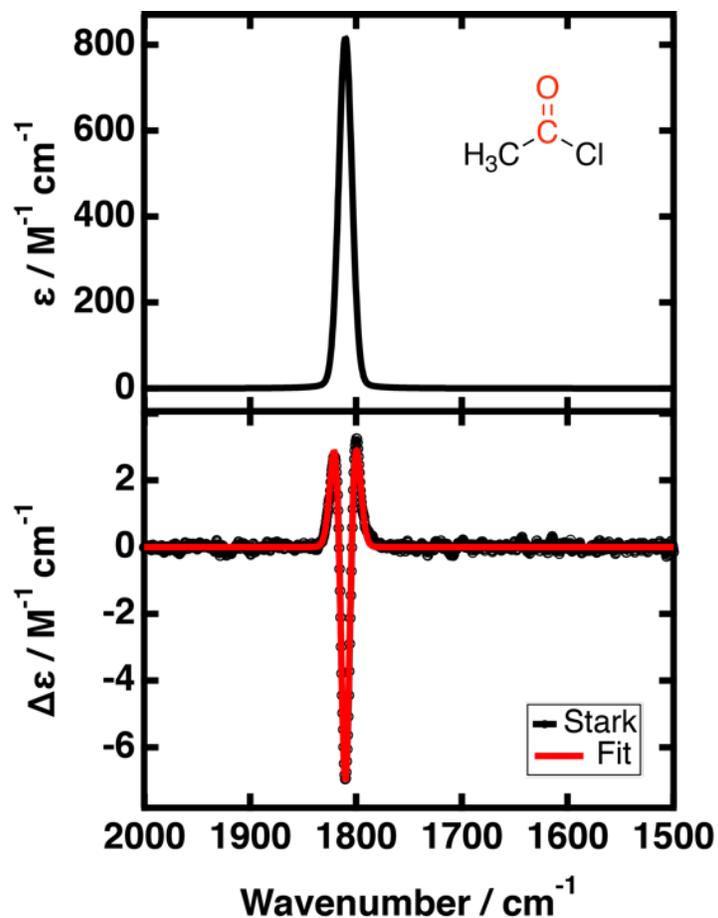
19  
 20  
 21



1  
 2 **Fig. 6.** (a) The molecular model of CXF-D with its electrostatic potential map. (b) the molecular  
 3 model of CXF-D with its molecular dipole moment.  
 4



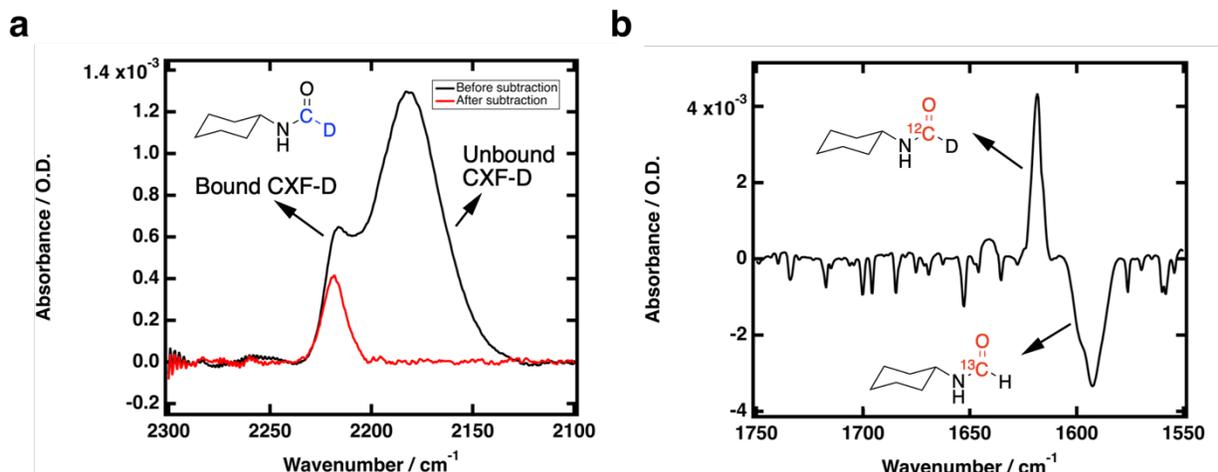
5  
 6 **Fig. 7. Vibrational solvatochromism of acetyl chloride.** 10 mM acetyl chloride was dissolved in  
 7 hexane, THF, and acetonitrile, and the FTIR spectra of C=O were recorded. (a) FTIR spectra of the  
 8 C=O in acetyl chloride. (b) Molecular model of acetyl chloride with its molecular dipole moment.  
 9 (c) Molecular model of acetyl chloride with its electrostatic potential map (warm/cold color  
 10 indicates negative/positive potential). The C=O of acetyl chloride shows a surprisingly small  
 11 redshift (2  $\text{cm}^{-1}$ ) going from hexane to acetonitrile (Supplementary Table 17) compared to the  
 12 shift in CXF-D (Fig. 3b), methyl formate-D (Supplementary Fig. 3), or any other simple carbonyl-  
 13 containing compound observed to date.



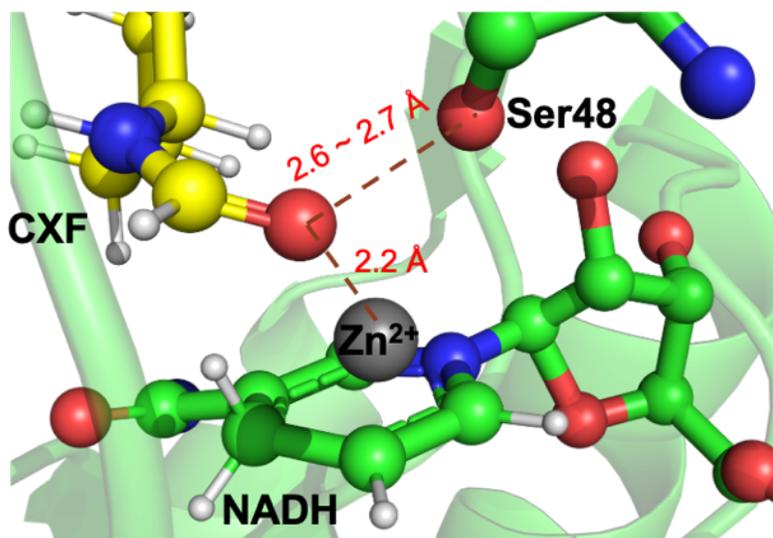
1

2 **Fig. 8. Vibrational Stark spectroscopy of C=O measured with 100 mM acetyl chloride in 2-**  
 3 **methyltetrahydrofuran at 77 K.** The upper panel is the absorption spectrum, and the bottom is  
 4 the vibrational Stark spectrum which is shown as a difference spectrum of field on minus field off  
 5 and scaled to an applied electric field of 1.0 MV/cm with best fits (red, see Supplementary Table  
 6 1 for full fitting details).

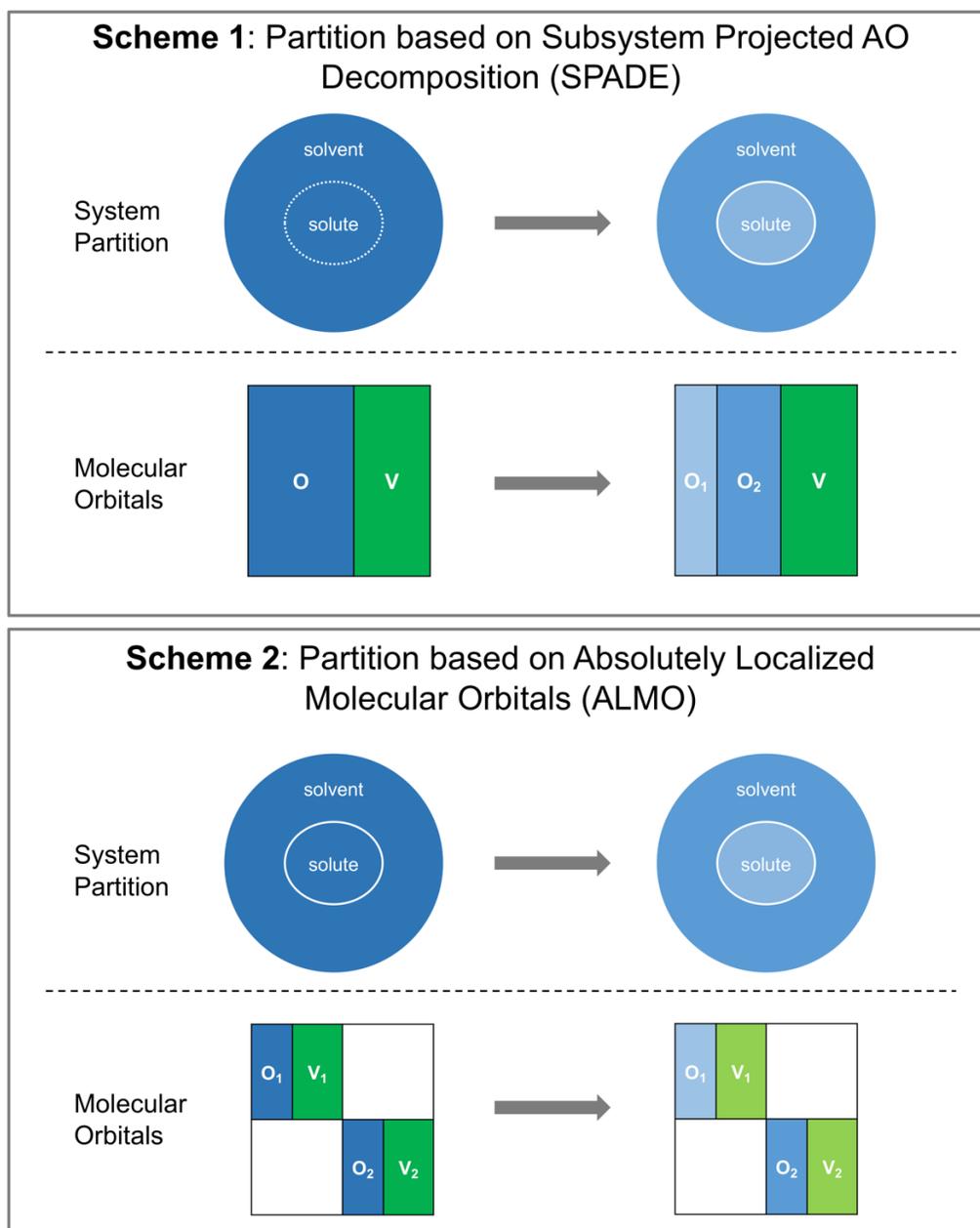
7



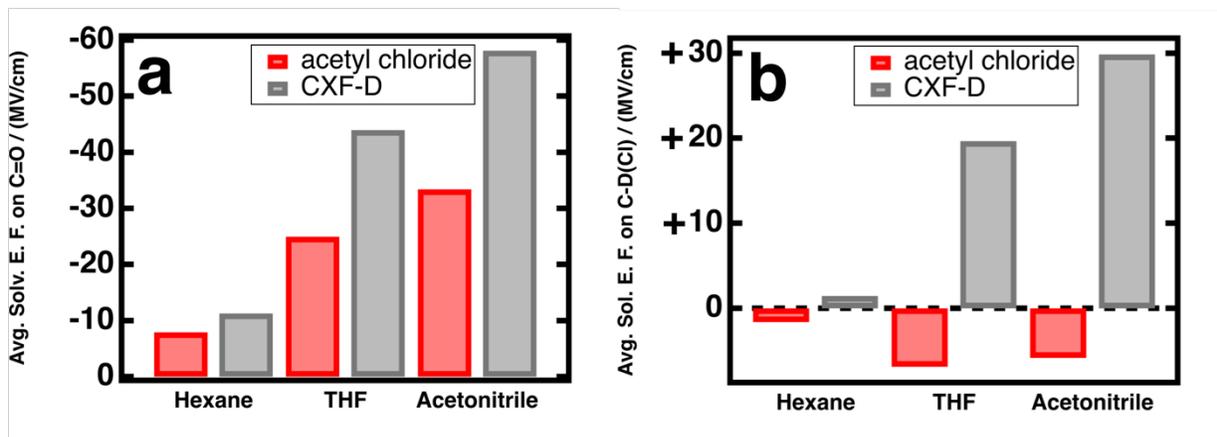
1  
 2 **Fig. 9. Unnormalized FTIR spectra of CXF-D in wild-type LADH.** (a) The absorption spectrum of  
 3 CXF-D's C-D in wild-type LADH. Bound and unbound CXF-D can be seen in the spectrum, and the  
 4 unbound peak could be completely subtracted, resulting in a blue-shifted C-D peak  
 5 corresponding to the bound CXF-D. (b) The difference absorption spectrum of C=O in CXF-D  
 6 minus that in N-[formyl-<sup>13</sup>C]-cyclohexylformamide (<sup>13</sup>C-CXF) in wild-type LADH. The frequency  
 7 shift between these two isotope-labeled species is ca. 25 cm<sup>-1</sup> (Supplementary Table 28). The  
 8 reason that the C=O frequency difference between CXF-D and <sup>13</sup>C-CXF is less than 40 cm<sup>-1</sup>, which  
 9 is the typical and expected isotope shift for a carbonyl group between <sup>12</sup>C- and <sup>13</sup>C-labeled species,  
 10 is that, substituting D for H on the aldehyde also causes some red shift (ca. 15-20 cm<sup>-1</sup>) in C=O  
 11 (Supplementary Table 28).



12  
 13 **Fig. 10. Illustration of the active site of the wild type LADH in complex with NADH and N-**  
 14 **cyclohexylformamide (CXF) based on the crystal structure resolved at 1.43 Å (PDB: 7RM6, see**  
 15 **Table 29).** The C=O of CXF forms a hydrogen bond with Ser48 and has a strong electrostatic  
 16 interaction with the catalytic Zn<sup>2+</sup> at the active site.

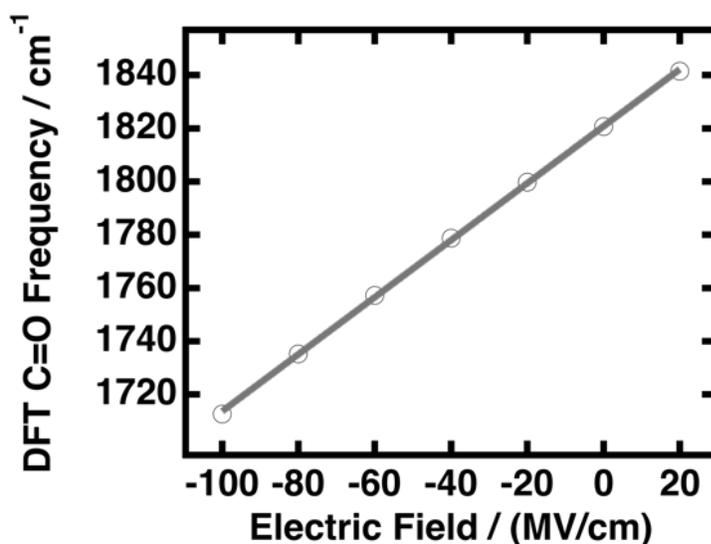


1  
 2 **Figure 11. Two electronic structure partitioning schemes that were employed for the QM**  
 3 **calculations of solvent electric fields exerted on the vibrational probes.** The first approach  
 4 (SPADE) is based on the singular value decomposition (SVD) of subsystem-projected molecular  
 5 orbitals (MOs), where the partition of the whole system is not invoked until the converged  
 6 canonical MOs are obtained (indicated by the contrast between dashed and solid lines that  
 7 separate solute and solvent on the top row of Scheme 1). The second approach is based on  
 8 absolutely localized molecular orbitals (ALMOs), where the system is partitioned into solute and  
 9 solvent fragments at the beginning and the MO coefficients that are being variationally optimized  
 10 remain fragment-block-diagonal. In this schematic illustration, different colors are used to  
 11 indicate changes in electronic wavefunctions (MOs), “O” and “V” denote occupied and virtual  
 12 orbitals, and “1” and “2” denote the solute and solvent fragments, respectively.



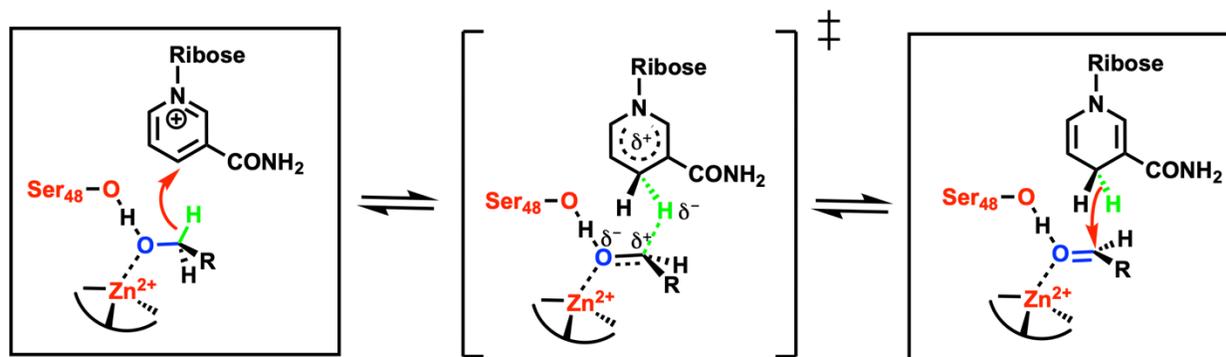
1  
 2 **Fig. 12. Electric fields acting on acetyl chloride and CXF-D in hexane, THF, and acetonitrile.** (a)  
 3 Solvent electric fields on C=O. (b) Solvent electric fields on C-D of CXF-D and C-Cl of acetyl chloride.  
 4 Data are from Supplementary Tables 6 and 23.

5



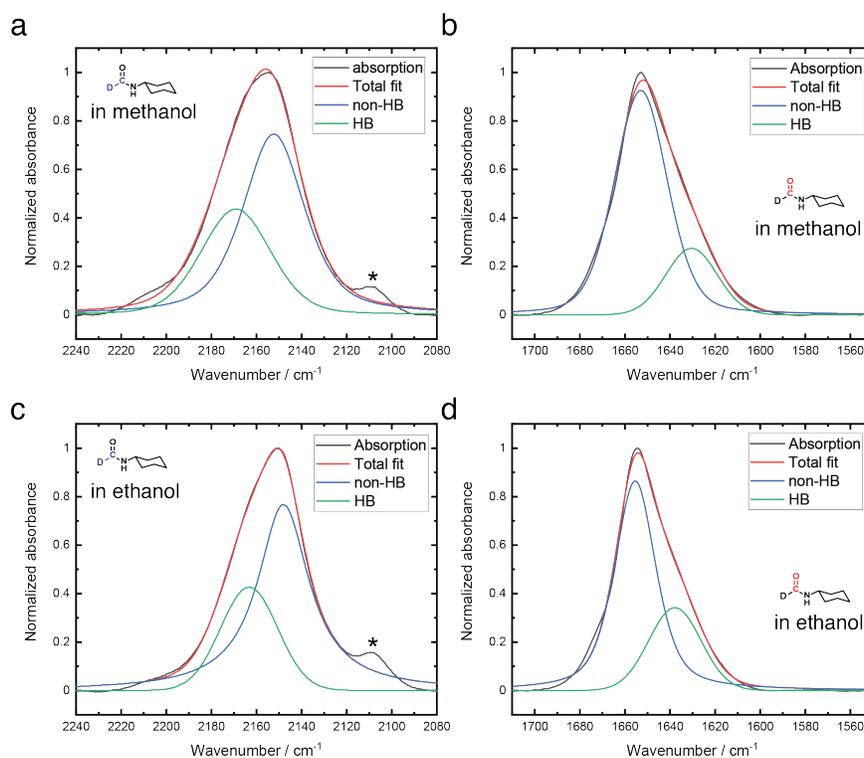
6  
 7 **Fig. 13. Vibrational Stark effect of acetyl chloride's C=O stretch mode revealed by "Stark in silico"**  
 8 **calculations performed in the gas phase.** A uniform external electric field was applied along the  
 9 bond axis of C=O, and *ab initio* frequency calculations were performed, at the B3LYP/6-31+G(d)  
 10 level with a scaling factor 0.964, to determine the vibrational frequencies under external electric  
 11 fields (the original data are shown in Supplementary Table 18). A linear fitting for the field-  
 12 frequency correlation of C=O yields  $\bar{\nu}_{C=O} = 1.07F_{C=O} + 1820.8$  ( $R^2 = 0.9998$ ).

13



1  
2  
3

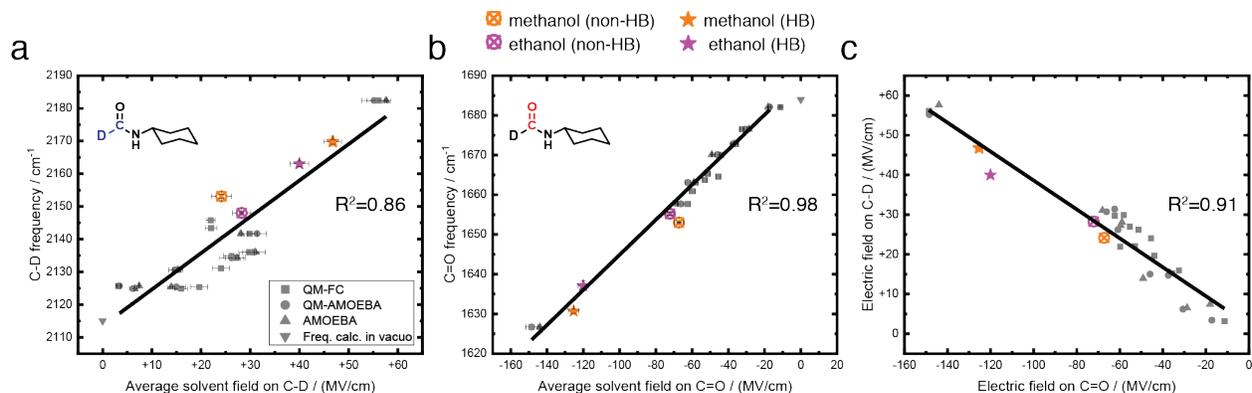
**Fig. 14. The reversible hydride transfer reaction catalyzed by liver alcohol dehydrogenase.**



4

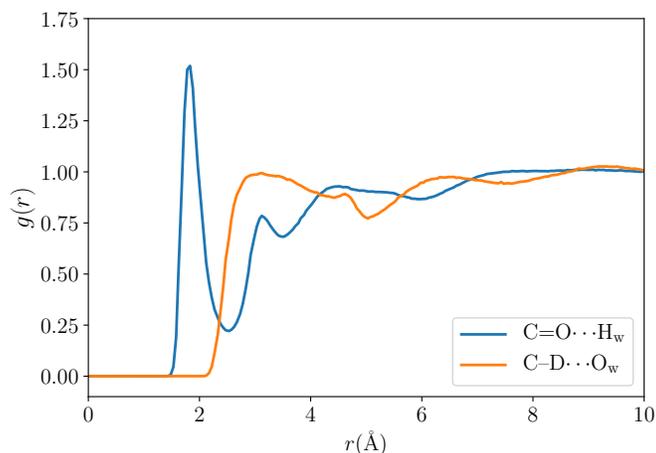
**Fig. 15. The IR spectra of CFX-D in methanol and ethanol.** (a) C-D of CFX-D in methanol. (b) C=O of CFX-D in methanol. (c) C-D of CFX-D in ethanol. (d) C=O of CFX-D in ethanol. The non-symmetric C=O and C-D IR bands indicate the heterogeneity of solvation environments in alcohols regarding the formation of H-bonds between CFX-D's C=O group and the OH groups of the alcohols. Here we fit these spectra with two peaks, which correspond to the HB (green) and non-HB (blue) populations. The asterisk-labeled peaks in (a) and (c) correspond to the Fermi resonance band. More details are shown in Supplementary Table 3.

12  
13



1  
 2 **Fig. 16. The field-frequency correlations and field-field correlation with the HB and non-HB**  
 3 **populations of CXF-D in methanol and ethanol taken into account.** Gray data points represent  
 4 the solvents shown in Fig. 4. Error bars represent the standard errors of electric field calculations.  
 5 (a) The field-frequency correlation of C-D in solvents. (b) The field-frequency correlation of C=O  
 6 in solvents. (c) The correlation between solvent electric fields on C=O and C-D.

7  
 8



9  
 10 **Fig. 17. Radial distribution functions (RDFs) for CXF-D in water obtained from 10ns fixed-charge**  
 11 **MD simulation.** Blue curve: RDF for the distance between the oxygen atom in CXF-D's C=O group  
 12 and the H atoms of solvent water molecules; orange curve: RDF for the distance between the  
 13 deuterium atom in CXF-D and water's O atoms.

14  
 15  
 16  
 17

# 1 Supplementary Tables

2 **Table 1. Vibrational Stark effect data**

Compound	Vibrational mode	Conc. M	Solvent	Peak position $\text{cm}^{-1}$	FWHM $\text{cm}^{-1}$	A <sup>a</sup>	B $\text{cm}^{-1}$	C $\text{cm}^{-2}$	$ \Delta\mu f$ $\text{cm}^{-1} / (\text{MV}/\text{cm})$
Methyl formate-D	C=O	0.1	2MeTHF	1684.6	8.3	$-4.58 \times 10^{-5}$	$-1.79 \times 10^{-3}$	$9.28 \times 10^{-2}$	0.96
	C-D	4	2MeTHF	2224.8	22.3	$-1.12 \times 10^{-6}$	$-5.88 \times 10^{-3}$	$8.39 \times 10^{-2}$	0.92
	C-D shoulder	4	2MeTHF	2192.0	15.5	$2.20 \times 10^{-3}$	$5.63 \times 10^{-3}$	$1.09 \times 10^{-1}$	1.04
N-[formyl- <sup>2</sup> H]cyclohexyl formamide	C=O	0.05	D <sub>2</sub> O:glycerol-D <sub>3</sub> (v/v 1/1)	1622.7	21.9	$-5.72 \times 10^{-5}$	$-2.25 \times 10^{-3}$	$1.27 \times 10^{-1}$	$1.13 \pm 0.01^b$
	C-D	1	2MeTHF	2157.1	37.1	$3.91 \times 10^{-5}$	$-8.91 \times 10^{-3}$	$1.67 \times 10^{-1}$	$1.32 \pm 0.04^b$
	C-D shoulder	1	2MeTHF	2110.2	12.9	$1.63 \times 10^{-5}$	$-9.24 \times 10^{-3}$	$3.20 \times 10^{-2}$	0.57
Acetyl chloride	C=O	0.1	2MeTHF	1810.1	14.7	$4.53 \times 10^{-5}$	$-8.02 \times 10^{-4}$	$3.24 \times 10^{-1}$	1.80

3 <sup>a</sup> A, B, and C are the fitting coefficients for the Stark spectra, corresponding to the 0<sup>th</sup> (A), 1<sup>st</sup> (B), and 2<sup>nd</sup> (C) derivative  
 4 contributions respectively<sup>4</sup>. The Stark tuning rates ( $|\Delta\mu|f$ ) can be extracted from the coefficient of the 2<sup>nd</sup> derivative  
 5 contribution (C) assuming the difference dipole of the vibrational mode is parallel with its transition dipole moment<sup>4</sup>.

6 <sup>b</sup> Error bars represent the standard deviations based on analyzing three separate VSS datasets.

7

8 **Table 2. FTIR data analysis of CFX-D**

Solvent	C=O		C-D	
	peak position $\text{cm}^{-1}$	FWHM $\text{cm}^{-1}$	peak position $\text{cm}^{-1}$	FWHM $\text{cm}^{-1}$
Hexane	1682.0	3.9	2126.4	18.0
Toluene	1672.9	10.7	2134.4	28.7
Dibutyl ether	1676.6	11.0	2123.8	20.4
Chloroform	1660.4	19.0	2146.1	21.6
THF	1670.1	14.4	2123.8	20.4
DCM	1664.1	16.7	2144.3	25.3
Cyclohexanone	1664.9	12.8	2126.3	39.1
Acetone	1665.1	16.6	2133.0	24.9
Acetonitrile	1662.5	14.7	2141.9	23.3
Dimethyl formamide	-	-	2132.0	27.3
DMSO	1658.7	23.0	2134.3	29.4
Water (or D <sub>2</sub> O)	1627.1	21.6	2183.1	32.7

9

10

11

1 **Table 3. FTIR data analysis of CXF-D by curve fit**

Solvent	C=O		C-D	
	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>
Hexane	1682.1	4.9	2125.7	20.8
Toluene	1672.8	13.1	2130.6	34.7
Dibutyl ether	1676.5	12.2	2124.9	32.1
Chloroform	1660.9	18.9	2145.8	21.5
THF	1670.1	15.3	2125.4	34.1
DCM	1663.7	16.9	2143.4	29.7
Cyclohexanone	1664.6	12.9	2131.1	40.4
Acetone	1665.3	18.1	2134.9	39.5
Acetonitrile	1663.1	18.4	2141.7	31.6
Dimethyl formamide	-	-	2134.3	40.3
DMSO	1657.7	23.8	2136.0	43.6
Water (or D <sub>2</sub> O)	1626.7	22.1	2182.4	34.6
Methanol (non-HB)	1653.0	27.3	2153.1	33.5
Methanol (HB)	1630.7	26.2	2169.8	37.4
Ethanol (non-HB)	1655.2	21.9	2148.0	27.7
Ethanol (HB)	1637.1	26.3	2163.0	30.5

2

3

4 **Table 4. FTIR data analysis of methyl formate-D**

solvent	C=O		C-D	
	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>
hexane	1702.9	11.2	2209.5	15.0
toluene	1696.8	8.9	2211.6	15.7
dibutyl ether	1700.0	11.9	2210.2	18.5
chloroform	1693.8	11.8	2216.6	13.2
THF	1694.3	11.6	2216.0	19.6
DCM	1693.9	9.1	2216.8	14.2
cyclohexanone	-	-	2217.4	20.0
acetone	1693.2	10.0	2220.0	18.1
acetonitrile	1694.0	9.3	-	-
dimethyl formamide	-	-	2219.4	43.6
DMSO	1685.8	13.7	2219.4	17.6
Water (or D <sub>2</sub> O)	1683.5	20.1	2233.0	16.8

1 **Table 5. FTIR data analysis of methyl formate-D by curve fit**

solvent	C=O		C-D	
	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>
hexane	1703.1	13.5	2209.6	18.6
toluene	1696.8	9.4	2211.9	19.1
dibutyl ether	1699.5	12.6	2211.0	23.2
chloroform	1693.2	11.9	2216.6	16.7
THF	1694.2	13.5	2216.2	24.7
DCM	1693.8	9.3	2216.7	17.0
cyclohexanone	-	-	2217.2	26.6
acetone	1692.4	9.8	2220.1	22.6
acetonitrile	1694.0	9.7	-	-
dimethyl formamide	-	-	2219.6	24.5
DMSO	1686.0	14.1	2219.6	21.4
Water (or D <sub>2</sub> O)	1683.0	20.6	2233.5	22.6

2

3

4

5

6

7

8

9

10

11

12

1 **Table 6. Solvent electric fields projected on the C=O and C-D in CXF-D computed using the**  
 2 **SPADE QM partitioning method, solute-solvent configurations sampled from fixed-charge MD,**  
 3 **and the averages of atomic electric field projections on each bond.**

Solvent	C=O			C-D		
	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Standard error (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Standard error (MV/cm)
Hexane	-11.30	3.73	0.373	+3.18	3.43	0.343
Toluene	-35.68	8.45	0.845	+15.30	10.29	1.029
Dibutyl ether	-32.57	8.61	0.861	+15.96	13.33	1.333
Chloroform	-59.82	16.82	1.683	+21.96	9.40	0.94
THF	-43.96	10.44	1.044	+19.66	16.84	1.684
DCM	-52.98	12.96	1.296	+22.06	11.80	1.18
Cyclohexanone	-45.54	13.51	1.351	+24.06	17.32	1.732
Acetone	-51.30	12.37	1.237	+26.20	15.91	1.591
Acetonitrile	-58.11	13.91	1.391	+29.87	15.65	1.565
DMF	-55.26	15.88	1.588	+26.99	20.20	2.02
DMSO	-62.25	14.25	1.425	+29.73	20.67	2.067
Water	-148.63	33.68	3.368	+56.11	24.84	2.484
Methanol (non-HB)	-67.22	20.76	2.68	+24.14	15.69	2.026
Methanol (HB)	-125.31	29.79	2.979	+46.78	18.20	1.82
Ethanol (non-HB)	-72.14	13.09	1.734	+28.27	14.05	1.861
Ethanol (HB)	-120.06	28.81	2.881	+39.99	18.88	1.888

4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15

1 **Table 7. Solvent electric fields projected on the C=O and C-D in CXF-D computed using the**  
 2 **SPADE QM partitioning method, solute-solvent configurations sampled from AMOEBA**  
 3 **polarizable MD, and the averages of atomic electric field projections on each bond.**

4

Solvent	C=O			C-D		
	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Standard error (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Standard error (MV/cm)
Hexane	-17.20	6.25	0.625	+3.43	3.6	0.36
Toluene	-37.40	9.97	0.997	+14.70	12.47	1.247
Dibutyl ether	-30.58	9.13	0.913	+6.20	7.67	0.767
THF	-45.91	12.74	1.274	+15.01	18.29	1.829
Acetonitrile	-62.34	16.58	1.658	+31.39	19.26	1.926
DMF	-60.79	15.37	1.537	+26.20	24.83	2.483
DMSO	-66.14	15.92	1.592	+30.72	23.32	2.332
Water	-148.45	32.22	3.222	+55.19	25.57	2.557

5  
 6  
 7 **Table 8. Solvent electric fields projected on the C=O and C-D in CXF-D computed using the**  
 8 **SPADE QM partitioning method, solute-solvent configurations sampled from fixed-charge MD,**  
 9 **and the gradient of electrostatic potential along each bond.**

Solvent	C=O		C-D	
	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-10.17	3.37	+0.34	3.48
Toluene	-34.61	8.38	+14.11	9.86
Dibutyl ether	-30.69	8.31	+14.78	12.59
Chloroform	-57.67	16.46	+20.69	9.53
THF	-42.00	10.33	+19.12	16.51
DCM	-51.16	12.62	+21.00	11.63
Cyclohexanone	-43.79	13.32	+23.32	16.80
Acetone	-49.55	12.11	+26.15	15.77
Acetonitrile	-57.14	14.12	+29.94	15.61
DMF	-53.89	15.66	+26.09	19.14
DMSO	-59.80	14.41	+28.65	19.90
Water	-144.01	32.79	+55.26	24.26

10

11

1 **Table 9. Solvent electric fields projected on the C=O and C-D in CXF-D computed using the**  
 2 **SPADE QM partitioning method, solute-solvent configurations sampled from AMOEBA**  
 3 **polarizable MD, and the gradient of electrostatic potential along each bond.**

Solvent	C=O		C-D	
	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-15.47	5.70	+1.98	3.49
Toluene	-35.53	9.57	+13.40	12.08
Dibutyl ether	-28.10	8.53	+5.04	7.66
THF	-43.44	12.61	+13.67	17.30
Acetonitrile	-60.70	16.11	+29.69	18.44
DMF	-57.81	15.60	+23.52	23.30
DMSO	-63.43	15.38	+28.27	21.73
Water	-145.26	31.14	+53.26	24.16

4

5

6 **Table 10. Solvent electric fields projected on the C=O and C-D in CXF-D computed using the**  
 7 **ALMO QM partitioning method, solute-solvent configurations sampled from fixed-charge MD,**  
 8 **and the averages of atomic electric field projections on each bond.**

Solvent	C=O		C-D	
	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-10.24	3.28	+3.63	3.47
Toluene	-34.95	8.08	+18.27	10.40
Dibutyl ether	-31.00	8.29	+19.31	13.70
Chloroform	-57.29	16.20	+24.28	9.56
THF	-42.45	10.42	+22.84	16.82
DCM	-50.68	12.55	+24.47	11.77
Cyclohexanone	-43.87	13.27	+26.94	17.49
Acetone	-49.58	12.00	+29.12	15.99
Acetonitrile	-56.63	13.63	+32.39	15.89
DMF	-53.22	15.34	+30.15	20.25
DMSO	-59.46	14.09	+33.75	20.85
Water	-141.19	31.60	+59.79	24.96

9

10

11

1 **Table 11. Solvent electric fields projected on the C=O and C-D in CXF-D computed using the**  
 2 **ALMO QM partitioning method, solute-solvent configurations sampled from AMOEBA**  
 3 **polarizable MD, and the averages of atomic electric field projections on each bond.**

Solvent	C=O		C-D	
	Average solvent electric field / (MV/cm)	Standard deviation / (MV/cm)	Average solvent electric field / (MV/cm)	Standard deviation / (MV/cm)
Hexane	-14.85	5.34	+5.58	3.97
Toluene	-35.40	9.49	+16.62	12.48
Dibutyl ether	-28.11	8.25	+9.57	7.50
THF	-41.70	11.79	+17.88	19.36
Acetonitrile	-59.33	15.76	+33.98	19.44
DMF	-56.56	14.65	+30.07	24.78
DMSO	-61.05	16.11	+31.88	22.88
Water	-148.09	30.57	+59.23	25.95

4

5

6 **Table 12. Solvent electric fields projected on the C=O and C-D in CXF-D computed via AMOEBA**  
 7 **polarizable MD using the intrinsic parameters of the AMOEBA force field.**

Solvent	C=O			C-D		
	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Standard error (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Standard error (MV/cm)
Hexane	-18.00	6.83	0.068	+7.43	5.77	0.058
Toluene	-36.67	10.38	0.1	+15.28	11.55	0.12
Dibutyl ether	-28.70	8.23	0.059	+6.53	11.09	0.08
THF	-49.11	11.30	0.14	+13.97	19.83	0.24
Acetonitrile	-58.95	14.47	0.14	+28.09	19.60	0.2
DMF	-59.61	13.42	0.13	+27.29	24.30	0.24
DMSO	-68.03	14.86	0.12	+31.07	24.39	0.2
Water	-143.96	27.89	0.28	+57.62	21.66	0.22

8

9

10

11

12

1 **Table 13. Solvent electric fields projected on the C=O and C-D in CXF-D computed via fixed-**  
 2 **charge MD using the intrinsic parameters of GAFF.**

Solvent	C=O		C-D	
	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-0.22	0.85	+0.25	1.07
Toluene	-25.69	7.83	+16.18	11.92
Dibutyl ether	-18.30	7.55	+10.86	10.59
Chloroform	-44.16	13.22	+21.10	9.03
THF	-31.94	9.51	+20.26	14.55
DCM	-45.72	12.25	+25.57	11.24
Cyclohexanone	-35.30	11.15	+23.22	17.18
Acetone	-43.20	11.08	+27.63	16.76
Acetonitrile	-47.20	11.93	+31.64	16.36
DMF	-45.24	12.41	+28.88	18.01
DMSO	-54.40	12.62	+39.38	19.22
Water	-108.98	21.69	+51.80	17.30

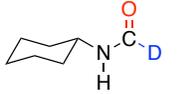
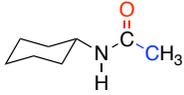
3

4

5 **Table 14. Vibrational Stark effect of CXF-D's C=O and C-D stretch modes revealed by “Stark *in***  
 6 ***silico*” calculations performed in the gas phase. A uniform external electric field was applied**  
 7 **along the bond axis of C=O or C-D, whose vibrational frequency was then calculated at the**  
 8 **B3LYP/6-31+G(d) level using 0.964 as the scaling factor.**

Electric Field on C=O (MV/cm)	DFT C=O Frequency (cm <sup>-1</sup> )	Electric Field on C-D (MV/cm)	DFT C-D Frequency (cm <sup>-1</sup> )
+20	1697.78	+80	2200.57
0	1687.09	+60	2186.01
-20	1674.35	+40	2167.96
-40	1661.00	+20	2146.23
-60	1647.37	0	2119.92
-80	1633.55	-20	2090.01
-100	1619.68		
-120	1605.77		
-140	1592.28		
-160	1578.46		

1 **Table 15. Electric fields experienced by CXF-D and *N*-cyclohexylacetamide in water computed**  
 2 **using the SPADE QM partitioning method, solute-solvent configurations sampled from fixed-**  
 3 **charge MD, and the averages of atomic electric field projections on each bond.**

Structure	Compound name	C=O		C-D/C-C	
		Average water field (MV/cm)	STDEV (MV/cm)	Average water field (MV/cm)	STDEV (MV/cm)
	<i>N</i> -[formyl- <sup>2</sup> H]cyclohexylformamide	-148.63	33.68	+56.11	24.84
	<i>N</i> -cyclohexyl acetamide	-147.36	41.20	+44.35	21.10

4  
5  
6  
7  
8  
9

10 **Table 16. Vibrational frequencies of the C=O and C-D stretches in CXF-D in solvents computed**  
 11 **at the B3LYP/6-31+G(d)/C-PCM level of theory using 0.964 as the scaling factor.**

Solvent	Dielectric constant	C=O frequency cm <sup>-1</sup>	C-D frequency cm <sup>-1</sup>
Hexane	1.88	1656.05	2137.03
Toluene	2.37	1647.08	2141.68
Dibutyl ether	3.05	1638.82	2145.93
Chloroform	4.71	1638.82	2145.93
THF	7.43	1619.85	2155.54
DCM	8.82	1617.59	2156.68
Cyclohexanone	15.62	1612.21	2159.35
Acetone	20.49	1610.51	2160.20
Acetonitrile	37.50	1607.98	2161.46
DMF	37.22	1608.01	2161.44
DMSO	46.83	1607.37	2161.76
Water	78.40	1606.38	2162.26

12  
13

1 **Table 17. FTIR of the C=O of acetyl chloride**

Solvent	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>	Curve fit cm <sup>-1</sup>	FWHM cm <sup>-1</sup>
Hexane	1809.2	10.6	1809.2	10.9
THF	1809.0	14.8	1808.8	15.3
Acetonitrile	1807.0	17.4	1807.1	18.2

2

3

4 **Table 18. Vibrational Stark effect of acetyl chloride's C=O stretch mode revealed by "Stark *in***  
 5 ***silico*"** calculations performed in the gas phase. A uniform external electric field was applied  
 6 **along the bond axis of C=O, whose vibrational frequency was then calculated at the B3LYP/6-**  
 7 **31+G(d) level using 0.964 as the scaling factor.**

Electric Field on C=O (MV/cm)	DFT C=O Frequency cm <sup>-1</sup>
+20	1841.47
0	1820.71
-20	1799.77
-40	1778.71
-60	1757.25
-80	1735.27
-100	1712.65

8

9

10 **Table 19. FTIR of CXF-D (C=O) in wild-type LADH (backgrounded by N-[formyl-<sup>13</sup>C]-**  
 11 **cyclohexylformamide, denoted as <sup>13</sup>C-CXF) and D<sub>2</sub>O**

	CXF-D				<sup>13</sup> C-CXF			
	Peak picking		Curve fit		Peak picking		Curve fit	
	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>						
WT <sup>a</sup>	1618.4 ± 0.1	5.0 ± 0.2	1618.7 ± 0.1	5.3 ± 0.2	1592.6 ± 0.1	13.7 ± 0.1	1592.3 ± 0.1	11.0 ± 0.4
D <sub>2</sub> O	1627.1	20.9	1626.5	22.1	1605.2	21.1	1604.8	21.7

12 <sup>a</sup> Values and errors are the mean and standard deviation over multiple repeats

13

1 **Table 20. FTIR of CXF-D (C-D) in wild-type LADH and H<sub>2</sub>O**

	Peak picking		Curve fit	
	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>
WT <sup>a</sup>	2218.3 ± 0.1	12.4 ± 0.3	2218.3 ± 0.3	13.4 ± 0.6
H <sub>2</sub> O	2182.2	32.3	2182.8	34.0
Hexane	2127.9	5.3	2125.7	20.8

2 <sup>a</sup> Values and errors are the mean and standard deviation over multiple repeats.

3  
4 **Table 21. Solvent electric fields projected on the C=O and C-D in CXF-D (averaged over 100**  
5 **frames) calculated using the SPADE QM partitioning approach with two different cutoff radii**  
6 **(7Å and 8Å) for the generation of truncated solute-solvent clusters from fixed-charge MD**  
7 **trajectories.**

Solvent	Cutoff radii	C=O		C-D	
		Average solvent electric field (MV/cm)	Standard deviation (MV/cm)	Average solvent electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	7Å	-11.30	3.73	+1.44	3.43
	8Å	-11.55	3.74	+1.58	3.46
Toluene	7Å	-35.69	8.45	+15.30	10.29
	8Å	-35.96	8.18	+15.50	10.21
DMSO	7Å	-62.25	14.25	+29.73	20.67
	8Å	-63.70	13.20	+31.37	21.27
Water	7Å	-145.05	32.82	+52.54	23.44
	8Å	-146.64	32.64	+54.40	22.87

8  
9 **Table 22. Vibrational frequencies of acetyl chloride's C=O stretch mode in solvents computed**  
10 **at the B3LYP/6-31+G(d)/C-PCM level of theory using 0.964 as the scaling factor.**

Solvent	Dielectric constant	C=O frequency cm <sup>-1</sup>
Hexane	1.88	1806.6
THF	7.43	1790.8
Acetonitrile	37.50	1785.8

11

1 **Table 23. Solvent electric fields projected on the C=O and C-Cl in acetyl chloride computed using**  
 2 **the SPADE QM partitioning method, solute-solvent configurations sampled from fixed-charge**  
 3 **MD, and the averages of atomic electric field projections on each bond.**

Solvent	C=O		C-Cl	
	Average electric field (MV/cm)	Standard deviation (MV/cm)	Average electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-7.91	3.51	-1.63	2.3
THF	-24.97	8.42	-6.90	7.19
Acetonitrile	-33.43	12.96	-5.83	10.47

4

5 **Table 24. Solvent electric fields projected on the C=O and C-Cl in acetyl chloride computed using**  
 6 **the SPADE QM partitioning method, solute-solvent configurations sampled from AMOEBA**  
 7 **polarizable MD, and the averages of atomic electric field projections on each bond.**

Solvent	C=O		C-Cl	
	Average electric field (MV/cm)	Standard deviation (MV/cm)	Average electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-10.40	4.48	-2.64	2.58
THF	-23.47	12.43	-10.82	10.33
Acetonitrile	-31.36	14.60	-9.73	13.40

8

9 **Table 25. Solvent electric fields projected on the C=O and C-Cl in acetyl chloride calculated via**  
 10 **AMOEBA polarizable MD simulations using the intrinsic parameters of the AMOEBA force field.**

Solvent	C=O		C-Cl	
	Average electric field (MV/cm)	Standard deviation (MV/cm)	Average electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-8.82	4.99	-2.69	3.6
THF	-21.04	10.16	-10.96	11.21
Acetonitrile	-27.88	12.70	-8.53	12.63

11

12 **Table 26. Solvent electric fields projected on the C=O and C-Cl in acetyl chloride calculated via**  
 13 **fixed-charge MD simulations using the intrinsic parameters of the GAFF force field.**

Solvent	C=O		C-Cl	
	Average electric field (MV/cm)	Standard deviation (MV/cm)	Average electric field (MV/cm)	Standard deviation (MV/cm)
Hexane	-0.060	0.79	+0.030	0.68
THF	-19.48	8.94	-3.57	8.57
Acetonitrile	-29.89	11.37	-3.14	10.22

14

1 **Table 27. Assignment of atomic charges to the aldehyde moiety of CFX-D based on several**  
 2 **charge population analysis schemes based on QM calculations at the B3LYP/6-31+G(d) level.**

Scheme	C	D	O
Mulliken	+0.328	+0.143	-0.480
Iterative Hirshfeld	+0.445	+0.053	-0.504
ChELPG	+0.557	+0.008	-0.570
Merz-Kollman ESP	+0.569	+0.021	-0.569
CM5	+0.150	+0.121	-0.352

3

4 **Table 28. FTIR of isotope-labeled cyclohexylformamide compounds.**

Compound	Peak picking		Curve fit	
	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>	Peak position cm <sup>-1</sup>	FWHM cm <sup>-1</sup>
<i>N</i> -cyclohexylformamide	1645.4	21.0	1645.0	21.8
<i>N</i> -[formyl- <sup>2</sup> H]-cyclohexylformamide (CFX-D)	1627.1	20.9	1626.5	22.1
<i>N</i> -[formyl- <sup>13</sup> C]-cyclohexylformamide ( <sup>13</sup> C-CXF)	1605.2	21.1	1604.8	21.7

5

6

7 **Table 29. X-ray diffraction data collection and refinement statistics<sup>a</sup>**

Protein	Wild type LADH complexed with NADH and CFX
PDB entry	7RM6
<b>Data collection statistics</b>	
Beamline	BL 12-2
Wavelength (Å)	0.979
Resolution range (Å)	37.94 – 1.43 (1.48 – 1.43)
Space group	P1
Unit cell dimensions a, b, c, (Å) α, β, γ (°)	44.14, 50.98, 92.86 92.41, 103.03, 108.94
Matthews coefficient	2.25
Solvent content (%)	45.40
Total observations	932,809 (43,206)

Unique observations	119,395 (5,789)
Multiplicity	7.8 (7.5)
Completeness (%)	87.2(85.5)
Mean I/ $\sigma$ (I)	19.7 (2.22)
Wilson B-factor ( $\text{\AA}^2$ )	15.5
Anisotropy	0.736
R <sub>merge</sub>	0.050 (1.522)
R <sub>meas</sub>	0.058 (1.771)
R <sub>pim</sub>	0.029 (0.901)
CC <sub>1/2</sub>	0.999 (0.754)
<b>Refinement statistics</b>	
Reflections used	117,399 (11,104)
Reflections used for R <sub>free</sub>	5,869 (555)
R <sub>work</sub>	0.190 (0.441)
R <sub>free</sub>	0.208 (0.458)
Number of non-H atoms	6466
Protein	5579
Ligand	110
Solvent	777
Protein residues	756
RMSD bond lengths ( $\text{\AA}$ )	0.030
RMSD bond angles ( $^\circ$ )	1.81
Ramachandran favored (%)	96.77
Ramachandran allowed (%)	3.23
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	1.94
Clashscore	3.22
Average B factor ( $\text{\AA}^2$ )	28.56
Protein	27.41
Ligand	22.78
Solvent	37.63

1 <sup>a</sup>Statistics for the highest-resolution shell are shown in parentheses.

2

3

1 **Table 30. Vibrational Stark effects of CXF-D's C=O and C-D stretch modes investigated using**  
 2 **"Stark *in silico*" calculations performed in vacuum and two solvents (2-MeTHF and**  
 3 **D<sub>2</sub>O:glycer(ol-D<sub>3</sub>) mixture (v/v = 1:1)) described by the C-PCM model. The vibrational**  
 4 **frequencies were calculated at the B3LYP/6-31+G(d) level and then scaled by 0.964, and the**  
 5 **values of  $|\Delta\mu_{CO}|$  and  $|\Delta\mu_{CD}|$  were determined using the frequencies calculated with the**  
 6 **electric field on and off. The local field factors were estimated as the ratio of the Stark tuning**  
 7 **rate in solvent to that in vacuum.**

	Vacuum	2-MeTHF ( $\epsilon = 6.97$ )	D <sub>2</sub> O:glycer(ol-D <sub>3</sub> ) (v/v = 1:1) ( $\epsilon = 62.0$ )
$\nu_{CO} (F_{CO} = 0)$ cm <sup>-1</sup>	1687.1	1620.9	1606.9
$\nu_{CO}(F_{CO} = -20 \text{ MV/cm})$ cm <sup>-1</sup>	1674.4	1601.7	1586.7
$ \Delta\mu_{CO} $ cm <sup>-1</sup> /(MV/cm)	0.64	0.96	1.01 <sup>a</sup>
Estimated $f_{CO}$	1.00	1.50	1.58
$\nu_{CD} (F_{CD} = 0)$ cm <sup>-1</sup>	2119.9	2155.3	2162.4
$\nu_{CD}(F_{CD} = -20 \text{ MV/cm})$ cm <sup>-1</sup>	2146.2	2179.8	2186.3
$ \Delta\mu_{CD} $ cm <sup>-1</sup> /(MV/cm)	1.32	1.22 <sup>b</sup>	1.20
Estimated $f_{CD}$	1.00	0.93	0.91

8 <sup>a</sup> Experimental value measured in the D<sub>2</sub>O/glycer(ol-D<sub>3</sub>) mixture (v/v = 1:1): 1.13 cm<sup>-1</sup>/(MV/cm)

9 <sup>b</sup> Experimental value measured in 2-MeTHF: 1.32 cm<sup>-1</sup>/(MV/cm)

10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23

1 **Table 31. Fermi resonance analysis for the C-D vibration of CXF-D.**

Solvent	$\nu_1^a / \text{cm}^{-1}$	$\nu_2 / \text{cm}^{-1}$	$\Delta / \text{cm}^{-1}$	R	W / $\text{cm}^{-1}$
Hexane	2125.7	2102.1	23.6	2.2	10.9
Toluene	2130.6	2101.2	29.4	12.6	7.7
Dibutyl Ether	2124.9	2100.8	24.1	5.7	8.6
Chloroform	2145.8	2103.6	42.2	29.8	7.5
THF	2127.0	2100.7	26.3	8.7	8.0
DCM	2143.5	2105.7	37.8	12.6	9.8
Cyclohexanone	2131.1	2101.2	29.9	6.2	10.4
Acetone	2134.9	2103.5	31.4	6.9	10.5
Acetonitrile	2144.1	2106.1	38.0	6.6	12.8
Dimethyl Formamide	2134.3	2101.6	32.7	4.1	13.0
DMSO	2136.9	2103.8	33.1	17.1	7.5

2 <sup>a</sup> $\nu_1$  refers to the perturbed frequency of C-D vibration and  $\nu_2$  refers to the perturbed frequency of the combination/overtone  
 3 mode that is in Fermi resonance with C-D.

4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18

# References

1. Deng H, Schindler JF, Berst KB, Plapp BV, Callender R. A Raman spectroscopic characterization of bonding in the complex of horse liver alcohol dehydrogenase with NADH and N-cyclohexylformamide. *Biochemistry* 1998, **37**(40): 14267-14278.
2. Boxer SG. Stark realities. *J Phys Chem B* 2009, **113**(10): 2972-2983.
3. Fried SD, Bagchi S, Boxer SG. 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.
4. Bublitz GU, Boxer SG. Stark spectroscopy: applications in chemistry, biology, and materials science. *Annu Rev Phys Chem* 1997, **48**: 213-242.
5. Schneider SH, Boxer SG. Vibrational Stark Effects of Carbonyl Probes Applied to Reinterpret IR and Raman Data for Enzyme Inhibitors in Terms of Electric Fields at the Active Site. *J Phys Chem B* 2016, **120**(36): 9672-9684.
6. Kozuch J, Schneider SH, Boxer SG. Biosynthetic Incorporation of Site-Specific Isotopes in beta-Lactam Antibiotics Enables Biophysical Studies. *ACS Chem Biol* 2020, **15**(5): 1148-1153.
7. Schneider SH, Kozuch J, Boxer SG. The Interplay of Electrostatics and Chemical Positioning in the Evolution of Antibiotic Resistance in TEM  $\beta$ -Lactamases. *bioRxiv* 2021: 2021.2005.2027.446023.
8. Park DH, Plapp BV. Isoenzymes of horse liver alcohol dehydrogenase active on ethanol and steroids. cDNA cloning, expression, and comparison of active sites. *J Biol Chem* 1991, **266**(20): 13296-13302.
9. Bahnson BJ, Colby TD, Chin JK, Goldstein BM, Klinman JP. A link between protein structure and enzyme catalyzed hydrogen tunneling. *Proc Natl Acad Sci U S A* 1997, **94**(24): 12797-12802.
10. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, *et al.* Gaussian 16, Revision C.01. Wallingford CT: Gaussian, Inc.; 2016.

- 1 11. D.A. Case, R.M. Betz, D.S. Cerutti, T.E. Cheatham I, T.A. Darden, R.E. Duke, *et al.* AMBER  
2 2016. University of California, San Francisco; 2016.
- 3
- 4 12. Caleman C, van Maaren PJ, Hong M, Hub JS, Costa LT, van der Spoel D. Force Field  
5 Benchmark of Organic Liquids: Density, Enthalpy of Vaporization, Heat Capacities,  
6 Surface Tension, Isothermal Compressibility, Volumetric Expansion Coefficient, and  
7 Dielectric Constant. *J Chem Theory Comput* 2012, **8**(1): 61-74.
- 8
- 9 13. Hess B, Kutzner C, van der Spoel D, Lindahl E. GROMACS 4: Algorithms for Highly  
10 Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput*  
11 2008, **4**(3): 435-447.
- 12
- 13 14. Berendsen HJC, Postma JPM, van Gunsteren WF, DiNola A, Haak JR. Molecular dynamics  
14 with coupling to an external bath. *J Chem Phys* 1984, **81**(8): 3684-3690.
- 15
- 16 15. Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular  
17 dynamics method. *J Appl Phys* 1981, **52**(12): 7182-7190.
- 18
- 19 16. Kozuch J, Schneider SH, Zheng C, Ji Z, Bradshaw RT, Boxer SG. Testing the Limitations of  
20 MD-Based Local Electric Fields Using the Vibrational Stark Effect in Solution: Penicillin G  
21 as a Test Case. *J Phys Chem B* 2021, **125**(17): 4415-4427.
- 22
- 23 17. Ponder JW, Wu C, Ren P, Pande VS, Chodera JD, Schnieders MJ, *et al.* Current status of  
24 the AMOEBA polarizable force field. *J Phys Chem B* 2010, **114**(8): 2549-2564.
- 25
- 26 18. Fried SD, Wang LP, Boxer SG, Ren P, Pande VS. Calculations of the electric fields in liquid  
27 solutions. *J Phys Chem B* 2013, **117**(50): 16236-16248.
- 28
- 29 19. Beeman D. Some multistep methods for use in molecular dynamics calculations. *J*  
30 *Comput Phys* 1976, **20**(2): 130-139.
- 31
- 32 20. Bussi G, Donadio D, Parrinello M. Canonical sampling through velocity rescaling. *J Chem*  
33 *Phys* 2007, **126**(1): 014101.
- 34
- 35 21. Claudino D, Mayhall NJ. Automatic Partition of Orbital Spaces Based on Singular Value  
36 Decomposition in the Context of Embedding Theories. *J Chem Theory Comput* 2019,  
37 **15**(2): 1053-1064.
- 38

- 1 22. Manby FR, Stella M, Goodpaster JD, Miller TF, 3rd. A Simple, Exact Density-Functional-  
2 Theory Embedding Scheme. *J Chem Theory Comput* 2012, **8**(8): 2564-2568.
- 3
- 4 23. Lee SJR, Welborn M, Manby FR, Miller TF, 3rd. Projection-Based Wavefunction-in-DFT  
5 Embedding. *Acc Chem Res* 2019, **52**(5): 1359-1368.
- 6
- 7 24. Waldrop JM, Windus TL, Govind N. Projector-Based Quantum Embedding for Molecular  
8 Systems: An Investigation of Three Partitioning Approaches. *J Phys Chem A* 2021.
- 9
- 10 25. Khaliullin RZ, Head-Gordon M, Bell AT. An efficient self-consistent field method for large  
11 systems of weakly interacting components. *J Chem Phys* 2006, **124**(20): 204105.
- 12
- 13 26. Mao Y, Ge Q, Horn PR, Head-Gordon M. On the Computational Characterization of  
14 Charge-Transfer Effects in Noncovalently Bound Molecular Complexes. *J Chem Theory  
15 Comput* 2018, **14**(5): 2401-2417.
- 16
- 17 27. Mao Y, Loipersberger M, Horn PR, Das A, Demerdash O, Levine DS, *et al.* From  
18 Intermolecular Interaction Energies and Observable Shifts to Component Contributions  
19 and Back Again: A Tale of Variational Energy Decomposition Analysis. *Annu Rev Phys  
20 Chem* 2021, **72**: 641-666.
- 21
- 22 28. Epifanovsky E, Gilbert ATB, Feng X, Lee J, Mao Y, Mardirossian N, *et al.* Software for the  
23 frontiers of quantum chemistry: An overview of developments in the Q-Chem 5  
24 package. *J Chem Phys* 2021, **155**(8).
- 25
- 26 29. Frisch MJ, Pople JA, Binkley JS. Self - consistent molecular orbital methods 25.  
27 Supplementary functions for Gaussian basis sets. *J Chem Phys* 1984, **80**(7): 3265-3269.
- 28
- 29 30. Becke AD. Density - functional thermochemistry. III. The role of exact exchange. *J Chem  
30 Phys* 1993, **98**(7): 5648-5652.
- 31
- 32 31. Johnson RD, III NIST Computational Chemistry Comparison and Benchmark Database,  
33 NIST Standard Reference Database Number 101, Release 21. August 2020 ed; 2020.
- 34
- 35 32. Truong TN, Stefanovich EV. A new method for incorporating solvent effect into the  
36 classical, ab initio molecular orbital and density functional theory frameworks for  
37 arbitrary shape cavity. *Chem Phys Lett* 1995, **240**(4): 253-260.
- 38

- 1 33. Barone V, Cossi M. Quantum Calculation of Molecular Energies and Energy Gradients in  
2 Solution by a Conductor Solvent Model. *J Phys Chem A* 1998, **102**(11): 1995-2001.
- 3
- 4 34. Cossi M, Rega N, Scalmani G, Barone V. Energies, structures, and electronic properties of  
5 molecules in solution with the C-PCM solvation model. *J Comput Chem* 2003, **24**(6): 669-  
6 681.
- 7
- 8 35. Bondi A. van der Waals Volumes and Radii. *J Phys Chem* 2002, **68**(3): 441-451.
- 9
- 10 36. Lange AW, Herbert JM. Polarizable Continuum Reaction-Field Solvation Models  
11 Affording Smooth Potential Energy Surfaces. *J Phys Chem Lett* 2009, **1**(2): 556-561.
- 12
- 13 37. Lange AW, Herbert JM. A smooth, nonsingular, and faithful discretization scheme for  
14 polarizable continuum models: the switching/Gaussian approach. *J Chem Phys* 2010,  
15 **133**(24): 244111.
- 16
- 17 38. Mulliken RS. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I. *J*  
18 *Chem Phys* 1955, **23**(10): 1833-1840.
- 19
- 20 39. Bultinck P, Van Alsenoy C, Ayers PW, Carbo-Dorca R. Critical analysis and extension of  
21 the Hirshfeld atoms in molecules. *J Chem Phys* 2007, **126**(14): 144111.
- 22
- 23 40. Breneman CM, Wiberg KB. Determining atom-centered monopoles from molecular  
24 electrostatic potentials. The need for high sampling density in formamide  
25 conformational analysis. *J Comput Chem* 1990, **11**(3): 361-373.
- 26
- 27 41. Besler BH, Merz KM, Kollman PA. Atomic charges derived from semiempirical methods. *J*  
28 *Comput Chem* 1990, **11**(4): 431-439.
- 29
- 30 42. Marenich AV, Jerome SV, Cramer CJ, Truhlar DG. Charge Model 5: An Extension of  
31 Hirshfeld Population Analysis for the Accurate Description of Molecular Interactions in  
32 Gaseous and Condensed Phases. *J Chem Theory Comput* 2012, **8**(2): 527-541.
- 33
- 34 43. Bottcher CJF. *Theory of Electric Polarization*. Elsevier: New York, 1952.
- 35
- 36 44. Fried SD, Boxer SG. Measuring electric fields and noncovalent interactions using the  
37 vibrational stark effect. *Acc Chem Res* 2015, **48**(4): 998-1006.

- 1  
2 45. Lin CY, Romei MG, Oltrogge LM, Mathews, II, Boxer SG. Unified Model for Photophysical  
3 and Electro-Optical Properties of Green Fluorescent Proteins. *J Am Chem Soc* 2019,  
4 **141**(38): 15250-15265.
- 5  
6 46. Bateman JB, Gabriel C. Dielectric properties of aqueous glycerol and a model relating  
7 these to the properties of water. *Journal of the Chemical Society, Faraday Transactions 2*  
8 1987, **83**(2).
- 9  
10 47. Schneider SH, Kratochvil HT, Zanni MT, Boxer SG. Solvent-Independent Anharmonicity  
11 for Carbonyl Oscillators. *J Phys Chem B* 2017, **121**(10): 2331-2338.
- 12  
13 48. Luzar A, Chandler D. Hydrogen-bond kinetics in liquid water. *Nature* 1996, **379**(6560):  
14 55-57.
- 15  
16 49. Bertran JF, Ballester L, Dobrihalova L, Sánchez N, Arrieta R. Study of Fermi resonance by  
17 the method of solvent variation. *Spectrochimica Acta Part A: Molecular Spectroscopy*  
18 1968, **24**(11): 1765-1776.
- 19  
20 50. McKean DC, Machray S. Analysis of the  $\nu_2 : \nu_3 + \nu_4$  Fermi resonance in CH<sub>3</sub>CN isotopic  
21 species in solution. *Spectrochimica Acta Part A: Molecular Spectroscopy* 1988, **44**(5):  
22 533-537.
- 23  
24 51. Rodgers JM, Abaskharon RM, Ding B, Chen J, Zhang W, Gai F. Fermi resonance as a  
25 means to determine the hydrogen-bonding status of two infrared probes. *Phys Chem*  
26 *Chem Phys* 2017, **19**(24): 16144-16150.
- 27  
28 52. Hush N, Reimers J. Vibrational stark spectroscopy. 1. Basic theory and application to the  
29 CO stretch. *J Phys Chem* 1995, **99**(43): 15798-15805.
- 30  
31  
32