Protein Electric Fields Enable Faster and Longer-Lasting Covalent Inhibition of β-Lactamases

Zhe Ji, Jacek Kozuch, Irimpan I. Mathews, Christian S. Diercks, Yasmin Shamsudin, Mirjam A. Schulz, and Steven G. Boxer*

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The widespread design of covalent drugs has focused on crafting reactive groups of proper electrophilicity and positioning toward targeted amino-acid nucleophiles. We found that environmental electric fields projected onto a reactive chemical bond, an overlooked design element, play essential roles in the covalent inhibition of TEM-1 β-lactamase by avibactam. Using the vibrational Stark effect, the magnitudes of the electric fields that are exerted by TEM active sites onto avibactam’s reactive C=O were measured and demonstrate an electrostatic gating effect that promotes bond formation yet relatively suppresses the reverse dissociation. These results suggest new principles of covalent drug design and off-target site prediction. Unlike shape and electrostatic complementarity which address binding constants, electrostatic catalysis drives reaction rates, essential for covalent inhibition, and deepens our understanding of chemical reactivity, selectivity, and stability in complex systems.

INTRODUCTION

The search for drugs that make covalent bonds to their targets has emerged as an important theme of drug discovery and development due to the advantages of increased potency and prolonged residence time.1–4 A typical covalent inhibitor bears an electrophilic warhead that, following specific binding to the target protein, is positioned to react rapidly with a proximal nucleophilic amino-acid residue, but minimally to off-target sites.5–7 Such selective reactivities are interpreted to be a combined consequence of using moderate electrophiles as the warhead, addressing protein residues that are highly nucleophilic and sterically accessible, and enforcing their proximity upon binding (Figure 1a).8,9 These design principles, based solely on intrinsic reactivity and atom locality, have limited power in predicting covalent inhibitors’ potency of action. Given the same covalent inhibitor, what makes the bond formation fast at one site, yet slow or even futile at another? This question also cannot be simply answered by shape and electrostatic complementarity, which are widely used to optimize binding constants rather than reactivity (Figure 1b). Beyond discoveries from screening, a deeper understanding of covalent inhibition is missing.

We chose TEM-1 β-lactamase, a culprit of antibiotic resistance, as the model enzyme because it has been extensively studied as a target of covalent inhibition. TEM-1 β-lactamase rapidly hydrolyzes penicillin G (PenG), a β-lactam substrate, through a two-step mechanism.10,11 The hydroxy group of S70 attacks PenG’s β-lactam carbonyl (C=O), generating an acyl-enzyme (Figure 2a), which is subsequently decacylated via hydrolysis for a catalytic turnover (Table S1). By contrast, a similar nucleophilic attack on the urea C=O of avibactam (AVB) forms a carbamyl-enzyme stable to hydrolysis, thus trapping the enzyme in the covalent complex (Figure 2b).12 The carbamylation is reversible through a slow recyclization reaction to reform the cyclic urea, making AVB a reversible covalent inhibitor.13 AVB has been identified as an effective covalent drug for targeting many serine β-lactamases,14,15 exhibiting desired kinetic parameters, including high rates of carbamylation (kcbm) for fast action, low rates of recyclization (krec) for prolonged residence time, and small overall dissociation constants (Kcbm) for high affinity (Tables 1 and S2). We sought to find a unifying physical basis for the outstanding performance of AVB that may inform the design of covalent inhibitors in general.

The concept of electrostatic catalysis has been proposed16 and experimentally17–19 and computationally20,21 demonstrated to account for a substantial fraction of the remarkable proficiency of a number of enzymes. Different from externally applied electric fields in modulating reactivities22–24 electric fields in enzyme active sites are exerted by the charges and dipoles organized by protein scaffolds and have been found to stabilize charge separation in transition states, thus lowering the free energy barrier (Figure 2c). To apply this concept beyond enzyme catalysis and use it to investigate covalent inhibition, we examined the electrostatic interactions in the
Covalent drug design based on intrinsic reactivity and proximity
Protein electric field: an untapped handle for modulating reactivity

![Diagram of protein electric field](image)

**Figure 1.** Electrostatic catalysis applied to covalent inhibition. (a) Common considerations for covalent drug design, using a carbonyl warhead as an example. This work presents a new design principle: the electric fields that are produced by protein dipoles and charged groups and act on chemical bonds undergoing reactions. (b) Shape and electrostatic complementarity in contrast to electrostatic catalysis. The former considers all of the protein groups and act on chemical bonds undergoing reactions. (b) Shape and electrostatic complementarity in contrast to electrostatic catalysis. The former considers all of the protein groups and act on chemical bonds undergoing reactions. The latter focuses on the electric fields experienced by the reactive bond and therefore contributes to $k_{cat}$ or $k_{cat'}$, as presented in this work.

**RESULTS**

**Removal of a Key H-Bond by Introducing a Backbone Ester.** We expressed TEM-1 (Table S3) and obtained crystal structures (Tables S4–S8) showing that the residue of A237 is largely solvent-exposed (Figure S1), suggesting the potential tolerance of the protein architecture to replacement of the methyl group with bulkier residues. The A237Y mutation yielded a similar protein structure overlapping well with that of the wild type (WT), except for the extra phenolic residue (Figure S2). Soaking TEM-1 crystals in AVB solutions resulted in TEM–AVB covalent complexes (Figure S3). The bound AVB forms a rich network of noncovalent interactions using its sulfate and amide groups (Figure S4) in addition to the two key H-bonds of the carbamate C=O (Figure 3a). The O–N distances between the C=O and the backbone amides of A237 and S70 are 2.98(12) and 2.69(12) Å, respectively (errors estimated using rearranged Cruickshank's formulae26). The A237Y mutation preserves the conformation of AVB (Figures 3a and S5 and S6) with comparable O–N distances of 2.90(11) and 2.67(11) Å.

To replace the amide backbone of A237Y with an ester, we used amber suppression to site-specifically incorporate p-hydroxy-1-phenylactic acid (HPLA) (Figure S7 and Table S9), a tyrosine analogue that bears the same phenolic residue but a hydroxy acid rather than an amino acid for making a backbone ester (Figure 3b).27 The obtained ester protein (A237Ye) shows a gain in mass by 1 Da captured by high-resolution mass spectrometry (MS), consistent with the extra mass carried by the O compared with the NH group in the original amide backbone (Figure 3b). The ester can be selectively hydrolyzed under alkaline conditions, generating two protein fragments of the expected masses (Figures 3b and S8), which confirms the position of the ester in the protein backbone. Their sum is larger than the mass of the intact ester protein by 18 Da, corresponding to the addition of water by hydrolysis. Although we did not obtain crystals of the A237Y mutant (Text S1 in the Supporting Information), 1H NMR spectroscopy shows high similarity between the spectrum of A237Y and that of A237Ye (Figure S9), including both the downfield amide/aromatic region (Figure S10) and the upfield aliphatic region (Figure S11). The consistent 1H NMR fingerprints confirm that the ester protein adopts the same tertiary structure as the amide protein does.

**Measurement of the Active-Site Electric Fields Experienced by the Reactive C=O in PenG and AVB.** To measure the electric fields at the active site, AVB's reactive C=O is not only a participant in the reactions but is also used as a vibrational (infrared) probe of the electric field it experiences.25,28,29 According to the linear vibrational Stark effect, vibrational frequencies shift in proportion to the magnitude of electric fields. The sensitivity of such frequency shifts to electric fields was determined by vibrational solvatochromism, molecular dynamics (MD) simulations, and vibrational Stark spectroscopy using molecular model compounds (Figures S12–S14 and Tables S10–S14). With the urea and the carbamate C=O probes calibrated, the readout of the vibrational frequency of the C=O can be mapped to the magnitude of electric fields projected onto the C=O for both TEM-AVB and TEM–AVB.
Samples of TEM-AVB for infrared spectroscopy were prepared by mixing AVB with TEM-S70G, a nucleophile-impaired mutant, which traps the noncovalent complex. To extract the AVB $\text{C}=\text{O}$'s vibrational peaks from the overwhelming protein background, we synthesized AVB with an isotope-labeled $^{13}\text{C}=\text{O}$ (Figures 3c and S15–S23), which displays the expected isotope redshift of 47 cm$^{-1}$ with respect to $^{12}\text{C}=\text{O}$ (Figure S24 and Table S15). By subtracting the infrared absorption spectrum of TEM$\cdot$AVB-$^{13}\text{C}$ from that of the unlabeled TEM$\cdot$AVB, the protein background cancels out, leaving only $^{13}\text{C}=\text{O}$ peaks with positive intensities and the corresponding $^{13}\text{C}=\text{O}$ peaks with negative intensities (Figures 3d and S25 and Table S16). The observed peak multiplicity indicates that the vibration experiences a heterogeneous electrostatic environment comprising a distribution of $\text{C}=\text{O}$ orientations and H-bond lengths. The electrostatic heterogeneity is also manifested in TEM$\cdot$AVB samples (Figures 3e and S26 and Table S17). The $^{12}\text{C}^{\text{C}}\text{=}^{13}\text{C}$ difference spectrum shows a peak envelope fitted to five pairs of positive–negative peaks (Text S2 in the Supporting Information). By contrast, the crystal structure of TEM$\cdot$AVB shows only one conformation of AVB (Figures 3a and S27), whether measured at 100 K (1.72 Å) or room temperature (2.45 Å). Complete conformational sampling is rarely accessible from crystallography data (Figure S28), but the associated local electrostatic effects can be sensitively probed by vibrational frequency shifts, and this is the likely origin of the multiple peaks we observe.

Figure 2. Reactions and key electrostatics in the active site of TEM $\beta$-lactamases. (a) Mechanism of acylation by PenG. The hydroxy residue of S70 acts as a nucleophile to attack PenG’s $\beta$-lactam $\text{C}=\text{O}$, forming an oxyanion intermediate before turning into an acyl-enzyme complex, which is subject to hydrolysis (not shown). (b) Mechanism of covalent inhibition by AVB. A similar nucleophilic attack of AVB’s urea $\text{C}=\text{O}$ by S70 traps the enzyme in the carbamate complex without turnover. Recyclization, the reverse reaction of carbamylation, regenerates AVB, making it a reversible covalent inhibitor. Proton movement is omitted in the electron-pushing mechanism. (c) Model of electrostatic catalysis. The conversion of a $\text{C}=\text{O}$ to an oxyanion intermediate passes through a transition state where charges are more separated between C and O atoms, generating a reaction difference dipole ($\Delta \bar{\mu}_{\text{rxn}}$), which interacts with the electric field in the enzyme active site $\bar{F}$ to reduce the free energy barrier by $\Delta \Delta G^\ddagger$. (d) H-bond donated by the A237 backbone amide, a key contributor to the enzyme electric field in $\beta$-lactamases (and many other enzymes), can be perturbed by the A237Y mutation and even removed by replacing the amide with an ester in A237Ye.
The $^{12}$C–O vibrational frequencies were translated into the magnitudes of electric fields projected onto the bond from the environment. To analyze the effects of mutation, we focus on each mutant’s largest-field population because it is most electrostatically activated for the reactions (Text S3 in the Supporting Information). For TEM–AVB, a field as large as $-150$ MV/cm (the negative sign represents stabilization of the bond dipole) was found (Figure 3f and Table S18). The A237Y mutation decreases the field to $-125$ MV/cm, which is remarkably further decreased to $-63$ MV/cm by the A237Ye mutation. Additionally, A237E, R, and W mutants were expressed, showing similar results to A237Y (Figure 3f). For TEM–AVB, a similar ensemble of electric fields was observed among WT and A237Y, E, R, and W (Figure 3g) with the

<table>
<thead>
<tr>
<th></th>
<th>PenG acylation rate, $k_{ac}$ (s$^{-1}$)</th>
<th>AVB carbamylation rate, $k_{cbm}$ (s$^{-1}$)</th>
<th>AVB recyclization rate, $k_{rec}$ (s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>4941 [4784, 5086]</td>
<td>23.1 ± 0.2</td>
<td>$(2.3 ± 0.3) \times 10^{-1}$</td>
</tr>
<tr>
<td>A237Y</td>
<td>2818 [2740, 2940]</td>
<td>6.19 ± 0.08</td>
<td>$(1.7 ± 0.3) \times 10^{-3}$</td>
</tr>
<tr>
<td>A237Y (ester)</td>
<td>86.1 [69.2, 97.0]</td>
<td>$(1.8 ± 0.1) \times 10^{-3}$</td>
<td>$(3.4 ± 0.1) \times 10^{-7}$</td>
</tr>
<tr>
<td>A237E</td>
<td>5857 [5288, 6012]</td>
<td>4.63 ± 0.23</td>
<td>$(8.7 ± 1.5) \times 10^{-3}$</td>
</tr>
<tr>
<td>A237R</td>
<td>3265 [3137, 3396]</td>
<td>3.63 ± 0.03</td>
<td>$(6.2 ± 0.9) \times 10^{-4}$</td>
</tr>
<tr>
<td>A237W</td>
<td>1390 [1366, 1415]</td>
<td>3.31 ± 0.05</td>
<td>$(6.7 ± 1.3) \times 10^{-4}$</td>
</tr>
</tbody>
</table>

$^a$k$_{ac}$ was derived from $k_{cat}$ using steady-state mass spectrometry. $k_{cbm}$ was determined by fitting full-time kinetic data to the integrated Michaelis–Menten equation. Three independent repeats and a boot-strapping algorithm provided the medians with uncertainties at 68.25% confidence interval. See the Supporting Information for experimental details and full kinetic data (Figure S1).

$^b$k$_{cbm}$ and $k_{rec}$ were measured by inhibition experiments from which the rate of substrate turnover’s initial value, final value, and decay rate were obtained. The errors were estimated based on the fitting results of 20–30 inhibition experiments with a series of AVB concentrations. See the Supporting Information for experimental details and full kinetic data (Figure S2).

Figure 3. Electrostatic perturbation by A237 mutations measured using isotope-edited infrared spectroscopy. (a) Key active site residues in the WT and A237Y crystal structures of TEM–AVB (Figure S5). The two key H-bonds between the carbamate C=O and the backbone amides are highlighted as red dashed lines. The electron density maps of AVB (2m$F_o$ − DF$_c$, 1.5$\sigma$) are depicted. (b) Substitution of a backbone amide to an ester by incorporation of p-hydroxy-L-phenyllactic acid (HPLA), the $\alpha$-hydroxy acid counterpart of tyrosine, and the high-resolution MS of A237Y, A237Ye, and the two protein fragments of hydrolyzed A237Ye (Table S3). (c) $^{13}$C-Labeled AVB and PenG. (d, e) $^{12}$C–$^{13}$C difference infrared absorption spectra for TEM–AVB (trapped by the S70G mutation) (d) and TEM–AVB (e). The experimental curve (black) is fitted to a sum (gray) of $^{12}$C (positive) and $^{13}$C (negative) peaks (curve fitting details in Tables S16 and S17). Peaks belonging to the same positive-negative pair are filled with the same color. (f–h) Fitted $^{12}$C peaks in infrared absorption spectra for WT and A237 mutants of TEM–AVB (f), TEM–AVB (g), and TEM–PenG (h). The top electric field axes are mapped from the bottom frequency axes based on calibration results as detailed in the Supporting Information.
largest fields dropped to ca. −100 MV/cm compared with those in TEM-AVB, as recapitulated by polarizable MD simulations (Figure S29). A237Y was observed with only the smallest field (−35 MV/cm) (Figure 3g) due to inaccessibility of the electrostatic environments involving the H-bond donated by the backbone amide of A237.

We also measured the isotope-edited infrared absorption spectra for TEM-PenG (Figure S30 and Table S19) using 13C-labeled PenG (Figure 3c). Similarly, a distribution of electric fields was observed (Figure 3h), and the A237Y mutation again diminished the field magnitude (Table S20 and Text S4 in the Supporting Information). Polarizable MD simulations show that PenG’s β-lactam C=O is still properly loaded into the active site of A237Yy albeit forming only one H-bond (Figure S31), a conformation likely enforced by many other noncovalent interactions. The calculated field magnitudes match well with experimentally measured ones (largest-field populations) and display an intimate correlation with H-bond number and length (Figure S31).

**Electrostatic Catalysis of PenG Acylation, AVB Carbamylation, and AVB Recyclization.** To correlate the magnitude of electric fields with free energy barriers, we carried out kinetic studies to obtain $k_\text{cat}$ for PenG (Figures S32 and S33) and $k_\text{diss}$ and $k_\text{rec}$ for AVB (Figures S34–S37 and Text S5 in the Supporting Information). Compared with A237Y, A237Yy follows the same kinetic models but shows a decrease in $k_\text{diss}$, $k_\text{rec}$, and $k_\text{cat}$ by factors of $3.5 \times 10^2$, $5.0 \times 10^3$, and 33, respectively (Tables 1 and S1 and S2). Based on the obtained rates, the free energy barrier of these reactions was calculated using transition state theory (Tables S21 and S22) and plotted against the largest electric fields measured for each mutant (Figure 4a).

Under the framework of electrostatic catalysis, and if we make a simplifying assumption that the electric field experienced by the C=O is the same for the reactant and transition state, linear fittings provide intercepts $\Delta G^{\text{f}}_{\text{intermediate}} = 25.0$, $28.6$, and $24.0$ kcal/mol for AVB carbamylation, recyclization, and PenG acylation, respectively, where $\Delta G^{\text{f}}_{\text{intermediate}}$ is the hypothetical activation barrier if there were no electric field in TEM active sites. These results indicate that without a stabilizing electric field, AVB and PenG would have similar reactivity for bonding with S70 while AVB recyclization is less reactive than carbamylation by $\Delta G^{\text{f}}_{\text{intermediate}} = 3.6$ kcal/mol (Figure 4a). We suggest this difference originates from three components: (1) the difference in intrinsic electrophilicity of the bonds, being a urea for carbamylation and a carbamate for recyclization, (2) the difference in nucleophilicity, and (3) the difference in nucleophile positioning (the nucleophile for recyclization is too far away as shown in Figure S38).

The slopes of the linear fittings are $|\Delta \mu_{\text{elec}}| = 1.40, 1.42,$ and $1.42$ D for AVB carbamylation, recyclization, and PenG acylation, respectively, showing that the three reactions have a similar increase in dipole moment upon passage from the reactant C=O to the transition states preceding the tetrahedral C−O$^-$ intermediates (Figure 2a,b). These $|\Delta \mu_{\text{elec}}|$ values are reasonably larger than that of the reaction in ketosteroid isomerase (1.05 D), where a C=O substrate is transformed to an enolate C=C−O$^-$ intermediate.17 $|\Delta \mu_{\text{elec}}|$ is also a measure of a reaction’s sensitivity to electrostatic catalysis, indicating that the three reactions have similar “catalyzability” by an electric field.19 As a result, the very large field in WT TEM-PenG (−171 MV/cm) provides an enhancement of $k_\text{cat}$ by 8.6 orders of magnitude, while a similarly large field in TEM-AVB (−150 MV/cm) enhances $k_\text{diss}$ by 7.5 orders of magnitude, making the bonding of the covalent inhibitor almost as fast as that of the substrate. The much smaller field in TEM−AVB (−97 MV/cm) only accelerates recyclization by 4.9 orders of magnitude. The 10$^{2.6}$-fold less acceleration corresponds to an electrostatic increase in the free energy barrier by $\Delta \Delta G^{\text{f}}_{\text{Le}} = 3.4$ kcal/mol (Figure 4b).

This analysis provides a quantitative model to assess the role of electric field in impairing the recyclization reaction. If the C=O in TEM−AVB experienced the same magnitude of electric field as that in TEM•AVB, the recyclization would run 320-fold faster, shortening the residence time from 72 min to 13 s. Such an undesired enhancement in recyclization would substantially shift $K_\text{diss}$ from 0.74 to 240 nM. Our quantitative experimental evidence suggests an electrostatic gating effect in TEM active sites that...

**Figure 4.** Role of electric fields in covalent inhibition and drug design. (a) Plot of free energy barrier ($\Delta G^{\text{f}}$) against the magnitude of electric fields (F) projected on the reactive C=O for AVB carbamylation, recyclization, and PenG acylation. Expressing $\Delta G^{\text{f}}$ in kcal/mol and F in units of kcal/mole (top axis), the linear regression lines are $\Delta G^{\text{f}} = 1.40F + 25.0$ (AVB carbamylation), $\Delta G^{\text{f}} = 1.42F + 28.6$ (AVB recyclization), and $\Delta G^{\text{f}} = 1.42F + 24.0$ (PenG acylation). (b) Contribution of electrostatic catalysis, $\Delta \mu_{\text{elec}}$ to the $\Delta G^{\text{f}}$ gap between WT AVB carbamylation and recyclization. (c) Conceptual illustration of using two handles together—bond electrophilicity and environmental electric field—to tune the rate of covalent inhibition (illustrated by the shade of orange). Ideally, only the high-field target reaches high reactivity (filled in orange) while the low-field, off-target sites remain under-reactive (unfilled).
promotes the forward reaction but relatively suppresses the reverse reaction, making AVB a faster, tighter-binding, and longer-acting covalent inhibitor.

**DISCUSSION**

Our studies of TEM/AVB provide a new physical basis for understanding the working principles of covalent inhibition. Currently, the most exploited chemical tool for modulating the reactivity of covalent inhibitors is to tune the electrophilicity of warheads. The concept of electrostatic catalysis as we demonstrated in this work adds an additional axis to guide reactivity optimization. If a target site provides large electric fields, it will be preferred for covalent bond formation against off-target sites with small fields even though the off-target sites are equally nucleophilic and positioned. The electrostatic effect can be coupled with electrophilicity tuning such that only the on-target reaction is promoted to a satisfactory level of pharmacodynamics while off-target sites are left with higher energy barriers and negligible rates.

We see the importance of targeting protein sites that are not only nucleophilic but also surrounded by charges and dipoles that can facilitate electrostatic catalysis. Given recent advances in structure prediction from which electrostatic potentials can be readily computed, an analysis of protein electric fields could be a valuable screen both for target sites and to avoid unintended nucleophiles that happen to be in regions where the fields are large and in orientations that are potentially reaction productive. To search for protein sites amenable to covalent bond formation, we propose the following workflow: (i) obtain the structure of the target protein, from either experimental data or structure prediction programs, (ii) generate an electrostatic potential map using MD simulations or more advanced theories (polarizable force fields appear to give more realistic results but at a considerable computational cost), (iii) convert the electrostatic potential map into an electric field map by taking the gradient, and (iv) rank protein sites based on (a) the magnitude of the electric field, (b) the nucleophilicity of the residues nearby, and (c) spatial accessibility/steric hindrance. We envision that protein electric fields can provide a general, quantitative descriptor to predict which site in a protein and which protein within a proteome is preferred for covalent bonding. This is fundamentally different from the concept of electrostatic complementarity, which addresses all of the charges in a ligand, such as the sulfate and amide side groups of AVB, and thus largely affects the binding constants for substrates and inhibitors. Electrostatic catalysis focuses on the fields that are experienced only by the reactive bond and therefore drives reaction rates for covalent inhibitors.

Covalent inhibitors should be structurally designed to engage in electrostatic interactions through proper positioning of the warheads. AVB inhibition naturally utilizes the same large fields fashioned by the catalytic apparatus. But when targeting noncatalytic nucleophiles, or when there is no such luxury of choosing from multiple candidate residues, a thorough examination of the vicinity of the chosen nucleophile to fully exploit polar residues, backbone amides, and structured water for their potential to exert electric fields may lead to new design options. An ideal configuration is to have a C==O (e.g., acryl) warhead not just be attacked by protein nucleophiles (e.g., cysteine and serine) from a perpendicular direction to the bond axis but also experience two coplanar, strong H-bonds. Indeed, H-bonded warheads are found in the crystal structures of BTK1—ibrutinib (Figure S39), KRAS-G12C—sotorasib (Figure S40), DPP-IV—saxagliptin (Figure S41), and NS3/4A-protease—telaprevir (Figure S42), where electric fields may contribute to the efficacy of these covalent inhibitors, though this factor was not recognized. Given that more inhibitor warheads are not H-bonded even though potential H-bond donors are available, there exists a vast space for improving their selectivity. As illustrated in Figure 1a and with the results presented here as an example for TEM/AVB interactions, these four fundamental components—electrophilicity, nucleophilicity, nucleophile positioning with respect to electrophile, and environmental electric field—together provide a basis for understanding reactivity, selectivity, and relative stability.

**ASSOCIATED CONTENT**

**Supporting Information**

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.2c09876. Experimental details on protein expression and purification, mass spectrometry, crystallography, protein 1H NMR, 13C-AVB synthesis, vibrational spectroscopy, enzyme kinetics, and covalent inhibition kinetics, as well as MD simulation methods and extended discussions (PDF)

**Accession Codes**

All X-ray density maps and atomic models for the proteins have been deposited in the Protein Data Bank. WT TEM-1, 7U6Q; A237Y TEM-1, 8DDZ; WT TEM—AVB, 8DE0; A237Y TEM—AVB, 8DE1; A237Y TEM—AVB at room temperature, 8DE2.

**AUTHOR INFORMATION**

**Corresponding Author**

Steven G. Boxer — Department of Chemistry, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0001-9167-4286; Email: sboxer@stanford.edu

**Authors**

Zhe Ji — Department of Chemistry, Stanford University, Stanford, California 94305, United States; orcid.org/0000-0002-8532-333X

Jacek Kozuch — Department of Physics, Freie Universität Berlin, D-14195 Berlin, Germany; Research Building SupraFAB, 14195 Berlin, Germany; orcid.org/0000-0002-2115-4899

Irpinan I. Mathews — Stanford Synchrotron Radiation Lightsource, Menlo Park, California 94025, United States; orcid.org/0000-0001-6254-3519

Christian S. Diercks — Department of Chemistry, Scripps Research, La Jolla, California 92037, United States; orcid.org/0000-0002-7813-0302

Yasmin Shamsudin — Department of Chemistry-BMC, Uppsala University, 752 37 Uppsala, Sweden; orcid.org/0000-0002-6249-9877

Mirjam A. Schulz — Department of Physics, Freie Universität Berlin, D-14195 Berlin, Germany; Research Building SupraFAB, 14195 Berlin, Germany; orcid.org/0000-0003-4201-8092

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.2c09876
ACKNOWLEDGMENTS

The authors thank C. Zheng, S. H. Schneider, and S. D. E. Fried for helpful discussions. They also thank P. Soumillon and P. G. Schultz for providing plasmids, T. McLaughlin and the Vincent Coates Foundation Mass Spectrometry Laboratory, Stanford University Mass Spectrometry core facility for the help with mass spectrometry, and S. Lynch from Stanford University NMR facility for assistance with 1H protein NMR. This work was supported in part by NIH Grant GM118044. J.K. and M.A.S were supported by German Research Foundation grant KO 5464/4. J.K. and M.A.S thank the support of the SupraFAB research building (funded by Federal Government and the State of Berlin) and the ZEDAT for computational resources (Freie Universität Berlin). Y.S. was supported by the Knut and Alice Wallenberg Foundation, Y.S. thanks the Swedish National Infrastructure for Computing (SNIC) for computational resources. Use of the Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, was 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.

REFERENCES


ABBREVIATIONS

PenG penicillin G
AVB avibactam
WT wild type
HPLA p-hydroxy-l-phenylactic acid
MS mass spectrometry
MD molecular dynamics
BTK Bruton’s tyrosine kinase
KRAS a protein translated from the gene KRAS (Kirsten rat sarcoma virus)
DPP dipeptidyl peptidase


Recommended by ACS

Alkaline State of the Domain-Swapped Dimer of Human Cytochrome c: A Conformational Switch for Apoptotic Peroxidase Activity
Huantian Lei, Bruce E. Bowler, et al.
NOVEMBER 08, 2022

Intramolecular C–C Bond Formation Links Anthraquinone and Enediyne Scaffolds in Tiancimycin Biosynthesis
Chun Gui, Ben Shen, et al.
OCTOBER 24, 2022

Cooperative Weak Dispersive Interactions Actuate Catalysis in a Shape-Selective Abiological Racemase
Yujia Wang, Jay S. Siegel, et al.
FEBRUARY 04, 2022

Total Synthesis of (+)-Mutilin: A Transannular [2+2] Cycloaddition/Fragmentation Approach
Han Chen, Tuoping Luo, et al.
AUGUST 22, 2022
Supporting Information

Protein electric fields enable faster and longer-lasting covalent inhibition of β-lactamases

Zhe Ji¹, Jacek Kozuch²,³, Irimpan I. Mathews⁴, Christian S. Diercks⁵, Yasmin Shamsudin⁶, Mirjam A. Schulz²,³ and Steven G. Boxer¹,*

¹Department of Chemistry, Stanford University, Stanford, CA 94305, USA.
²Department of Physics, Experimental Molecular Biophysics, Freie Universität Berlin, Arnimallee 14, D-14195 Berlin, Germany.
³Research building SupraFAB, Altensteinstr. 23a, 14195 Berlin, Germany.
⁴Stanford Synchrotron Radiation Lightsource, Menlo Park, CA 94025, USA.
⁵Department of Chemistry, Scripps Research, La Jolla, CA 92037, USA.
⁶Department of Chemistry-BMC, Uppsala University, 752 37 Uppsala, Sweden.
*Email: sboxer@stanford.edu
Materials and Methods

Plasmid Construction

The pBAD plasmid containing the gene for TEM-1 β-lactamase was generously provided by the Patrice Soumillion lab at the Université Catholique de Louvain\textsuperscript{1,2}. The pBAD vector is a derivative of pBAD/myc-HisB (Invitrogen), with the gene cloned between the Ncol and Xbal sites, carrying a tetracycline resistance gene. The TEM-1 gene contains a A184V stabilizing mutation. In our previous publication\textsuperscript{3}, we have incorporated a thrombin cleavage site into the myc tag. This modified plasmid served as the starting point for studying mutants of this work. Point mutations were made using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent) according to the manufacturer’s protocols.

The site-specific incorporation of a backbone ester was achieved using an aminoacyl-tRNA synthetase (aaRS)\textsuperscript{4}. A pUltra plasmid containing the aaRS gene was generously provided by the Peter G. Schultz lab at the Scripps Research Institute.

To make TEM-1 free of the His tag for crystallization, we redesigned the gene by moving the His tag from the C-terminus to the N-terminus, right after the signal peptide. A TEV protease cleavage sequence was incorporated after the His tag. We used M182T for protein stabilization rather than A184V because of M182T’s higher stabilization effect\textsuperscript{5} and its wide use for the crystallization of TEM proteins. The gene was inserted between the Ncol and Xbal sites into the original pBAD vector provided by the Soumillion lab.

DNA Sequence

TEM-1-A184V-His = TEM-1. Codon for S70 and A237 in bold

\[
\begin{align*}
\text{ATGGGTAGTCAACATTTTCCGTGCCTTATTTCCCTTTTTTTCCGCGCATTTTGCCTTCTGTTT} \\
\text{TTGCACCCAGAAACGCTGGTGAAGATTAAGAGTGAGTCAGTTGGGTGCACGAGTGGG} \\
\text{TTACATCGAATCAGACCAAGCCTAGTCCCTGAGATTTTTCCGCCCAAGAAGCGTGT} \\
\text{CAATGTGAGACACTTTTAAAGTTCTGCTATGTGGAACGCGCATTTATCCGTGTTGAGC} \\
\text{AGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGACAGTTGGCTGCGTATCAACATGAGT} \\
\text{GATAAACACTGCGCCCAACTTTCTGACAACAGCAGTGGCAGGAGGAGCTTAACCCGCTTTTT} \\
\text{TGCACACATGGGGATCATGTAACCTCGCTTTGATCGTTGGGAAACCAGCAGCTGAAATGACCAT} \\
\text{ACCAAAACACGAGCAGTCCCAACACGGGTCTGCTGAGTGAATGGAACAACTIAAAGCGCACAATAT} \\
\text{ACTGGGGACTACTTACTACTTCTAGGCTTTCCGGCAACAATTAAAGACTGAGTGGAGGAGCGGAATAAG} \\
\text{TTGCAGGACCACCTCTGCGCTGCTCCCTTCGCTGGCTGTGTATTTATTTGCTGATAAAATCTTGGAGC} \\
\text{C GGTGACGCGTGGTCTCGCGGTATCATGGCAGCAGTGTTGGGACGAGTGAAGCCCTCCCGTATC} \\
\text{GTATTATACGACTACGGGAGTCAGCAGGAATATTAGGATGAACGAATAAGACAGACGACTGCTGAGA} \\
\text{TAGGTGCTCCCTACGTGAGATTTGAGCCTGCTAGTACCAAGAGGCAGCTAGCAAGAGAGATCTGAA} \\
\text{TACGCGCCGTCACCATCATCATCATCATCATTTA}
\end{align*}
\]
**Protein Sequence**

**TEM-1-A184V-His** ≡ TEM-1. S70 and A237 in bold; Signal peptide in purple

```
MGSQHFRVALIFPFFAFFCLPVFAHPETLVKVKDAEDQLGARVGYIELDLNSGKILESFRPEERF
PMSTFKVLLCGAVLSRVDAGQEQLGRIHYSQNDLVEYSVPTEKHTLDGTMVRELCSAAITMS
DNTAANLLTTTIGGPKETLAFHLHMGDHTVRLDREWPELENIAIPNDERDTTMPVAMATLKRLL
TGELLTLASRQQLIDWMEADKVAPLPRLSALPAGWFIADKSAGERGSRGIIAALGPGKPSRIVVI
YTTGSQATMDERNQIAEIGASLIKWALVPGRSSEEDLNSAVDHHHHHH
```

**TEM-1-M182T-G25** ≡ TEM-1-native. S70 and A237 in bold; Signal peptide in purple; Cleavable His tag in green

```
MGSQHFRVALIFPFFAFFCLPVFAARETHHHHHHGAENLYFQGHPETLVLVKDAEDQLGARVGYI
ELDLNSGKILESFRPEERFPMSTFKVLLCGAVLSRVDAGQEQLGRIHYSQNDLVEYSVPTEK
HTLDGTMVRELCSAAITMSDNTAANLLTTTIGGPKETLAFHLHMGDHTVRLDREWPELENIAIPN
DERDTTPAAMATTLLRLLTGELLTLASRQQLIDWMEADKVAPLPRLSALPAGWFIADKSAGERGS
RGIIAALGPGKPSRIVVIYTTGSQATMDERNQIAEIGASLIKW
```

**Protein Expression and Purification**

**The general protocol for TEM-1:** The pBAD plasmid was transformed into chemically competent TOP10 *Escherichia coli* cells (Invitrogen) using selection with 15 μg/mL tetracycline (Sigma) on Luria Broth (Fisher) agar plates. A single colony of transformed cells was grown into 1 mL culture for 8 h and further into 50 mL cultures overnight using Luria Broth with 15 μg/mL tetracycline at 37°C. 50 mL of an overnight culture was used to inoculate 1 L of Terrific Broth (Fisher) with 10 μg/mL tetracycline shaking at 200 rpm and 37°C until they reached an OD$_{600}$ ~ 0.6, at which point protein expression was induced with 2 g/L *L-*(+)-arabinose (Sigma), and grown for 6 h at 27°C. Cells were harvested by centrifugation at 6,000 × g for 20 mins and resuspended
in a lysis buffer (50 mM potassium phosphate (KPi), 20 mM imidazole, 500 mM NaCl, 10% (v/v) glycerol, pH 7.4). Cells were lysed with a homogenizer (Avestin EmulsiFlex-C3), and the lysate was centrifuged twice for 90 min each at 15,000 × g. The crude protein in the supernatant was filtered through a 0.45 μm filter membrane before being loaded onto a Ni-NTA affinity resin (QIAGEN) column equilibrated with the lysis buffer. The protein was purified using 6 column volumes of a washing buffer (50 mM KPi, 50 mM imidazole, 500 mM NaCl, pH 7.4) and 2 column volumes of an elution buffer (50 mM KPi, 200 mM imidazole, 500 mM NaCl, pH 7.4). Further purification was performed using anion exchange chromatography on a 5 mL HiTrap-Q HP (GE Healthcare) column and eluted using a 0-25% gradient of buffer A (25 mM Tris-HCl, 25 mM NaCl, pH 8.4) and buffer B (25 mM Tris-HCl, 1 M NaCl, pH 8.4) over 25 column volumes. Purified protein was exchanged into a cryoprotecting storage buffer (50 mM KPi, 100 mM NaCl, 10% (v/v) glycerol, pH 7.4) for long-term storage at -80°C. A typical yield is 5 mg of protein from 1 L of media. The identity and purity of all proteins were confirmed with mass spectrometry (MS). The expected and observed masses are summarized in Table S3.

**A237Ye:** For the expression of A237Ye, the genes *aspC* and *tyrB* in DH10B cells were knocked out by using phage λ-red recombinase, in order to impair their capability to metabolize the non-canonical hydroxy acid, *p*-hydroxy-L-phenyllactic acid (HPLA, AstaTech). The success of gene knockout was confirmed by testing the dependence of cell growth on an external tyrosine source. Cells failed to grow in a media depleted with tyrosine due to the removal of *tyrB* (Table S9). Furthermore, growing the modified DH10B cells with HPLA overnight after OD₆₀₀ had reached 0.6 led to a concentration drop of HPLA only from 1 mM to 0.7 mM in the media, as quantified by HPLC, while growing the native DH10B cells led to a final HPLA concentration of merely 0.04 mM (Figure S7). These results confirmed that the modified DH10B cells have lost the capability to metabolize HPLA, allowing for prolonged existence of HPLA in the media for protein expression.

The modified DH10B cells were made chemically competent and sequentially transformed with the aaRS plasmid (selection with 50 μg/mL spectinomycin) and the pBAD plasmid (selection with 15 μg/mL tetracycline). The protein expression was carried out in the same way except for additionally using 50 μg/mL spectinomycin as antibiotics and using 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG, Santa Cruz Biotechnology) and 2 mM HPLA for induction. The protein purification was carried out in the same way except for using buffers of lower pH for anion exchange chromatography: buffer C (12.5 mM KPi, 6.25 mM NaCl, pH 7.4) and buffer D (50 mM KPi, 1 M NaCl, pH 7.4). The lower pH reduces the hydrolysis of the backbone ester. A similar yield of 5 mg of protein from 1 L of media can be achieved.

The incorporation of the backbone ester was confirmed first by a hydrolysis test. 0.1 mg of A237Ye was incubated with 0.5 M NaOH for 1 h at room temperature (RT). The reaction mixture was neutralized by acetic acid before a MS measurement. MS spectra displayed a complete decomposition of A237Ye into two peptide fragments, corresponding to the hydrolysis of the very backbone ester (Figures 3b and S8, and Table S3). High-resolution MS further characterized the extra mass carried by the backbone ester. The O atom is heavier than the NH group by 1 Da and this difference was captured (Figure 3b). The hydrolysis fragments were also measured by high-resolution MS and showed a gain in their total mass by 18 Da (Figure 3b), corresponding to the addition of H₂O associated with ester hydrolysis. The amide protein A237Y cannot be hydrolyzed.
under the same conditions. The incorporation fidelity was estimated to be ~ 99% according to kinetic tests (Table S1).

**TEM-1-native**: For the expression of TEM-1-native proteins using the corresponding plasmids, protein expression and purification were carried out in the same way except for adding two extra steps. After the crude protein was collected from a Ni-NTA affinity column in the elution buffer, TEV protease (Gene and Cell Technologies) was added (0.2 mg TEV protease per 1 mg crude protein) for the cleavage of the His tag. Complete cleavage (> 95%) was observed after incubation at 30 min at RT by MS. The mixture was diluted 10 times by water before being loaded onto the Ni-NTA affinity column equilibrated with the same 10-fold dilution of the elution buffer. The column was washed with 2 column volumes of the washing buffer. All the solution that has flowed through the column by this point was collected and used for the anion exchange chromatography as described above. A similar final yield of 5 mg of protein from 1 L of media can be achieved.

**Mass Spectrometry**

Mass spectrometry (MS) was performed at the Stanford University Mass Spectrometry facility on a Waters Single Quadrupole Mass Spectrometer and Acquity H Class UPLC equipped with a reverse-phase C8 column (Zorbax 300SB-C8, Rapid Resolution 2.1 × 50 mm 3.5u with 300SB-C8 Narrow=Bore Guard column 2.1 × 12.5mm 5u). The LC mobile phases consist of 0.1% formic acid in water and 0.1% formic acid in acetonitrile. ESI was performed under standard cone voltage. The mass was recorded in ESI+ mode in the range of 50-2000 m/z. Data deconvolution and analysis was performed using the Intact Mass software (Protein Metrics).

High-resolution MS was performed at the Stanford University Mass Spectrometry facility on the Waters Acuity UPLC and Bruker MicroTOF-Q II. The column was a 2.1 x 100 mm, 3.5u Zorbax Stablebond 300-C8 column from Agilent, the temperature was 50°C, and the flow rate was 0.2ml/min. The injection volume was 2 µL. The LC mobile phases consist of 0.1% formic acid in water and 0.1% formic acid in acetonitrile. Spectra were collected in full scan MS mode with a mass range of 600-4000 Da and collision RF setting of 800 Vpp.

**Protein Crystallization and Crystallography**

Crystallization was carried out based on the hanging drop method using VDX Plates with sealant (Hampton Research). Each well was loaded with 0.5 mL reservoir solution. 1 µL of protein solution was added onto a Fisherbrand unbreakable cover slip (Fisher Scientific). This was repeated to form two initial drops. Into each drop was added 1 µL of reservoir solution. The cover slips were flipped and sealed onto the wells. The plates were left at RT (296 K). When seeding was implemented, the cover slip was unsealed after 20 h. Microcrystals in a seed stock were transferred using a cat whisker. The whisker was streaked through the two drops consecutively without reloading. Afterwards, the cover slip was re-sealed onto the well. Crystals grew at RT.

**Seed-a**: TEM-1-native solution was prepared in 50 mM KPi (pH 7.0) at a concentration of 20 mg/mL. The reservoir solution contains 0.1 M HEPES (pH 7.5) and 2.0 M ammonium sulfate.
Clusters of small needle crystals appeared after 2 days. The crystals were collected and fragmented using ceramic beads and diluted 1000-fold with the reservoir solution as the stock of seed-a.

**Seed-b**: TEM-1-native solution was prepared in 50 mM KPi (pH 7.0) at a concentration of 10 mg/mL. The reservoir solution contains 0.1 M HEPES (pH 7.0) and 1.8 M ammonium sulfate. The drops were seeded with seed-a. Clusters of small needle crystals appeared after 7 days. The needle crystals disappeared, and thin plate crystals emerged after 14 days. The crystals were collected and fragmented using ceramic beads and diluted 1000-fold with the reservoir solution as the stock of seed-b.

**Seed-c**: TEM-1-native solution was prepared in 50 mM KPi (pH 7.0) at a concentration of 20 mg/mL. The reservoir solution contains 0.1 M HEPES (pH 7.0) and 1.6 M ammonium sulfate. The drops were seeded with seed-a. Needle crystals appeared after 14 days. The crystals were collected and fragmented using ceramic beads and diluted 100-fold with the reservoir solution as the stock of seed-c.

**WT**: TEM-1-native solution was prepared in 10 mM HEPES (pH 7.0) at a concentration of 20 mg/mL. The reservoir solution contains 50 mM HEPES (pH 7.0), 50 mM KPi (pH 7.0) and 1.6 M ammonium sulfate. The drops were seeded with seed-b. Thin plate crystals appeared after 1 month and were harvested for diffraction measurements.

**A237Y**: A better crystallization condition was found. TEM-1-native-A237Y solution was prepared in 10 mM HEPES (pH 7.0) at a concentration of 20 mg/mL. The reservoir solution contains 25% or 27.5% PEG3350 (Hampton Research). The drops were seeded with seed-c. Thicker plate crystals appeared after 7 days and were harvested for diffraction measurements. This condition was also used to obtain WT crystals for AVB binding.

**WT–AVB, A237Y–AVB**: For obtaining structures with bound AVB, the crystals were soaked in reservoir solution containing 2 mM AVB for 24 h before dipping into the cryoprotectant.

Single crystals were looped and dipped into the cryoprotectant (reservoir solution supplemented with 28% ethylene glycol) before flash cooling in liquid nitrogen. The X-ray diffraction data were collected at 100 K at the Stanford Synchrotron Radiation Lightsource (Menlo Park, CA) at beamline BL12-2. The room-temperature structures were measured at beamline BL12-1. The data were indexed, reduced, and scaled using X-ray Detector Software (XDS). The structures were solved using the Molecular Replacement Module in PHENIX, where the initial model was a published structure of the M182T mutant of TEM-1 (PDB: 1jwp). Several rounds of manual and automatic structure refinement were performed using COOT and PHENIX. The resulting data collection and refinement statistics are summarized in Tables S4-S8.

**Protein 1H-NMR**

Protein samples of the A237Y and A237Ye mutants of TEM-1 were prepared by exchanging 3 mg of the proteins into 500 µL of a buffer containing 20 mM KPi at pH 6.0 with 10% D2O. The protein concentration is ~ 0.2 mM. Proton nuclear magnetic resonance (1H-NMR) spectra were
recorded on a Varian Inova 600 MHz spectrometer with a 5 mm Z-gradient HCN probe. The spectra were acquired at 15 °C. Water suppression was achieved using the WET method. The data were processed with MestReNova software using a line-broadening parameter of 2 Hz. The spectra are shown in Figures S9-S11.

**Synthesis of Avibactam-13C**

(2S,5R)-Ethyl 5-((benzyloxy)amino)piperidine-2-carboxylate (1) oxalate was purchased from ChemScene LLC. N,N-Diisopropylethylamine, phosgene-13C solution (99 atom % 13C, 20% in toluene), anhydrous sodium sulfate, tetrahydrofuran, lithium hydroxide monohydrate, methanol-d4 (99.8 atom % D), isobutyl chloroformate, palladium on carbon (10 wt.% loading), n-butyl acetate, acetic acid, and 4-methyl-2-pentone were purchased from Sigma-Aldrich. Dichloromethane (extra dry) and CDCl3 (99.8 atom % D) were purchased from Acros Organics. Ethyl acetate, hexanes, phosphoric acid, sodium bicarbonate, sodium chloride, hydrogen chloride, ammonium hydroxide, methanol, isopropanol, and dichloromethane were purchased from Fisher Scientific. Triethylamine was purchased from Millipore. Sulfur trioxide trimethylamine complex was purchased from Matrix Scientific. Tetrabutylammonium acetate was purchased from Alfa Aesar.

1H-NMR spectra were recorded on a Varian VNMRS-400 MHz instrument and are reported in ppm using solvent as an internal standard (CDCl3 at 7.26 ppm, CD3OD at 3.31 ppm). Proton-decoupled carbon nuclear magnetic resonance spectra (13C-NMR) were recorded on a Varian VNMRS-400 MHz instrument and are reported in ppm using solvent as an internal standard (CDCl3 at 77.16 ppm). Mass spectra were recorded on a Waters Single Quadrupole Mass Spectrometer at the Stanford University Mass Spectrometry facility. ESI was performed under a low cone voltage. The mass was recorded in both positive and negative modes in the range of 50-2000 m/z.

The overall synthesis route for AVB-13C is shown in Figure S15.

(2S,5R)-Ethyl 6-(benzyloxy)-7-oxo-1,6-diazabicyclo[3.2.1]octane-2-carboxylate-13C (2)

A solution of 1 oxalate (100 mg, 0.27 mmol) and N,N-diisopropylethylamine (188 µL, 1.08 mmol) in anhydrous dichloromethane (DCM, 2 mL) was cooled to 0°C under argon. A solution of phosgene-13C (20%, 99 atom% 13C) in toluene (234 µL, 0.41 mmol) was added dropwise. The reaction mixture was stirred at room temperature overnight, and the progress was monitored by thin layer chromatography (TLC) using EtOAc/hexanes (1:1, with 0.1% triethylamine).
reaction mixture was washed with H$_3$PO$_4$ (10%) twice, sat. NaHCO$_3$ twice, and brine. The organic layer was dried over anhydrous Na$_2$SO$_4$, concentrated under reduced pressure, and dried under vacuum, affording 2 (69 mg, 83%) as a white solid. $^1$H-NMR (400 MHz, CDCl$_3$) as shown in Figure S16: $\delta$ 7.40-7.31 (m, 5H), 5.01 (dd, $J = 11.5, 2.8$ Hz, 1H), 4.86 (dd, $J = 11.5, 2.6$ Hz, 1H), 4.20 (qd, $J = 7.2, 2.3$ Hz, 2H), 4.07-4.04 (m, 1H), 3.32-3.28 (m, 1H), 3.05-3.00 (m, 1H), 2.90 (ddd, $J = 11.8, 6.4, 2.3$ Hz, 1H), 2.09-2.01 (m, 2H), 2.00-1.97 (m, 1H), 1.68-1.59 (m, 1H), 1.28-1.23 (td, $J = 7.2, 3.0$ Hz, 3H); $^{13}$C-NMR (100 MHz, CDCl$_3$) as shown in Figure S17: $\delta$ 168.08 (strong), 168.05, 135.68, 129.20, 128.68, 128.48, 78.20, 61.60, 59.65, 58.22, 47.39, 21.12, 19.89, 14.12.

\[
\text{(2S,5R)-6-(benzyloxy)-7-oxo-1,6-diazabicyclo[3.2.1]octane-2-carboxylic acid-^{13}C (3)}^{11}: \text{To a solution of 2 (150 mg, 0.49 mmol) in tetrahydrofuran (3 mL) was added dropwise LiOH-H$_2$O (20.6 mg, 0.49 mmol) in water (1 mL). The reaction mixture was stirred at room temperature and monitored by TLC using EtOAc/hexanes (1:1, with 0.1% triethylamine). After the precursor disappeared (~1 h), the reaction mixture was washed with EtOAc twice, and then the aqueous solution was adjusted to pH = 3 with 1N HCl. The resulting mixture was extracted with DCM three times. The combined organic layer was dried over anhydrous Na$_2$SO$_4$, concentrated under reduced pressure, and dried under vacuum, affording 3 (112 mg, 82%) as a white solid. $^1$H-NMR (400 MHz, CDCl$_3$) as shown in Figure S18: $\delta$ 7.44-7.37 (m, 5H), 5.06 (d, $J = 11.3$ Hz, 1H), 4.91 (d, $J = 11.3$ Hz, 1H), 4.16 (d, $J = 7.1$ Hz, 1H), 3.42-3.38 (m, 1H), 3.18 (d, $J = 11.8$ Hz, 1H), 2.92 (dd, $J = 11.9, 6.4$ Hz, 1H), 2.26-2.21 (m, 1H), 2.11-2.01 (m, 2H), 1.72-1.63 (m, 1H); $m/z$ (ESI-MS) 276.1 [M – H]$^+$.}

\[
\text{(2S,5R)-6-(Benzyloxy)-7-oxo-1,6-diazabicyclo[3.2.1]octane-2-carboxamide-^{13}C (4)}^{12}: \text{A solution of 3 (112 mg, 0.40 mmol) and triethylamine (74 $\mu$L, 0.53 mmol) in DCM (2 mL) was cooled to 0°C under argon. Isobutyl chloroformate (67 mg, 0.49 mmol) was added quickly to the mixture, followed by stirring at 0°C for 40 min. Aqueous ammonia (0.3 mL) was added slowly to the reaction mixture, followed by stirring at room temperature for 2 h. Subsequently, water (10
(2S,5R)-6-hydroxy-7-oxo-1,6-diazabicyclo[3.2.1]octane-2-carboxamide-13C (5): To a solution of 4 (95 mg, 0.34 mmol) in methanol (3.5 mL) was added palladium on carbon (17 mg, 10 wt.% loading), followed by stirring under H₂ atmosphere. The reaction was monitored by TLC using methanol/DCM (1:20) and took ~45 min for the complete consumption of the precursor. The reaction mixture was filtered through celite, concentrated under reduced pressure, and dried under vacuum, affording 5 (60 mg, 94%) as a white solid. ¹H-NMR (400 MHz, CD₃OD) as shown in Figure S21: δ 3.85 (d, J = 7.8 Hz, 1H), 3.72-3.68 (m, 1H), 3.19-3.14 (m, 1H), 2.97 (dd, J = 11.7, 6.4 Hz, 1H), 2.27 (dd, J = 15.0, 6.9 Hz, 1H), 2.09-2.04 (m, 1H), 1.93-1.87 (m, 1H), 1.77-1.69 (m, 1H).

Tetrabutylammonium [(2S,5R)-2-Carbamoyl-7-oxo-1,6-diazabicyclo[3.2.1]octan-6-yl] Sulfate-13C (AVB-13C)¹³: 5 (60 mg, 0.32 mmol) was mixed with sulfur trioxide trimethylamine complex (53.6 mg, 0.39 mmol) in isopropanol (2 mL) and water (2 mL). Triethylamine (11.2 µL, 0.08 mmol) was added to the reaction mixture, followed by stirring at room temperature for 24 h. The reaction mixture was washed with n-butyl acetate, from which the aqueous phase was collected. A solution of tetrabutylammonium acetate (163 mg, 0.54 mmol) and acetic acid (2 µL)
in water (400 µL) was prepared. Next, 300 µL of the tetrabutylammonium acetate solution was added to the reaction mixture, which was then extracted with 2 mL of DCM. The remaining 100 µL of the tetrabutylammonium acetate solution was added to the aqueous phase, which was then extracted with 2 mL of DCM. The organic extracts were combined, concentrated under reduced pressure, and dried under vacuum, affording a crude product (189 mg) containing ~42 wt.% AVB-\textsuperscript{13}C (49% yield), the rest being tetrabutylammonium acetate and residual solvents. The crude product was recrystallized from 4-methyl-2-pentone by cooling to 0°C to precipitate the desired product. The product was collected by filtration, washed with 4-methyl-2-pentanone, and dried under vacuum, affording white crystalline solids of AVB-\textsuperscript{13}C (24 mg, 15%). \textsuperscript{1}H-NMR (400 MHz, CDCl\textsubscript{3}) as shown in Figure S22: δ 6.72 (s, 1H), 5.59 (s, 1H), 4.36 (s, 1H), 3.92 (d, J=7.8 Hz, 1H), 3.38 (d, J = 11.7 Hz, 1H), 3.32-3.27 (m, 8H), 2.87 (dd, J = 11.5, 6.5 Hz, 1H), 2.39 (dd, J = 15.0, 7.0 Hz, 1H), 2.20-2.14 (m, 1H), 1.92-1.82 (m, 1H), 1.71-1.62 (m, 9H), 1.49-1.40 (m, 8H), 1.00 (t, J = 7.4 Hz, 12H); \textsuperscript{13}C-NMR (100 MHz, CDCl\textsubscript{3}) as shown in Figure S23: δ 165.67 (strong), 60.61, 58.83, 57.93, 48.27, 24.08, 20.87, 19.83, 17.37, 13.81; m/z (ESI-MS) 508.8 [M + H]^+, 265.0 [M – NBu\textsubscript{4}]; FTIR, see Figure S24.

Calibration of Vibrational Stark Probes

Overview: Instead of directly using AVB for probe calibration, we employed (2S,5R)-6-(Benzyloxy)-7-oxo-1,6-diazabicyclo[3.2.1]octane-2-carboxamide (BODAC), the counterpart of 4 without \textsuperscript{13}C labeling, as a model compound which replaces the charged sulfate of AVB with a neutral benzyloxy group to avoid overestimation of electric fields in MD simulations (Figure S12, Tables S10 and S11)\textsuperscript{14}. When AVB is covalently linked to S70, its urea C=O is converted to a carbamate C=O (Figure 2b), whose frequency sensitivity to electric field was separately calibrated using methyl piperidine-1-carboxylate (MPC), a simple, rigid carbamate, as an ideal model compound (Figures S13 and S14, Tables S12-S14).

In Figure 3f-h, the top electric field axes are mapped from the bottom frequency axes according to the calibration results: \(\bar{\nu} = 0.76F + 1779.8\) (f) based on BODAC; \(\bar{\nu} = 0.75F + 1714.0\) (g) based on MPC; \(\bar{\nu} = 0.78F + 1796.8\) (h) based on penam\textsuperscript{14}. \(\bar{\nu}\) is the wavenumber (cm\textsuperscript{-1}) of the C=O vibrations, and \(F\) is the magnitude of active-site electric fields (MV/cm) projected on the C=O.

Materials: BODAC was purchased from Enovation Chemicals. MPC was purchased from AA Blocks. All organic solvents were obtained from Sigma-Aldrich (anhydrous and at > 99% purity). D\textsubscript{2}O (99 atom % D) was obtained from Cambridge Isotope.

FTIR Spectroscopy—Vibrational Solvatochromism: BODAC and MPC were dissolved in organic solvents and D\textsubscript{2}O at a concentration of 10 mM. In cases where solubility became a limiting factor, a saturated solution was prepared and diluted by a factor of least 2 before injection into the cell, in order to prevent solute aggregation. Six organic solvents of varying polarity were chosen: hexanes, dibutylether, tetrahydrofuran (THF), toluene, dimethylsulfoxide (DMSO), dichloromethane (DCM). The exception is that BODAC has only trace solubility in hexanes and no absorption peaks were obtained for this combination.
Infrared spectra were recorded using a Bruker Vertex 70 equipped with a liquid nitrogen-cooled mercury cadmium telluride detector and under constant purging of the sample chamber with dry air, similarly to previous works\textsuperscript{15}. 20 μL of the samples were loaded into a demountable cell composed of two CaF\textsubscript{2} optical windows (19.05 mm diameter, 3 mm thickness, Lambda Research Optics). The windows were separated by two Teflon spacers, one 25 μm thickness and the other 50 μm. Before FTIR measurement, a 5 min delay was applied to purge gaseous CO\textsubscript{2}, and then 256 scans were acquired and averaged to obtain each transmission interferogram. Spectra were recorded in window of 4000–1000 cm\textsuperscript{-1} with 1 cm\textsuperscript{-1} resolution. Blank solvent spectra were recorded under the same condition and used as a reference for subtracting the background. Data collection and processing were performed using the software OPUS 5.0. Peak positions and the full width at half maximum (FWHM) were determined using the built-in programs PeakPick and CurveFit (combination of Lorentzian and Gaussian) in OPUS 5.0 (Tables S10 and S12).

**Fixed-Charge MD Simulations:** Fixed-charge MD simulations were carried out similarly to previous works\textsuperscript{14,16,17}. The molecule of interest was constructed in Gaussview 6.0, and the geometry was optimized by Density Functional Theory (DFT) in Gaussian 16 (ref.\textsuperscript{18}) at the B3LYP/6-311++G(2d,2p) level. The resultant structure was then used for parameterization through the Antechamber program of AmberTools16 (ref.\textsuperscript{19}) based on the GAFF force field. The force field parameters for organic solvents were taken from Caleman et al\textsuperscript{20}, and water was parameterized using the TIP3P model. All simulations were performed by GROMACS 2018 (ref.\textsuperscript{21}). A single solute molecule was placed in the center of a cubic solvation box that has a size of 4×4×4 nm\textsuperscript{3} filled with solvent molecules. Under periodic boundary conditions, the system was first energy minimized until the maximum force is less than 1000 kJ/mol/nm, followed by NVT and NPT equilibration runs (100 ps for each). MD production runs were performed over 1 ns with 2 fs steps, with a van der Waals cutoff of 10 Å (with analytical VdW correction), an electrostatics cutoff of 10 Å (using the particle mesh Ewald method), an SD (leap-frog stochastic dynamics) integrator, the Berendsen thermostat\textsuperscript{22} and the Parrinello-Rahman barostat\textsuperscript{23}. The electric field vectors on the relevant atoms (C and O) were obtained via dividing the electrostatic forces acting on these atoms by the corresponding atomic charges. To extract the solvent contributions to the electric fields, the MD frames obtained were recalculated with atomic charges on all solvent molecules zeroed out, providing the field contributions from the solute atoms themselves which were then subtracted from the electric field values obtained in the original production run. The electric field projections along the C=O directions are then evaluated using the following equation:

\[
F_{CO} = \frac{1}{2} (\vec{F}_C \cdot \hat{r}_{CO} + \vec{F}_O \cdot \hat{r}_{CO})
\]

where \(F_{CO}\) denotes the field projection along the C=O bond, \(\vec{F}_C\) and \(\vec{F}_O\) are the electric fields acting on the C and O atoms, respectively, and \(\hat{r}_{CO}\) refers to a unit vector along the direction of the C=O bond. The magnitudes of the fields projected on the C=O bond in BODAC and MPC are listed in Tables S11 and S13. The calibration of vibrational frequency against the calculated electric field is shown in Figures S12 and S13.

**Vibrational Stark Spectroscopy:** Vibrational Stark spectroscopy was carried out similarly to previous works\textsuperscript{24,25}. MPC was dissolved in 2-methyl tetrahydrofuran (Sigma-Aldrich), a glass-forming solvent, at a concentration of 50 mM. 10 μL of sample was injected into a Stark cell assembled using two offset CaF\textsubscript{2} optical windows (12.7 mm diameter, 1 mm thickness, FOCtek Photonics), which were coated with a 4.5 nm layer of nickel metal and separated by two Teflon
spacers (~26 µm thickness). The cell was immediately immersed into a custom-built cryostat filled with liquid nitrogen. The Stark spectra were recorded by a Bruker Vertex 70 spectrometer at 1 cm⁻¹ resolution for 64 scans in the presence and absence of an externally applied electric field, respectively. The external field ranged from 0.5 to 1.4 MV/cm, corresponding to 1.5 to 3.6 kV voltage, which was applied via a Trek 10/10 high-voltage power amplifier. Stark spectra were obtained by taking the difference between the field-on and field-off transmission spectra (Figure S14). The linear Stark tuning rates were obtained from the contribution of the second derivative of the low temperature absorbance spectrum to the Stark spectrum (Table S14)¹⁶. The determined Stark tuning rate is reported as a product of the local field factor, which reflects the difference between the applied electric field and the actual local electric field that is experienced by the vibrational probe, with the |Δμ| extracted from solvatochromism.

**Isotope-Edited FTIR Spectroscopy**

The TEM proteins were concentrated to a volume of 60 µL at a concentration of 2.3 mM using a 10,000 Da MWCO Amicon spin filter (Millipore). The proteins were then solvent exchanged with a D₂O buffer containing 100 mM NaCl and 50 mM KPi (pD 7.4) using Micro Bio-Spin P-6 gel column (Bio-Rad). The protein solution was allowed to sit at 4°C overnight for sufficient H-D exchange. Afterwards, the protein solution was solvent exchanged again with the D₂O buffer before splitting into two portions of equal volume (~30 µL). The ligands (AVB or PenG), both ¹³C-labeled and unlabeled, were prepared in the same D₂O buffer at a concentration of 25 mM, respectively. 4 µL of the ¹³C-labeled and unlabeled ligand solution were added into each portion of the protein solution, respectively. The two samples were each mixed using pipette and spun down. For proteins of large K_M, 8 µL of the ligand solutions were used. The protein samples were directly used for FTIR measurements, except for A237Y e–AVB, which was incubated at room temperature for 2 h, due to the slow formation of the covalent bond. Mass spectra were taken for samples of covalent complexes. These samples were quenched by using MeOH/H₂O (1:1, v/v) both before and after the FTIR measurements, which confirmed the formation and stability of the covalent bonds (Table S3).

Infrared spectra were taken similarly to the protocol described in the section of FTIR Spectroscopy — Vibrational Solvatochromism, except for the following differences. The dry air purging was carried out for 10 min. 512 scans were acquired on the interferometer and averaged. Nine spectra were collected for each sample. For every three spectra, the sample cell mounted on the cell holder was rotated by 120°. The FTIR spectra of free AVB and Pen-G, both ¹³C-labeled and unlabeled, were also collected for use as references, including AVB sodium (AVB-Na, Advanced ChemBlocks), AVB tetrabutylammonium (AVB-NBu₄, synthesized from BODAC using the same protocol in the section of Synthesis of ¹³C-Avibactam), AVB-¹³C-NBu₄ (Figure S24 and Table S15), PenG sodium (Sigma Aldrich), and PenG-¹³C (ref.²⁷).

The procedure for data processing has been described in detail in previous works. Briefly, the raw transmission spectra were converted to absorption spectra. To remove protein background, difference FTIR spectra were generated by subtracting the protein spectra of the ¹³C-labeled ligand from that of the unlabeled one. A scaling factor close to unity was typically used to compensate for slight differences in sample concentration and pathlength. The resulting spectra were further
processed by removing the broad peaks associated with the unbound, free ligands, and the weak, broad peaks associated with protein amide backbone experiencing H-D (back) exchange. The latter were obtained by taking the difference from the 9 spectra of the same sample (either the $^{13}$C-labeled or the unlabeled). Finally, the spectra were baselined guided by the offset (35-55 cm$^{-1}$) between the positive and negative peaks, the frequency shift due to $^{13}$C-labeling. This process was repeated for the 9 spectra taken for each sample to ensure consistency. To deconvolve the difference spectra where positive $^{12}$C peaks overlap with the negative $^{13}$C peaks, the difference spectra was fitted into component peaks as Gaussian-Lorentzian sums$^{28}$, using a script we wrote (Script 1).

In parallel to the half-manual-half-automatic procedure, we further wrote a script (Script 2) that can perform the whole procedure from spectrum subtraction, baselining, all the way to peak fitting. In this way, human bias can be minimized, especially when peak fitting is sensitive to the choice of baselines. We used this script to analyze the spectra for TEM–AVB, which have high intensity of the difference spectra and thus are suitable for the fully automated data processing. All the nine $^{12}$C spectra and nine $^{13}$C spectra were loaded into the script, generating nine difference spectra, which are all used as the input for the fitting program. The fitting generated nine baselines, each for one difference spectra, and a single set of $^{12}$C and $^{13}$C peaks as Gaussian-Lorentzian sums.

Polarizable MD Simulations of Proteins

**Protein and ligand preparation:** A total of 5 protein/ligand complexes, WT TEM•AVB, WT TEM–AVB, WT TEM•PenG, A237Y TEM•PenG, and A237Y• PenG were modeled for molecular docking and molecular dynamics simulations. We chose the crystal structure of AVB covalently bound to WT TEM-1-native (PDB: 8DE0) as the template. We used the Mutate Residue function in the Maestro$^{29}$ Build menu to introduce the T182M and A184V mutations to match the amino acid sequences of the models with those of the proteins measured by FTIR. To create the non-covalent complexes, the covalently bound AVB was deleted, followed by the S70G mutation. Non-covalently bound AVB was built using the 2D-builder in Maestro. Non-covalently bound PenG was extracted from the crystal structure 1FXV$^{30}$. Both ligands were minimized using the OPLS4 force field$^{31}$ in Macromodel before docking. To make the A237Yc mutant of TEM•PenG, atom point mutation was used to convert the corresponding amide in A237Y to an ester.

**Molecular docking and fixed-charge molecular dynamics simulations:** AVB and PenG were manually docked into the oxyanion hole in the TEM binding pockets, ensuring optimal distance between the carbonyl oxygen and the S70G and A237 or A237Y amides. To reduce steric clashes between the docked ligands and the proteins in both the non-covalently and the covalently bound complexes, they were minimized using Macromodel. The systems were then solvated with TIP3P waters and neutralized with Na$^+$ in an optimized truncated octahedron with a 10 Å buffer distance between protein and boundary. Each complex was subjected to minimization and heating according to the built-in protocol in Desmond, and five replicates of 100 ns molecular dynamics simulations using the OPLS4 force field and the NPT ensemble at 298 K.

The distances between carbonyl oxygen and amide nitrogen were measured and plotted for all simulation trajectories. Trajectories where these distances were stable throughout the
simulation were manually inspected and single frames were chosen as starting structures for the following simulations.

**Polarizable molecular dynamics simulations:** Polarizable molecular dynamics simulations were performed using Tinker9 (ref. 32) using the AMOEBA BIO18 force field 33,34. The ligands, PenG and AVB, as well as the residues of A237Ye, the glycine adjacent to A237Ye, and the serine-AVB covalent adduct were parameterized using Poltype2 (ref. 35). Electrostatic parameters were obtained from fitting the electrostatic potential of 6 conformers at the MP2/aug-cc-pvtz level, and torsion parameters were refined based on dihedral scans at the WB97x-d/aug-cc-pvtz level with steps of 20–30° (depending on the number of parameters to be fit). The residues of A237Ye, the glycine adjacent to A237Ye, and the serine-AVB covalent adduct were capped using acetyl and dimethyl amine groups with parameters taken from the AMOEBA BIO18 force field (kept fixed during the parameterization) to ensure compatibility with the force field. For the serine-AVB covalent adduct, AMOEBA BIO18 parameters were used to describe the entire capped serine backbone, while the parameters for the sidechain were refined. To ensure that the residues of A237Ye and the serine-AVB covalent adduct carry a total charge of 0 and -1, respectively, the charge of Cα atoms was adjusted (by ca. ±0.05). Structures from fixed-charge molecular dynamics were minimized and equilibrated as NVT (Bussi thermostat; Ewald and vdw cutoffs of 7 and 12 Å; induced dipole threshold of \(10^{-4}\) D; 1 fs steps) and NPT (Monte Carlo barostat; otherwise, similar conditions) ensembles for 500 ps each. Three production runs were performed over 10 ns for each system. Electric fields along the β-lactam C=O bond were determined as the averaged fields on the C and O atoms projected along the C=O bond axis, as reported previously 14.

**Kinetic Measurement on PenG Hydrolysis by β-Lactamases**

**Overview:** The rate of PenG acylation (\(k_{ac}\)) was determined by first fitting full-time kinetic data to the integrated Michaelis-Menten equation 36,37 (Figure S32), followed by decoupling \(k_{ac}\) from \(k_{cat}\) using steady-state MS (ref. 3) (Figure S33).

**Full-Time Kinetics for Obtaining \(k_{cat}\) and \(K_M\):** Conventional Michaelis-Menten kinetic measurements require multiple steady-state reactions with varying substrate concentrations. For each reaction, only the initial slope is used, and the rest of the data is wasted. \(k_{cat}\) and \(K_M\) can be extracted from a single run of the reaction because the concentration vs time function is just the integrated form of the rate vs concentration function (see below), and thus also fully governed by \(k_{cat}\) and \(K_M\) (ref. 36,37). However, using the integrated function in an explicit form requires access to the Lambert function (\(W\)), which is only recently incorporated into data processing software such as Matlab 2. A full-time kinetic measurement that runs all the way to completion, although seldomly used, offers great advantages over conventional Michaelis-Menten kinetics, including fewer reactions to be run and more robust data fitting 38 against hundreds of data points that can be easily collected by measuring the concentration over time. The only requirement of full-time kinetics is that the reaction has to start from a substrate concentration at least comparable to \(K_M\) and continues until the substrate is fully exhausted. Therefore, the reaction cannot be too slow (reasonable \(k_{cat}/K_M\)) so that the data can be collected in a practical period. In the following we recapitulate the derivation 36,37 in order to define parameters that are derived and connect to the reactions described in this manuscript.
To derive the integrated form of the Michaelis-Menten equation, consider a typical enzymatic reaction:

\[
E + S \xrightleftharpoons[k_i]{k_f} E\cdot S \xrightarrow{k_{cat}} E + P \quad \text{eq. 1},
\]

where E is the enzyme, S is the substrate, E•S is the Michaelis complex, and P is the product. The rate of the reaction, \(v\), can be expressed as:

\[
v = -\frac{d[S]}{dt} = \frac{k_{cat}[E]_0[S]}{K_M + [S]} \quad \text{eq. 2},
\]

where \([E]_0\) is the initial concentration of the free enzyme, \(t\) is time, \(K_M\), the Michaelis constant, equals to \(k_{-1}/k_1\). If we integrate eq. 2, we will obtain:

\[
K_M \ln \left( \frac{[S]_0}{[S]_t} \right) + [S]_0 - [S]_t - k_{cat}[E]_0 t = 0 \quad \text{eq. 3},
\]

where \([S]_0\) is the initial concentration of the substrate, and \([S]_t\) is the concentration of the substrate at time \(t\). If we rearrange all the terms involving \([S]_t\) to the left side and \(t\) to the right side, we obtain:

\[
\frac{[S]_t}{K_M} + \ln \frac{[S]_t}{[S]_0} = \ln \left( \frac{[S]_0}{K_M} e^{-\frac{k_{cat}[E]_0 t}{K_M}} \right) \quad \text{eq. 4}.
\]

To solve \([S]_t\) as a function of \(t\), we need to write this equation in a closed form. This requires using the Lambert function \(W(x)\), which satisfies:

\[
W(x) + \ln(W(x)) = \ln(x) \quad \text{eq. 5}.
\]

If we let \(x\) to be the term in \(\ln()\) on the right side of eq. 4, \(W(x)\) becomes \([S]_t/K_M\):

\[
[S]_t = K_M W \left( \frac{[S]_0}{K_M} e^{-\frac{k_{cat}[E]_0 t}{K_M}} \right) \quad \text{eq. 6}.
\]

Now there is only one term involving \([S]_t\) on the left side, so \([S]_t\) can be determined at any \(t\) given known values of \(k_{cat}\) and \(K_M\). Conversely, if we measure the change of substrate concentration \(([S]_t)\) over time \((t)\), we can fit the data to eq. 6 and obtain \(k_{cat}\) and \(K_M\).

TEM-1 β-lactamases were prepared in 50 mM potassium phosphate (KPi, pH 7.0) and their concentrations were determined using the absorbance at 280 nm measured by a Lambda 365 (Perkin-Elmer) UV-Vis spectrometer. The extinction coefficients are listed in Table S3. The TEM-1 samples were then diluted into a reaction buffer containing 0.02% (w/v) bovine serum albumin (BSA) and 50 mM KPi (pH 7.0). The dilution fold was chosen such that the hydrolysis reactions to be run can finish in a suitable period (longer than 10 s but shorter than 10 min). BSA in the reaction buffer acts to passivate the surfaces of cuvettes and mixing tubing, thus minimizing the adhesion of the TEM-1 β-lactamases. PenG (Sigma-Aldrich) solutions were also prepared in the reaction buffer. The final concentration of PenG was chosen at least 2-fold higher than the estimated \(K_M\).
The hydrolysis reaction was monitored on a Cary 6000i UV-Vis-NIR spectrometer equipped with a stopped-flow accessory (SFA-20 Rapid Mixing Accessory), which is not always necessary but makes fast reactions (< 1 min) convenient to measure. Prior to use, all cuvettes and mixing tubing were passivated with a solution containing 0.1% (w/v) BSA and 50 mM KPi (pH 7.0) for at least 30 min before washing with the reaction buffer. The TEM-1 solution and the PenG solution were loaded into the syringes of the stopped-flow accessory and quickly mixed in the cuvette. The consumption of PenG during the reaction was monitored at 232 nm (Δε_{232} = -940 M^{-1} cm^{-1}) every 0.2 s until completion of the reaction. The experiment from mixing to monitoring was repeated 5 times.

With the measured time-dependent change in PenG concentration, \( k_{\text{cat}} \) and \( K_M \) were determined using the closed-form (explicit) solution to the integrated Michaelis-Menten equation (eq. 6), which utilizes the Lambert function. The corresponding Matlab script is presented in our previous work. A modified scripted that automatically reads the initial substrate concentration and makes an initial guess for \( k_{\text{cat}} \) and \( K_M \) is provided in Script 3. Note that the final concentration of TEM-1 should be half of that of the prepared TEM-1 samples due to the 1:1 mixing with the PenG solutions. An example of data fitting for PenG hydrolysis by the WT is shown in Figure S32. The average and standard deviation are based on 5 repeats and presented in Table S1.

**Steady-State Mass Spectrometry for Obtaining \( k_{\text{ac}} \) and \( k_3 \):** PenG hydrolysis by class A \( \beta \)-lactamases follows a two-step mechanism:

\[
\begin{align*}
E + S \quad &\underset{k_1}{\overset{k_2}{\rightleftharpoons}} \quad E \cdot S \quad \underset{k_3}{\rightarrow} \quad E - S \quad \rightarrow \quad E + P
\end{align*}
\]

(eq. 7)

After formation of the Michaelis complex, TEM•PenG (E•S), PenG first acylates Ser70 in \( \beta \)-lactamases, producing a covalent acyl complex, TEM–PenG (E–S), which is then hydrolyzed to give the product and the regenerated \( \beta \)-lactamases. \( k_2 \) is the rate constant of acylation, also referred to as \( k_{\text{ac}} \) in the main text. \( k_3 \) is the rate constant of deacylation.

The relationship between \( k_2 \), \( k_3 \), and \( k_{\text{cat}} \) is:

\[
k_{\text{cat}} = \frac{k_2 k_3}{k_2 + k_3}
\]

(eq. 8)

In steady state, the concentration of \( E-S \) stays constant, meaning that its formation and conversion have the same rate:

\[
k_2[E\cdot S] = k_3[E-S]
\]

(eq. 9)

Bringing eq. 9 into eq. 8, we obtain:

\[
k_2 = k_{\text{cat}} \left(1 + \frac{[E-S]}{[E\cdot S]}\right)
\]

(eq. 10)

To deconvolve \( k_{\text{cat}} \) into \( k_2 \) and \( k_3 \), the molar ratio between \( E-S \) and \( E\cdot S \) in steady state was measured using mass spectrometry (Figure S33). To prepare a steady state that can last long
enough for manual mixing and quenching (ideally >20 s), the concentrations of TEM-1 and PenG were carefully chosen according to the determined $k_{cat}$ values. For example, the WT was prepared in 500 mM KPi buffer (pH 7.0) at a concentration of 10 mg/mL. 5 µL of the TEM-1 solution was diluted with 450 µL of 500 mM KPi (pH 7.0), into which 450 µL of 200 mM PenG in 500 mM KPi (pH 7.0) was quickly added and mixed by pipette. Immediately, 900 µL of methanol was added into the reaction mixture to quench the steady state on time. A 1:1 (v/v) mixture of methanol and water is effective to rapidly denature TEM-1 and keeps it denatured without turning over PenG substrate. It is important to use a high concentration of PenG in a small volume rather than a low concentration in a big volume, because high concentrations minimize (1) the impact of diffusion and (2) the free enzyme’s concentration which leads to overestimation of the non-covalent complex’s concentration.

Afterwards, the mixture was diluted into 40 mL of methanol/water (1:1, v/v) and concentrated into a final volume of 50 µL using a 3,000 Da MWCO Amicon spin filter (Millipore). The purpose of this step is to reduce the concentration of PenG in the final sample for mass spectrometry, because a high concentration of PenG enables the acylation of even denatured proteins during UPLC/ionization, leading to artificial enhancement of the covalent complex. The dilution before the concentration is necessary because it brings KPi’s concentration down, otherwise the denatured proteins tend to leak through the spin filter membrane. Finally, the denatured proteins are in a suitable concentration (33 µM) for mass spectrometry. Three independent samples were prepared for each mutant. The above protocol was optimized based on previous works\textsuperscript{3,40}.

In the deconvoluted mass spectra, the ratio between non-covalent and covalent complexes (mass listed in Table S3) was quantified based on their peak area in the deconvoluted mass spectra (Figure S33 as an example of WT) and averaged over three independent repeats. The bootstrapping method for determining the median and confidence intervals of $k_2$ and $k_3$ was presented in our previous work\textsuperscript{3} (see “RateDistribution.m”) and the results are displayed in Table S1.

**Kinetic Measurement on AVB Inhibition of β-Lactamases**

**Overview:** $k_{cbm}$ is for the first time determined separately from the second-order rate constant $k_{cbm}/K_i$ by using the stopped-flow technique (Figure S34) to record the initial rate of substrate turnover in the inhibition assay\textsuperscript{41,42} (Figures S35-S37). $k_{rec}$ was obtained from the final rate of substrate turnover when the formation of TEM–AVB reaches equilibrium, rather than being estimated by $k_{off}$ as in prior literature\textsuperscript{43} (Text S5).

**Mechanism of Covalent Inhibition:** Avibactam (AVB) inhibits TEM β-lactamases by fast binding to the active site, forming a non-covalent complex, TEM•AVB (E•I, where I stands for the inhibitor), which is followed by slow formation of a covalent bond with S70. The generated carbamoyl complex, TEM–AVB (E–I), can undergo a reverse reaction to regenerate the non-covalent E•I complex; however, the equilibrium highly favors the covalent side. The mechanism can be written as:

$$E + I \overset{k_4}{\rightleftharpoons} E\cdot I \overset{k_5 \cdot (k_{cbm})}{\rightarrow} E\cdot I \overset{k_6 \cdot (k_{rec})}{\rightarrow} E\cdot I$$

**eq. 11,**
where $k_5$ is the rate constant of carbamylation, equivalent to $k_{cb}$ used in the main text, and $k_{-5}$ is the rate constant of recyclization, equivalent to $k_{rec}$ used in the main text. The equilibrium constant for the non-covalent binding is $K_i$, following $K_i = k_{-4}/k_4$, similar to the $K_i$ for non-covalent inhibitors and $K_M$ for substrates. The overall dissociation constant, $K_i^*$ follows:

$$K_i^* = \frac{[E][I]}{[E\cdot I] + [E\cdot I]} = \frac{K_i k_{-5}}{k_5 + k_{-5}}$$

The kinetics of covalent inhibition have been thoroughly studied in the literature\textsuperscript{42}. Here we recapitulate the derivation and present the equations specific to the kinetic behavior of AVB. To measure AVB inhibition kinetics, a substrate is mixed with AVB sodium (Advanced ChemBlocks), and then together mixed with β-lactamases. The real-time rate of the substrate’s conversion to product (eq. 1) characterizes the amount of the β-lactamases that has not been inhibited by AVB yet and is still available at that moment. The slowing down of product formation over time describes the progression of AVB inhibition:

The product-time trace can be written as\textsuperscript{42}

$$[P] = v_s t + \frac{(v_o - v_s)(1 - e^{-k_{obs}t})}{k_{obs}}$$

where $v_0$ stands for the initial rate and $v_s$ for the final rate of product formation. The physical meaning of $k_{obs}$ is clearer in the derivative form of eq.13:

$$\frac{d[P]}{dt} = v_s + (v_0 - v_s)e^{-k_{obs}t}$$

Eq. 14 shows an exponential decay of the rate in product formation from the initial rate $v_0$ to the final rate $v_s$. $k_{obs}$ is the constant for how fast the exponential decay is.

The initial rate $v_0$ is the highest rate during the reaction time course because no E–I has been generated. At that moment, the concentration of the total enzyme, $[E]_0$, is partitioned between that of E•I, E•S, and the free enzyme E. The proportion of E•S in the total enzyme dictates the rate in production formation:

$$v_0 = k_{cat}[E]_0 \frac{[E\cdot S]}{[E] + [E\cdot I] + [E\cdot S]}$$
Given that:

\[
\frac{[E\cdot S]}{[E] + [E\cdot I] + [E\cdot S]} = \frac{[S]}{K_M} \quad \text{eq. 16,}
\]

we obtain:

\[
v_0 = k_{cat}[E]_0 \frac{[S]}{1 + \frac{[I]}{K_i} + \frac{[S]}{K_M}} \quad \text{eq. 17.}
\]

The rate reaches the final rate \(v_s\) when the equilibrium between \(E\cdot I\) and \(E-I\) is established. Now the total enzyme has to be additionally partitioned into \(E-I\), along with \(E\cdot I\), \(E\cdot S\), and the free enzyme:

\[
v_s = k_{cat}[E]_0 \frac{[E\cdot S]}{[E] + [E\cdot I] + [E-I] + [E\cdot S]} \quad \text{eq. 18.}
\]

Given that:

\[
\frac{[E\cdot S]}{[E] + [E\cdot I] + [E-I] + [E\cdot S]} = \frac{[S]}{K_M} \quad \text{eq. 19,}
\]

We obtain:

\[
v_s = k_{cat}[E]_0 \frac{[S]}{1 + \frac{[I]}{K_i^*} + \frac{[S]}{K_M}} \quad \text{eq. 20.}
\]

The decay from \(v_0\) to \(v_s\) therefore reflects the decrease of the enzyme that is available to form \(E\cdot S\) due to the conversion from \(E\cdot I\) to \(E-I\). This conversion is first order, and that is why the decay takes an exponential form in eq. 14. This conversion is also reversible, so its rate constant \(k_{obs}\) is the sum of the reverse and the forward reaction rate constant, \(k_S\) and \(k_S^{-}\), with the former modified by the proportion of \(E\cdot AVB\) in total enzyme:

\[
k_{obs} = k_S + k_S^{-} \frac{[E\cdot I]}{[E] + [E\cdot I] + [E\cdot S]} \quad \text{eq. 21.}
\]

Given that:

\[
\frac{[E\cdot I]}{[E] + [E\cdot I] + [E\cdot S]} = \frac{[I]}{K_i} \quad \text{eq. 22,}
\]

where \(\Theta\) is defined as the partition ratio, we obtain:

\[
k_{obs} = k_S + k_S^{-} \Theta \quad \text{eq. 23.}
\]

The key to obtaining \(k_S, K_i,\) and \(k_S^{-}\) in this work is to measure \(v_0\), from which \(K_i\) is fitted using eq. 17. With the obtained \(K_i\) and \(\Theta\), \(k_S^{-}\) is determined from eq. 23. In parallel, by measuring \(v_s\), \(K_i^*\) is fitted using eq. 20. Finally, \(k_S\) is obtained using \(K_i, k_S,\) and \(K_i^*\) according to eq. 12. To accurately measure \(v_0\), we used the stopped-flow technique to capture early data points; in addition, the dead time was determined to precisely position \(t_0\) on the product \(v_s\) time trace.
Choice of the Substrate for Probing the Inhibition: We chose nitrocefin (Cayman Chemical) as the substrate for probing the inhibition of WT due to the large absorbance change in nitrocefin hydrolysis. Its $\Delta \varepsilon_{486}$ was determined as 15217 M$^{-1}$ cm$^{-1}$ through complete hydrolysis of nitrocefin of known concentrations. The large $\Delta \varepsilon_{486}$ allows for sensitive probing of the inhibition progression. The values of $k_{cat}$ and $K_M$ were measured as $1000 \pm 16$ s$^{-1}$ and $65 \pm 1$ µM for the WT using full-time kinetics as described for PenG hydrolysis.

We chose PenG as the substrate for probing the inhibition of mutants other than the WT because their $K_M$ for nitrocefin hydrolysis is too large. Therefore, nitrocefin is no longer a suitable substrate to use in these inhibition experiments.

Dead-Time Measurement for the Stopped-Flow Setup: To precisely fit the initial rate $v_0$, it is important to know the delay in collecting the first data point. When solutions are mixed in the stopped-flow setup, there is a delay in the change of absorbance due to the mixing (Figure S34a). Such delay is the sum of the intrinsic dead time of the stopped-flow apparatus and the step of time in absorbance reading (every 33 ms on the Cary 6000i UV-Vis-NIR spectrometer). In analysis of absorbance-time traces, we started from the last data point that has not been affected by mixing (“-3” in Figure S34a) and count for another 3 data points forward as the arbitrarily chosen time zero (“0” in Figure S34a). To measure how much the time zero that we chose is delayed from the actual time zero, we developed an assay based on nitrocefin hydrolysis by the WT TEM-1.

The WT TEM-1 was prepared at a series of concentrations (0, 60, 120, 180, 240, and 300 nM) in a reaction buffer containing 0.02% (w/v) BSA and 50 mM KPi (pH 7.0). Nitrocefin was first prepared as a stock solution in DMSO at 160 mM and then diluted to a final concentration of 400 µM in the reaction buffer. The inhibition reaction was monitored on a Cary 6000i UV-Vis-NIR spectrometer equipped with a stopped-flow accessory. Prior to use, all cuvettes and mixing tubing were passivated with a solution containing 0.1% (w/v) BSA and 50 mM KPi (pH 7.0) for at least 30 min before washing with the reaction buffer. The TEM-1 solution and the nitrocefin solution were loaded into the syringes of the stopped-flow accessory and quickly mixed in the cuvette. The hydrolysis of nitrocefin was monitored at 486 nm ($\Delta \varepsilon_{486} = 15217$ M$^{-1}$ cm$^{-1}$) every 33 ms.

Compared with the blank control (0 WT concentration), other absorbances at the first data point are larger because the hydrolysis of nitrocefin has already taken place for the period of the deadtime. Such excess absorbance can be converted into the concentration of the initial product. The concentration of initial product varies between experiments because they were run with varying $[E]_0$. The different $[E]_0$ is manifested by the different initial rates, which were obtained as the slope of the linear fitting of the first 7 data points. The higher the $[E]_0$, the faster the initial rates. Instead of $[E]_0$, we directly used the initial rates for the following analysis. We drew a linear correlation between the initial product concentration and the initial rates and obtained the dead time as the slope (Figure S34b), which is 58 ms.

General Protocol of the Inhibition Experiment: Let us first use the WT as an example to demonstrate the inhibition experiments. The WT was prepared at a concentration of 20 nM in a reaction buffer containing 0.02% (w/v) BSA and 50 mM KPi (pH 7.0). Nitrocefin was first
prepared as a stock solution in DMSO at 160 mM and then diluted to a final concentration of 400 µM in the reaction buffer. In this way, the final concentration of DMSO in the reaction mixture is only 0.12%, minimizing its effect on the kinetics. AVB was added into the nitrocefin solutions at a series of concentrations of 10, 20, 30, 40, 50, 60, and 70 µM. In addition, a blank control without AVB was prepared. In total, there are 8 nitrocefin samples containing varying concentrations of AVB.

The inhibition reaction was monitored on a Cary 6000i UV-Vis-NIR spectrometer equipped with a stopped-flow accessory. Prior to use, all cuvettes and mixing tubing were passivated with a solution containing 0.1% (w/v) BSA and 50 mM KPi (pH 7.0) for at least 30 min before washing with the reaction buffer. The TEM-1 solution and the nitrocefin solution were loaded into the syringes of the stopped-flow accessory and quickly mixed in the cuvette. The hydrolysis of nitrocefin was monitored at 486 nm (Δε = 15217 M⁻¹ cm⁻¹) every 33 ms until the absorbance increases linearly and enough data points have been collected in the linear region. The experiment from mixing to monitoring was repeated 5 times. The same procedure was repeated for all the 8 nitrocefin samples. A representative product vs time trace of each [I] is shown in Figure S35a,b.

To analyze the experimental data, which consist of product vs time traces at each [I], we first used Script 4 to fit the inhibition features, including v₀, kₖobs and vₕ, according to eq.13 (Figure S35a,b). With the obtained v₀ and kₖobs, we then used Script 5 to fit Kᵢ and kₕ, according to eq. 17 and eq. 23, respectively (Figure S35c,d). Not all the obtained values of v₀ were used for fitting Kᵢ. Those obtained from experiments using high [I] and exhibiting large kₖobs could be unreliable, evidenced by a large variance at the same [I], and these values of v₀ are not used for fitting Kᵢ. Note that [E]₀, [S], and [I] should be half of those in preparation of the samples due to the 1:1 mixing of two solutions in the stopped-flow apparatus. The script only allows for using a constant [E]₀, therefore the product concentration can be normalized against the same [E]₀ if multiple [E]₀ values are used, as described later for A237Y. In parallel, using the obtained vₕ from Script 4, we run Script 6 to fit Kᵢ* based on eq. 20, and finally to calculate kₕ⁻¹ from kₕ, Kᵢ, and Kᵢ* based on eq. 12 (Figure S35e). The results for all the mutants are listed in Table S2.

**Choice of Concentration:** The concentrations of the WT, nitrocefin, and AVB are judiciously chosen. In order to use eq. 17 to fit Kᵢ, [I]/Kᵢ should be large enough compared with [S]/K_M to cause measurable changes of v₀ by varying [I]. The side effect of having a large [I]/Kᵢ is an accompanying large kₖobs, according to eq. 21, especially when kₕ is large. This means that the product vs time trace will quickly enter the linear region and only limited data are collected for fitting kₖobs. Therefore, [I]/Kᵢ should be designed to vary in a suitable range, which is large enough to change v₀ (by > 5%) but not too large such that kₖobs remains < 3 s⁻¹. To help satisfy the former requirement, [S]/K_M should avoid being too large. But [S] cannot be too low, otherwise the substrate consumption cannot be neglected, disallowing the treatment of [S] as a constant, which is a key assumption. In our experiments, the final [P] never exceeds 5% of [S]. This requirement can also be met by using a low [E]₀. However, [E]₀ should also be large enough to produce a sizable vₕ, otherwise a prolonged accumulation of data in the linear region is necessary. Along this line, a variation of the above protocol is to use a series of [E]₀ for different nitrocefin samples. For those nitrocefin samples with a low [I], a low [E]₀ is used to avoid a high [P]; for those with a high [I], a high [E]₀ is used to produce a sizable vₕ. In addition, [E]₀ should always be << [I].
When the inhibition constants ($K_i$, $k_5$, and $k_{-5}$) were unknown, we started with an estimation of their values based on other mutants or simply by guessing. Then we tried out $[E]_0$, $[S]$, and $[I]$ using Script 7 that simulates the corresponding $[P]$, $v_0$, $k_{\text{obs}}$ and $v_s$. If we were satisfied with the simulated $[P]$, $v_0$, $k_{\text{obs}}$ and $v_s$, we would proceed to use the $[E]_0$, $[S]$, and $[I]$ for carrying out kinetic measurements. From the experimental data, we used Scripts 4-6 to obtain the fitted $K_i$, $k_5$, and $k_{-5}$. If there is a big deviation from our guess, we may have used unsuitable values of $[E]_0$, $[S]$, and $[I]$ for the experiments. In such cases, the data may not be able to provide an estimation of $K_i$, $k_5$, and $k_{-5}$, but at least some hints on the range of their values. We then made new guesses based on the hints from the last experiment and repeated this process until we arrived at a suitable $[E]_0$, $[S]$, and $[I]$, and the experimental data allowed for reliable fitting of $K_i$, $k_5$, and $k_{-5}$. In our experience, no more than two rounds were required to obtain the final results.

**Using PenG as the Substrate:** For mutants that use PenG as the substrate to probe their inhibition kinetics, all the procedures are the same, except for monitoring the reaction at 232 nm ($\Delta \varepsilon_{232} = -940 \text{ M}^{-1} \text{ cm}^{-1}$). Due to the small $\Delta \varepsilon_{232}$, the targeted $v_0$ and $v_s$ should be correspondingly larger. Here is an example. The A237Y mutant was prepared in the reaction buffer containing 0.02% (w/v) BSA and 50 mM KPi (pH 7.0) at a series of concentrations of 0.5 (for blank control), 1, 2, 3, 4, 5, and 6 µM. PenG was prepared at a concentration of 10 mM in the reaction buffer, into which AVB was added at a series of concentrations of 0 (blank control), 100, 200, 300, 400, 500, and 600 µM, respectively. In total, there are 7 solutions of A237Y and 7 solutions of PenG containing varying concentrations of AVB. They are used in pairs for mixing in the stopped-flow apparatus. A representative product $v_s$ time trace for each $[I]$ and the fitting of the whole data is shown in Figure S36.

**A237Y Inhibition Experiment:** The ester protein exhibits relatively slow kinetics for AVB inhibition. Therefore, rather than waiting for the equilibrium of E–I formation, we performed its kinetic measurements in two separate experiments. The first experiments only monitored the first 1 h of the product $v_s$ time trace, allowing us to collect enough data for fitting $k_{\text{obs}}$ and $v_0$ using Script 4 (conditioned that $v_s = 0$), which are in turn used for the fitting of and $K_i$ and $k_5$ using Script 5. The results are shown in Figure S37a-c. A separate set of experiments was performed on samples that contain various concentrations of $[I]$ and were equilibrated for 2 days before the addition of PenG and the measurement of product $v_s$ time trace. A linear fitting produced $v_s$, which was subsequently used for fitting $K_i^*$ and $k_{-5}$ using Script 6. The results are shown in Figure S37d.
Supporting Text

Text S1. Attempts to crystallize the A237Ye mutant of TEM-1

We expressed the A237Ye variant of TEM-1-native and found this protein is very susceptible to the hydrolysis of the backbone ester. Incubation at pH 7.0 overnight led to >50% protein fragmentation at the site of the ester, as characterized by HPLC-MS. To use a stable protein for crystallization, we expressed the E166N mutant of TEM-1-native-A237Ye, which showed <1% hydrolysis of the backbone ester under the same incubation condition. We tried to crystallize this protein using both crystallization conditions (using ammonium sulfate and PEG3350 as precipitants, respectively) as described in the section Protein Crystallization and Crystallography but did not obtain crystals. We further tried to co-crystallize this protein with 5 mM AVB, which forms a covalent complex, but obtained no crystals. We noticed that E166N-A237Ye is not thermally stable enough for crystallization at room temperature. After being left at room temperature for two days, it cannot react with AVB any more as measured by mass spectrometry. We also tried to grow the crystal at 4°C using the above conditions and to vary the pH between 6.0 and 7.0 but still obtained no crystals.

E166N is a general destabilizing mutation for TEM proteins. The A237Ye mutation is more destabilizing than A237Y given the success in crystallizing E166N-A237Y. This makes sense because the ester backbone not only cannot act as a H-bond donor, but also becomes a weaker H-bond acceptor compared to a normal amide backbone. This impacts the two H-bonds from G244 and R243 to the 237 backbone, leading to an overall destabilization/conformational flexibility of the A237Ye mutant. For this reason, we chose to use 1H NMR as an alternative to characterize the structure of A237Ye.

Text S2. Rich conformational heterogeneity for TEM–AVB

An envelope of five 12C=O peaks were observed in the isotope-edited infrared spectra of TEM–AVB covalent complexes, while there are only three peaks for TEM•AVB. This is a counterintuitive observation because covalently bound ligands might be expected to be more conformationally rigid than the noncovalently bound counterparts. Although the covalent linkage is an anchor that limits the movement of AVB, it is not the only anchor through which AVB interacts with the protein active site. The sulfate and amide side groups in AVB are the other two anchors, forming significant noncovalent interactions whose contribution to the binding energy is essential. Due to competition among the anchors, multiple local minima are possible.

Comparing TEM–AVB with TEM•AVB, it is also interesting to see that the sulfate group in AVB becomes more flexible in TEM–AVB. As shown in Figure 2b, the sulfate is connected to the ring through a flexible NH bridge which was part of the rigid ring in TEM•AVB. The extra flexibility of the sulfate anchor may also lead to heterogenous positioning of the C=O.

Comparing TEM–AVB with TEM–PenG, we found that PenG experiences less conformational heterogeneity evidenced by fewer 12C=O peaks present in the isotope-edited infrared spectrum of WT TEM–PenG (ref.3). We posit that the three anchors can act more synergistically, favoring the same conformations. This is a reasonable hypothesis because TEM
proteins have evolved for billions of years to become a PenG-hydrolysis specialist\textsuperscript{44}, while AVB is a human-made compound less optimized to sit in the active site as described above.

Text S3. Largest field contributes dominantly to rate enhancement

When multiple $^{12}$C=O peaks are found, the peak of the lowest frequency was used for analyzing the electric field’s contribution to catalysis, because the lowest frequency corresponds to the largest electric field and thus contributes the most to the lowering of free energy barrier. Consider a field lower than the largest field by 20 MV/cm, or in a more relevant unit, 1 kcal/mol/D. Given a reaction difference dipole of 1.4 D, this field provides a lowering of free energy barrier 1.4 kcal/mol smaller than that by the largest field. Accordingly, the rate enhancement conferred by this field is only 0.09-fold of that by the largest field. That means the population of C=O that experiences this field can only react 0.09-fold as fast as that experiencing the largest field. Therefore, the largest field plays a dominant role in promoting reactions.

A more rigorous treatment is to sum the contribution of each observed field population. However, the weight of each field population is not simply their peak area because the C=O bond in different electrostatic environments may have somewhat different transition dipole moments for the vibration\textsuperscript{45}. The peak area analysis may be further complicated by the baselining of the difference spectra. Therefore, we chose to only pick the largest field for a more straightforward data analysis.

Text S4. Change in magnitude of electric field due to the ester mutation

Compared with A237Y, A237Ye is found to exert electric fields onto the C=O that are smaller in magnitude by 61 MV/cm for TEM•AVB and by 67 MV/cm for TEM–AVB. However, for TEM•PenG, the field decrease due to the ester backbone is only 25 MV/cm. The remaining field (-138 MV/cm) is still very large. To explain this observation, we hypothesize that the remaining large fields arise from not only the H-bond from S70 backbone amide but also oriented dipoles across the whole enzyme scaffold whose preorganization has been optimized through evolution. This hypothesis is supported by polarizable MD simulations on TEM•PenG where frames displaying no H-bonds for PenG’s β-lactam C=O are still calculated with electric fields more negative than -100 MV/cm (Figure S31d-f).

Text S5. Comparison with previous kinetic studies of AVB inhibition

Previous kinetic studies on AVB inhibition used eq. 22 & 23 to determine the value of $k_{cbm}/K_i$, but not their individual values, as exemplified by the original work\textsuperscript{43}. Also, $k_{rec}$ has never been measured, but instead $k_{off}$ has been a common target of measurement. $k_{off}$ is often measured as the value of $k_{obs}$ in eq. 22 & 23 when [I] is very low, such as that in a jump-dilution experiment. However, the contribution from $k_{cbm}$ still cannot be neglected when $k_{cbm} >> k_{rec}$, which is exactly the case here. Therefore, the $k_{off}$ values reported in the literature are not a good approximation of $k_{rec}$, but rather depend on how low [I] is in a specific experiment.
Figure S1. Crystal structure of WT TEM-1 (PDB: 7U6Q) with A237 highlighted. The methyl group of A237 is largely exposed to solvent.
Figure S2. Overlays of the crystal structures of WT (yellow, PDB: 7U6Q) and A237Y (cyan, PDB: 8DDZ) TEM-1. (a) Global structures. (b) Key amino acids and structured water molecules in the active sites. The position of the A237 residue is slightly perturbed, with the methylene group in A237Y shifted away from the original methyl group in WT.
Figure S3. Overlays of the crystal structures of WT TEM-1 (yellow, PDB: 7U6Q) and TEM–AVB (green, PDB: 8DE0). (a) Global structures. (b) AVB, key amino acids, and structured water molecules in the active sites.
Figure S4. Noncovalent interactions of AVB bound to TEM-1 (PDB: 8DE0). (a) Comprehensive view. (b) View from the left side, focusing on the sulfate group. (c) View from the back side. (d) View from the right side, focusing on the amide side group. The electron density maps of AVB (2mFo-DFc, 1.5σ) are depicted.
**Figure S5.** Overlays of the crystal structures of WT (green, PDB: 8DE0) and A237Y (magenta, PDB: 8DE1) TEM–AVB. (a) Global structures. (b) AVB, key amino acids, and structured water molecules in the active sites.
Figure S6. Overlays of the crystal structures of A237Y TEM (cyan, PDB: 8DDZ) and TEM–AVB (magenta, PDB: 8DE1). (a) Global structures. (b) AVB, key amino acids, and structured water molecules in the active sites.
Figure S7. HPLC chromatograms of *E. coli* lysates for testing the metabolism of HPLA (starting from 1 mM). (a) DH10B cells were incubated at 37°C until OD$_{600}$ 1.2 before cell lysis. The peak area gives a HPLA concentration of 0.684 mM. (b) DH10B cells were incubated at 20°C overnight before cell lysis. The peak area gives a HPLA concentration of 0.039 mM. (c) DH10B ΔaspC ΔtyrB cells were incubated at 37°C until OD$_{600}$ 1.2 before cell lysis. The peak area gives a HPLA concentration of 1.078 mM. (d) DH10B ΔaspC ΔtyrB cells were incubated at 20°C overnight before cell lysis. The peak area gives a HPLA concentration of 0.657 mM. The method of HPLC is the same as described for HPLC-MS. The peak area of HPLA is calibrated by using HPLA solutions of known concentrations.
Figure S8. Characterization of A237Y using HPLC-MS (lower resolution). (a) HPLC chromatogram showing three components. (b) MS spectra of the three components, with i being the smaller fragment, ii being the whole protein, and iii being the larger fragment. The formation of the fragments is due to the hydrolysis of the ester bond during protein purification. (c,d) The ester bond is completely hydrolyzed by treating the protein with 0.5 M NaOH for 1 h, evidenced by the disappearance of the whole protein peak in the chromatogram (c) and the MS spectra (d).
Figure S9. $^1$H NMR spectra of the A237Y and A237Ye mutants of TEM-1 (0.2 mM) in 20 mM potassium phosphate (pH 6.0, 10% D$_2$O, 15°C). Water suppression was achieved using the WET method$^{10}$. These spectra and the enlarged versions of the aromatic and aliphatic regions (Figures S10 and S11, respectively) are nearly identical in every detail suggesting that the solution structure of A237Ye is very similar to A237Y for which a crystal structure was obtained (see Figure S2).
**Figure S10.** Downfield amide/aromatic region of $^1$H NMR spectra of the A237Y and A237Ye mutants of TEM-1. See Figure S9 caption for experimental details.
**Figure S11.** Upfield aliphatic region of $^1$H NMR spectra of the A237Y and A237Y<sup>e</sup> mutants of TEM-1. See Figure S9 caption for experimental details.
Figure S12. Calibration of BODAC as a vibrational Stark probe. (a) Structural analogue between TEM•AVB and BODAC. The charged sulfate group is replaced with a neutral benzyl group to avoid complicated MD simulations and to allow a broad range of solubility. (b) Infrared spectra of the urea C=O in BODAC dissolved in solvents of varying polarity. (c) Urea C=O infrared frequencies (by peak picking) plotted against the calculated electric fields based on MD simulations. A linear correlation provides the sensitivity of frequency shift to field (slope) as 0.76 ± 0.02 cm⁻¹/(MV/cm) and the vibrational frequency in vacuum (intercept) as 1779.8 ± 0.8 cm⁻¹. (d) A wider range of the infrared spectra of BODAC dissolved in solvents of varying polarity. The peaks at lower frequencies come from the vibration of the C=O in the side amide group. The peak intensities were normalized based on the intensities of the urea C=O.
Figure S13. Calibration of MPC as a vibrational Stark probe. (a) Structural analogue between TEM–AVB and MPC. The charged sulfate group is removed to avoid complicated MD simulations and to allow a broader range of solubility. The amide side group is also removed for simplification. (b) The infrared spectra of MPC dissolved in solvents of varying polarity. (c) C=O infrared frequencies (by peak picking) were plotted against the calculated electric fields based on MD simulations. A linear correlation provides the sensitivity of frequency shift to field (slope) as 0.75 ± 0.03 cm⁻¹/(MV/cm) and the vibrational frequency in vacuum (intercept) as 1714.0 ± 1.1 cm⁻¹.
Figure S14. Infrared absorption and Stark spectra of MPC (50 mM in 2-methyltetrahydrofuran) at 77 K. (a) Infrared absorption spectrum. (b) Vibrational Stark spectrum scaled to an applied field of 1.0 MV/cm. The Stark spectrum was fitted to a sum of the derivatives of the absorption spectrum. See Table S14 for the fitting details.
Figure S15. The synthetic route to $^{13}$C-labeled AVB.
Figure S16. Liquid $^1$H NMR spectrum of 2 in CDCl$_3$. 
Figure S17. Liquid $^{13}$C NMR spectrum of 2 in CDCl$_3$. 
Figure S18. Liquid $^1$H NMR spectrum of 3 in CDCl$_3$. 
Figure S19. Liquid $^1$H NMR spectrum of 4 in CDCl$_3$. 
Figure S20. Liquid $^{13}$C NMR spectrum of 4 in CDCl$_3$. 
Figure S21. Liquid $^1$H NMR spectrum of 5 in CD$_3$OD.
Figure S22. Liquid $^1$H NMR spectrum of $^{13}$C-labeled AVB in CDCl$_3$. 
Figure S23. Liquid $^{13}$C NMR spectrum of $^{13}$C-labeled AVB in CDCl$_3$. 
Figure S24. Infrared spectra of 10 mM AVB variants in D$_2$O buffer (100 mM NaCl, 50 mM KPi, pD 7.4) in the carbonyl region. (a) Overlay of the infrared absorption spectra for AVB sodium salt, AVB tetrabutylammonium salt, and the $^{13}$C-labeled AVB tetrabutylammonium salt. (b) Difference spectrum between AVB sodium salt and the $^{13}$C-labeled AVB tetrabutylammonium salt. (c) The spectrum of AVB sodium salt fitted to two peaks, corresponding to the urea C=O and the side amide C=O. (d) The spectrum of the $^{13}$C-labeled AVB tetrabutylammonium salt fitted to two peaks, corresponding to the urea C=O and the side amide C=O.
Figure S25. $^{12}$C-$^{13}$C difference infrared absorption spectra for TEM•AVB (trapped by S70G mutation). (a) WT. (b) A237Y. (c) A237Ye. (d) A237E. (e) A237R. (f) A237W. The experimental curve (black) is fitted to a sum (grey) of $^{12}$C (positive) and $^{13}$C (negative) peaks. Peaks belonging to the same positive-negative pair are filled with the same color.
Figure S26. $^{12}$C-$^{13}$C difference infrared absorption spectra for TEM–AVB. (a) WT. (b) A237Y. (c) A237Yc. (d) A237E. (e) A237R. (f) A237W. The experimental curve (black) is fitted to a sum (grey) of $^{12}$C (positive) and $^{13}$C (negative) peaks. Peaks belonging to the same positive-negative pair are filled with the same color.
Figure S27. Conformation of AVB bound to TEM. (a) WT (PDB: 8DE0). (b) A237Y (PDB: 8DE1). (c) A237Y at room temperature (PDB: 8DE2). The electron density maps of AVB and S70 (2mFo-DFc, 1.5σ) are depicted.
Figure S28. Conformations of AVB bound to CTX-M-14 solved at 0.83 Å (PDB: 6mz2)\(^{36}\). (a) View focusing on three conformations of the sulfate group. (b) View showing the two key H-bonds. The O–N distances between the carbamate C=O and the backbone amides vary from 2.79(1) to 3.01(1) for the H-bond to S237 and from 2.80(1) to 2.55(1) for the H-bond to S70. The electron density maps (2mF\(_o\)−DF\(_c\), 1.0\(\sigma\)) are depicted.
Figure S29. Polarizable MD simulations and electric field calculations on WT TEM•AVB (trapped by S70G mutation) and TEM–AVB. (a) Representative frames of TEM$^{S70G}$•AVB (dark green) and TEM–AVB (bright green) overlapped based on the whole protein structures. The orientation of AVB’s C=O changes after the covalent bond formation. (b,c) Plots of H-bond lengths and electric field magnitudes for AVB’s reactive C=O in TEM$^{S70G}$•AVB (b) and TEM–AVB (c). H-bonds were defined as those with O-N distance < 3.5 Å and H-N-O angle < 30°. For those missing a qualified H-bond, the data are presented on the top and right side with bond lengths shown as “NA”. The electric fields were calculated using the AMOEBABIO18 force field. (d,e) Magnitudes of the electric fields experienced by AVB’s reactive C=O, as calculated using polarizable force fields on MD frames (d), and as obtained from infrared spectroscopy (the largest-field population duplicated from Figure 3f,g) (e). The calculated average fields for TEM$^{S70G}$•AVB and TEM–AVB are -136 and -99 MV/cm, respectively. The experimentally measured fields for TEM$^{S70G}$•AVB and TEM–AVB are -150 and -97 MV/cm, respectively.
Figure S30. $^{12}$C-$^{13}$C difference infrared absorption spectra for TEM•PenG (trapped by S70G mutation). (a) WT. (b) A237Y. (c) A237Ye. (d) A237E. (e) A237R. (f) A237W. The experimental curve (black) is fitted to a sum (grey) of $^{12}$C (positive) and $^{13}$C (negative) peaks. Peaks belonging to the same positive-negative pair are filled with the same color.
**Figure S31.** Polarizable MD simulations and electric field calculations on TEM PenG (trapped by S70G mutation). (a-c) Representative frames of WT (a), A237Y (b), and A237Ye (c). The key H-bonds are highlighted in red dashed lines and labeled with their bond lengths and electric fields exerted onto PenG’s β-lactam C=O. The electric fields were calculated using the AMOEBABIO18 force field. (d-f) Plots of H-bond lengths and electric field magnitudes for PenG’s β-lactam C=O in WT (d), A237Y (e), and A237Ye (f). H-bonds were defined as those with O-N distance < 3.5 Å and H-N-O angle < 30°. For those missing a qualified H-bond, the data are presented on the top and right side with bond lengths shown as “NA”. For those missing both H-bonds, the data are presented on the top right corner. All the datapoints in (f) are arranged in one dimension on the top because the O-N^A237Y distance for the ester protein is always “NA” and thus the O-N^S70G distance becomes the only variable. Compared with the WT, the A237Y mutant displays slightly elongated H-bonds and correspondingly smaller electric fields. A237Ye possesses at most one H-bond and shows correspondingly large drops in field magnitudes. (g,h) Magnitudes of the electric fields experienced by PenG’s β-lactam C=O in WT, A237Y, and A237Ye, as calculated using polarizable force fields on MD frames (g), and as obtained from infrared spectroscopy (the largest-field population duplicated from Figure 3h) (h). The calculation shows similar trends compared to the experimental results. The calculated average fields for WT, A237Y, and A237Ye are -169, -164, and -127 MV/cm, respectively. The experimentally measured fields for WT, A237Y, and A237Ye are -171, -163, and -138 MV/cm, respectively.
Figure S32. Full-time kinetic measurement on PenG hydrolysis catalyzed by TEM-1. The experimental curve was fitted to eq. 6, giving a fitted value of $k_{\text{cat}}$ and $K_M$. 

$[\text{WT}] = 0.02 \, \mu\text{M}$

$k_{\text{cat}} = 1742 \, \text{s}^{-1}$

$K_M = 20.43 \, \mu\text{M}$
Figure S33. Mass spectrum of a mixture of PenG and TEM-1 quenched at a steady state. The peak area confers the concentration ratio between TEM and TEM–PenG, which can be used to determine $k_{ac}$ according to eq. 10.
Figure S34. Dead-Time measurement of the stopped-flow apparatus. (a) Jump in absorbance when reactants were mixed using the stopped-flow technique in a typical kinetic experiment. Counting from the last data point that has not been affected by mixing ("-3"), we choose the datapoint "0" as the point of time zero. (b) Plot of the product concentration at time zero $[P]_0$ against the reaction rate at time zero $v_0$ for nitrocefin hydrolysis by the WT TEM-1. The slope of a linear regression affords the value of the dead time $t_{dead}$. 

$$[P]_0 = t_{dead} \cdot v_0$$ 

$t_{dead} = 58 \text{ ms}$
Figure S35. Kinetic measurement on AVB inhibition of TEM-1 using nitrocefin as the substrate. (a) Formation of product over time as a function of AVB concentration ranging from 5 µM to 35 µM. For each concentration of AVB, multiple experiments were run but shown here is only one experimental result for simplicity. The thicker, lighter curves are experimental data while the thinner, darker curves are the essentially identical fitting results according to eq. 13, giving $v_0$, $v_s$, and $k_{obs}$. (b) Selected window of (a) highlighting the initial product formation. (c) Plot of $v_0$ against AVB concentration, which was fitted to eq. 17, giving $K_i$. (d) Plot of $k_{obs}$ against $\Theta$, the partition term defined in eq. 22. The plot was fitted to eq. 23, giving $k_5$. (e) Plot of $v_s$ against AVB concentration, which was fitted to eq. 20, giving $K_i^*$, which in turn was used to calculate $k_{5}$. 
Figure S36. Kinetic measurement on AVB inhibition of A237Y TEM-1 using PenG as the substrate. (a) Formation of product over time as a function of AVB concentration ranging from 50 µM to 300 µM. The [P] was scaled to an [E]₀ of 1 µM. For each concentration of AVB, multiple experiments were run but shown here is only one experimental result for simplicity. The thicker, lighter curves are experimental data while the thinner, darker curves are the fitting results according to eq. 13, giving $v_0$, $v_s$, and $k_{obs}$. (b) Selected window of (a) highlighting the initial product formation. (c) Plot of $v_0$ against AVB concentration, which was fitted to eq. 17, giving $K_i$. (d) Plot of $k_{obs}$ against $\Theta$, the partition term defined in eq. 22. The plot was fitted to eq. 23, giving $k_5$. (e) Plot of $v_s$ against AVB concentration, which was fitted to eq. 20, giving $K_{i^*}$, which in turn was used to calculate $k_5$. 
Figure S37. Kinetic measurement on AVB inhibition of A237Y<sup>c</sup> TEM-1 using PenG as the substrate. (a) Formation of product over time as a function of AVB concentration ranging from 4 mM to 20 mM. The [P] was scaled to an [E]<sub>0</sub> of 1 μM. For each concentration of AVB, multiple experiments were run but shown here is only one experimental result for simplicity. The thicker, lighter curves are experimental data while the thinner, darker curves are the fitting results according to eq. 13, giving v<sub>0</sub>, v<sub>s</sub>, and k<sub>obs</sub>. (b) Plot of v<sub>0</sub> against AVB concentration, which was fitted to eq. 17, giving K<sub>i</sub>. (c) Plot of k<sub>obs</sub> against Θ, the partition term defined in eq. 22. The plot was fitted to eq. 23, giving k<sub>s</sub>. (d) Plot of v<sub>s</sub> against AVB concentration, which was fitted to eq. 20, giving K<sub>i</sub>*<sup>s</sup>, which in turn was used to calculate k<sub>s</sub>. 
**Figure S38.** Crystal structure of WT TEM–AVB (PDB: 8DE0) showing the bad positioning of the N atom as the nucleophile for AVB recyclization. The electron density maps (2mFo-DFc, 1.5σ) are depicted.
Figure S39. Crystal structure of ibrutinib bound to BTK1 (PDB: 5P9J)\textsuperscript{47}.
Figure S40. Crystal structure of sotorasib bound to KRAS-G12C (PDB: 6OIM)\textsuperscript{48}.
Figure S41. Crystal structure of saxagliptin bound to DPP-IV (PDB: 3BJM)\textsuperscript{49}.
Figure S42. Crystal structure of telaprevir bound to NS3/4A-protease (PDB: 3SV6)\textsuperscript{50}. 
Table S1. Kinetic parameters for PenG hydrolysis by TEM-1

<table>
<thead>
<tr>
<th></th>
<th>$K_M$ (µM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{ac}$ (s$^{-1}$)$^a$</th>
<th>$k_3$ (s$^{-1}$)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>22 ± 1</td>
<td>1774 ± 43</td>
<td>4941 [4784, 5086]</td>
<td>2769 [2697, 2843]</td>
</tr>
<tr>
<td>A237Y</td>
<td>261 ± 3</td>
<td>793 ± 4</td>
<td>2838 [2740, 2940]</td>
<td>1101 [1084, 1118]</td>
</tr>
<tr>
<td>A237Ye</td>
<td>1539 ± 1$^b$</td>
<td>4.54 ± 0.04$^b$</td>
<td>86.1 [69.2, 97.0]</td>
<td>4.80 [4.75, 4.87]</td>
</tr>
<tr>
<td>A237E</td>
<td>1967 ± 20</td>
<td>1168 ± 14</td>
<td>5857 [5288, 6012]</td>
<td>1466 [1442, 1500]</td>
</tr>
<tr>
<td>A237R</td>
<td>258 ± 15</td>
<td>2044 ± 64</td>
<td>3265 [3137, 3396]</td>
<td>5462 [5191, 5773]</td>
</tr>
<tr>
<td>A237W</td>
<td>366 ± 5</td>
<td>332 ± 3</td>
<td>1390 [1366, 1415]</td>
<td>436 [432, 441]</td>
</tr>
</tbody>
</table>

$^a$ $k_{ac}$ and $k_3$ are shown as median values and their uncertainties are shown as 68.25% confidence interval, equivalent to 1σ.

$^b$ Measured after the addition of 50 µM AVB, which inhibits any amide mutants (such as A237Y) due to the imperfect fidelity in the incorporation of the non-canonical amino acid. Without adding AVB, the apparent $k_{cat}$ was measured as 9.25 ± 0.55 s$^{-1}$, indicating a 99% fidelity assuming all mis-incorporated amino acid is Tyr. The apparent $K_M$ was measured as 194 ± 10 mM. AVB at 50 µM concentration only binds to A237Ye at a negligible level in the measurement time, according to the inhibition kinetics.
**Table S2.** Kinetic parameters for AVB inhibition of TEM-1

<table>
<thead>
<tr>
<th></th>
<th>$K_i$ (µM)</th>
<th>$K_{i*}$ (nM)</th>
<th>$k_{cbm}$ (s$^{-1}$)</th>
<th>$k_{rec}$ (10$^{-4}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>76 ± 10</td>
<td>0.741 ± 0.007</td>
<td>23.1 ± 0.2</td>
<td>2.3 ± 0.3</td>
</tr>
<tr>
<td>A237Y</td>
<td>58 ± 11</td>
<td>15.9 ± 0.2</td>
<td>6.19 ± 0.08</td>
<td>17 ± 3</td>
</tr>
<tr>
<td>A237Y$^c$</td>
<td>(5.5 ± 2.0) × 10$^3$</td>
<td>(9.9 ± 1.3) × 10$^2$</td>
<td>(1.8 ± 0.1) × 10$^{-3}$</td>
<td>(3.4 ± 0.1) × 10$^{-3}$</td>
</tr>
<tr>
<td>A237E</td>
<td>(1.48 ± 0.25) × 10$^3$</td>
<td>27.7 ± 0.8</td>
<td>4.63 ± 0.23</td>
<td>0.87 ± 0.15</td>
</tr>
<tr>
<td>A237R</td>
<td>47 ± 7</td>
<td>8.03 ± 0.04</td>
<td>3.63 ± 0.03</td>
<td>6.2 ± 0.9</td>
</tr>
<tr>
<td>A237W</td>
<td>54 ± 11</td>
<td>10.91 ± 0.05</td>
<td>3.31 ± 0.05</td>
<td>6.7 ± 1.3</td>
</tr>
</tbody>
</table>
### Table S3. Mass and extinction coefficient of TEM-1 mutants

<table>
<thead>
<tr>
<th>Mutant</th>
<th>$\varepsilon_{280}$ (M$^{-1}$ cm$^{-1}$)$^a$</th>
<th>Expected Mass (Da)$^a$</th>
<th>Observed Mass (Da)$^b$</th>
<th>Observed Mass at high resolution (Da)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>28085</td>
<td>31499</td>
<td>31508</td>
<td></td>
</tr>
<tr>
<td>S70G</td>
<td>28085</td>
<td>31469</td>
<td>31479</td>
<td></td>
</tr>
<tr>
<td>WT-AVB</td>
<td></td>
<td>31763</td>
<td>31775</td>
<td></td>
</tr>
<tr>
<td>WT-PenG</td>
<td></td>
<td>31832</td>
<td>31845</td>
<td></td>
</tr>
<tr>
<td>A237Y</td>
<td>29575</td>
<td>31591</td>
<td>31600</td>
<td>31588.3</td>
</tr>
<tr>
<td>A237Y S70G</td>
<td></td>
<td>31561</td>
<td>31564</td>
<td></td>
</tr>
<tr>
<td>A237Y-AVB</td>
<td></td>
<td>31855</td>
<td>31862$^e$</td>
<td></td>
</tr>
<tr>
<td>A237Y-PenG</td>
<td></td>
<td>31924</td>
<td>31934</td>
<td></td>
</tr>
<tr>
<td>A237E</td>
<td>28085</td>
<td>31557</td>
<td>31568</td>
<td></td>
</tr>
<tr>
<td>A237E S70G</td>
<td></td>
<td>31527</td>
<td>31533</td>
<td></td>
</tr>
<tr>
<td>A237E AVB</td>
<td></td>
<td>31821</td>
<td>31829$^e$</td>
<td></td>
</tr>
<tr>
<td>A237E-PenG</td>
<td></td>
<td>31890</td>
<td>31901</td>
<td></td>
</tr>
<tr>
<td>A237W</td>
<td>33585</td>
<td>31614</td>
<td>31622</td>
<td></td>
</tr>
<tr>
<td>A237W S70G</td>
<td></td>
<td>31584</td>
<td>31588</td>
<td></td>
</tr>
<tr>
<td>A237W-AVB</td>
<td></td>
<td>31878</td>
<td>31891$^e$</td>
<td></td>
</tr>
<tr>
<td>A237W-PenG</td>
<td></td>
<td>31947</td>
<td>31956</td>
<td></td>
</tr>
<tr>
<td>A237R</td>
<td>28085</td>
<td>31584</td>
<td>31589</td>
<td></td>
</tr>
<tr>
<td>A237R S70G</td>
<td></td>
<td>31554</td>
<td>31556</td>
<td></td>
</tr>
<tr>
<td>A237R-AVB</td>
<td></td>
<td>31848</td>
<td>31859$^e$</td>
<td></td>
</tr>
<tr>
<td>A237R-PenG</td>
<td></td>
<td>31917</td>
<td>31924</td>
<td></td>
</tr>
<tr>
<td>A237Ye</td>
<td>29575</td>
<td>31591$^{d}$</td>
<td>31602$^{d}$</td>
<td>31589.4$^{d}$</td>
</tr>
<tr>
<td>A237Ye S70G</td>
<td></td>
<td>31561$^{d}$</td>
<td>31572$^{d}$</td>
<td>(23401.2, 8205.7)</td>
</tr>
<tr>
<td>A237Ye AVB</td>
<td></td>
<td>31855$^{d}$</td>
<td>31878$^{d}$</td>
<td>(23683, 8212)$^c$</td>
</tr>
<tr>
<td>A237Ye PenG</td>
<td></td>
<td>31924$^{d}$</td>
<td>31937$^{d}$</td>
<td>(23746, 8209)</td>
</tr>
<tr>
<td>TEM-1 native</td>
<td></td>
<td>28085</td>
<td>28934</td>
<td>28941</td>
</tr>
<tr>
<td>A237Y</td>
<td>29575</td>
<td>29026</td>
<td>29034</td>
<td></td>
</tr>
</tbody>
</table>


$^b$ Proteins of ~ 30 kDa have a systematic error within +15 Da, depending on instrument status.

$^c$ Higher mass might be observed because the infrared samples were used for measuring mass spectra, which have been partially deuterium exchanged.

$^d$ Values in brackets correspond to the protein fragments generated by backbone ester hydrolysis.
Table S4. X-ray diffraction data and refinement statistics of WT TEM-1

<table>
<thead>
<tr>
<th>Protein</th>
<th>WT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB entry</td>
<td>7U6Q</td>
</tr>
</tbody>
</table>

**Data collection statistics**

<table>
<thead>
<tr>
<th>Beamline</th>
<th>BL 12-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength (Å)</td>
<td>0.97946</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>39.23 – 1.90 (1.95 – 1.90)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 2₁ 2₁ 2₁ (No.19)</td>
</tr>
<tr>
<td>Unit cell dimensions a, b, c, (Å)</td>
<td>60.73, 96.33, 154.20</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Matthews coefficient</td>
<td>2.59</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>52.60</td>
</tr>
<tr>
<td>Total observations</td>
<td>948,280 (70,789)</td>
</tr>
<tr>
<td>Unique observations</td>
<td>71,785 (5,232)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>13.210 (13.530)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.6 (99.9)</td>
</tr>
<tr>
<td>Mean I/s(I)</td>
<td>8.23 (1.50)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>22.0</td>
</tr>
<tr>
<td>Anisotropy</td>
<td>0.192</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.288 (2.512)</td>
</tr>
<tr>
<td>Rmeas</td>
<td>0.300 (2.610)</td>
</tr>
<tr>
<td>CC₁/₂</td>
<td>0.997 (0.640)</td>
</tr>
</tbody>
</table>

**Refinement statistics**

| Reflections used | 71,756 (2,565) |
| Reflections used for Rfree | 3,588 (134) |
| Rwork | 0.1911 (0.3077) |
| Rfree | 0.2385 (0.3395) |
| Number of non-H atoms | 6,960 |
| Protein | 6,116 |
| Ligand | 30 |
| Solvent | 814 |
| Protein residues | 790 |
| RMSD bond lengths (Å) | 0.014 |
| RMSD bond angles (°) | 1.78 |
| Ramachandran favored (%) | 98.21 |
| Ramachandran allowed (%) | 1.79 |
| Ramachandran outliers (%) | 0.00 |
| Rotamer outliers (%) | 2.12 |
| Clashscore | 2.52 |
| Average B factor (Å²) | 24.40 |
| Protein | 23.18 |
| Ligand | 38.83 |
| Solvent | 32.99 |
Table S5. X-ray diffraction data and refinement statistics of A237Y TEM-1

<table>
<thead>
<tr>
<th></th>
<th>A237Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB entry</td>
<td>8DDZ</td>
</tr>
</tbody>
</table>

**Data collection statistics**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamline</td>
<td>BL 12-2</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97946</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>38.37 – 1.45 (1.50 – 1.46)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 1 2 1 (No.4)</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td></td>
</tr>
<tr>
<td>a, b, c, (Å)</td>
<td>60.61, 83.77, 95.66</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90.23, 90</td>
</tr>
<tr>
<td>Matthews coefficient</td>
<td>2.09</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>41.12</td>
</tr>
<tr>
<td>Total observations</td>
<td>2,273,538 (164,682)</td>
</tr>
<tr>
<td>Unique observations</td>
<td>164,606 (11,707)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>13.812 (14.067)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.6 (95.7)</td>
</tr>
<tr>
<td>Mean I/s(I)</td>
<td>9.51 (1.56)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>15.5</td>
</tr>
<tr>
<td>Anisotropy</td>
<td>0.254</td>
</tr>
<tr>
<td>R&lt;sub&gt;merge&lt;/sub&gt;</td>
<td>0.167 (2.002)</td>
</tr>
<tr>
<td>R&lt;sub&gt;meas&lt;/sub&gt;</td>
<td>0.174 (2.076)</td>
</tr>
<tr>
<td>CC&lt;sub&gt;1/2&lt;/sub&gt;</td>
<td>0.998 (0.729)</td>
</tr>
</tbody>
</table>

**Refinement statistics**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflections used</td>
<td>164,606 (3,699)</td>
</tr>
<tr>
<td>Reflections used for R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>8,086 (206)</td>
</tr>
<tr>
<td>R&lt;sub&gt;work&lt;/sub&gt;</td>
<td>0.1769 (0.1858)</td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.2121 (0.2282)</td>
</tr>
<tr>
<td>Number of non-H atoms</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>8,900</td>
</tr>
<tr>
<td>Ligand</td>
<td>8,183</td>
</tr>
<tr>
<td>Solvent</td>
<td>0</td>
</tr>
<tr>
<td>Solvent</td>
<td>717</td>
</tr>
<tr>
<td>Protein residues</td>
<td>1,048</td>
</tr>
<tr>
<td>RMSD bond lengths (Å)</td>
<td>0.018</td>
</tr>
<tr>
<td>RMSD bond angles (°)</td>
<td>1.70</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>97.88</td>
</tr>
<tr>
<td>Ramachandran allowed (%)</td>
<td>2.12</td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
<td>0.00</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>1.25</td>
</tr>
<tr>
<td>Clashscore</td>
<td>2.26</td>
</tr>
<tr>
<td>Average B factor (Å²)</td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>18.62</td>
</tr>
<tr>
<td>Ligand</td>
<td>17.92</td>
</tr>
<tr>
<td>Solvent</td>
<td>NA</td>
</tr>
<tr>
<td>Solvent</td>
<td>26.58</td>
</tr>
</tbody>
</table>
Table S6. X-ray diffraction data and refinement statistics of WT TEM-1–AVB

<table>
<thead>
<tr>
<th>Protein</th>
<th>WT–AVB</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB entry</td>
<td>8DE0</td>
</tr>
</tbody>
</table>

### Data collection statistics

| Beamline | BL 12-2 |
| Wavelength (Å) | 0.97946 |
| Resolution range (Å) | 38.33 – 1.72 (1.77 – 1.73) |
| Space group | P 1 2_1 1 (No.4) |
| Unit cell dimensions | |
| a, b, c, (Å) | 60.65, 83.69, 95.57 |
| α, β, γ (°) | 90, 90.19, 90 |
| Matthews coefficient | 2.08 |
| Solvent content (%) | 40.77 |
| Total observations | 1,347,709 (80,926) |
| Unique observations | 99,749 (7,221) |
| Multiplicity | 13.51 (11.21) |
| Completeness (%) | 99.0 (97.7) |
| Mean I/s(I) | 9.060 (1.620) |
| Wilson B-factor (Å²) | 21.07 |
| Anisotropy | 0.044 |

### Refinement statistics

| Reflections used | 99,749 (2,557) |
| Reflections used for Rfree | 5,045 (118) |
| Rwork | 0.1727 (0.1604) |
| Rfree | 0.1981 (0.2141) |
| Number of non-H atoms | 9,284 |
| Protein | 8,149 |
| Ligand | 68 |
| Solvent | 1,067 |
| Protein residues | 1,048 |
| RMSD bond lengths (Å) | 0.013 |
| RMSD bond angles (°) | 1.65 |
| Ramachandran favored (%) | 97.98 |
| Ramachandran allowed (%) | 2.02 |
| Ramachandran outliers (%) | 0.00 |
| Rotamer outliers (%) | 1.26 |
| Clashscore | 3.52 |
| Average B factor (Å²) | 21.71 |
| Protein | 20.37 |
| Ligand | 21.23 |
| Solvent | 31.94 |
Table S7. X-ray diffraction data and refinement statistics of A237Y TEM-1–AVB

<table>
<thead>
<tr>
<th></th>
<th>A237Y–AVB</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PDB entry</strong></td>
<td>8DE1</td>
</tr>
</tbody>
</table>

**Data collection statistics**

<table>
<thead>
<tr>
<th></th>
<th>BL 12-2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Beamline</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Wavelength (Å)</strong></td>
<td>0.97946</td>
</tr>
<tr>
<td><strong>Resolution range (Å)</strong></td>
<td>38.30 – 1.56 (1.60 – 1.56)</td>
</tr>
<tr>
<td><strong>Space group</strong></td>
<td>P 1 2 1 (No.4)</td>
</tr>
<tr>
<td><strong>Unit cell dimensions</strong></td>
<td>60.61, 83.57, 95.86</td>
</tr>
<tr>
<td><strong>α, β, γ (°)</strong></td>
<td>90, 90.44, 90</td>
</tr>
<tr>
<td><strong>Matthews coefficient</strong></td>
<td>2.07</td>
</tr>
<tr>
<td><strong>Solvent content (%)</strong></td>
<td>40.64</td>
</tr>
<tr>
<td><strong>Total observations</strong></td>
<td>1,825,431 (123,641)</td>
</tr>
<tr>
<td><strong>Unique observations</strong></td>
<td>135,097 (9,774)</td>
</tr>
<tr>
<td><strong>Multiplicity</strong></td>
<td>13.51 (12.65)</td>
</tr>
<tr>
<td><strong>Completeness (%)</strong></td>
<td>99.5 (97.4)</td>
</tr>
<tr>
<td><strong>Mean I/s(I)</strong></td>
<td>7.97 (1.59)</td>
</tr>
<tr>
<td><strong>Wilson B-factor (Å²)</strong></td>
<td>16.4</td>
</tr>
<tr>
<td><strong>Anisotropy</strong></td>
<td>0.546</td>
</tr>
<tr>
<td><strong>R_{merge}</strong></td>
<td>0.201 (1.829)</td>
</tr>
<tr>
<td><strong>R_{meas}</strong></td>
<td>0.209 (1.905)</td>
</tr>
<tr>
<td><strong>CC_{1/2}</strong></td>
<td>0.998 (0.733)</td>
</tr>
</tbody>
</table>

**Refinement statistics**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reflections used</strong></td>
<td>135,091 (6,217)</td>
</tr>
<tr>
<td><strong>Reflections used for R_{free}</strong></td>
<td>6,817 (319)</td>
</tr>
<tr>
<td><strong>R_{work}</strong></td>
<td>0.2065 (0.3257)</td>
</tr>
<tr>
<td><strong>R_{free}</strong></td>
<td>0.2214 (0.3410)</td>
</tr>
<tr>
<td><strong>Number of non-H atoms</strong></td>
<td>9,507</td>
</tr>
<tr>
<td><strong>Protein</strong></td>
<td>8,169</td>
</tr>
<tr>
<td><strong>Ligand</strong></td>
<td>68</td>
</tr>
<tr>
<td><strong>Solvent</strong></td>
<td>1,270</td>
</tr>
<tr>
<td><strong>Protein residues</strong></td>
<td>1,052</td>
</tr>
<tr>
<td><strong>RMSD bond lengths (Å)</strong></td>
<td>0.013</td>
</tr>
<tr>
<td><strong>RMSD bond angles (°)</strong></td>
<td>1.20</td>
</tr>
<tr>
<td><strong>Ramanchandran favored (%)</strong></td>
<td>97.99</td>
</tr>
<tr>
<td><strong>Ramanchandran allowed (%)</strong></td>
<td>2.01</td>
</tr>
<tr>
<td><strong>Ramanchandran outliers (%)</strong></td>
<td>0.00</td>
</tr>
<tr>
<td><strong>Rotamer outliers (%)</strong></td>
<td>1.48</td>
</tr>
<tr>
<td><strong>Clashscore</strong></td>
<td>7.28</td>
</tr>
<tr>
<td><strong>Average B factor (Å²)</strong></td>
<td>18.87</td>
</tr>
<tr>
<td><strong>Protein</strong></td>
<td>17.60</td>
</tr>
<tr>
<td><strong>Ligand</strong></td>
<td>20.51</td>
</tr>
<tr>
<td><strong>Solvent</strong></td>
<td>26.97</td>
</tr>
</tbody>
</table>
Table S8. X-ray diffraction data and refinement statistics of A237Y TEM-1–AVB at RT

<table>
<thead>
<tr>
<th>Protein</th>
<th>A237Y–AVB at RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PDB entry</td>
<td>8DE2</td>
</tr>
</tbody>
</table>

**Data collection statistics**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Beamline</td>
<td>BL 12-1</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.97946</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>39.33 – 2.45 (2.51 – 2.45)</td>
</tr>
<tr>
<td>Space group</td>
<td>P 1 2 1 (No.4)</td>
</tr>
<tr>
<td>Unit cell dimensions</td>
<td></td>
</tr>
<tr>
<td>a, b, c, (Å)</td>
<td>60.93, 86.07, 96.94</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90.00, 90</td>
</tr>
<tr>
<td>Matthews coefficient</td>
<td>2.17</td>
</tr>
<tr>
<td>Solvent content (%)</td>
<td>43.31</td>
</tr>
<tr>
<td>Total observations</td>
<td>211,429 (11,647)</td>
</tr>
<tr>
<td>Unique observations</td>
<td>35,092 (2,004)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>6.025 (5.812)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>94.9 (73.5)</td>
</tr>
<tr>
<td>Mean I/s(I)</td>
<td>8.54 (1.79)</td>
</tr>
<tr>
<td>Wilson B-factor (Å²)</td>
<td>36.7</td>
</tr>
<tr>
<td>Anisotropy</td>
<td>0.853</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.152 (0.978)</td>
</tr>
<tr>
<td>Rmeas</td>
<td>0.167 (1.073)</td>
</tr>
<tr>
<td>CC1/2</td>
<td>0.998 (0.851)</td>
</tr>
</tbody>
</table>

**Refinement statistics**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Reflections used</td>
<td>35,092 (1,971)</td>
</tr>
<tr>
<td>Reflections used for Rfree</td>
<td>1,789 (102)</td>
</tr>
<tr>
<td>Rwork</td>
<td>0.2045 (0.3512)</td>
</tr>
<tr>
<td>Rfree</td>
<td>0.2495 (0.4194)</td>
</tr>
<tr>
<td>Number of non-H atoms</td>
<td>8,302</td>
</tr>
<tr>
<td>Protein</td>
<td>8,132</td>
</tr>
<tr>
<td>Ligand</td>
<td>68</td>
</tr>
<tr>
<td>Solvent</td>
<td>102</td>
</tr>
<tr>
<td>Protein residues</td>
<td>1,052</td>
</tr>
<tr>
<td>RMSD bond lengths (Å)</td>
<td>0.004</td>
</tr>
<tr>
<td>RMSD bond angles (°)</td>
<td>0.68</td>
</tr>
<tr>
<td>Ramachandran favored (%)</td>
<td>97.99</td>
</tr>
<tr>
<td>Ramachandran allowed (%)</td>
<td>2.01</td>
</tr>
<tr>
<td>Ramachandran outliers (%)</td>
<td>0.00</td>
</tr>
<tr>
<td>Rotamer outliers (%)</td>
<td>3.44</td>
</tr>
<tr>
<td>Clashscore</td>
<td>9.34</td>
</tr>
<tr>
<td>Average B factor (Å²)</td>
<td>43.53</td>
</tr>
<tr>
<td>Protein</td>
<td>43.61</td>
</tr>
<tr>
<td>Ligand</td>
<td>40.29</td>
</tr>
<tr>
<td>Solvent</td>
<td>39.35</td>
</tr>
</tbody>
</table>
Table S9. Growth of DH10B ΔaspC ΔtyrB in the minimal media

<table>
<thead>
<tr>
<th>Incubation Time (h)</th>
<th>OD$_{600}$ (no Tyr)</th>
<th>OD$_{600}$ (1 mM Tyr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>0.009</td>
<td>0.218</td>
</tr>
<tr>
<td>42</td>
<td>0.012</td>
<td>1.855</td>
</tr>
</tbody>
</table>
Table S10. FTIR data analysis of BODAC

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Urea Carbonyl</th>
<th></th>
<th>Amide Carbonyl</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Peak Picking</td>
<td>Curve Fitting</td>
<td>Peak Picking</td>
<td>Curve Fitting</td>
</tr>
<tr>
<td></td>
<td>Position (cm⁻¹)</td>
<td>FWHM (cm⁻¹)</td>
<td>Position (cm⁻¹)</td>
<td>FWHM (cm⁻¹)</td>
</tr>
<tr>
<td>dibutylether</td>
<td>1771.8</td>
<td>14.2</td>
<td>1771.3</td>
<td>14.7</td>
</tr>
<tr>
<td>toluene</td>
<td>1767.5</td>
<td>13.1</td>
<td>1767.2</td>
<td>13.8</td>
</tr>
<tr>
<td>THF</td>
<td>1764.7</td>
<td>15.2</td>
<td>1764.4</td>
<td>15.6</td>
</tr>
<tr>
<td>DCM</td>
<td>1757.1</td>
<td>21.8</td>
<td>1756.0</td>
<td>23.2</td>
</tr>
<tr>
<td>DMSO</td>
<td>1751.6</td>
<td>20.0</td>
<td>1751.1</td>
<td>21.5</td>
</tr>
<tr>
<td>water</td>
<td>1722.3</td>
<td>31.7</td>
<td>1722.4</td>
<td>39.6</td>
</tr>
</tbody>
</table>
Table S11. Solvent electric fields projected on the C=O (Urea) in BODAC based on MD simulation

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Average (MV/cm)</th>
<th>Standard Deviation (MV/cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>dibutylether</td>
<td>-10.6</td>
<td>5.2</td>
</tr>
<tr>
<td>toluene</td>
<td>-15.3</td>
<td>6.0</td>
</tr>
<tr>
<td>THF</td>
<td>-20.3</td>
<td>7.1</td>
</tr>
<tr>
<td>DCM</td>
<td>-32</td>
<td>12</td>
</tr>
<tr>
<td>DMSO</td>
<td>-35.0</td>
<td>9.8</td>
</tr>
<tr>
<td>water</td>
<td>-75</td>
<td>22</td>
</tr>
</tbody>
</table>
Table S12. FTIR data analysis of MPC

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Peak Picking</th>
<th></th>
<th>Curve Fitting</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Position</td>
<td>FWHM</td>
<td>Position</td>
<td>FWHM</td>
</tr>
<tr>
<td></td>
<td>(cm(^{-1}))</td>
<td>(cm(^{-1}))</td>
<td>(cm(^{-1}))</td>
<td>(cm(^{-1}))</td>
</tr>
<tr>
<td>hexanes</td>
<td>1715.4</td>
<td>7.5</td>
<td>1715.3</td>
<td>7.6</td>
</tr>
<tr>
<td>dibutylether</td>
<td>1711.6</td>
<td>10.1</td>
<td>1711.4</td>
<td>10.3</td>
</tr>
<tr>
<td>THF</td>
<td>1705.8</td>
<td>10.7</td>
<td>1705.7</td>
<td>10.8</td>
</tr>
<tr>
<td>toluene</td>
<td>1705.3</td>
<td>11.6</td>
<td>1705.2</td>
<td>11.7</td>
</tr>
<tr>
<td>DMSO</td>
<td>1694.1</td>
<td>16.4</td>
<td>1694.0</td>
<td>16.4</td>
</tr>
<tr>
<td>DCM</td>
<td>1691.7</td>
<td>18.8</td>
<td>1692.0</td>
<td>19.6</td>
</tr>
<tr>
<td>water</td>
<td>1658.8</td>
<td>34.0</td>
<td>1657.0</td>
<td>34.7</td>
</tr>
</tbody>
</table>
Table S13. Solvent electric fields projected on the C=O in MPC based on MD simulation

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Average (MV/cm)</th>
<th>Standard Deviation (MV/cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hexanes</td>
<td>-0.04</td>
<td>0.75</td>
</tr>
<tr>
<td>dibutylether</td>
<td>-6.4</td>
<td>5.3</td>
</tr>
<tr>
<td>THF</td>
<td>-11.8</td>
<td>8.6</td>
</tr>
<tr>
<td>toluene</td>
<td>-10.4</td>
<td>6.8</td>
</tr>
<tr>
<td>DMSO</td>
<td>-22.7</td>
<td>9.7</td>
</tr>
<tr>
<td>DCM</td>
<td>-26</td>
<td>15</td>
</tr>
<tr>
<td>water</td>
<td>-75</td>
<td>22</td>
</tr>
</tbody>
</table>
Table S14. The fitting results of the vibrational Stark spectrum of MPC

| Peak position $(\text{cm}^{-1})$ | FWHM $(\text{cm}^{-1})$ | $A^a$ $(\text{cm}^{-1})$ | $B^a$ $(\text{cm}^{-1})$ | $C^a$ $(\text{cm}^{-2})$ | $|\Delta \mu f|$ $[\text{cm}^{-1} / (\text{MV/cm})]$ | $|\Delta \mu|^b$ $[\text{cm}^{-1} / (\text{MV/cm})]$ | $f$ |
|-------------------------------|------------------------|-------------------------|-------------------------|-------------------------|----------------------------------|----------------------------------|---|
| 1696.9                        | 10.5                   | 2.24                    | 1.97                    | 1.85                    | $1.36$                           | $0.754$                          | 1.80 |

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

$^b$ $|\Delta \mu|$ was obtained from solvatochromism and MD simulations, see Figure S13.
Table S15. FTIR data analysis of 10 mM AVB solutions in D$_2$O buffer$^a$

<table>
<thead>
<tr>
<th>Solute</th>
<th>Urea Carbonyl</th>
<th>Amide Carbonyl</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Peak Picking</td>
<td>Curve Fitting</td>
</tr>
<tr>
<td></td>
<td>Position (cm$^{-1}$)</td>
<td>FWHM (cm$^{-1}$)</td>
</tr>
<tr>
<td>AVB-Na</td>
<td>1744.7</td>
<td>40.0</td>
</tr>
<tr>
<td>AVB-NBu$_4$</td>
<td>1745.0</td>
<td>39.9</td>
</tr>
<tr>
<td>AVB-$^{13}$C-NBu$_4$</td>
<td>1697.6</td>
<td>21.7</td>
</tr>
</tbody>
</table>

$^a$ The D$_2$O buffer is the same used for FTIR spectroscopy for proteins, containing 100 mM NaCl and 50 mM KPi (pD 7.4).
Table S16. Peak fitting of the difference FTIR spectra of TEM•AVB

<table>
<thead>
<tr>
<th></th>
<th>(^{12}\text{C})</th>
<th></th>
<th>(^{13}\text{C})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Position (cm(^{-1}))</td>
<td>FWHM (cm(^{-1}))</td>
<td>Lorentzian (%)(^a)</td>
</tr>
<tr>
<td>WT</td>
<td>1729.0</td>
<td>20.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1681.2</td>
<td>6.9</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1666.1</td>
<td>7.9</td>
<td>18.8</td>
</tr>
<tr>
<td>A237Y</td>
<td>1735.2</td>
<td>12.6</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td>1697.9</td>
<td>9.8</td>
<td>27.8</td>
</tr>
<tr>
<td></td>
<td>1685.4</td>
<td>9.8</td>
<td>0.0</td>
</tr>
<tr>
<td>A237Ye</td>
<td>1744.1</td>
<td>12.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1732.3</td>
<td>14.1</td>
<td>0.0</td>
</tr>
<tr>
<td>A237E</td>
<td>1736.6</td>
<td>7.0</td>
<td>30.0</td>
</tr>
<tr>
<td></td>
<td>1698.1</td>
<td>11.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1684.7</td>
<td>10.6</td>
<td>13.3</td>
</tr>
<tr>
<td>A237R</td>
<td>1732.7</td>
<td>13.0</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td>1698.0</td>
<td>13.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1684.7</td>
<td>12.7</td>
<td>27.8</td>
</tr>
<tr>
<td>A237W</td>
<td>1735.9</td>
<td>11.2</td>
<td>18.6</td>
</tr>
<tr>
<td></td>
<td>1697.2</td>
<td>9.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1685.9</td>
<td>12.6</td>
<td>8.2</td>
</tr>
</tbody>
</table>

\(^a\)Gaussian-Lorentzian sum
Table S17. Peak fitting of the difference FTIR spectra of TEM–AVB

<table>
<thead>
<tr>
<th></th>
<th>$^{12}$C</th>
<th>$^{13}$C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Position (cm$^{-1}$)</td>
<td>FWHM (cm$^{-1}$)</td>
</tr>
<tr>
<td>WT</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1683.9</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>1672.4</td>
<td>12.8</td>
</tr>
<tr>
<td></td>
<td>1661.6</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>1651.6</td>
<td>8.7</td>
</tr>
<tr>
<td></td>
<td>1641.0</td>
<td>11.3</td>
</tr>
<tr>
<td>A237Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1684.3</td>
<td>15.0</td>
</tr>
<tr>
<td></td>
<td>1673.6</td>
<td>14.0</td>
</tr>
<tr>
<td></td>
<td>1661.3</td>
<td>10.2</td>
</tr>
<tr>
<td></td>
<td>1650.4</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td>1637.6</td>
<td>15.0</td>
</tr>
<tr>
<td>A237Y$^e$</td>
<td>1687.4</td>
<td>14.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1684.2</td>
<td>15.0</td>
</tr>
<tr>
<td>A237E</td>
<td>1673.3</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td>1661.4</td>
<td>11.6</td>
</tr>
<tr>
<td></td>
<td>1650.5</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td>1642.4</td>
<td>13.5</td>
</tr>
<tr>
<td>A237R</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1684.2</td>
<td>14.5</td>
</tr>
<tr>
<td></td>
<td>1674.4</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>1662.0</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>1650.6</td>
<td>8.8</td>
</tr>
<tr>
<td></td>
<td>1640.8</td>
<td>13.1</td>
</tr>
<tr>
<td>A237W</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1684.2</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td>1674.8</td>
<td>11.9</td>
</tr>
<tr>
<td></td>
<td>1662.1</td>
<td>10.7</td>
</tr>
<tr>
<td></td>
<td>1650.5</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td>1640.8</td>
<td>13.3</td>
</tr>
</tbody>
</table>

$^a$ Gaussian-Lorentzian sum
**Table S18.** The largest field experienced by AVB in TEM

<table>
<thead>
<tr>
<th></th>
<th>$F$ (TEM•AVB) (MV/cm)$^a$</th>
<th>$F$ (TEM–AVB) (MV/cm)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>-150 ± 4</td>
<td>-97 ± 4</td>
</tr>
<tr>
<td>A237Y</td>
<td>-124 ± 3</td>
<td>-102 ± 4</td>
</tr>
<tr>
<td>A237Y$^c$</td>
<td>-63 ± 2</td>
<td>-35 ± 2</td>
</tr>
<tr>
<td>A237E</td>
<td>-125 ± 3</td>
<td>-95 ± 4</td>
</tr>
<tr>
<td>A237R</td>
<td>-125 ± 3</td>
<td>-98 ± 4</td>
</tr>
<tr>
<td>A237W</td>
<td>-124 ± 3</td>
<td>-98 ± 4</td>
</tr>
</tbody>
</table>

$^a$ The error of the field was calculated based on the error of probe calibration (see the captions of Figures S12 and S13).
Table S19. Peak fitting of the difference FTIR spectra of TEM•PenG

<table>
<thead>
<tr>
<th></th>
<th>$^{12}$C</th>
<th></th>
<th></th>
<th>$^{13}$C</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Position (cm$^{-1}$)</td>
<td>FWHM (cm$^{-1}$)</td>
<td>Lorentzian (%)$^a$</td>
<td>Position (cm$^{-1}$)</td>
<td>FWHM (cm$^{-1}$)</td>
<td>Lorentzian (%)$^a$</td>
</tr>
<tr>
<td>WT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1746.4</td>
<td>13.4</td>
<td>4.0</td>
<td>1701.9</td>
<td>9.7</td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>1732.0</td>
<td>12.7</td>
<td>0.0</td>
<td>1689.3</td>
<td>13.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1668.8</td>
<td>12.8</td>
<td>100.0</td>
<td>1631.5</td>
<td>6.4</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>1663.3</td>
<td>6.3</td>
<td>0.0</td>
<td>1624.3</td>
<td>6.2</td>
<td>0.0</td>
</tr>
<tr>
<td>A237Y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1772.3</td>
<td>16.1</td>
<td>0.0</td>
<td>1725.3</td>
<td>13.5</td>
<td>40.0</td>
</tr>
<tr>
<td></td>
<td>1693.2</td>
<td>4.9</td>
<td>0.0</td>
<td>1651.4</td>
<td>5.4</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1683.9</td>
<td>12.7</td>
<td>0.0</td>
<td>1639.6</td>
<td>6.3</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1669.5</td>
<td>6.3</td>
<td>43.3</td>
<td>1630.4</td>
<td>8.0</td>
<td>3.3</td>
</tr>
<tr>
<td>A237Y$^e$</td>
<td>1770.3</td>
<td>20.0</td>
<td>39.9</td>
<td>1728.1</td>
<td>24.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1689.3</td>
<td>9.1</td>
<td>0.0</td>
<td>1648.8</td>
<td>6.7</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>1768.7</td>
<td>20.0</td>
<td>12.1</td>
<td>1727.5</td>
<td>21.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1692.1</td>
<td>7.0</td>
<td>0.0</td>
<td>1660.0</td>
<td>11.4</td>
<td>0.0</td>
</tr>
<tr>
<td>A237E</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1685.5</td>
<td>5.8</td>
<td>0.0</td>
<td>1650.4</td>
<td>5.7</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1672.3</td>
<td>6.0</td>
<td>0.0</td>
<td>1627.9</td>
<td>5.4</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1664.9</td>
<td>7.9</td>
<td>13.4</td>
<td>1623.4</td>
<td>6.9</td>
<td>0.0</td>
</tr>
<tr>
<td>A237R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1770.5</td>
<td>20.0</td>
<td>30.7</td>
<td>1726.9</td>
<td>26.5</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1672.1</td>
<td>7.8</td>
<td>0.0</td>
<td>1630.3</td>
<td>5.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1664.6</td>
<td>5.9</td>
<td>0.7</td>
<td>1624.8</td>
<td>9.2</td>
<td>0.0</td>
</tr>
<tr>
<td>A237W</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1693.1</td>
<td>5.2</td>
<td>0.0</td>
<td>1650.4</td>
<td>4.6</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1686.7</td>
<td>8.9</td>
<td>0.0</td>
<td>1654.7</td>
<td>9.1</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>1665.0</td>
<td>6.2</td>
<td>0.0</td>
<td>1624.8</td>
<td>8.5</td>
<td>0.0</td>
</tr>
</tbody>
</table>

$^a$ Gaussian-Lorentzian sum
**Table S20.** The largest field experienced by PenG in TEM•PenG

<table>
<thead>
<tr>
<th></th>
<th>$F$ (MV/cm)$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>-171 ± 4</td>
</tr>
<tr>
<td>A237Y</td>
<td>-163 ± 4</td>
</tr>
<tr>
<td>A237Ye</td>
<td>-138 ± 4</td>
</tr>
<tr>
<td>A237E</td>
<td>-169 ± 4</td>
</tr>
<tr>
<td>A237R</td>
<td>-169 ± 4</td>
</tr>
<tr>
<td>A237W</td>
<td>-169 ± 4</td>
</tr>
</tbody>
</table>

$a$ The error of the field was calculated based on the error of probe calibration. The calibration was performed in our previous work using penam as a model compound$^{14}$, providing $|\Delta \mu| = 0.78 \pm 0.02 \text{ cm}^{-1/(\text{MV/cm})}$ and the frequency in vacuum as $1796.8 \pm 0.6 \text{ cm}^{-1}$.
Table S21. Free energy barrier for PenG hydrolysis by TEM-1

<table>
<thead>
<tr>
<th></th>
<th>$\Delta G_{\text{cat}}^{\ddagger}$ (kcal/mol)$^{a,b}$</th>
<th>$\Delta G_{\text{ac}}^{\ddagger}$ (kcal/mol)$^{a,b}$</th>
<th>$\Delta G_{3}^{\ddagger}$ (kcal/mol)$^{a,b}$</th>
</tr>
</thead>
</table>

$^a\Delta G^{\ddagger} = -RT \ln[k/(k_B T / h)]; R =$ gas constant; $T = 296$ K; $k$ is the measured rate constant; $k_B =$ Bolzmann’s constant; $\hbar =$ Planck’s constant.

$^b$ The uncertainties of $\Delta G_{\text{cat}}^{\ddagger}$, $\Delta G_{\text{ac}}^{\ddagger}$, and $\Delta G_{3}^{\ddagger}$ were calculated based on the standard error and uncertainty of $k_{\text{cat}}$, $k_{\text{ac}}$, and $k_3$ in Table S1, respectively.
**Table S22.** Free energy barrier for AVB inhibition of TEM-1

<table>
<thead>
<tr>
<th></th>
<th>$\Delta G_{cbm}^\ddagger$ (kcal/mol)$^{a,b}$</th>
<th>$\Delta G_{rec}^\ddagger$ (kcal/mol)$^{a,b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>WT</td>
<td>15.48 [15.47, 15.48]</td>
<td>22.25 [22.18, 22.33]</td>
</tr>
</tbody>
</table>

$^a$ $\Delta G^\ddagger = -RT \ln[k/(k_BT/\hbar)]$; $R$ = gas constant; $T = 296$ K; $k$ is the measured rate constant; $k_B$ = Bolzmann’s constant; $\hbar$ = Planck’s constant.

$^b$ The uncertainties of $\Delta G_{cbm}^\ddagger$ and $\Delta G_{rec}^\ddagger$ were calculated based on the standard error of $k_{cbm}$ and $k_{rec}$ in Table S2, respectively.
Script 1. Data processing of difference IR spectra

%-----------------------------------------------------------Description---------------------------% This script is used to fit 12C-13C difference FTIR spectra into multiple 12C and 13C peaks to account for multiple probe conformations. % Users can specify the wavenumber window for fitting and the number of 12C-13C peak pairs to fit with. % Author Zhe Ji

%-----------------------------------------------------------Input-----------------------------% A 12C-13C spectra that has been baselined: The first column for wavenumber, the second column for absorbance. % The wavenumber window that the spectrum needs to be cut into. % The number of 12C-13C peak pairs to fit the data into. % User can suggest the peak positions to start the fitting from.

%-----------------------------------------------------------Output-----------------------------% result.txt containing the fitting parameters: each row for a 12C-13C pair. % 1st column: 12C frequency (cm⁻¹) % 2nd column: 12C FWHM (cm⁻¹) % 3rd column: 12C intensity (mOD) % 4th column: 12C Lorentzian percentage % 5th column: 13C frequency (cm⁻¹) % 6th column: 13C FWHM (cm⁻¹) % 7th column: 13C intensity (mOD) % 8th column: 13C Lorentzian percentage % 9th column: frequency shift from 12C to 13C % 10th column: FWHM ratio from 12C to 13C % 11th column: Intensity ratio from 12C to 13C % 12th column: change in Lorentzian percentage from 12C to 13C % Fit.txt containing the absorbance-wavenumber curve for each peak component of fitting. % 1st column is wavenumber, followed by n columns of 12C and n columns of 13C absorbance, n=the number of 12C-13C pairs. % These are further followed by the sum of the fitting peaks, the experimental difference spectrum, and the fitting residue.

%-----------------------------------------------------------Main-------------------------------% clear
cle
scaf=10000;%A scaling factor that adjust the spectra intensity for better data fitting.

%Load a 12C-13C difference spectrum
[fname,pname]=uigetfile('*.*', 'Load IR Spectra','MultiSelect','on');
cd(pname);
Data=readtable(fname,'Delimiter','','','Format','%f %f');
w=Data(:,1);%Wavenumber, cm⁻¹
a=Data(:,2)*1000;%Absorbance, mOD
plot(w,a)
% Input wavenumber window. The data outside of the window will be cut and not used for fitting
% also input the number of conformations. For instance, "5" means to fit 5 12C positive peaks and 5 13C negative peaks
dlgtitle='Input';
dim=1;
prompt1={'Wavenumber upper limit','Wavenumber lower limit','Number of C=O conformations'};
def1={'2000','1500','3',};
answer1=inputdlg(prompt1,dlgtitle,dim,def1);
drawnow

% suggest the initial value of the peak positions. If there are 5 conformations, the user can suggest 5 12C-peak positions and 5 13C-peak positions.
pair=str2num(char(answer1(3)));% the number of positive-negative pairs (conformations)
dlgtitle='Input Wavenumber';
dim=1;
prompt2=cell(1,pair*2);
prompt2(1,1:pair)={'12C Wavenumber'};
prompt2(1,pair+1:pair*2)={'13C Wavenumber'};
def2=cell(1,pair*2)
def2(1,1:pair*2)={'0'};% the default position is "0", if no suggestion is given.
answer2=inputdlg(prompt2,dlgtitle,dim,def2);
drawnow

%--------------------------Data cutting and peak position guess-----------------------
if str2num(char(answer1(1)))>=w(1)
    wul=w(1);
    ul_index=1;% the index of the upper limit of the wavenumber
else
    wul=str2num(char(answer1(1)));% the number of positive-negative pairs (conformations)
    for j=1:length(w)
        if w(j)>wul
            ul_index=j;
        end
    end
    ul_index=ul_index+1;
end
if str2num(char(answer1(2)))<=w(end)
    wll=w(end);
    ll_index=length(w);% the index of the lower limit of the wavenumber
else
    wll=str2num(char(answer1(2)));% the number of positive-negative pairs (conformations)
    for j=length(w):-1:1
        if w(j)<wll
            ll_index=j;
        end
    end
    ll_index=ll_index-1;
end
guess=zeros(pair*2,1);  
guessinput=zeros(pair*2,1);  
for i=1:pair*2  
guessinput(i)=str2num(char(answer2(i)));%user-suggested peak positions  
end  
if ismember(0,guessinput(1:pair))% if there is no suggestion for 12C positions  
gap=(wul-wll)/2/(pair+1);  
for i=1:pair  
guess(i)= wul-gap*i;% the initial 12C positions are set to evenly span the upper half of the wavenumber window  
end  
guess(pair+1:pair*2)=guess(1:pair)-47;  
else % if there is suggestion for 12C positions  
guess(1:pair)=guessinput(1:pair);  
if ismember(0,guessinput(pair+1:pair*2)) %if there is no suggestion for the 13C positions  
guess(pair+1:pair*2)=guess(1:pair)-47;  
else  
guess(pair+1:pair*2)=guessinput(pair+1:pair*2);%if there is suggestion for the 13C positions  
end  
end  
w=w(ul_index:ll_index);%cut the data according to the user-defined wavenumber window  
a=a(ul_index:ll_index,:);  

%--------------------------------------------------Peak Fitting----------------------------  
scaling=1/max(a)*scaf;  
a=a*scaling;%scale absorbance such that the built-in fitting tolerance is suitable  
a=[w,a];  

fun=@ModFun;  
x0=zeros(pair*8,1);%the initial fitting parameters. There are 8 parameters for each 12C-13C pair.  
for i=0:(pair-1)  
x0(i*8+1)=guess(i+1);%peak position by guess  
x0(i*8+2)=10;%FWHM  
x0(i*8+3)=0.5*scaling;%Intensity  
x0(i*8+4)=0.2;%percentage of Lorentzian  
x0(i*8+5)=guess(i+1+pair)-guess(i+1);%frequency shift from 12C to 13C  
x0(i*8+6)=1;%widening ratio from 12C to 13C  
x0(i*8+7)=1;%Amplification ratio from 12C to 13C  
x0(i*8+8)=0;%change in Lorentzian percentage  
end  

lb=zeros(pair*8,1);%the lower bound for the parameters  
for i=0:(pair-1)  
    lb(i*8+1)=min(w);  
    lb(i*8+2)=2;%the narrowest FWHM  
    lb(i*8+3)=0.1*scaling;%the lowest intensity  
    lb(i*8+4)=0;%Lorentzian percentage from 0 to 1  
end
lb(i*8+5)=-62;% the largest frequency shift allowed from 12C to 13C
lb(i*8+6)=0.8;% the narrowing limit of FWHM from 12C to 13C
lb(i*8+7)=-1.5;% the upper limit of intensity enhancement from 12C to 13C. Take negative sign for 13C peaks
lb(i*8+8)=-0.3;% the largest decrease of Lorentzian percentage allowed from 12C to 13C

end

ub=zeros(pair*8,1);% the upper bound for the fitting parameters
for i=0:(pair-1)
  ub(i*8+1)=max(w);%the widest FWHM
  ub(i*8+2)=20;%the highest intensity
  ub(i*8+3)=20*scaling;%the highest intensity
  ub(i*8+4)=1;% Lorentzian percentage from 0 to 1
  ub(i*8+5)=-32;% the smallest frequency shift allowed from 12C to 13C
  ub(i*8+6)=1.25;% the broadening limit of FWHM from 12C to 13C
  ub(i*8+7)=-0.67;% the lower limit of intensity enhancement from 12C to 13C. Take negative sign for 13C peaks
  ub(i*8+8)=0.3;% the largest increase of Lorentzian percentage allowed from 12C to 13C
end

TargetFun=zeros(length(a(:,1)),1);
options = optimset('MaxFunEvals',100000);
options = optimset(options,'MaxIter',100000,'TolX',1E-18);
[x,resnorm,residual] = lsqcurvefit(fun,x0,a,TargetFun,lb,ub,options);% the fitting
a(:,1)=[];
a=a/scaling;% recover the original Absorbance, unit: mOD
residual=-residual/scaling;
result=zeros(pair,12);% the matrix that stores the fitting parameters
for i=0:(pair-1)% each row for one conformation
  result(i+1,1)=x(i*8+1);% Peak position for 12C
  result(i+1,2)=x(i*8+2);% FWHM for 12C
  result(i+1,3)=x(i*8+3)/scaling;% Intensity for 12C
  result(i+1,4)=x(i*8+4)*100;% Lorentzian percentage for 12C
  result(i+1,5)=x(i*8+1)+x(i*8+5);% Peak position for 13C
  result(i+1,6)=x(i*8+2)*x(i*8+6);% FWHM for 13C
  result(i+1,7)=x(i*8+3)*x(i*8+7)/scaling;% Intensity for 13C
  result(i+1,8)=(x(i*8+4)+x(i*8+8))*100;% Lorentzian percentage for 13C
  if result(i+1,8)<0 % the Lorentzian percentage must be [0,100]
    result(i+1,8)=0;
  elseif result(i+1,8)>100
    result(i+1,8)=100;
  end
end
result(:,9)=result(:,5)-result(:,1);% frequency shift from 12C to 13C
result(:,10)=result(:,6)./result(:,2);% FWHM ratio
result(:,11)=result(:,7)./result(:,3);% Intensity ratio
result(:,12)=result(:,8)-result(:,4);% change in Lorentzian percentage
result=sortrows(result,1,'descend');

PeakFit=zeros(length(w),pair*2+1);% The matrix that stores the absorbance- wavenumber curve for each peak component of fitting.
PeakFit(:,1)=w;
for i = 1:pair
E = result(i,1);
F = result(i,2);
h = result(i,3);
m = result(i,4)/100;
E2 = result(i,5);
F2 = result(i,6);
h2 = result(i,7);
m2 = result(i,8)/100;
for j = 1:length(w)
    PeakFit(j,(i-1)*2+2)=PeakFit(j,(i-1)*2+2)+h*(1-m)*exp(-4*log(2)*(w(j)-E)^2/F^2)+h*m/(1+4*(w(j)-E)^2/F^2);
end
for j = 1:length(w)
    PeakFit(j,(i-1)*2+3)=PeakFit(j,(i-1)*2+3)+h2*(1-m2)*exp(-4*log(2)*(w(j)-E2)^2/F2^2)+h2*m2/(1+4*(w(j)-E2)^2/F2^2);
end
end

PeakFitTotal=zeros(length(w),1);
for i = 1:pair*2
    PeakFitTotal=PeakFitTotal+PeakFit(:,i+1);
end

figure(1)
plot(w,PeakFit(:,2:(pair*2+1)),w,a,'--')
xlabel('Wavenumber (cm\textsuperscript{-1})')
ylabel('Difference Absorbance (mOD)')

figure(2)
plot(w,a,'--',w,PeakFitTotal,w,residual)
xlabel('Wavenumber (cm\textsuperscript{-1})')
ylabel('Difference Absorbance (mOD)')

Fit=[PeakFit,PeakFitTotal,a,residual];

mkdir new
cd new/
writematrix(Fit);
writematrix(result);
cd ..

function Res=ModFun(x,a)
pairnum=length(x)/8;%every positive-negative pair has 8 parameters
Abs=zeros(length(a(:,1)),1);
for i = 0:(pairnum-1)
    E=x(i*8+1);%peak position
    F=x(i*8+2);%FWHM
    h=x(i*8+3);%Intensity
    m=x(i*8+4);%the percentage of Lorentzian
    shift=x(i*8+5);%the frequency shift from 12C to 13C
    widen=x(i*8+6);%the FWHM widening ratio from 12C to 13C
    Amplif=x(i*8+7);%the Intensity amplification ratio from 12C to 13C
    reshape=x(i*8+8);%the change in Lorentzian percentage from 12C to 13C
E2=E+shift;
F2=F*widen;
h2=h*Amplif;
if reshape<0     %the Lorentzian percentage of 13C should be in the range of 0-1
    m2=max(m+reshape,0);
else
    m2=min(m+reshape,1);
end
for j=1:length(a(:,1))
    Abs(j)=Abs(j)+h*(1-m)*exp(-4*log(2)*(a(j,1)-E)^2/F^2)+h*m/(1+4*(a(j,1)-E)^2/F^2);%the Gaussian-Lorentzian sum of 12C peak
    Abs(j)=Abs(j)+h2*(1-m2)*exp(-4*log(2)*(a(j,1)-E2)^2/F2^2)+h2*m2/(1+4*(a(j,1)-E2)^2/F2^2);%the Gaussian-Lorentzian sum of 13C peak
end
Res=Abs-a(:,2);
Script 2. Data processing of raw isotope-edited IR spectra

%------------------------------------Description-------------------------------
% This script is used to cut, baseline, and fit 12C and 13C FTIR spectra.
% Loaded into the script are raw 12C and 13C spectra. They can be repeats of
% multiple experiments.
% The 12C spectra will be subtracted by the corresponding 13C spectra.
% A baseline will be fitted for each difference spectrum.
% A single set of 12C-13C peak pairs will be fitted for all the difference
% spectra.
% Author Zhe Ji

%------------------------------------Input-----------------------------------
% n 12C spectra and n 13C spectra (n=number of experimental repeats): The
first column for wavenumber, the second column for absorbance.
% The wavenumber window that the spectrum needs to be cut into.
% The number of 12C-13C peak pairs to fit the data into.
% For baselines, whether add a linear function in addition to the two default
Gaussian functions.

%------------------------------------Output----------------------------------
% Spec_13Scaled.txt containing a column of wavenumber followed by n columns
of 12 absorbance and n columns of 13C absorbance. The 13C spectra has been
scaled.
% Spec_DiffBase.txt containing a column of wavenumber followed by n columns
of difference absorbance, n columns of baseline absorbance, and n columns of
baselined difference absorbance.
% result.txt containing the fitting parameters: each row for a 12C-13C pair.
% 1st column: 12C frequency (cm⁻¹)
% 2nd column: 12C FWHM (cm⁻¹)
% 3rd column: 12C intensity (mOD)
% 4th column: 12C Lorentzian percentage
% 5th column: 13C frequency (cm⁻¹)
% 6th column: 13C FWHM (cm⁻¹)
% 7th column: 13C intensity (mOD)
% 8th column: 13C Lorentzian percentage
% 9th column: frequency shift from 12C to 13C
% 10th column: FWHM ratio from 12C to 13C
% 11th column: Intensity ratio from 12C to 13C
% 12th column: change in Lorentzian percentage from 12C to 13C
% Fit.txt containing the absorbance-wavenumber curve for each peak component
of fitting.
% 1st column is wavenumber, followed by n columns of 12C and n columns of 13C
absorbance,
% a column of the absorbance sum of the fitting peaks, a column of the
% average absorbance of the baselined difference spectra,
% a column of averaged fitting residue, and n columns of fitting residue.
% BaseResult.txt containing the fitting parameters: each row for a baseline.
% 1st column: 1st Gaussian frequency (cm⁻¹)
clear
clc
scaling=1/100;%A scaling factor that adjust the spectra intensity for better
data fitting.

% Load a 12C and 13C spectra. For 9 experimental repeats, the input will be 9
12C spectra followed by 9 13C spectra.
% The spectra is in the format of wavenumber vs absorbance.
[fname,pname]=uigetfile('*.*','Load IR Spectra','MultiSelect','on');
cd(pname);
SpecNum=length(fname)/2;
Data=readtable(fname{1},'Delimiter',',','Format','%f %f');
w=Data{:,1};%Wavenumber, cm-1
a=Data{:,2}*1000;%Absorbance, mOD
if SpecNum>0.5
    for i = 2:SpecNum*2
        Data=readtable(fname{i},'Delimiter',',','Format','%f %f');
        atemp=Data{:,2}*1000;%Absorbance, mOD
        a=[a,atemp];
    end
end

% Input wavenumber window. The data outside of the window will be cut and not
used for fitting.
% Input the number of conformations. For instance, "5" means to fit 5 12C
positive peaks and 5 13C negative peaks.
% Input the baseline parameter: whether add a linear component? "0" means no;
"1" means yes.
dlgtitle='Input';
dim=1;
prompt1={'Wavenumber upper limit','Wavenumber lower limit','Number of C=O
conformations','Baseline add linear?'};
defl={'1704.5','1580.5','5','0'};
answer1=inputdlg(prompt1,dlgtitle,dim,defl);
drawnow
bslLinear=str2num(char(answer1(4)));% the number of positive-negative pairs

if str2num(char(answer1(1)))>=w(1)
    wul=w(1);%the index of the upper limit of the wavenumber
    ul_index=1;
else
    wul=str2num(char(answer1(1)));
    for j=1:length(w)
if \( w(j) > wul \)
    \( \text{ul}_{\text{index}} = j; \)
end
\( \text{ul}_{\text{index}} = \text{ul}_{\text{index}} + 1; \)
end
if \( \text{str2num} \text{char} (\text{answer1}(2)) \leq w(\text{end}) \)
    \( wll = w(\text{end}); \) % the index of the lower limit of the wavenumber
    \( \text{ll}_{\text{index}} = \text{length}(w); \)
else
    \( wll = \text{str2num} \text{char} (\text{answer1}(2)); \)
    for \( j = \text{length}(w):1:-1:1 \)
        if \( w(j) < wll \)
            \( \text{ll}_{\text{index}} = j; \)
        end
    end
    \( \text{ll}_{\text{index}} = \text{ll}_{\text{index}} - 1; \)
end
guess = \text{zeros}(\text{pair}, 1);
gap = (wul - wll) / 2 / (\text{pair} + 1);
for \( i = 1:\text{pair} \)
    guess(i) = wul - \text{gap}\cdot i; % the initial 12C positions are set to evenly span the upper half of the wavenumber window
end
\( w = w(\text{ul}_{\text{index}}:\text{ll}_{\text{index}}); \)
\( a = a(\text{ul}_{\text{index}}:\text{ll}_{\text{index}}, :); \)

%%% 12C subtraction by 13C %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
for \( i = 1:\text{SpecNum} \)
    % A scaling factor is used to adjust the intensity such that the area under the 13C curve is the same as that of the 12C curve
    \( a(:, i + \text{SpecNum}) = \text{sum}(a(:, i)) / \text{sum}(a(:, i + \text{SpecNum})) \cdot a(:, i + \text{SpecNum}); \)
end
\( \text{Spec}_{13\text{Scaled}} = a; \)
for \( i = 1:\text{SpecNum} \)
    \( a(:, i + \text{SpecNum}) = a(:, i) - a(:, i + \text{SpecNum}); \)
end
\( a(:, 1:\text{SpecNum}) = []; \)
\( \text{Spec}_{\text{Subtracted}} = a; \)

%%% Peak Fitting %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
a = a * \text{scaling}; % scale absorbance such that the built-in fitting tolerance is suitable
\( a = [w, a]; \)
\( \text{fun} = @\text{ModFun}; \)
\( \text{x0} = \text{zeros}(\text{pair}\cdot 8 + \text{SpecNum}\cdot 8, 1); \) % the initial fitting parameters. There are 8 parameters for each 12C-13C pair plus 8 parameters for each baseline.
for \( i = 0:(\text{pair} - 1) \)
    \( \text{x0}(i\cdot 8 + 1) = \text{guess}(i + 1); \) % peak position
    \( \text{x0}(i\cdot 8 + 2) = 10; \) % FWHM
    \( \text{x0}(i\cdot 8 + 3) = 5\cdot \text{scaling}; \) % Intensity
    \( \text{x0}(i\cdot 8 + 4) = 0.5; \) % percentage of Lorentzian
    \( \text{x0}(i\cdot 8 + 5) = -47; \) % frequency shift from 12C to 13C
x0(i*8+6)=1; % widening ratio from 12C to 13C
x0(i*8+7)=1; % Amplification ratio from 12C to 13C
x0(i*8+8)=0; % change in Lorentzian percentage
end

% also the fitting parameters for the baselines. There is one baseline for each difference spectrum.
% The baseline contains two gaussian functions and a linear function (optional)
for i=0:SpecNum-1
    x0(8*pair+8*i+1)=1620; % peak position of the 1st gaussian
    x0(8*pair+8*i+2)=40; % FWHM of the 1st gaussian
    x0(8*pair+8*i+3)=15*scaling; % The peak intensity of the 1st gaussian
    x0(8*pair+8*i+4)=1660; % peak position of the 2nd gaussian
    x0(8*pair+8*i+5)=40; % FWHM of the 2nd gaussian
    x0(8*pair+8*i+6)=-5*scaling; % The peak intensity of the 2nd gaussian
end

lb=zeros(pair*8+8,1); % the lower bound for the fitting parameters
for i=0:(pair-1)
    lb(i*8+1)=min(w)+20;
    lb(i*8+2)=2; % the narrowest FWHM
    lb(i*8+3)=1*scaling; % the lowest intensity
    lb(i*8+4)=0; % Lorentzian percentage from 0 to 1
    lb(i*8+5)=-52; % the largest frequency shift allowed from 12C to 13C
    lb(i*8+6)=0.80; % the narrowing limit of FWHM from 12C to 13C
    lb(i*8+7)=-1.25; % the upper limit of intensity enhancement from 12C to 13C. Take negative sign for 13C peaks
    lb(i*8+8)=-0.2; % the largest decrease of Lorentzian percentage allowed from 12C to 13C
end

for i=0:SpecNum-1
    lb(8*pair+8*i+1)=min(w)+20; % 1st Gaussian
    lb(8*pair+8*i+2)=20;
    lb(8*pair+8*i+3)=-30*scaling;
    lb(8*pair+8*i+4)=min(w)+20; % 2nd Gaussian
    lb(8*pair+8*i+5)=15;
    lb(8*pair+8*i+6)=-30*scaling;
    lb(8*pair+8*i+7)=-0.1*scaling*bslLinear; % optional linear slope
    lb(8*pair+8*i+8)=-10*scaling*bslLinear; % optional linear intercept
end

ub=zeros(pair*8+8,1); % the upper bound for the parameters
for i=0:(pair-1)
    ub(i*8+1)=max(w)-20;
    ub(i*8+2)=15; % the widest FWHM
    ub(i*8+3)=30*scaling; % the highest intensity
    ub(i*8+4)=1; % Lorentzian percentage from 0 to 1
    ub(i*8+5)=-42; % the smallest frequency shift from 12C to 13C
    ub(i*8+6)=1.25; % the broadening limit of FWHM from 12C to 13C
    ub(i*8+7)=-0.80; % the lower limit of intensity enhancement from 12C to 13C. Take negative sign for 13C peaks
    ub(i*8+8)=0.2; % the largest increase of Lorentzian percentage allowed from 12C to 13C
end
for i=0:SpecNum-1
    ub(8*pair+8*i+1)=max(w)-20;%1st Gaussian
    ub(8*pair+8*i+2)=200;
    ub(8*pair+8*i+3)=30*scaling;
    ub(8*pair+8*i+4)=max(w)-20;%2nd Gaussian
    ub(8*pair+8*i+5)=200;
    ub(8*pair+8*i+6)=30*scaling;
    ub(8*pair+8*i+7)=0.1*scaling*bslLinear;%optional linear slope
    ub(8*pair+8*i+8)=10*scaling*bslLinear;%optional linear intercept
end

TargetFun=zeros(length(a(:,1)),SpecNum);
options = optimset('MaxFunEvals',100000);
options = optimset(options,'MaxIter',100000,'TolX',1E-18);
[x,resnorm,residual] = lsqcurvefit(fun,x0,a,TargetFun,lb,ub,options);%the fitting
a(:,1)=[];
a=a/scaling;%recover the original Absorbance, unit: mOD
residual=residual/scaling;
result=zeros(pair,12);%the matrix that stores the fitting parameters for the 12C and 13C peaks
for i=0:(pair-1)%each row for one pair of 12C-13C peaks
    result(i+1,1)=x(i*8+1);%Peak position for 12C
    result(i+1,2)=x(i*8+2);%FWHM for 12C
    result(i+1,3)=x(i*8+3)/scaling;%Intensity for 12C
    result(i+1,4)=(x(i*8+4)+x(i*8+5))/100;%Lorentzian percentage for 12C
    result(i+1,5)=x(i*8+1)+x(i*8+5);%Peak position for 13C
    result(i+1,6)=(x(i*8+2)+x(i*8+6))/100;%FWHM for 13C
    result(i+1,7)=x(i*8+3)*x(i*8+7)/scaling;%Intensity for 13C
    result(i+1,8)=(x(i*8+4)+x(i*8+8))*100;%Lorentzian percentage for 13C
    if result(i+1,8)<0
        result(i+1,8)=0;
    elseif result(i+1,8)>100
        result(i+1,8)=100;
    end
end
result(:,9)=result(:,5)-result(:,1);%frequency shift from 12C to 13C
result(:,10)=result(:,6)/result(:,2);%FWHM ratio
result(:,11)=result(:,7)/result(:,3);%Intensity ratio
result(:,12)=result(:,8)-result(:,4);%change in Lorentzian percentage
result=sortrows(result,1,'descend');

PeakFit=zeros(length(w),pair*2+1);%The matrix that stores the absorbance-wavenumber curve for each peak component of fitting.
PeakFit(:,1)=w;
for i = 1:pair
    E=result(i,1);
    F=result(i,2);
    h=result(i,3);
    m=result(i,4)/100;
    E2=result(i,5);
    F2=result(i,6);
    h2=result(i,7);
    m2=result(i,8)/100;
    for j=1:length(w)
PeakFit(j,(i-1)*2+2)=PeakFit(j,(i-1)*2+2)+h*(1-m)*exp(-4*log(2)*(w(j)-E)^2/F^2)+h*m/(1+4*(w(j)-E)^2/F^2);
end
for j=1:length(w)
    PeakFit(j,(i-1)*2+3)=PeakFit(j,(i-1)*2+3)+h2*(1-m2)*exp(-4*log(2)*(w(j)-E2)^2/F2^2)+h2*m2/(1+4*(w(j)-E2)^2/F2^2);
end

PeakFitTotal=zeros(length(w),1);%The sum of the absorbance-wavenumber curves for all peak component of fitting.
for i =1:pair*2
    PeakFitTotal=PeakFitTotal+PeakFit(:,i+1);
end

BaseResult=zeros(SpecNum,8);%the matrix that stores the fitting parameters for the baselines
for i=0:SpecNum-1
    for j=1:8
        if ismember(j,[3 6 7 8])
            BaseResult(i+1,j)=x(8*pair+i*8+j)/scaling;
        else
            BaseResult(i+1,j)=x(8*pair+i*8+j);
        end
    end
end

Baseline=zeros(length(w),SpecNum);%The matrix that stores the absorbance-wavenumber curve for each baseline
for i=1:SpecNum
    E=BaseResult(i,1);%peak position for the 1st Gaussian
    F=BaseResult(i,2);%FWHM
    h=BaseResult(i,3);%Intensity
    E2=BaseResult(i,4);%peak position for the 2nd Gaussian
    F2=BaseResult(i,5);%FWHM
    h2=BaseResult(i,6);%Intensity
    slope=BaseResult(i,7);%slope for the linear
    intercept=BaseResult(i,8);%intercept for the linear
    for j=1:length(w)
        Baseline(j,i)=h*exp(-4*log(2)*(w(j)-E)^2/F^2)+h2*exp(-4*log(2)*(w(j)-E2)^2/F2^2);
    end
    Baseline(:,i)=Baseline(:,i)+slope*w+intercept;
end

Spec_Baselined=Spec_Subtracted-Baseline;
Spec_Averaged=sum(Spec_Baselined,2)/SpecNum;
Residual_Averaged=sum(residual,2)/SpecNum;

figure(1)
pplot(w,Baseline,w,Spec_Subtracted,'--')
xlabel('Wavenumber (cm-1)')
ylabel('Difference Absorbance (mOD)')

figure(2)
pplot(w,Spec_Baselined,'--',w,Spec_Averaged)
xlabel('Wavenumber (cm-1)')
ylabel('Difference Absorbance (mOD)')

figure(3)
plot(w,PeakFit(:,2:(pair*2+1)),w,Spec_Averaged,'--')
xlabel('Wavenumber (cm-1)')
ylabel('Difference Absorbance (mOD)')

figure(4)
plot(w,Spec_Averaged,'--',w,PeakFitTotal,w,residual)
xlabel('Wavenumber (cm-1)')
ylabel('Difference Absorbance (mOD)')

Spec_13Scaled=[w,Spec_13Scaled];
Spec_DiffBase=[w,Spec_Subtracted,Baseline,Spec_Baselined];
Fit=[PeakFit,PeakFitTotal,Spec_Averaged,Residual_Averaged,residual];

mkdir new
cd new/
writematrix(Spec_13Scaled);
writematrix(Spec_DiffBase);
writematrix(Fit);
writematrix(result);

function Res=ModFun(x,a)
specnum=length(a(1,:))-1;
pairnum=(length(x)-8*specnum)/8;% every positive-negative pair has 8
parameters
Abs=zeros(length(a(:,1)),1);
for i = 0:(pairnum-1)
    E=x(i*8+1);%peak position
    F=x(i*8+2);%FWHM
    h=x(i*8+3);%Intensity
    m=x(i*8+4);%the percentage of Lorentzian
    shift=x(i*8+5);%the frequency shift from 12C to 13C
    widen=x(i*8+6);%the FWHM widening ratio from 12C to 13C
    Amplif=x(i*8+7);%the Intensity amplification ratio from 12C to 13C
    reshape=x(i*8+8);%the change in Lorentzian percentage from 12C to 13C
    E2=E+shift;
    F2=F*widen;
    h2=h*Amplif;
    if reshape<0 %the Lorentzian percentage of 13C should be in the range
        m2=max(m+reshape,0);
    else
        m2=min(m+reshape,1);
    end
    for j=1:length(a(:,1))
        Abs(j)=Abs(j)+h*(1-m)*exp(-4*log(2)*(a(j,1)-E)^2/F^2)+h*m/(1+4*(a(j,1)-E)^2/F^2);%the Gaussian-Lorentzian sum of 12C peak
        Abs(j)=Abs(j)+h2*(1-m2)*exp(-4*log(2)*(a(j,1)-E2)^2/F2^2)+h2*m2/(1+4*(a(j,1)-E2)^2/F2^2);%the Gaussian-Lorentzian sum of 13C peak
    end
baseline=zeros(length(a(:,1)),specnum);
for i=0:specnum-1
    E=x(pairnum*8+i*8+1);%peak position
    F=x(pairnum*8+i*8+2);%FWHM
    h=x(pairnum*8+i*8+3);%Intensity
    E2=x(pairnum*8+i*8+4);%peak position
    F2=x(pairnum*8+i*8+5);%FWHM
    h2=x(pairnum*8+i*8+6);%Intensity
    slope=x(pairnum*8+i*8+7);
    intercept=x(pairnum*8+i*8+8);
    for j=1:length(a(:,1))
        baseline(j,i+1)=h*exp(-4*log(2)*((a(j,1)-E)^2/F^2)+h2*exp(-4*log(2)*((a(j,1)-E2)^2/F2^2));
    end
    baseline(:,i+1)=baseline(:,i+1)+slope*a(:,1)+intercept;
end

baselined=a(:,2:specnum+1)-baseline;
Res=Abs-baselined;
end
Script 3. Full-time kinetic data fitting to integrated Michaelis-Menten

function
[Km,kcat,dS,StErr,y,residual,KMGuess,kcatGuess]=MMLam_v2(Data,E0input)

%------------------------------Description-----------------------------
% This script generates Michaelis-Menten constants by fitting full-time
% kinetic data

%------------------------------Version Info------------------------------
% The first script was written by Samuel H. Schneider; See ACS Central
% Science, 7, 1996-2008
% The 2nd version was written by Zhe Ji, which directly reads the initial
% susbtrate concentration and makes guesses for kcat and KM.

%------------------------------Input Data--------------------------------
% Data is a 2-column matrix of x-values (1st column; time; sec), and y-values
% (2nd column; substrate concentration; uM)
% E0input is enzyme concentration (uM)

%------------------------------Output Data--------------------------------
% Km (uM) and kcat (s-1) are the Michaelis-Menten constants
% dS is a fitting parameter that adjust any deviation of the final substrate
% concentration from 0; it should be a small value
% StErr is the fitting error of Km, kcat, and dS
% y is the fitted substrate concentration as a function of time
% residual is the fitting residual of substrate concentration as a function
% of time
% KMGuess is the initial guess for KM
% kcatGuess is the initial guess for kcat

%------------------------------Main-----------------------------------
E0=E0input;
t=Data(:,1);
St=Data(:,2);
S0=St(1); %Starting substrate concentration (uM)

function St=Lambert(beta,t)
Km=beta(1);
kcat=beta(2);
dS=beta(3);
S0new=S0-dS;
St=Km*lambertw((exp((S0new-kcat*E0*t)./Km)).*(S0new/Km))+dS;
end

dSGuess=0;
stepmin=20;%the step for sampling slopes
partnum=20;%the number of slopes to be sampled
partleng=fix(length(t)/(partnum+1));
step=min(stepmin,partleng);
slope=zeros(partnum,1);
for i=1:partnum
    start=(i-1)*partleng+1;
    stop=start+step;

LinFit=polyfit(t(start:stop),St(start:stop),1);
slope(i)=LinFit(1);
end
kcatGuess=-slope(1)/E0;
[MinValue,ClosestIndex]=min(abs(slope-slope(1)/2));
KMGuess=St((ClosestIndex-1)*partleng+11);

% Fitting with the closed form of the integrated Michaelis Menten kinetics
fun=@Lambert;
beta0=[KMGuess, kcatGuess, dSGuess];
options=statset('FunValCheck','off');
[beta,R,J,CovB]=nlinfit(t,St,fun,beta0,options);

S0new=S0-beta(3);
y=beta(1)*lambertw((exp((S0new-
beta(2)*E0*t)./beta(1)))*(S0new/beta(1))); % the fitted results
format long
Km=beta(1);
kcat=beta(2);
dS=beta(3);
residual=R;
plot(t,(St-beta(3)),t,y,t,residual);
StErr=(sqrt(diag(CovB)));
end
Script 4. Inhibition kinetic data fitting to eq. 13

```matlab
function [Time,Product,V0,Vs,Kobs,ProductFit,D] = InhibitionFittingPart1(Data,I)

%------------------------------Description-------------------------------
% This script fits the inhibition features (v0,vs, and kobs) from
% experimental data (P) using eq. 13
% Author: Zhe Ji

%------------------------------Input Data------------------------------
% Data is a (n+1)-column matrix. The first column is time (s). The rest
% columns are the concentrations of product for each I. n = the number of
% I. The n experiments are allowed to have different length but the time
% column should cover the longest.
% I is a 1-column matrix containing the concentration of the inhibitor (uM)
% for each experiment

%------------------------------Output Data-------------------------------
% Time is a 1-column matrix (s)
% Product is a n-column matrix of product concentration (uM) as a function of
% time and I
% V0 is a 1-column matrix of the initial rate of product formation as a
% function of I
% Vs is a 1-column matrix of the final rate of product formation as a
% function of I
% Kobs is a 1-column matrix of the apparent first order rate constant for
% forming E-AVB as a function of I
% ProductFit is a n-column matrix of fitted product concentration
% D is a 1-column matrix of a displacement parameter that adjusts the absolute
% product concentration

n=length(Data(1,:))-1; % n is the number of experiments
Time=Data(:,1);
Product=Data(:,2:n+1);
LinLeng1st=7; % the number of data points used for making the guess of the
            % initial slope
LinLeng2nd=300; % the number of data points used for making the guess of the
            % final slope
delay=0.058; % the dead time of the stopped-flow apparatus(s)

function P=ModFun(beta,t)
v0=beta(1);
vs=beta(2);
kobs=beta(3);
d=beta(4); % d is a displacement parameter that adjusts the absolute product
            % concentration
P=vs*(t+delay)+(v0-vs)*(1-exp(-kobs*(t+delay)))/kobs-(vs*delay+(v0-vs)*(1-
            % exp(-kobs*delay))/kobs)+d; % eq. 13
end

fun=@ModFun;
V0=zeros(n,1);```

107
Vs=zeros(n,1);
Kobs=zeros(n,1);
D=zeros(n,1);
Residual=zeros(length(Data(:,1)),n);

for i=1:n
    temp=nonzeros(Product(2:end,i)); % remove the "0" artifacts due to the
different length of data
    ProductNew=[0;temp]; % bring back the first data point "0" that was mis-
removed
    TimeNew=Time(1:length(ProductNew)); % cut the time column
    LinFit=polyfit(TimeNew(1:LinLeng1st),ProductNew(1:LinLeng1st),1); % the
initial slope
    v0Guess=LinFit(1);
    vguess=LinFit((end-LinLeng2nd):end),ProductNew((end-
    LinLeng2nd):end),1); % the final slope
    vsGuess=LinFit(1);
    kGuess=(v0Guess-vsGuess)/LinFit(2); % make guess for kobs
    beta0=[v0Guess, vsGuess, kGuess,0];
    options=statset('FunValCheck','off');
    [beta,R]=nlinfit(TimeNew,ProductNew,fun,beta0,options);
    V0(i)=beta(1);
    Vs(i)=beta(2);
    Kobs(i)=beta(3);
    D(i)=beta(4);
    Residual(1:length(R),i)=R;
end

Product(Product==0)=nan;
Product(1, :) = 0;
ProductFit=Product-Residual; % fitted product concentration

figure(1)
plot(Time,Product,'--',Time,ProductFit);
xlabel('Time (sec)')
ylabel('[Product] (uM)'

figure(2)
scatter(I,V0);
xlabel('[Inhibitor] (uM)')
ylabel('v0 (s-1)'

figure(3)
scatter(I,Vs);
xlabel('[Inhibitor] (uM)')
ylabel('vs (s-1)'

figure(4)
scatter(I,Kobs);
xlabel('[Inhibitor] (uM)')
ylabel('kobs')

end
Script 5. Inhibition kinetic data fitting to eq. 17 and eq. 23

function [Ki, StErrKi, k5, StErrk5, V0disp, IFit, V0Fit, xFit, KobsFit] = InhibitionFittingPart2(V0, Kobs, I, E, S, kcat, KM)

% DESCRIPTION
% This script is used following InhibitionFittingPart1. It fits the inhibition parameters ( Ki and k5) from inhibition features (v0 and kobs) using eq. 17 and eq. 23
% Author: Zhe Ji

% INPUT DATA
% V0 is a 1-column matrix of the initial rate of product formation as a function of I
% Kobs is a 1-column matrix of the apparent first order rate constant for forming E-AVB as a function of I
% I is a 1-column matrix containing the concentration of the inhibitor (uM)
% E is the concentration of the enzyme (uM)
% S is the concentration of the substrate (uM)
% kcat and KM are the Michaelis-Menten constants for the substrate

% OUTPUT DATA
% Ki is the binding constant of the inhibitor for the non-covalent binding, StErrKi is the corresponding fitting error
% k5 is the rate constant for forming E-AVB, StErrk5 is the corresponding fitting error
% V0disp is a 1-column matrix of a displacement parameter that adjusts the absolute initial slope
% IFit is a 1-column matrix of x for the fitted trace of eq. 17
% V0Fit is a 1-column matrix of y for the fitted trace of eq. 17
% xFit is a 1-column matrix of x for data and the fitted trace of eq. 23
% KobsFit is a 1-column matrix of y for the fitted trace eq. 23

% MAIN
vmax = kcat*E;

% FIT Ki FROM v0 USING EQ. 17
function output = ModFun(beta, I)
output = zeros(length(I), 1); % output is v0
V0disp = beta(2); % a displacement parameter that adjusts the absolute initial slope
for i = 1:length(I)
    output(i) = vmax*S/KM/(1+I(i)/beta(1)+S/KM)+V0disp; % eq. 17
end
end

fun = @ModFun;
beta0 = [100, 0]; % make guess for Ki (uM) and the displacement parameter (uM)
options = statset('FunValCheck', 'off');
[beta, R, J, CovB] = nlinfit(I, V0, fun, beta0, options);
Ki = beta(1);
V0disp = beta(2);
StErrKi = sqrt(diag(CovB));
IFit = (min(I):(max(I)-min(I))/100:max(I)).'; % x for the fitted trace
V0Fit = zeros(length(IFit),1); % y for the fitted trace
for i = 1:length(IFit)
    V0Fit(i) = vmax*S/KM/(1+IFit(i)/Ki+S/KM)+V0disp;
end

figure(1)
plot(IFit,V0Fit)
hold on
scatter(I,V0)
xlabel('([Inhibitor] (uM)')
ylabel('v0 (s-1)')
hold off

%------------------------Linear Fitting of k5 using eq. 23-----------------

xFit = I/Ki./(1+S/KM+(I/Ki)); % x for the linear fitting
mdl = fitlm(xFit,Kobs);
k5 = mdl.Coefficients.Estimate(2);
StErrk5 = mdl.Coefficients.SE(2);
KobsFit = mdl.Coefficients.Estimate(1)+k5*xFit; % y for the fitted trace

figure(2)
plot(xFit,KobsFit)
hold on
scatter(xFit,Kobs)
xlabel('I/Ki/(1+S/KM+(I/Ki))')
ylabel('kobs')
hold off

end
Script 6. Inhibition kinetic data fitting to eq. 20

function [Kii, StErrKii, Vsdisp, k5r, StErrk5r, IFit, VsFit] =
InhibitionFittingPart3(Vs, I, E, S, kcat, KM, Ki, StErrKi, k5, StErrk5)

% This script is used following InhibitionFittingPart2. It fits the
% inhibition parameters (Ki* and k-5) from inhibition features (vs) using eq.20
% Author: Zhe Ji

%---------------------------------Input Data--------------------------------
% Vs is a 1-column matrix of the initial rate of product formation as a
% function of I
% I is a 1-column matrix containing the concentration of the inhibitor (uM)
% E is the concentration of the enzyme (uM)
% S is the concentration of the substrate (uM)
% kcat and KM are the Michaelis-Menten constants for the substrate
% Ki is the binding constant of the inhibitor, StErrKi is the corresponding
% fitting error
% k5 is the rate constant for forming E-AVB, StErrk5 is the corresponding
% fitting error

%---------------------------------Output Data--------------------------------
% Kii is the binding constant of the inhibitor for both non-covalent and
% covalent binding (Ki*), StErrKii is the corresponding fitting error
% Vsdisp is a 1-column matrix of a displacement parameter that adjusts the
% absolute final slope
% k5r is the rate constant for the reverse reaction of forming E-AVB, StErrk5r is the corresponding fitting error
% IFit is a 1-column matrix of x for the fitted trace of eq. 20
% VsFit is a 1-column matrix of y for the fitted trace of eq. 20

vmax=kcat*E;

%---------------Fit Ki* from vs using eq. 20-----------------------------
function output=ModFun(beta,I)
    output=zeros(length(I),1);
    d=beta(2);
    for i=1:length(I)
        output(i)=vmax*S/KM/(1+I(i)/beta(1)+S/KM)+d; % eq. 20
    end
end

fun=@ModFun;
beta0=[0.0001,0]; % make guess for Ki* and Vsdisp
options=statset('FunValCheck','off');
[beta, R, J, CovB]=nlinfit(I,Vs,fun,beta0,options);
Kii=beta(1);
Vsdisp=beta(2);
StErr=(sqrt(diag(CovB)));
StErrKii=StErr(1);
IFit= (min(I):(max(I)-min(I))/100:max(I)).'; % x for the fitted trace
VsFit=zeros(length(IFit),1); % y for the fitted trace
for i=1:length(IFit)
    VsFit(i)=vmax*S/KM/(1+IFit(i)/Kii+S/KM)+Vsdisp;
end

figure(1)
plot(IFit,VsFit)
hold on
scatter(I,Vs)
xlabel('Inhibitor (uM)')
ylabel('vs (s$^{-1}$)')
hold off

%------------------------Fit $k_5r$ from $K_i$, $K_{i^*}$, and $k_5$------------------------
k5r=Kii*k5/(Ki-Kii);
StErrk5r=k5*((StErrKii/Kii)^2+(StErrk5/k5)^2+(StErrKi(1)^2+StErrKii^2)/((Ki-Kii)^2))^0.5;

end
Script 7. Simulation of Inhibition from kinetic parameters

```matlab
function [t,P,V0, Vs,Kobs] = InhibitionSimulation(E,S,KM,kcat, I, Ki,k5,k5r)

%------------------------------Description-------------------------------
% This script simulates the inhibition data (P) and features (v0, vs, and kobs) from the kinetic parameters (Ki, k5, k-5)
% Author: Zhe Ji

%-------------------------------Input Data-------------------------------
% E is a 1-column matrix containing the concentration of the enzyme (uM) for each experiment
% S is the value of substrate concentration (uM)
% KM (uM) and kcat (s-1) are the Michaelis-Menten constants for the substrate
% I is a 1-column matrix containing the concentration of the inhibitor (uM) for each experiment
% Ki is the binding constant of the inhibitor for non-covalent binding
% k5 is the rate constant for forming E-AVB
% k5r is the rate constant for the reverse reaction of E-AVB formation (k-5)

%-------------------------------Output Data-------------------------------
% t is a 1-column matrix of time (s)
% P is a x-column matrix of product concentration (uM) as a function of time and I, x= number of I
% V0 is a 1-column matrix of the initial rate of product formation as a function of I
% Vs is a 1-column matrix of the final rate of product formation as a function of I
% Kobs is a 1-column matrix of the apparent first order rate constant for forming E-AVB as a function of I

%---------------------------------Main-----------------------------------

\text{t}=0:0.1:250; \ % the range of time
\text{P}=[];
\text{V0}=[];
\text{Vs}=[];
\text{Kobs}=[];

\text{for\ num = 1:length(I)}
\text{i}=I(num);
\text{V=E(num)\cdot kcat; \ % Vmax}
\text{v0=V\cdot S/(KM\cdot (1+i/Ki)+S);}
\text{Kii=Ki\cdot k5r/(k5+k5r);}
\text{vs=V\cdot S/(KM\cdot (1+i/Kii)+S);}
\text{kobs=k5r+k5\cdot (i/Ki/(1+S/KM+i/Ki));}
\text{p=vs\cdot t+(v0-vs)*\cdot (1{-exp(-kobs\cdot t)})/kobs;}
\text{P=[P; p];}
\text{V0=[V0; v0];}
\text{Vs=[Vs; vs];}
\text{Kobs=[Kobs; kobs];}
\text{end}

\text{legendcell=\text{strcat('I=',string(num2cell(I))));}
```
```matlab
figure(1)
plot(t,P)
xlabel('Time (sec)')
ylabel('[Product] (uM)')
legend(legendcell)

figure(2)
plot(I,V0./E.*E(1))
xlabel('[Inhibitor] (uM)')
ylabel('V0 (uM s-1)')

figure(3)
plot(I,Vs./E.*E(1))
xlabel('[Inhibitor] (uM)')
ylabel('Vs (uM s-1)')

figure(4)
plot(I,Kobs)
xlabel('[Inhibitor] (uM)')
ylabel('kobs (s-1)')
end
```
References


118

Fakih, M. G.; O’Neil, B. H.; Price, T. J.; Falchook, G. S.; Desai, J.; Kuo, J.; Govindan,
R.; Hong, D. S.; Ouyang, W.; Henary, H.; Arvedson, T.; Cee, V. J.; Lipford, J. R. The
Clinical KRAS(G12C) Inhibitor AMG 510 Drives Anti-Tumour Immunity. Nature 2019,
575, 217-223.

Corbett, M.; Tamura, J. K.; He, B.; Hamann, L. G.; Kirby, M. S. Involvement of DPP-IV
Catalytic Residues in Enzyme-Saxagliptin Complex Formation. Protein Sci. 2008, 17,
240-250.

Cao, H.; Newton, A.; Petropoulos, C. J.; Huang, W.; Schiffer, C. A. The Molecular Basis
2012, 8, e1002832.