Supporting Information for

Ortho-Substituents Govern Aryl Aldehyde Reactivity: Toward Lysine-Targeted, Tunable Inhibitors of Glucose-6-Phosphate Dehydrogenase

*Bronwyn E. Rowland[†], Kelsey D. Roy[‡], Adil Alkaş[†‡], Mirele Barsoum[†], Tangweth Kuot[‡], Jesse C. Fuller[†], Adrien J. Naudet[#], Elisa Ospanow[†], David L. Jakeman[†‡]**

† Department of Chemistry, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada.

‡ College of Pharmacy, Dalhousie University, Halifax, Nova Scotia, B3H 4R2, Canada.

Faculté de Pharmacie, Université Grenoble Alpes, Grenoble, 38400, France.

* Author to whom correspondence should be addressed.
email: David.jakeman@dal.ca

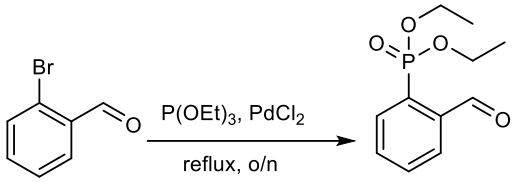
TABLE OF CONTENTS

1. SYNTHESIS AND CHARACTERIZATION

1.1 General methods for synthesis

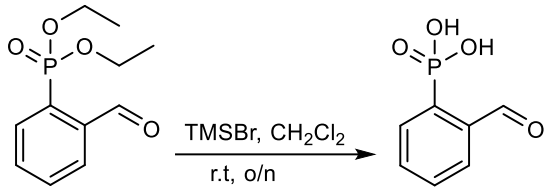
All starting materials and solvents were used as received from commercial sources unless otherwise noted. Nuclear magnetic resonance (NMR) spectra were recorded using a 400 MHz spectrometer. ^{1}H chemical shifts are reported in ppm relative to tetramethylsilane, referenced to the resonances of $CDCl_3$ (δ = 7.26 ppm), D_2O (δ = 4.69 ppm), or DMSO-d_6 (δ = 2.50 ppm) as internal standards. ^{13}C chemical shifts are proton decoupled and reported in ppm relative to tetramethylsilane, referenced to the resonances of $CDCl_3$ (δ = 77.16 ppm) or DMSO-d_6 (δ = 39.52 ppm) as internal standards. ^{31}P chemical shifts are reported in ppm, externally referenced to phosphoric acid (H_3PO_4, δ = 0.00 ppm). ^{1}H NMR splitting patterns are indicated as follows: s, singlet; d, doublet; dd, doublet of doublet; t, triplet; q, quartet; m, multiplet. All coupling constants (*J*) are reported in Hertz (Hz). Thin-layer chromatography was performed using commercially prepared silica gel plates and visualized using long- or short-wave UV lamps. Column chromatography was performed using 40-63 micron particle size (60 Å) silica gel.

1.2 Synthetic procedures and characterization



Diethyl-ortho-formylphenylphosphonate. 2-Bromobenzaldehyde (1.85 g, 10.00 mmol) was combined with triethyl phosphite (2.23 mL, 13.00 mmol) and $PdCl_2$ (89 mg, 0.5 mmol), and the mixture was refluxed under nitrogen atmosphere overnight. The reaction mixture was then allowed to cool to room temperature. The crude material was columned over silica using gradient hexane:EtOAc (v:v; 95:5→70:30) as eluent. The semi-pure product was purified over silica using gradient hexane:diethylether (v:v; 50:50 → 0:100) as eluent to obtain diethyl-ortho-formylphenylphosphonate as a colourless oil (775 mg, 32%). Note: *This material is not stable over long term storage. It rapidly decomposes in the presence of H_2O and MeOH.*

^{1}H NMR (400 MHz, $CDCl_3$) δ: 10.68 (s, 1H), 8.56-7.96 (m, 2H), 7.90-7.57 (m, 2H), 4.45-3.89 (m, 4H), 1.35 (td, *J* = 7.1, 0.5 Hz, 6H) ppm.
^{13}C NMR (101 MHz, $CDCl_3$) δ:191.94 (d, *J* = 3.9 Hz), 138.05 (d, *J* = 9.9 Hz), 134.11 (d, *J* = 8.9 Hz), 133.08 (d, *J* = 14.2 Hz), 132.74 (d, *J* = 3.0 Hz), 130.57 (d, *J* = 182.6 Hz), 128.06 (d, *J* = 12.6 Hz), 62.75 (d, *J* = 5.9 Hz), 16.34 (d, *J* = 6.3 Hz) ppm.
^{31}P NMR (162 MHz, $CDCl_3$) δ: 16.33 ppm. Data is in accordance with literature.[1]



***2-Formylphenylphosphonic acid* (11).** Following a literature procedure,[2] diethyl-ortho-formylphenylphosphonate (740 mg, 3.06 mmol) was dissolved in dichloromethane (10 mL). To this solution, bromotrimethylsilane (2.00 mL, 15.16 mmol) was added under a nitrogen atmosphere and the reaction mixture was stirred at room temperature overnight. Solvent was evaporated in vacuo, deionized water (20 mL) was added, and the solution stirred for 1 h. Aqueous phase was washed with diethylether (10 mL, X2). After removal of water, 2-formylphenylphosphonic acid was obtained as a pale orange solid (545 mg, 96%).

^{1}H NMR (400 MHz, DMSO-d$_6$) δ: 10.73 (d, *J* = 0.8 Hz, 1H), 7.97-7.87 (m, 2H), 7.76 (tdd, *J* = 7.6, 2.9, 1.5 Hz, 1H), 7.74-7.64 (m, 1H) ppm.
^{13}C NMR (101 MHz, DMSO-d$_6$) δ: 193.40 (d, *J* = 4.5 Hz), 137.67 (d, *J* = 12.3 Hz), 136.76 (d, *J* = 151.0 Hz), 133.86 (d, *J* = 12.8 Hz), 132.93 (d, *J* = 7.9 Hz), 131.96 (d, *J* = 2.8 Hz), 127.26 (d, *J* = 11.8 Hz) ppm.
^{31}P NMR (162 MHz, DMSO-d$_6$) δ: 9.34 ppm. Data is in accordance with the literature data.[1]

A pH titration of 2-formylphenylphosphonic acid (**11**) was performed by preparing a 50 mM solution in D_2O. A pH range of about 2 to 10 (adjusted to account for the use of deuterated solvent)[3] was chosen to observe changes in the ^{1}H and ^{31}P NMR shifts. NaOD and DCl were used to adjust the solution to the desired pH before NMR spectra were recorded. Titration curves were fit to the equation below to derive pK_a, where δH and δH are the fully protonated and fully deprotonated chemical shifts, respectively.[4]

$$\mathbf{Y} = \frac{\boldsymbol{\delta}\mathbf{H}(\mathbf{10}^{\boldsymbol{pK_a - pH}}) + \boldsymbol{\delta}\mathbf{L}}{\mathbf{1} + \mathbf{10}^{\boldsymbol{pK_a - pH}}}$$

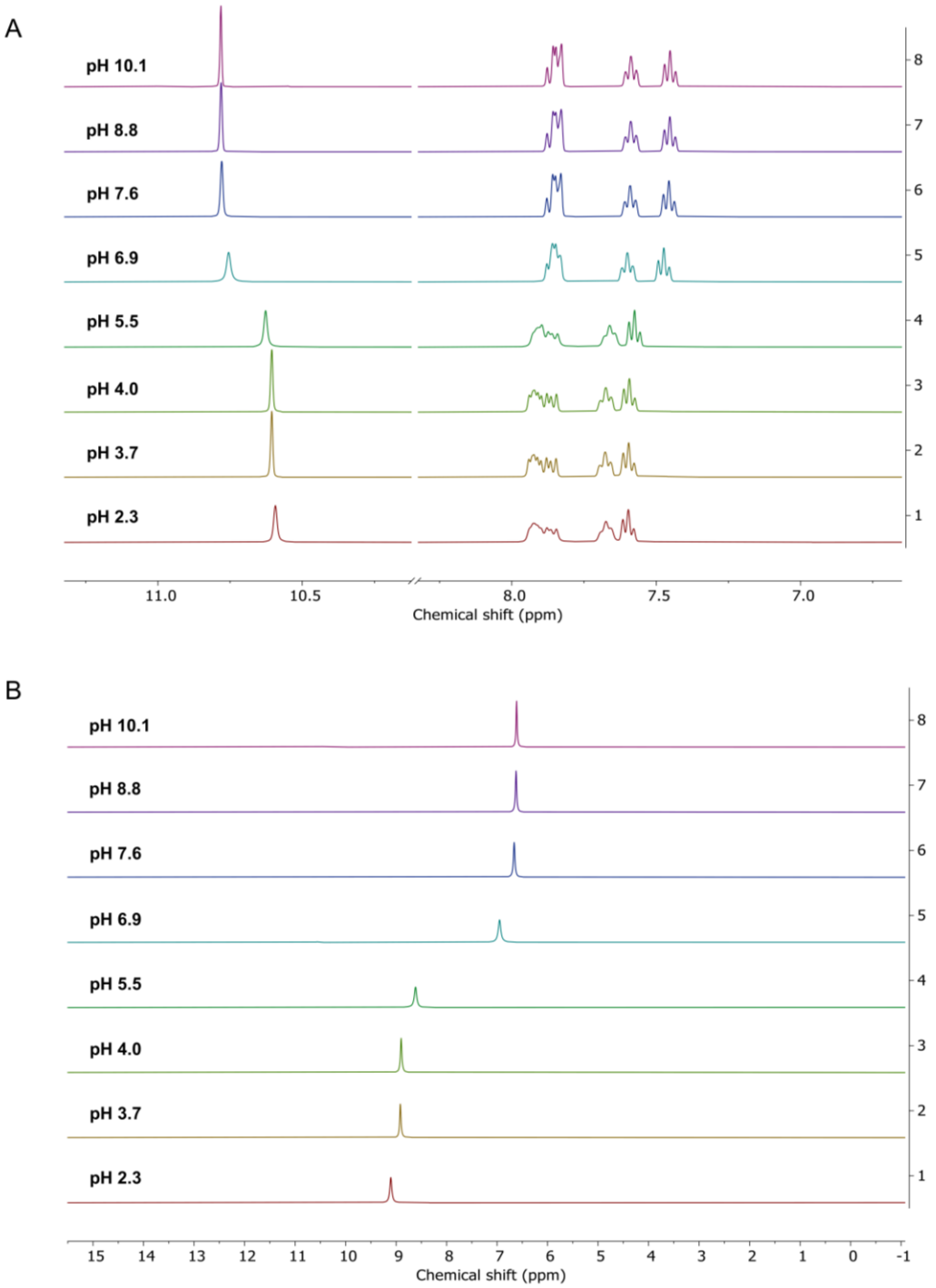


Figure S1. (A) ^{1}H NMR and (B) ^{31}P NMR spectra of 2-formylphenylphosphonic acid (compound **11**) across a pH titration.

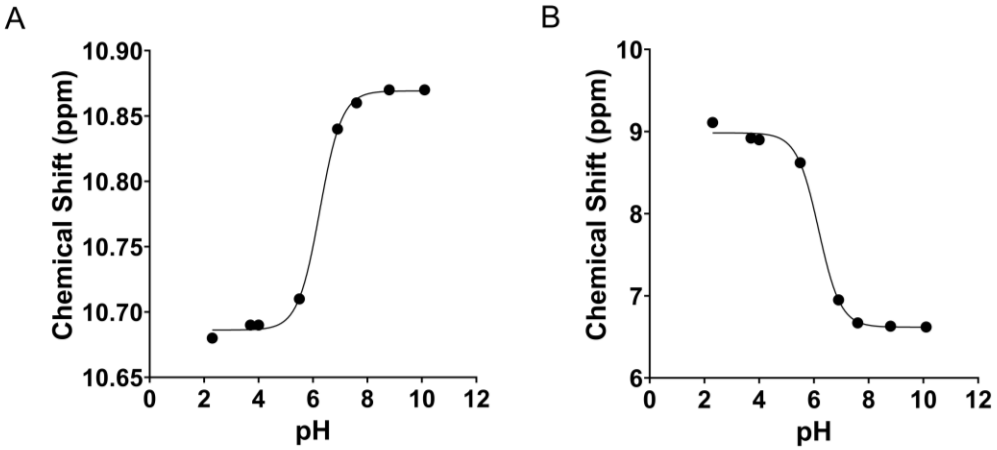
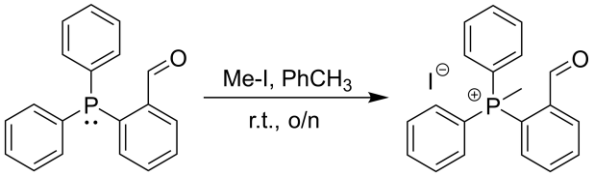


Figure S2. Titration curve for 2-formylphenylphosphonic acid (compound **11**) derived from (A) ^{1}H NMR and (B) ^{31}P NMR spectra. The pK_a was calculated as the inflection point of the fitted curve, yielding values of 6.3 and 6.2, respectively.



***(2-Formylphenyl)(methyl)diphenylphosphonium iodide* (12).** Synthesis was performed as previously described.[5] Iodomethane (0.70 mL, 10.3 mM) was added to a solution of 2-(diphenylphosphino)benzaldehyde (500 mg, 1.72 mM) in toluene (40 mL), and the mixture was stirred overnight at room temperature. The yellow precipitate was filtered, washed with toluene (5 x 5 mL), and dried under vacuum. (2-Formylphenyl)(methyl)diphenylphosphonium iodide was obtained as a yellow solid (360 mg, 48.5%). ^{1}H NMR (400 MHz, D_2O) and ^{31}P NMR (162 MHz, D_2O) were in accordance with the literature data.[5]

1.3 NMR spectra of synthesized compounds

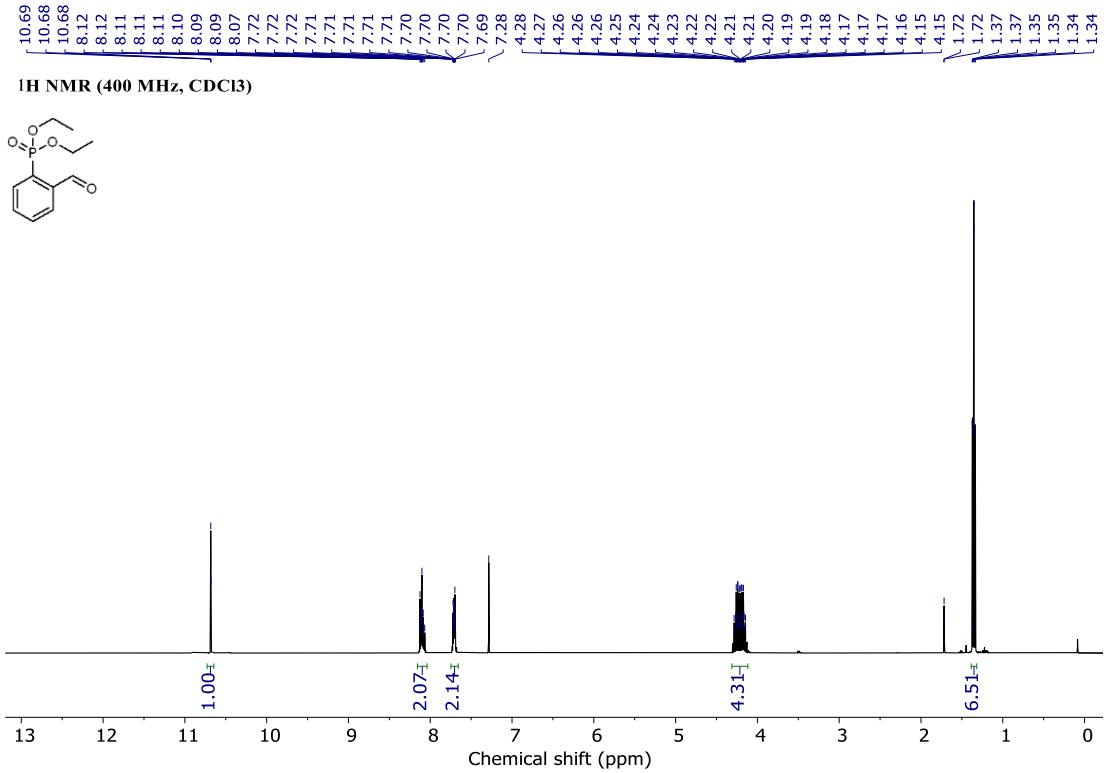


Figure S3. ^{1}H NMR spectrum of diethyl-ortho-formylphenylphosphonate.

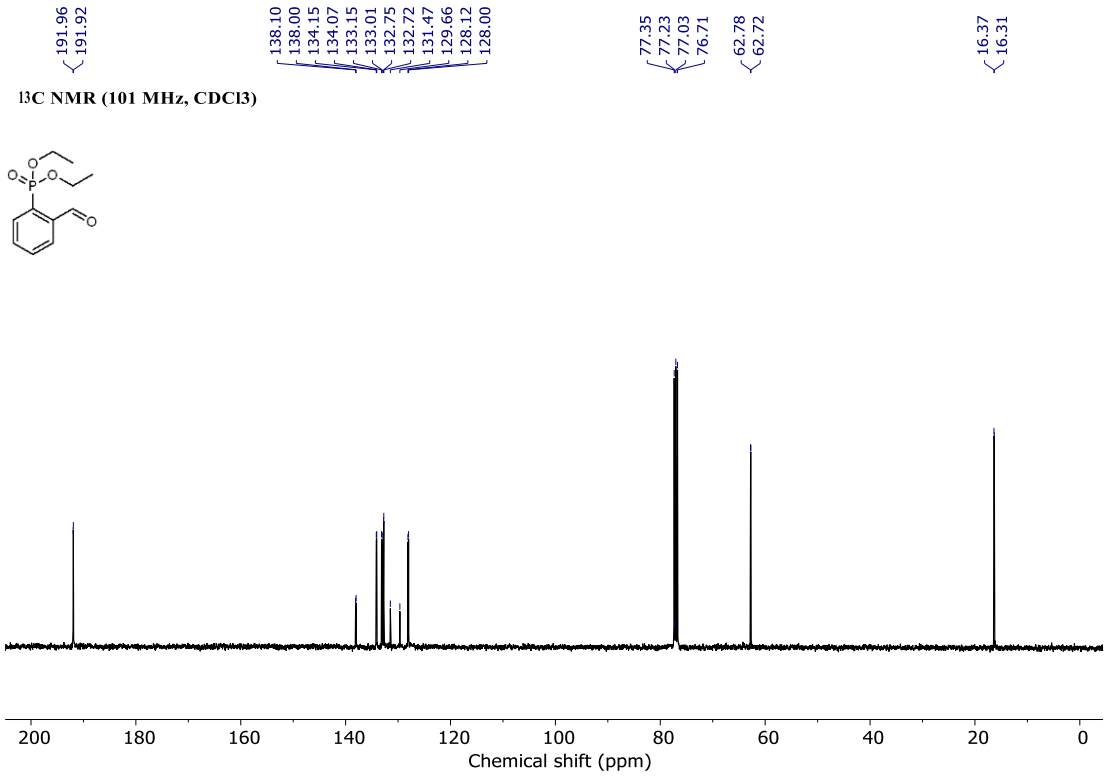


Figure S4. ^{13}C NMR spectrum of diethyl-ortho-formylphenylphosphonate.

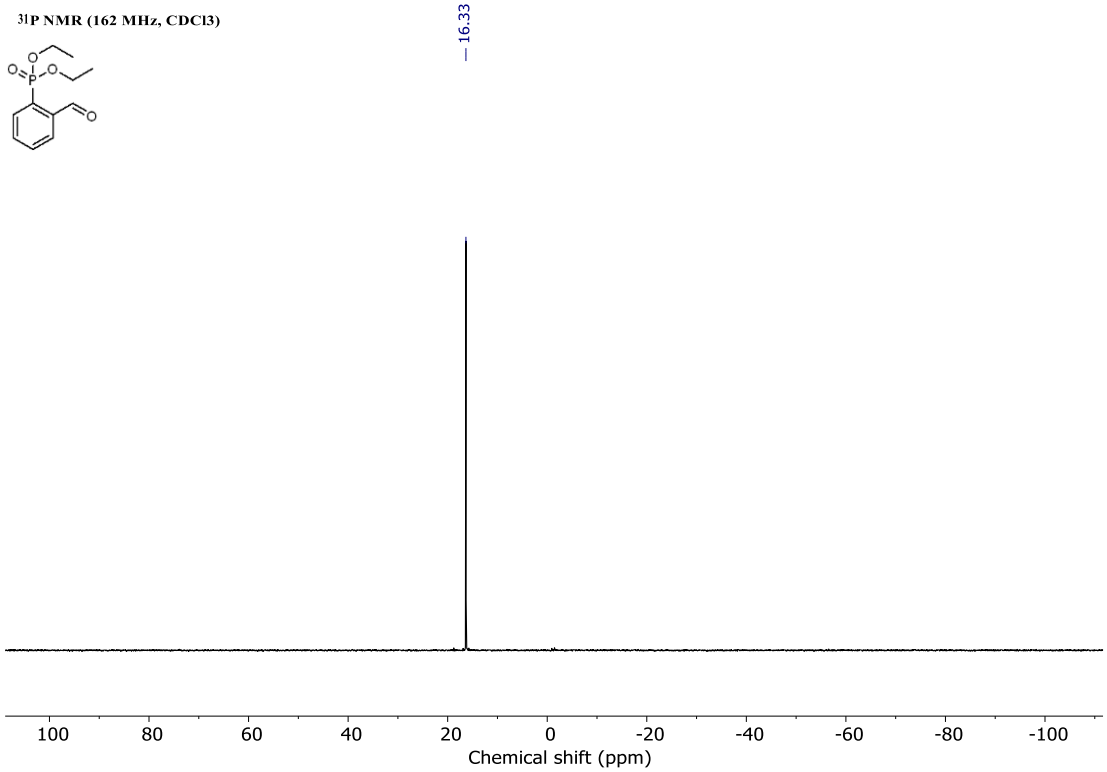


Figure S5. ^{31}P NMR spectrum of diethyl-ortho-formylphenylphosphonate.

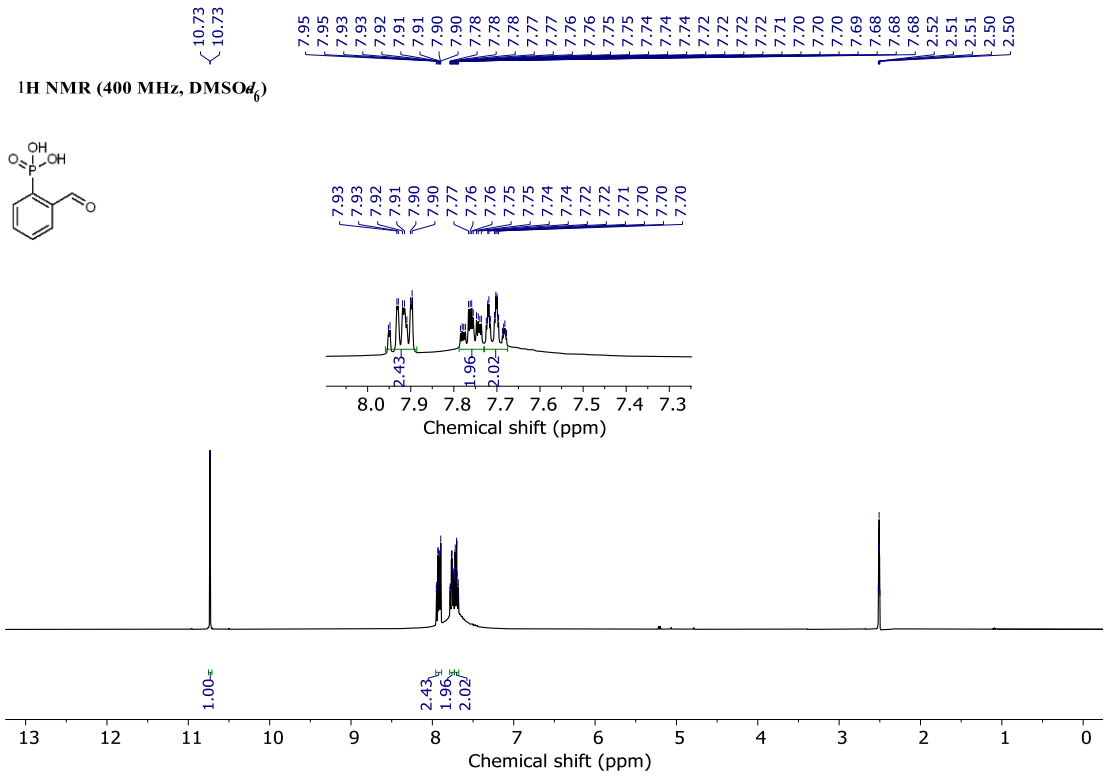


Figure S6. [1]H NMR spectrum of 2-formylphenylphosphonic acid (**11**).

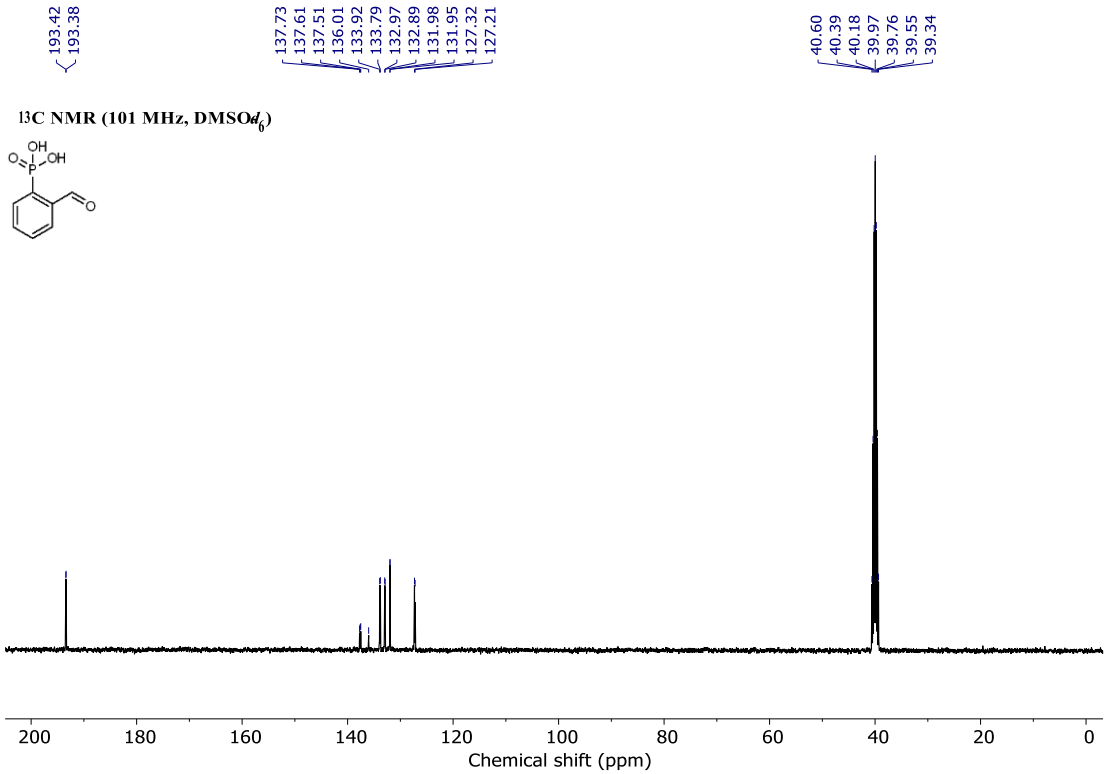


Figure S7. [13]C NMR spectrum of 2-formylphenylphosphonic acid (**11**).

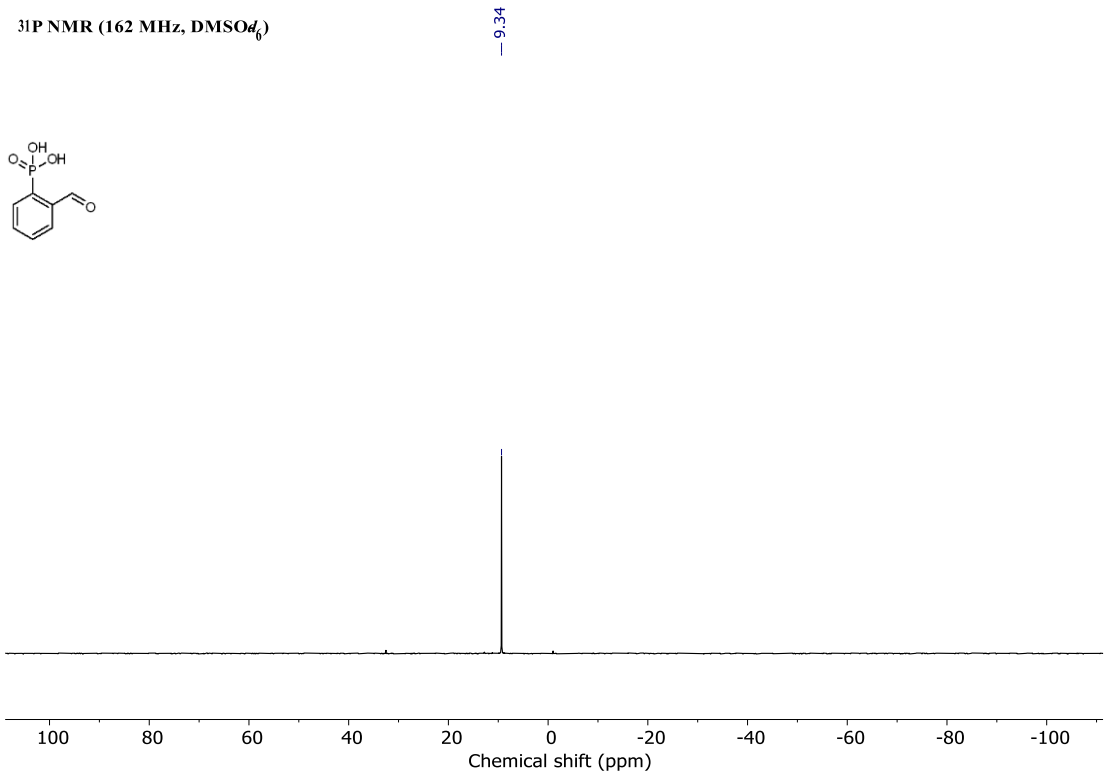


Figure S8. ^{31}P NMR spectrum of 2-formylphenylphosphonic acid (**11**).

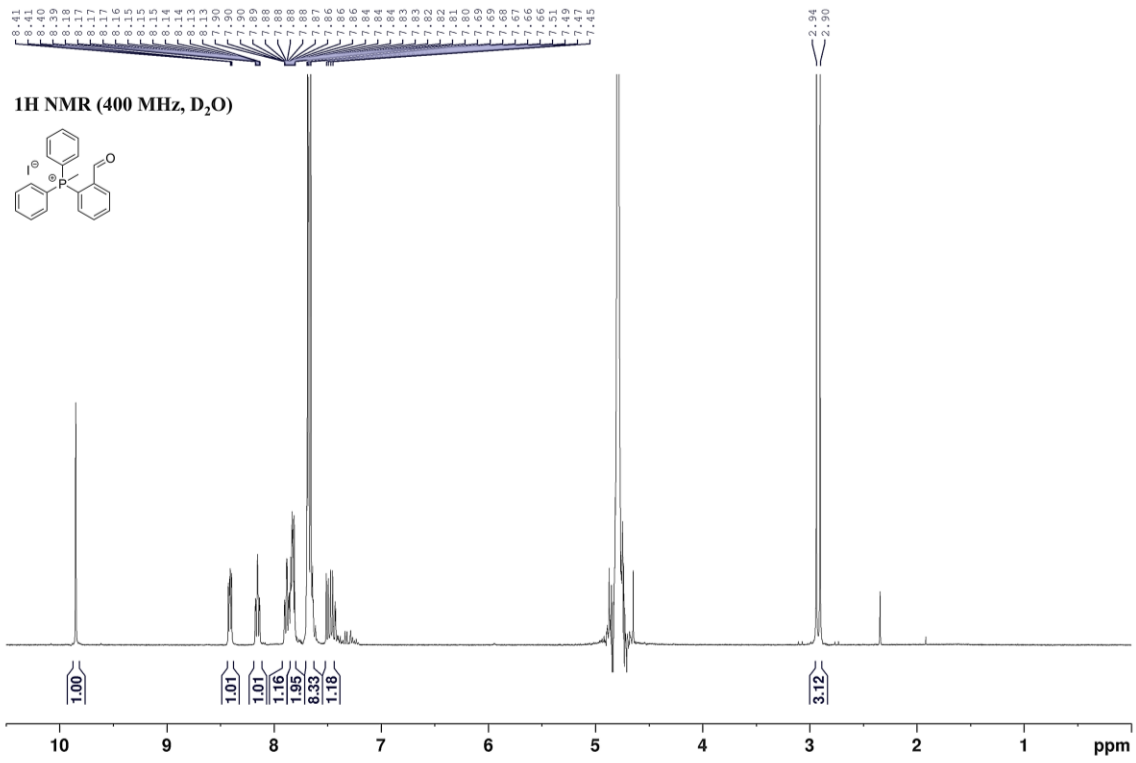


Figure S9. ^{1}H NMR spectrum of (2-formylphenyl)(methyl)diphenylphosphonium iodide (**12**).

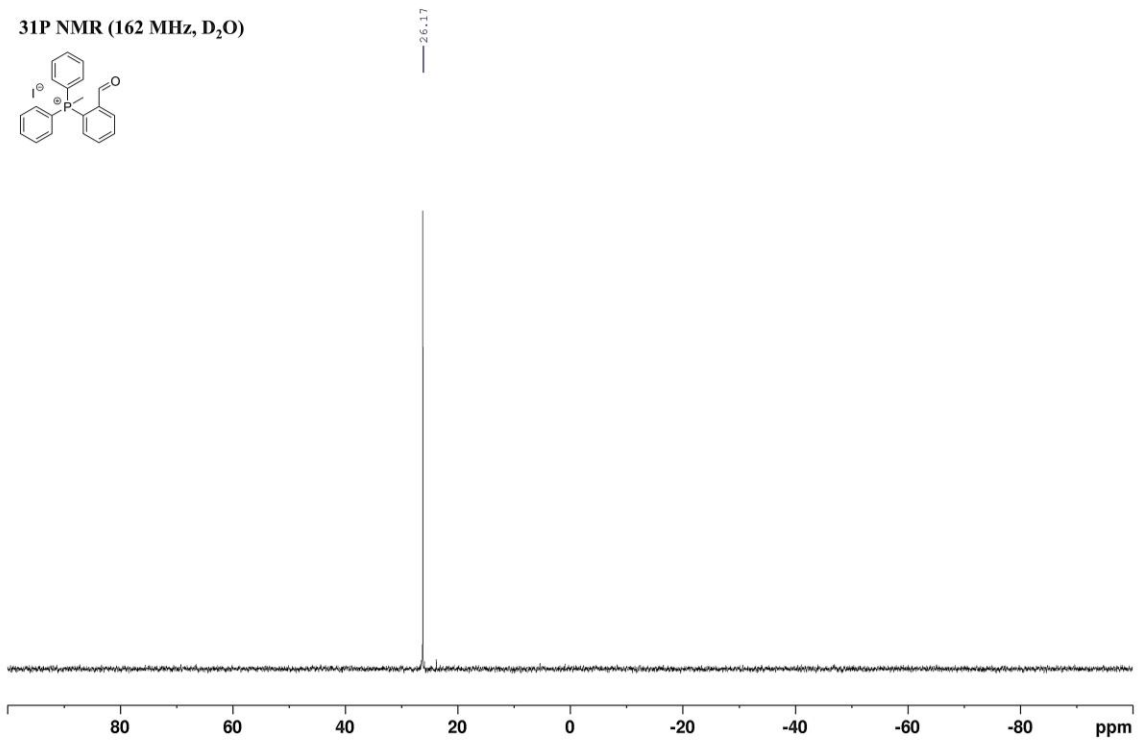


Figure S10. ^{31}P NMR spectrum of (2-formylphenyl)(methyl)diphenylphosphonium iodide (**12**).

2. NMR AND UV-VIS STUDIES OF IMINE FORMATION

2.1 Calculation of imine yield

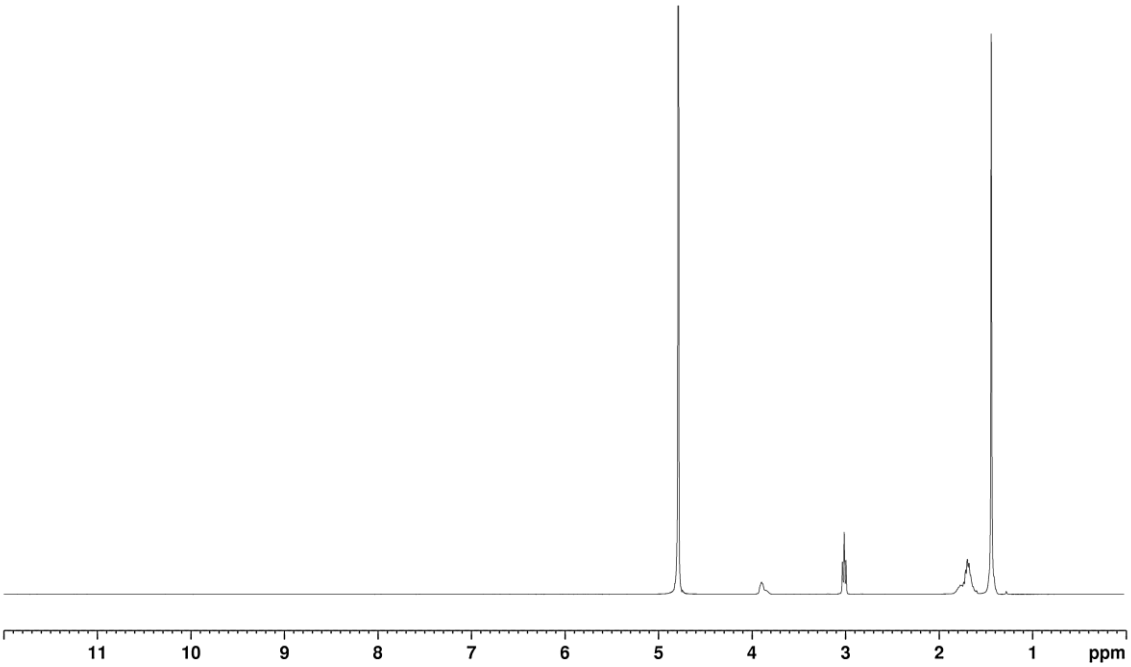


Figure S11. ^{1}H NMR spectrum of Boc-Lys in PBS (pH 7.4), recorded in D_2O.

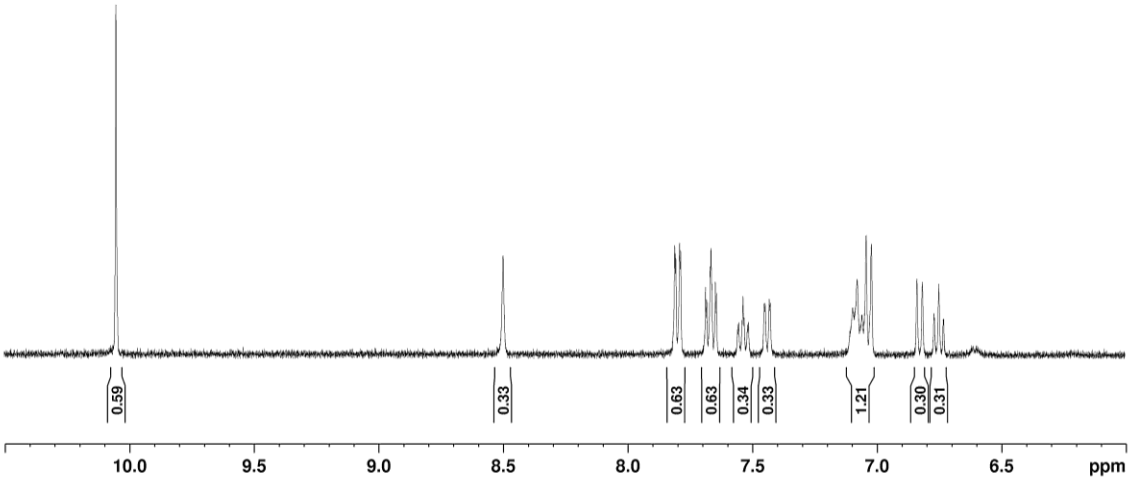


Figure S12. Calculation of aldehyde, imine, and hydrate abundance. Reaction between compound **3** and Boc-Lys used as example. Peak at 10.05 ppm corresponds to the unreacted aldehyde. Peak at 8.50 corresponds to the imine product. No peak observed for aldehyde hydrate (usually found between 6-6.5 ppm).[6] Proportion of aldehyde is therefore 0.59/(0.59+0.33)*100 = 64.1%. Proportion of imine is 0.33/(0.59+0.33)*100 = 35.9%.

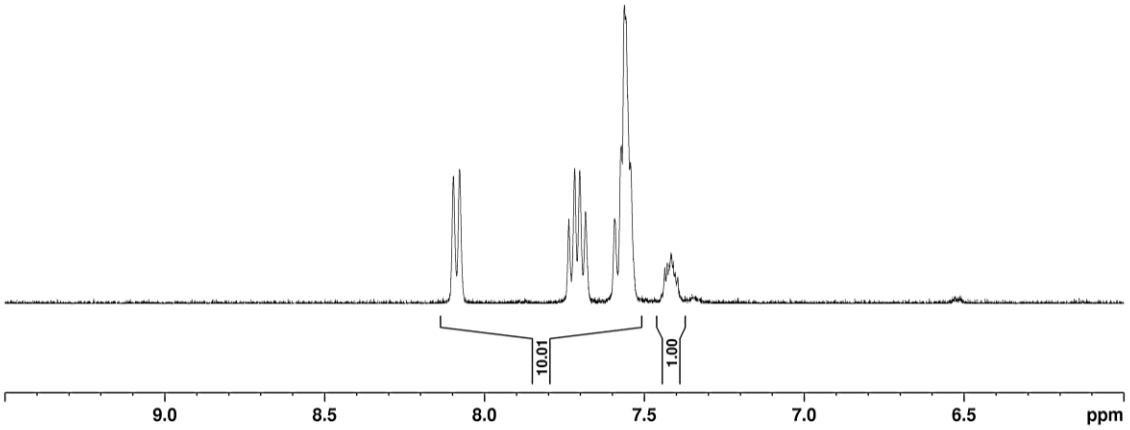


Figure S13. Calculation of ketone, imine, and hydrate abundance for reaction between compound **6** and Boc-Lys. Spectrum is calibrated to the peak at 7.42, corresponding to an aromatic hydrogen from the imine product. The rest of the peaks correspond to the 3 other aromatic hydrogens from the imine, and 4 from the unreacted ketone. No indication of hydrate formation was found in the spectrum. As such, $10.01 = 3H_{imine} + (4H_{ketone})x$, where x is the ratio of ketone to imine. Proportion of ketone to imine is therefore x=1.75, or 64% ketone and 36% imine.

2.2 NMR studies

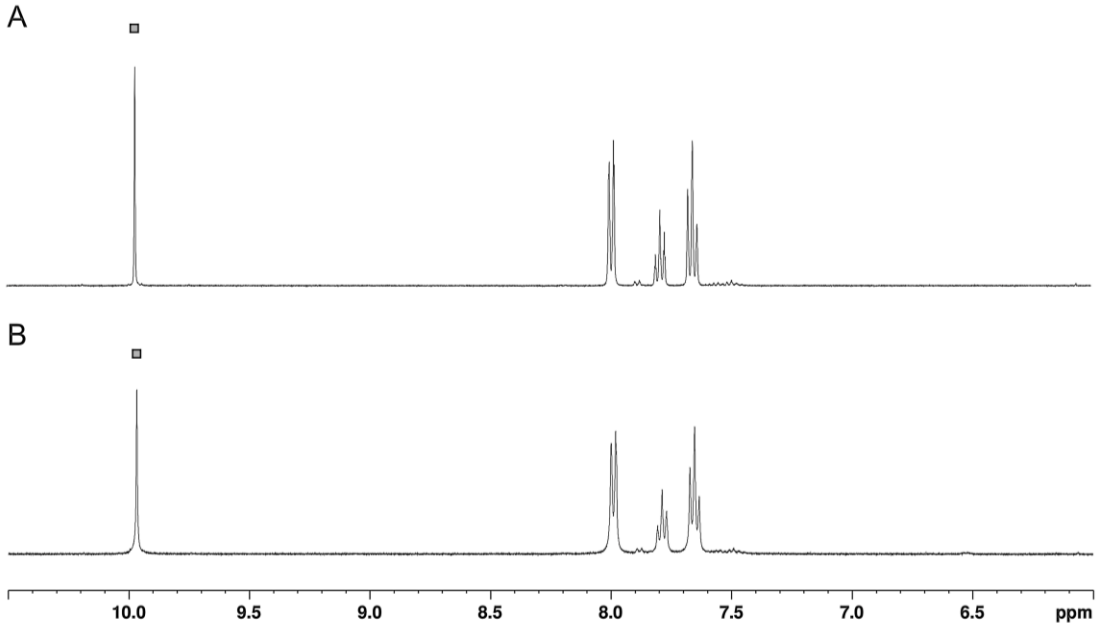


Figure S14. ^{1}H NMR spectra of (A) compound **1** and (B) the reaction mixture of **1** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde.

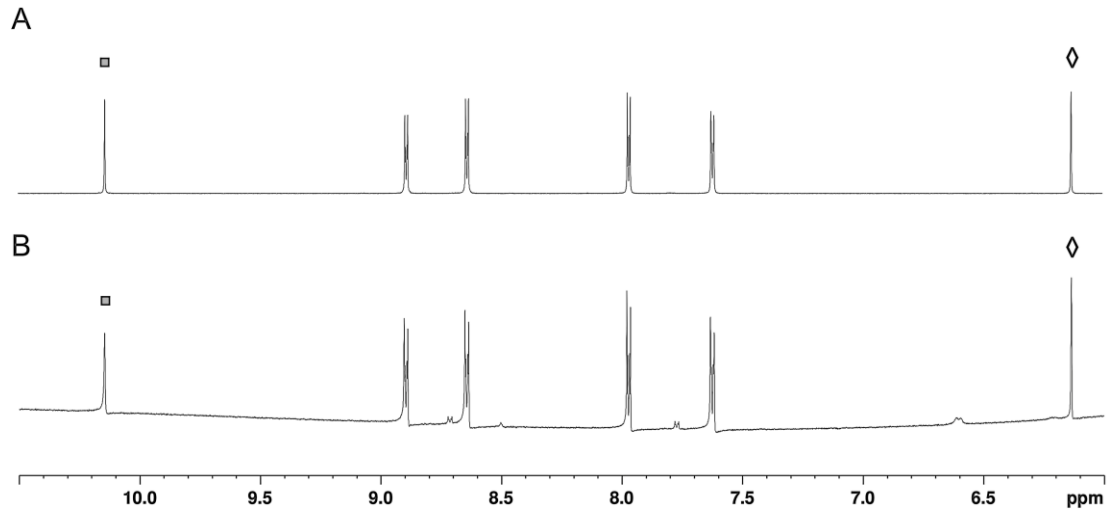


Figure S15. ^{1}H NMR spectra of (A) compound **2** and (B) the reaction mixture of **2** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde, open diamonds (◊) correspond to the aldehyde hydrate.

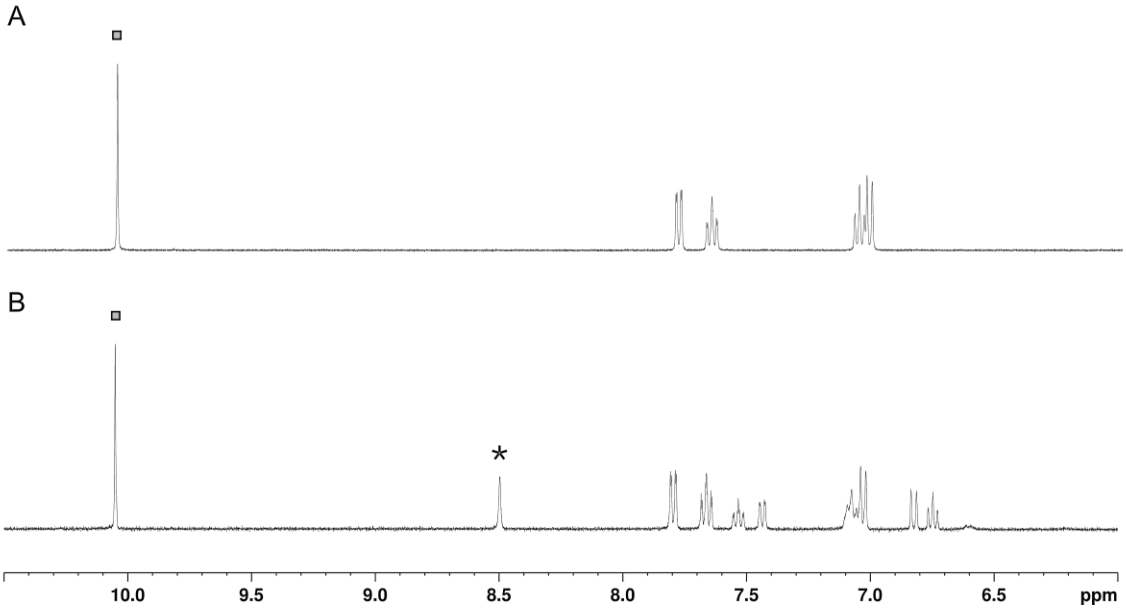


Figure S16. ^{1}H NMR spectra of (A) compound **3** and (B) the reaction mixture of **3** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde, asterisk (*****) corresponds to the imine.

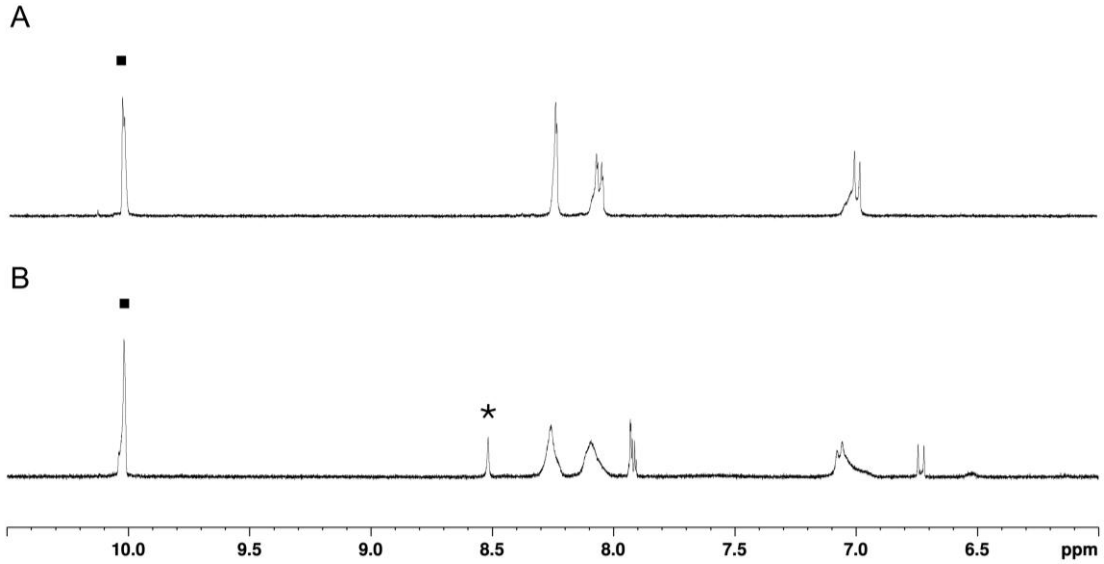


Figure S17. ^{1}H NMR spectra of (A) compound **4** and (B) the reaction mixture of **4** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde, asterisk (*****) corresponds to the imine.

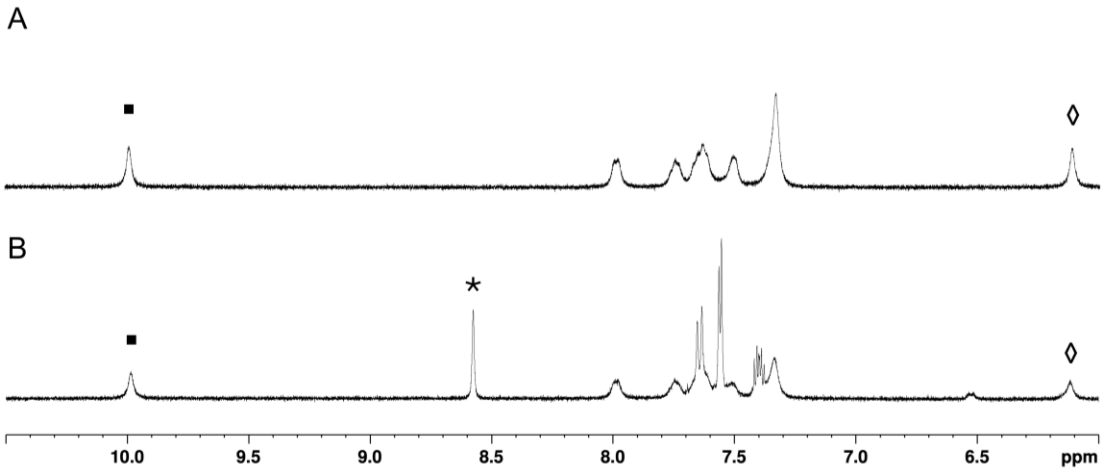


Figure S18. ^{1}H NMR spectra of (A) compound **5** and (B) the reaction mixture of **5** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde, asterisk (*) corresponds to the imine. The hash symbols (#) correspond to a tautomer of compound **5** where a cyclization occurs between the aldehyde and boronic acid.[7]

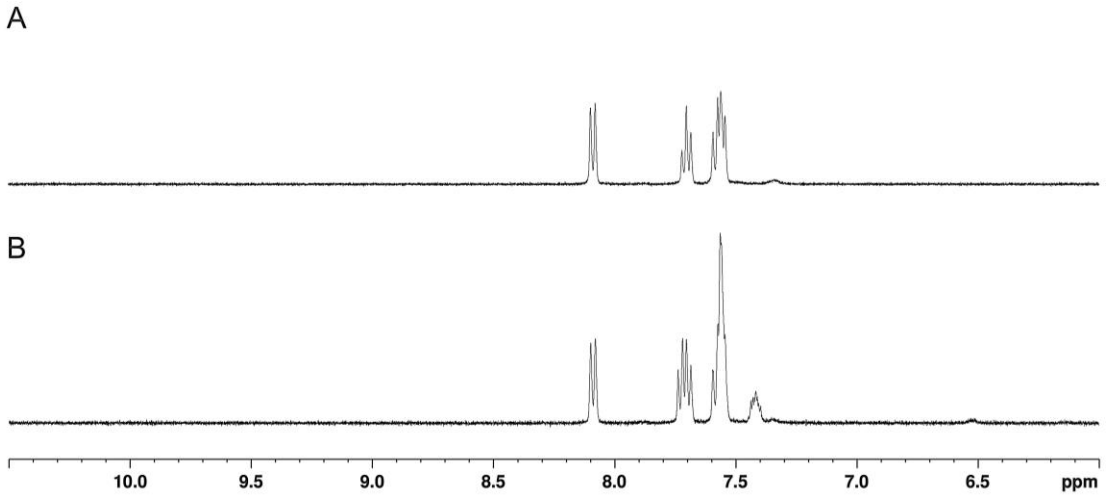


Figure S19. ^{1}H NMR spectra of (A) compound **6** and (B) the reaction mixture of **6** with Boc-Lys in PBS (pH 7.4), recorded in D_2O.

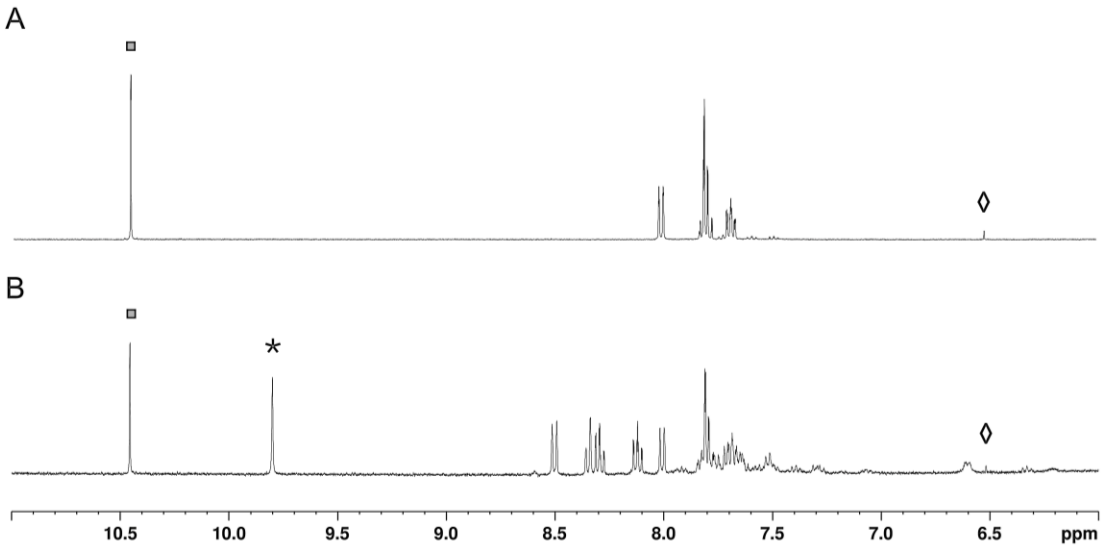


Figure S20. ^{1}H NMR spectra of (A) compound **7** and (B) the reaction mixture of **7** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde, open diamonds (◊) correspond to the aldehyde hydrate, asterisk (*****) corresponds to the imine.

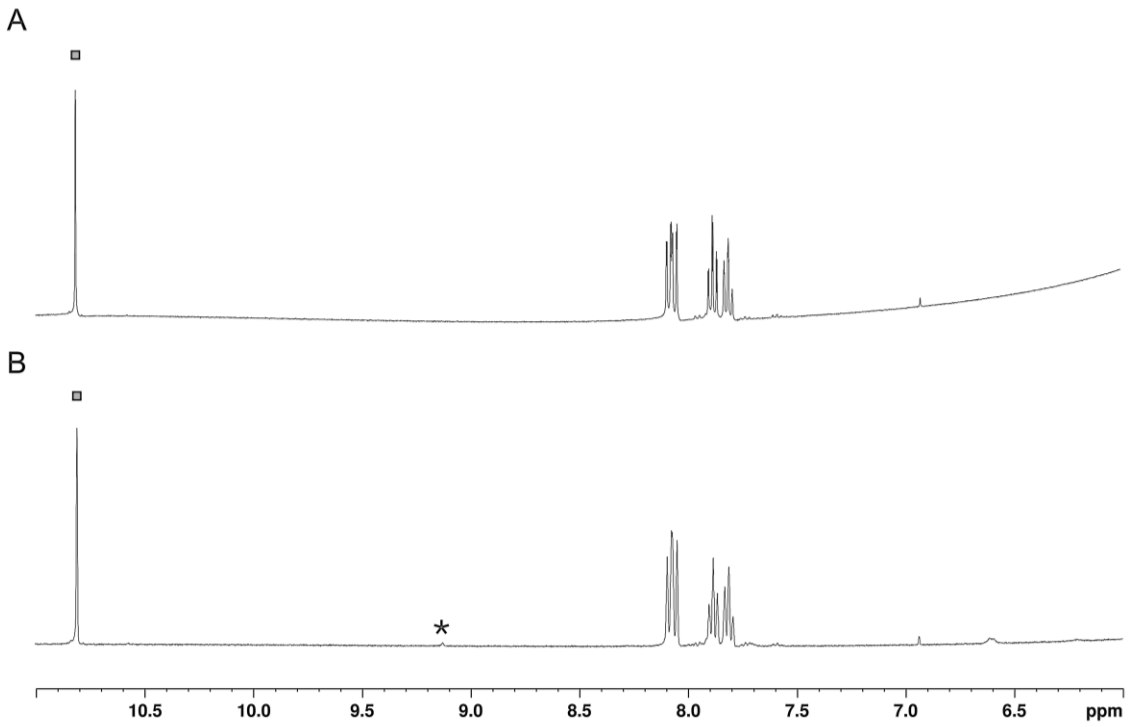


Figure S21. ^{1}H NMR spectra of (A) compound **8** and (B) the reaction mixture of **8** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde, asterisk (*****) corresponds to the imine.

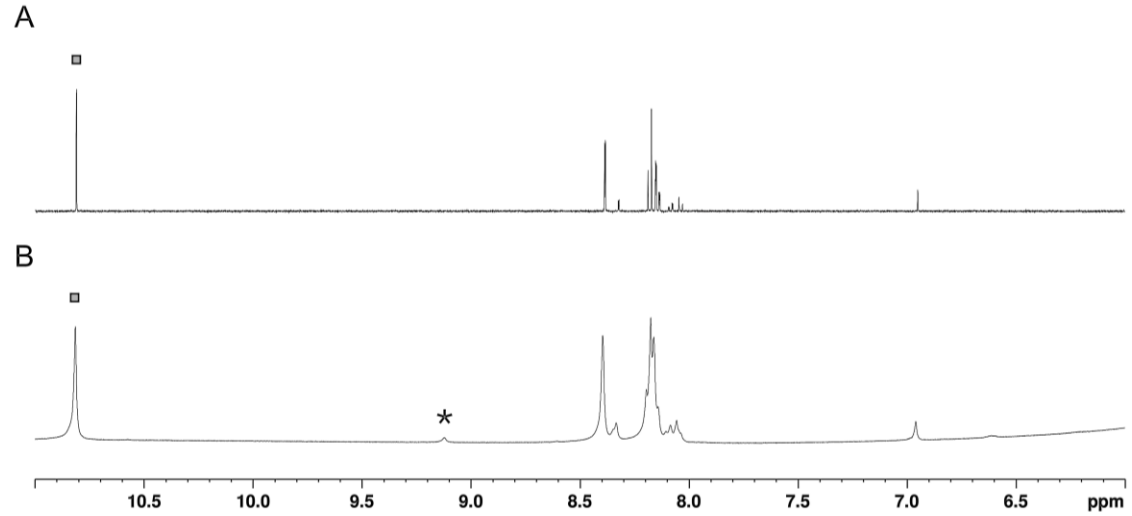


Figure S22. ^{1}H NMR spectra of (A) compound **9** and (B) the reaction mixture of **9** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde, asterisk (*****) corresponds to the imine.

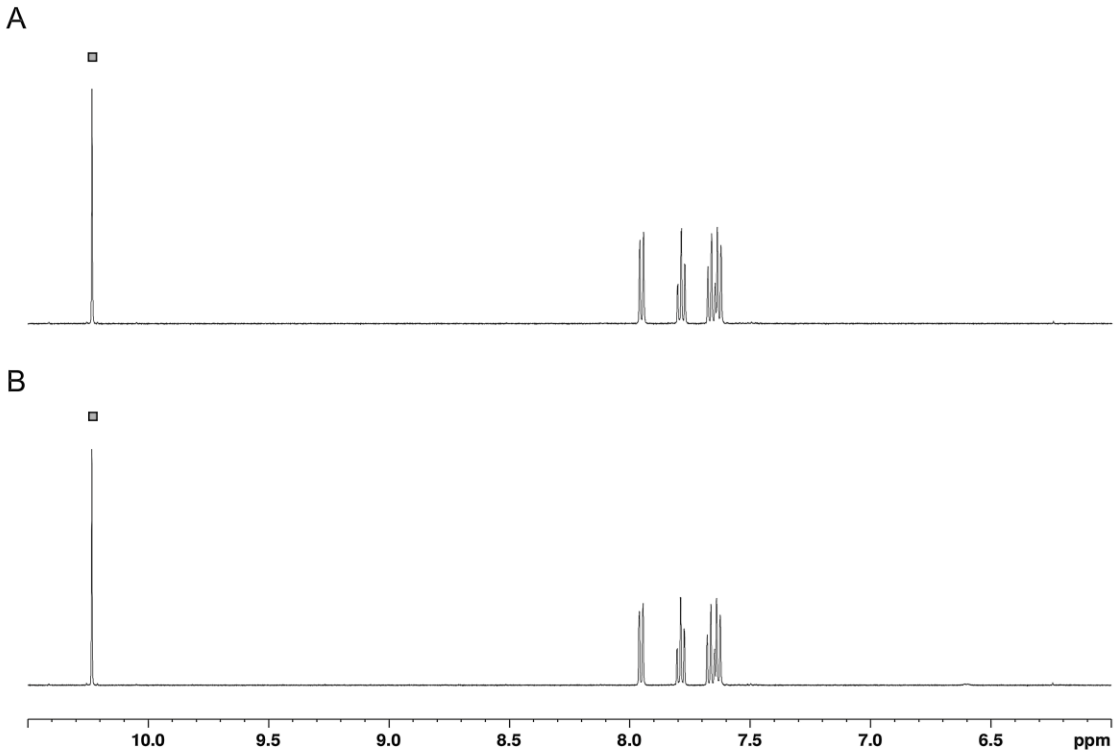


Figure S23. ^{1}H NMR spectra of (A) compound **10** and (B) the reaction mixture of **10** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde.

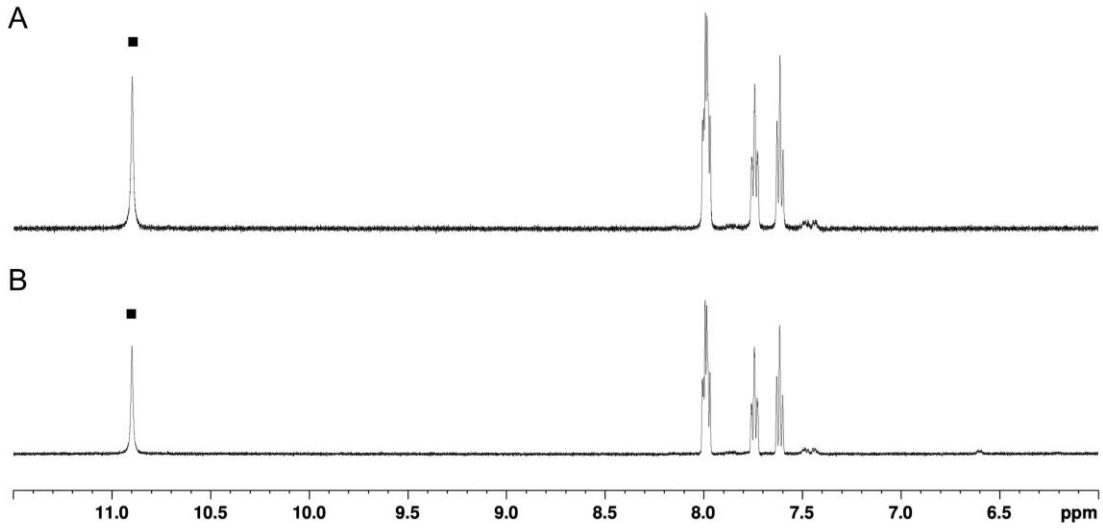


Figure S24. ^{1}H NMR spectra of (A) compound **11** and (B) the reaction mixture of **11** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde.

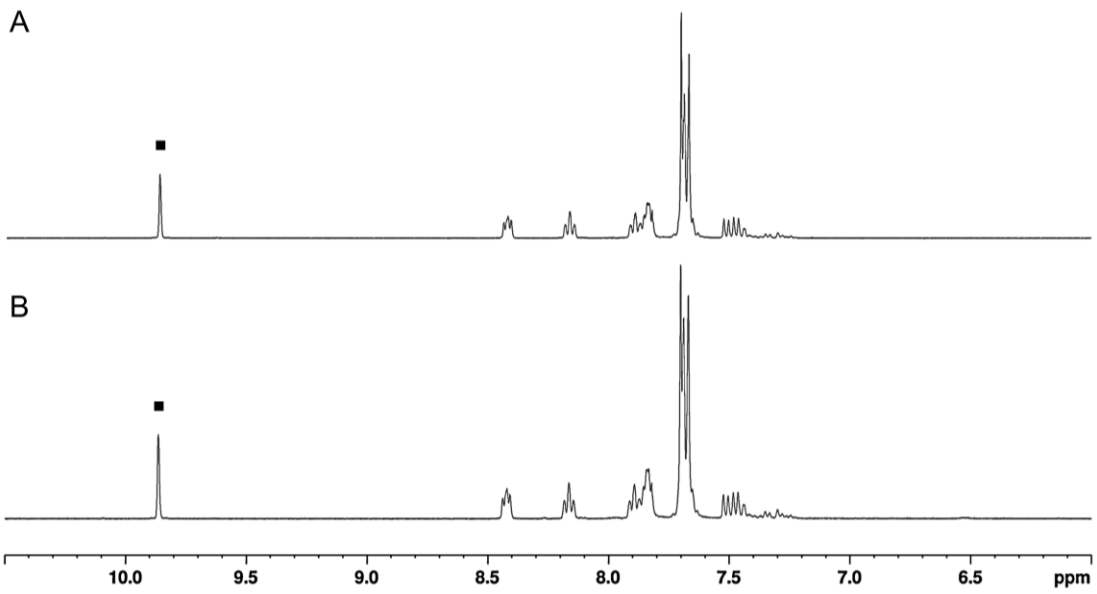


Figure S25. ^{1}H NMR spectra of (A) compound **12** and (B) the reaction mixture of **12** with Boc-Lys in PBS (pH 7.4), recorded in D_2O. Solid squares (■) correspond to the aldehyde.

2.3 Additional figures for UV-Vis studies

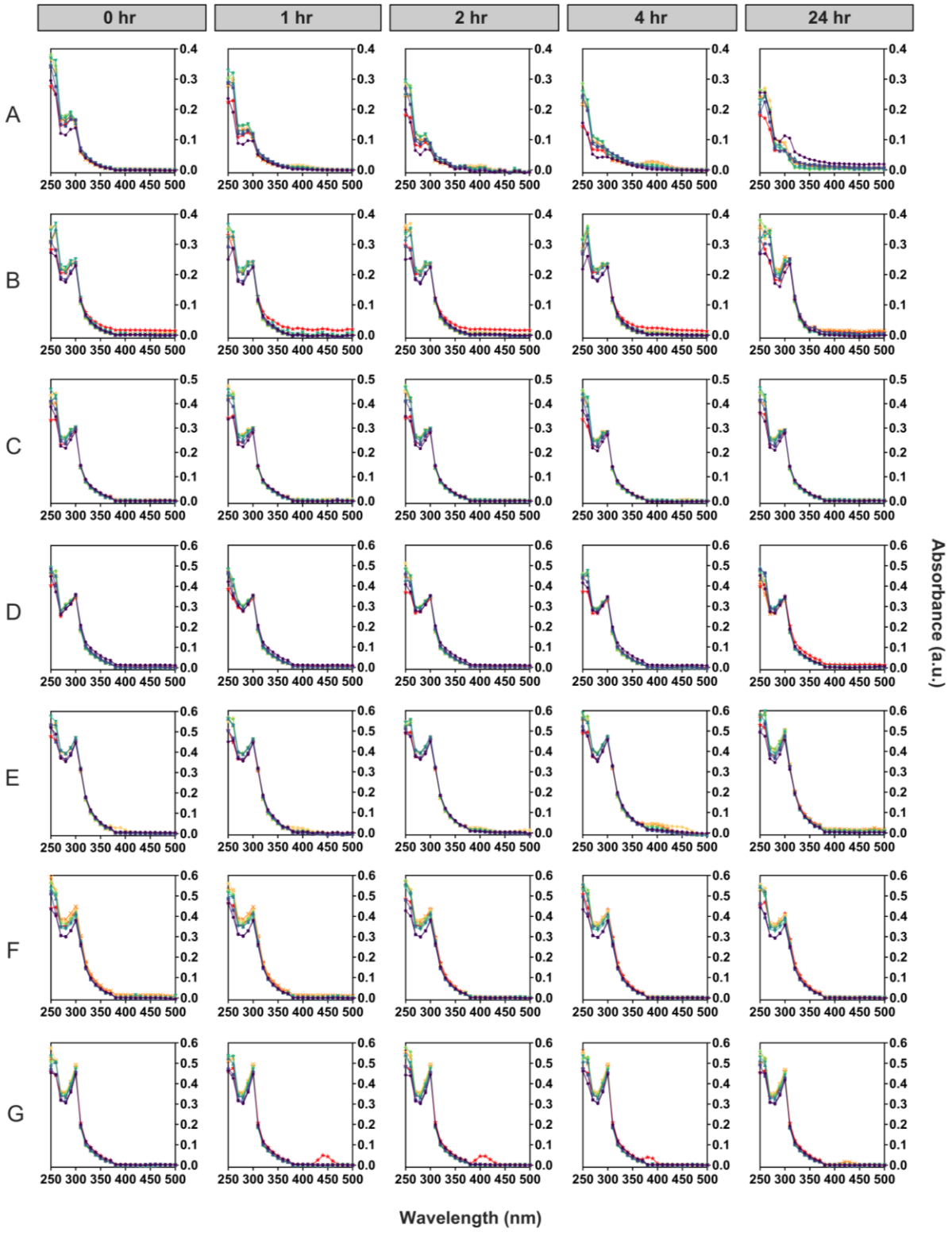


Figure S26. Spectrophotometric titrations of Boc-Lys with (A) benzaldehyde (**1**), (B) 4-pyridinecarboxaldehyde (**2**), (C) 2-formylbenzenesulfonic acid (**8**), (D) benzaldehyde-2,4-disulfonic acid (**9**), (E) 2-carboxybenzaldehyde (**10**), (F) 2-formylphenylphosphonic acid (**11**), and (G) (2-formylphenyl)(methyl)diphenylphosphonium iodide (**12**). Reaction mixtures contained 1 mM aryl aldehyde/ketone with 100 mM (★), 60 mM (×), 40 mM (✳), 20 mM (◆), 10 mM (▼), 5 mM (▲), 2 mM (■), or 0 mM (●) Boc-Lys in PBS (pH 7.4). Reactions carried out at room temperature and spectra recorded at 0, 1, 2, 4, and 24 hrs post-initiation. Data shown correspond to one replicate.

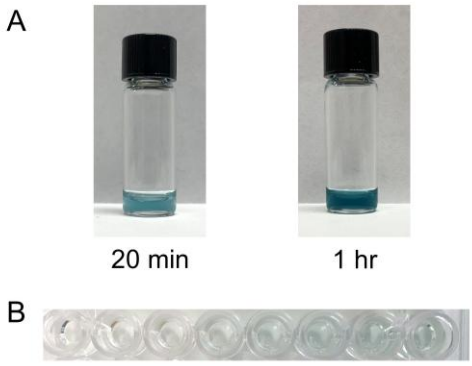


Figure S27. Isoquinolinium adduct production in reaction between **7** and Boc-Lys. (A) Colour change over time in a reaction mixture of 2 mM compound **7** and 20 mM Boc-Lys in PBS (pH 7.4). Solution turns from clear to blue-green within 20 min. (B) Colour change after 24 hours in reaction mixtures containing 1 mM compound **7** and varying concentrations of Boc-Lys (0 mM (left), 2 mM, 5 mM, 10 mM, 20 mM, 40 mM, 60 mM, 100 mM (right)).

3. ADDITIONAL FIGURES

3.1 Kinetic characterization of G6PD

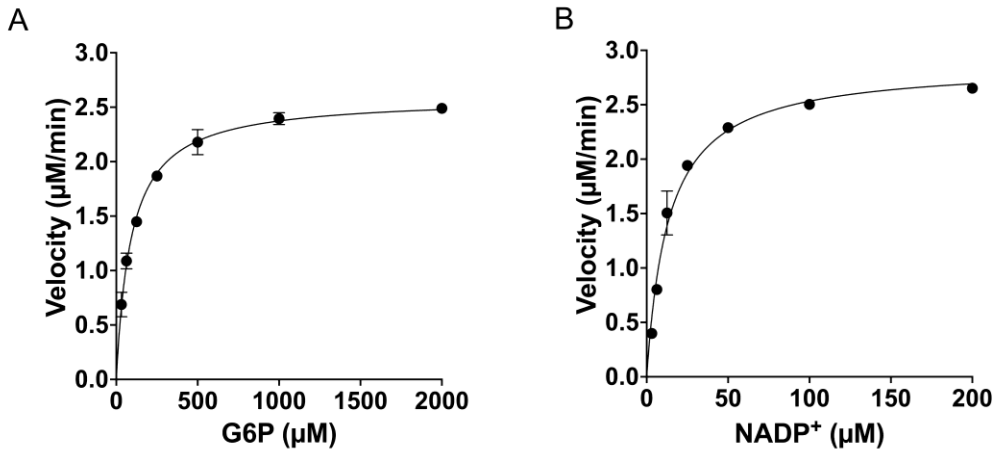


Figure S28. Michaelis-Menten curves of (A) G6P substrate and (B) $NADP^+$ coenzyme for *L. mesenteroides* G6PD. G6PD activity was determined via NADPH production across varying substrate/coenzyme concentrations. Initial velocities were plotted as a function of substrate/coenzyme concentration and fit to Eqn. 3. Data shown are mean ± SD (n=2). Some error bars are smaller than the data points.

Table S1. Apparent kinetic parameters of *L. mesenteroides* G6PD.

Substrate	K_m^{app} (mM)[a]	V_{max}^{app} (µM/min)[a]
G6P	92.0 (± 5.1)	2.6 (± 0.1)
$NADP^+$	13.4 (± 1.2)	2.9 (± 0.1)

[a]Data reported as parameter estimates ± SE (n=2), derived from nonlinear least-squares regression.

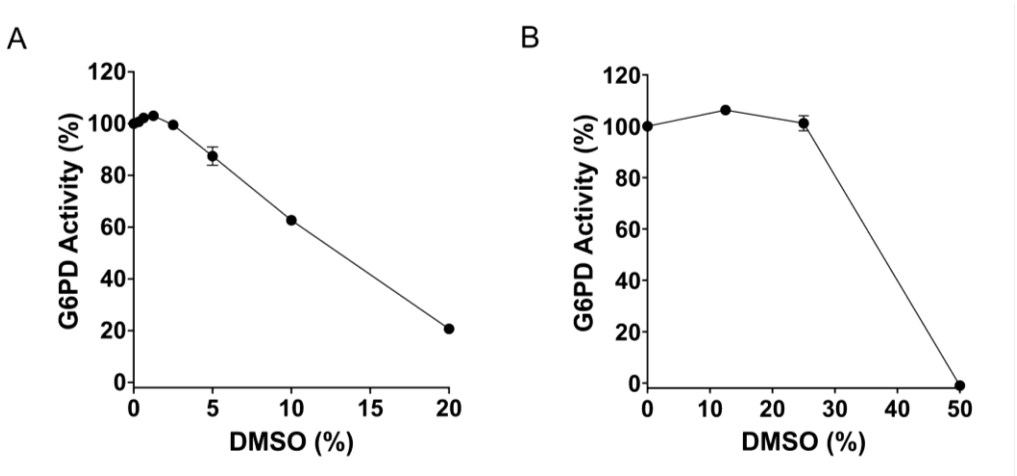


Figure S29. *L. mesenteroides* G6PD tolerance to DMSO. (A) Effect of DMSO in a reaction mixture of 0.001 U/mL G6PD, 100 µM $NADP^+$, and 500 µM G6P in MOPS buffer (pH 7.4). (B) Effect of DMSO incubated with 0.025 U/mL G6PD in MOPS buffer (pH 7.4) (30 min, r.t.). Enzyme activity measured after a 50-fold dilution with substrate and coenzyme, such that final concentrations were 0.0005 U/mL G6PD, 100 µM $NADP^+$, 500 µM G6P, and 0-1% DMSO. Data shown are mean ± SD (n=2). Some error bars are smaller than the data points.

3.2 Determination of IC_{10}s

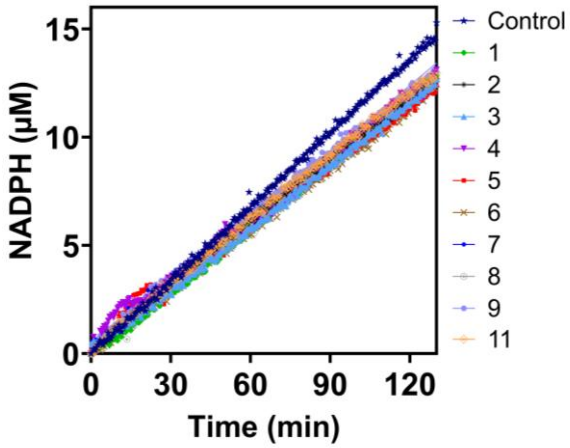


Figure S30. Verification of estimated IC_{10}s for L. mesenteroides G6PD inhibitors. G6PD activity was measured via NADPH production in reaction mixtures containing 0.001 U/mL G6PD, 100 µM $NADP^+$, 500 µM G6P, and aryl aldehyde (compounds **1-9, 11**). Aryl aldehyde concentration was chosen such that ~90% G6PD activity remained compared to the uninhibited control (**Table S2**). Enzyme activity was estimated by linear regression of the progress curves. Data shown correspond to one replicate.

Table S2. Estimated IC_{10}s of *L. mesenteroides* G6PD inhibitors.

Aryl aldehyde/ketone	Concentration (mM)	Velocity (μM/min)	G6PD activity (%)
Control	–	0.113	100.0
1	0.2	0.097	85.7
2	4.0	0.099	87.3
3	0.01	0.096	85.1
4	0.2	0.099	87.5
5	0.3	0.096	84.9
6	0.4	0.094	83.3
7	0.05	0.099	87.3
8	5.0	0.097	85.6
9	2.0	0.103	91.1
11	6.0	0.102	89.8

3.3 Calculation of inhibition constants

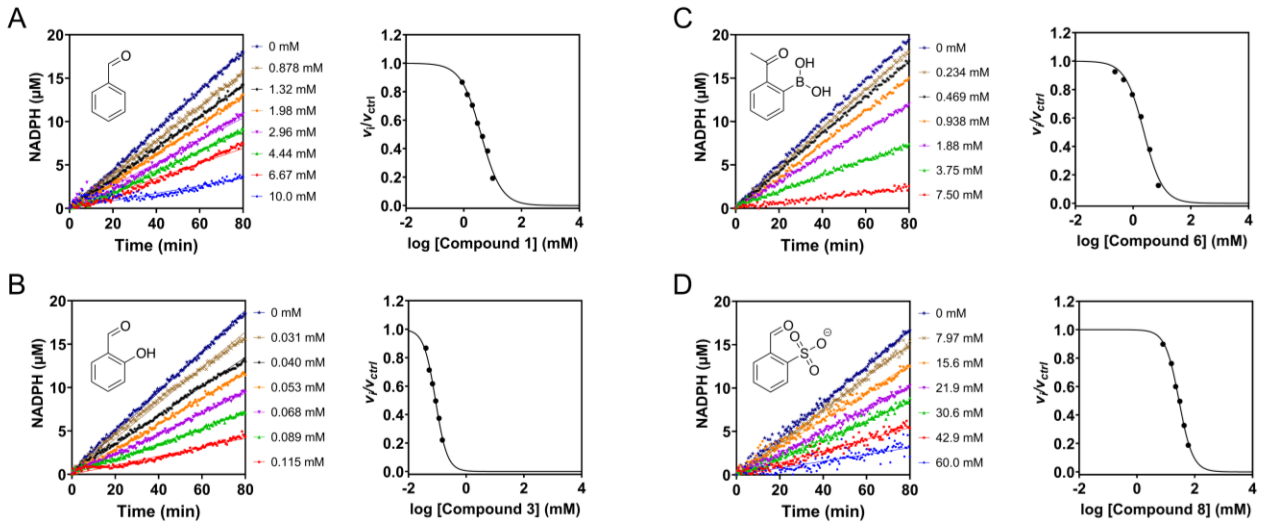


Figure S31. Determination of K_i^{app} for rapid-equilibrium inhibitors of *L. mesenteroides* G6PD. Compounds are (A) benzaldehyde (**1**), (B) salicylaldehyde (**3**), (C) 2-acetylphenylboronic acid (**6**), and (D) 2-formylbenzenesulfonic acid (**8**). Progress curves produced by inhibited and uninhibited reactions were fit by linear regression to obtain v_i and v_{ctrl}, respectively. Normalized velocities (v_i/v_{ctrl}) were plotted against log-transformed inhibitor concentration and then fit to Eqn 6 to derive K_i^{app}. Error bars from propagated uncertainty of v_i and v_{ctrl} SEs are smaller than the data points, derived from nonlinear least-squares regression of one experimental replicate.

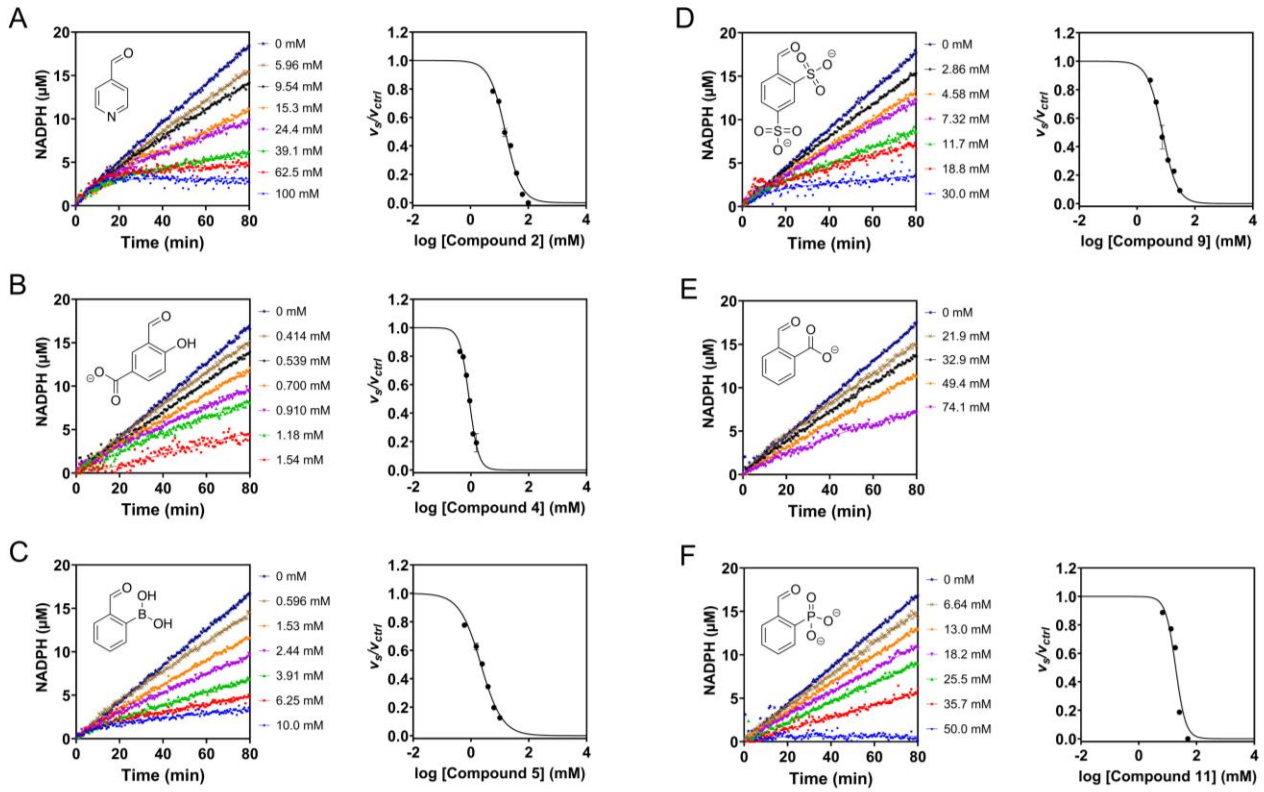


Figure S32. Determination of K_i*^{app} for reversible-covalent inhibitors of *L. mesenteroides* G6PD. Compounds are (A) 4-pyridinecarboxaldehyde (**2**), (B) 3-formyl-4-hydroxybenzoic acid (**4**), (C) 2-formylphenylboronic acid (**5**), (D) benzaldehyde-2,4-disulfonic acid (**9**), (E) 2-carboxybenzaldehyde (**10**), and (F) 2-formylphenylphosphonic acid (**11**). Progress curves produced by inhibited and uninhibited reactions were fit using Eqn 5 and linear regression, respectively, to obtain v_s and v_{ctrl}, respectively. Normalized velocities (v_s/v_{ctrl}) were plotted against log-transformed inhibitor concentration, and then fit to Eqn 6 to derive K_i*^{app}. A dose-response curve and thus K_i*^{app} weren't obtained for 10 as inhibition was limited within the compound's range of solubility. Error bars show propagated uncertainty from v_s and v_{ctrl} SEs, derived from nonlinear least-squares regression of one experimental replicate.

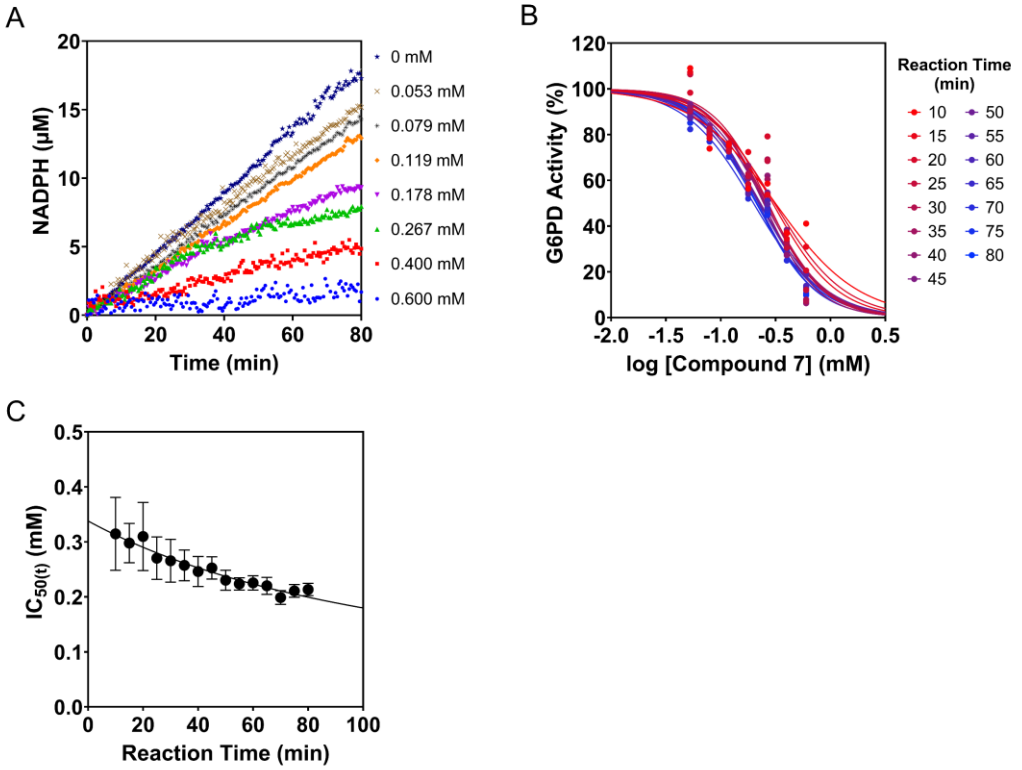


Figure S33. Determination of K_I^{app} and k_2 for irreversible-covalent inhibitor of *L. mesenteroides* G6PD, 2-ethynylbenzaldehyde (7). (A) End-point concentrations of NADPH were sampled from inhibitor progress curves at 15 timepoints. (B) Concentrations were plotted against log-transformed inhibitor concentration and fit to Eqn 8 to obtain time-dependent IC_{50} values ($IC_{50(t)}$). (C) A plot of $IC_{50(t)}$ against time was fit to Eqn 9 to derive K_I^{app} and k_2. Error bars represent $IC_{50(t)}$ SE**,** derived from nonlinear least-squares regression of one experimental replicate.

Table S3. Estimated pK_as of *L. mesenteroides* G6PD lysine residues[ab]

Residue number	pK_a	Residue number	pK_a
5	11.7	**265**	11.3
19	11.0	**273**	12.1
21	10.4	**282**	12.1
31	12.3	**298**	11.5
32	12.6	**311**	11.4
37	11.2	**338**	14.5
55	12.0	**343**	10.3
63	11.2	**352**	11.7
96	12.0	**376**	10.7
105	11.8	**382**	12.1
128	11.8	**386**	10.6
131	11.3	**408**	10.6
148	12.0	**409**	11.4
182	11.9	**441**	11.8
204	11.9	**454**	11.2
208	13.1	**461**	11.5
252	11.3	**472**	11.4
259	11.0	**484**	11.2

[a]pK_as estimated using PypKa (pKAI prediction model).[8,9]

[b]The two lysine residues with the lowest pK_a values are highlighted in green.

REFERENCES

1 J. L. Turnbull, B. R. Benlian, R. P. Golden and E. W. Miller, *J. Am. Chem. Soc.*, 2021, **143**, 6194–6201.
2 K. Qian, S. M. Shepard, T. Xin, G. Park and C. C. Cummins, *J. Am. Chem. Soc.*, 2023, **145**, 6045–6050.
3 A. Krezel and W. Bal, *J. Inorg. Biochem.*, 2004, **98**, 161–166.
4 G. Schenck, K. Baj, J. A. Iggo and M. Wallace, *Anal. Chem.*, 2022, **94**, 8115–8119.
5 I. Neira, A. Blanco-Gómez, J. M. Quintela, C. Peinador and M. D. García, *Org. Lett.*, 2019, **21**, 8976–8980.
6 C. Godoy-Alcántar, A. K. Yatsimirsky and J.-M. Lehn, *J. Phys. Org. Chem.*, 2005, **18**, 979–985.
7 P. M. S. D. Cal, J. B. Vicente, E. Pires, A. V. Coelho, L. F. Veiros, C. Cordeiro and P. M. P. Gois, *J. Am. Chem. Soc.*, 2012, **134**, 10299–10305.
8 P. B. P. S. Reis, D.-A. Clevert and M. Machuqueiro, *Bioinformatics*, 2021, **38**, 297–298.
9 P. B. P. S. Reis, M. Bertolini, F. Montanari, W. Rocchia, M. Machuqueiro and D.-A. Clevert, *J. Chem. Theory Comput.*, 2022, **18**, 5068–5078.