SUPPLEMENTARY INFORMATION

A Generalized Platform for Artificial Intelligence-powered Autonomous Enzyme Engineering

Nilmani Singh[1,2]†, Stephan Lane[1,2,3,]†, Tianhao Yu[1,3,4], Jingxia Lu[1,2,4], Adrianna Ramos[1], Haiyang Cui[3,4], and Huimin Zhao[1,2,3,4]*

[1]Carl R. Woese Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA

[2]DOE Center for Advanced Bioenergy and Bioproducts Innovation, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA

[3]NSF Molecule Maker Lab Institute, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA

[4]Department of Chemical and Biomolecular Engineering, University of Illinois Urbana-Champaign, Urbana, IL, 61801, USA

†These authors contributed equally to this work

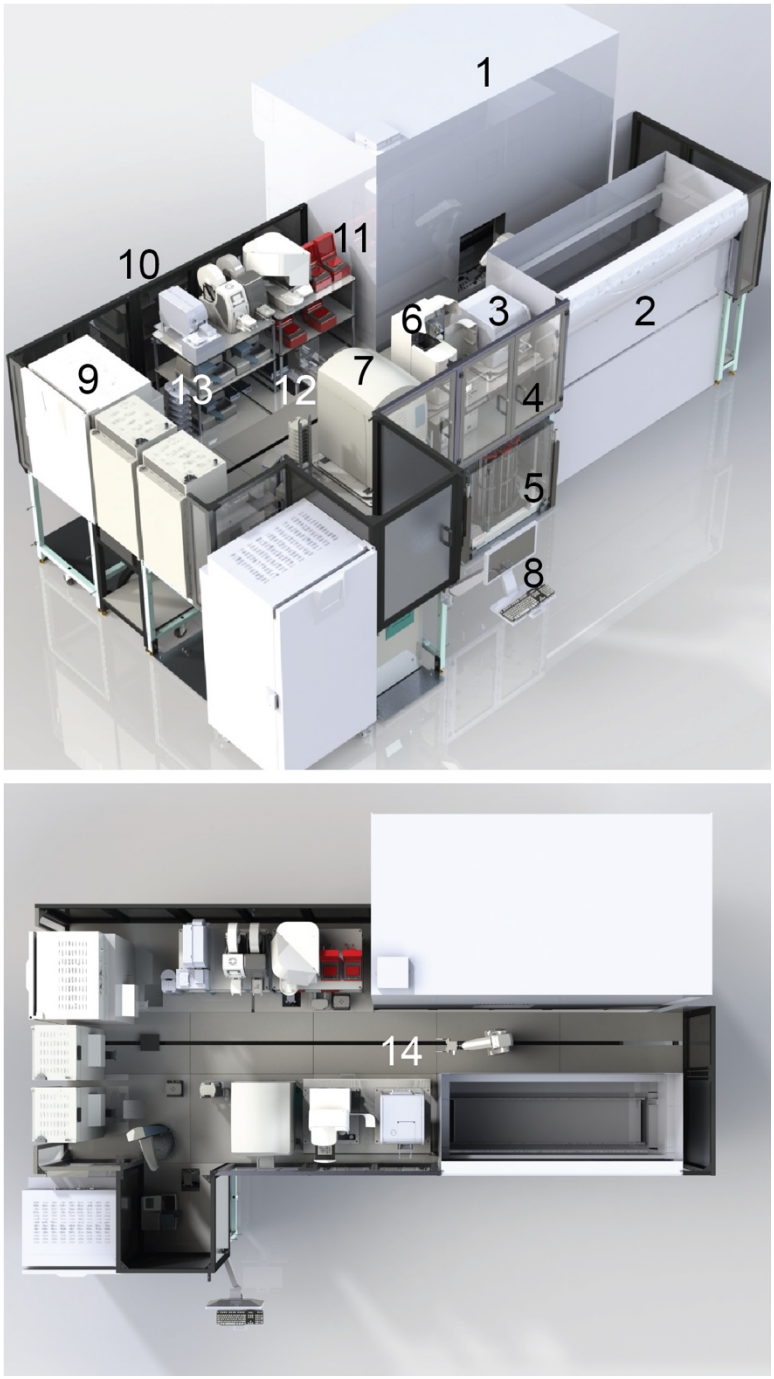
*Corresponding author. Email: zhao5@illinois.edu

Contents

Supplementary Figures 2

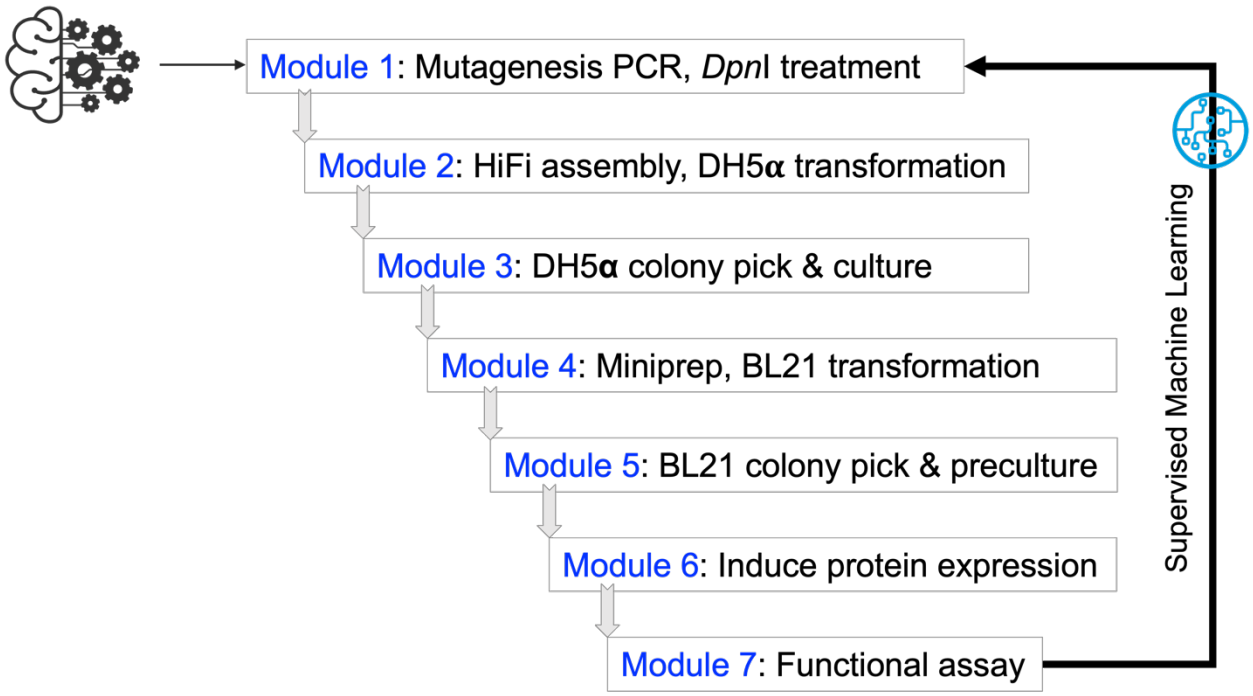
Supplementary Tables..... 29

Supplementary Figures

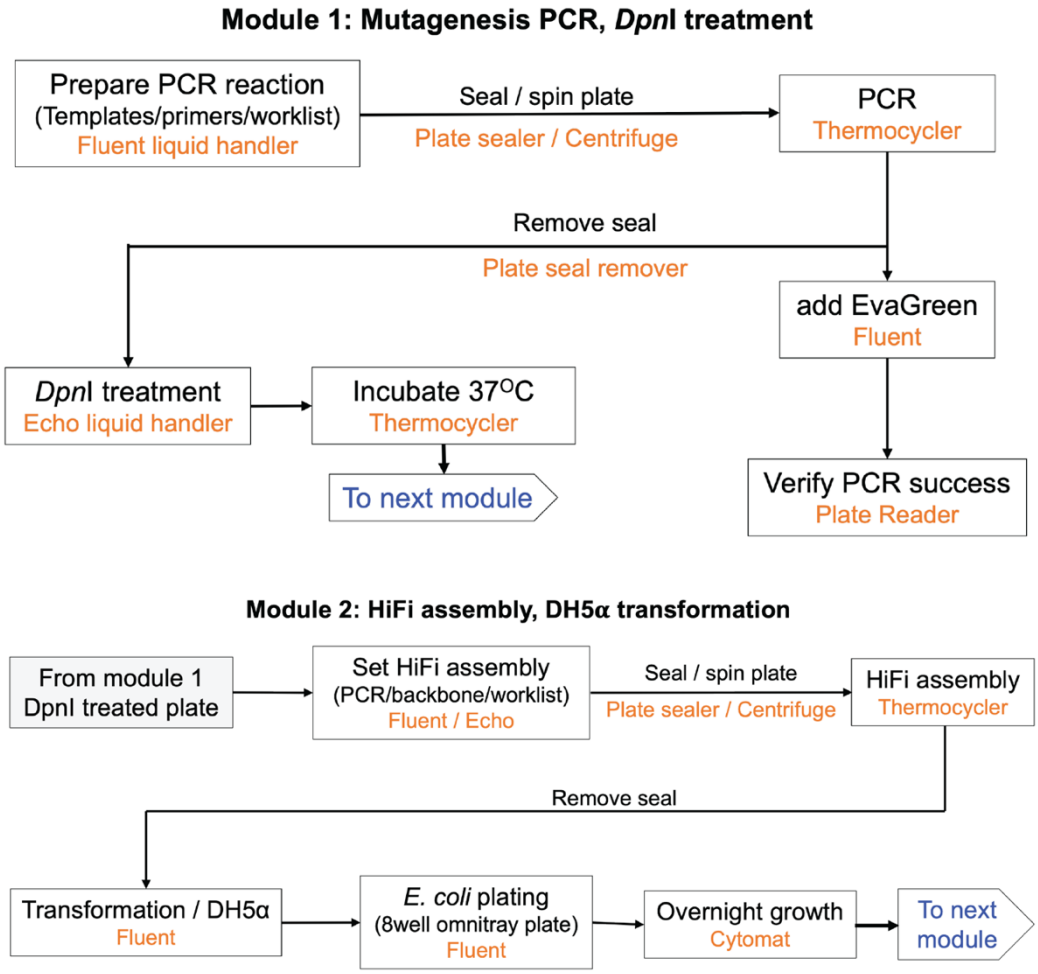


1. Robotic liquid handler I
2. Robotic liquid handler II
3. Cytation 5 imaging platform
4. Tecan infinite plate reader
5. Plate/tips handling hotel
6. Agilent Fragment Analyzer
7. Echo liquid handler
8. Momentum scheduling software
9. Cytomat automated incubators
10. Peripherals (sealers/unsealer)
11. Thermocyclers
12. Agilent automated centrifuge
13. Media dispensers
14. Central robotic arm and track

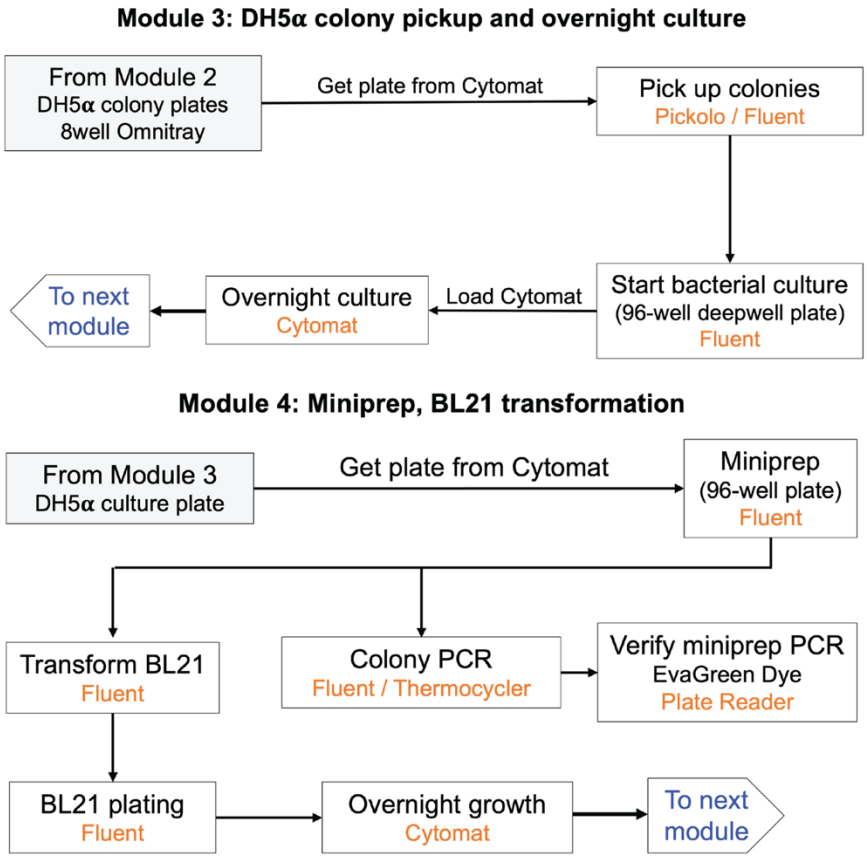
Supplementary Fig. 1 | An overview of the iBioFAB automation platform. Isometric and top-down views of the iBioFAB automation platform are shown. Various core and peripheral instruments are integrated using Thermo Fisher Momentum scheduling software and a Thermo Fisher F5 central robotic arm. A description of the component instruments is listed alongside the image.



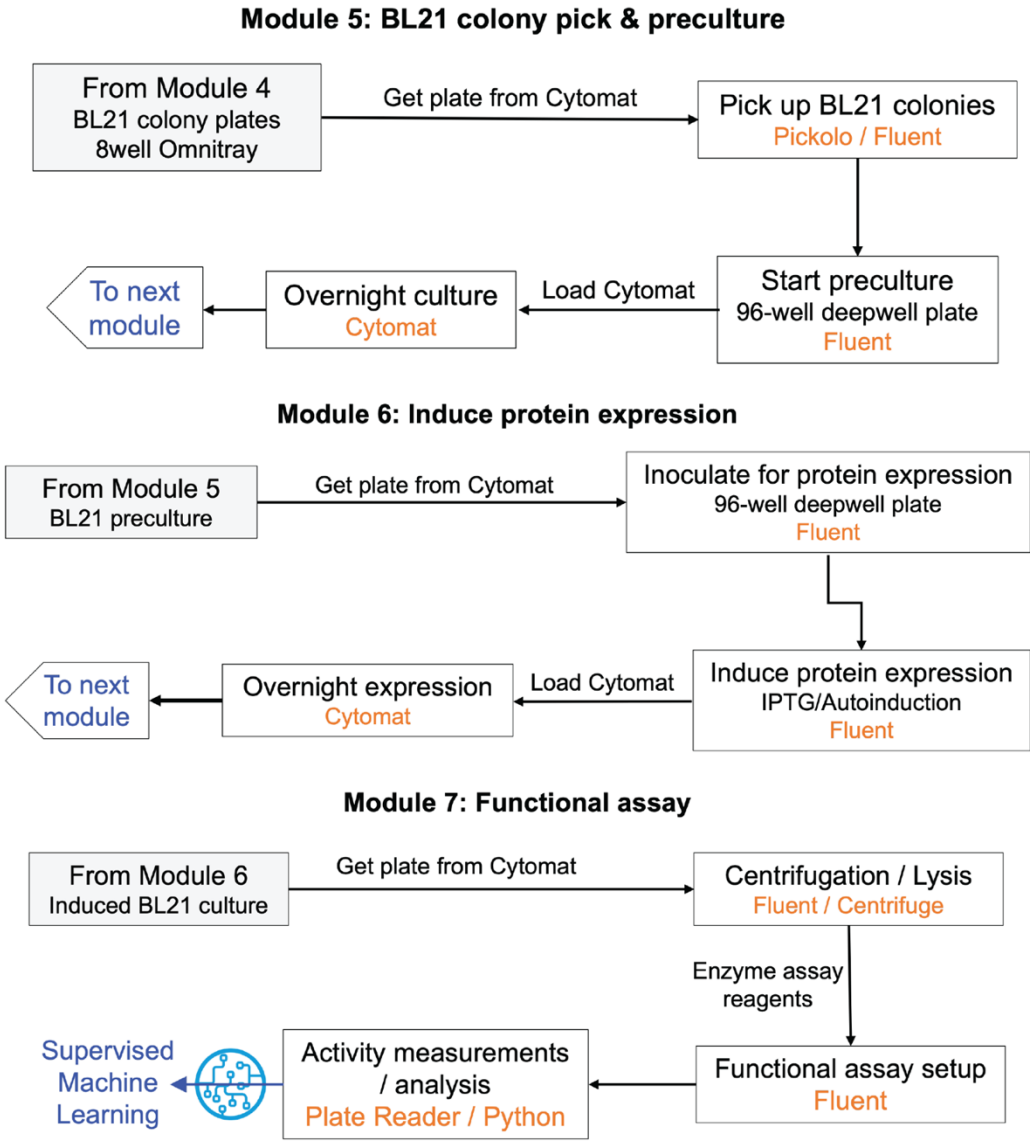
Supplementary Fig. 2 | Overview of the laboratory automation workflow for autonomous protein engineering. The experimental workflow was divided into seven modules, with each designed to be completed in one day. Each module consisted of multiple experiments, coordinated and integrated using Thermo Fisher Momentum scheduling software across one or more instruments on the iBioFAB. The first two rounds of HMT screening served as a test case to design, develop, and troubleshoot the workflow, addressing various experimental and instrumental issues for a seamless, interconnected process. Detailed operations for each module are described in the subsequent Supplementary figures. Supplementary Figure 2 was created with BioRender.



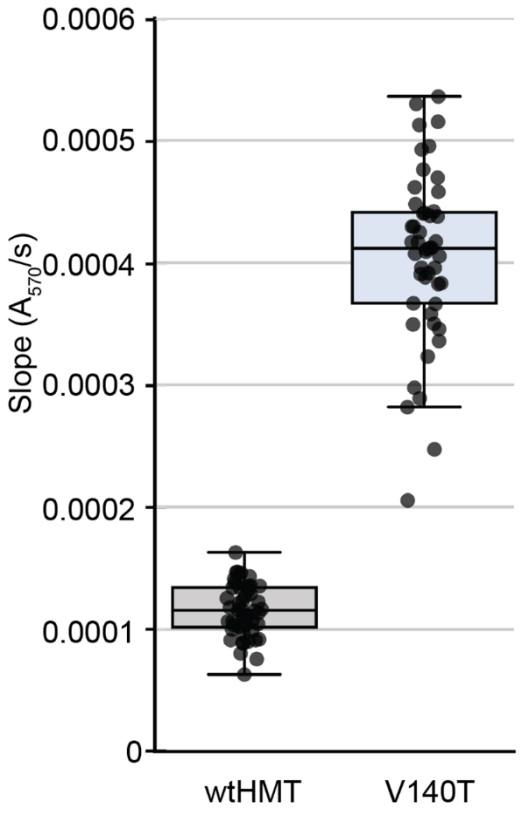
Supplementary Fig. 3 | Detailed overview of modules 1 and 2 for automated protein engineering. Module 1 of the workflow prepared mutagenesis PCR for 96 mutants in a 96-well PCR plate. The templates and primers for PCR were mixed using worklists on the Tecan Fluent and Echo liquid handler. PCR success was measured by adding 2.5 µL of the PCR reaction to 47.5 µL of 1x Evagreen dye (Biotium #31000) and measuring fluorescence (λ_{ex} = 498 nm / λ_{em} = 535 nm) with the Tecan Infinite plate reader. 25 µL PCR product was transferred to a new PCR plate, and 1 µL of *Dpn*I (NEB #R0176) was added, followed by incubation at 37 °C. The *Dpn*I-treated PCR products were then transferred to a 384-well plate, along with the vector backbone, and a worklist guided the mixing of the correct fragments in the Echo liquid handler for HiFi assembly. After a 30-minute HiFi assembly at 50 °C, competent DH5α cells in a 96-well plate were transformed by heat shock on the Tecan Fluent using onboard heating/cooling blocks. The cells were plated on 8-well omnitray agar plates containing LB + 50 µg/mL kanamycin and incubated overnight at 37 °C in a Cytomat automated shaking incubator.



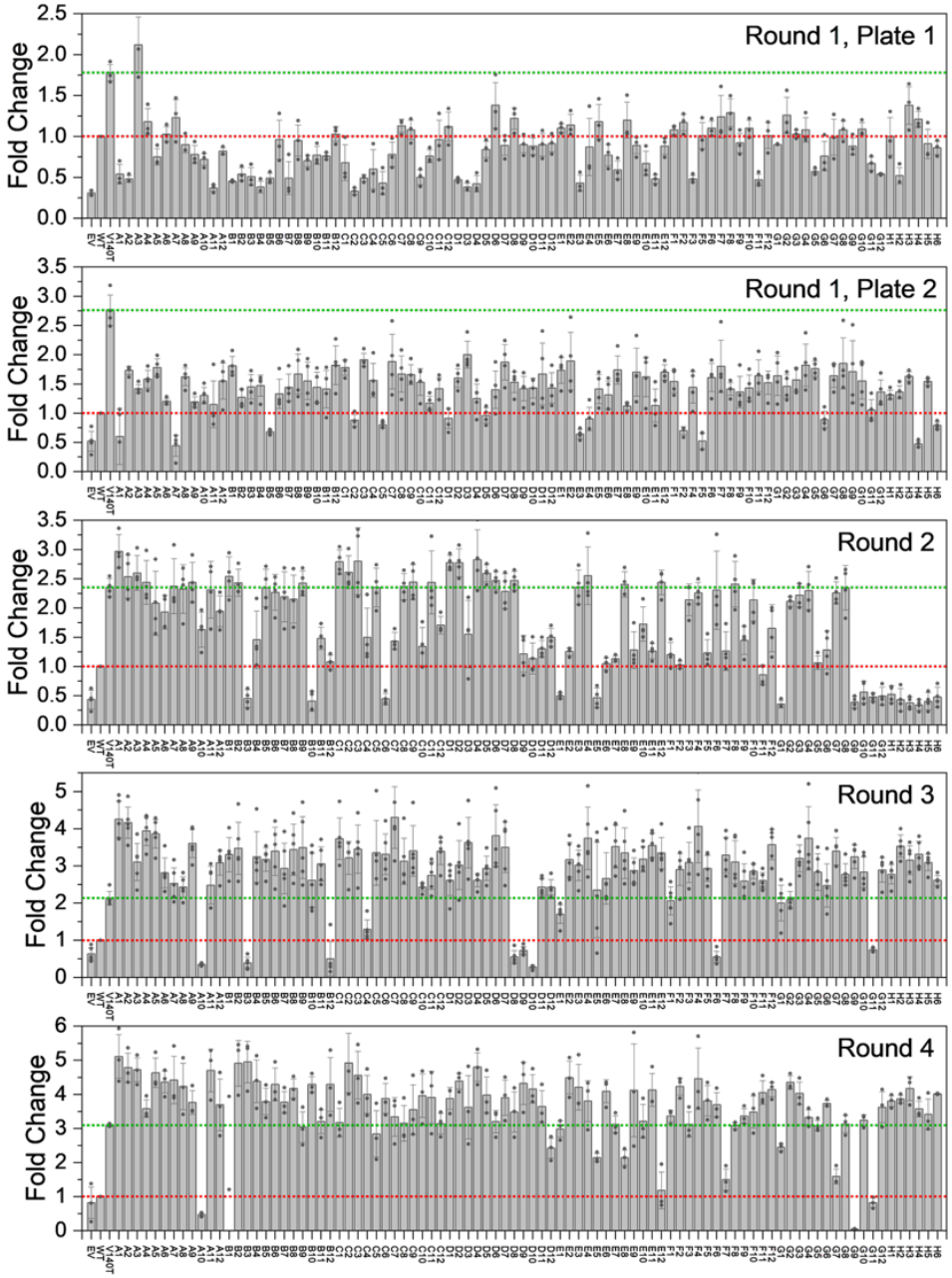
Supplementary Fig. 4 | Detailed overview of modules 3 and 4 of laboratory automation workflow. In Module 3, the Pickolo (SciRobotics) system on the Tecan Fluent was used to pick DH5α colonies from 8-well omnitray plates and inoculate them into 1 mL of TB media (kanamycin 50 µg/mL) in a 96-deepwell plate. The inoculation process maintained the order of mutations in the 96-well plate to match the initial placements in the PCR plate. After overnight culture in an automated Cytomat incubator at 37 °C and 900 rpm, miniprep was performed using the vacuum module on the Tecan Fluent with the PureLink Pro Quick96 Plasmid Purification Kit (Invitrogen #K211004A). 5 µL miniprep DNA was used to transform competent BL21 cells in a 96-well plate by heat shock, using the heating/cooling block onboard the Tecan Fluent. The BL21 cells were plated on 8-well LB/agar omnitray plates (kanamycin 50 µg/mL) and incubated overnight at 37 °C in the Cytomat. We also performed a colony PCR and verified successful PCRs using Evagreen dye to measure the success of HiFi assembly. We consistently achieved a >90% success rate for HiFi assembly. Since we started with 96 mutants and only needed 90 for screening, the <10% failed HiFi reactions did not impact the workflow. The minipreps of some mutants were randomly selected and sent for sequencing. Throughout all evolutionary cycles of both HMT and phytase, we observed 100% correct mutagenesis for the randomly chosen mutants that were sent for sequencing.



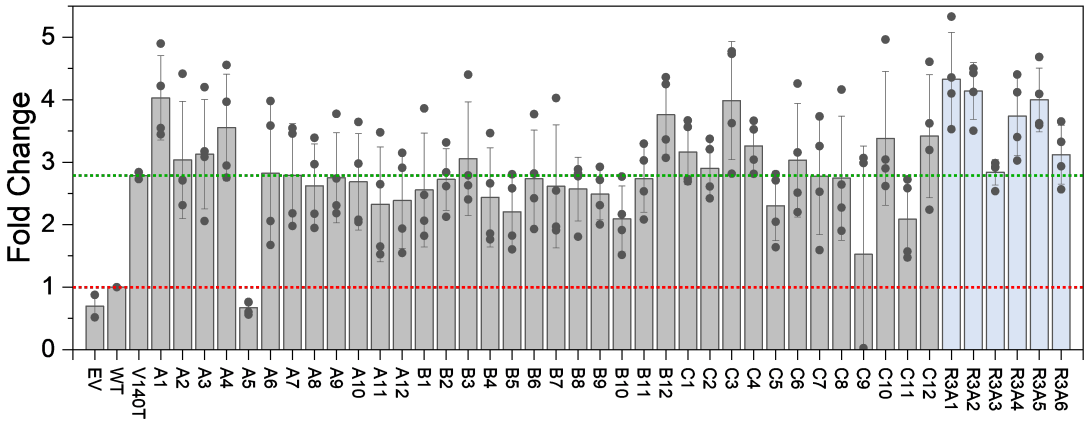
Supplementary Fig. 5 | Detailed overview of modules 5, 6 and 7 of laboratory automation workflow. In Module 5, BL21 colonies were picked using the Pickolo (SciRobotics) system on the Tecan Fluent and inoculated into LB media (kanamycin 50 µg/mL) in a 96-deepwell plate for preculture. Control wells, including wild type, empty vector, and positive control, were also added and inoculated at this stage. The next day, Module 6 inoculated a 96-deepwell plate for protein expression using the preculture, and the plates were loaded into the Cytomat under the appropriate conditions for protein expression. In Module 7, the cells were centrifuged, followed by lysis, a second centrifugation, and the collection of lysates for the activity assay. The protein activity assays were optimized to be compatible with liquid handling systems. Protein activity was measured using the Tecan Fluent plate reader, and the data were analyzed using Python scripts to calculate the relative activity of mutants compared to the wild type.



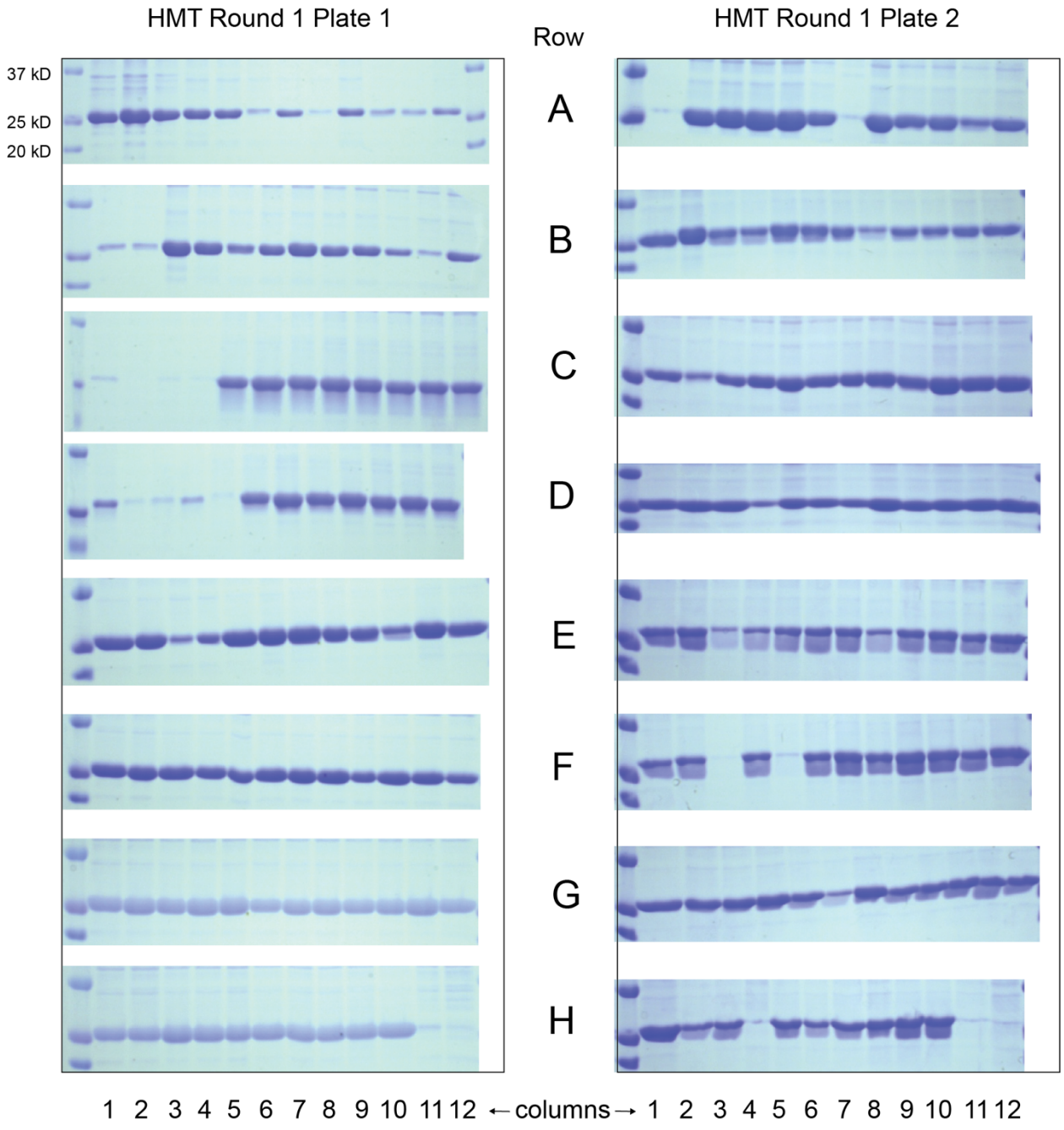
Supplementary Fig. 6 | Variance of iodide detection assay for HMT activity in 96-well plate. In a 96-well plate, wtHMT and the positive control V140T mutant were each expressed in 48 wells. Subsequently, the lysates from the plate were used to detect HMT activity using the iodide detection assay. Absorbance at 570 nm was measured and the rate of increase in A_{570} was calculated. A box plot with overlaying individual values is shown for both the wild type and V140T. For wtHMT [Q1 (25%): 1.02e-04, Median (50%): 1.15e-04, Q3 (75%): 1.34e-04, IQR: 3.26e-05, Whisker Min: 6.32e-05, Whisker Max: 1.63e-04], and V140T [Q1 (25%): 3.67e-04, Median (50%): 4.12e-04, Q3 (75%): 4.41e-04, IQR: 7.43e-05, Whisker Min: 2.82e-04, Whisker Max: 5.36e-04].



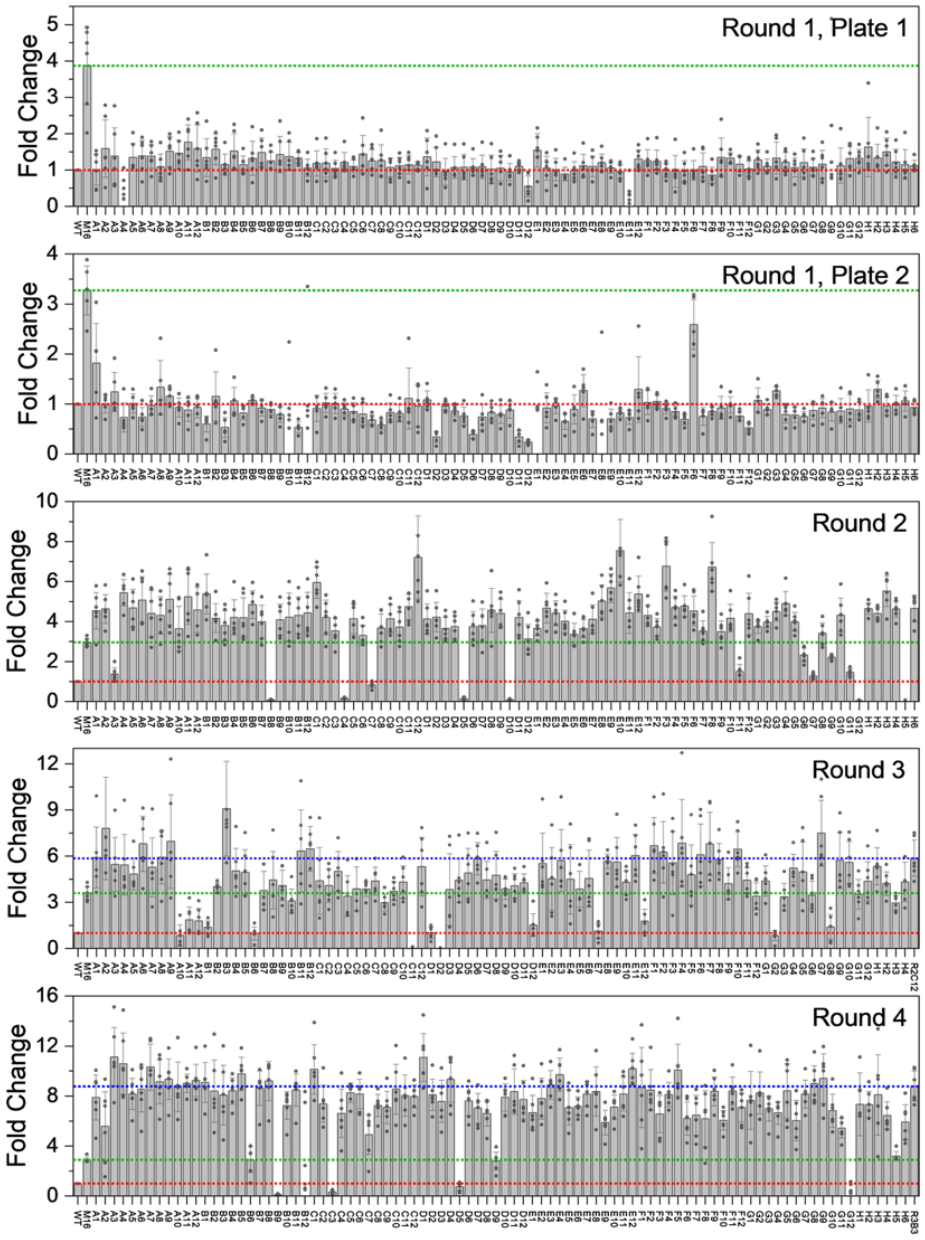
Supplementary Fig. 7 | Crude lysate screening from all four rounds of autonomous protein engineering of *Arabidopsis thaliana* halide methyltransferase (*At*HMT). *At*HMT mutants were screened for activity in 96-well format using the iodide assay, and the fold change compared to the wild type was plotted. Each plate contained ~90 new mutants. Round 1 screening involved two 96-well plates for 180 single mutants, followed by 90 mutants with double, triple, and quadruple mutations in rounds 2, 3, and 4, respectively. Each screening plate included the wild type (red dotted line), the V140T mutant (green dotted line), and an empty vector (EV). The plots represent the average of n ≥ 3 experiments with standard deviations indicated by error bars. Mutants are labeled according to the wells of the 96-well plate, and the names of the mutations are provided in Supplementary Data File 3 along with the original data.



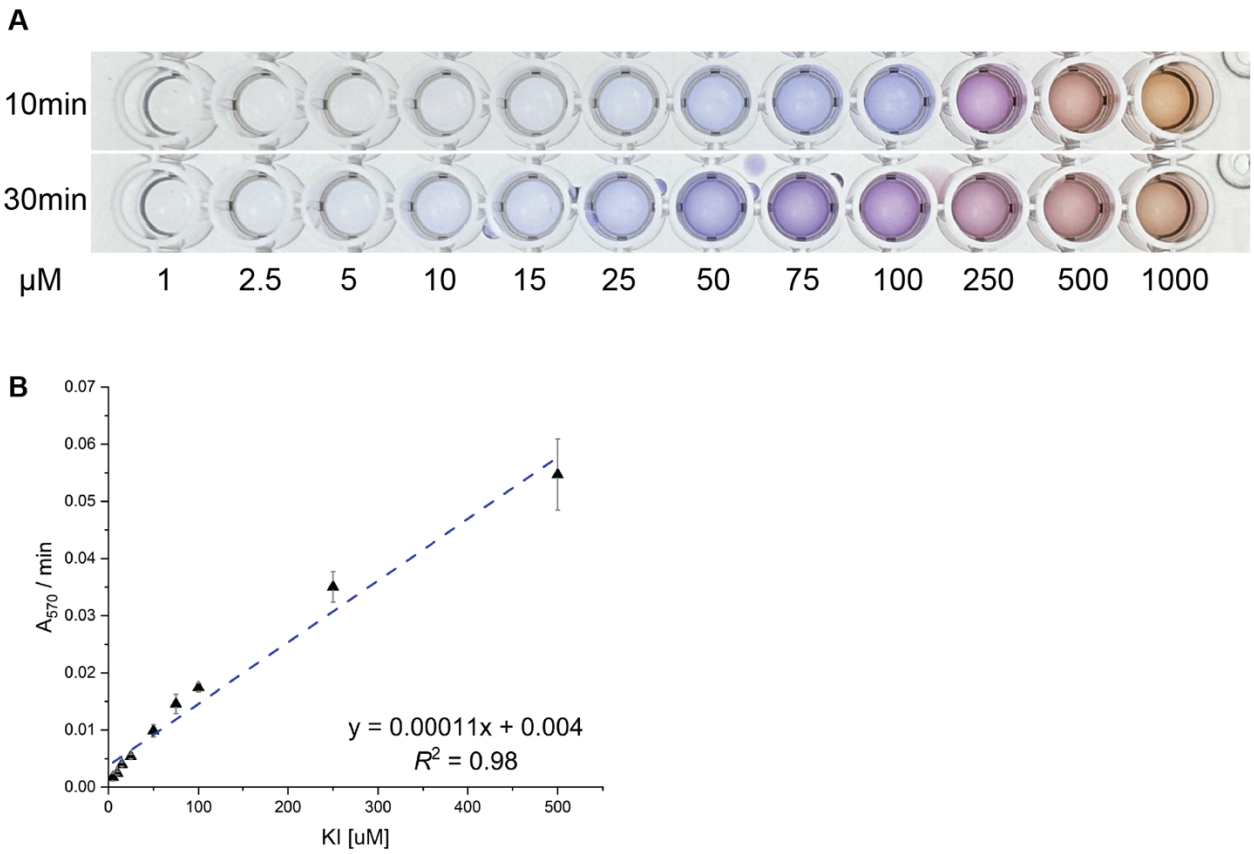
Supplementary Fig. 8 | Comparing human intuition with machine learning predictions for *At*HMT variants. The predicted mutants for third round (triple mutants) did not include mutants with a V140T/S99T combination in the top 500 predictions, despite V140T/S99T being one of the best double mutants from the second evolutionary round. We generated the top 36 triple mutants containing V140T/S99T (grey bars) and screened them for *At*HMT activity, along with the top six predicted triple mutants for the third round (light blue bars). The screening plate also contained the wild type (red dotted line), V140T (green dotted line), and empty vector controls. The fold change compared to the wild type is plotted as average of four independent experiments with standard deviations indicated by error bars. The names of the mutations corresponding to each well are listed in Supplementary Data File 5 along with the original data.



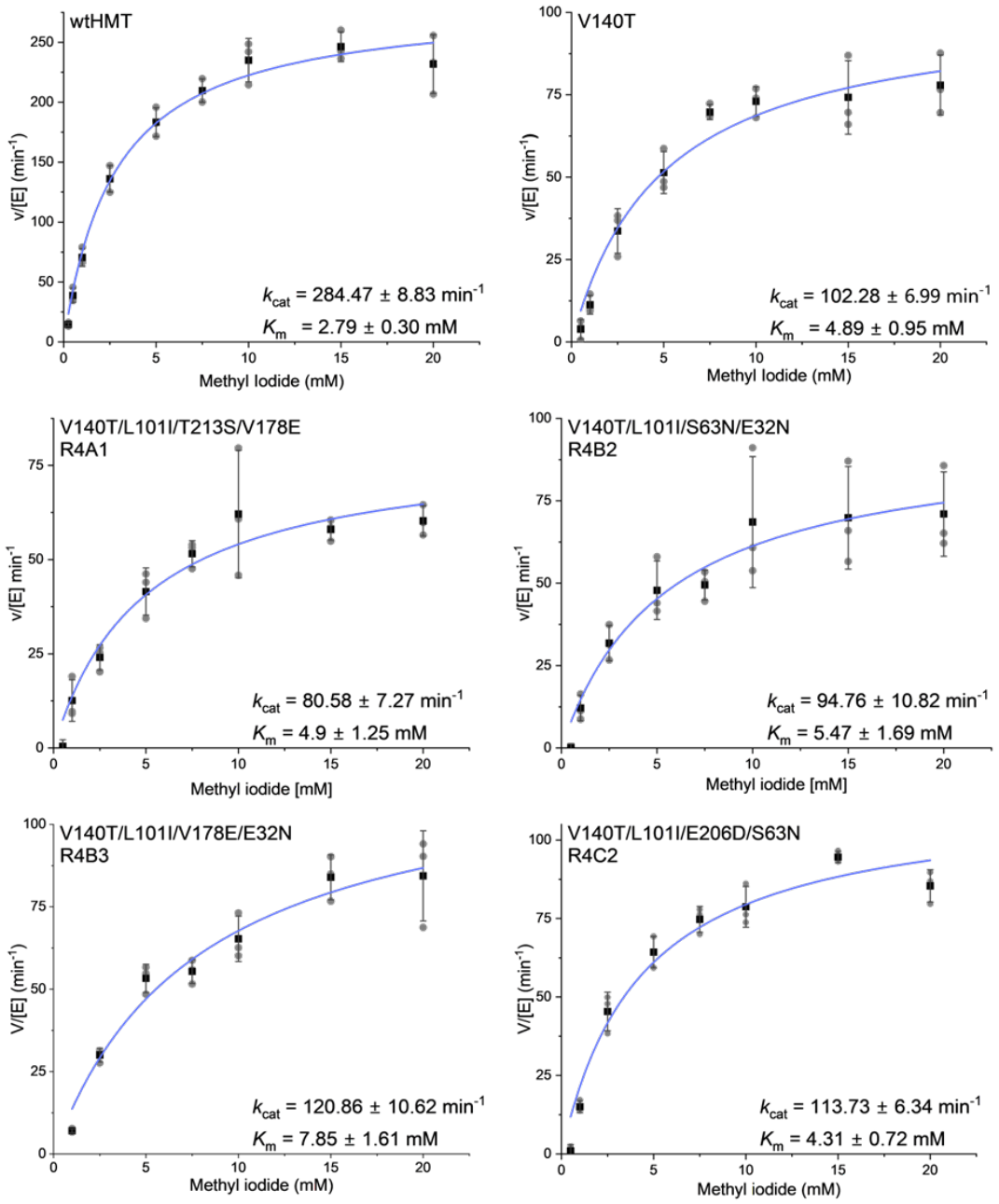
Supplementary Fig. 9 | SDS-PAGE of cell lysates from plates 1 and 2 from the first round of autonomous protein engineering of *At*HMT. Equal amounts of lysates from three replicate plates were combined and mixed with 2x Laemmli buffer (+5% BME) and boiled for five minutes at 95 °C. Subsequently, 15 µL of each sample was run on an 10% SDS-PAGE gel and stained with Coomassie Blue (Bio-Rad #1610436). Images were acquired after thoroughly washing away the stain. The newly generated mutants were loaded in wells A1 to H6, wtHMT in H7 and H8, V140T in H9 and H10, and the empty vector in H11 and H12. The names of the mutations are available in Supplementary Data File 3.



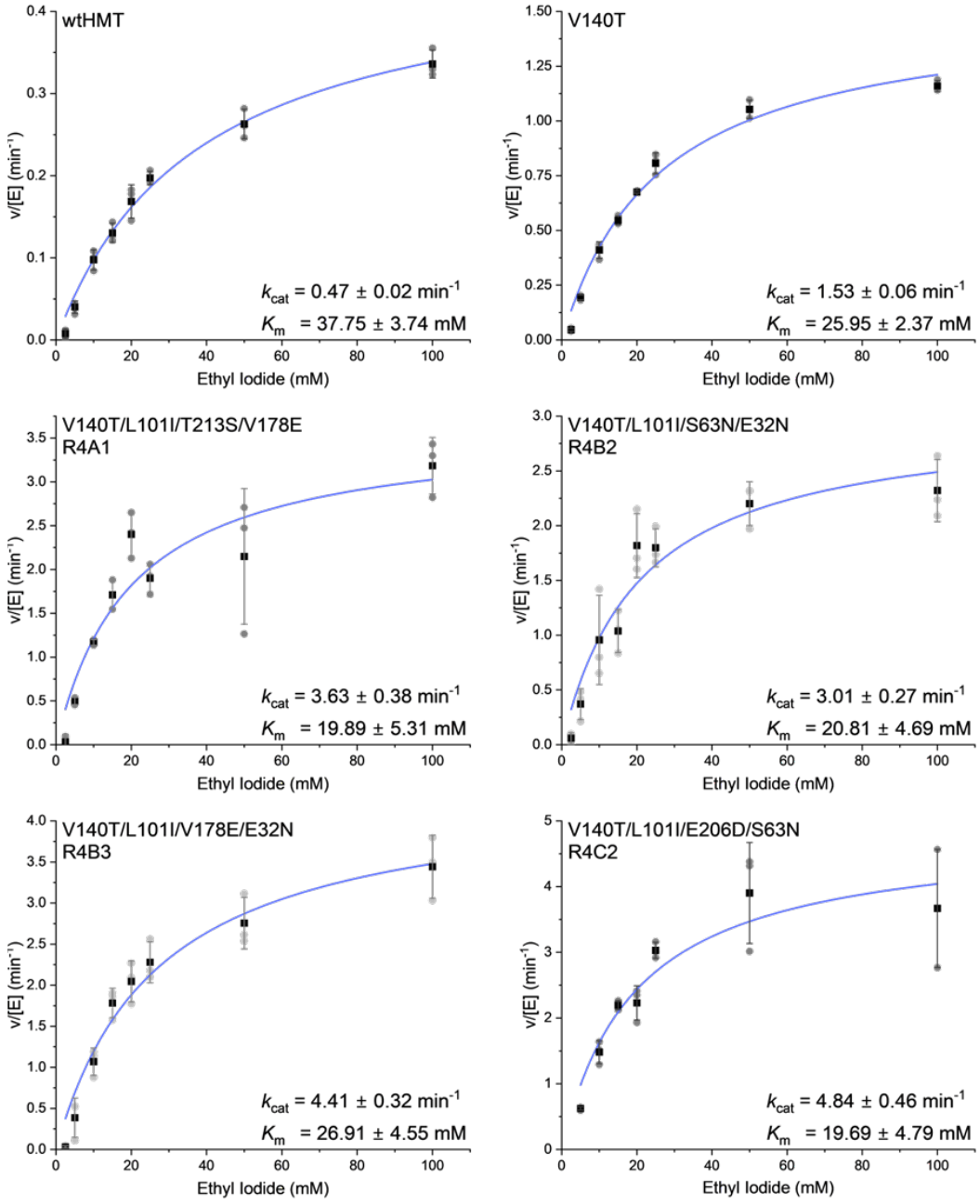
Supplementary Fig. 10 | Crude lysate screening from all four rounds of autonomous protein engineering of *Yersinia mollaretii* phytase (*Ym*Phytase). *Ym*Phytase mutants were screened in 96-well format using the 4-MUP assay, and the fold change compared to the wild type was plotted. Each 96-well plate contained 90 new mutants and six controls, with the first round of evolution involving two 96-well plates for 180 single mutants and 12 controls, followed by approximately 90 mutants with double, triple, and quadruple mutations in the second, third, and fourth rounds, respectively. Each screening plate included the wild type (red dotted line), a previously reported M16 variant containing a T44V/K45E double mutation *(7)* (green dotted line), and an empty vector (EV). To calculate the slope, the fluorescence background for EV was subtracted at each time point. For the third and fourth rounds, one of the best mutants from the previous round was included (blue dotted line). The plots represent the average of $n \geq 3$ experiments with standard deviation. Mutants are labeled according to the wells of the 96-well plate, and the names of the mutations are listed in Supplementary Data File 4 along with the original data.



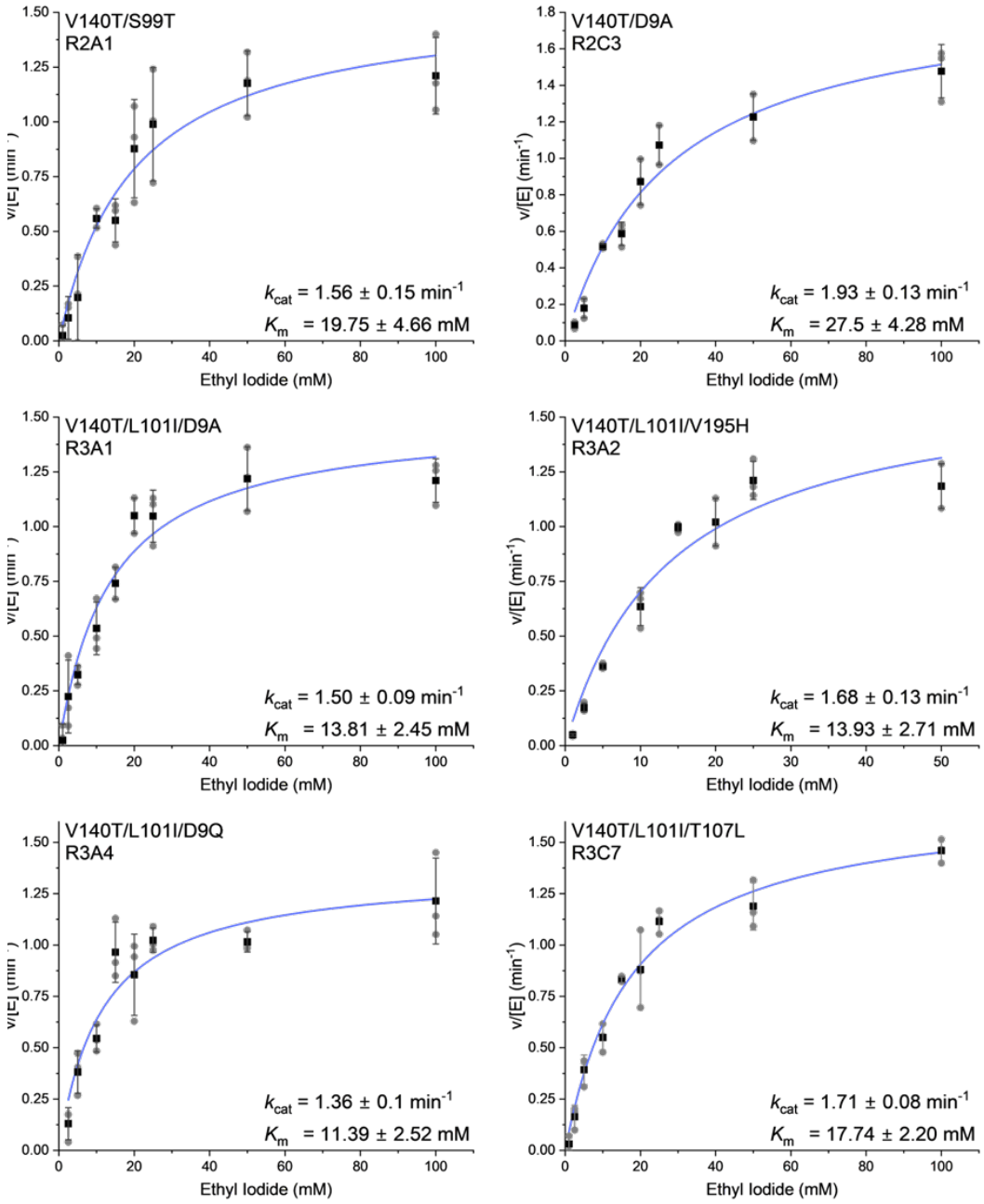
Supplementary Fig. 11 | Iodide detection assay and the standard curve for iodide detection. (**A**) Image of color formation after adding different concentrations of potassium iodide (KI) to the iodide detection reagent at 10 minutes and 30 minutes (**B**) Standard curve for iodide detection assay. In a 50 µL reaction, 2.5 µL of potassium iodide (KI) [5 µM to 500 µM] was added to the iodide detection reagent containing TMB (47 µL) and CiVCPO (0.5 µL / 0.35 mg). The initial increase in Absorbance at 570 nm was measured and slope per minute was plotted against KI concentration (n=4).



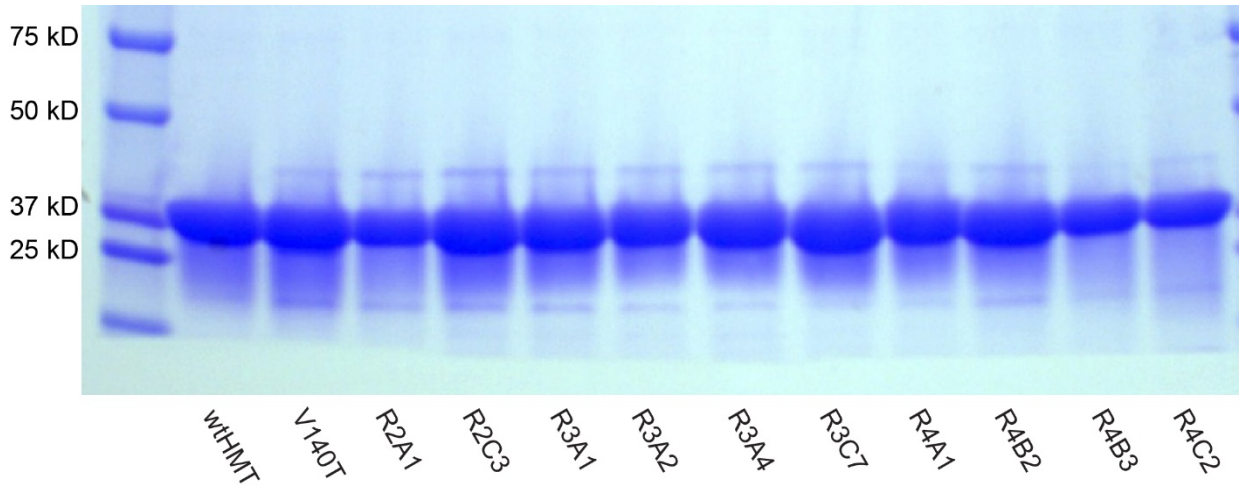
Supplementary Fig. 12 | Michaelis-Menten kinetics of wild type *At*HMT and its round 4 variants using methyl iodide. The methyltransferase activity for purified variants of *At*HMT was determined using different methyl iodide concentrations. wtHMT was added at 200 ng (0.145 μM) per reaction while the round 4 variants were added at 400 ng (0.29 μM) per reaction. The curves were fit using Michaelis Menten fit on OriginPro 2023. The values (circle), mean (square), fitted plot (blue), and standard deviation are plotted (n=3). The name of the purified protein is labeled on each graph. The notation R4C2 represents well C2 of the fourth round of autonomous protein engineering.



Supplementary Fig. 13 | Michaelis-Menten kinetics of wild type *At*HMT and its round 4 variants using ethyl iodide. The ethyltransferase activity for purified variants of *At*HMT was determined using different ethyl iodide concentrations. wtHMT was added at 87 µM (2.4 mg/mL), V140T at 29 µM, and round 4 variants at 8.7 µM per reaction. The curves were fit using Michaelis Menten fit on OriginPro 2023. The values (circle), mean (square), fitted plot (blue), and standard deviation are plotted (n=3). The name of the purified protein is labeled on each graph. The notation R4C2 represents well C2 of the fourth round of autonomous protein engineering.

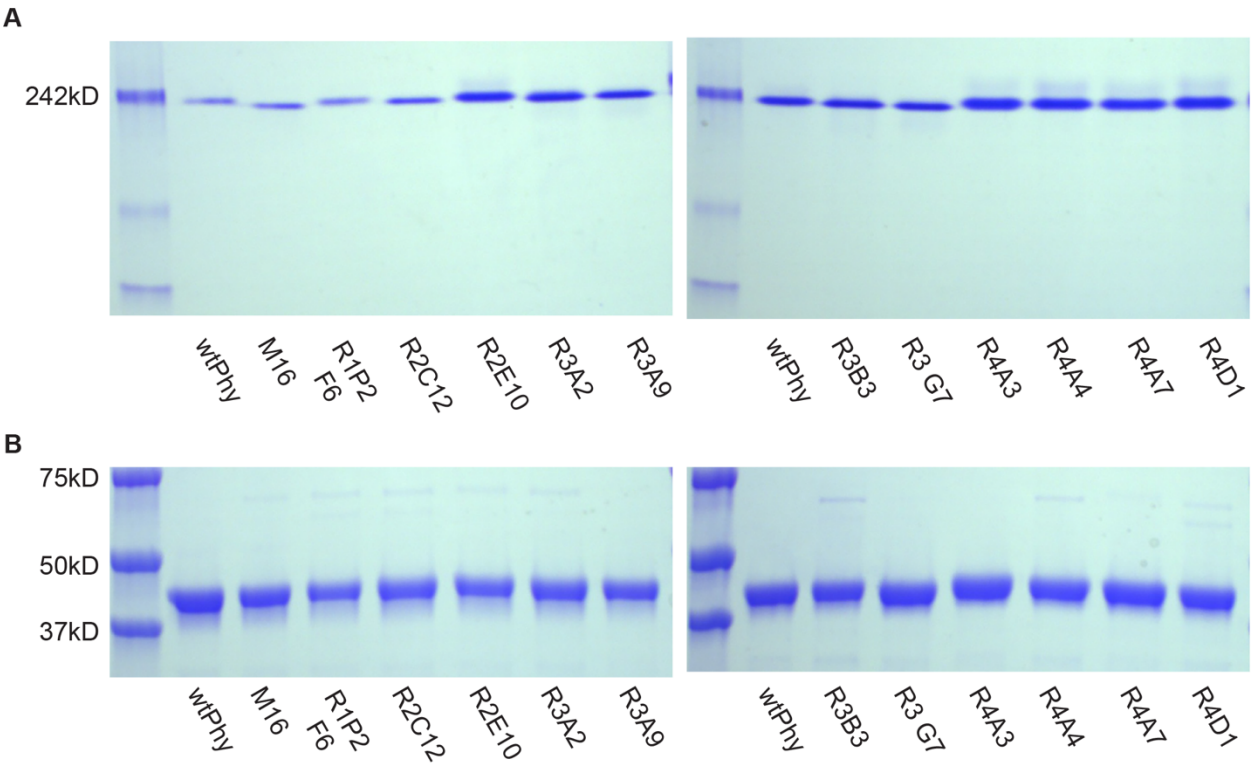


Supplementary Fig. 14 | Michaelis-Menten kinetics of wild type *At*HMT and its round 2 and round 3 variants using ethyl iodide. The ethyltransferase activity for purified variants of *At*HMT was determined using different ethyl iodide concentrations. Second and third round variants were added at final concentrations of 29 μM per reaction. The curves were fit using Michaelis Menten fit on OriginPro 2023. The values (circle), mean (square), fitted plot (blue), and standard deviation are plotted (n=3). The name of the purified protein is labeled on each graph. The notation R3C7 represents well C7 of the third round of autonomous protein engineering.



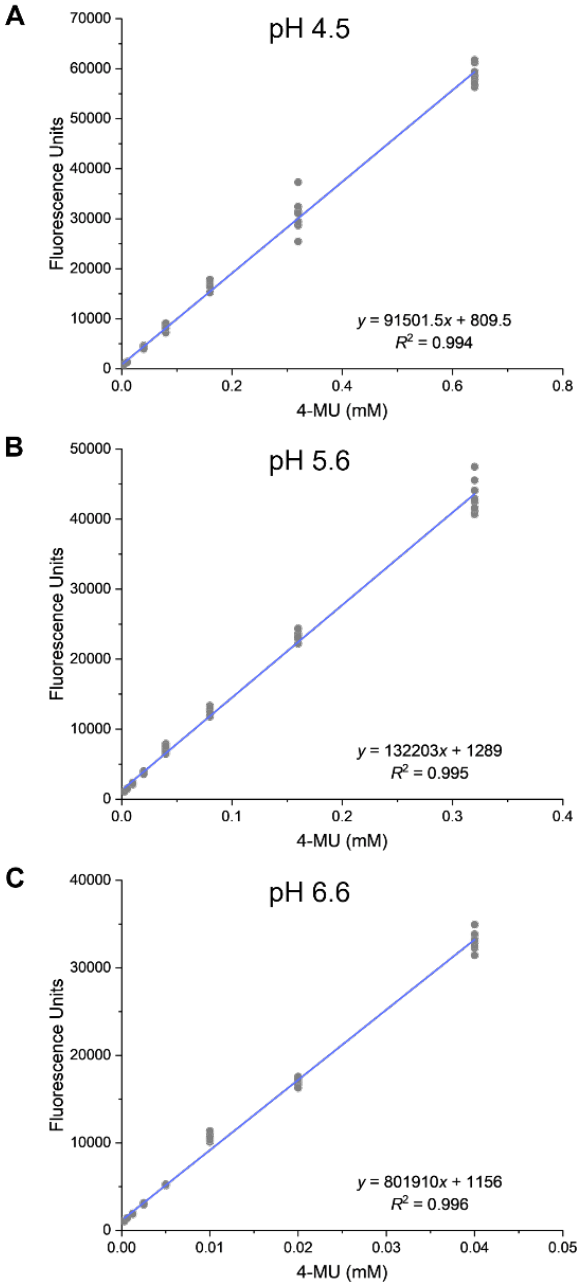
Supplementary Fig. 15 | SDS PAGE of purified *At*HMT proteins. The purified proteins were mixed with 1x Laemmli buffer (+5% BME) and boiled at 95 °C for five minutes. An equal amount of each sample was loaded on a 12% SDS-PAGE gel along with a protein ladder (BioRad # 1610374). After Coomassie staining, the gels were thoroughly washed before imaging.

Enzyme	Description
wtHMT	Wild type HMT from *Arabidopsis thaliana*
V140T	Positive control
R2 A1	V140T / S99T
R2 C3	V140T / D9A
R3 A1	V140T / L101I / D9A
R3 A2	V140T / L101I / V195H
R3 A4	V140T / L101I / D9Q
R3 C7	V140T / L101I / T107L
R4 A1	V140T / L101I / T213S / V178E
R4 B2	V140T / L101I / S63N / E32N
R4 B3	V140T / L101I / V178E / E32N
R4 C2	V140T / L101I / E206D / S63N

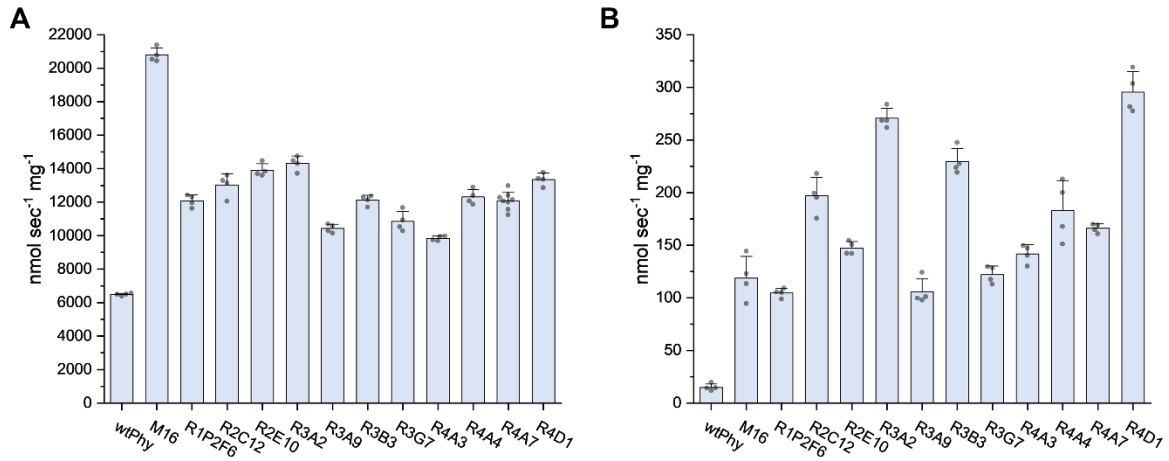


Supplementary Fig. 16 | Native and SDS PAGE gel of purified *Ym*Phytase proteins. (**A**) For native PAGE, purified protein were mixed in native sample buffer and run in Tris/Glycine buffer at 7.5% TGX gel (BioRad #4568024) along with native protein ladder (Invitrogen # LC0725). As reported previously, the native gel confirmed that the phytase protein forms a tetramer. (**B**) For SDS-PAGE, purified proteins were mixed with 1x Laemmli buffer (+5% BME) and boiled at 95°C for five minutes. An equal amount of each sample was then loaded onto a 12% SDS-PAGE gel along with a protein ladder (BioRad # 1610374). After Coomassie staining, the gels were thoroughly washed before imaging.

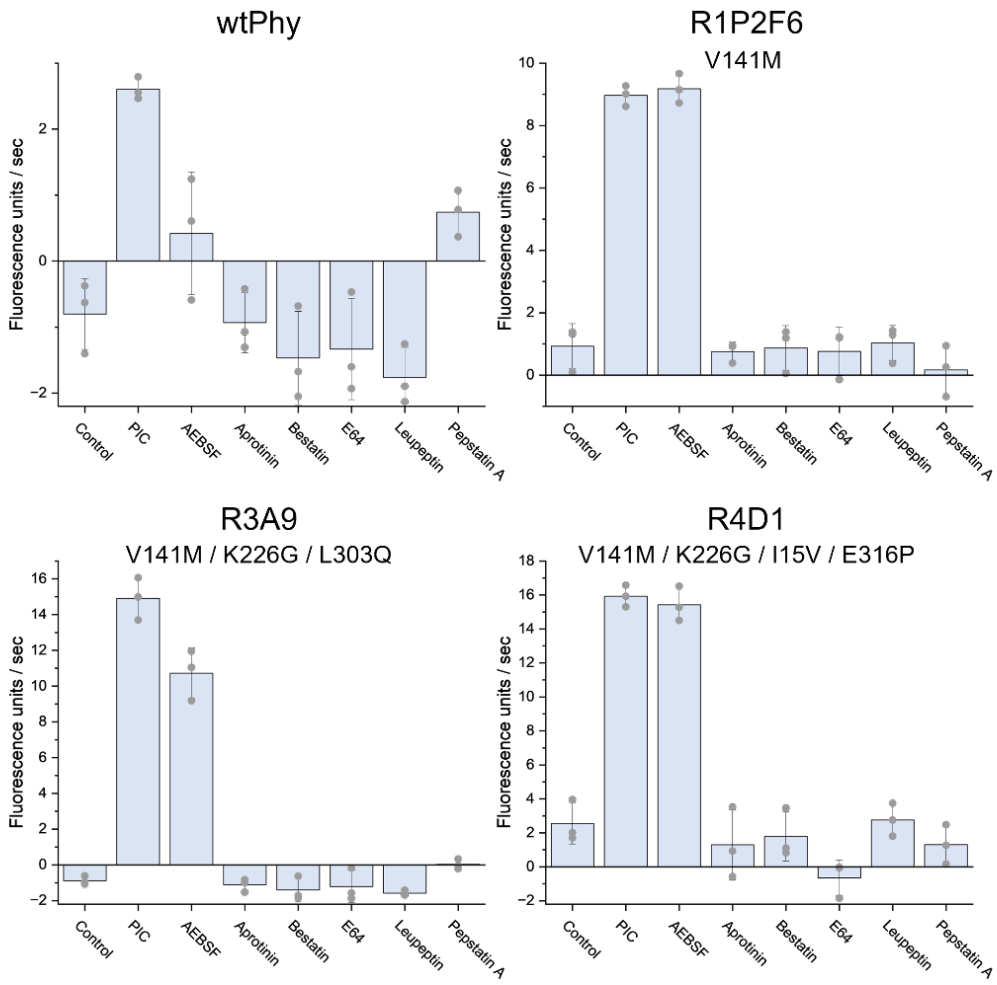
Enzyme	Description
wtPhy	Wild type *Yersinia mollaretii* phytase
M16	T44V / K45E
R1 P2 F6	V141M
R2 C12	V141M / K226G
R2 E10	V141M / E316D
R3 A2	V141M / K226G / T25S
R3 A9	V141M / K226G / L303Q
R3 B3	V141M / K226G / I15V
R3 G7	V141M / K226G / L303P
R4 A3	V141M / K226G / I15V / Q167R
R4 A4	V141M / K226G / I15V / Q362R
R4 A7	V141M / K226G / I15V / V266R
R4 D1	V141M / K226G / I15V / E316P



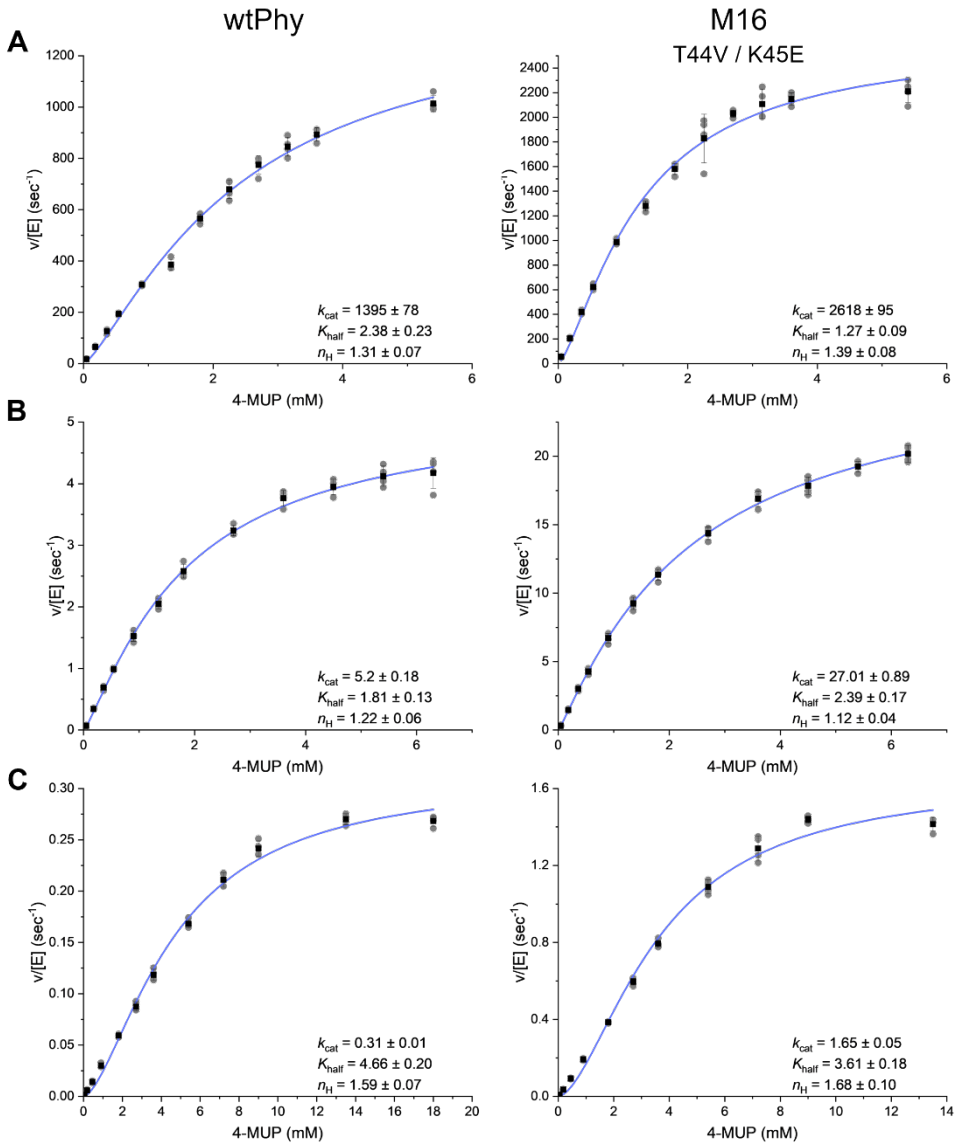
Supplementary Fig. 17 | Standard curves for kinetic analysis of *Ym*Phytase at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Fluorescence values (λ_{ex} = 354 nm / λ_{em} = 465 nm) were correlated with 4-methylumbelliferone (4-MU) concentrations through making a standard curve. For pH 4.5, standards are comprised of eight different concentrations of 4-MU with eight independently prepared samples each, which are represented by grey dots. For pH 5.6 and 6.6, eight different concentrations of 4-MU were prepared in 12 independently prepared samples, which are represented by grey dots. 4-MU was dissolved into methanol and 50 µL of dissolved 4-MU solution was mixed with 50 µL of appropriate pH buffer. For pH 4.5, the buffer used was 0.25 M sodium acetate, 1 mM calcium chloride, 0.01% Tween-20 buffer. For pH 5.6 and 6.6, 0.2 M tris maleate buffer at the appropriate pH was used.



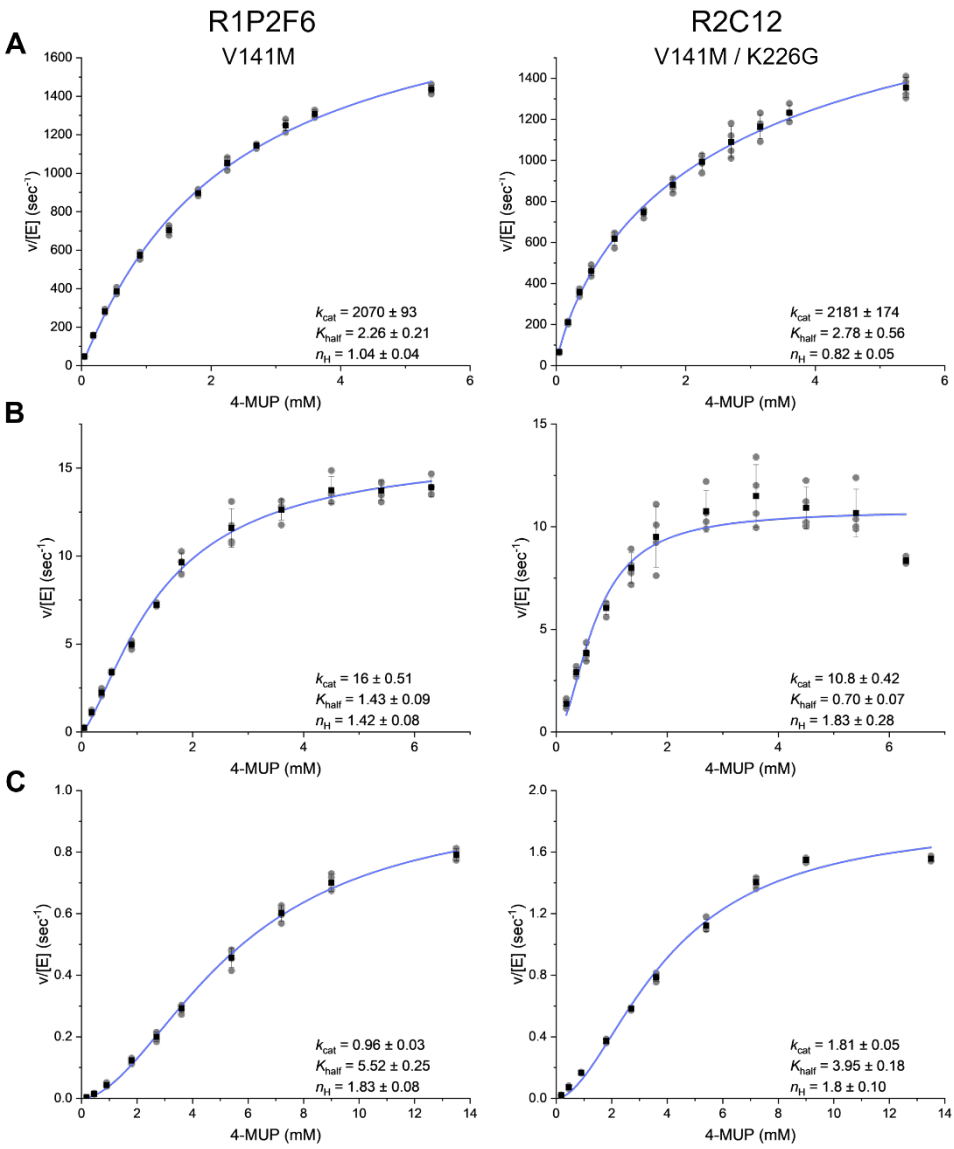
Supplementary Fig. 18 | Specific enzyme activity of purified *Ym*Phytase variants at A) pH 4.5 and B) pH 5.6. Purified *Ym*Phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to assays. Specific activity assays were performed with 0.9 mM 4-MUP. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6, assays were performed in 0.2 M tris maleate buffer. Bars indicate the average of four independent samples with standard deviations signified by error bars and individual data points as grey dots.



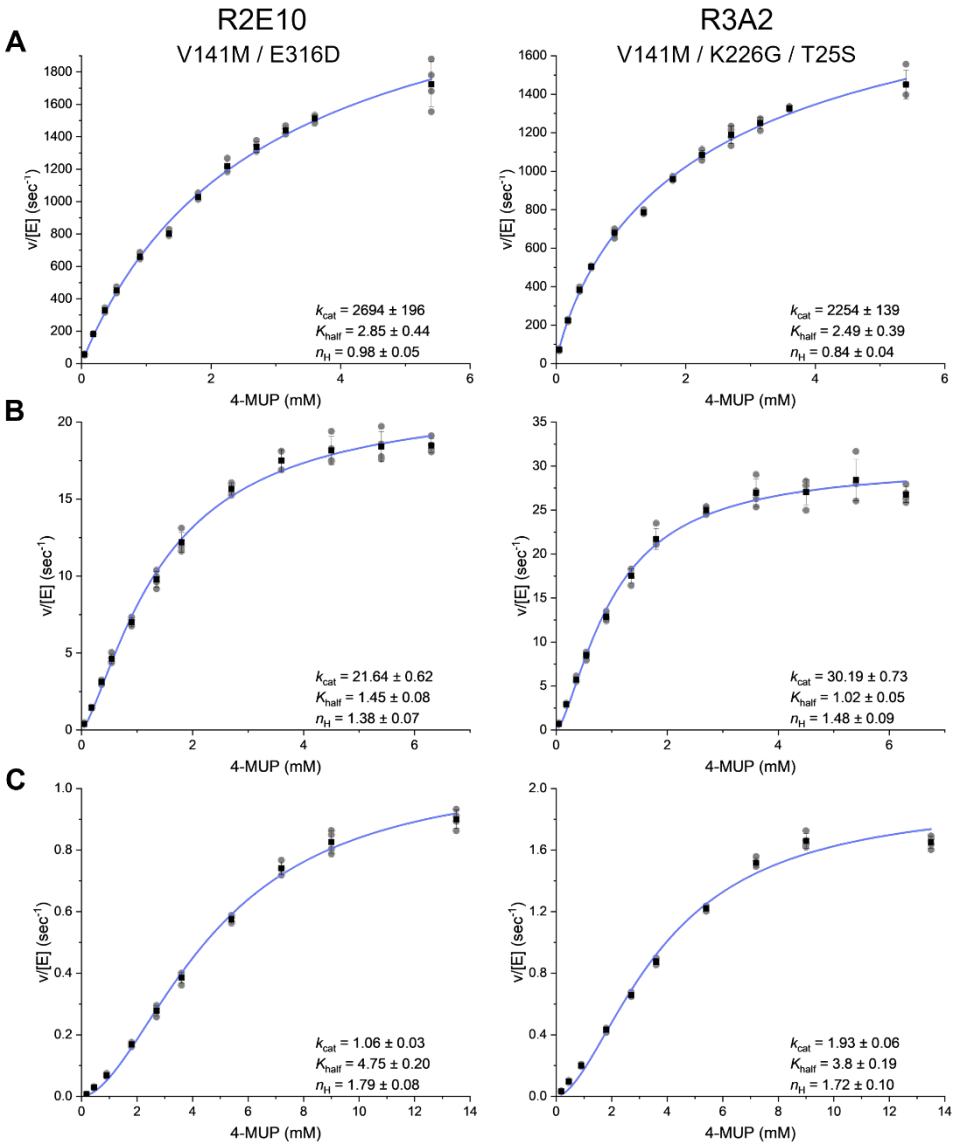
Supplementary Fig. 19 | Effect of protease inhibitor additives to relative activity of purified phytase enzymes. Purified enzymes were mixed with Protease Inhibitor Cocktail (PIC, Sigma #P8849), 4-(2-aminoethyl)benzenesulfonyl fluoride hydrochloride (AEBSF), aprotinin, bestatin, protease inhibitor E64, leupeptin, or pepstatin A and incubated for 30 minutes at 37 °C, then used in a 4-MUP activity assay in 0.2 M tris maleate buffer at pH 6.6. Bars represent the average of three independent samples which are shown as grey dots.



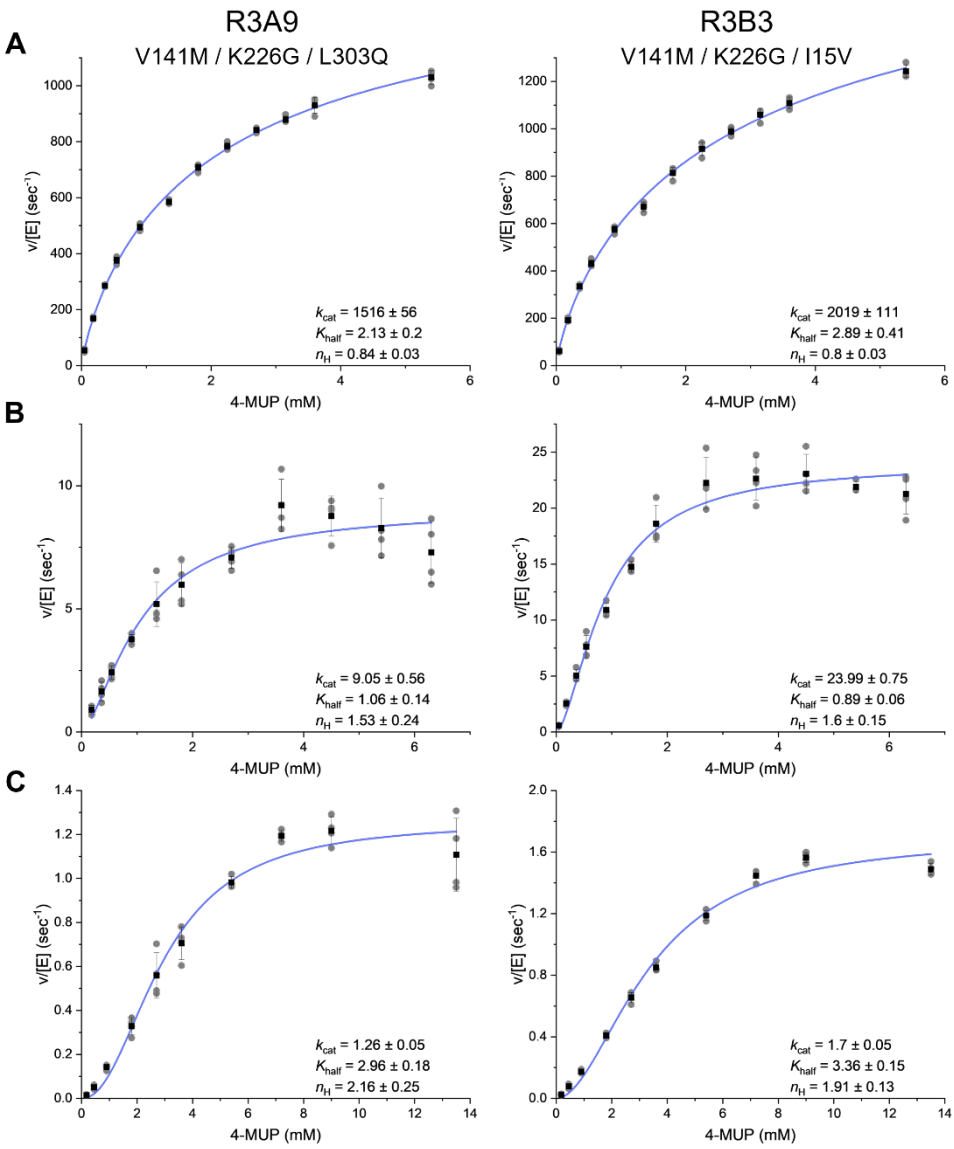
Supplementary Fig. 20 | Kinetic analysis of wild type *Ym*Phytase (wtPhy, left) and positive control (M16, right) at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Purified phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to kinetic assays. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6 and 6.6, assays were performed in 0.2 M tris maleate buffer. Black dots indicate the average of four independent samples with standard deviations signified by error bars and individual datapoints by grey dots.



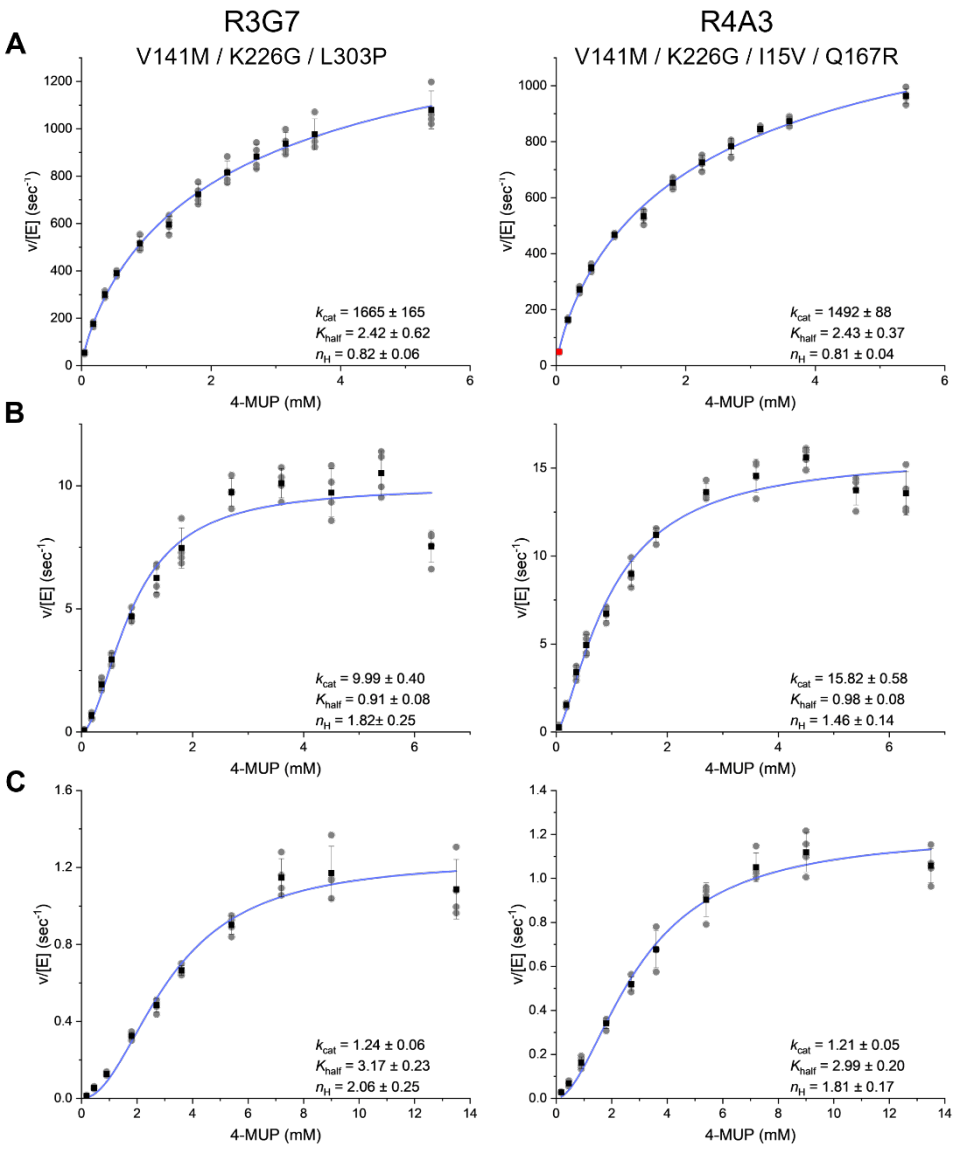
Supplementary Fig. 21 | Kinetic analysis of *Ym*Phytase variants R1P2F6 (left) and R2C12 (right) at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Purified *Ym*Phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to kinetic assays. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6 and 6.6, assays were performed in 0.2 M Tris maleate buffer. Black dots indicate the average of four independent samples with standard deviations signified by error bars and individual datapoints by grey dots.



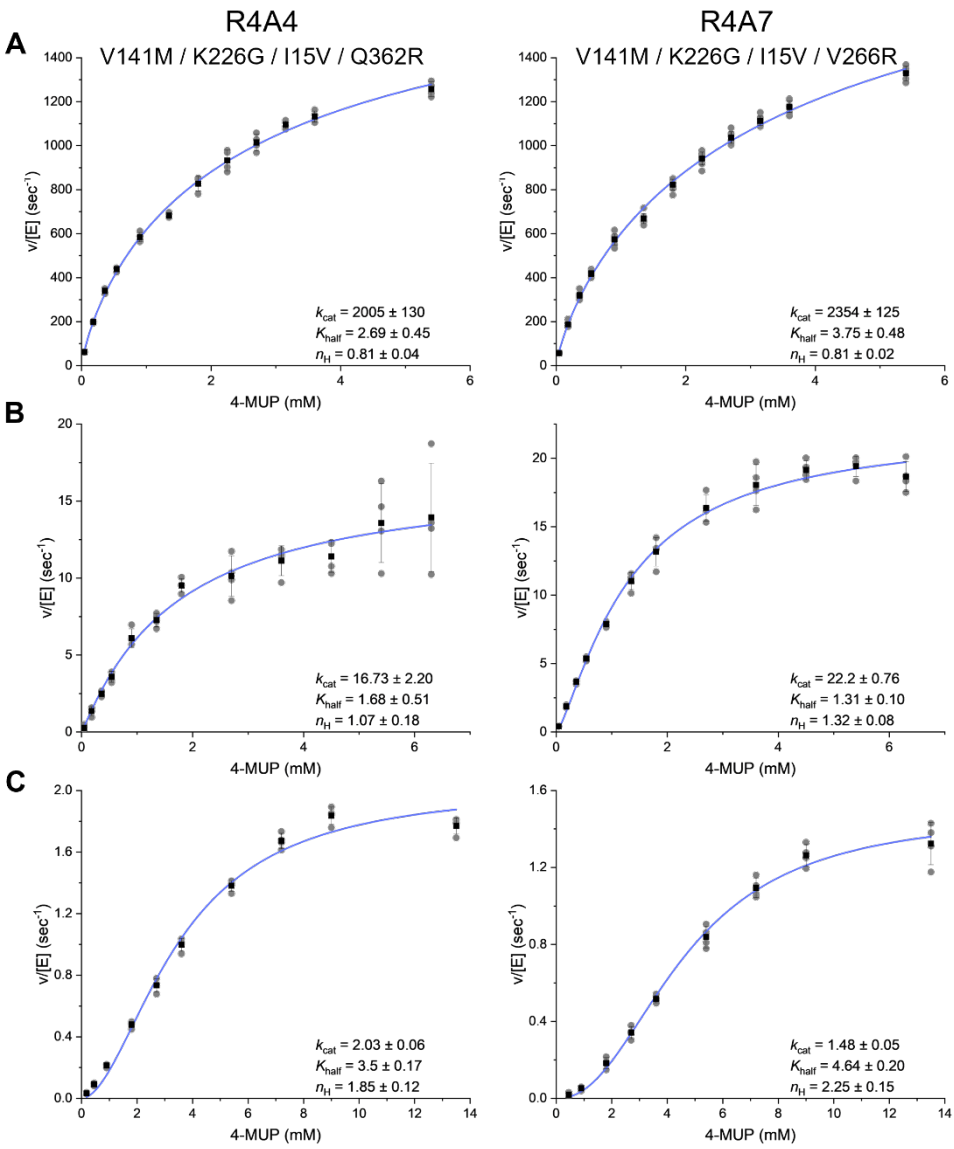
Supplementary Fig. 22 | Kinetic analysis of *Ym*Phytase variants R2E10 (left) and R3A2 (right) at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Purified *Ym*Phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to kinetic assays. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6 and 6.6, assays were performed in 0.2 M Tris maleate buffer. Black dots indicate the average of four independent samples with standard deviations signified by error bars and individual datapoints by grey dots.



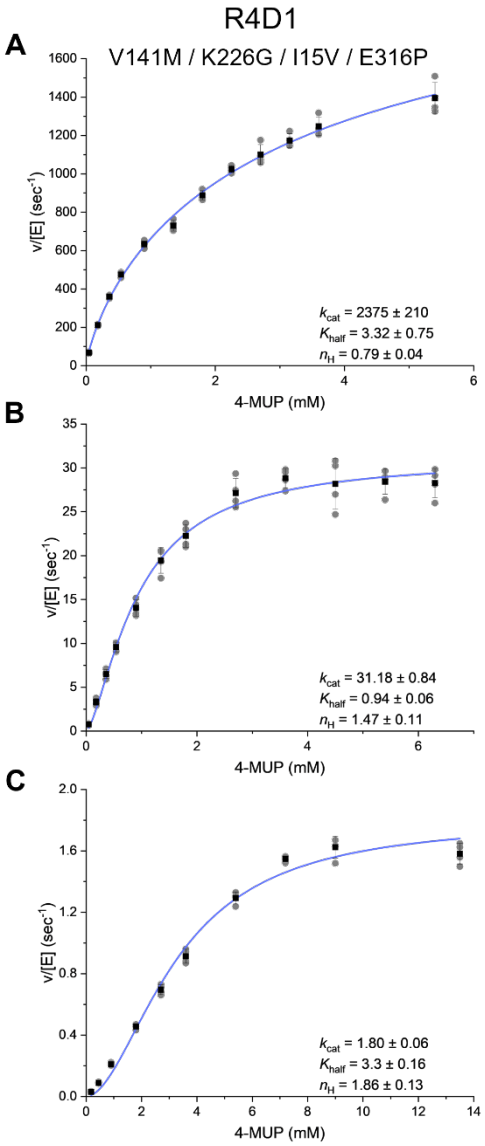
Supplementary Fig. 23 | Kinetic analysis of *Ym*Phytase variants R3A9 (left) and R3B3 (right) at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Purified *Ym*Phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to kinetic assays. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6 and 6.6, assays were performed in 0.2 M Tris maleate buffer. Black dots indicate the average of four independent samples with standard deviations signified by error bars and individual datapoints by grey dots.



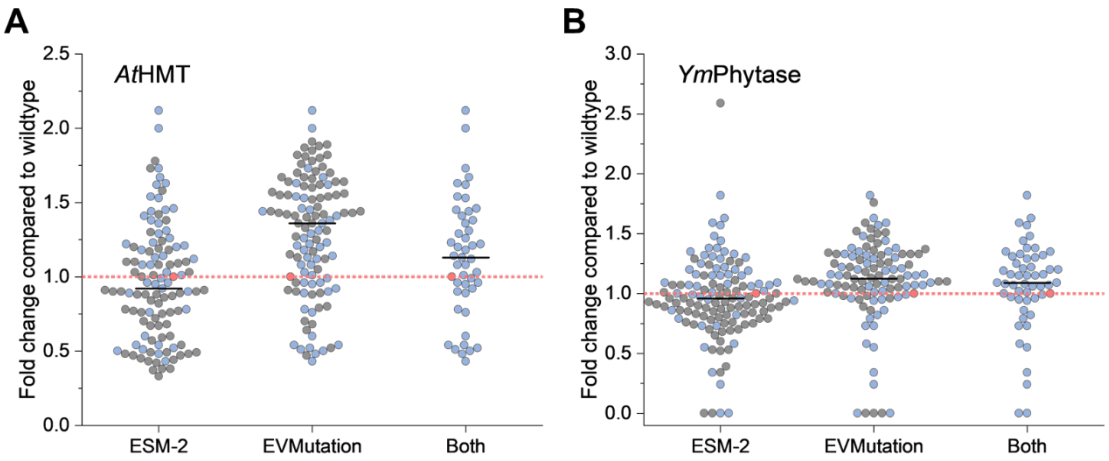
Supplementary Fig. 24 | Kinetic characterization of *Ym*Phytase variants R3G7 (left) and R4A3 (right) at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Purified *Ym*Phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to kinetic assays. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6 and 6.6, assays were performed in 0.2 M Tris maleate buffer. Black dots indicate the average of four independent samples with standard deviations signified by error bars and individual datapoints by grey dots.



Supplementary Fig. 25 | Kinetic characterization of *Ym*Phytase variants R4A4 (left) and R4A7 (right) at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Purified *Ym*Phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to kinetic assays. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6 and 6.6, assays were performed in 0.2 M Tris maleate buffer. Black dots indicate the average of four independent samples with standard deviations signified by error bars and individual datapoints by grey dots.



Supplementary Fig. 26 | Kinetic characterization of *Ym*Phytase variant R4D1 at (A) pH 4.5, (B) pH 5.6, and (C) pH 6.6. Purified phytase variants were mixed with 1 mM AEBSF and incubated at 37 °C for 30 minutes prior to kinetic assays. For pH 4.5, assays were performed in 0.25 M sodium acetate, 1 mM calcium chloride, and 0.01% Tween-20 buffer. For pH 5.6 and 6.6, assays were performed in 0.2 M Tris maleate buffer. Black dots indicate the average of four independent samples with standard deviations signified by error bars and individual datapoints by grey dots.



Supplementary Fig. 27 | Comparison of models used for prediction of initial variants of A) *At*HMT and B) *Ym*Phytase. Circles indicate variant performance relative to the wild-type enzyme (pink circle and dotted line). Mutants ranked within the top 200 predictions of both ESM-2 and EVmutation are pictured as blue circles. Mutants predicted solely by the indicated model or ranked lower than 200 by the opposing model are shown as grey circles. The black line indicates the median of all samples within the column. For *At*HMT, 26.1% (46 of 176 total predictions) of variants were predicted by both ESM-2 and EVmutation. For *Ym*Phytase, 29.4% (53 of 180 total predictions) of variants were predicted by both models.

Supplementary Table 1 | Michaelis-Menten kinetics of wild type *At*HMT and its variants using ethyl iodide as a substrate.

HMT variant	K_m	k_{cat}	Mutations
wtHMT	37.75 ± 3.74	0.47 ± 0.02	Wild type HMT from *Arabidopsis thaliana*
V140T	25.95 ± 2.37	1.53 ± 0.06	V140T
R2 A1	19.75 ± 4.66	1.56 ± 0.15	V140T / S99T
R2 C3	27.50 ± 4.28	1.93 ± 0.13	V140T / D9A
R3 A1	13.81 ± 2.45	1.50 ± 0.09	V140T / L101I / D9A
R3 A2	13.93 ± 2.71	1.68 ± 0.13	V140T / L101I / V195H
R3 A4	11.39 ± 2.52	1.36 ± 0.1	V140T / L101I / D9Q
R3 C7	17.74 ± 2.2	1.71 ± 0.08	V140T / L101I / T107L
R4 A1	19.89 ± 5.31	3.63 ± 0.38	V140T / L101I / T213S / V178E
R4 B2	20.81 ± 4.69	3.01 ± 0.27	V140T / L101I / S63N / E32N
R4 B3	26.91 ± 4.55	4.41 ± 0.32	V140T / L101I / V178E / E32N
R4 C2	19.69 ± 4.79	4.84 ± 0.46	V140T / L101I / E206D / S63N

*The notation R4 C2 signifies well C2 from the fourth round of evolution.

Supplementary Table 2 | Michaelis-Menten kinetics of wild type *At*HMT and its variants using methyl iodide as a substrate.

HMT variant	K_m	k_{cat}	Mutations
wtHMT	2.79 ± 0.3	284.47 ± 8.83	Wild type HMT from *Arabidopsis thaliana*
V140T	4.89 ± 0.95	102.28 ± 6.99	V140T
R4 A1	4.90 ± 1.25	80.58 ± 7.27	V140T / L101I / T213S / V178E
R4 B2	5.47 ± 1.69	94.76 ± 10.82	V140T / L101I / S63N / E32N
R4 B3	7.85 ± 1.61	120.86 ± 10.62	V140T / L101I / V178E / E32N
R4 C2	4.31 ± 0.72	113.73 ± 6.34	V140T / L101I / E206D / S63N

*The notation R4 C2 signifies well C2 from the fourth round of evolution.

Supplementary Table 3 | Specific activities of wild type *At*HMT and its variants using ethyl iodide a substrate.

HMT variant	Specific activity (nmol min^{-1} mg^{-1})	Mutations
wtHMT	4.72 ± 0.42	Wild type HMT from *Arabidopsis thaliana*
V140T	19.78 ± 0.72	V140T
R2A1	19.94 ± 3.56	V140T / S99T
R2C3	21.27 ± 2.31	V140T / D9A
R3A1	26.87 ± 2.63	V140T / L101I / D9A
R3A2	36.12 ± 0.72	V140T / L101I / V195H
R3A4	34.97 ± 5.32	V140T / L101I / D9Q
R3C7	30.15 ± 0.50	V140T / L101I / T107L
R4A1	62.02 ± 6.06	V140T / L101I / T213S / V178E
R4B2	37.63 ± 7.06	V140T / L101I / S63N / E32N
R4B3	64.63 ± 6.52	V140T / L101I / V178E / E32N
R4C2	79.57 ± 2.57	V140T / L101I / E206D / S63N

Supplementary Table 4 | Substrate preference of wild type *At*HMT and its variants for ethyl iodide (EtI) and methyl iodide (MeI).

HMT variant	k_{cat}/K_m Ethyl iodide	k_{cat}/K_m Methyl iodide	k_{cat}/K_m ratio	Relative substrate preference (EtI/MeI)	Mutations
wtHMT	0.01	102.34	1.22E-04	1	Wild type HMT from *Arabidopsis thaliana*
V140T	0.06	20.92	2.80E-03	23.01	V140T
R4 A1	0.18	16.44	1.11E-02	91.2	V140T / L101I / T213S / V178E
R4 B2	0.14	17.32	8.35E-03	68.64	V140T / L101I / S63N / E32N
R4 B3	0.16	15.4	1.06E-02	87.47	V140T / L101I / V178E / E32N
R4 C2	0.25	26.39	9.32E-03	76.55	V140T / L101I / E206D / S63N

*The relative substrate preference is calculated as fold change, with respect to the wtHMT, in the ratio of k_{cat}/Km for ethyl iodide and methyl iodide.

Supplementary Table 5 | Specific activities of wild type *Ym*Phytase and its variants. The specific activity was determined at pH 4.5, 5.6, and 6.6 with a substrate concentration of 0.9 mM 4-MUP. Displayed values are the average of at least four independent samples with standard deviations. Specific activities are expressed as nmol 4-MU / second / mg phytase.

Enzyme	pH 4.5	pH 5.6	pH 6.6	Mutations
wtPhy	6482 ± 74	14.9 ± 3.3	0.17 ± 0.11	Wild type *Yersinia mollaretii* phytase
M16	20783 ± 426	118.8 ± 20.7	1.87 ± 0.12	T44V / K45E
R1P2F6	12075 ± 360	104.5 ± 4.3	0.91 ± 0.12	V141M
R2C12	13023 ± 679	197 ± 17.4	3.51 ± 0.06	K226G / V141M
R2E10	13904 ± 391	147.3 ± 6.2	1.43 ± 0.11	E316D / V141M
R3A2	14313 ± 439	270.7 ± 9.4	4.22 ± 0.16	V141M / K226G / T25S
R3A9	10427 ± 252	105.6 ± 12.4	2.98 ± 0.24	V141M / K226G / L303Q
R3B3	12120 ± 295	229.5 ± 12.4	3.64 ± 0.24	V141M / K226G / I15V
R3G7	10852 ± 605	122 ± 8.2	2.65 ± 0.19	V141M / K226G / L303P
R4A3	9837 ± 144	141.7 ± 8.7	3.42 ± 0.56	V141M / K226G / I15V / Q167R
R4A4	12305 ± 446	182.9 ± 28.4	4.5 ± 0.22	V141M / K226G / I15V / Q362R
R4A7	12084 ± 525	166.1 ± 4.2	1.11 ± 0.19	V141M / K226G / I15V / V266R
R4D1	13352 ± 378	295.5 ± 19.5	4.42 ± 0.2	V141M / K226G / I15V / E316P

Supplementary Table 6 | Enzyme kinetics of wild type *Ym*Phytase and its variants using the 4-methylumbelliferyl phosphate (4-MUP) assay.

Enzyme	pH 4.5			pH 5.6			pH 6.6		
	k_{cat}	K_{half}	n_H	k_{cat}	K_{half}	n_H	k_{cat}	K_{half}	n_H
wtPhy	1396 ± 78	2.38 ± 0.23	1.31 ± 0.07	5.2 ± 0.18	1.81 ± 0.13	1.22 ± 0.06	0.31 ± 0.01	4.66 ± 0.2	1.59 ± 0.07
M16	2653 ± 74	1.27 ± 0.07	1.41 ± 0.06	27.01 ± 0.89	2.39 ± 0.17	1.12 ± 0.04	1.65 ± 0.05	3.61 ± 0.18	1.68 ± 0.1
R1P2F6	2071 ± 93	2.26 ± 0.21	1.04 ± 0.04	16 ± 0.51	1.43 ± 0.09	1.42 ± 0.08	0.96 ± 0.03	5.52 ± 0.25	1.83 ± 0.08
R2C12	2182 ± 174	2.78 ± 0.56	0.82 ± 0.05	10.8 ± 0.42	0.7 ± 0.07	1.83 ± 0.28	1.81 ± 0.05	3.95 ± 0.18	1.8 ± 0.1
R2E10	2089 ± 107	1.74 ± 0.18	1.13 ± 0.06	21.64 ± 0.62	1.45 ± 0.08	1.38 ± 0.07	1.06 ± 0.03	4.75 ± 0.2	1.79 ± 0.08
R3A2	2176 ± 118	2.29 ± 0.31	0.85 ± 0.04	30.19 ± 0.73	1.02 ± 0.05	1.48 ± 0.09	1.93 ± 0.06	3.8 ± 0.19	1.72 ± 0.1
R3A9	1517 ± 56	2.13 ± 0.2	0.84 ± 0.03	9.05 ± 0.56	1.06 ± 0.14	1.53 ± 0.24	1.26 ± 0.05	2.96 ± 0.18	2.16 ± 0.25
R3B3	2020 ± 111	2.89 ± 0.41	0.8 ± 0.03	23.99 ± 0.75	0.89 ± 0.06	1.6 ± 0.15	1.7 ± 0.05	3.36 ± 0.15	1.91 ± 0.13
R3G7	1666 ± 165	2.42 ± 0.62	0.82 ± 0.06	9.99 ± 0.4	0.91 ± 0.08	1.82 ± 0.25	1.24 ± 0.06	3.17 ± 0.23	2.06 ± 0.25
R4A3	1493 ± 88	2.43 ± 0.37	0.81 ± 0.04	15.82 ± 0.58	0.98 ± 0.08	1.46 ± 0.14	1.21 ± 0.05	2.99 ± 0.2	1.81 ± 0.17
R4A4	2006 ± 130	2.69 ± 0.45	0.81 ± 0.04	16.73 ± 2.2	1.68 ± 0.51	1.07 ± 0.18	2.03 ± 0.06	3.5 ± 0.17	1.85 ± 0.12
R4A7	2355 ± 125	3.75 ± 0.48	0.81 ± 0.02	22.2 ± 0.76	1.31 ± 0.1	1.32 ± 0.08	1.48 ± 0.05	4.64 ± 0.2	2.25 ± 0.15
R4D1	2376 ± 210	3.32 ± 0.75	0.79 ± 0.04	31.18 ± 0.84	0.94 ± 0.06	1.47 ± 0.11	1.8 ± 0.06	3.3 ± 0.16	1.86 ± 0.13

*****An enzyme label of R4D1 signifies the variant in well D1 from the fourth round of autonomous protein engineering. Mutations present in the variants listed above can be found in Table S5.

Supplementary Table 7 | Comparison of various reported autonomous protein engineering platforms

Automation platform	Target	Prediction & design	Library construction	Expression & screening	Functional assay	Iterations & improvement	Reference
In-house automation at iBioFAB facility (Suppl Fig. 1)	*At*HMT and *Ym*Phytase enzymes	AI/ML tools; ESM-2, EVmutation, low-N supervised learning	High fidelity site-directed mutagenesis (SDM)	BL21 *E. Coli* expression and 96-well plate assays for screening with crude cell lysates	Absorbance (*At*HMT) and Fluorescence (*Ym*Phytase) using enzyme assays	4 iterations, ~450 variants; 16-fold (*At*HMT) and 26-fold (*Ym*Phytase) activity improvement	This Study
In-house automation at iBioFAB facility (Suppl Fig. 1)	T7 promoter optimization	Bayesian optimization	Golden gate assembly from a library of DNA fragments	*E. Coli* expression in 96-well plates and DMSO extraction of lycopene	Absorbance of DMSO extracted lycopene from *E. Coli*	3 iterations, 136 variants; Improved lycopene production by 9-fold	HamediRad et al.,[1]
Strateos Cloud Lab; a commercial laboratory automation vendor	Glycoside hydrolase enzyme	Bayesian optimization	Golden gate assembly from a library of gene fragments	Cell-free protein expression	Thermostability using thermal cycler and fluorescent probe	20 iterations, 3 variants/iteration; Improved thermostability by at least 12^{O}C	Rapp et al., [2]
Automation designed using in-house robotic platform.	RhlA enzyme for rhamnolipid production	Bayesian optimization	Golden gate assembly, SDM, site-saturation mutagenesis	*E. Coli* expression in 96-well plate and organic solvent extraction in 24-well plates	MALDI-ToF MS analysis of *E. Coli* extracted rhamnolipid	4 iterations, 1536 variants; Improved RL congener production by 4.8-fold	Hu et al., [3]

References

1. HamediRad, M. *et al.* Towards a fully automated algorithm driven platform for biosystems design. *Nat. Commun.* **10**, 5150 (2019).
2. Rapp, J. T., Bremer, B. J. & Romero, P. A. Self-driving laboratories to autonomously navigate the protein fitness landscape. *Nat Chem Eng* **1**, 97–107 (2024).
3. Hu, R. *et al.* Protein engineering via Bayesian optimization-guided evolutionary algorithm and robotic experiments. *Brief. Bioinform.* **24**, bbac570 (2023).