Supplementary Information

High-throughput analysis and protein engineering using microcapillary arrays

Authors and affiliations:

Bob Chen[1], Sungwon Lim[1*], Arvind Kannan[2*], Spencer C. Alford[1*], Fanny Sunden[3], Daniel Herschlag[3,4], Ivan Dimov[5], Thomas M. Baer[6], Jennifer R. Cochran[1,2]

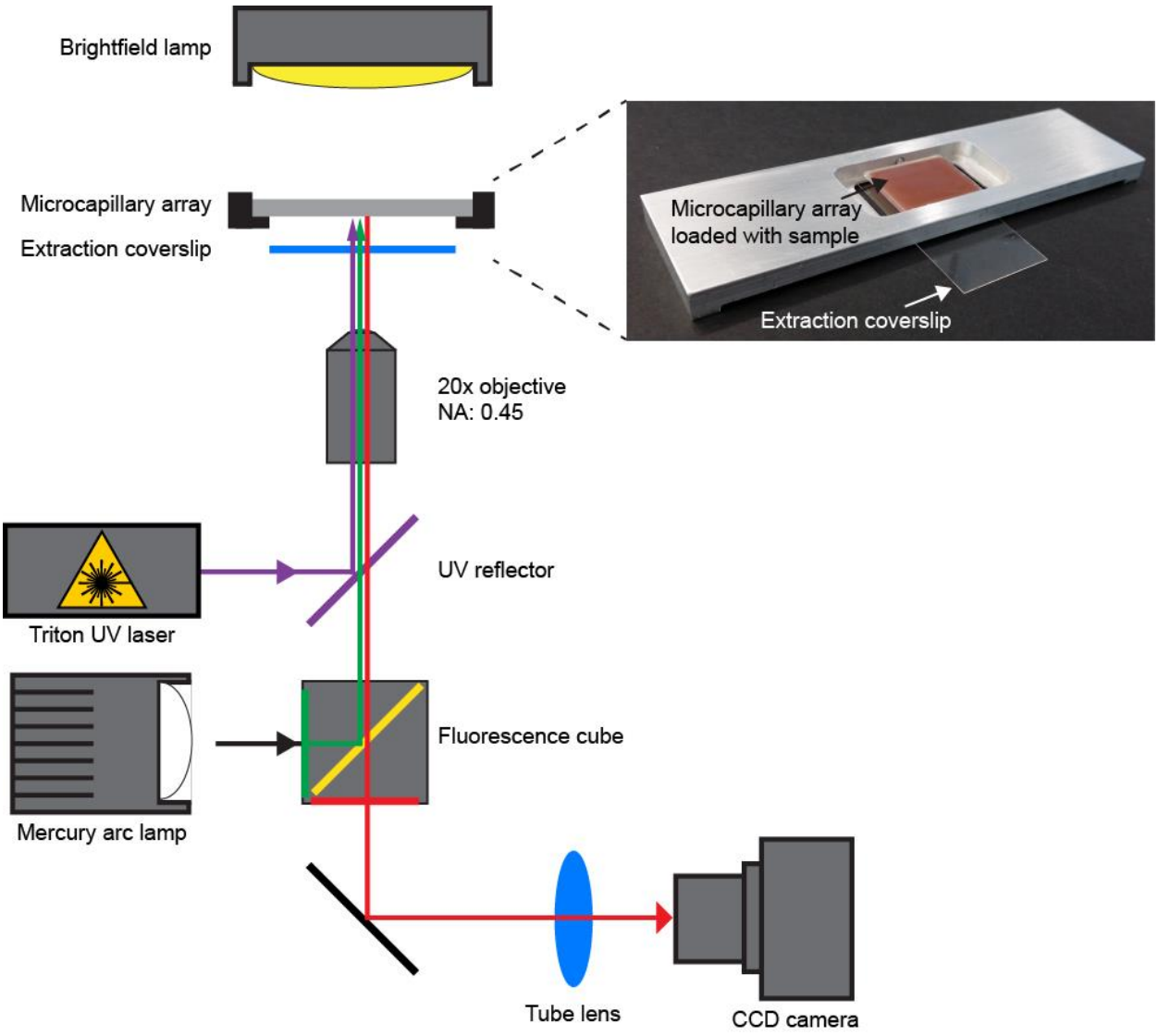
[1]Department of Bioengineering, [2]Chemical Engineering, [3]Biochemistry, [4]Chemistry, [5]Institute for Stem Cell Biology and Regenerative Medicine, [6]Stanford Photonics Research Center

[*]These authors contributed equally

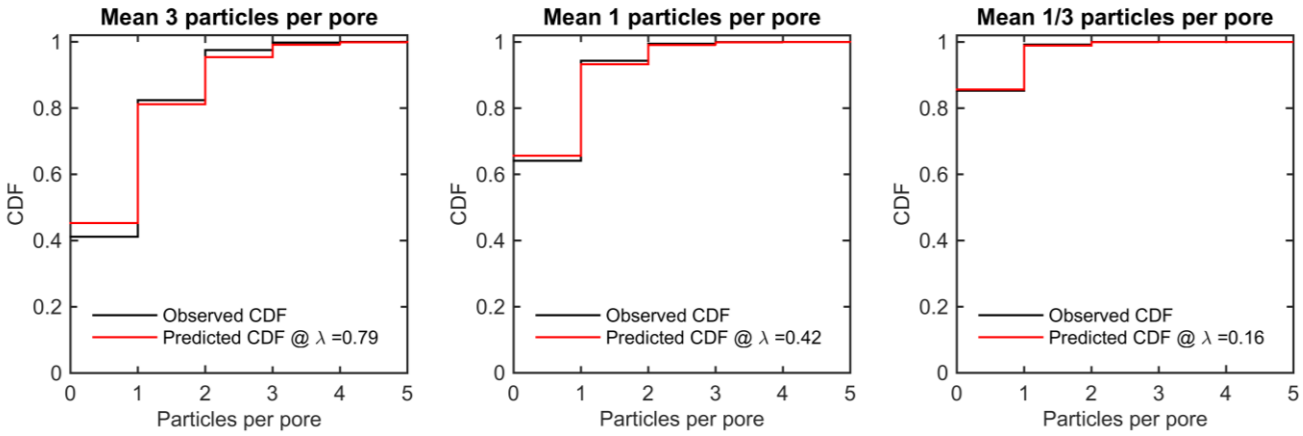
To whom correspondence should be addressed:
tmbaer@stanford.edu, jennifer.cochran@stanford.edu

Supplementary Results

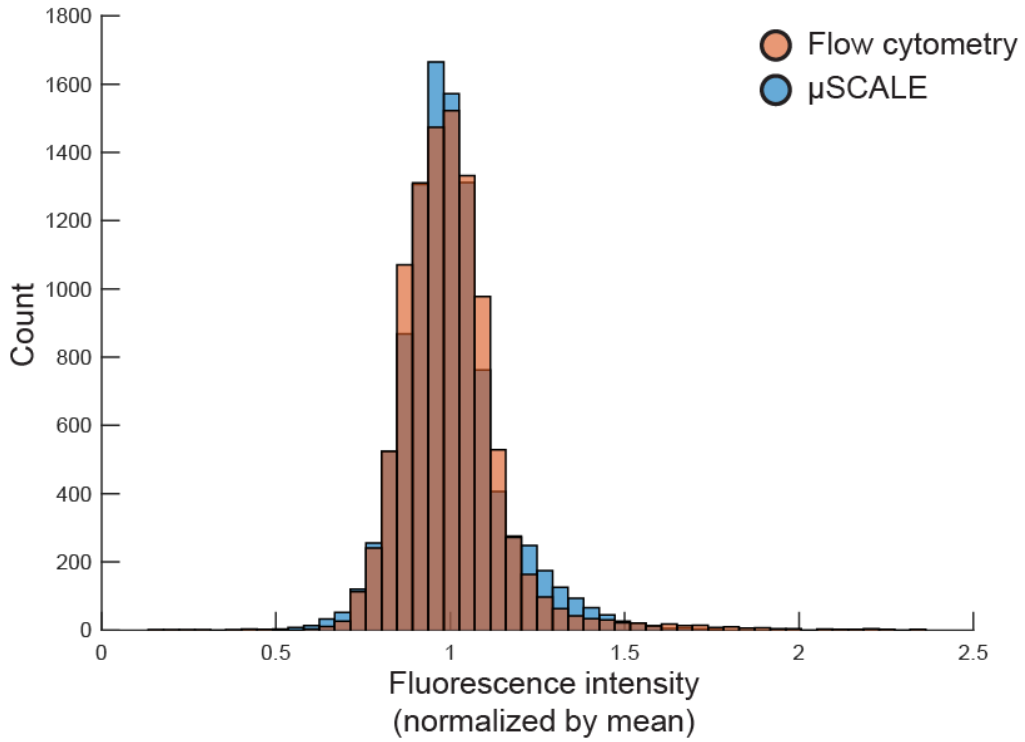
Supplementary Figures



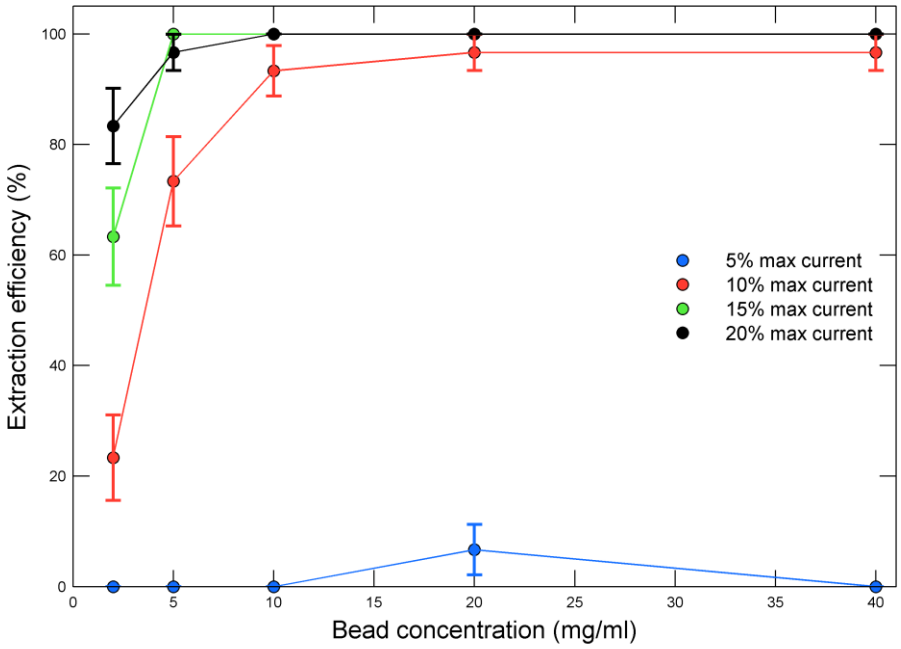
Supplementary Figure 1: µSCALE platform diagram. The µSCALE platform is incorporated around the automated inverted fluorescence microscope and the Triton UV laser system in the Arcturus Veritas Laser Capture Microdissection (LCM) System. The instrument has both bright-field and fluorescence imaging capabilities. The 3-axis stage allows for rapid alignment of the laser with the desired microcapillary. A similar configuration could be recreated with a wide range of automated inverted microscopes with appropriate fluorescence illumination system and laser source. **Inset:** Sample (brown color imparted by beads) is loaded into the microcapillary array and placed into a holder, which is then analyzed using the instrument. An extraction coverslip is placed below the holder for sample recovery.



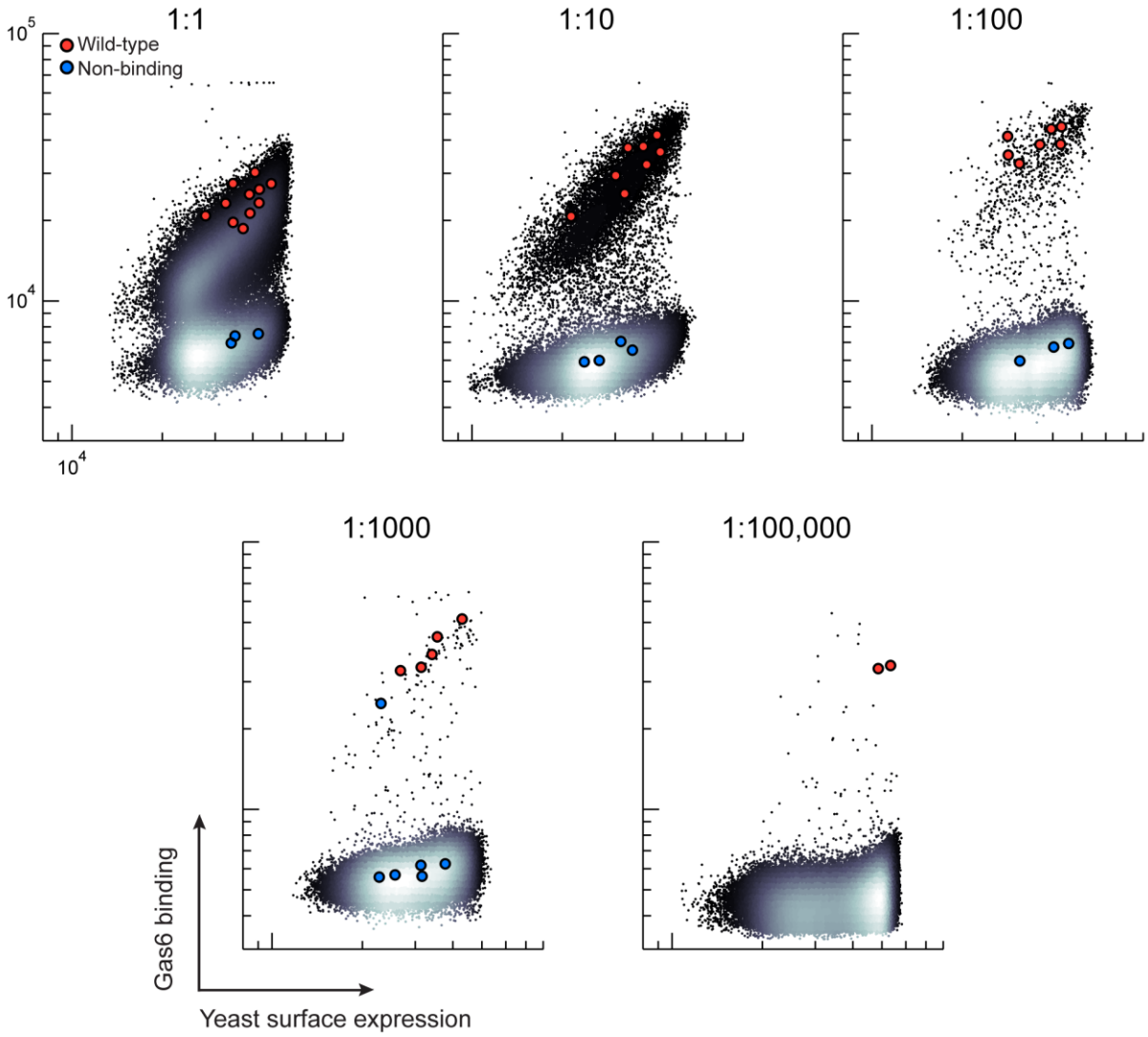
Supplementary Figure 2: Loading of microcapillaries follows Poisson statistics. Fluorescent beads were load into the array at three concentrations corresponding to means of 3, 1, and 1/3 particles per microcapillary (λ = 3,1,1/3). The contents of 20,000 microcapillaries were counted, and the cumulative distribution functions (CDF) for each condition are plotted as red lines. Maximum likelihood fits to the Poisson distribution with λ = 0.79, 0.42, and 0.16 are overlaid above the data as black lines. While the distributions closely follow Poisson statistics (KS distance< 0.04), the observed and expected means differ by 2–3 fold.[1] This difference likely arises due to the high aspect ratio of the microcapillaries such that some particles may not travel into the imaging plane.



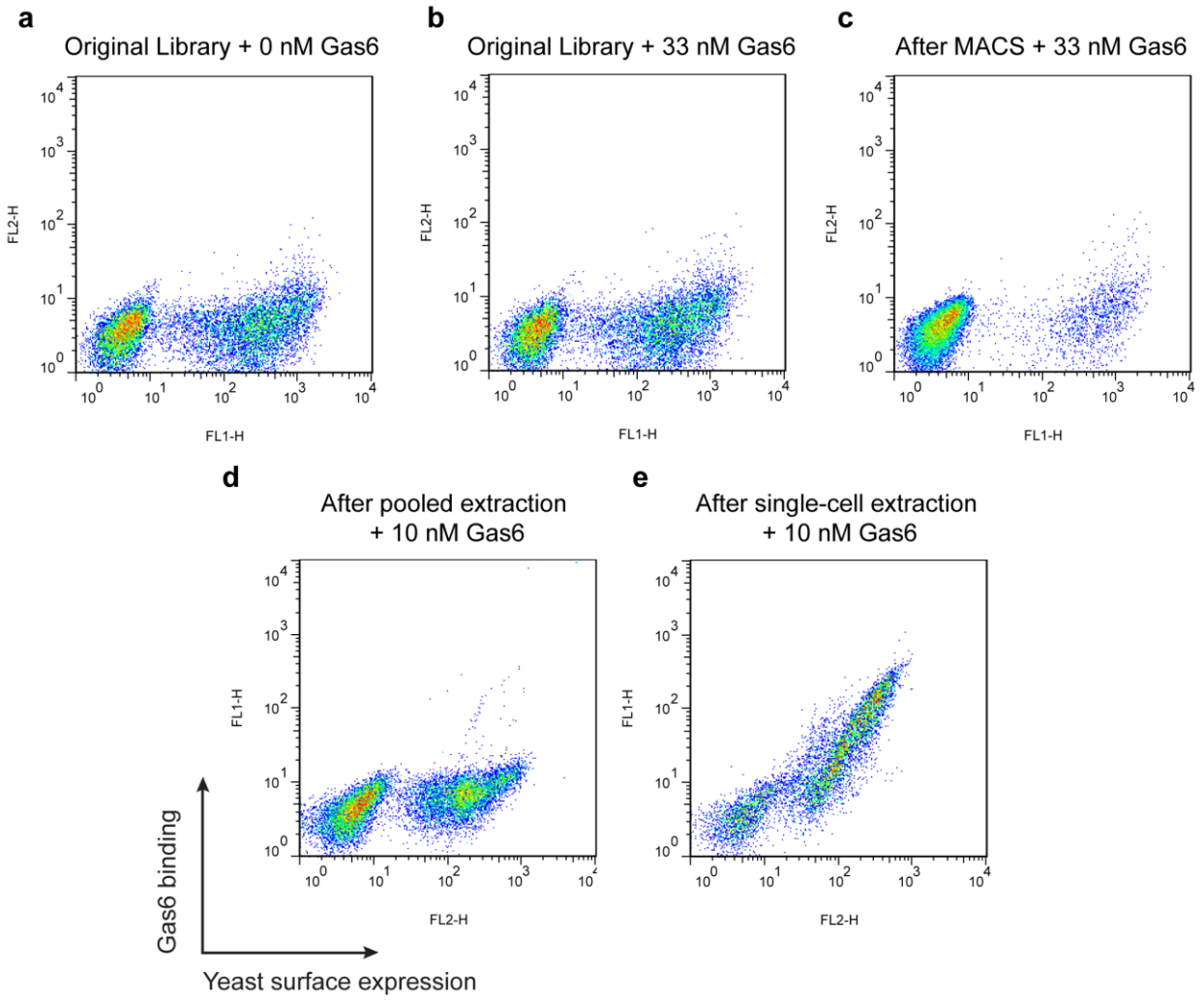
Supplementary Figure 3: Assessment of inter-capillary variability in µSCALE fluorescence measurements. Fluorescent beads (Sphero, FP-4052-2, Spherotech) were analyzed on both µSCALE and a Guava easyCyte flow cytometer (Millipore) (n = 10,000). For µSCALE measurements, the fluorescent beads were loaded into a 20 µm microcapillary array without magnetic particles. Both populations were normalized by the mean of their respective distributions. The coefficients of variation of the bead fluorescence analyzed by flow cytometry and by µSCALE were 0.15 and 0.14, respectively.



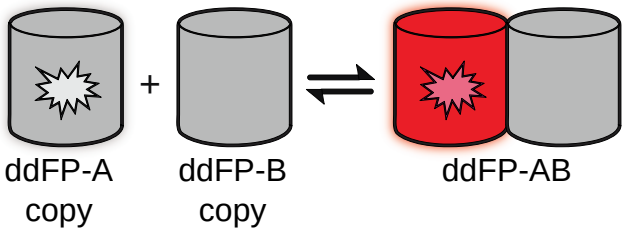
Supplementary Figure 4: Extraction efficiency at different laser powers and bead concentrations. For each extraction, the Triton cutting laser of the LCM was pulsed for approximately 18 msec at a range of different laser powers, ranging from 5-20% of the maximum current as set by the laser manufacturer. Thirty extractions were performed for each condition. Error bars represent the binomial sampling error for each condition.



Supplementary Figure 5: μSCALE scatter plots from mock library screens. Ratio of Axl wild-type : non-binding clones in each mock library = 1:1 (a), 1:10 (b), 1:100 (c), 1:1,000 (d), 1:10,000 (e), and 1:100,000 (f). Extracted microcapillaries are indicated as red dots (identified as wild-type) and blue dots (identified as non-binding mutant). The total number of analyzed microcapillaries was 74,373, 72,021, 64,931, 87,254, and 276,384, respectively.



Supplementary Figure 6: Flow cytometry scatter plots showing enrichment of Gas6 binders from the yeast-displayed naïve scFv library following screening by MACS and μSCALE. Yeast cell surface expression levels of scFv proteins are measured by primary and secondary antibodies against the C-terminal c-myc epitope tag (x axis). Gas6 binding is measured using primary and secondary antibodies against a hexahistidine tag (y axis). (a) Original scFv naïve library (library size: 7×10^8) without addition of Gas6. (b) Original scFv naïve library incubated with 33 nM Gas6. (c) Analysis of scFv clones after MACS (new library size: 2.8×10^6). Cells were incubated with 33 nM Gas6. (d) Analysis of scFv clones isolated after first μSCALE sort via auto-extraction (library size: 143). (e) A Gas6-binding scFv clone isolated after second μSCALE sort via single-cell extractions. For (d) and (e), yeast cells were incubated with 10 nM Gas6. The total number of analyzed cells was 12,754, 13,096, 12,344, 13,373, and 8,064 respectively.

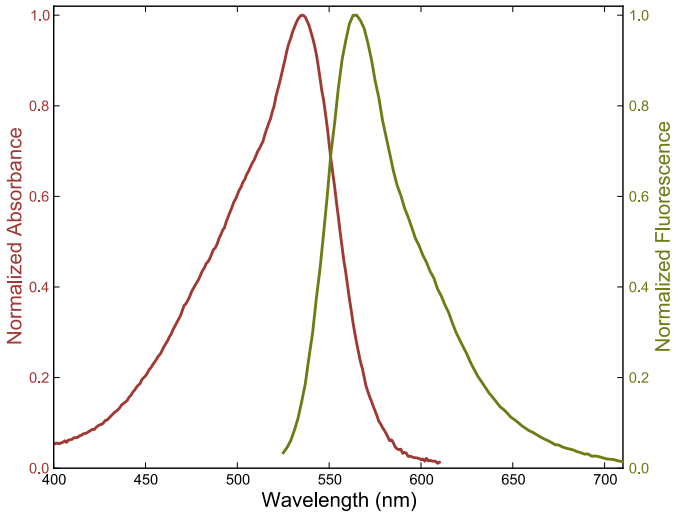


Supplementary Figure 7: ddFP technology. Weak or nonfluorescent protein monomers, designated ddFP-A (A copy) and ddFP-B (B copy), reversibly bind to form a bright fluorescent heterodimer of characteristic hue. The chromophore is indicated by a starburst.

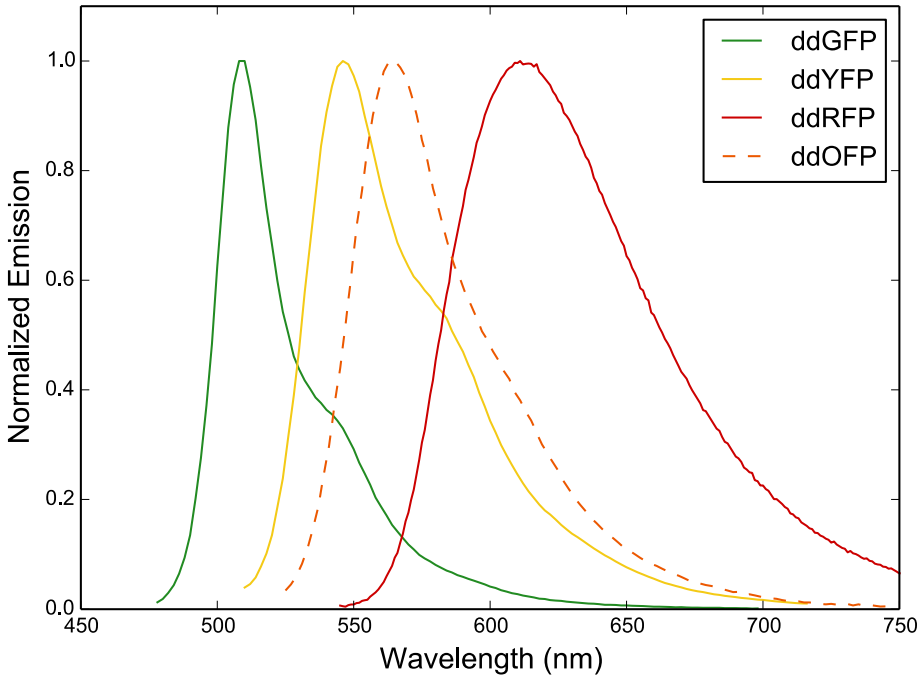
a)

```
ddOFP     1 -------VIKEFMRFKVRLEGSMNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWD
ddRFP     1 MVSKSEEVIKEFMRFKVRLEGSMNGHEFEIEGEGEGRPYEGTQTAKLKVTKGGPLPFAWD

ddOFP       ILSPQITYGSKAYVKHPADVPDYMKLSFPEGFKWERVMHFEDGGLVTVTQDTSLQDGTLI
ddRFP       ILSPQIMYGSKAYVKHPADVPDYMKLSFPEGFKWERVMHFEDGGLVTVTQDTSLQDGTLI

ddOFP       YKVKMRGTNFPPDGPVMQRKTLGWDYSTERLYPENGVLKGELLGRLKLKDGGLYLVEFKT
ddRFP       YKVKMRGTNFPPDGPVMQKKTLGWDYATERLYPEDGVLKGELLGRLKLKDGGLNLVEFKT

ddOFP       IYMAKKPVQLPGYYFVDTKLDITSHNEDYTIVEQYERSEGRHHLGMDELYK
ddRFP       IYMAKKPVQLPGYYFVDTKLGITSHNEDYTIVEQYERSEGRHHLGMDELYK
```

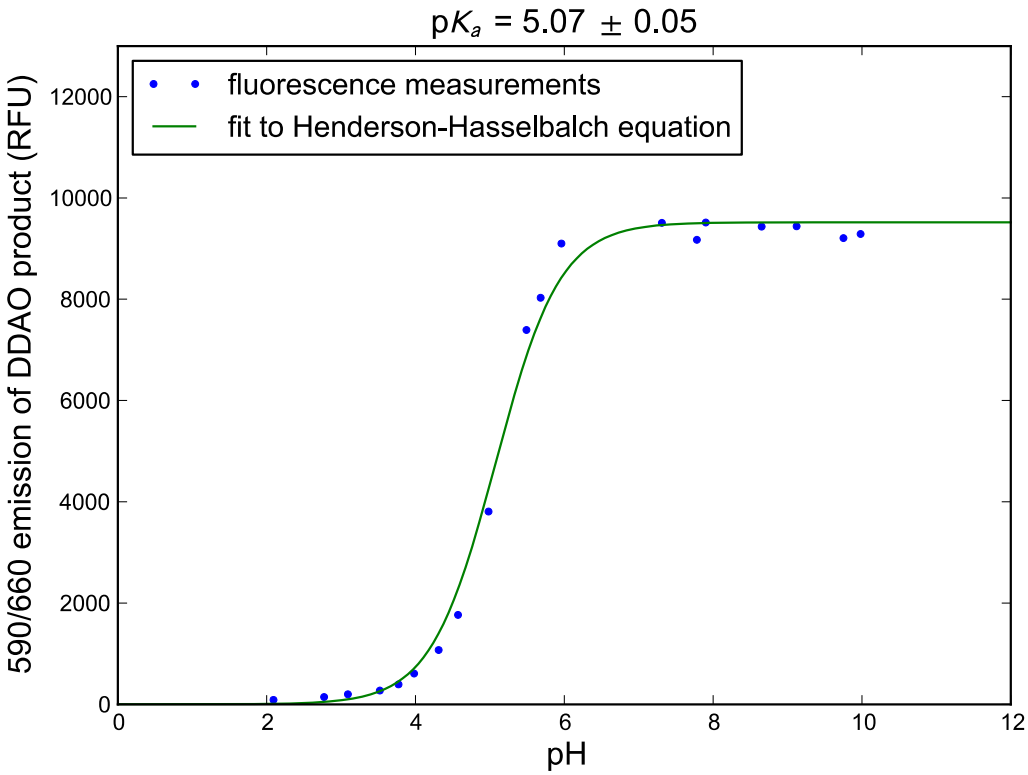
b)



Supplementary Figure 8: Characterization of ddOFP. (a) Sequence alignment of ddOFP with its parent, ddRFP. The original M66T mutation is boxed and colored blue. Mutations acquired with three rounds of directed evolution are highlighted in grey. (b) Normalized absorbance (red line) and emission (green line) spectra for ddOFP.

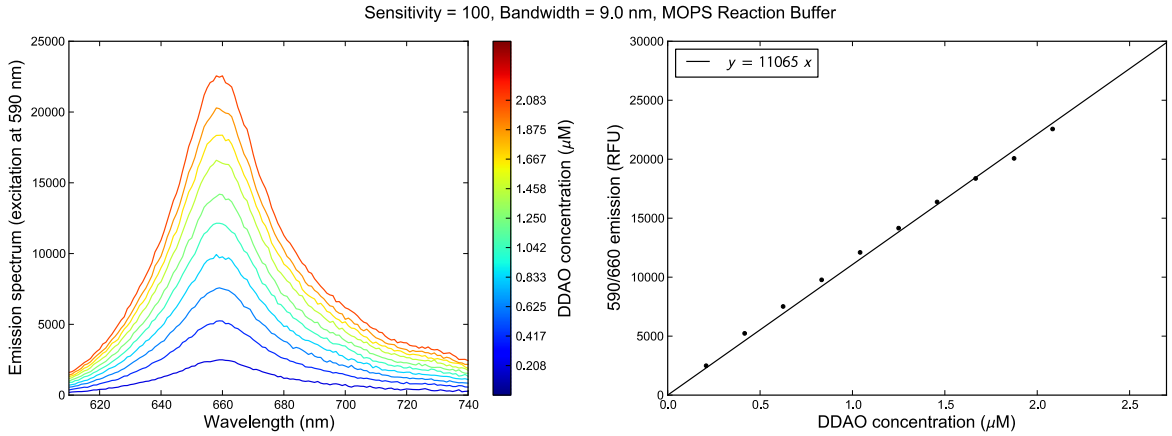


Supplementary Figure 9: ddOFP fills the spectral gap between ddYFP and ddRFP. Normalized emission spectra for ddGFP (green line), ddYFP (yellow), ddOFP (orange dashed line), and ddRFP (red). Reference spectra for ddGFP, ddYFP, and ddRFP taken from previous work.[2]

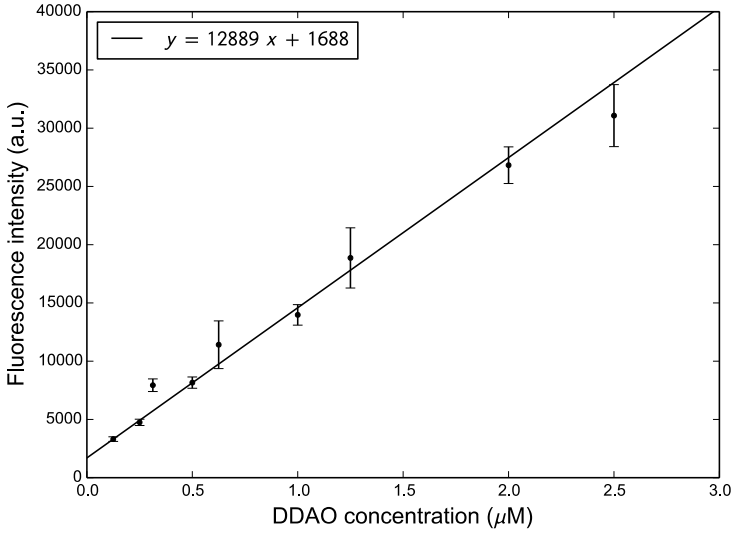


Supplementary Figure 10: Determination of the pK_a of the fluorescent DDAO product used in high-throughput screens of alkaline phosphatase mutants. Error corresponds to a 1-σ (68%) confidence interval for the fit pK_a value, and was determined by bootstrapping using case resampling of the data. Subsequent enzymatic assays were performed at pH 8.0 (3 units above the product pK_a).

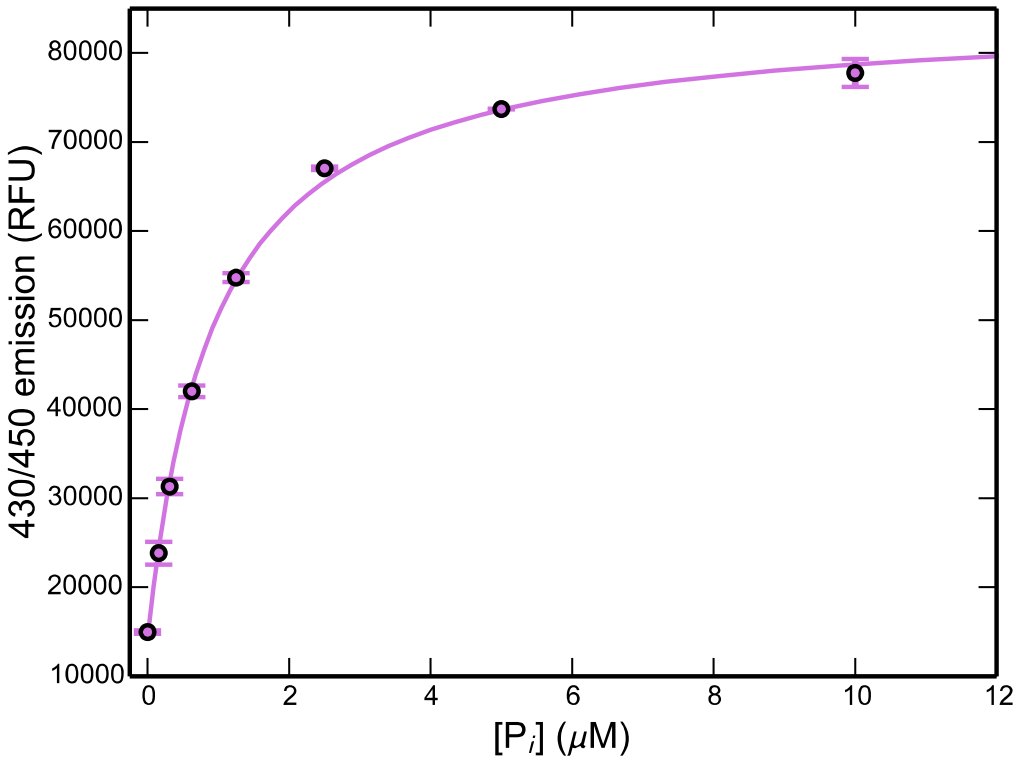
(a)



(b)

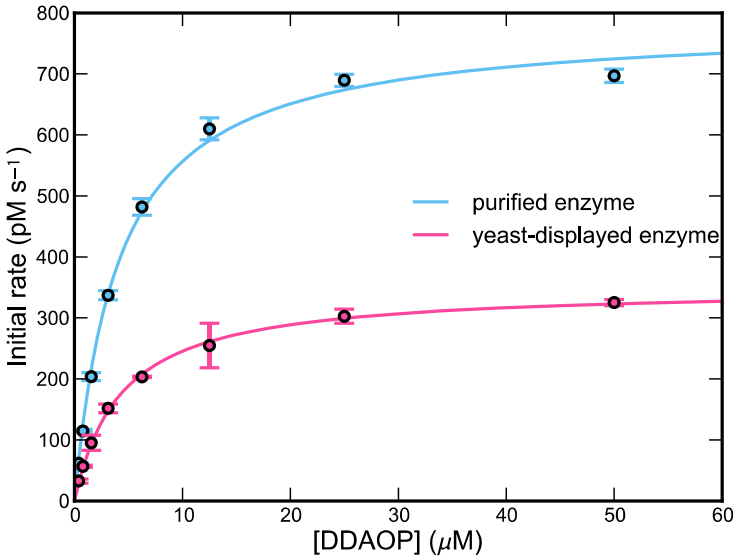


Supplementary Figure 11: Product calibration curves for quantifying enzyme kinetics. (a) Emission spectra for the DDAO product were measured as a function of product concentration on a Biotek Synergy H4 plate reader spectrophotometer, and the peak emissions were quantified to yield the calibration line on the right. The same fluorescence acquisition parameters (sensitivity = 100 and bandwidth = 9.0 nm) for the spectrophotometer were used for all bulk enzymatic assays. (b) A similar calibration line was generated on the μSCALE platform by loading product standards of variable concentration onto isolated spots on a single microcapillary array and then measuring the fluorescence intensity of each microcapillary on the array (n ≈ 3500 microcapillaries for each concentration). Error bars reflect the standard deviation across all microcapillaries in each spot and were used to fit a calibration line using weighted least squares.

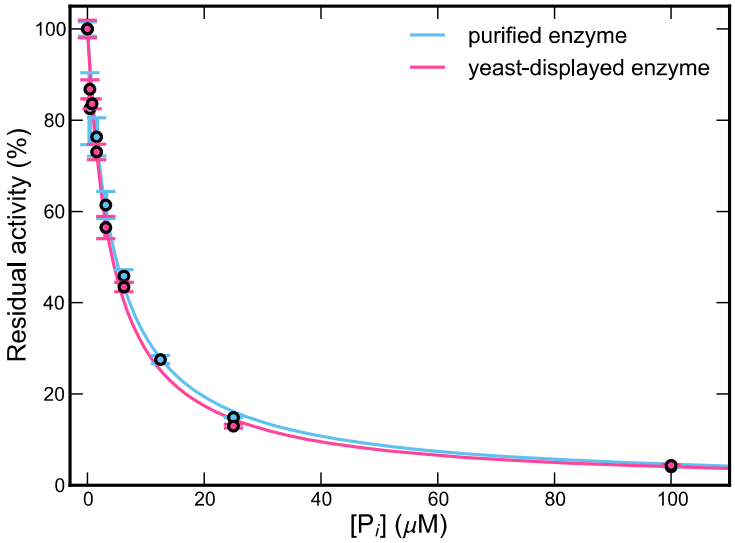


Supplementary Figure 12: Calibration curve for quantifying levels of inorganic phosphate present in alkaline phosphatase reactions using fluorescently labeled phosphate-binding protein from *E. coli*. Error bars reflect the standard deviation over three measurements. A fit to a two-state binding isotherm is overlaid above the data as a solid line.

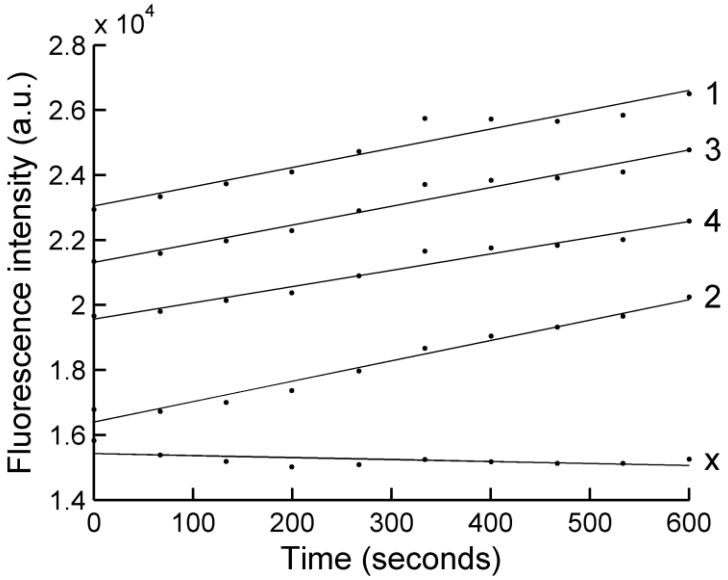
a)



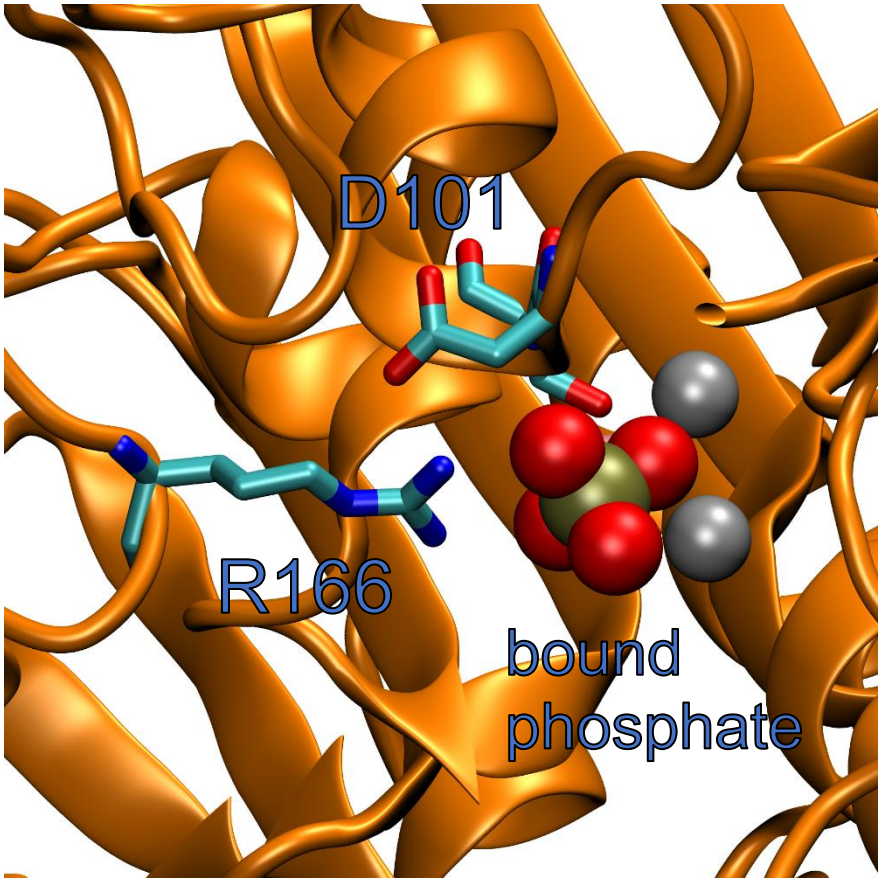
b)



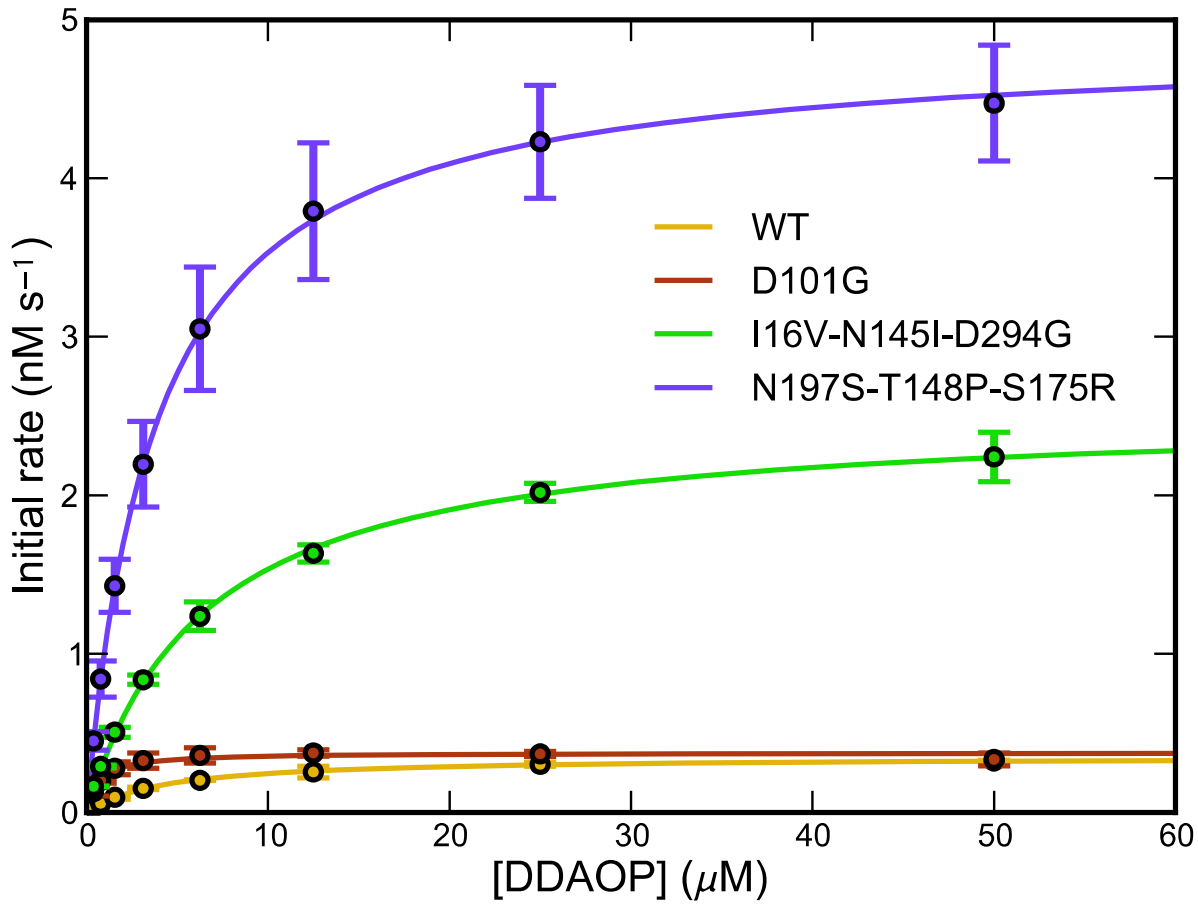
Supplementary Figure 13: a) Kinetics of WT alkaline phosphatase displayed on the surface of yeast compared to purified WT enzyme harvested from *E. coli*. Reactions were run at a 200 µl scale in 96-well microtiter plates with either 50 pM of purified enzyme or 6.4 × 10^5 yeast cells from a culture induced for expression. Fits to the standard Michaelis-Menten model are overlaid above the data as solid lines. b) Comparison of the phosphate inhibition profile of yeast-displayed WT alkaline phosphatase with that of purified WT enzyme harvested from *E. coli*. All reactions were run at a DDAOP concentration of 400 nM (10-fold below the apparent K_M for WT alkaline phosphatase). Fits to a competitive inhibition model are overlaid above the data as solid lines. For both parts, error bars correspond to the standard deviation of three technical replicates.



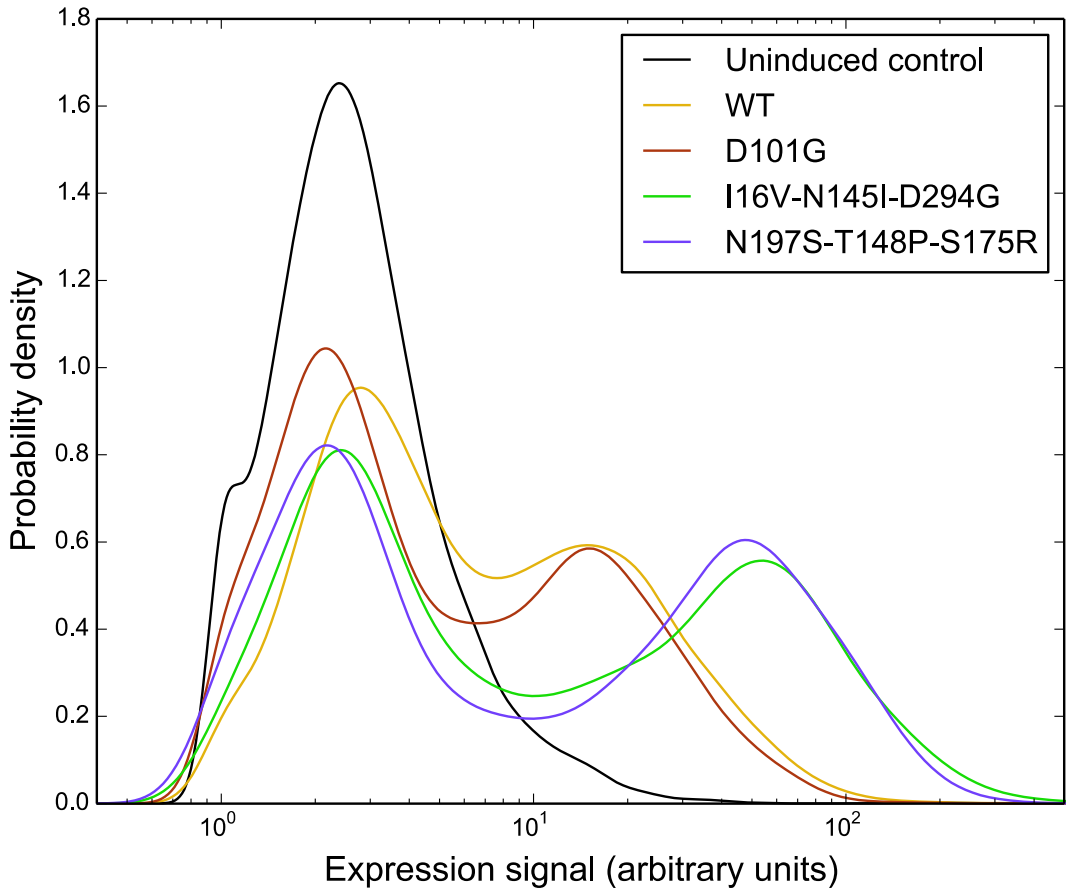
Supplementary Figure 14: Quantification of the microcapillaries corresponding to the images in Figure 4a. The numbering scheme corresponds with Figure 4a, indicating four microcapillaries harboring yeast cells displaying active enzyme (1-4) and a nearby empty microcapillary ('x') for comparison. A constant accumulation of product in each microcapillary is observed, which enables the calculation of reaction rates by linear regression.



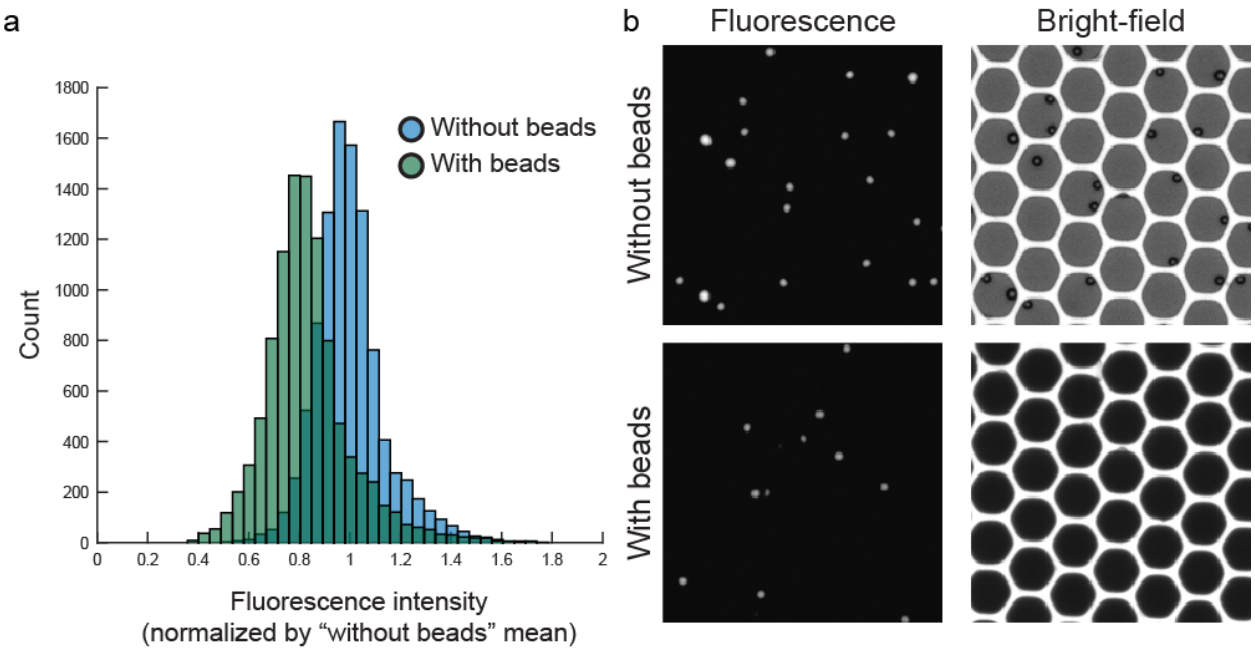
Supplementary Figure 15: Structural basis for reduced phosphate inhibition in the D101G mutant. Figure shows active site of WT alkaline phosphatase (PDB *1ALK*) with bound inorganic phosphate (shown as red and gold spheres). Active site zinc ions are shown as gray spheres, D101 and R166 residues are rendered as sticks, and all other protein residues are depicted in ribbon representation. The D101 position is known to be important for positioning the R166 residue that coordinates the phosphate ion within the alkaline phosphatase active site. As such, the D101G mutation uncovered in our screens likely alters the positioning of R166 in a manner that lowers the affinity of the enzyme for inorganic phosphate.



Supplementary Figure 16: Kinetics of yeast-displayed WT alkaline phosphatase and the three variants isolated from μSCALE screens as a function of DDAOP concentration. Fits to the standard Michaelis-Menten model are overlaid above the data as solid lines. All reactions were performed at a cell loading of 3200 cells/μl (equivalent to the catalyst concentration within a single 20 μm microcapillary), so the resulting kinetic parameters can be interpreted as per-cell specific activities. Error bars correspond to the standard deviation of three technical replicates.



Supplementary Figure 17: Expression profiles of alkaline phosphatase mutants displayed on the surface of yeast (*n* = 10,000 cells for each sample). Protein display levels were measured via antibody labeling of the C-terminal c-myc epitope tag fused to each enzyme and quantified using flow cytometry.



Supplementary Figure 18: Effect of magnetic beads on µSCALE fluorescence measurements. Fluorescent beads (Sphero, FP-4052-2, Spherotech) were analyzed on µSCALE with and without the presence of the magnetic particles used for extraction (n = 10,000). a) Histograms of the two conditions. The mean of the normalized "with beads" population is 0.83, demonstrating that the magnetic beads partially occlude the fluorescence intensity of the microcapillaries. The coefficients of variation of the fluorescent bead distributions with and without magnetic particles are 0.2 and 0.14, respectively, demonstrating that the use of magnetic beads for extraction does not significantly increase the variability of fluorescence measurements on µSCALE. b) Fluorescence and bright-field images of the two conditions. The presence of magnetic beads reduces the fluorescence intensity and obscures the bright-field images.

Supplementary Tables

Initial Ratio		Post-sort		
WT	Non-binding	WT	Non-binding	Enrichment Ratio
1	1	11	0	2
1	10	8	0	11
1	100	7	0	101
1	1,000	5	1	834
1	10,000	4	1	8,001
1	100,000	2	0	100,001

Supplementary Table 1: Enrichment ratios from mock libraries composed of wild-type (WT) and non-binding Axl clones. A series of mock library extractions were performed to quantify the efficiency of sorting. Enrichment ratio is defined as: $\eta = \frac{N_{+,1}}{N_{+,1}+N_{-,1}} / \frac{N_{+,0}}{N_{+,0}+N_{-,0}}$ where η is enrichment ratio, $N_{+,0}$ is the initial number of wild-type cells, $N_{+,1}$ is the number of wild-type cells after sorting, and $N_{-,0}$and $N_{-,1}$ are the equivalent values for non-binding cells, respectively.

	EC	**QY**	**Brightness**	**p*K*$_a$**
ddFP-A copy	32200	0.08	2576	7.9
ddFP-B copy	NA	NA	NA	NA
tdOFP	64500	0.31	19995	6.9

Supplementary Table 2: Summary of spectral properties of ddOFP and its constituent monomers, A copy and B copy (ddFP technology is described in Supplementary Figure 7). EC, extinction coefficient ($M^{-1}cm^{-1}$); QY, quantum yield; NA, not applicable. QY determined relative to mCitrine (QY = 0.78). Brightness is the product of extinction coefficient and quantum yield. The B copy is non-fluorescent. A genetic fusion of A and B copies (tandem dimer OFP [tdOFP]) was used to determine spectral properties of the ddOFP system.

Sample	[P_i]
AP Reaction Buffer	1.8 ± 0.2 μM
0.2% BSA (in Milli-Q H_2O)	150 ± 10 nM
0.5 μM DDAOP (in Milli-Q H_2O)	< 10 nM
5 μM DDAOP (in Milli-Q H_2O)	< 10 nM
50 μM DDAOP (in Milli-Q H_2O)	< 10 nM
50 pM WT Enzyme (purified)	< 10 nM

Supplementary Table 3: Measurement of inorganic phosphate levels present in alkaline phosphatase reactions using fluorescently labeled phosphate-binding protein from *E. coli*. Fluorescence measurements were converted to phosphate concentrations using the calibration curve in Supplementary Figure 11 (see Methods). Errors were propagated from uncertainties in the calibration curve using the parametric bootstrap method. Limit of detection for the assay is 10 nM P_i.

Mutant	Construct	k_{cat} (s^{-1})	K_M (μM)	k_{cat} / K_M (M^{-1} s^{-1})	K_I (μM)
WT	yeast displayed	7.0 ± 0.2	4.3 ± 0.4	(1.6 ± 0.1) × 10^6	4.2 ± 0.3
WT	purified protein	15.7 ± 0.2	4.1 ± 0.1	(3.8 ± 0.1) × 10^6	4.8 ± 0.3
R166S	yeast displayed	0.4 ± 0.1	55 ± 7	(7 ± 2) × 10^3	1500 ± 100
R166S	purified protein	0.7 ± 0.2	51 ± 4	(1.4 ± 0.4) × 10^4	1580 ± 20

Supplementary Table 4: Comparison of kinetic parameters obtained for yeast-displayed and purified alkaline phosphatase constructs using DDAOP and inorganic phosphate as the substrate and inhibitor, respectively. For substrate titrations performed with yeast-displayed enzymes, initial rates were converted to enzyme turnovers by assuming an average of 10^4 displayed protein copies per yeast cell. Errors correspond to 1-σ (68%) confidence intervals for the fit parameters, and were estimated with the parametric bootstrap method using data from three technical replicates of each fit point. Reported K_M and K_I values should be interpreted as apparent values, as they are inflated by the presence of inorganic phosphate contamination in the reaction buffer (see Supplementary Table 3).

Mutant	k_{rel} (yeast display)	k_{rel} (purified protein)	Relative yeast expression levels
WT	1.00 ± 0.07	1.00 ± 0.03	1.0
D101G	10.39 ± 0.07	49 ± 6	0.6
I16V-N145I-D294G	6.4 ± 0.1	0.0039 ± 0.0006	2.1
N197S-T148P-S175R	19.2 ± 0.5	0.011 ± 0.002	1.9

Supplementary Table 5: Relative rates of WT alkaline phosphatase and the three variants isolated from μSCALE screens in both yeast-displayed and purified enzyme formats. All reactions were conducted using 1 μM DDAOP in the presence of 15 μM added inorganic phosphate. All yeast assays were carried out at a cell density of 3200 cells/μl using freshly transformed cells. Purified enzyme reactions with WT alkaline phosphatase and the D101G mutant were carried out using 50 pM enzyme, while reactions with the I16V-N145I-D294G and N197S-T148P-S175R triple mutants were carried out using 50 nM enzyme (with rates scaled accordingly). Relative rates are defined on a per-cell basis for yeast-displayed constructs and on a per-mole basis for purified constructs, and are normalized to the WT values. Uncertainties in reported rates reflect the standard deviation from three technical replicates. Relative yeast surface expression levels of each mutant are shown alongside the rate data for comparison, demonstrating that improved performance of the I16V-N145I-D294G and N197S-T148P-S175R triple mutants in cell-based assays can be partially attributed to increased expression of these variants. Relative expression is calculated as (<mutant> - <uninduced>)/(<WT> - <uninduced>), where <> refers to the geometric mean of the data in Supplementary Figure 17.

Mutant	Zn / protein	Mg / protein	P / protein
WT	2.50	0.82	0.85
I16V-N145I-D294G	0.08	0.01	0.62
N197S-T148P-S175R	0.08	0.01	0.10

Supplementary Table 6: Metal ion and phosphorous content in WT alkaline phosphatase and the two triple mutants isolated from μSCALE screens, as measured by atomic emission spectroscopy (AES). The expected Zn/Mg/protein stoichiometry for an alkaline phosphatase monomer is 2:1:1.[3] In addition, WT alkaline phosphatase expresses with inorganic phosphate bound in the active site in a 1:1 stoichiometry. Deviations from these values in the WT alkaline phosphatase sample presumably reflect errors in sample preparation and determination of protein concentration. The lack of significant metal occupancy in the active sites of the I16V-N145I-D294G and N197S-T148P-S175R triple mutants helps explain the poor activity of these variants. The AES data, along with the crystal structure of WT alkaline phosphatase (PDB *1ALK*), support a model in which mutations at N145 and T148 disrupt the folding of the soluble protein and hence its ability to bind the metal cofactors required for efficient catalysis.

	Experiment	Extracted pores	Observed Growth	Viability
Yeast	Mock library – *S. cerevisiae*			
	1:1	20	14	70%
	1:10	20	12	60%
	1:100	20	10	50%
	1:1000	20	11	55%
	1:10000	20	9	45%
	1:100000	5	2	40%
	scFv library – *S. cerevisiae*			
	Sort 2	15	12	80%
	AP library – *S. cerevisiae*			
	Sort 1	15	12	80%
			Mean	61% ± 14%
E. coli	FP library – *E. coli*			
	Sort 1 & 2	24	24	100%
	Sort 3	10	10	100%
			Mean	100% ± 0%
B. subtilis	Extraction survival test – *B. subtilis*			
	Survival	20	20	100%
			Mean	100% ± 0%

Supplementary Table 7: Viability of cells after laser extraction. This table shows the number of clones that grew into viable cultures after single-cell extraction for both yeast and bacterial cells. Viability is defined as the ratio of cells with observed growth over extracted microcapillaries.

References:

1. Arnold, T. & Emerson, J. Nonparametric Goodness-of-Fit Tests for Discrete Null Distributions. *R J.* 34–39 (2011).

2. Alford, S. C., Ding, Y., Simmen, T. & Campbell, R. E. Dimerization-dependent green and yellow fluorescent proteins. *ACS Synth. Biol.* **1,** 569–575 (2012).

3. Zalatan, J. G., Fenn, T. D. & Herschlag, D. Comparative enzymology in the alkaline phosphatase superfamily to determine the catalytic role of an active-site metal ion. *J. Mol. Biol.* **384,** 1174–1189 (2008).