

**Supporting Information for**

**Ultrahigh-Throughput Directed Enzyme Evolution by**

**Absorbance-Activated Droplet Sorting (AADS)**

Fabrice Gielen,<sup>†§</sup> Raphaëlle Hours,<sup>†§</sup> Stéphane Emond,<sup>†</sup> Martin Fischlechner,<sup>†‡</sup> Ursula Schell,<sup>#</sup> Florian Hollfelder<sup>†\*</sup>

<sup>†</sup> Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, CB2 1GA, United Kingdom

<sup>‡</sup> Institute for Life Sciences, University of Southampton, Southampton SO17 1BJ, United Kingdom

<sup>#</sup> Johnson Matthey Catalysis and Chiral Technologies, 28, Cambridge Science Park, Milton Road, Cambridge CB4 0FP, United Kingdom

\* Corresponding author: E-mail: [fh111@cam.ac.uk](mailto:fh111@cam.ac.uk). Phone: +44-1223766048

§ These authors contributed equally to this work.

## 1. SI Text

### 1.1 Detailed Materials and Methods

#### Bacterial strains, plasmids, enzymes and chemicals

Proteins were expressed in the *E. coli* strain BL21(DE3). Electrocompetent *E. coli* strain E.cloni 10G (Lucigen) was used for the construction of directed evolution libraries and DNA recovery from single emulsions. Plasmid pASK-IBA63b-plus (High copy plasmid, IBA, Germany) was used for the expression of *Rhodococcus* sp. M4 PheDH (wtPheDH; Uniprot ID Q59771) in fusion with a C-terminal Strep-tag. Unless indicated otherwise, all DNA modifying enzymes were from Thermo Fisher Scientific and used according to the manufacturer's instructions. Chemicals used in this study were all purchased from Sigma-Aldrich unless indicated otherwise.

#### Construction of directed evolution libraries

In the first round of directed evolution, the wtPheDH gene was randomly mutated by error-prone PCR, using Mutazyme II low-fidelity polymerase (Stratagene) with an average of 1.7 mutations per gene. The resulting PCR product was ligated into pASK-IBA63b-plus at NcoI and XhoI sites. The ligation products were then transformed into electrocompetent *E. coli* E.cloni 10G cells, yielding library **Lib0** (~10<sup>6</sup> transformants). For the second round of evolution, a shuffling library (dubbed **Lib1**) was prepared using StEP PCR (1) by recombining six hits (V26I, Q45H, T13N/L193M, T121I and N122S/T339I) from the first round of screening. The reaction was performed using 0.15 pmol of the total template plasmid pool with BioTaq DNA polymerase (Bioline), 0.6 μM of each primer, 0.2 mM of each dNTPs and 1.5 mM of MgCl<sub>2</sub>. The PCR program was performed as previously described (1). Subsequent cloning steps were carried out as described above to yield library **Lib1** (~10<sup>6</sup> transformants).

#### Preparation of cells for compartmentalization and lysis in droplets

For the expression of wtPheDH and its variants, *E. coli* BL21(DE3) cells were initially grown overnight at 37°C in 5 mL LB broth supplemented with ampicillin (Amp; 100 μg/mL). 5 mL LB-Amp medium containing ampicillin (100 μg/mL) were inoculated with this starter culture at a ratio 1:20. After 2 h of growth at 37°C,

protein expression was induced by addition of anhydrotetracyclin (AnhTet; 200 ng/mL) and growth at 20 °C for 16 h. After expression, OD600 was measured to determine the cell density (an OD600 of 1 is equivalent to  $\sim 5 \times 10^8$  cells/mL) and the culture was diluted to enable encapsulation at occupancies of 1 cell per droplet. This diluted cell suspension was then mixed with the density-matching agent Percoll (final concentration: 25%, v/v) in 100 mM glycine-KOH buffer, pH 10 and taken up into a syringe. A second syringe contained a mixture of substrate (10 mM L-Phe, 10 mM NAD<sup>+</sup>, 5 mM WST-1 and 5 µg/mL mPMS and 1 mM of WST-1 formazan as an offset, in 100 mM glycine-KOH buffer, pH 10) and cell lysis agents (rLysozyme 4 µL/mL Novagen (30 KU/µl), and CelLytic<sup>TM</sup> B (4% (v/v) of 10× stock) in 100 mM glycine-KOH buffer pH10). This procedure was used for the screening of directed evolution libraries. For model enrichment experiments, two cultures were grown separately: one for the expression of wtPheDH and the other for the expression of P91 (not active on L-Phe) (2). After determining the cell density, these cultures were mixed to a ratio of 1:5000 (wtPheDH : P91) and diluted prior to droplet encapsulation as described.

### **DNA recovery from droplets by transformation**

Sorted droplets (representing a typical total volume of  $\sim 180$  pL) were de-emulsified by addition of 200 µL 1H,1H,2H,2H-perfluorooctanol (PFO), to achieve phase separation. After the removal of the top layer (i.e. the aqueous layer containing DNA) the bottom layer (containing PFO in oil) was again extracted with 200 µL double-distilled water to recover any remaining DNA. The harvested plasmid DNA was recovered from the combined aqueous layers using a spin column (DNA Clean & Concentrator-5, Zymo Research), eluting in 7 µL of elution solution from the kit. The purified DNA was then transformed into *E.coli* E. cloni 10G cells by applying one electric pulse of 1.80 kV (using an *E. coli* Pulser Cuvette, 0.1 cm electrode; Bio-Rad and MicroPulser). Sterile S.O.C Medium (275 µL, Invitrogen) was added and the transformed cells were incubated for 40 min at 37 °C, before being plated on LB-Amp agar and grown overnight at 37 °C. The transformation efficiency was determined by comparison of the number of colonies obtained after transformation to the number of sorted droplets. The plasmid DNA corresponding to the collected

wtPheDH variants was then extracted from the colonies and used to transform *E. coli* BL21(DE3) cells for further screens on colonies or in microplate format.

### **Screening procedures in microplates and on colonies**

For screening in microplate format, *E. coli* BL21(DE3) transformants were individually picked and grown overnight in 300  $\mu$ L LB-Amp at 37 °C in 96 deep-well plates. 25  $\mu$ L of these cultures were used to inoculate 500  $\mu$ L LB-Amp in 96 deep-well plates. After 2 hours of growth at 37 °C, protein expression was induced by adding AnhTet (to a final concentration of 200 ng/mL) and cells were further grown overnight at 20 °C. Cells were then pelleted by centrifugation at 4°C at maximum speed (3320 $\times$ g) for 5-10 minutes and the supernatant removed. Pellets were frozen at -80 °C for 30 min and, after thawing, lysed by addition of 200  $\mu$ L of lysis buffer (25 mM Tris-HCl pH 7.5, 0.1% Triton X100, 100  $\mu$ g/mL rLysozyme (Novagen) and 0.8 U/mL benzonase (Novagen)). After 30 minutes of lysis, cell debris was spun down at 4 °C at 3320 $\times$ g for 20 minutes. Enzyme assays were performed in 96-well plates containing a volume of 200  $\mu$ L per well (20  $\mu$ L pre-diluted lysate + 180  $\mu$ L of 10 mM L-Phe, 10 mM NAD<sup>+</sup>, 5 mM INT and 5  $\mu$ g/mL mPMS in 100 mM glycine-KOH buffer pH10). The reactions were monitored by recording the absorbance at 495 nm.

For colony screens, transformed *E. coli* BL21(DE3) cells (on average 500-2,000 CFU) were plated on LB-amp agar. The colonies were replicated with a sheet of filter paper (BioTrace NT Pure Nitrocellulose Transfer Membrane 0.2  $\mu$ m, PALL Life Sciences) that was applied to the plate to print a replica and then placed onto a second plate containing LB-amp agar with 200 ng/mL AnhTet. After expression overnight at 20°C, the filter paper was placed into an empty Petri dish and cells were lysed prior to the activity assay by alternating three times between storage at -20°C and 37°C. Subsequently, a layer of top agar (0.5% agar in 100 mM glycine-KOH pH 10) containing the substrates (10 mM L-Phe, 10 mM NAD<sup>+</sup>, 5 mM INT and 5  $\mu$ g/mL mPMS) was applied and the emergence of red color (resulting from the formation of INT-Formazan) monitored after 30 minutes.

These procedures were the basis of the determination of enrichment ratios and also used to identify hits after droplet sorting during directed evolution experiments.

### **Protein expression analysis**

The mutants were transformed and expressed in *E. coli* BL21(DE3) overnight at 20 °C in 5 mL liquid cultures (containing anhydrotetracyclin, 200 ng/mL). 500 µL of the resulting cell suspension were pelleted. The pellets were resuspended in 50 µL of lysis solution (BugBuster protein extraction reagent from Merck Millipore (10:1 diluted in water) and lysonase bioprocessing reagent from Merck Millipore (3 µL/m)), and incubated at room temperature for one hour. Cell debris was pelleted and the soluble fractions were collected. 2 µL of each soluble fraction were analysed by SDS-PAGE (12%, run at 200 V for 1 hour; see Fig. S11). The pellets containing the insoluble fractions were washed with a 5-fold diluted lysis solution (see above) ; and 2 µL of each were analysed side-by-side by SDS-PAGE. The amount of expressed protein was quantified by densitometric analysis (ImageJ) of the corresponding band on the gel in Fig. S11.

### **Protein purification procedures**

After expression in 200 mL liquid cultures, *E. coli* BL21(DE3) cells expressing wtPheDH or its variants were harvested by centrifugation and resuspended in 10 mL of lysis buffer (100 mM Tris-HCl pH 8.0, 150 mM NaCl, 1× Bugbuster™ protein extraction reagent (Merck Millipore), 3 µL/mL lysonase™ bioprocessing reagent (Merck Millipore) and EDTA-free protease inhibitors). Cell debris was removed by centrifugation (30,000×g, 1 h, 4 °C) and the clarified lysate was directly loaded onto Strep-Tactin Superflow resin (IBA Life). Strep-tagged wtPheDH proteins were eluted with Elution buffer (100 mM, pH 8.0, containing 150 mM NaCl and 2.5 mM d-desthiobiotin) according to the manufacturer's instructions. Eluted proteins were concentrated to a final volume of 1 ml and buffer exchange was performed using PD MiniTrap G-25 Spin columns from GE Healthcare with 100 mM phosphate buffer pH 7. All the identified hits were purified using this procedure, with the exception of two 2<sup>nd</sup> round variants (V26I/N122S/T339I and Q45H/N122S/L193M), which aggregated during the purification process.

### **Kinetic characterization of PheDH variants**

Enzymatic assays were performed at 25 °C in a final volume of 200 µL of 100 mM glycine-KOH pH 10 under saturating conditions of NAD<sup>+</sup> (5 mM) and a range of L-Phe concentrations (0.2-60 mM). Purified protein was diluted to a concentration of 10

nM. Initial velocities ( $v_0$ ) were determined by monitoring NADH formation at 340 nm. Kinetic parameters were obtained by fitting initial rates  $v_0$  to the Michaelis-Menten with substrate inhibition using Kaleidagraph (Synergy Software):  $v = k_{\text{cat}} [S] / (K_M + [S] + ([S]^2 / K_i))$  (3). Example data and details of equations used for fitting are shown in Fig. S12.

### **Differential scanning fluorimetry**

2  $\mu\text{M}$  of the purified enzymes (wtPheDH and variants V26I/L193M and V26I/N122S/L193M/T339I) were mixed with Sypro Orange protein gel stain (from Invitrogen), in two different buffers: 2 $\times$  in MOPS (100 mM, NaCl 150 mM, pH 8) or Glycine-KOH (100 mM, pH 10). The samples were denatured by increasing the temperature from 25 to 80  $^{\circ}\text{C}$  using the Corbett Life Science Rotor-gene 6000, and the fluorescence of the Sypro orange was measured ( $\lambda_{\text{excitation}} = 410 \text{ nm}$ ,  $\lambda_{\text{emission}} = 610 \text{ nm}$ ). The  $T_m$  (defined as the temperature at which half of the enzyme population is denatured) corresponds to the first derivative for each temperature-fluorescence curve.

### **Kinetics of thermal inactivation**

Half-lives of thermal inactivation were determined for purified wtPheDH and the two most thermostable variants (V26I/L193M and V26I/N122S/L193M/T339I) by incubating the enzymes (2  $\mu\text{M}$ ) at 50 $^{\circ}\text{C}$  for various time intervals. Initial and residual activities were measured at 20 $^{\circ}\text{C}$  in glycine-KOH buffer 100 mM, pH 10, by measuring the NADH production at 340 nm with a spectrophotometer. The first-order rate constant,  $k_d$ , of irreversible thermal denaturation was obtained from the slope of the linear plots of  $\ln(\text{initial } v_0 / \text{residual } v_0)$  versus time (measured at [L-Phe]=10 mM), and the half-lives ( $t_{1/2}$ ) were calculated as  $\ln 2 / k_d$ .

### **Activity versus temperature profiles**

Initial specific activities of the purified wtPheDH and the two most thermostable variants (V26I/L193M and V26I/N122S/L193M/T339I) were measured at temperatures ranging from 30 to 65  $^{\circ}\text{C}$ , in glycine-KOH buffer (100 mM, pH 10, with 10 mM L-Phe and 5 mM  $\text{NAD}^+$ ), by following the NADH production at 340 nm with a spectrophotometer.

## 2. Supplementary Data

### A. Operation of the droplet sorter

#### A.1 Theoretical sensitivity of the absorbance detection

The percentage of transmission that could be reliably detected was 0.3% based on the noise level (30 mV) when working close to the saturation voltage of the detector (10 V). Therefore, at fixed pathlength (here 50  $\mu\text{m}$ ), given the extinction coefficient of WST-1 formazan ( $\epsilon^{455\text{nm}} = 34660 \text{ M}^{-1} \text{ cm}^{-1}$  at pH 10 in glycine-KOH buffer, 100 mM, measured at 455 nm), and applying the Beer-Lambert law, the estimated detection limit would be  $\sim 7.5 \mu\text{M}$ .

#### A.2 Fluidics

Sorting large droplets require higher electric fields because they experience a higher drag force. Although the dielectrophoretic force scales with the volume of the droplets, the large distance between droplets and electric field maximum ( $>100 \mu\text{m}$ ) means this bulk force is not significantly higher than for smaller droplets flowing closer to the electrodes. Additionally, increases in the flow rate of the spacing oil was found to push droplets further away from the central separation wall. This means that the emulsion-to-oil ratio could be increased at high oil flow rates without affecting the sorting operation and, for instance, was 1:10 at 40  $\mu\text{L}/\text{min}$ . This is useful for reducing the flow rate of the oil and therefore increase throughput further while reducing the total volume of respacing oil needed.

#### A.3 Assessment of leakage

Droplets with buffer only and droplets containing WST-1 (5 mM), mPMS (5  $\mu\text{g}/\text{mL}$ ) and NADH (5 mM) were co-generated in a microfluidic device with two flow-focussing channels converging to a single outlet and incubated together in tubing. The carrier phase was HFE-7500 containing 1.5% Picosurf 1 (w/w). The absorbance of droplets was measured several times during incubation for 0-4 hours by passing them through the detection module at arbitrary intervals. No increase in the absorbance of buffer droplets was observed within detection sensitivity ( $\sim 10 \mu\text{M}$ ) confirming the absence of detectable leakage.

#### **A.4 Triggering droplet selection based on absorbance signals**

Droplets containing glycine buffer (pH 9.85) only and buffer with WST-1 formazan (100  $\mu$ M) were co-generated (See A.3 above) and their absorbance measured in the microfluidic sorting chip (Fig. 1A). A threshold was applied based on the second scattering edge to distinguish droplets containing WST-1 from those that did not. Below 100  $\mu$ M WST-1 the edges mask the real absorbance value and droplets cannot be distinguished, which defines this concentration as the threshold for triggering a sorting signal. To circumvent this limitation, WST-1 formazan (1 mM) was added to the samples, resulting in a signal which can be triggered more sensitively (see Fig. 2B and Fig. S4).

#### **A.5 Code for the Arduino Due microcontroller**

The following code was used to read the analog signal coming from the photodetector and compare it to an arbitrary value (sortValue). When the sensor value (sensorValue) was lower than the sort value, it triggered a digital output (D13) to the high state.

```
int triggerPin = 13;

float sortValue=7.5; // arbitrary voltage threshold

void setup() {

  pinMode(triggerPin, OUTPUT);

  Serial.begin(115200);

  analogReadResolution(12);

}

void loop() {

  // read the input on analog pin 0:

  int sensorValue = analogRead(A0);

  float voltage = sensorValue * (3.33 / 4096.0);

  if (voltage<(sortValue/3))

  {
```

```
digitalWrite(triggerPin,HIGH);  
  
delay(2);  
  
}  
  
else  
  
digitalWrite(triggerPin,LOW);  
  
}
```

## **B. Calculation of the number of soluble enzyme molecules per cell**

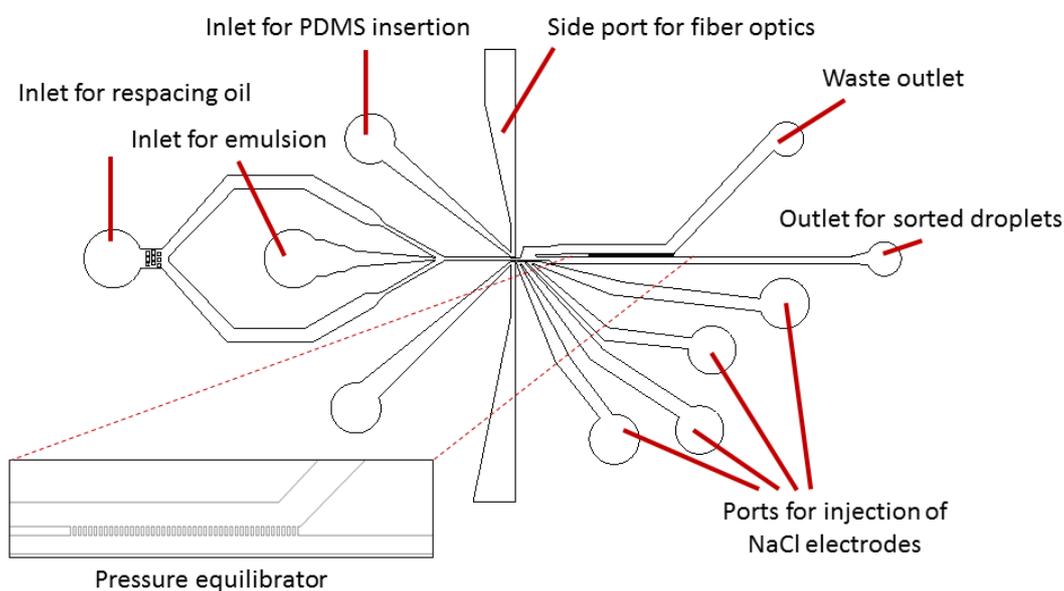
The initial rates  $v_0$  were measured (Table S1, column<sup>a</sup>) for different dilutions of cell lysate (*E. coli* BL12 (DE3)) expressing wtPheDH. The quantity of pure enzyme corresponding to each initial rate was determined according to a titration curve (Fig. S14). The number of molecules for each enzyme concentration was then calculated and divided by the number of cells to assess the average number of enzyme molecules per cell (Table S1, columns<sup>b,c,d,e</sup>).

## **References**

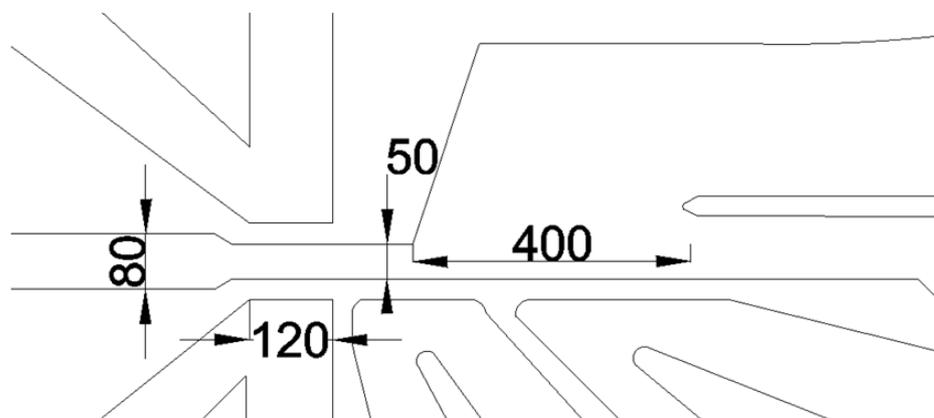
1. Zhao H, Giver L, Shao Z, Affholter JA, & Arnold FH (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat Biotechnol* 16(3):258-261.
2. Colin P-Y, *et al.* (2015) Ultrahigh-throughput discovery of promiscuous enzymes by picodroplet functional metagenomics. *Nature Communications* 6:10008.
3. Cornish-Bowden A (2004) *Fundamentals of Enzyme Kinetics* (Portland Press (London)) p 438.

### 3. Supplementary Figures and /Tables

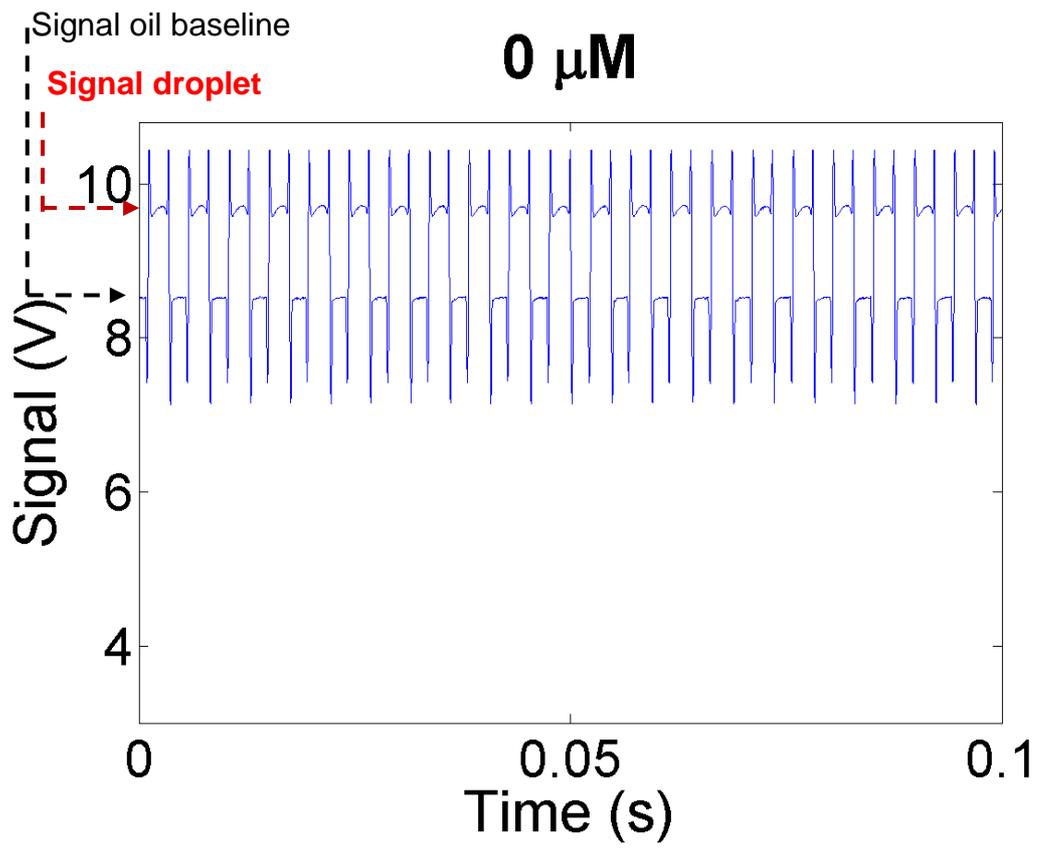
**A**



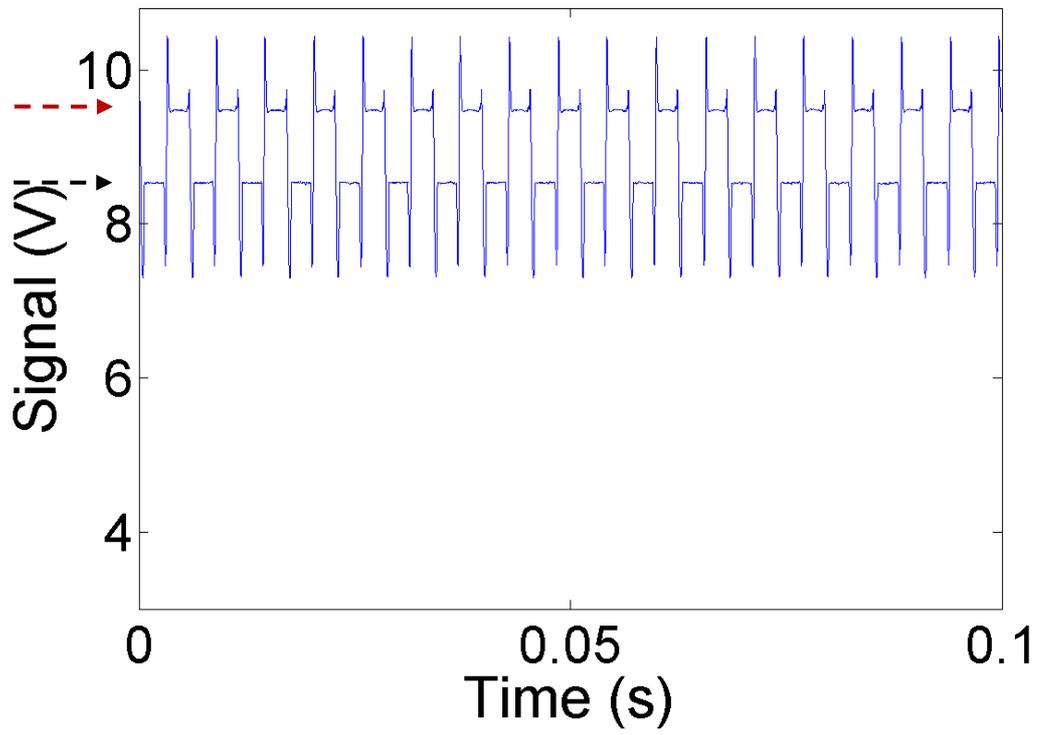
**B**



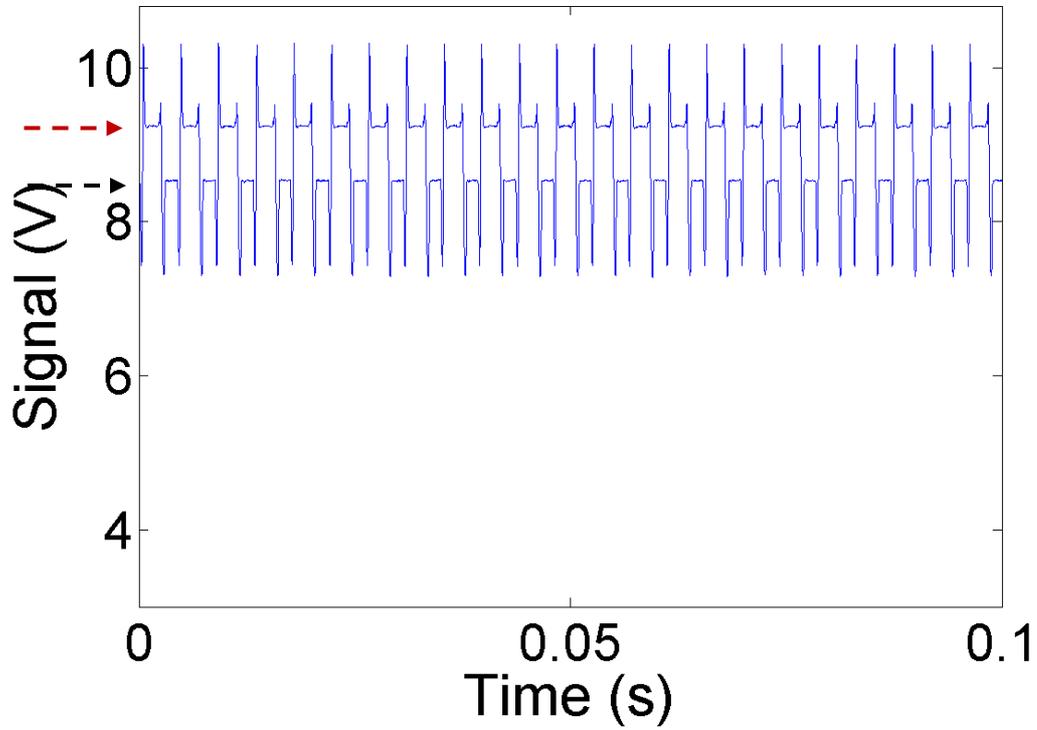
**Fig. S1 Chip design. A.** CAD drawing of the absorbance sorting chip displaying the side access ports for the fiber optics, inlets for spacing oil and emulsion as well as electrode channels and outlets for both waste and sorted droplets. **B.** Close-up view of the sorting junction with dimensions quoted in microns.



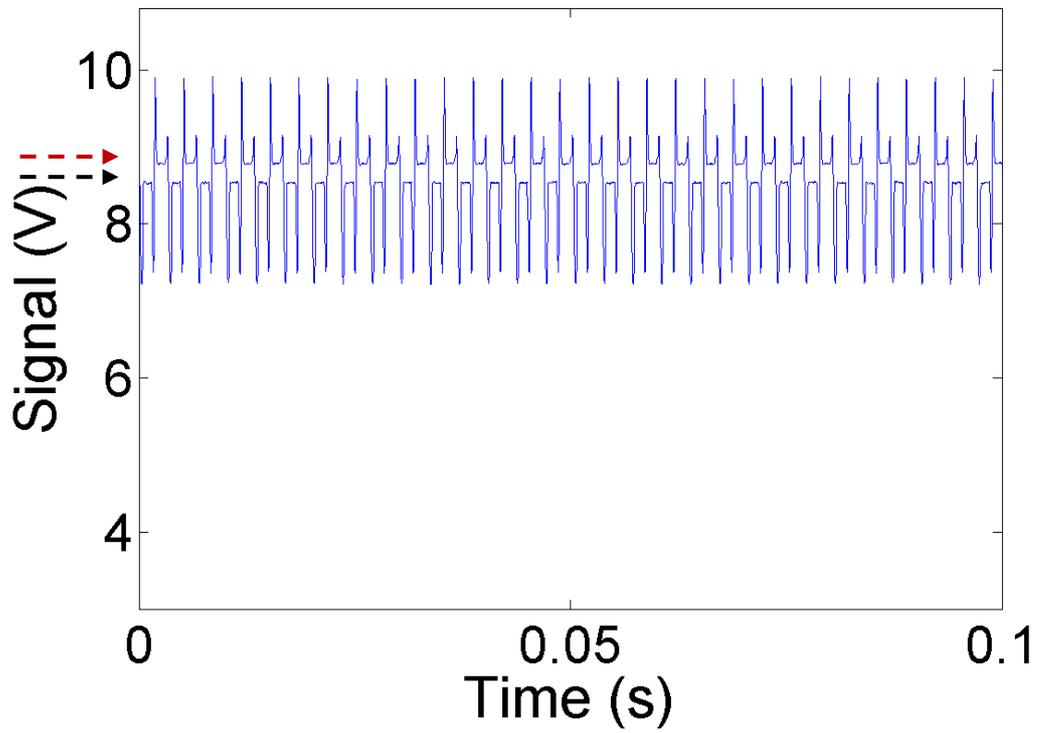
**100  $\mu\text{M}$**



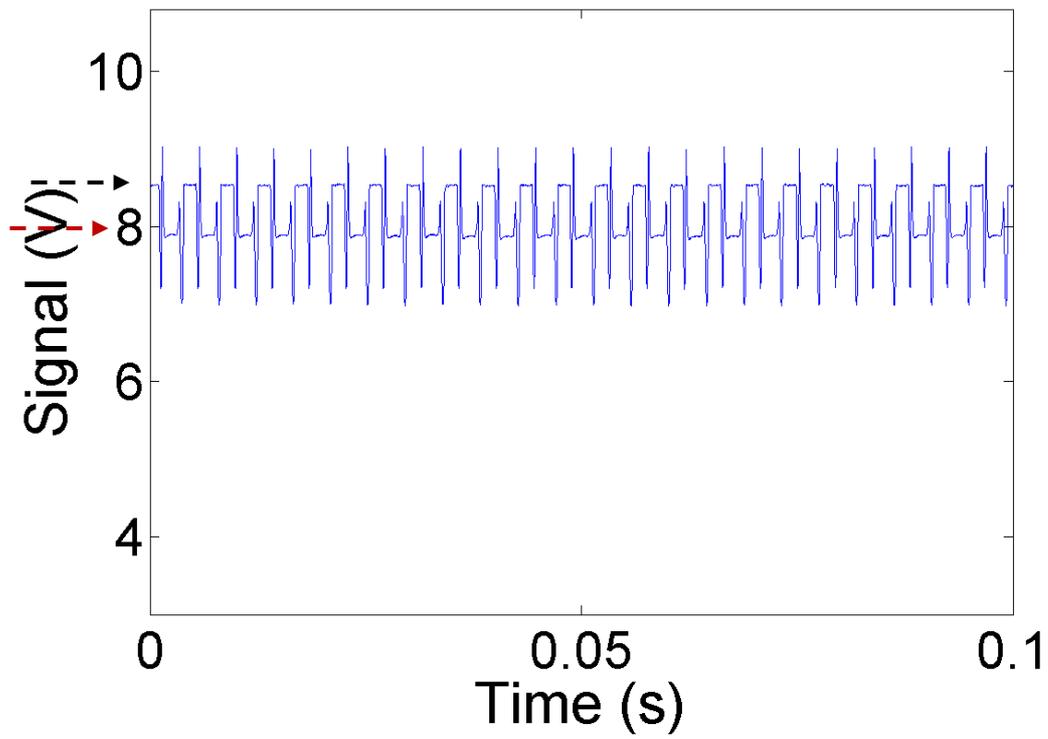
**200  $\mu\text{M}$**



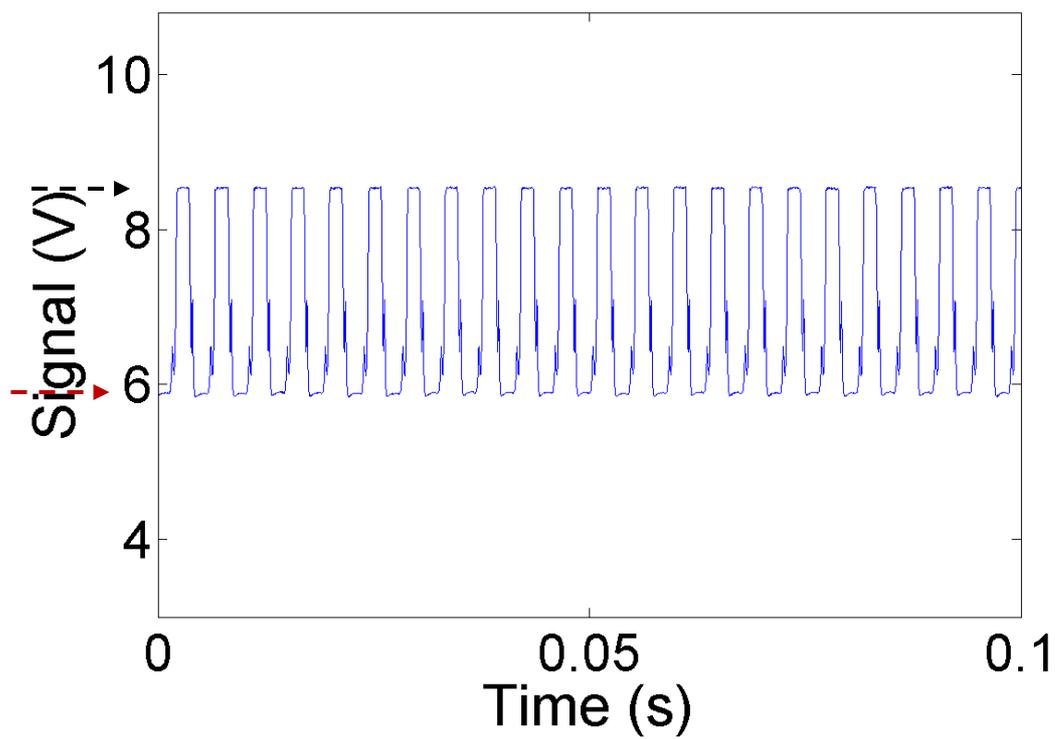
**500  $\mu\text{M}$**



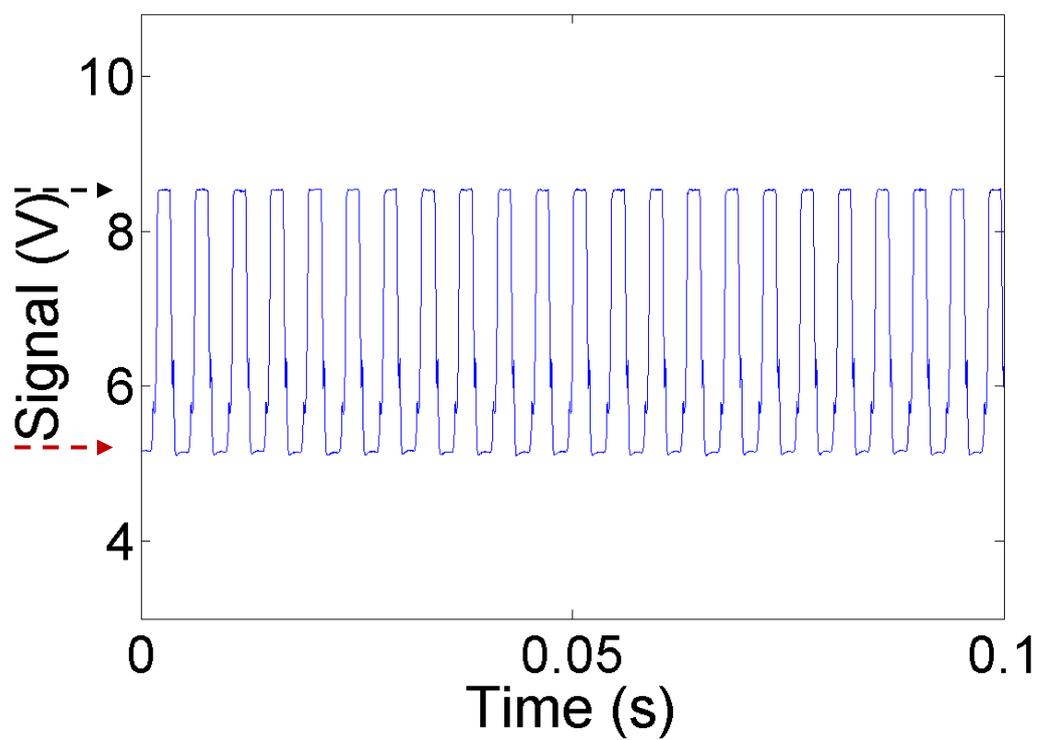
**1 mM**



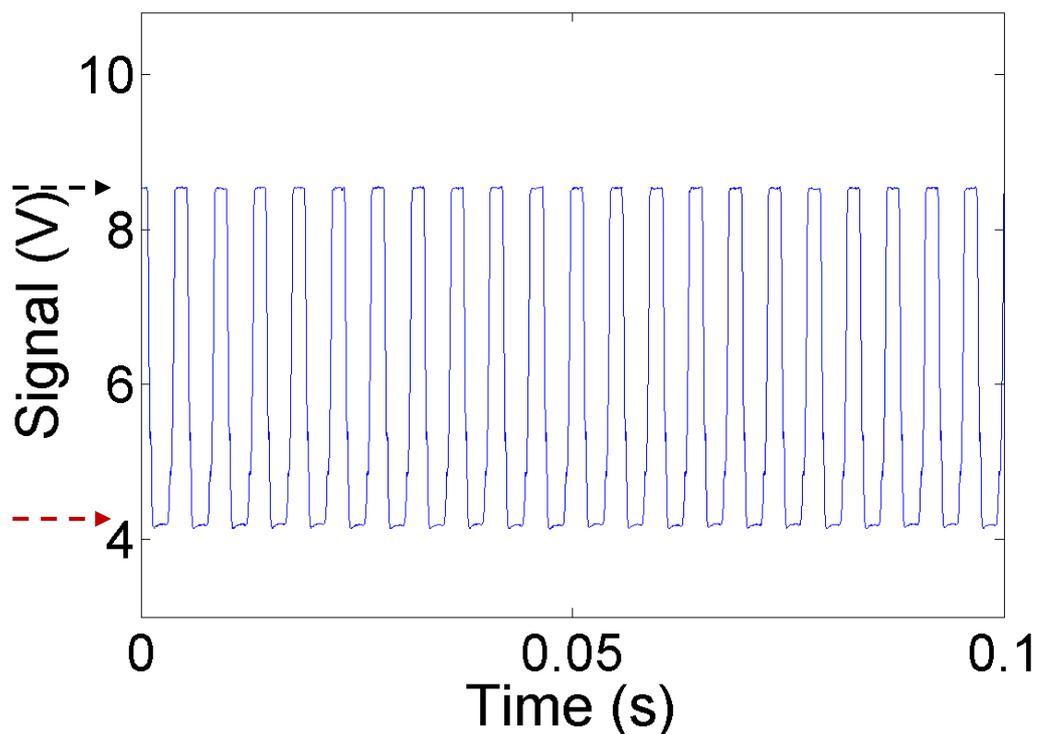
**2 mM**



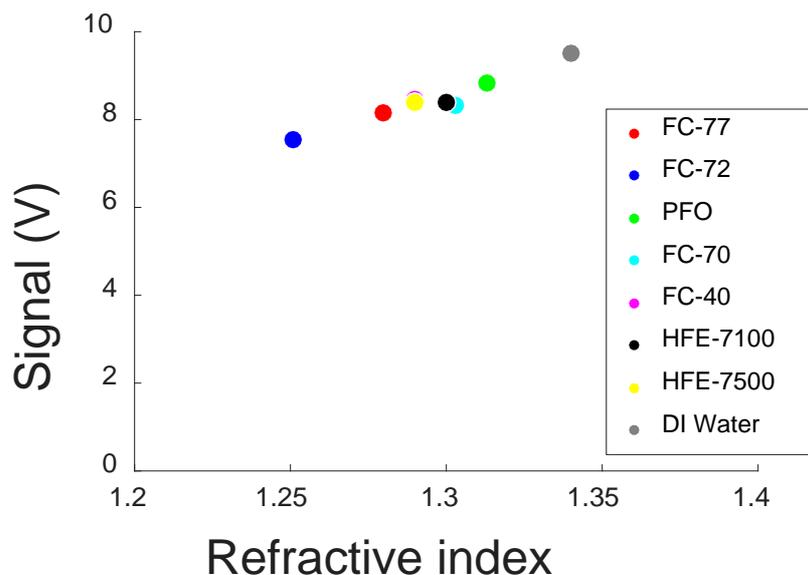
**3 mM**



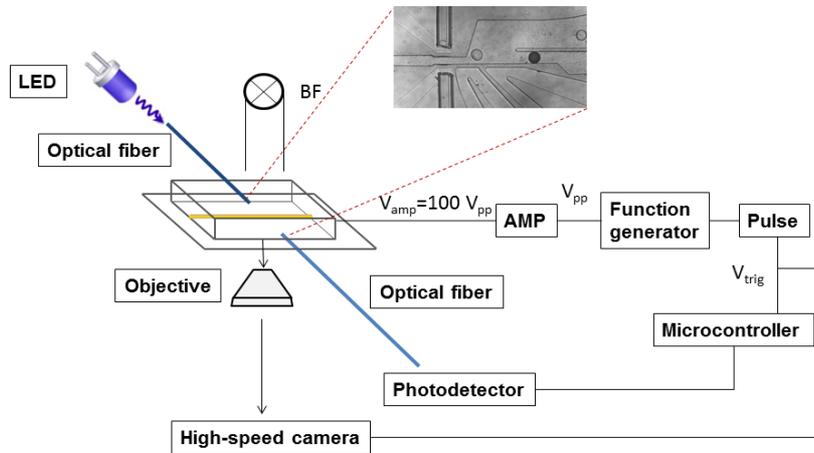
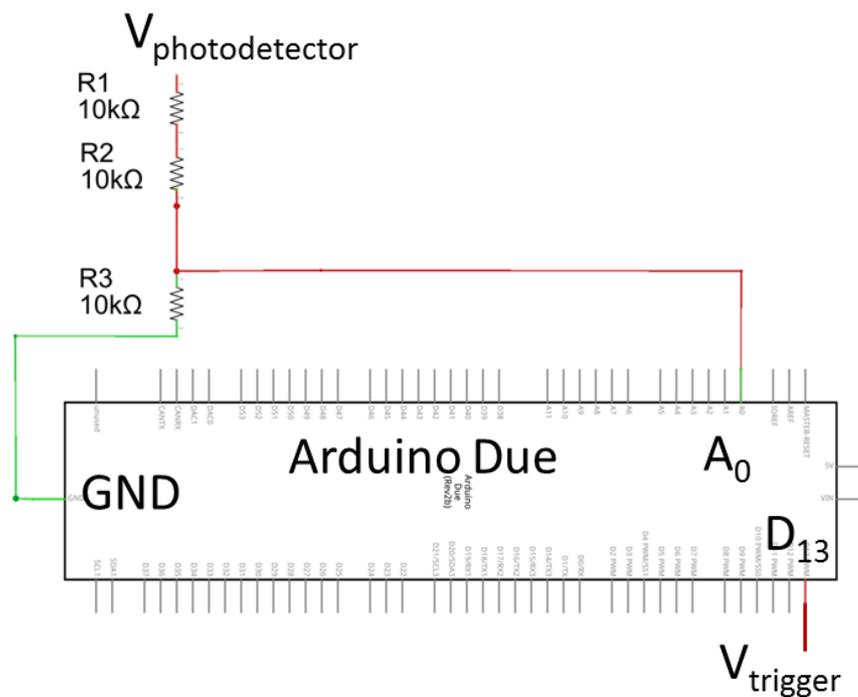
## 4 mM



**Fig S2.** Time traces for calibration of the absorbance detector. Oil baseline and droplet signal are represented by dashed black and red arrows, respectively. Between 500  $\mu\text{M}$  and 1 mM the positions of the black and red arrows are inverted, as the oil baseline is constant, but the droplet signal decreases (with increasing product concentration). The regularity of the time traces suggests stability of the readout and therefore suitability of quantitative assessments.

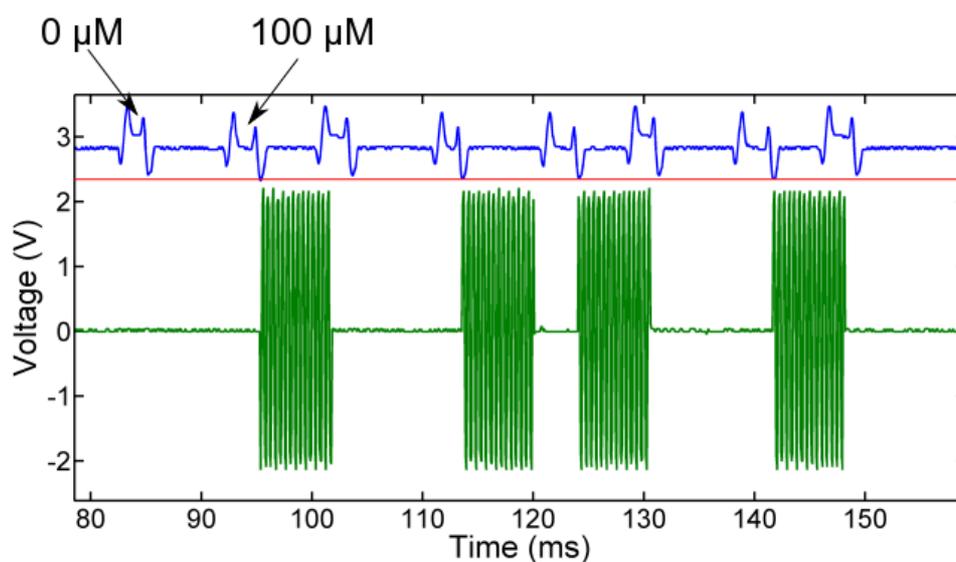


**Fig S3.** Ascribing the invariant signal monitored by the photodiode detectors to the oil phase (marked with black arrows on the left of all panels in Fig. S2) may be initially puzzling, because it seems to suggest that the colourless oil absorbs with similar intensity as the evidently coloured WST-1 formazan dye. However, we interpret the output signal (in V) as a combination of the total amount of light directed towards the detection fiber for a given solvent (which depends mainly on how much it scatters light, including scatter as a function of its refractive index) as well as the amount of light absorbed. To probe the contribution of the refraction index to the output signal we measured the voltage signal for a number of pure carrier oil phases at constant LED power at 455 nm and plotted this signal as a function of their refraction index. The clear correlation between the refractive index of the phase used and the detected output intensity at constant LED power observed in this figure explains why HFE-7500 appears to have a different signal than de-ionized water, even though it does not absorb light at 455 nm. Using one oil phase, only the concentration of the chromophore is determining the signal output and the linearity of this signal against product concentration (Fig. 2c) suggests that it is faithfully reflecting product absorbance (refraction being constant for the buffer used).

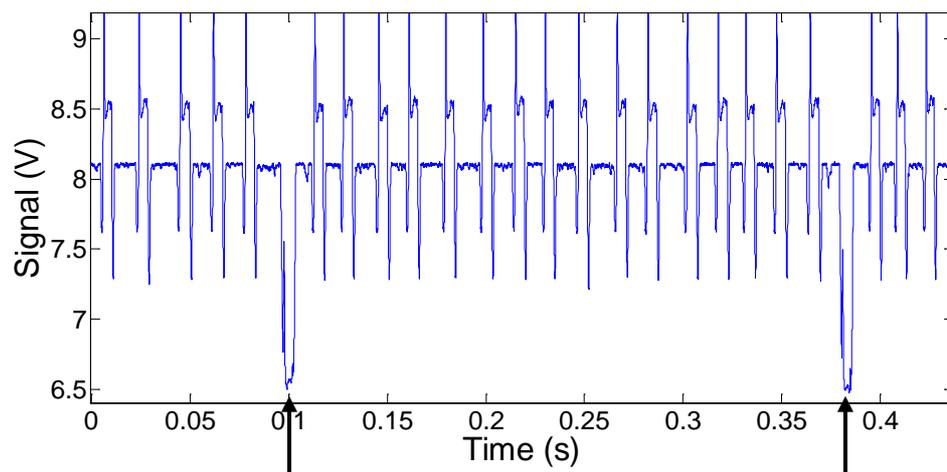
**A****B**

**Fig. S4** Schematic of the AADS setup: the measurement of droplet transmittance is performed *via* the voltage measured by the photodetector. **A.** An Arduino Due microcontroller converts the analog voltage to a digital signal that can be used to generate a trigger signal. This signal will activate a 5 V pulse of typical width 5 ms which in turn triggers a function generator generating a 10 kHz square wave of amplitude 600  $V_{pp}$ . **B.** Wiring of the Arduino Due: the input signal from the

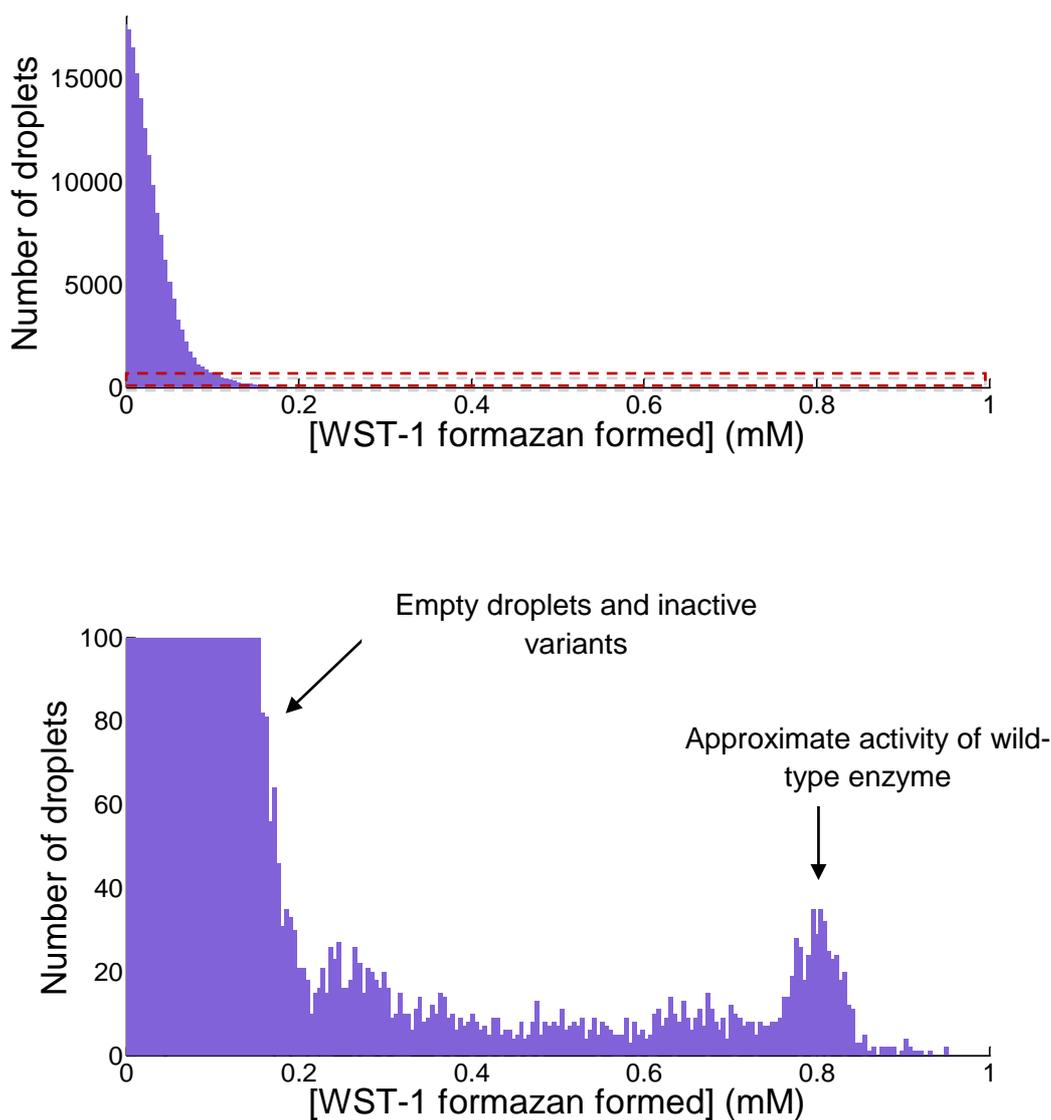
photodetector is divided 3 fold by resistors (to match the maximum voltage tolerable for the microcontroller board, i.e. 3.3V), and connected to an analog input of the Arduino ( $A_0$ ). The trigger signal is exported to a digital pin ( $D_{13}$ ) that is connected to the pulse generator.



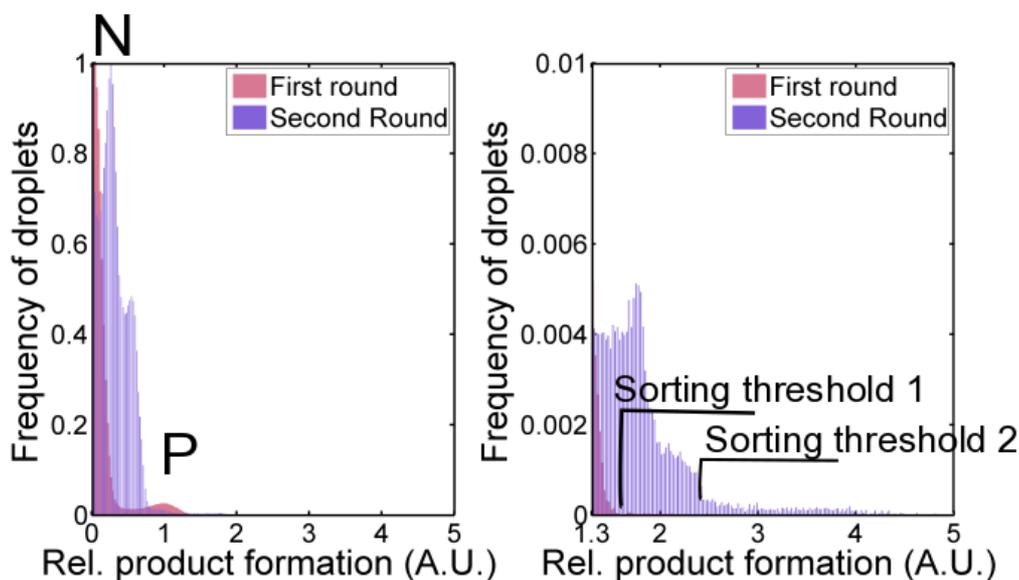
**Fig. S5.** Example of the raw signal from the photodetector measuring droplets containing 0 or 100  $\mu\text{M}$  WST-1 formazan (in blue, divided by 3) and the corresponding AC wave used to trigger sorting before 100x amplification (in green). The sorting threshold is indicated by the red line. Only the droplets containing 100  $\mu\text{M}$  WST-1 formazan result in a triggered signal.



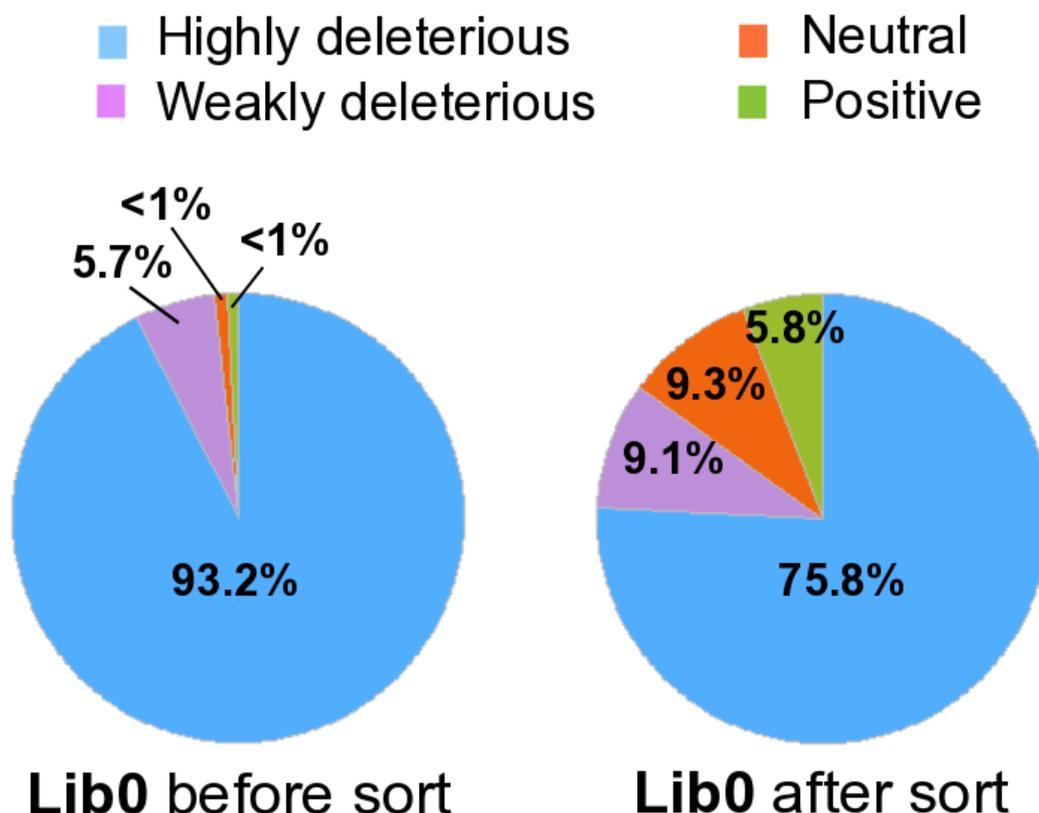
**Fig. S6.** Time-dependent single cell lysate in droplets for the reaction of wtPheDH. Example read-out after 4 hours incubation at an occupancy 0.1 cell/droplet. Black arrows indicate droplets containing a single-cell and where a reaction product is apparent.



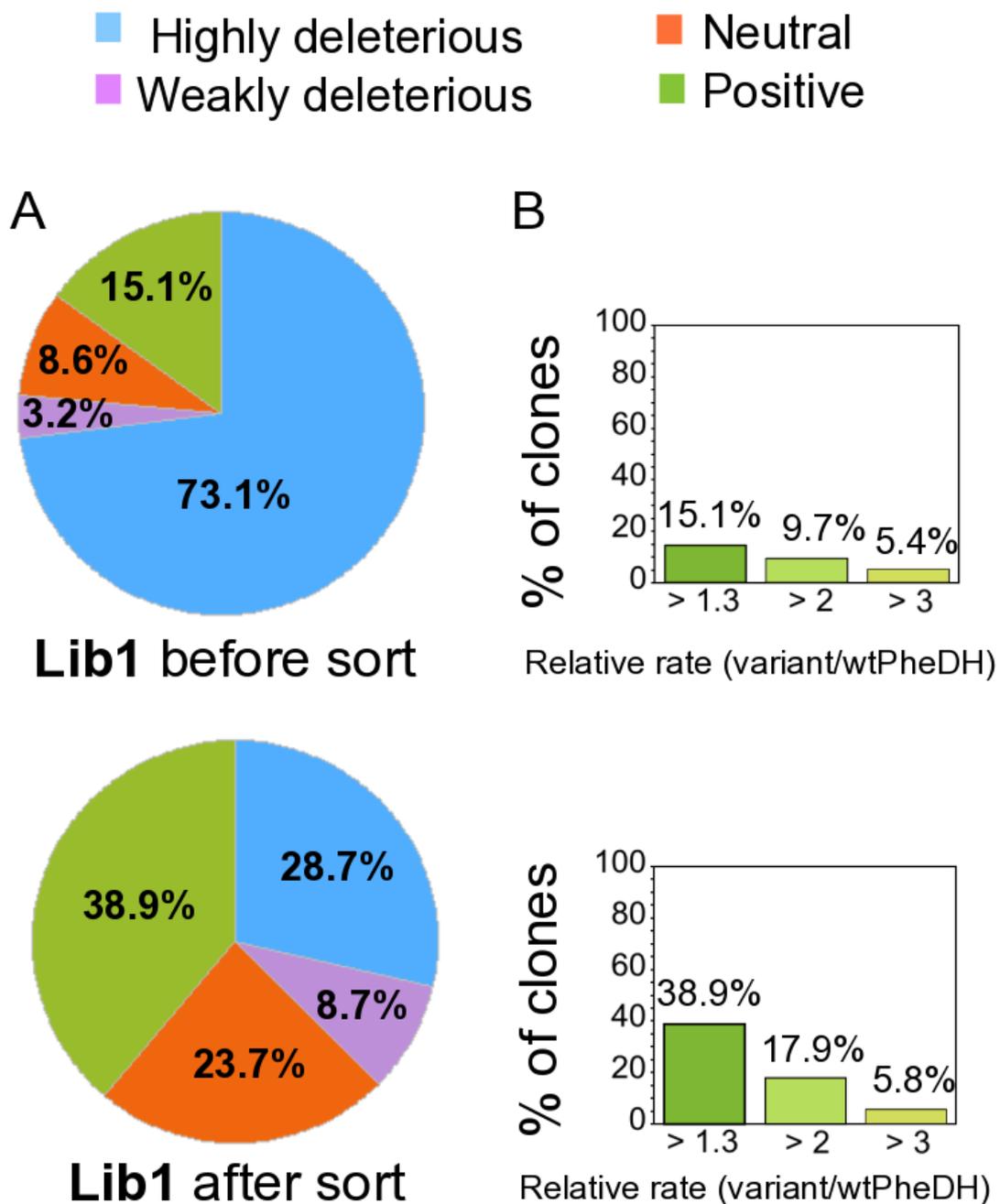
**Fig. S7.** Distribution function of a screening at  $[S] = 1$  mM. A library generated by ep-PCR starting with PheDH<sup>V26I/N122S/L193M/T339I</sup> as the template was screened by AADS. A clear separation between active and inactive mutants is evident, suggesting that selections at relatively low substrate concentrations (1/10 of the substrate concentration used elsewhere in this paper) are feasible. For such selections, each of the bars inbetween the mutants with near wild-type and inactive could be chosen as a threshold (as in Fig. S7, right panel). *Conditions:* [glycine-KOH buffer] = 100 mM, pH 10, 20 °C, saturating concentrations of  $[NAD^+] = 5$  mM.



**Fig. S8.** Distribution function of **Lib0** (EpPCR:  $5 \times 10^5$  transformants encapsulated into  $10^6$  droplets) in pink and **Lib1** (shuffled variants:  $1 \times 10^5$  transformants encapsulated into  $10^6$  droplets) in purple; libraries screened for turnover of L-phenylalanine. The two histograms are derived from approximately 1 million droplets each. Poisson distribution dictates that most droplets (60% for the first histogram, and 90% for the second) do not contain a cell, so they form the left peak (N peak) together with droplets containing an inactive enzyme variant. The smaller peak on the right (P peak) and its perimeter corresponds to wtPheDH or variants with mutations that are silent or neutral. Low frequency peaks with higher product formation correspond to either multiple encapsulated cells (10% for the first histogram and 0.5% for the second, following Poisson statistics) or improved variants (0.01% after screening for each round). The arrows represent the sorting thresholds, corresponding to 1.5 and 2.5-fold improvements relative to the parental enzyme for the first and the second round of directed evolution, respectively.



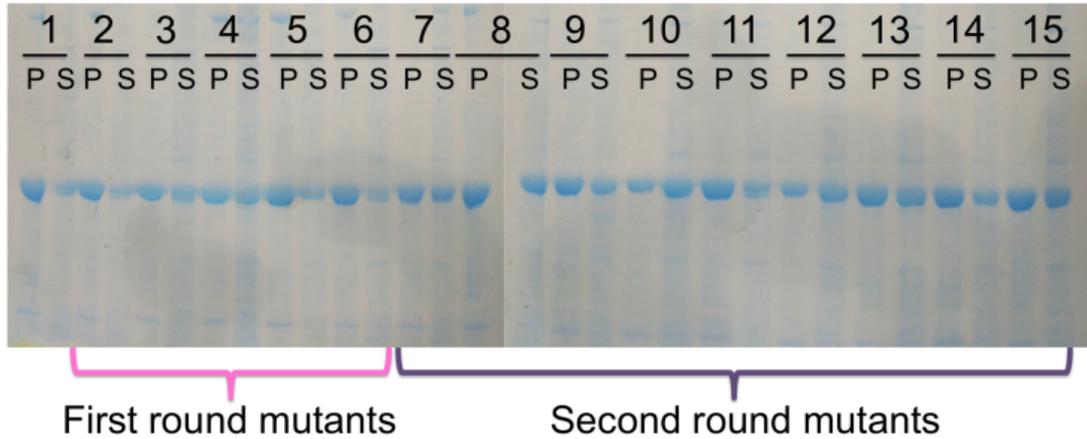
**Fig. S9.** Distribution functions of the dehydrogenase activities measured in cell lysate for **Lib0** before and after sorting with AADS. The activities are expressed relative to the activity of wtPheDH in cell lysate and assigned to four categories: (1) highly deleterious mutants (i.e. those with a >2-fold decrease in activity compared to wtPheDH), (2) weakly deleterious mutants (i.e. those with a 2-1.3 fold decrease), (3) neutral mutants (i.e. those ranging between a 1.3 fold decrease and a 1.3 fold increase), (4) positive mutant (i.e. those with a >1.3 fold increase). The increase in positives with increased activity and the decreasing fraction of mutants with deleterious mutations suggests that AADS selects for a catalytically relevant criterion (i.e. product formation), providing evidence for the utility of this new sorting module for directed evolution.



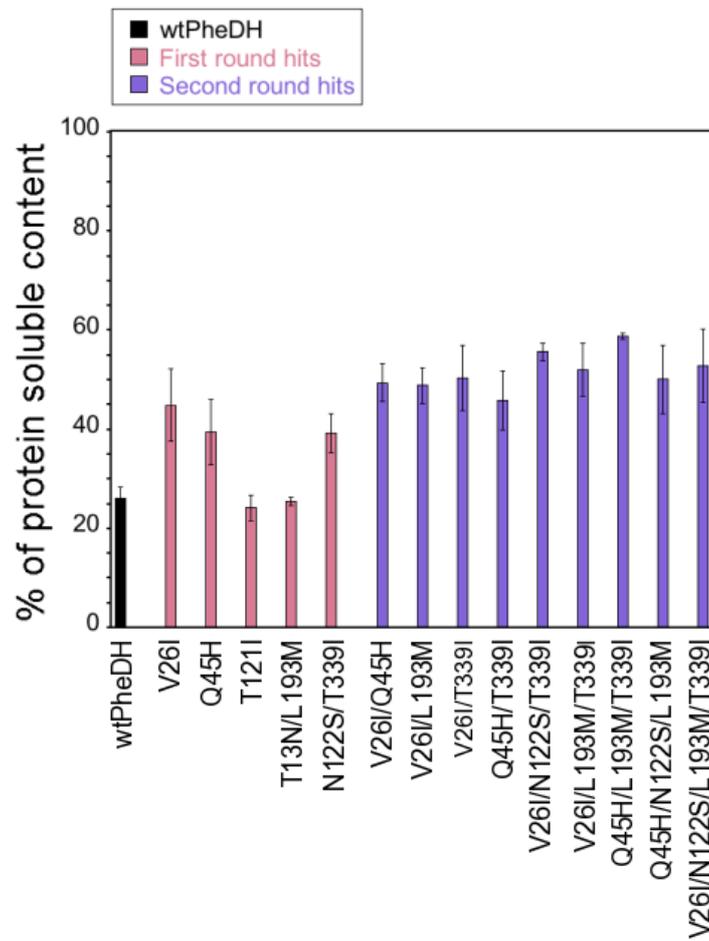
**Fig. S10.** Distribution functions of the dehydrogenase activities (measured as above) for **Lib1** before and after sorting with AADS. **(A)** The selected mutants (before and after sorting) cell lysate activities in 96 well plate are expressed relative to wtPheDH cell lysate activity, and clustered in four categories: the assignment to the four categories has been described in the caption to Figure S9. **(B)** Percentage of clones (from **Lib1**) with a relative activity *higher* than 1.3-, 2- and 3-fold compared to wtPheDH before and after sorting. The activity profile of the sorted population

showed that variants with >2-fold improved relative activity had been enriched by 2-fold during the sorting, reflecting the not very stringent screening regime.

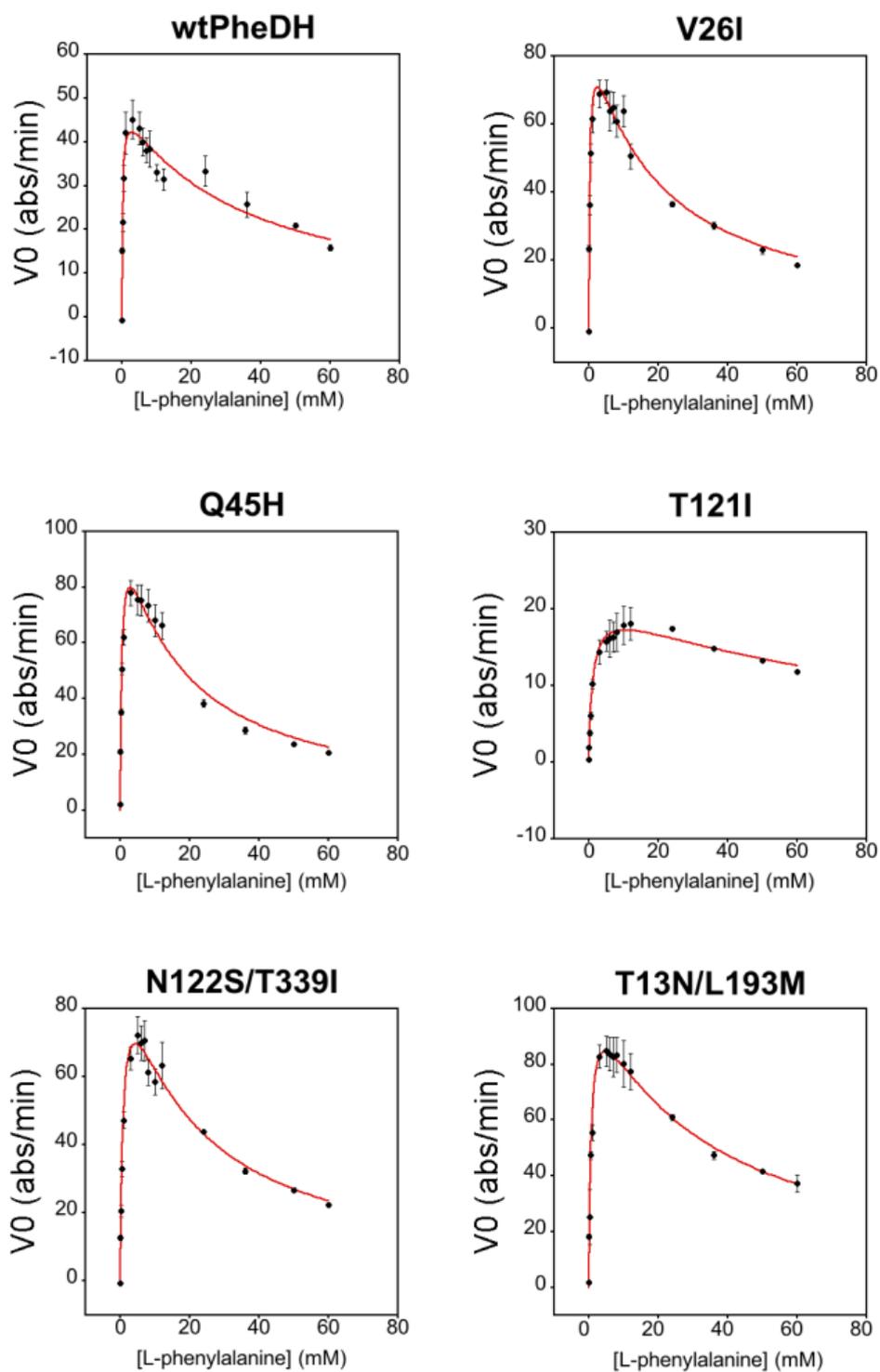
**A**



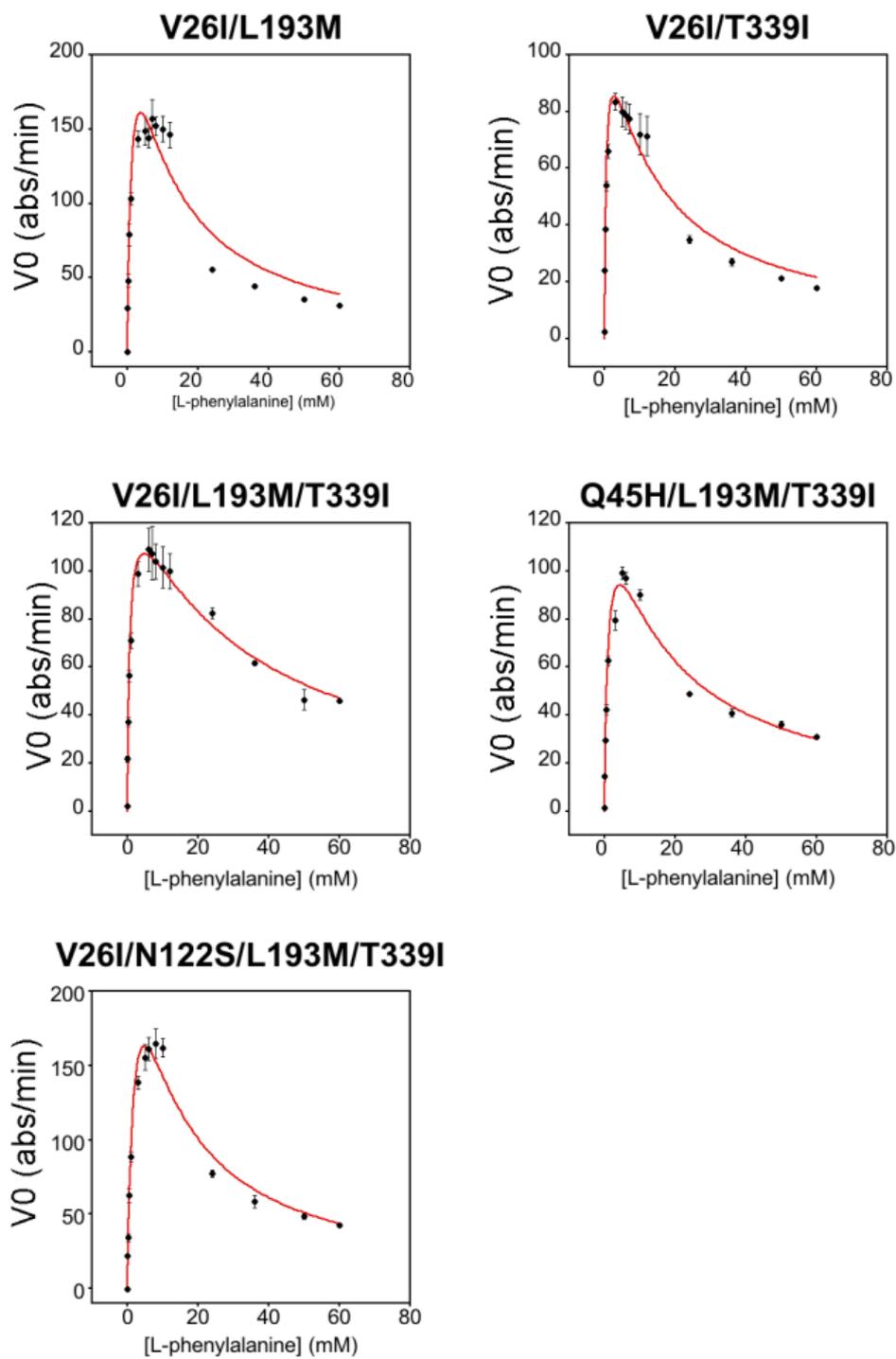
**B**



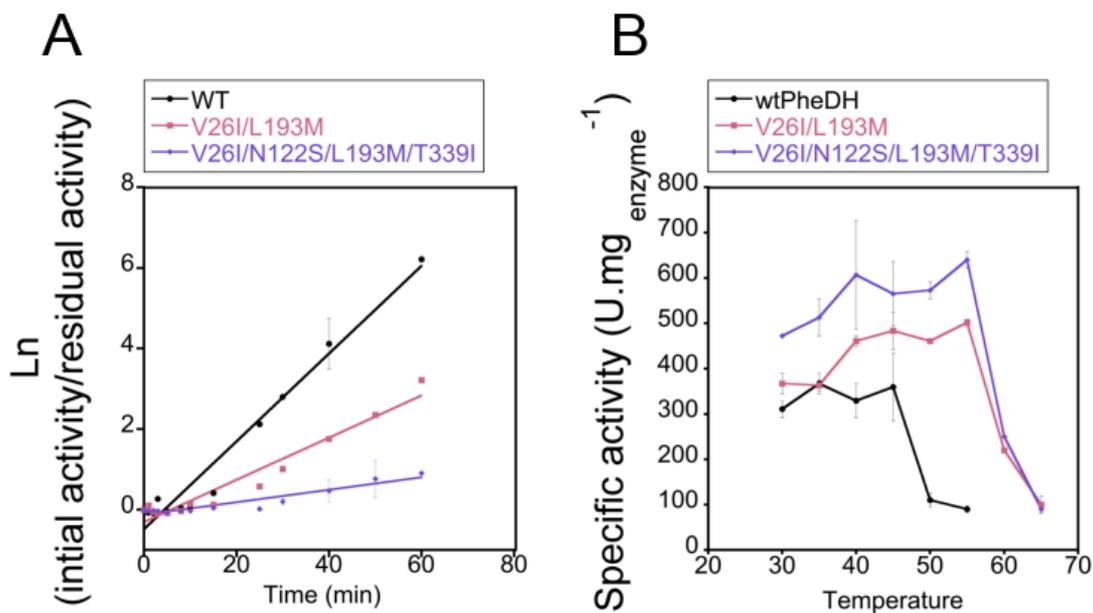
**Fig. S11.** Expression tests of the mutants selected by the first and second round of directed evolution. **(A)** Soluble expression levels of wtPheDH and selected variants. After protein expression, the soluble and insoluble fractions of cell lysates were analyzed by SDS-PAGE (12%). S: soluble fraction, P: pellet. Lane 1: wtPheDH, Lanes 2-6: first round mutants. Lane 2: clone T13N/L193M, 3: clone Q45H, 4: clone V26I, 5: clone N122S/T339I, 6: clone T121I, Lanes 7-15: second round mutants. Lane 7: clone V26I/L193M, 8: clone V26I/N122S/T339I, 9: clone V26I/N122S/L193M/T339I, 10: clone V26I/Q45H, 11: clone Q45H/N122S/L193M, 12: clone V26I/T339I, 13: Q45H/T339I, 14: clone Q45H/L193M/T339I, 15: clone V26I/L193M/T339I. **(B)** The percentage of enzyme in the soluble fraction was determined by the relative intensities of the supernatant and pellet bands. Error bars represent SEM of three independent measurements.



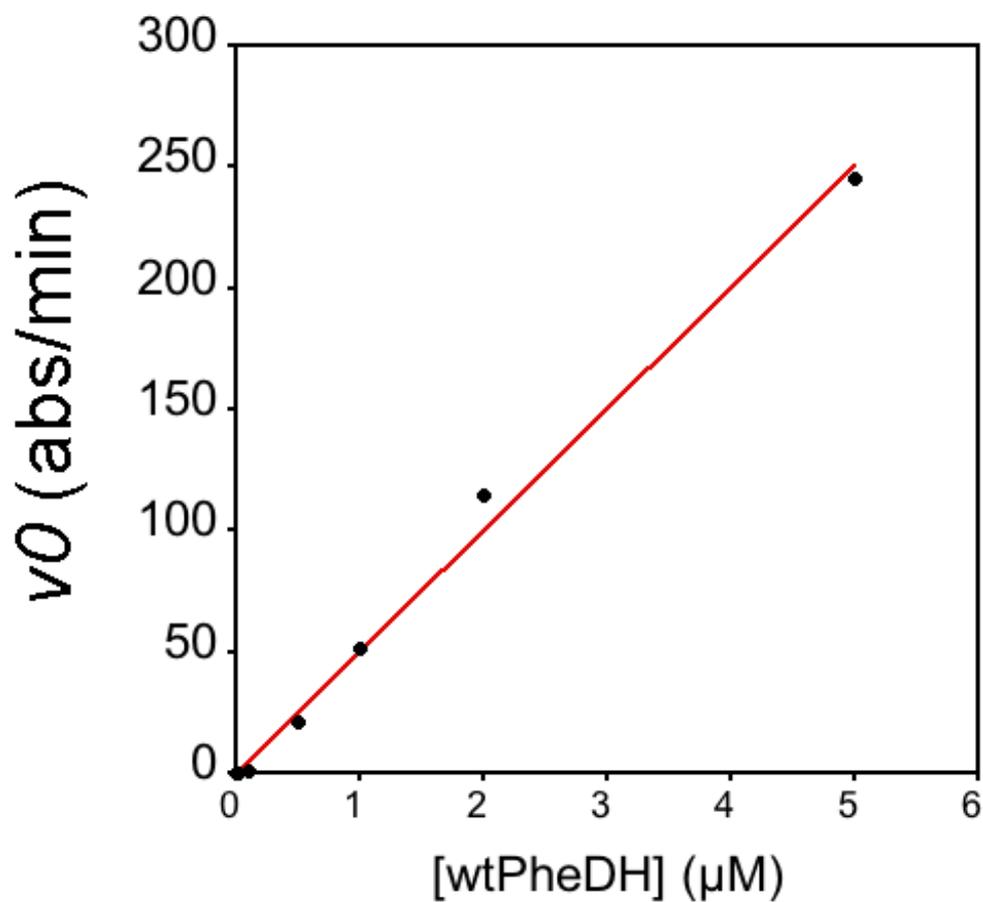
**Fig. S12A.** Michaelis Menten plots for first round mutants. *Conditions:* [glycine-KOH buffer] = 100 mM, pH 10, 20 °C, saturating concentrations of  $[NAD^+] = 5$  mM.



**Fig. S12B.** Michaelis Menten plots for second round mutants. *Conditions:* [glycine-KOH buffer] 100 mM, pH 10, 20 °C, saturating concentrations of  $[NAD^+] = 5$  mM.



**Fig. S13.** (A) Thermal stability of wtPheDH and variants V26I/L193M and V26I/N122S/L193M/T339I at 50°C. wtPheDH and its variants were incubated at 50 °C, and samples were removed at regular time intervals and assayed for NADH production at 20 °C. (B) Temperature dependency of wild-type and variants V26I/L193M and V26I/N122S/L193M/T339I. Specific activities were measured over a range of temperatures from 30 °C to 65 °C in [glycine-KOH buffer] = 100 mM, pH 10, with 10 mM [L-Phe] and 5 mM [NAD<sup>+</sup>].



**Fig. S14.** Correlation of enzyme initial rates  $v_0$  (wtPheDH) with enzyme concentration (measured with the spectrophotometer at 340 nM with [L-Phe] = 10 mM and [NAD<sup>+</sup>] = 1 mM; [glycine-KOH buffer] = 100 mM, pH 10, T = 20 °C).

**Table S1.** Calculation of the number of enzyme molecules produced per cell for wtPheDH.

| Dilution of lysate | $v_0$ (abs/min) <sup>a</sup> | Quantity of pure enzyme (wtPheDH) ( $\mu\text{M}$ ) <sup>b</sup> | Number of molecules/L <sup>c</sup> | Number of cells/L <sup>d</sup> | Number of molecules/cell <sup>e</sup> |
|--------------------|------------------------------|--|------------------------------------|--------------------------------|---------------------------------------|
| <b>1/1000</b>      | 2.88                         | 0.48   | $2.9 \times 10^{16}$               | $5 \times 10^{10}$             | $5.8 \times 10^5$                     |
| <b>1/100</b>       | 55.33                        | 1  | $6.6 \times 10^{17}$               | $5 \times 10^{11}$             | $1.3 \times 10^6$                     |
| <b>1/10</b>        | 222.47                       | 4.4  | $2.7 \times 10^{18}$               | $5 \times 10^{12}$             | $5.3 \times 10^5$                     |
| <b>Average</b>     |                              |  |                                    |                                | $8.1 \times 10^5$                     |

<sup>a</sup> Spectrophotometrically determined  $v_0$  of NADH production at 340 nm. <sup>b</sup> Determined based on the titration shown in Fig. S14. <sup>c</sup> Calculated: (quantity of pure enzyme (in  $\mu\text{M}$ )<sup>b</sup> \*  $N_A$  (Avogadro constant))/ $10^6$ ; <sup>d</sup> Measured:  $\text{OD}_{600}$ , <sup>e</sup> Calculated: Number of molecules/L<sup>c</sup> / Number of cells/L<sup>d</sup>.

The detection limit of the sorter is [WST-1 formazan]=10  $\mu\text{M}$ , which corresponds to  $6.022 \times 10^{18}$  molecules/L (=10  $N_A \times 10^{-6}$  =  $10^{-5} N_A$ ; with  $N_A = 6.022 \times 10^{23}$  molecules). The volume of one droplet is 180 pL, which means that at least  **$1.08 \times 10^9$**  molecules WST-1 formazan (=6.022  $\times 10^{18}$  molecules/L x 180  $\times 10^{-12}$  L) must be present in one droplet.

Therefore,  $1.08 \times 10^9$  molecules WST-1 formazan need to be turned over by  $8.1 \times 10^5$  enzyme molecules, which means that 1333 turnovers (=1.08  $\times 10^9$  /  $8.1 \times 10^5$ ) are necessary per enzyme molecule to generate 10  $\mu\text{M}$  of product WST-1 formazan.

**Table S2.** Michaelis-Menten parameters of the purified mutants from first and second round screens. Initial rate data ( $[S]=[L\text{-Phe}] = 0\text{-}60$  mM) was plotted against substrate concentrations and fit to an equation containing a term for substrate inhibition ( $v = k_{cat} [S] / (K_M + [S] + ([S]^2 / K_i))$ ). *Conditions:* [glycine-KOH buffer] = 100 mM, pH 10, 20 °C, saturating concentrations of  $[NAD^+] = 5$  mM. Measured in a plate reader (SpectraMax 190, Molecular Devices).

| <b>Mutant</b>              | $K_M$<br>(mM)    | $k_{cat}$<br>(s <sup>-1</sup> ) | $K_i$<br>(mM) | $k_{cat}/K_M$<br>(10 <sup>3</sup> M <sup>-1</sup> s <sup>-1</sup> ) | % protein<br>soluble<br>contents |
|----------------------------|------------------|---------------------------------|---------------|---|----------------------------------|
| Wt                         | 0.23 ±0.043      | 16.4 ±2.2                       | 34 ±5.8       | 71.4 ±3.7   | 29 ±2.1                          |
| <b>First round</b>         |                  |                                 |               |   |                                  |
| V26I                       | 0.32 ±0.04       | 24 ±3.6                         | 18 ±2         | 75±7  | 52 ±7.3                          |
| Q45H                       | 0.44 ±0.14       | 29.4 ±5                         | 17 ±5         | 66.9±6.12   | 46 ±6.2                          |
| T13N/L193M                 | 0.73 ±0.07       | 29.7 ±3.4                       | 30 ±2.8       | 40.7±9.89   | 27 ±2.5                          |
| T121I                      | 21 ±0.07         | 2.8 ±0.73                       | 90 ±13        | 0.13±0.81   | 25 ±0.8                          |
| N122S/T339I                | 1 ±0.14          | 30 ±5.6                         | 18 ±2.4       | 30.3±4.54   | 36 ±3.8                          |
| <b>Second round</b>        |                  |                                 |               |   |                                  |
| V26I/L193M                 | 1.5 ±0.57        | 40 ±4                           | 9 ±3.4        | 26.7±6.45   | 46 ±3.8                          |
| V26I/N122S/<br>L193M/T339I | 2.6 ±0.03        | 45 ±5.6                         | 10 ±3.8       | 17.3±1.45   | 58 ±6.5                          |
| V26I/L193M/T33<br>9I       | 0.66 ±0.16       | 17 ±1.02                        | 35.7 ±7.9     | 26.4±0.62   | 52 ±6.0                          |
| V26I/Q45H                  | 0.24<br>±0.0001  | 9.2 ±1.0                        | 18.7 ±2.0     | 56.7±4.26   | 55 ±1.7                          |
| V26I/T339I                 | 0.58 ±0.04       | 33 ±2.7                         | 13 ±2.7       | 56.8±0.58   | 60 ±0.6                          |
| Q45H/T339I                 | 0.39<br>±0.00039 | 21.9 ±1.0                       | 43 ±11.6      | 56.3±4.26   | 57±6.8                           |
| Q45H/L193M/T3<br>39I       | 1.7 ±0.53        | 23.3 ±1.2                       | 13.1 ±3.8     | 13.7±6.45   | 60 ±7.4                          |

**Table S3.**  $T_m$  of the mutants selected by the two rounds of evolution. Directed evolution increased the  $T_m$  by up to 12.4 °C. Conditions: [MOPS buffer] = 100 mM with NaCl 150 mM, pH 8, and [glycine-KOH buffer] = 100 mM pH 10, [enzyme] = 2  $\mu$ M.

|                        | $T_m$                       |                    | $\Delta T_m$ |              |
|------------------------|-----------------------------|--------------------|--------------|--------------|
|                        | (°C)                        |                    | (°C)         |              |
|                        | <b>pH 8</b>                 | <b>pH 10</b>       | <b>pH 8</b>  | <b>pH 10</b> |
| wtPheDH                | 54.6 ( $\pm$ 0.059)         | 47 ( $\pm$ 0.03)   | -            | -            |
|                        | <b>First round mutants</b>  |                    |              |              |
| V26I                   | 55.2 ( $\pm$ 0.053)         | 47 ( $\pm$ 0.1)    | 0.6          | 0            |
| Q45H                   | 53.4 ( $\pm$ 0.08)          | 43 ( $\pm$ 0.54)   | -1.2         | -4           |
| T13N/L193M             | 57.2 ( $\pm$ 0.094)         | 49 ( $\pm$ 1.9)    | 2.6          | 2            |
| T121I                  | N.D                         | N.D                | N.D          | N.D          |
| N122S/T339I            | 56.3 ( $\pm$ 0.07)          | 48.8 ( $\pm$ 0.15) | 1.7          | 1.8          |
|                        | <b>Second round mutants</b> |                    |              |              |
| V26I/L193M             | 61.7 ( $\pm$ 0.66)          | 50.5 ( $\pm$ 0.12) | 7.1          | 3.5          |
| V26I/N122S/L193M/T339I | 67 ( $\pm$ 2.56)            | 54.7 ( $\pm$ 0.07) | 12.4         | 7.7          |
| V26I/L193M/T339I       | 55 ( $\pm$ 0.3)             | 51 ( $\pm$ 0.05)   | 0.4          | 4            |
| V26I/Q45H              | N.D                         | N.D                | N.D          | N.D          |
| V26I/T339I             | 58 ( $\pm$ 0.27)            | 48.4 ( $\pm$ 0.1)  | 3.4          | 1.4          |
| Q45H/T339I             | 55.6 ( $\pm$ 0.085)         | 48 ( $\pm$ 0.06)   | 1            | 1            |
| Q45H/L193M/T339I       | 42.1 ( $\pm$ 0.08)          | 51.7 ( $\pm$ 0.11) | -12.5        | 4.7          |

**Table S4.** Half-life of inactivation for wtPheDH and mutants V26I/L193M and V26I/N122S/L193M/T339I at 50 °C. Directed evolution increased the half life of inactivation by up to 7.5-fold.

| <b>Mutant</b>          | <b>Half-life<br/>(min)</b> | <b>Half-life<br/>increase<br/>(n-fold)</b> |
|------------------------|----------------------------|--|
| wtPheDH                | 6                          | -  |
| V26I/L193M             | 13                         | 2.2  |
| V26I/N122S/L193M/T339I | 45                         | 7.5  |