Exploration of Enzyme Diversity: High-Throughput Techniques for Protein Production and Microscale Biochemical Characterization

Michal Vasina\textsuperscript{1,2,†}, Pavel Vanacek\textsuperscript{1,2,†}, Jiri Damborsky\textsuperscript{1,2} and Zbynek Prokop\textsuperscript{1,2,*}

Authors’ affiliations:

\textsuperscript{1} Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Kamenice 5, 625 00 Brno, Czech Republic

\textsuperscript{2} International Clinical Research Center, St. Anne's University Hospital, Pekarska 53, 656 91 Brno, Czech Republic

† M.V. and P.V. contributed equally.

* Author for correspondence: Zbynek Prokop, zbynek@chemi.muni.cz, ORCID 0000-0001-9358-4081.
Abstract

Enzymes are being increasingly utilized for acceleration of industrially and pharmaceutically critical chemical reactions. The strong demand for finding robust and efficient biocatalysts for these applications can be satisfied via the exploration of enzyme diversity. The first strategy is to mine the natural diversity, represented by millions of sequences available in the public genomic databases, by using computational approaches. Alternatively, metagenomic libraries can be targeted experimentally or computationally to explore the natural diversity of a specific environment. The second strategy, known as directed evolution, is to generate man-made diversity in the laboratory using gene mutagenesis and screen the constructed library of mutants. The selected hits must be experimentally characterized in both strategies, which currently represent the rate-limiting step in the process of diversity exploration. The traditional techniques used for biochemical characterization are time-demanding, cost, and sample volume ineffective, and low-throughput. Therefore, the development and implementation of high-throughput experimental methods are essential for discovering novel enzymes. This chapter describes the experimental protocols employing the combination of robust production and high-throughput microscale biochemical characterization of enzyme variants. We validated its applicability against the model enzyme family of haloalkane dehalogenases. These protocols can be adapted to other enzyme families, paving the way towards the functional characterization and quick identification of novel biocatalysts.

Keywords: enzyme diversity, genomic databases, metagenomics, directed evolution, high-throughput protein production, microscale characterization, microfluidics

1. Introduction

Enzymes are increasingly used in industrial and biomedical applications thanks to their ability to operate under mild reaction conditions in aqueous media (Sheldon & Woodley, 2018). Yet the widespread use of enzymes is often limited by their low catalytic efficiency, narrow substrate specificity, insufficient stability under operating conditions, or poor stereo- and/or regioselectivity (Reetz, 2013). Therefore, it is essential to either discover novel enzymes or to improve the existing enzymes by techniques of protein engineering (Wahler & Reymond, 2001) (Fig. 1).

Fig. 1 here.

Nature represents an inexhaustible source of biocatalysts with around 250 million non-redundant protein sequences available in the public databases (release 2020-02-10) (Sayers et al., 2020), out of which only about 560,000 are experimentally characterized (release 2020_01) (“UniProt,” 2019). Bioinformatic tools and genome mining approaches, together with low-cost de novo gene synthesis, enable fast in-silico screening of this massive unmapped sequence space and selection of best candidates for experimental characterization (Lin et al., 2019). Recently we released the publicly accessible genome mining web-server EnzymeMiner (https://loschmidt.chemi.muni.cz/enzymeminer/), which identifies putative members of the studied enzyme families, facilitates their prioritization and rationally select the
candidates for experimental characterization (Hon et al., 2019). The server is extremely easy to use. It provides two conceptually different ways to define the input of the workflow using: (i) curated sequences from the UniProtKB/Swiss-Prot database, and (ii) custom sequences (Fig. 2). The output of the calculation is the target selection table, which is organized into eleven sheets that summarize the results from different perspectives. The table can be filtered using solubility and identity sliders, and transmembrane and extra domain exclusion switches.

Another option to find interesting enzyme activities in nature is to screen metagenomic libraries, notably with ultra-high-throughput microfluidic sorting devices (Mair et al., 2017). Although most of the enzymes discovered in this way are not suitable for the important catalytic processes, there were many industrially relevant hits (Berini et al., 2017), mostly coming from extremophiles (Madhavan et al., 2017). To survive under extreme conditions, extremophilic organisms must have evolved a unique cellular metabolism, harboring enzyme repertoire with corresponding physiological and biochemical properties. Finally, these discovered enzymes can represent a good starting point for subsequent protein engineering (Trudeau & Tawfik, 2019).

Directed evolution is the method-of-choice for protein engineers to explore man-made diversity. Directed evolution improves a target enzyme property via a combination of mutagenesis and functional screening. Over the 30 years of its existence, directed evolution has developed in a powerful technique (Moore et al., 2018). The randomness of the mutagenesis was reduced by using smart libraries (Moore et al., 2018; Sebestova et al., 2014; Sumbalova et al., 2018), employing reduced amino acid alphabets (Li et al., 2019) and recently also machine learning (K. K. Yang et al., 2019). Similarly, classical screening techniques using agar-plates and microtiter plates (offering medium throughput of $10^3 – 10^5$ assays/day) (Markel et al., 2020) were gradually outperformed by ultra-high-throughput techniques ($10^6 – 10^9$ assays/day) (Bunzel et al., 2018), for example by microfluidic droplet sorting devices (Baret et al., 2009; Gielen et al., 2016).

The resulting set of genes pre-selected based on the genome mining approach or pre-screened from directed evolution studies needs to be further produced and biochemically characterized. The size of such a set shall fall in the range of tens to a hundred genes/variants. The conventional protein production and characterization techniques suffer from low-throughput and relatively high material consumption (Vanacek et al., 2018). Here outstands the need for the combination of high-throughput microscale techniques for both robust production and microscale characterization of novel biocatalysts. (Fig. 3).

The protocols for protein production do not require investments in any specific equipment, whereas the enzyme characterization protocols are designed to reach the maximum throughput. Thus, some protocols will require commercial or customized analytical instruments, robotic systems, and microfluidic devices. We validated most methods on the group of model enzymes haloalkane dehalogenases (HLDs, E.C 3.8.1.5), yet we discuss here also the applicability to other enzyme families.
The old-fashioned practice, where kilograms of organic tissues or large volumes of biological fluids lead to the small quantity of a given protein, is the bygone era. Thankfully, the process of extraction and multi-step purification was replaced by the recombinant DNA technology, enabling the production of a large number of recombinant proteins in plethora types of heterologous host systems. In the post-genomic era, the soluble production of large sets of proteins in a microscale manner is perfectly in line with the current challenges in structural and functional characterization of proteins identified via exploration of natural and/or man-made diversity. Microscale enzyme production consists of 3 steps (i) obtaining target genes, (ii) robust gene expressions, and (III) high-throughput protein purification.

Traditionally, collections of genes require (i) reverse transcription to create cDNA library, and (ii) direct cloning genes from created libraries as a pool into specific vectors. However, there are two limitations to the robust application of the expression libraries. First, untranslated regions at both ends of clones prevent elegant attachment of fusion tags to either end of the proteins of interest. Second, genes of interest must be frequently fished out of an expression library for experimental testing. Although gene retrieval based on PCR is still the most widely used technique for obtaining genes of interest, gene retrieval based on PCR is hard to perform in a high-throughput manner. Obtaining a set of genes requires: (i) a batch of primers need to be designed, and (ii) high-throughput PCR and PCR-product purification. The latter is often accomplished by automated laboratory workstations (Otto et al., 2016). However, problems such as the absence of a band (faint band) in gels, non-specific bands, and primer-dimers may occur after PCR and slow the experimental process (Jia Baolei & Jeon Che Ok, 2016). The exploration of natural diversity requires a robust way for the preparation of gene candidates (10-100) identified in sequence databases. The de novo gene synthesis has opened a new route for obtaining the large set of target genes as it allows overcoming many obstacles commonly imposed using natural templates. The significant advantage of de novo gene synthesis is the coupling with in silico gene design, enabling the optimization of the gene of interest by codon usage to the specific host system. The codon-optimized gene often leads to effective protein translation by the heterologous cellular translational machinery.

Gene synthesis represents an attractive future perspective, while it creates a challenge for proper throughput for protein characterization. These enhancements can be achieved by the further advent of new microarray-based nucleotide synthesis technologies including ink-jet printing (Agilent, Protogene), photosensitive 5′ deprotection (Nimblegen, Affymetrix, Flexgen), photo-generated acid deprotection (Atactic/Xeotron/Invitrogen, LC Sciences), and electrolytic acid/base arrays (Oxamer, Combimatrix/Customarray). Thanks to them, the costs of synthesis might be reduced by 3–5 orders of magnitude to 100 bp per dollar, making de novo gene synthesis a favored approach for obtaining genetic information from ever-increasing DNA databases (Klein et al., 2016; Xiong et al., 2006; Yehezkel et al., 2008). Similar progress has been achieved in molecular cloning, where the large sets of DNA constructs can be generated by ligation-independent cloning methods, allowing simple, fast, and highly efficient cloning in 96-well format (Aslanidis & de Jong, 1990; Vincentelli et al., 2003).
The *Escherichia Coli* (*E. Coli*) dominates among the others as the preferable-choice host system in robust expression of large sets of proteins/variants in a high-throughput manner for structural and functional characterization of proteins (Jia Baolei & Jeon Che Ok, 2016). Small-scale expression screening is most useful for quick evaluation of the folding state of the proteins. The obtained solubility profiles facilitate the prioritization of those proteins/clones, which are soluble enough to proceed with subsequent downstream applications, including microscale characterization. Besides that, the results from small-scale expression screening are often used for further buffer/additives screening to tune the expression conditions, which corresponds with the protein of interest. Evaluation of the folding state is typically assessed by fractionation of a lysate, resulting in the relative quantification of soluble protein by SDS–PAGE (Shih et al., 2002), ELISA (Lesley et al., 2002), mass spectrometry (Chance et al., 2002), or capillary electrophoresis (Hilser & Freire, 1995). Several reporter based assays for high-throughput protein solubility screening have been developed to quickly assess the solubility of recombinant proteins (Dixon et al., 2016; Lockard et al., 2011).

### 2.1 Small-scale Enzyme Expression and Purification

The following small-scale protocol employs established strategies for small-scale expression with minor adjustments for handling medium collection of genes (Saez & Vincentelli, 2014; Vincentelli et al., 2011; Zerbs et al., 2014). It is primarily designed to robustly express the set of genes cloned into an expression vector with an IPTG-inducible T7 promoter using an *E. coli* host. The general features of the protocol can be easily adapted to other organisms and expression systems. Similarly to other expression screening pipelines, small-scale expression screening (1–5 mL of culture) is carried out to identify the optimal conditions that will allow sufficient protein production upon scaling up for follow-up downstream processes such as micro-characterization. It can be performed either in 5 mL of culture volume in individual 50-mL tubes for a small number of gene candidates, in 2 mL of culture volume in 24-deep-well plates for a large number of gene candidates, or eventually in 1 mL of culture volume in 48-deep-well plates for high-throughput screening. The subsequent high-throughput purification is based on the affinity between his-tagged proteins and paramagnetic precharged nickel particles. It employs magnetic beads provided in the MagneHis™ Protein Purification System (Promega), but various purification systems can be used. In general, the main benefit of various strategies of small-scale expression screening is the wide applicability without the need for special equipment. At this point, we would like to refer to an excellent book chapter (Saez & Vincentelli, 2014). The authors describe the high-throughput expression protocols sharing similar features like transformation in a 96-well format, followed by small-scale expression screening using auto-induction medium in a 24-well format, finishing with purification in a 96-well format.

### Solutions and Media

- **Luria-Broth (2xLB) media**: 40 g of LB broth dissolved in 1 L of deionized water, pH adjusted to 7.2 with 1 M sodium hydroxide, sterilized by autoclaving before use
- **Starting Medium**: 2xLB supplemented with 0.5 % glucose and antibiotics
• **Induction Medium**: 2xLB supplemented with 0.6% lactose, 50 mM HEPES (pH 7.4), 0.5 mM IPTG and antibiotics

• **Wash buffer**: 100 mM Tris-HCl, pH 7.5, with 10 mM imidazole

• **Elution buffer**: 100 mM Tris-HCl, pH 7.5, with 500 mM imidazole

• **DNase solution**: 5 units µL⁻¹ of DNase I (New England Biolabs, USA)

**Biological Material**

• **Competent cells**: *E. coli* BL21(DE3) cells (New England Biolabs, USA) or their derivatives containing T7 polymerase under the control of a lactose-inducible promoter

• **dsDNA (~100 ng µL⁻¹)**

**Material and Equipment**

• MagneHis™ Ni-Particles (Promega, USA) or MagneHis™ Protein Purification System (Promega, USA)

• Agar multiwell plates in 6-, 12-, or 24-well formats, or single agar plate

• Centrifuge Sigma 6K-15 (Sigma Laborzentrifugen, Germany) or another preferably refrigerated

• Orbital Incubator Innova 44 (New Brunswick Scientific, USA)

• Stationary incubator NB-205 (N-Biotek, South Korea)

• Air-pore tape sheet (Qiagen, Germany)

• Magnetic separation stand/MTP to accommodate small-scale purifications using microcentrifuge tubes/MTP (Promega, USA)

**Procedure**

**Step 1: Preparation Step and Construct Transformation**

1. Transform individually 1 µL of the plasmid DNA (~100 ng µL⁻¹) into the competent expression cells *E. coli* by easy-to-parallel heat-shock method (follow the manufacturer’s protocol). Series of dilution can be performed to obtain the separate colonies on agar plates.

2. In parallel, prepare sterile a single, 6-, 12-, or 24-well 2xLB-agar plates supplemented with antibiotics depending on the size of the gene expression set.

3. Transfer the multiple cells on the 2xLB-agar supplemented with antibiotics plate, cover the plate, and gently rotate (shake) it horizontally (to spread the mixtures homogeneously).

4. Incubate at 37 °C overnight, bottom oriented down.

5. Sterilize 24-deep-well plates for the next day.

**Step 2: Expression in Small-scale Bacterial Cultures**

1. Prepare culture media (Starting or induction media) for cell cultivation in small-scale.
Scrape several *E. coli* colonies off the agar plate and transfer them into 2 mL of Starting medium in a 24-deep-well plate.

*Note 1*: For every single gene candidate inoculate a duplicate of minicultures to withdraw the control samples: Either (i) sample “before induction” and (ii) sample “after harvesting” for 5 mL of culture volume in individual 50-mL tubes or only sample “after harvesting” for 2(1) mL of culture volume in 24-(48)-deep-well plates.

2. Cover the plate carefully with an air-pore tape sheet.
3. Incubate at 37 °C for 4.5–5 hours with shaking (200 rpm).
4. Decrease the temperature to 22 °C and let the culture incubate for an additional 30 minutes.
5. Add 2 mL of Induction medium to each well.
6. Continue incubation at 22 °C overnight with shaking (200 rpm).

*Note 2*: Time to time check the bacterial growth by visual inspection.

*Note 3*: Withdraw the control samples “after harvesting” and measure OD$_{600}$ from the second out of two cultures in duplicate.

**Step 3: Harvesting and Lysis of Pelleted Bacterial Cells**

1. Harvest cells by centrifugation the 24-deep-well plate for 10 min, 4,000 g at 4 °C.
2. Decant the supernatant and remove remaining drops by tissue.
3. Re-suspended in desired wash buffer (100 mM Tris-HCl, pH 7.5, with 10 mM imidazole) to reach a final OD$_{600}$ 6.0.
4. Transfer a maximum of 1 mL of bacterial culture with OD$_{600}$ 6.0 to the 1.5-mL tubes and freeze the 1-mL aliquots at -80 °C for at least 30 minutes.

*Note 4*: The freeze-thaw cycle can be repeated to achieve higher lysis efficiency.
5. Add 1-2 µL 1 unit µL$^{-1}$ of DNase I (follow the manufacturer’s recommendation) to the lysed bacterial culture (New England Biolabs, USA).
6. Break the cells by small-tip sonication for 3x2 min (amplitude 55 %, cycle 0.5).

*Note 5*: The 24 Tip Horn can be mounted to process each well of a standard-sized 24 well plate simultaneously, or alternatively, bacterial cells can be lysed directly using cell lysis reagent without the need for centrifugation or mechanical cell disruption.

**Step 4: Purification of Polyhistidine-Proteins Using Magnetic Ni-Particles**

1. Vortex the MagneHis™ Ni-Particles to make a uniform suspension.
2. Add 30 µL of MagneHis™ Ni-Particles to the cell lysate.

*Note 6*: In the case of highly-expressing proteins, the amount of MagneHis™ Ni-Particles might be increased.
Note 7: The final concentration of 500mM NaCl can be added to lysate (i.e., 0.03g NaCl per 1.0 mL of lysate) to improve binding to MagneHis™ Ni-Particles.

3. Incubate mixtures for 10 minutes on a roller (40 rpm) at 4 °C. Make sure the magnetic particles are well mixed.

4. Place the 1.5-mL tube in the appropriate magnetic stand for approximately 60 seconds to capture the Magnetic particles at room temperature.

5. Carefully remove the unbound proteins by pipetting the supernatant out.

6. Remove the tube from the magnetic stand. Add 150 µL of wash Buffer to the pellet and pipet to mix.

Note 8: If 500mM NaCl in the final concentration was added to the lysate, also add the same concentration of NaCl during washing.

7. Place the 1.5-mL tube in the appropriate magnetic stand for approximately 60 seconds to capture the magnetic particles.

8. Carefully remove the unbound proteins by pipetting the supernatant out.

9. Repeat the wash step three times.

10. Remove the tube from the magnetic stand. Add 50-100 µL of elution buffer, and properly mix it by pipetting.

Note 9: The volume of elution buffer depends on the desired protein concentration.

11. Incubate for 5 minutes at room temperature.

12. Place the 1.5-mL tube in the appropriate magnetic stand.

13. Carefully collect the target proteins by pipetting the supernatant out.

Step 5: Expression Analysis

For every single HLD gene candidate, a duplicate of minicultures was inoculated. The duplicate of minicultures was used to withdraw the control samples: either (i) sample “before induction” and (ii) sample “after harvesting” for 5 mL of culture volume in individual 50-mL tubes or only sample “after harvesting” for 2(1) mL of culture volume in 24-(48)-deep-well plates (Fig. 4).

Fig. 4. here.

1. Measure the OD₆₆₀ of the culture “before the induction” or “after harvesting” of cell culture. Withdraw 0.5 mL-samples of cell culture to microtube labeled as I and H (induction/harvest).

2. Centrifuge the cell cultures for 15 min at 6,000 g and discards the supernatant subsequently. Store the cell pellets in the freezer.

3. Defreeze the cells before the preparation of SDS samples.

4. Add V µL of harvesting buffer to the cells to reach OD₆₆₀ of 10. Mix properly.
Note 10: If 0.5 mL cell culture reached OD₆₀₀ 5, add (V) 0.25 mL of harvesting buffer to reach OD₆₀₀ of 10.

5. Sonicate the cells in the sonication bath two-times for 5 min.

6. Mix 10 µL of sonicated cells taken before induction (microtube I) with 10 µL of sterile water. Name the sample PBG.

7. Mix 10 µL of disintegrated cells taken at harvesting (microtube H) with 10 µL of sterile water. Name the sample TOT (total cell protein).

8. Centrifuge rest of cell suspension in microtube H in a small centrifuge for 10 min at 20,000 g.

9. Mix 20 µL of the supernatant. Name the sample SOL (soluble protein).

10. Discard the rest of the supernatant.

11. Add V µL of 8 M urea to the pellet, mix, and vortex properly.

12. Centrifuge the solution for 10 min at 20,000 g.

13. Mix 20 µL of the supernatant. Name the sample INS (insoluble protein – solubilized by 8 M urea).

14. Add 5 µL of SDS dye to each sample, incubate 5 min at 95 °C.

15. Load 10 µL of mixture onto the SDS gel, run for 80 min at 120 V and 300 mA.

16. Stain the gel with Coomassie Brilliant Blue (Bio-Rad, USA) solution.

17. Destain the gel by using the destaining solution,store the gel in distilled water.

Tips & Troubleshooting

- A fresh stock of the appropriate antibiotics corresponding with the expression vector is used as the supplement into liquid and plate cultures. Dilute antibiotics at recommended working concentration.

- Rich media 2xLB, Terrific Broth, or Super Broth medium enable the growth and maintenance of high density of cells for a long growth period and thus affect the quality of the end production of a homogeneous population of proteins culture (Kram & Finkel, 2015).

- Cell-free protein synthesis represents the future approach for high-throughput functional and structural studies of proteins.

3. Microscale Enzyme Characterization

The final step in the exploration of natural and/or man-made diversity – biochemical characterization of the identified hits – is the rate-limiting step as the traditional biochemical techniques are time-demanding, volume, and cost-ineffective, and low-throughput. As an answer, novel high-throughput experimental techniques employing miniaturization and automation have been developed to speed up biochemical characterization. Note that the methods described below can also be used for functional screening in either metagenomic or directed evolution studies. These methods do not necessarily offer
the highest possible throughput. Yet, they are robust and reliable and thus satisfy the needs of biochemical characterization. This section provides detailed protocols for the acquisition of both structural (e.g., stability, solubility, and aggregation propensity) and functional characteristics (e.g., activity, specificity, temperature profiles) with a particular accent on the method applicability.

3.1. Structural Characterization

Natural enzymes have been evolving for billions of years of evolution in delicate biomolecules possessing marginal stability. Poor stability, misfolding or aggregation reduce the total yield of functional protein for use in downstream applications upon purification. Proper assessment of both stability and aggregation propensity is, therefore, necessary for protein screening pipelines to minimize potential time- and economic losses in further development. Two of the most common and well-established techniques for measurement of protein stability are arguably circular dichroism spectroscopy (CD) and differential scanning calorimetry (DSC), which provide information about the changes in secondary structure and internal energy during denaturation, respectively. Unfortunately, their use for extensive stability screening is hampered mostly by their low throughput and relatively large sample consumption. Automated versions of the DSC and CD instruments utilizing multiple 96-well plate format with autosampler offer the benefit of unattended operations. However, the lack of parallelization and high investment cost still restrict their general usage to larger biotechnological companies or core-facilities (Table 1). For stability screening, when only a comparison between individual proteins is necessary, techniques monitoring either intrinsic or extrinsic fluorescence are more suitable. The following protocols describe the high-throughput thermal denaturation of recombinant proteins using (i) robust thermal shift assay (TSA) utilizing the extrinsic fluorescence (SYPRO orange dye), and (ii) various versions of differential scanning fluorimetry (DSF) employing the intrinsic fluorescence measurement.

Table 1. here.

Alternatively, thermal inactivation assay focused on monitoring enzymatic activity at a specific temperature can be employed. It usually provides more accurate information about the structure-function relationship, but it lacks sufficient throughput and robustness. More specifically, thermal inactivation assay needs to be optimized for specific enzyme family, which often results in discontinuous measurement. In terms of its applicability, it is suitable for the evaluation of the set of mutants with the specific temperature threshold. Lastly, the preliminary screening of protein stability-activity can be achieved by assaying residual activity. This type of assay is especially important in protein engineering studies focused on the improvement of protein stability, where a functional state of a protein must be maintained under harsh conditions, such as the presence of organic co-solvents or high temperature, over a certain period of time. In this fashion, stability is often correlated to the residual activity after incubation at a certain temperature for a given time.

3.1.1. Thermal Shift Assay
The thermal shift assay, also referred to as Thermofluor, is a widely-used technique for high-throughput determination of protein stability by monitoring the extrinsic fluorescence using environment-sensitive fluorescent dyes, e.g., SYPRO orange (Pantoliano et al., 2001). In principle, SYPRO orange dye undergoes a significant increase in quantum yield upon binding on the exposed hydrophobic core regions of the protein during protein unfolding. Since fluorescence properties of SYPRO orange dye (λ<sub>ex</sub> 470 nm /λ<sub>em</sub> 570 nm) are compatible with common filter sets found on real-time PCR instruments, it allows easy adaptation of protein thermal denaturation assays. The further advantages of TSA include low sample consumption and high-throughput analysis as it is performed in 96-well format.

TSA is commonly used for high-throughput screens of thermal denaturation and ligand affinity (Bai et al., 2019), or optimization of buffer conditions. However, the quality of the thermal denaturation curve is highly protein-dependent. Attention should be paid to possible caveats (Deller et al., 2016), e.g., potential adverse effects of protein-dye interaction on protein stability, a common occurrence of false-positive or aberrant thermal shifts. It is estimated that 15–25% of recombinant proteins might provide artificial profiles upon thermal denaturation (Crowther et al., 2010). Huynh and Partch provided a detailed description of the assay’s principle in a comprehensive methodological review elsewhere (Huynh & Partch, 2015).

The following protocol describes the high-throughput measurement of thermal denaturation using SYPRO Orange dye.

**Material and Equipment**

- 96-well real-time PCR detection system (e.g., StepOne™, Applied Biosystems)
- SYPRO Orange Fluorescent Dye (5,000x concentrated in DMSO) (e.g., ThermoFisher Scientific, USA)
- Optical 96-well plate for real-time PCR (e.g., Thermo Fisher Scientific, USA) or other 96-well plates with low profile, unskirted and clear
- Optically clear sealing film for 96-well plate
- Swinging bucket centrifuge equipped with plate adapters for 96-well plates (Beckman, USA)
- Protein in a final concentration of (5–10 µM) (protein dialyzed against reaction buffer)

**Procedure**

**Step 1: Sample Preparation**

1. Make a 50x working concentration of SYPRO Orange Dye by diluting 2.5 µL of 5,000x concentrate into 250 µL of deionized water (calculated for 100 reactions or 1 real-time PCR plate) into black micro centrifuge tube for light-sensitive samples.

   *Note 1: Before opening, allow the stock solution of SYPRO orange to come to room temperature and then briefly centrifuge in a microfuge.*

2. Prepare reaction mixture to the final volume of 25 µL per well with the following composition:
17.5 µL of reaction buffer (protein dialysis buffer)

2.5 µL of SYPRO Orange Fluorescent Dye (50x stock concentration)

5 µL of protein (20 µM)

**Note 2**: We recommend calculating the amount of protein and dye so that their final concentrations in the reaction mixture are approx. 5 µM and 5x, respectively.

**Note 3**: The optimization of SYPRO Orange dye and protein concentrations might be necessary to get optimal assay performance. The final assay concentrations of SYPRO Orange dye should fall in the range from 10x to 0.5x, whereas 2 to 20 µM of protein sample.

3. Pipet the individual components of the reaction mixture in the following order: (i) reaction buffer, (ii) protein sample, (iii) SYPRO Orange dye. Alternatively, prepare a master mix of the reaction buffer with diluted SYPRO Orange dye in the final concentration.

**Note 4**: We recommend performing each experimental condition in at least triplicate to ensure reproducibility.

4. Mix the well content by pipetting up and down several times.

5. Cover the assay plate tightly with a sheet of an optically clear adhesive seal to avoid sample evaporation.

6. Centrifuge the assay plate at 750 × g for 1 min to collect reaction solutions in the bottom of the well and to remove bubbles.

**Step 2: Temperature Denaturation Experiment and Data Analysis**

- Place the assay plate into the real-time PCR system and set a temperature gradient program for protein thermal denaturation.

- Generalized protocol for monitoring thermal denaturation:
  - Start temperature: 25 °C
  - Initial equilibration: 2 min
  - Temperature ramp: increments of 1 °C min⁻¹
  - Final temperature: 95 °C.

**Note 5**: Low fluorescence signal at the starting temperature indicates a well-folded protein.

- The inflex point of the resulting sigmoid curve is generally used as a parameter for stability, i.e., apparent melting temperature ($T_{\text{m}, \text{app}}$). It is determined by fitting Boltzmann or other sigmoid curves to the data or by taking a maximum of the first derivative of the fluorescence emission as a function of temperature ($-dF/dT$).

**Note 6**: Measured $T_{\text{m}, \text{app}}$ is identical to the “true” melting temperature only in cases when the protein denaturation is two-state and reversible, which is not valid for many proteins. However, it is still a
suitable indicator of protein stability and is commonly used in the scientific literature. We recommend monitoring $T_{\text{onset}}$ besides monitoring the solely $T_m^{\text{app}}$ value. $T_{\text{onset}}$ depicts the temperature at which the thermal decomposition of the protein structure begins. Importantly, it usually correlates well with the temperature optima of an enzyme.

**Note**: The baseline can be defined by subtracting buffer conditions (thermal denaturation of solely buffer with identical experimental conditions).

**Tips & Troubleshooting**

- Not all proteins give “typical” sigmoidal profiles upon thermal denaturation as they may lack the compact, globular fold (i.e., intrinsic disorder) or hydrophobic core. They may also contain hydrophobic patches on the solvent-exposed surface in the folded state or have poor stability at room temperature (Huynh & Partch, 2015).

- The cases mentioned above might result in high fluorescence background even at low temperatures or low fluorescence signals caused by the protein lacking a cooperative unfolding transition. Both cases might preclude the analysis of protein stability by TSA.

- To optimize the thermal denaturation assay, perform the systematic titration of the concentration of SYPRO Orange dye (0.5 – $10 \times$ final concentration) versus the protein sample (1–20 µM). If optimal conditions for the denaturation assay can’t be found, a poor melting curve is most likely associated with the specific protein target. We highly recommend the control measurement of at least few variants with another technique and comparison of the melt curves.

- The selection of reaction buffer greatly depends on a relatively stable environment for the protein of interest. In general, the buffer concentration should be maintained rather low along with low ionic strength.

- Careful data analysis is recommended using tools such as Calfitter (Mazurenko et al., 2018).

**3.1.2. Capillary-Based Differential Scanning Fluorimetry**

The capillary-based differential scanning fluorimetry (DSF) also referred to as nano differential scanning fluorimetry (nanoDSF), is a label-free technique for the determination of protein stability by measuring the intrinsic change in tryptophan and/or tyrosine fluorescence of a protein upon unfolding (Magnusson et al., 2019). Due to the capillary-based set-up used in the Prometheus NT.48 instrument (NanoTemper Technologies, GmbH), up to 48 unique measurements can be carried out in parallel in high-throughput format with very low protein consumption. In contrast to TSA, the protein stability measurements can be performed within a broad concentration range of protein samples from 5 µg mL$^{-1}$ to 150 mg mL$^{-1}$ without the need of any optimization. The major set-back is the requirement of the specialized instrument for the analysis. Additionally, the label-free regime precludes the study of proteins lacking aromatic residues, especially tryptophan.

**Procedure**
Protein samples can be scanned from 15 to 98 °C at 0.1 – 5 °C min⁻¹ scan rate. Fluorescence signal excited at 295 nm is followed at 335 and 350 nm. The ratio of fluorescence intensities at both excitation wavelengths (corresponding to the “redshift” of the tryptophan fluorescence upon protein unfolding) is plotted as a function of temperature, and the inflex point of the resulting curve is used as a stability parameter for the comparison between conditions/variants.

**Tips & Troubleshooting**

- Efficient protein stability screening was demonstrated directly on overexpressed proteins in crude extracts by hypotonic extraction of overexpressed protein from bacterial host cells, enabling the determination of several hundred melting temperatures per day with high precision and reproducibility (Magnusson et al., 2019; Wedde et al., 2017).
- This label-free technique enables the screening of protein stability in the presence of co-solvents.
- Scattering of the incident light is utilized as the indicator of aggregation during unfolding, which provides the user with an additional layer of information about protein conformational stability.

### 3.1.4. Applicability

The techniques mentioned above can be used to determine various protein stability-related processes (Fig. 5), e.g., thermal denaturation, chemical denaturation, ligand affinity binding, buffer optimization, protein aggregation, or long-term storage.

*Fig. 5. here*
3.2. Functional Characterization

3.2.1. Activity-Based Robotic Assays

Most activity assays are based on spectrophotometric measurements (Bisswanger, 2014). In terms of throughput, microtiter plate-based methods, e.g. (Rachinskiy et al., 2009) and their robotized versions, e.g. (Dörr et al., 2016) provide a reproducible output with the moderate throughput (Bunzel et al., 2018).

We set up a robotic platform employing a 96-well microtiter plate format for fast screening of the dehalogenase activity. All liquid handling steps are automated, which allows performing sample preparation and analysis of the coupled chemical reaction producing a colorimetric signal. Robotic dehalogenase activity measurement employs a spectrophotometric detection of halides released during dehalogenase reactions originally developed in 1952 by Iwasaki (Iwasaki et al., 1952). Kunka et al. provided a detailed description of the assay’s principle and the calibration for measurement in a comprehensive methodological review elsewhere (Kunka et al., 2018).

The core of our robotic platform consists of a commercial liquid handling robot Micro-Star (Hamilton Robotics, Bonaduz, Switzerland) equipped by the 8-channel pipetting head and by several add-on modules including shaking module, cooling module, heating module, wash station, disposable tips carrier, steel tips carrier, reagent carrier, and microtiter plate carrier. Additionally, external orbital thermoshaker (Biosan, Latvia) and microplate reader SUNRISE RC (Tecan, Switzerland) are used.

The following paragraph is dedicated to the description of automated dehalogenase activity measurement rather than the detailed protocol.

Procedure

The method is designed to run up to 24 individual enzymatic reactions in 2-mL glass vials placed in the shaking-heating module. Before the initiation of the measurement, the operator is asked to set the following variable reaction parameters: (i) the number of assayed enzymes in triplicates (from 1 to 8), (ii) the amount of added enzyme (from 10 to 100 μL), (iii) the periodical intervals of the sample withdrawal (from 1 to 5 minutes), and (iv) the incubation temperature (from 5 to 90 °C). The 1 mL of 100 mM glycine buffer, pH 8.6, and 1 μL of the halogenated substrate are pipetted by 8-pipetting channel head. The reaction mixtures are incubated with shaking for 30 minutes. The reaction is initiated by adding 10-100 μL of the enzyme. The reaction progress is monitored by periodically withdrawing 75 μL of samples from the reaction mixture and immediately mixing these samples with 10 μL of 35 % (v/v) nitric acid to terminate the reaction. Released halide ions are detected by pipetting 10 μL of using Hg(SCN)$_2$ (9.5 mM in ethanol) and 20 μL of NH$_4$Fe(SO$_4$)$_2$ (250 mM in 9 M HNO$_3$). Finally, dehalogenation activities are quantified as the rate of product formation over time.

Tips & Troubleshooting

- Use reaction buffers with buffering capacity in the mild alkali pH region corresponding with the pH profile of HLDs, e.g., glycine buffer (100 mM, pH 8.6).
Spontaneous hydrolysis of the halogenated substrate might occur during long time incubation or at elevated temperature. Thus, the abiotic dehalogenation needs to be subtracted to quantify the dehalogenase activity accurately.

**Applicability**

The applicability of our robotic platform is demonstrated by the broad scope of dehalogenase activity measurements including substrate specificity (Babkova et al., 2017), temperature and pH profiles (Vanacek et al., 2018), functional half-lives at elevated temperatures (Kotik et al., 2017), and effects of organic solvents on HLD activity. The employment of robotic liquid handling enables speeding up the substrate specificity profiling for a single enzyme towards 30 halogenated substrates about 8-fold with a 10-fold reduction in the sample volume requirement compared to manual measurement (Buryska et al., 2019).

### 3.2.2. Activity-Based Microfluidic assays

Droplet-based microfluidics offers a further level of miniaturization with incomparable throughput (Tran et al., 2013). It manipulates small, uniform, and spatial aqueous reaction droplets encapsulated in oil, constrained within a microenvironment having internal dimensions on a scale of microns or lower (Seemann et al., 2011). As every single generated droplet represents individual biochemical microreactor, droplet microfluidics is perfectly suitable for the high-throughput biochemical characterization of the newly discovered/improved biocatalysts. Unfortunately, the hydrophobic compounds (e.g., substrates, fluorophores, drugs/drug-leads, vitamins) encapsulated into water droplets often tend to leak from water compartments to the carrier (Y. Chen et al., 2012).

To overcome this significant limitation, we have developed a novel way of delivery of hydrophobic compounds into microfluidic droplets (Buryska et al., 2019). This concept was utilized for high-throughput characterization of enzymes’ activity, substrate specificity, temperature optima, and thermodynamics, and it was validated on 8 well-known enzymes from the HLD family.

The setup of the capillary-based droplet microfluidic platform consists of a commercial robotic sampler Dropix (Dolomite Microfluidics, UK), an in-house constructed incubation chamber, a temperature controller, and an optical setup for monitoring the biochemical reactions in droplets (Fig. 6).

**Fig. 6 here.**

**Assay principle**

The principle of substrate delivery was described in detail in the original methodology paper (Buryska et al., 2019). In brief, the delivery is based on two fundamental principles: microdialysis and oil/water partitioning. The hydrophobic substrate molecules pass through the tubing walls via microdialysis, dissolve in the fluorinated oil and finally reach the aqueous droplets containing enzymes by partitioning equilibrium. At equilibrium, the final concentration of the compounds in the water compartments is determined by its oil/buffer partitioning coefficient (LogPoil/buf) and the concentration of the substrate in the oil phase. Based on theoretical calculation, the equilibrium can be reached within milliseconds, and based on the experiment, the final substrate concentration in aqueous phase reaches the level of several...
mM. Furthermore, it is possible to predict the LogP_{oil/buf} value, as it correlates well with the partition coefficients standardly determined for octanol/water system (LogP_{oct/wat}).

A pH-based fluorescence assay was employed to monitor the dehalogenase activity in the microfluidic droplets, containing a weakly buffered system, e.g., HEPES buffer, and a complementary fluorescent indicator, e.g., 8-hydroxypyrene-1,3,6-trisulfonic acid (HPTS) for the observation of small changes in the pH (Nevolova et al., 2019). The reaction buffer (1 mM HEPES, 20 mM Na_2SO_4, and 50 μM HPTS) was used in the range from pH 6.8 to 8.2 (with a linear signal response for this range).

**Data Collection**

Reaction progress was analyzed as an end-point measurement recorded after passing through the incubation chamber. Excitation was achieved by 450 nm blue laser (12V, 200mW), focused by a spherical lens into the multimode optical fiber Y-bundle suitable for reflection and backscatter spectroscopy (SQS, Czech Republic). The excitation wavelength was filtered on a dichroic mirror with cut off at 490 nm (ThorLabs, Germany). The analog signal was collected on Si-detector (Thor-Labs, Germany), converted to the digital signal by Ni-DAQ 6009 module (National Instruments, USA), and processed by LabView 12 (National Instruments, USA).

**Solutions and Media**

- **HEPES buffer**: 1 mM HEPES, 20 mM Na_2SO_4, pH 8.2, ultrasonicated and filtered (0.22 μm pore size)
- **Reaction buffer**: 1 mM HEPES, 20 mM Na_2SO_4, 50 μM HPTS, pH 8.2, ultrasonicated and filtered (0.22 μm pore size)
- **Oil**: 0.5 % Pico-Surf 1 (Dolomite, UK) in FC-40 oil (3M, USA)
- **Enzyme variant of interest dialyzed into HEPES buffer, pH 8.2**
- **100 mM NaOH** in deionized H_2O
- **100 mM HCl** in deionized H_2O

**Material and Equipment**

- **Mitos Dropix droplet generator** (Dolomite, UK)
- **Chemyx Fusion 200 pump** (Chemyx, USA)
- **Fixed Luer tip glass syringes**
- **Copper block coupled with a Peltier element** (in-house fabrication)
- **Temperature controlling element TEC-1089** (Meerstetter, Switzerland)
- **Fluidic connectors**: P-662, PK-112, XP-230 (all from IDEX Health & Science, USA)
- **2-mL glass vials with screwed caps**
- **Polyethylene Tubing O.D. 0.8 mm, I.D. 0.4mm** (Smiths Medical, UK)
- **Laser 450 nm 200 mW, 12V**
- **Optical fiber bundle 660 μm ID** (SQS fiberoptics, Czech Republic)
- **Detection cell** (in-house fabrication)
- **DMLP490R Longpass Dichroic Mirror, 490 nm Cut-On** (ThorLabs, Germany)
• Kinematic Fluorescence Filter Cube (Thorlabs, Germany)
• Si Amplified Photodetector (Thorlabs, Germany)
• NI USB-6009 DAQ module (National Instruments, USA)

Software
• Dropix control software (Dolomite, UK)
• TEC Service Software (Meerstetter, Switzerland)
• LabView (National Instruments, USA)
• MATLAB (Mathworks, USA)
• Statistica (TIBCO, USA)

Procedure

Step 1: Calibration
1. Prepare the 100 mM NaOH solution in deionized water and determine its factor using titration against a known amount of oxalic acid.
2. Prepare a solution of approximately 100 mM HCl from the concentrated HCl and determine the exact concentration by titration against 100 mM NaOH solution.
3. Prepare a semi-stock solution of 10 mM and 4 mM HCl solution in the reaction buffer.
4. Prepare a series of calibration samples by serial dilution. Example working concentrations are: 1.75 mM, 1.4 mM, 1.05 mM, 0.7 mM, and 0.35 mM.
5. The calibration samples shall be applied within the activity measurement and should be included in every activity measurement except for blank enzyme measurement.

Note 1: Work with the calibration series only for about 2 h. The chlorides enhance the fluorescence of HPTS in the time course. After 2 h, a new calibration series should be prepared out of 4 mM HCl semi-stock solution as well as a fresh reaction buffer.

Step 2: Activity Measurement
1. Fill the 2-mL glass vial with a pure halogenated substrate. Cut approximately 60 cm of polyethylene tubing with outer diameter (O.D.) of 0.8 mm and inner diameter (I.D.) of 0.4 mm. Wreathe ca. 20 cm of tubing in the middle into three loops around a cut pipetting tip and thread it through the punched holes in the screwed cap. Then, insert this tubing bundle into the glass vial with the halogenated substrate. Let the tubing soak with the substrate at least 12 hours.
2. Turn on the temperature controller and set the target temperature. Also, turn on the laser and detector. Align one part of the tubing to the excitation/emission optical fiber within the detection cell and attach it to the glass syringe (prefilled with FC-40) via P-662 adapter, including PK-112 ferrule (replacing the original F-152 ferrule).
3. The second part of the tubing should be thread through the Dropix sample hook and fixed by Dropix sample hook fitting. Only about 2-3 mm of tubing should come out of the hook.
4. Fill the whole tubing with FC-40 by gentle infusion from the attached syringe. Then fix the glass syringe to the pump. Before placing the hook into the oil bath, immerse it into about 5 mL of FC-40 and let the pump draw at least 100 μL of FC-40 to prevent contamination of fluid reservoir oil with the halogenated substrate. Then attach the hook to the Mitos Dropix instrument and place the glass vial into the metal block with the temperature controller.

Note 2: Be ensured that there are absolutely no air bubbles in the syringe. These will cause instability of droplets and also of the reaction time.

5. Prefill the fluid reservoir with ca. 16-20 mL of FC-40 oil with 0.5 % Pico-Surf 1 surfactant. Run the pump in a “withdraw mode” at a flow rate of 10 μL min⁻¹. Let the pump run for at least 2 min before starting the experiment to equilibrate the flow.

6. Load 20 μL of each sample (calibration solutions containing defined concentrations of HCl and enzyme dissolved in the reaction buffer) into the 24-well Dropix sample strip and place it above the fluid reservoir.

7. Set the sample sequence (order of samples, droplet volume, oil spacing, and the number of droplets) on the Dropix control software. Then start the sequence and thus the experiment.

Note 3: The individual calibration or enzyme samples should be separated by plain reaction buffer within the sample sequence set on Dropix. For the polyethylene tubing O.D. 0.8 mm, I.D. 0.4mm, the typical volumes are 150 nL of droplets and 300 nL of oil spacing. These are the minimum values. The standard number of droplets per any kind of sample is 10.

8. Start the data acquisition on LabView right at the moment when the first droplet of the sequence enters the glass vial. The time delay between the start of acquisition and the time of first droplet detection is the reaction time. After the detection of the whole sequence, stop the measurement.

Step 3: Blank Measurement

1. Fill the 2-mL glass vial with HEPES buffer. Insert a new tubing bundle prepared the same way as for activity measurement.

2. Carry out the same procedure as for activity measurement but skip the measurement of calibration samples. Include only plain reaction buffer and individual enzymes solutions.

Step 4: Data Analysis

A droplet detection script was written in MATLAB 2017b (Mathworks, USA) processes the raw signal providing the average signal for individual droplets. The code is accessible in Supplementary information (Buryška et al., 2019). Specific activities can be calculated based on the reaction time, enzyme concentration, calibration curve, and average signal for individual droplets. Principal component analysis (PCA) can be carried out to assess the substrate specificity. A detailed description of the PCA was described previously (Koudelaková et al., 2011)

Tips & Troubleshooting
Adjust the pH of the reaction buffer using Na$_2$CO$_3$ and H$_2$SO$_4$. Avoid using HCl for pH adjustment.

It takes about 5-15 minutes for the temperature controller to equilibrate the initial temperature. The change of temperature of ± 5 °C requires approximately 5 min of equilibration time.

**Applicability**

Our capillary-based microfluidic method uses a fluorescence probe to monitor a change of pH. This type of measurement is universal for numerous enzyme families that are important for industrial or medical applications (Table 2). Primarily, the platform serves to characterize enzymes converting hydrophobic compounds (log P > 0), but hydrophilic substrates can be used as well by using triggered droplet merging (Gielen et al., 2013). Alternatively, absorbance detection can also replace fluorescence detection, as demonstrated previously (Gielen et al., 2013, 2015; T. Yang et al., 2017).

Table 2 here.

### 4. Conclusions and Perspectives

Enzymes are prominent biocatalysts, offering environmental and economic benefits in the acceleration of industrially and pharmaceutically important chemical reactions (Sheldon & Woodley, 2018). The identification and acquisition of novel enzymes have been addressed by both genome mining approaches and protein engineering employing ultra-high-throughput screening techniques accelerating directed evolution (Devine et al., 2018). The rate-limiting step of the process remains to be the development of high-throughput methodologies for complete biophysical characterization of a large set of protein candidates (Truppo, 2017).

This book chapter addresses this issue by providing protocols for high-throughput protein production and microscale characterization of enzymes of interest. These methods are perfectly suitable for characterization of smaller sets of novel enzymes/variants (tens to hundreds) such as enzymes discovered via genome mining, or the hits originating from the screening in a directed evolution experiment. The integration of the high-throughput techniques for protein production and microscale characterization of enzymes should lead to a robust workflow, which accelerates the complete biophysical characterization of novel enzymes or improved variants.

**Acknowledgments**

The authors would like to thank the RECETOX research infrastructure of the Czech Ministry of Education, Youth and Sports (LM2018121), and CETOCOEN EXCELLENCE Teaming 2 project supported by Horizon 2020 (857560) and the Czech Ministry of Education, Youth and Sports (02.1.01/0.0/0.0/18_046/0015975). Michal Vasina gratefully acknowledges the financial support of his doctoral study by the scholarship Brno Ph.D. Talent. The authors would like to express gratitude to Antonin Kunka (Masaryk University) for his critical comments on the manuscript.
References


Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., Carver, T.,
https://doi.org/10.1177/108705710100600609

a high-throughput enzyme characterization technique including the long-term stability. *Biotechnology
and Bioengineering, 103*(2), 305–322. https://doi.org/10.1002/bit.22242

*Journal of the American Chemical Society, 135*(34), 12480–12496.
https://doi.org/10.1021/ja405051f

Sayers, E. W., Beck, J., Brister, J. R., Bolton, E. E., Canese, K., Comeau, D. C., Funk, K., Ketter,
A., Kim, S., Kimchi, A., Kitts, P. A., Kuznetsov, A., Lathrop, S., Lu, Z., McGarvey, K., Madden, T. L.,
Murphy, T. D., O’Leary, N., Phan, L., … Ostell, J. (2020). Database resources of the National Center
https://doi.org/10.1093/nar/gkz899

Smart Libraries. In E. M. J. Gillam, J. N. Copp, & D. Ackerley (Eds.), *Directed Evolution Library
Creation: Methods and Protocols* (pp. 291–314). Springer. https://doi.org/10.1007/978-1-4939-1053-
3_20

Reviews, 118*(2), 801–838. https://doi.org/10.1021/acs.chemrev.7b00203

https://doi.org/10.1110/ps.0205202

Web server for automated design of mutations and smart libraries based on sequence input information.

https://doi.org/10.1088/0022-3727/46/11/114004


**Tables**

**Table 1.** Comparison of the conventional and microscale techniques for the determination of protein stability.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Throughput (assays/day)</th>
<th>Sample requirement (µg)</th>
<th>Cost</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD spectroscopy</td>
<td>~10</td>
<td>~35-100(^1)</td>
<td>$$$$</td>
<td>(Greenfield, 2006)</td>
</tr>
<tr>
<td>DSC</td>
<td>~10</td>
<td>~50-500(^2)</td>
<td>$$$$</td>
<td>(Johnson, 2013)</td>
</tr>
<tr>
<td>TSA</td>
<td>~1,000</td>
<td>~0.1-20</td>
<td>$</td>
<td>(Huynh &amp; Partch, 2015)</td>
</tr>
<tr>
<td>Microscale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capillary DSF</td>
<td>~500</td>
<td>~0.05-2,500</td>
<td>$$$$</td>
<td>(Magnusson et al., 2019)</td>
</tr>
<tr>
<td>Microcuvette DSF</td>
<td>~500</td>
<td>~0.45-1,350</td>
<td>$$$$</td>
<td>(“Uncle,” n.d.)</td>
</tr>
</tbody>
</table>

\(^1\) calculated for standard 1.0 mm the path length of the cell; \(^2\) calculated for differential Scanning Micro calorimeter VP-Capillary DSC equipped with 130 µl standard cell volume
Table 2. The list of enzyme families suitable for the screening of enzymatic activity using pH assay. The theoretical log P values were extracted from the PubChem database (Kim et al., 2019).

<table>
<thead>
<tr>
<th>Enzyme group</th>
<th>E.C. number</th>
<th>Reaction</th>
<th>Log P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acetylcholinesterases</td>
<td>3.1.1.7</td>
<td>acetylic ester + H₂O ⇌ alcohol + acetate</td>
<td>0.2–2.0</td>
</tr>
<tr>
<td>Arylesterases</td>
<td>3.1.1.2</td>
<td>phenylacetate + H₂O ⇌ phenol + acetate</td>
<td>1.5–5.4</td>
</tr>
<tr>
<td>Carboxylesterases</td>
<td>3.1.1.1</td>
<td>carboxylic ester + H₂O ⇌ alcohol + carboxylate</td>
<td>0.1–2.5</td>
</tr>
<tr>
<td>Glucose oxidases</td>
<td>1.1.3.4</td>
<td>β-D-glucose + O₂ → D-glucono-1,5-lactone + H₂O₂</td>
<td>-3.24</td>
</tr>
<tr>
<td>Glutathione transferases</td>
<td>2.5.1.18</td>
<td>R - sulfate/nitrile/halide (R =Aliphatic/aromatic chain) + glutathione ⇌</td>
<td>1.4–3.3</td>
</tr>
<tr>
<td>Glycogen phosphorylases</td>
<td>2.4.1.1</td>
<td>(1,4-α-D-glucosyl)_n + H₃PO₄ ⇌ (1,4-α-D-glucosyl)ₙ₋₁ + α-D-glucose-1-phosphate</td>
<td>-8.5</td>
</tr>
<tr>
<td>Haloacid dehalogenases</td>
<td>3.8.1.–</td>
<td>haloacid + H₂O ⇌ hydroxyacid + halide ion + proton</td>
<td>0.2–0.9</td>
</tr>
<tr>
<td>Haloalkane dehalogenases</td>
<td>3.8.1.5</td>
<td>haloalkane + H₂O ⇌ haloalcohol + halide ion + proton</td>
<td>-1.5–3.0</td>
</tr>
<tr>
<td>Halohydrin dehalogenases</td>
<td>4.5.1.–</td>
<td>halohydrin ⇌ epoxide + halide ion + proton</td>
<td>0.4–1.6</td>
</tr>
<tr>
<td>Phospholipases A2</td>
<td>3.1.1.4</td>
<td>phosphatidylcholine + H₂O ⇌ 1-acylglycerophosphocholine + a carboxylate</td>
<td>2–15</td>
</tr>
<tr>
<td>Pyruvate kinases</td>
<td>2.7.1.40</td>
<td>pyruvate + ATP ⇌ phosphoenolpyruvate + ADP</td>
<td>-0.5</td>
</tr>
<tr>
<td>Triacylglycerol lipases</td>
<td>3.1.1.3</td>
<td>triacylglycerol + H₂O → diacylglycerol + carboxylate</td>
<td>2–15</td>
</tr>
<tr>
<td>Ureases</td>
<td>3.5.1.5</td>
<td>urea + H₂O → CO₂ + 2NH₃</td>
<td>-2.11</td>
</tr>
</tbody>
</table>
**Figure captions**

**Fig. 1** The exploration of natural and man-made diversity. The figure highlights the most important steps leading to the acquisition of new and improved biocatalysts. The final step in either type of enzyme diversity exploration is the production and biochemical characterization of the selected variants. The dashed arrow indicates that the characterized enzymes can serve as a template for subsequent directed evolution.

**Fig. 2** Graphical user interface of the automated enzyme mining web server EnzymeMiner. The server is easy to use and is freely available to the community at https://loschmidt.chemi.muni.cz/enzymeminer/. A detailed description of the server functionality and its experimental validation has been published recently (Hon et al., 2019).

**Fig. 3** The workflow of high-throughput production (beige), structural (blue), and functional characterization of novel or engineered enzyme variants. The techniques are shown in italics. The used abbreviations are as follows: SDS PAGE – sodium dodecyl sulfate–polyacrylamide gel electrophoresis, TSA – thermal shift assay, and DSF – differential scanning fluorimetry.

**Fig. 4** Workflow of expression analysis of expressed proteins with three distinct phases of the workflow: (1) sample withdrawing and preparation, (2) SDS-PAGE, and (3) protein band analysis. PBG – protein sample before induction (protein background), TOT – protein sample after induction containing the total content of proteins, SOL - protein sample after induction containing the soluble fraction of proteins, INS - protein sample after induction containing the insoluble fraction of proteins.

**Fig. 5** Applicability of high-throughput techniques for quantitative protein stability determination. Thermal shift assay (TSA, orange), capillary-based DSF (cDSF, blue), and microcuvette-based DSF (mDSF, green) are compared. The typical output of the technique is shown in the upper part as the graph of protein unfolding vs. temperature. The lower part shows individual applications. The pentagons below each application determine the applicability of the technique. Grey color stands for no applicability.

**Fig. 6** Capillary-based microfluidic platform. (A) A scheme of the platform. Aqueous samples are loaded in a bottomless rack (1), which is placed at the top of an oil bath of the Dropix instrument. Droplets are generated by a vertical movement of a hook (2) with a polythene tubing up and down between the oil bath and the rack with samples. Different enzyme samples are exchanged by a horizontal movement of the hook. Polythene tubing connected to a syringe pump in a withdrawal mode (5) is going through an incubation chamber (3) and a black Delrin (DuPont) cube (4) that serves as the detection point. Excitation light from a laser source (6) is brought to the tubing inside the detection cube. The reflected emission light is collected after passing through a dichroic mirror (7) at a photodetector (8). The temperature is controlled by a custom-made heating block (9). The schematic view of substrates delivery (bottom left). A substrate (orange) passes through the tubing wall (10), dissolves in a carrier oil (11), and finally reaches aqueous droplets (12) by oil/water partitioning. (B) The equilibrium distribution
of substrates between oil and aqueous phases. The FC40/HEPES buffer partitioning coefficients ana-
alyzed by gas chromatography for 25 halogenated compounds compared to octanol/water partitioning
coefficients retrieved from the ChemSpider (http://www.chemspider.com). (C) A raw data record from
the calibration sequence. The blue color represents a signal above the threshold (black color) used for
droplet detection. The red color represents a signal peak averaged during the droplet analysis. A dilution
series of a hydrochloric acid run in ten repetitions was used for the calibration (D). Reprinted with
permission from (Buryska et al., 2019). Copyright 2019 American Chemical Society.