

# Towards the Directed Evolution of Artificial Metalloenzymes

Jaicy Vallapurackal\*

SCS-Metrohm Award for best oral presentation in Catalysis Science & Engineering (runner-up)

**Abstract:** Artificial metalloenzymes (ArMs) are a class of enzymes holding great promise. In contrast to natural enzymes, the core of ArMs is a synthetic metallocofactor, with potential for bio-orthogonal reactivity, incorporated within a host protein. Next to chemical optimization of the metallocofactor, genetic optimization of the protein allows the further improvement of the ArM. Genetic optimization through directed evolution requires extensive screening of a large sequence-scape to enable the optimization of a desired phenotype. The process is however mostly limited by the throughput of the tools and methods available for screening. In recent years, versatile methods based on droplet microfluidics have been developed to address the need for higher throughput. This article aims to give an introduction into ArMs and the recent technological developments allowing high-throughput directed evolution of enzymes.

**Keywords:** Artificial metalloenzymes · Directed evolution · Droplet microfluidics · Screening · High-Throughput



**Jaicy Vallapurackal** completed her Bachelor and Master studies at the Department of Chemistry, University of Basel and is currently a PhD candidate at the Department of Chemistry under the supervision of Prof. Dr. Thomas R. Ward. Her research focuses on the high-throughput directed evolution of artificial metalloenzymes and *in vivo* screening of *E. coli* using a microfluidics-based screening assay.

## 1. Artificial Metalloenzymes

Incorporation of a synthetic metal cofactor within a protein scaffold enables the formation of artificial metalloenzymes (ArMs). In this strategy, the metal catalyst provides the first-coordination sphere with non-natural reactivities and larger substrate scopes. The host protein offers the second coordination sphere, increasing activity and selectivity through the protein environment.

ArMs are formed by mainly three different anchoring strategies: covalent, dative, and supramolecular anchoring. In the covalent approach, a functional group on the cofactor makes an irreversible bond with an amino acid side chain within the protein (Fig. 1A). For the dative anchoring, electrostatic interactions between the metal cofactor and the protein are exploited (Fig. 1B). Supramolecular anchoring, on the other hand, employs the high affinity of certain molecules for their target enzymes (*e.g.* biotin-streptavidin). Hereby, the metal cofactor is covalently linked to the high-affinity molecule before it is incorporated into the target enzyme (Fig. 1C).<sup>[1]</sup>

Such combination of a traditional homogeneous catalyst with a protein offers numerous advantages. Catalysis formerly limited to organic solvents becomes available in physiological conditions, facilitating the implementation of non-natural reactivities *in vivo*. The catalyst is shielded from the solvent by the often deep incorporation of the metallocofactor inside the pro-

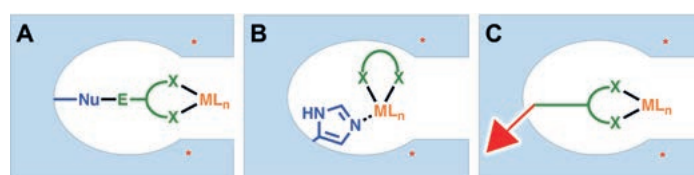


Fig. 1. Anchoring strategies to form ArMs: A. Covalent anchoring between a nucleophilic amino acid side chain and an electrophilic functional group of the metal-cofactor. B. Dative anchoring using electrostatic interactions between an amino acid and the metal. C. Supramolecular anchoring exploiting high affinity of certain molecules for their target protein. (\*) genetic modification.

tein, extending its life-time. A well-defined second coordination sphere, often well-characterized by X-ray crystallography, is an additional advantage. Above all, the protein environment offers a handle for genetic optimization (Fig. 1), ultimately enabling Darwinian-type evolution of a completely synthetic construct.

## 2. The Biotin–Streptavidin Technology

The first incorporation of an abiotic cofactor within a protein scaffold was achieved by Wilson and Whitesides in 1978. They employed the high affinity of biotin for avidin to anchor Rh(I)-diphosphine and perform hydrogenation of an  $\alpha$ -acetamidoacrylic acid in an aqueous environment. This was not only the first proof-of-concept for supramolecular anchoring to build an artificial metalloenzyme, but also the first example for asymmetric catalysis using an ArM. The developed ArM was significantly more efficient and showed moderate stereoselectivity (40% *ee*) when compared to the free metallocofactor in aqueous solution.<sup>[2]</sup>

Inspired by these results, the Ward laboratory developed ArMs based on the biotin-streptavidin technology, relying on supramolecular anchoring of biotin within streptavidin (Sav). Sav is a homo-

\*Correspondence: J. Vallapurackal, E-mail: jaicy.vallapurackal@unibas.ch  
Department of Chemistry, University of Basel, Mattenstrasse 24a, BPR 1096, CH-4058 Basel

tetrameric protein from the bacterium *Streptomyces avidinii*, showing high affinity towards biotin ( $K_d \sim 10^{-14}$  M) and able to bind up to four molecules of biotin per tetramer. It is a highly stable protein with melting temperatures of 73 °C for the apo enzyme and 112 °C for the holo enzyme. It can be lyophilized and tolerates denaturing agents such as guanidinium chloride (GuCl) and urea at a wide range of pH, making it an easy-to-handle protein. Furthermore, it was shown that Sav can tolerate extensive mutations, starting from point mutations to insertions of whole new loop regions. The active site of Sav consists of four loops between the  $\beta$ -sheets  $\beta_{3,4}$ ,  $\beta_{4,5}$ ,  $\beta_{5,6}$  and  $\beta_{7,8}$ . This provides a defined three-dimensional cavity which can accommodate a biotinylated complex and its substrate (Fig. 2). As the pocket is formed by loops, the cavity is rather dynamic, which allows fluctuations and gives enough room for the catalysis to take place.<sup>[3,4]</sup> To date, over a dozen reactions based on this technology have been reported, among which: metathesis,<sup>[5]</sup> transfer hydrogenation,<sup>[6]</sup> Suzuki cross-coupling,<sup>[7]</sup> allylic substitution<sup>[8]</sup> and more recently hydroxylation.<sup>[9]</sup>

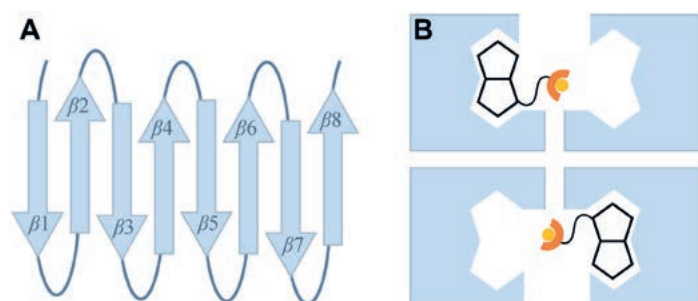


Fig. 2. A. Schematic representation of Sav  $\beta$ -sheets. B. Schematic representation of a Sav-tetramer with two incorporated biotinylated metal-locofactors.

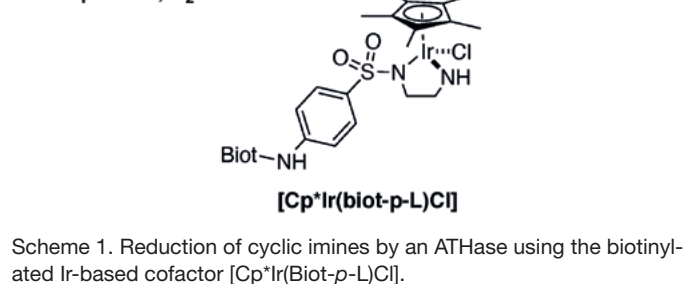
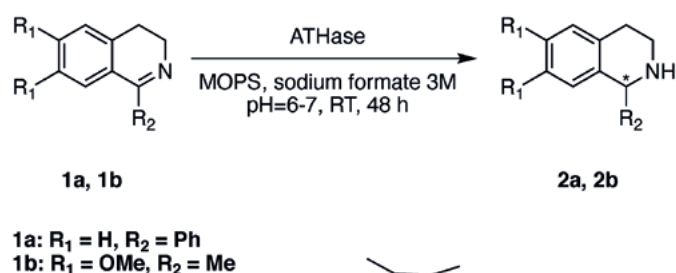
### 3. Directed Evolution of ArMs

Optimization of ArMs offers mainly two different approaches: either chemical modification to optimize the catalyst or genetic evolution to change the protein's properties. Directed evolution of enzymes is a method providing control to alter biological systems by genetic means allowing to finetune biological activities to optimize certain phenotypes.<sup>[10]</sup> Frances Arnold was awarded the Nobel prize in 2018 for the directed evolution of enzymes, jointly with George P. Smith and Sir Gregory P. Winter who were honored for their work on phage display of peptides and antibodies. Selected examples of directed evolution performed on ArMs using the biotin-streptavidin technology are described in detail below.

Different strategies were followed to develop a robust screening platform for the directed evolution of ArMs based on the biotin-streptavidin technology. Initial studies were based on an artificial transfer hydrogenase (ATHase). A small set of purified Sav isomers were screened, and improvements in conversion and differences between different mutants in enantioselectivity were observed.<sup>[11]</sup> However, obtaining lyophilized purified Sav is a time-consuming and laborious process and the screening is strongly limited by the number of variants that can be purified.

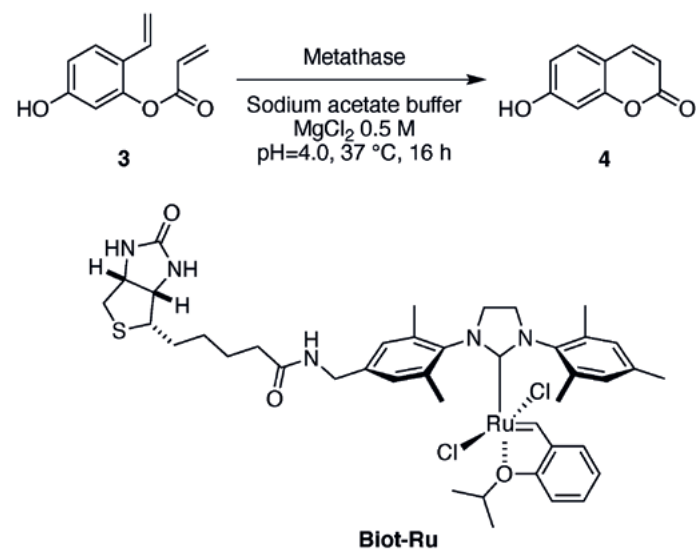
Applying *in vivo* directed evolution would highly increase the throughput of the screening but the incompatibility of the metal catalyst with the cellular environment makes it challenging. This incompatibility is mainly due to glutathione (GSH), a thiol present inside cells, which can poison the metal-catalyst. It was observed that the addition of diamide (1,1 azobis(*N,N*-dimethylformamide)) to cell free extracts (CFEs) prevents the poisoning of the Ir-based catalysts for transfer hydrogenation.<sup>[12]</sup> Based on these results, a directed evolution campaign of an artificial trans-

fer hydrogenase was carried out. Four positions in close vicinity to the embedded metallocofactor [ $\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}$ ] were selected and screened for the transfer hydrogenation of cyclic imines (Scheme 1). After iterative screening, the two best mutants led to high enantioselectivities of the (*R*)-enantiomer (95% ee) and (*S*)-enantiomer (86% ee) using the same cofactor.<sup>[13]</sup> However, diamide is not a universal solution since it is not compatible with all transition-metal based catalysts. Moreover, CFEs-based screening is challenging.



Scheme 1. Reduction of cyclic imines by an ATHase using the biotinylated Ir-based cofactor [ $\text{Cp}^*\text{Ir}(\text{Biot-}p\text{-L})\text{Cl}$ ].

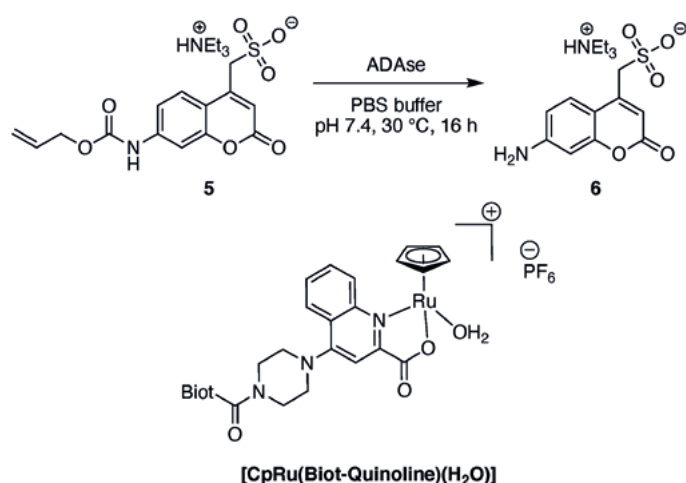
Using a periplasmic approach where Sav is expressed in the cytosol and secreted in the periplasm of *E. coli* ( $\text{Sav}^{\text{peri}}$ ), Jeschek *et al.* optimized a metathase, an ArM catalyzing the ring-closing metathesis (Scheme 2).<sup>[14]</sup> A 96-well plate assay was devised to screen single clones of *E. coli* expressing Sav in the periplasm. Fourteen positions in close vicinity of the biotinylated Hoveyda-Grubbs second-generation catalyst were subjected to saturation mutagenesis and were iteratively screened for improved metathesis activity. Five rounds of evolution yielded a mutant of Sav with 5 single point mutations and about 5-fold increase in catalytic activity.<sup>[14]</sup>



Scheme 2. Ring closing metathesis of non-fluorescent substrate **3** to umbelliferon (**4**) by a methathase. The incorporated metallocofactor is a biotinylated second generation Hoveyda-Grubbs catalyst.

Based on the same approach, three reactions catalyzed by ArMs (ring closing metathesis and two deallylation reactions) were recently screened. Variants displaying two mutations each (400 variants in total) were screened using Sav<sup>peri</sup> and a 96-well plate screening assay. In all three cases, the best mutants significantly improved the catalytic activity of the respective ArM and reached 5- to 20-fold higher TONs than the free cofactor in solution.<sup>[15]</sup>

More accessibility for the cofactor and the substrate can be achieved by the surface-display of Sav (Sav<sup>SD</sup>). In this case, Sav expressed in the cytosol is exported and displayed on the surface of *E. coli*. In a recent study, Heinisch, Schwizer, *et al.* improved the catalytic activity of an ArM for allylic deallylation (ADAse). Deprotection of an allyl-carbamate-protected coumarin (Scheme 3), served as a model system for the first directed evolution campaign. Libraries at positions K121 and S112 were introduced employing the 22-codon trick and were iteratively screened for catalytic activity using a 96-well plate assay. In the *in vivo* assay, the best mutant (S112Y-K121S) had a 25-fold improvement over WT, whereas, in purified protein, the improvement factor for the best variant (S112M-K121A) was 5.7-fold.<sup>[8]</sup>



Scheme 3. Allylic deallylation by an ADAse with the incorporated metal-locofactor [CpRu(Biot-Quinoline)(H<sub>2</sub>O)] leads to the uncaging of the fluorescent aminocoumarin (6).

However remarkable the aforementioned results are, they do not yet use the full potential of directed evolution. Every example, except one, targets one position per round of evolution, effectively missing out on potentially cooperative beneficial mutations.

This is mainly due to the screening throughout: a protein is a string of amino acids and each amino acid is encoded by three nucleotides. This gives 64 possibilities to code all 20 natural amino acids. To optimize the protein at one position, this corresponds to a screening effort of at least 190 different clones, considering it is a completely random library. To afford for cooperative effects, at least two or more positions would have to be screened simultaneously, in which case the screening effort exponentially increases to ~12'000, ~800'000 and over 50 Mio. clones for two, three and four simultaneous positions respectively.<sup>[16,17]</sup> Thereby, to truly exploit the potential of genetic optimization and eventually evolution, ultra-high-throughput screening is indispensable.

#### 4. High-Throughput Screening

With the advance of fluorescence activated cell sorting (FACS) and next generation sequencing (NGS), the throughput to investigate large populations of cells for molecular evolution and the possibility to understand cellular mechanisms on a genetic level

have tremendously increased. Together with large-scale screening, this holds the potential for faster and more efficient directed evolution campaigns.

There are two main requirements to develop such systems: (i) a tight linkage of the genetic information (genotype) with the activity of the target (phenotype) and (ii) an assay based on activity.<sup>[18,19]</sup> The genotype can be determined by gene sequencing (Sanger, Nanopore or NGS) and the phenotype activity can be measured by various analytical tools, such as mass spectroscopy, NMR, HPLC, UV-vis, and colorimetric or fluorometric assays. The genotype-phenotype linkage can be achieved by compartmentalizing distinct reactions in separate tubes, microtiter plates (MTPs), or by carrying out enzymatic reactions within a cell and thereby retaining the analytic information within or on the cell. The drawback of these types of assays is the throughput in the former and cross contamination or incompatibility with live cells in the latter case.<sup>[20]</sup>

*In vitro* compartmentalization (IVC) has emerged as an alternative approach to preserve the phenotype-genotype linkage. In this approach, the reaction to be screened can be encapsulated within micrometer-sized emulsions or immobilized within gel microbeads (Fig. 3), effectively creating distinct compartments. This not only drastically reduces screening time, but also decreases the amount of material needed.<sup>[21]</sup> The emulsion compartments can either be (i) single emulsion droplets, aqueous compartments surrounded by an oil phase, or (ii) double emulsion droplets, water-in-oil-in-water compartments generated by emulsifying single emulsions a second time. Single emulsions formed with a gellable liquid are the templates for gel microbeads.

Contrary to single emulsions, double emulsions and gel beads have an aqueous outer phase, making them FACS-compatible, and thereby an ideal solution for high-throughput *in vivo* directed evolution.<sup>[22,23]</sup>

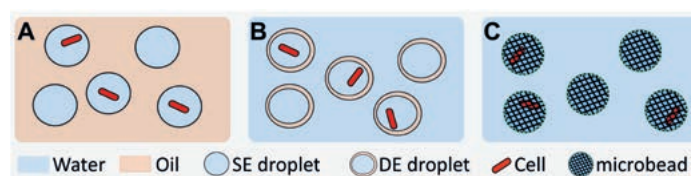


Fig. 3. A. Single emulsions (SE), B. Double emulsions (DE) and C. Gel microbeads.

The use of microfluidics tools for the formation of double emulsions affords more control over the double emulsions' size and reactant encapsulation. The Fordyce group recently illustrated the sort of double emulsions formed on-chip and encapsulating a variety of mammalian cells using a commercially available FACS machine at sorting rates over 10 kHz. They proved that the double emulsions were stable under the sorting conditions and could retrieve the genetic material from these for downstream analysis, demonstrating the versatility of IVC in double emulsions.<sup>[24,25]</sup>

#### 5. Application of IVC in Double Emulsions for Directed Evolution

In the past two decades, the technology advances in the field of IVC had tremendous impact on directed evolution. Here selected applications using double emulsions will be mentioned. A more in-depth summary of the field and comprehensive lists of applications can be found elsewhere.<sup>[20,26]</sup>

First applications consisted of simple enrichment assays of enzymes such as DNA methyltransferases<sup>[27]</sup> and arylsulfatases,<sup>[28]</sup> highlighting the throughput and potential of double emulsions. The first *in vivo* directed evolution campaign involved *E. coli*

surface-displayed serum paraoxonase 1 (PON1), which was evolved to achieve up to hundredfold improved catalytic activity. Following applications involved, for example, the directed evolution of  $\beta$ -glucosidase,<sup>[29]</sup> a model protease<sup>[30]</sup> and a cellulase.<sup>[31]</sup> More recently, a high fidelity polymerase for  $\alpha$ -L-threofuranosyl nucleic acid (TNA) was evolved *in vivo* in double emulsions.<sup>[32]</sup>

## 6. Outlook

Combining ArMs and droplet microfluidics, can be a powerful tool for propelling directed evolution-based research forward. Systematic and high-throughput screening of ArMs *in vivo* using double emulsions could allow the screening of a much bigger sequence space, which is to date not feasible. Identifying cooperative effects to improve catalysis or even remodelling whole enzymes to achieve completely new-to-nature reactivities are only two potential examples. Reactions based on ArMs could ultimately provide aqueous, environmentally friendly reaction pathways for industrial applications. Additionally, such big data sets could also be used as an input for machine learning applications, to further study active site elasticity, reaction pathways, or even protein folding mechanisms.

## Acknowledgements

The author is grateful to the Swiss Chemical Society and Metrohm for the best oral presentation award. The author thanks Prof. Thomas R. Ward for his guidance and support throughout the PhD, and Ariane Stucki for proofreading. The author acknowledges the Swiss National Science Foundation (Grant 200020\_182046), the NCCR Molecular Systems Engineering and the ERC (DrEAM - Advanced Grant 694424) for their generous financial support.

Received: February 5, 2021

- [1] J. C. Lewis, *ACS Catal.* **2013**, *3*, 2954, <https://doi.org/10.1021/cs400806a>.
- [2] M. E. Wilson, G. M. Whitesides, *J. Am. Chem. Soc.* **1978**, *100*, 306, <https://doi.org/10.1021/ja00469a064>.
- [3] T. R. Ward, *Acc. Chem. Res.* **2011**, *44*, 47, <https://doi.org/10.1021/ar100099u>.
- [4] C. M. Dundas, D. Demonte, S. Park, *Appl. Microbiol. Biotechnol.* **2013**, *97*, 9343, <https://doi.org/10.1007/s00253-013-5232-z>.
- [5] V. Sabatino, J. G. Rebelein, T. R. Ward, *J. Am. Chem. Soc.* **2019**, *141*, 24, <https://doi.org/10.1021/jacs.9b07193>.
- [6] Y. Okamoto, V. Köhler, T. R. Ward, *J. Am. Chem. Soc.* **2016**, *138*, 5781, <https://doi.org/10.1021/jacs.6b02470>.
- [7] A. Chatterjee, H. Mallin, J. Klehr, J. Vallapurackal, A. D. Finke, L. Vera, M. Marsh, T. R. Ward, *Chem. Sci.* **2016**, *7*, 673, <https://doi.org/10.1039/c5sc03116h>.
- [8] T. Heinisch, F. Schwizer, B. Garabedian, E. Csibra, M. Jeschek, J. Vallapurackal, V. B. Pinheiro, P. Marlière, S. Panke, T. R. Ward, *Chem. Sci.* **2018**, *9*, 5383, <https://doi.org/10.1039/c8sc00484f>.
- [9] J. Serrano-Plana, C. Rumo, J. G. Rebelein, R. L. Peterson, M. Barnet, T. R. Ward, *J. Am. Chem. Soc.* **2020**, *142*, 10617, <https://doi.org/10.1021/jacs.0c02788>.
- [10] H. Renata, Z. J. Wang, F. H. Arnold, *Angew. Chemie Int. Ed.* **2015**, *54*, 3351, <https://doi.org/10.1002/anie.201409470>.

- [11] T. Quinto, F. Schwizer, J. M. Zimbron, A. Morina, V. Köhler, T. R. Ward, *ChemCatChem* **2014**, *6*, 1010, <https://doi.org/10.1002/cctc.201300825>.
- [12] Y. M. Wilson, M. Dürrenberger, E. S. Nogueira, T. R. Ward, *J. Am. Chem. Soc.* **2014**, *136*, 8928, <https://doi.org/10.1021/ja500613n>.
- [13] M. Hesticová, T. Heinisch, L. Alonso-Cotchico, J.-D. Maréchal, P. Vidossich, T. R. Ward, *Angew. Chemie Int. Ed.* **2018**, *57*, 1863, <https://doi.org/10.1002/anie.201711016>.
- [14] M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke, T. R. Ward, *Nature* **2016**, *537*, 661, <https://doi.org/10.1038/nature19114>.
- [15] T. Vornholt, F. Christoffel, M. M. Pellizzoni, S. Panke, T. R. Ward, M. Jeschek, *Sci. Adv.* **2021**, *7*, eabe4208, <https://doi.org/10.1126/sciadv.abe4208>.
- [16] G. Qu, A. Li, C. G. Acevedo-Rocha, Z. Sun, M. T. Reetz, *Angew. Chem. Int. Ed.* **2020**, *59*, 13204, <https://doi.org/10.1002/anie.201901491>.
- [17] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, *ACS Synth. Biol.* **2013**, *2*, 83, <https://doi.org/10.1021/sb300037w>.
- [18] H. Leemhuis, V. Stein, A. D. Griffiths, F. Hoffelder, *Curr. Opin. Struct. Biol.* **2005**, *15*, 472, <https://doi.org/10.1016/j.sbi.2005.07.006>.
- [19] A. Rothe, R. N. Surjadi, B. E. Power, *Trends Biotechnol.* **2006**, *24*, 587, <https://doi.org/10.1016/j.tibtech.2006.10.007>.
- [20] U. Markel, K. D. Essani, V. Besirlioglu, J. Schiffels, W. R. Streit, U. Schwaneberg, *Chem. Soc. Rev.* **2020**, *49*, 233, <https://doi.org/10.1039/c8cs00981c>.
- [21] A. Autour, M. Ryckelynck, *Micromachines* **2017**, *8*, 128, <https://doi.org/10.3390/mi8040128>.
- [22] A. K. Price, B. M. Paegel, *Anal. Chem.* **2016**, *88*, 339, <https://doi.org/10.1021/acs.analchem.5b04139>.
- [23] F. W. Y. Chiu, S. Stavakis, *Electrophoresis* **2019**, *40*, 2860, <https://doi.org/10.1002/elps.201900222>.
- [24] K. K. Brower, C. Carswell-Crumpton, S. Klemm, B. Cruz, G. Kim, S. G. K. Calhoun, P. M. Fordyce, *Lab Chip* **2020**, *20*, 2062, <https://doi.org/10.1039/D0LC00261e>.
- [25] K. K. Brower, M. Khariton, P. H. Suzuki, C. Still, G. Kim, S. G. K. Calhoun, L. S. Qi, B. Wang, P. M. Fordyce, *Anal. Chem.* **2020**, *92*, 13262, <https://doi.org/10.1021/acs.analchem.0c02499>.
- [26] A. Stucki, J. Vallapurackal, T. R. Ward, P. S. Dittrich, *Angew. Chem. Int. Ed.* **2021**, <https://doi.org/10.1002/anie.202016154>.
- [27] K. Bernath, M. Hai, E. Mastrobattista, A. D. Griffiths, S. Magdassi, D. S. Tawfik, *Anal. Biochem.* **2004**, *325*, 151, <https://doi.org/10.1016/j.ab.2003.10.005>.
- [28] A. Zinchenko, S. R. A. Devenish, B. Kintses, P. Y. Colin, M. Fischlechner, F. Hoffelder, *Anal. Chem.* **2014**, *86*, 2526, <https://doi.org/10.1021/ac403585p>.
- [29] E. Hardiman, M. Gibbs, R. Reeves, P. Bergquist, *Appl. Biochem. Biotechnol.* **2010**, *161*, 301, <https://doi.org/10.1007/s12010-009-8794-6>.
- [30] R. Tu, R. Martinez, R. Prodanovic, M. Klein, U. Schwaneberg, *J. Biomol. Screen.* **2011**, *16*, 285, <https://doi.org/10.1177/1087057110396361>.
- [31] G. Körfer, C. Pitzler, L. Vojcic, R. Martinez, U. Schwaneberg, *Sci. Rep.* **2016**, *6*, 1, <https://doi.org/10.1038/srep26128>.
- [32] A. C. Larsen, M. R. Dunn, A. Hatch, S. P. Sau, C. Youngbull, J. C. Chaput, *Nat. Commun.* **2016**, *7*, 1, <https://doi.org/10.1038/ncomms11235>.

## License and Terms



This is an Open Access article under the terms of the Creative Commons Attribution License CC BY 4.0. The material may not be used for commercial purposes.

The license is subject to the CHIMIA terms and conditions: (<http://chimia.ch/component/sppagebuilder/?view=page&id=12>).

The definitive version of this article is the electronic one that can be found at <https://doi.org/10.2533/chimia.2021.257>