Organic & Biomolecular Chemistry

Development and Application of a High Throughput Assay System for the Detection of Rieske Dioxygenase Activity

Journal:	*Organic & Biomolecular Chemistry*
Manuscript ID	OB-COM-12-2020-002412.R1
Article Type:	Communication
Date Submitted by the Author:	26-Dec-2020
Complete List of Authors:	Preston-Herrera, Cristina; Ball State University, Chemistry Jackson, Aaron; Ball State University, Chemistry Bachmann, Brian; Vanderbilt University, Chemistry Department Froese, Jordan; Ball State University, Chemistry

COMMUNICATION

Development and Application of a High Throughput Assay System for the Detection of Rieske Dioxygenase Activity

Cristina Preston-Herrera,[a] Aaron S. Jackson,[a] Brian O. Bachmann,[b] Jordan T. Froese*[a]

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Herein we report the development of a new periodate-based reactive assay system for the fluorescent detection of the *cis*-diol metabolites produced by Rieske dioxygenases. This sensitive and diastereoselective assay system successfully evaluates the substrate scope of Rieske dioxygenases and determines the relative activity of a rationally designed Rieske dioxygenase variant library. The high throughput capacity of the assay system enables rapid and efficient substrate scope investigations and screening of large dioxygenase variant libraries.

Rieske dioxygenases that play a role in the bacterial metabolism of aromatic compounds in the soil demonstrate the potential of these organisms to adapt to the carbon sources available in their environment.[1,2] The pathways responsible for the metabolism of aromatic compounds in such organisms typically begin with the regio- and stereoselective dihydroxylation of the aromatic "pollutants", catalyzed by Rieske dioxygenases (**Figure 1**).[1-3] Following his discovery of this enzymatic transformation, Gibson developed a blocked mutant (*P. putida* 39D),[3(c)] and later a transgenic *E. coli* strain (JM109 pDTG601A),[4] which were capable of accumulating the *cis*-diol metabolites in fermentation cultures. To date, the action of Rieske dioxygenases remains the only known enzymatic method for the preparation of *cis*-diol metabolites of this type from aromatic compounds.[5]

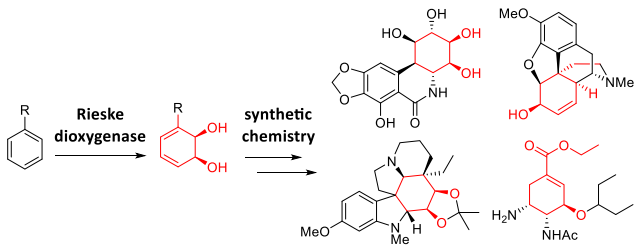

Figure 1: Asymmetric dihydroxylation of aromatic substrates catalyzed by Rieske dioxygenases, and examples of the synthetic utility of the resultant *cis*-diol metabolites.[7(g,n,q,r,y)]

With the availability of organisms capable of accumulating *cis*-diols in fermentation cultures, these metabolites became available in large quantities. The relative promiscuity of Rieske dioxygenases allows these enzymes to recognize and convert a wide range of substrates, resulting in the identification of over 400 *cis*-diol metabolites from toluene dioxygenase alone,[6] the most commonly utilized of these enzymes for synthetic purposes. These factors have facilitated the application of Rieske dioxygenase metabolites in the synthesis of a wide variety of valuable compounds (**Figure 1**).[7]

Despite the relative promiscuity of Rieske dioxygenases, the substrate scopes of these enzymes are limited by the sterics and electrostatics of their potential substrates.[8-10] These limitations have prevented Rieske dioxygenases from being applied in many potential synthetic sequences or have forced the development of circuitous synthetic routes to introduce the large or polar groups which cannot be accommodated by the enzyme. Furthermore, the limitations on the substrate scopes of these enzymes have also precluded their application in the bioremediation of many common aromatic pollutants. To alleviate these restrictions, several research groups have applied enzyme engineering techniques to expand the reactivity of Rieske dioxygenases.[10] These studies demonstrate the power of enzyme engineering to broaden the substrate scope and alter the selectivity of these enzymes. Despite the success of these reports, their scope has been limited by the throughput of the dioxygenase variant assay systems utilized or by the range of potential substrates that are amenable to these screening systems.[10] The available Rieske dioxygenase variant screening methods, including the colorimetric indigo assay, detection of metabolites through High Performance Liquid Chromatography (HPLC) or Gas Chromatography (GC), and the

[a] *Department of Chemistry, Ball State University, 2000 W Riverside Ave, Muncie, IN, 47306, USA.*
[b] *Department of Chemistry, Vanderbilt University, 7330 Stevenson Center, Station B 351822, Nashville, TN, 37235, USA.*
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

detection of real-time activity through the use of oxygen electrodes have recently been reviewed.[11] These methods, while effective, provide limited information on the variant's activity for various substrates (in the case of the indigo assay), or require highly specialized equipment which increases cost and limits throughput.[11] To establish a universal platform for the development of improved Rieske dioxygenases through enzyme engineering, a highly sensitive, accessible, cost-effective and generalizable screening system for these enzyme systems is required.

Over time, technological advances in expression vectors,[12] gene synthesis and cloning,[13] mutagenesis platforms,[14] and high throughput screening[15] have significantly lowered the barrier to entry into engineering enzymes through structure-based design and directed evolution. To facilitate the engineering of Rieske dioxygenase enzymes through directed evolution, the development of a highly sensitive and selective assay system which can detect the relative activity of Rieske dioxygenase variants is required. The successful development of such an assay system will enable the identification of enzymes with expanded substrate scopes from large variant libraries and the facile determination of the substrate scope of any Rieske dioxygenase enzyme.

Results and Discussion

Assay Development

In addition to the Rieske dioxygenase-specific screening methods described,[11] previous enzyme engineering studies have employed various other methods for the detection of enzymatic metabolites with *cis*-diol moieties, including saccharides. These methods include the direct detection of these compounds using fluorescent, boronic acid-functionalized probes,[16] and techniques that rely on the derivatization of these compounds to produce a detectable analyte such as the Gibbs assay.[10(g,h,m),17] Our exploratory studies employing fluorescent boronic acids to detect Rieske dioxygenase metabolites did not demonstrate sufficient sensitivity for practical applications with cell extracts. With the many reports of highly sensitive detection of aldehyde-functionalized compounds using fluorescent probes,[18] we determined to employ the well-known conversion of *cis*-diol metabolites to their corresponding dialdehydes through oxidation with sodium metaperiodate[19] to produce a detectable analyte, as has previously been used in the detection of sialic acid.[19c] This reaction was shown to proceed rapidly and to completion in buffered aqueous solutions using excess sodium metaperiodate (**Figure 2A**). This approach provides the additional advantage of producing a bifunctional dialdehyde analyte, with the theoretical potential to form multiple conjugates with a fluorescent probe. To facilitate initial studies on the proposed assay system, 1 L-scale fermentation of toluene was performed with *E. coli* (JM109 pDTG601A)

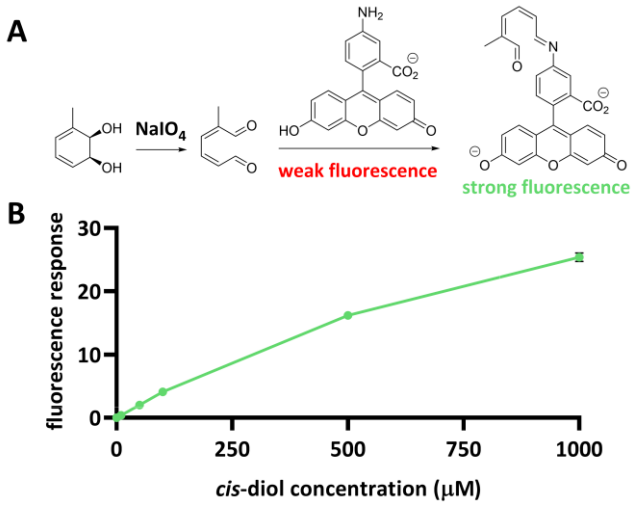

Figure 2: (**A**) Coupled reactions employed by the fluorescence-based assay system for the detection of *cis*-diol metabolites; (**B**) Concentration-dependent fluorescence response of the assay to the presence of *cis*-diol metabolites. All studies were performed in triplicate; all measurements demonstrated S. D. <5%; fluorescence responses normalized to negative control ($[I - I_0] / I_0$).

expressing toluene dioxygenase,[4] the extraction of which provided sufficient quantities of the *cis*-diol metabolite for initial studies. These exploratory studies demonstrated a concentration-dependent fluorescence response when the dialdehyde produced by *cis*-diol oxidation was treated with the fluorescent reporter fluoresceinamine (**Figure 2B**),[21] which has previously been shown to exhibit markedly increased fluorescence upon derivatization of the primary amine,[22] and to readily form fluorescent imine conjugates with aldehydes in aqueous solution.[18(e)]

To maximize the sensitivity of this *cis*-diol detection system, the assay parameters were varied to determine the reaction conditions which would afford the maximum level of sensitivity. Investigation of various buffering systems as the medium for the assay indicated that the greatest fluorescence response could be achieved when the reaction was performed in 0.1 M HEPES (*N*-2-hydroxyethylpiperazine-*N*'-2-ethanesulfonic acid) buffer, while significantly lower fluorescence responses were observed when the reaction was performed in 0.1 M TRIS (tris(hydroxymethyl)aminomethane) buffer (**Figure 3A**). This lowered response in TRIS buffer can be accounted for by the ability of the three hydroxyl moieties of TRIS buffer to coordinate to the metaperiodate oxidant, limiting the reactivity of the oxidant towards the *cis*-diol metabolites.[23] Having identified HEPES buffer as the best-performing buffer medium, the fluorescence response was investigated at various pH levels within the HEPES buffering range.[23] These studies indicated a correlation between the acidity of the HEPES buffer medium and the magnitude of the fluorescence response (**Figure 3B**). A similar result was observed when 0.1 M phosphate buffer was employed as the medium for the assay (**Supplementary Information, Fig. S1**). These results can be attributed to the acid-catalyzed imine formation required to form the fluorescent dialdehyde-probe conjugate. Although fluorescein derivatives

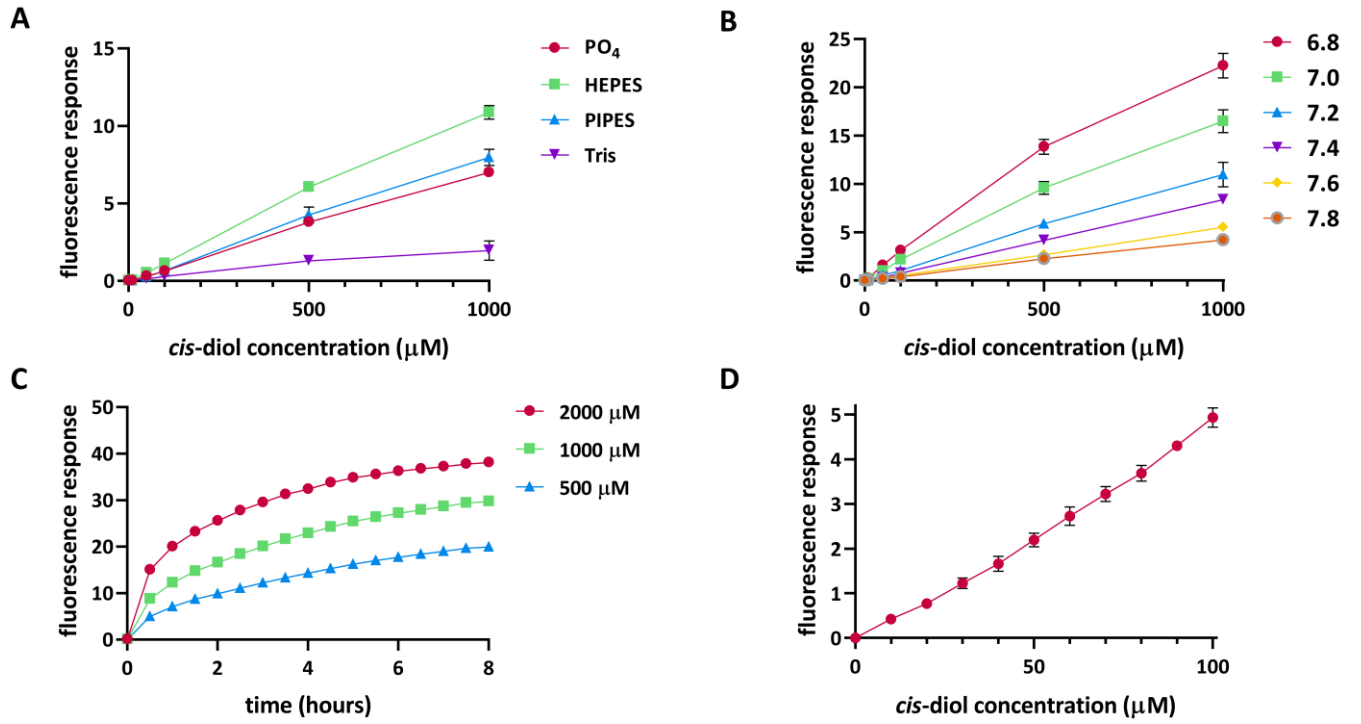

Figure 3: Assay Optimization Studies; (**A**) Fluorescence response in varied buffer solutions (0.1 M, pH 7.2); (**B**) Fluorescence response in 0.1 M HEPES buffer at varied pH levels; (**C**) Time-course fluorescence response in 0.1 M HEPES buffer (pH 6.8); (**D**) Fluorescence response in 0.1 M HEPES buffer (pH 6.8) with low *cis*-diol concentrations. All studies were performed in triplicate, with 0.1 mM fluoresceinamine and 10 mM $NaIO_4$. Data points with error bars not visible demonstrated S. D. <5%. Fluorescence responses normalized to negative control ($[I - I_0] / I_0$).

are known to form multiple, pH-dependent ionic equilibria (pKa of the monoanionic form of fluorescein is ~6.7),[23] and the various ionic states of fluorescein are known to exhibit differing spectral properties,[24,25] only the monoanion and dianion species of fluorescein derivatives are known to exhibit fluorescence.[24(b)] Therefore, this data likely indicates an improved formation of the fluoesceinamine-dialdehyde conjugate under acidic conditions as opposed to an improved fluorescence output for the protonated conjugate. Further studies identified the optimal pH for the assay system, striking a balance between the required acid-catalyzed imine formation and the decreased fluorescence of the protonated conjugate, to be ~6.5 (**Supplementary Information, Fig. S2**). To facilitate the performance of the assay under these conditions without risking acid-catalyzed aromatization of the *cis*-diol metabolites, the oxidation step was performed in buffered solution at pH 7.2, and the subsequent conjugation performed at pH 6.5. pH values were adjusted upon addition of the fluorescent probe solutions, acidified with calibrated amounts of HCl. Time course studies indicated an increase in fluorescence response over extended reaction times (**Figure 3C**), henceforth standardized at 5 hours. Studies varying the concentration of the fluorescent probe confirmed previous reports which reported optimal fluorescence response with the use of a final fluoresceinamine concentration of 0.1 mM (**Supplementary Information, Fig. S3**),[18(e)] and these studies also indicated a maximal fluorescence response at sodium metaperiodate concentrations in large excess from 10 mM - 50 mM (**Supplementary Information, Fig. S4**). These optimization studies demonstrated the capacity of the assay system to produce significant increases in fluorescence with *cis*-diol concentrations at or below 10 μM under the optimized conditions (**Figure 3D**). A relative increase in fluorescence response at analyte concentrations below that of the probe, which could indicate the formation of a bis-probe/dialdehyde adduct in the presence of excess probe was not observed, nor was a relative decrease in fluorescence response at low analyte concentrations that could indicate internal self-quenching[26] of such a bis-adduct (**Figure 3D, Supplementary Information, Fig. S5**). Henceforth, the described optimized assay system is referred to as the (meta)periodate fluorescein *cis*-diol assay (MPFCD).

Having established the utility of MPFCD for detecting *cis*-diol metabolites in buffered aqueous solutions, it remained to determine whether this system could detect the *cis*-diol metabolites produced by respiring cells expressing Rieske dioxygenases. To facilitate increased expression of the toluene dioxygenase genes, novel expression vectors were generated using the pCDF-Duet-1 template. The requisite ferredoxin (todB), reductase (todA) and dioxygenase β (structural)-subunit (todC2) were cloned from pDTG601A[4] into pCDF-Duet-1 MCS-2 to generate the novel pCP-01 vector. This vector, lacking the catalytic dioxygenase subunit, acted as an assay control in all future experiments. The catalytic dioxygenase (todC1) was then cloned into MCS-1 of pCP-01 to generate the catalytically active pCP-02 vector. Using this expression system, a checkerboard assay was performed, wherein a 96-well plate was alternately inoculated with single colonies of *E. coli* (BL21 (DE3) pCP-02)

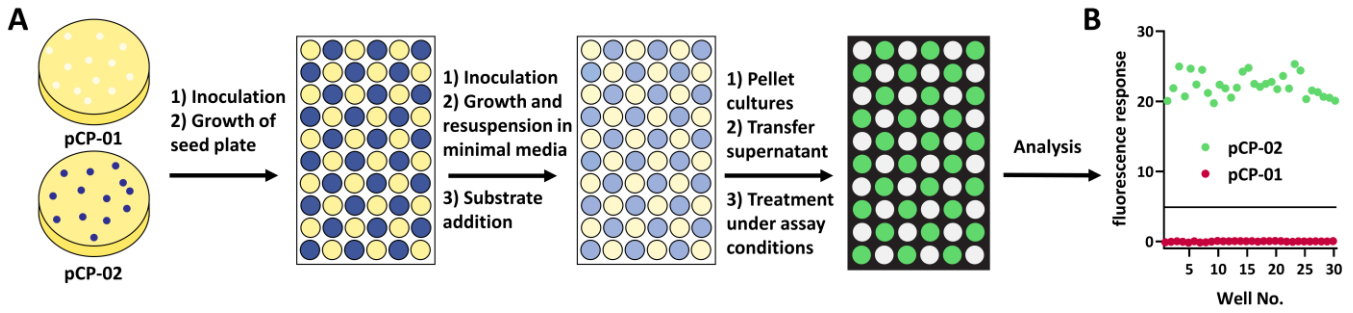

Figure 4: (**A**) Schematic overview of the procedure for the whole-cell checkerboard assay; (**B**) Normalized fluorescence response data for checkerboard assay with 3σ threshold indicated (4.83). The statistical effect size (Z') for this assay was calculated as 0.78. Fluorescence response of cultures expressing toluene dioxygenase (*E. coli* BL21 (DE3) pCP-02) (n = 30) were normalized to the mean fluorescence response of cultures not expressing toluene dioxygenase (*E. coli* BL21 (DE3) pCP-01) (n = 30; $[I - I_0] / I_0$).

expressing the catalytic dioxygenase, and *E. coli* (BL21 (DE3) pCP-01) not expressing the catalytic dioxygenase and grown overnight (**Figure 4**). These seed plates were then used to inoculate fresh 96-well plates where they were grown to exponential phase, pelleted and resuspended in minimal media,[20] following the Marley method.[27] Cultures were subsequently induced and incubated with 2 mM toluene substrate. Following incubation with the substrate, the cultures were pelleted, and the supernatant transferred to a new 96-well assay plate and treated under the MPFCD conditions described. The significant increase in the fluorescence response observed from cultures expressing the full toluene dioxygenase system over those lacking the catalytic dioxygenase (**Figure 4B**) demonstrated the power of this assay system to detect *cis*-diol metabolites produced by respiring cells. Additional studies performed with purified *cis*-diol metabolites and "spent" minimal media confirmed that the fluorescence response detected in this assay was within the linear range of the assay under these conditions (0.5-30 RFU, 10-500 μM, **Supplementary Information, Fig. S5**).

Application in Substrate Scope Determination

The ability to rapidly determine the substrate scope of a given Rieske dioxygenase without the need for enzyme isolation and anaerobic reaction conditions, or the need for medium-scale fermentation and metabolite isolation, has practical applications in synthesis, biocatalyst engineering, and bioremediation. When planning synthetic routes, it is often necessary to determine whether an aromatic compound will be accepted as a substrate, a process which typically necessitates medium-scale fermentation, extraction, and subsequent analysis.[8] Furthermore, genomic studies have putatively identified many dioxygenases of completely untested substrate scope to be present in a variety of organisms.[28] As our MPFCD allows for 96 individual cultures to be grown simultaneously per assay plate, with the potential for multiple plates to be used, we envisioned a valuable application for this assay system in determining the substrate scope of Rieske dioxygenase enzymes. To evaluate this application of the assay system, a 96-well plate of *E. coli* (BL21 (DE3) pCP-02) cultures expressing toluene dioxygenase was prepared, including *E. coli* (BL21 (DE3) pCP-01) cultures as a negative control. These cultures were grown and incubated with a small library of aromatic substrates (**Figure 5A**) as described. The results of this substrate scope analysis (**Figure 5B**) were consistent with the established substrate scope parameters of toluene dioxygenase,[6,8-10] including decreasing activity being detected with an increase in the steric size of alkyl substrates (toluene, ethylbenzene, *n*-propylbenzene, *n*-butylbenzene), and the lack of a significant fluorescent response from highly polar substrates (phenethylamine, benzamide, etc.). The ability to rapidly evaluate a wide range of potential aromatic substrates for Rieske dioxygenases will provide a valuable tool for synthetic chemists in evaluating potential synthetic sequences, and in

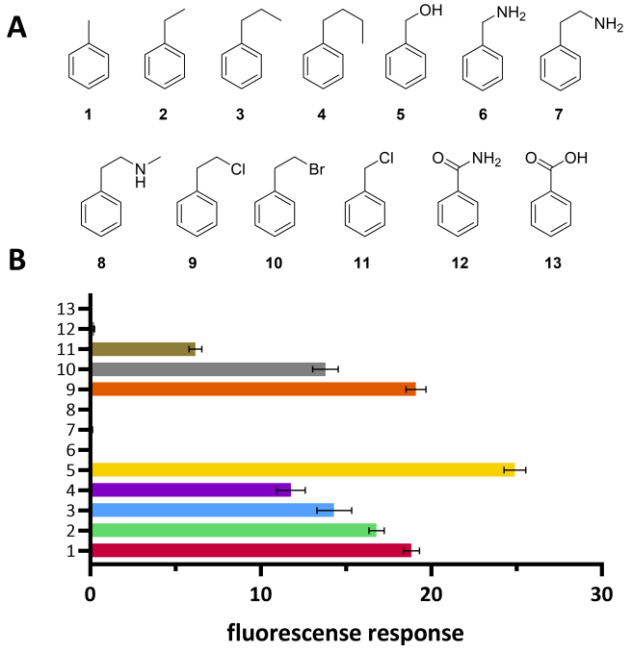

Figure 5: Toluene dioxygenase substrate scope is analyzed using the designed assay system (n = 4 for each substrate). Fluorescence responses normalized to negative control ($[I - I_0] / I_0$)].

determining the substrate scope for as yet uncharacterized Rieske dioxygenases.

Application in Enzyme Engineering

The primary goal of the assay development described was to provide a high throughput means of assessing the relative activity of Rieske dioxygenase variants. Such an assay system would afford the ability to perform directed evolution studies with the goal of expanding the substrate scope of Rieske dioxygenases. To evaluate the utility of MPFCD for this application, saturation mutagenesis was applied to produce a library of toluene dioxygenase variants.[29] For proof of concept studies, the leucine-272 active site residue was selected for saturation mutagenesis for its proximity to bound substrates and for its potential to be mutated to a more polar residue, which could potentially coordinate and stabilize more polar substrates (**Figure 6**).[30,31] A plasmid library of mutations at this position was generated via PCR using the 'NNK' protocol. Following transformation of the plasmid library bearing toluene dioxygenase leucine-272 mutants, single colonies from this library were inoculated into individual wells in a 96-well plate, and these seed plates were incubated overnight. Seed plate cultures were used to inoculate fresh assay plates and then stored in 60% glycerol at -80 °C. Assay plates were grown according to the Marley method as described,[26] and were incubated with the native substrate toluene. Following incubation with the aromatic substrate, assay plates were treated under MPFCD conditions to detect the relative amount of *cis*-diol metabolite produced by leucine- 272 variants (**Figure 7B**). To confirm the initial results of this assay, representative examples of plasmids containing both active and inactive mutant dioxygenases (as determined by MPFCD) were isolated from cultures stored in seed plates. Isolated mutant plasmids were retransformed into *E. coli* BL21 (DE3), which were cultured and carried through the assay system again. Repetition of the assay protocol with isolated mutant plasmids confirmed the relative activity of the enzyme variants identified as active for the native substrate (toluene) in the first analysis (**Figure 7C**). Results for the variant identified as inactive in the first analysis were similarly confirmed. Mutants applied in validation studies were submitted for sequencing, which confirmed mutations from leucine to phenylalanine, and leucine to valine as retaining activity towards the native substrate (**Figure 7C**). A mutation from leucine to proline was shown to abolish enzyme's activity for the native toluene substrate. To determine whether the identified active site mutations resulted in alterations to the substrate selectivity of the enzyme, the L272F and L272V variants were subjected to screening with a small library of substrates (**Figure 8A**) as described. Although the L272F variant demonstrated a similar reactivity profile to the parent enzyme (**Supplementary Information, Fig. S7**), the L272V mutation was shown to significantly alter the substrate selectivity relative to the parent (**Figure 8B**). This active site mutation resulted in a shift in substrate selectivity, which favoured sterically larger substrates (ethylbenzene, *n*-propylbenzene, (2-chloroethyl)benzene) over the native substrate (toluene), and in a near complete loss of activity for the benzyl alcohol substrate. These results represent the first identification of the leucine-272 active site residue of toluene dioxygenase as playing a significant role in determining the substrate selectivity of the enzyme. These initial studies clearly demonstrate the utility of MPFCD in determining the relative activity of dioxygenase variants, and determining the effects of specific point mutations on the enzyme's activity. Further, these results also demonstrate the potential of this assay system in identifying Rieske dioxygenase variants with significantly altered reactivity.

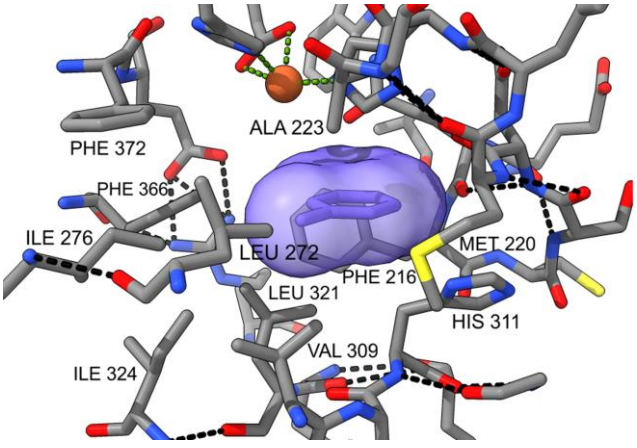

Figure 6: Visualization of the toluene dioxygenase active site with bound substrate (toluene, purple).[30] Residues within 7.0 Å of the bound substrate are included. Image was generated with ChimeraX software.[32]

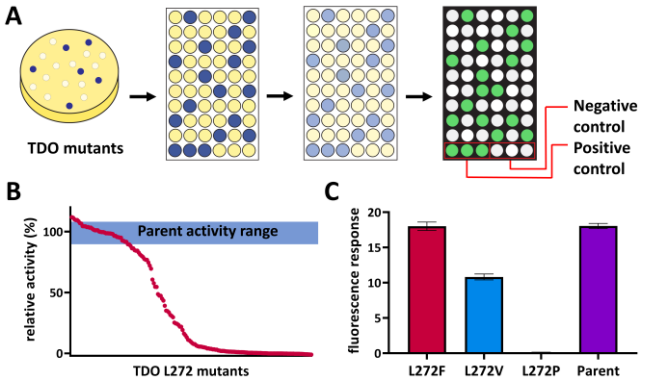

Figure 7: (**A**) Schematic overview of the procedure for the whole-cell toluene dioxygenase L272 variant screening assay; (**B**) Activity of toluene dioxygenase L272 variants relative to the parent enzyme (n = 172). Fluorescence response of each variant was normalized to the mean fluorescence response of the negative control ($[I - I_0] / I_0$)]. Parent activity range was determined from positive controls (*E. coli* BL21 (DE3) pCP-02), with the fluorescence response normalized to the negative control ($[I - I_0] / I_0$)]; (**C**) Validation of initial assay results confirms the activity of L272F and L272V variants, and the inactivity of the L272P variant (n = 8). Fluorescence responses normalized to the mean fluorescence response of the negative control ($[I - I_0] / I_0$)].

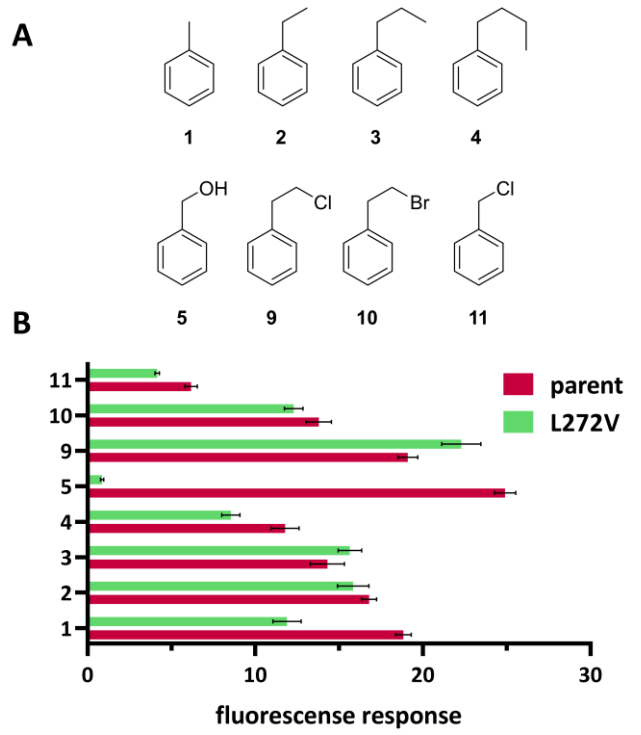

Figure 8: Toluene dioxygenase L272V variant substrate scope is analyzed using the designed assay system and compared to the activity observed with the parent enzyme (n = 4 for each substrate). Fluorescence responses normalized to negative control ([I – I_0] / I_0)].

Conclusions

In summary, we have developed a sensitive and general assay system for the diastereoselective detection of Rieske dioxygenase activity, which we have herein named as MPFCD. Optimization studies identified conditions under which the greatest levels of sensitivity can be achieved from MPFCD, including the use of non-coordinating buffering systems as the reaction medium and mildly acidic conditions for conjugation. Proof of concept studies confirmed the utility of MPFCD in analyzing the substrate scope of Rieske dioxygenases and in determining the relative activity of Rieske dioxygenase variants. In the latter application, the assay system facilitated the identification of active site mutations that are or are not tolerated and the identification of enzyme variants with significantly altered reactivity. In comparison to existing screening systems employed in Rieske dioxygenase engineering studies that rely on the intrinsic fluorescence/absorbance properties of potential substrates[10(d,e,l)], the tendency of metabolites to rearomatize to produce a detectable analyte,[10(g,h,m)] or the variant's specific activity for indole,[11] MPFCD offers substantially enhanced sensitivity and additional versatility due to its reliance only on the presence of a vicinal diol moiety for analyte detection. Further, MPFCD is suitable for high-throughput applications unlike most GC/HPLC-based or oxygen electrode-based screening platforms and can be performed without the need for specialized equipment unlike the Flow-Injection Analysis (FIA) techniques that have been applied in high-throughput screening.[11] The use of only readily available and low cost reagents, and the lack of need for specialized equipment makes MPFCD an accessible and economical option for the high-throughput screening of Rieske dioxygenase variants. Unlike oxygen electrode-based screening platforms however,[11] the detection of the fluorescent analyte by MPFCD is temporally separated from the production of the desired metabolites owing to the need for the removal of the respiring cells and pH adjustment prior to analysis. For this reason, MPFCD does not provide information on the rate of oxidation by Rieske dioxygenases. Future studies will expand the mutagenesis and screening work, with MPFCD being applied in identifying Rieske dioxygenase variants with expanded substrate scope or improved selectivity. The development of such Rieske dioxygenases will provide valuable synthetic tools by broadening the range of *cis*-diol metabolites that are available for synthetic applications, provide a starting point for engineering, and determine the scope of environmental remediation pathways.

Experimental Section

General experimental

E. coli JM109 (pDTG601A)[4] was generously donated by Prof. David T. Gibson (1938–2014) and Prof. Rebecca Parales (University of California, Davis). *E. coli* JM109 competent cells were obtained from Agilent Technologies. *E. coli* BL21 (DE3) competent cells were obtained from ThermoFisher. Plasmid isolation/purification was performed using New England Biolabs Monarch® miniprep kit. Transformations of electrocompetent cells were performed on an Eppendorf Eporator®. Whole cell assay cultures were grown in Greiner Bio-One polystyrene clear, round-bottom 96-well plates. Fluorescence analyses were performed using a Biotek® SynergyTM H1 monochromator-based multi-mode plate reader, using Corning® polystyrene black, opaque flat-bottom 96-well plates. All reagents were obtained from MilliporeSigma unless otherwise stated. Media were made at pH 7.2 and ampicillin or streptomycin were added at 100 and 50 µg/mL respectively. All *E. coli* cultures were maintained at 37 °C unless otherwise stated.

Preparative scale production of (1*S*,2*R*)-3-methylcyclohexa-3,5-diene-1,2-diol[20,27]

E. coli JM109 electrocompetent cells were transformed with isolated pDTG601A plasmid containing the toluene dioxygenase genes,[4] and selected on LB agar containing ampicillin overnight. Single colonies were inoculated into 5 mL LB medium containing ampicillin and incubated with shaking overnight. Cultures (2 x 500 mL LB with ampicillin in 2000 mL Erlenmeyer flasks) were inoculated with 5 mL overnight culture each and incubated with

shaking. Growth of the cultures was monitored via optical density measurement at 600 nm. Upon reaching an OD_{600} of 0.5 - 0.6 AU, 500 mL cultures were pelleted and resuspended in minimal media (KH_2PO_4 – 7.5 g/L; citric acid – 2 g/L; $MgSO_4{\cdot}7H_2O$ - 5 g/L; trace metal solution - 2 mL/L [Na_2SO_4 - 1 g/L; $MnSO_4$ - 2 g/L; $ZnCl_2$ - 2 g/L; $CoCl_2{\cdot}6H_2O$ - 2 g/L; $CuSO_4{\cdot}5H_2O$ - 0.3 g/L; $FeSO_4{\cdot}7H_2O$ - 10 g/L; pH 1.0]; conc. H_2SO_4 - 1.2 mL/L; ferric ammonium citrate - 0.3 g/L; glucose – 4 g/L; thiamine (34 mg/L); pH 7.2)[7k] containing ampicillin. After 1 h of recovery in minimal media, the cultures were induced to a final concentration of 1 mM IPTG and the incubation temperature decreased to 30 °C. After a 2 h induction period, neat toluene substrate was added directly via pipette to the cultures, in two portions, to a final concentration of 2 mM. The cultures were incubated with the toluene substrate for 3 h, and subsequently pelleted and the supernatant decanted. The combined supernatant was then extracted with 3 x 500 mL EtOAc, and the combined extracts were dried over anhydrous $MgSO_4$. The dried extract was concentrated and the (1*S*,2*R*)-3-methylcyclohexa-3,5-diene-1,2-diol metabolite was purified by flash column chromatography (3:2 EtOAc:hexanes) (~120 mg isolated yield).

Assay protocol for optimization studies

Stock solutions of (1*S*,2*R*)-3-methylcyclohexa-3,5-diene-1,2-diol were prepared in 0.1 M specified buffer systems. Stock solutions of $NaIO_4$ were freshly prepared immediately prior to use in the assay. Stock solutions of fluoresceinamine were prepared in methanol. 100 µL of (1*S*,2*R*)-3-methylcyclohexa-3,5-diene-1,2-diol stock solutions were transferred in 96-well black opaque assay plates (Greiner Bio-One, Cat. # CLS3915). Negative controls were included for each assay plate (buffer media without (1*S*,2*R*)-3-methylcyclohexa-3,5-diene-1,2-diol was used, negative controls otherwise treated in an identical fashion). Oxidative cleavage was initiated by adding 50 µL of $NaIO_4$ stock solution to each well, and the assay plate was incubated with shaking at room temperature for 30 mins. Aldehyde detection was performed by the addition of 50 µL of fluoresceinamine stock solution (0.4 mM) to each well. Assays plates were incubated with shaking at room temperature for 5 h (unless otherwise stated). Fluorescence response from each well was analyzed at 485 nm (ex), 520 nm (em), and normalized to the fluorescence response of the negative controls ($[I - I_0] / I_0$)].

Vector Construction

A 2.1-kbp DNA fragment including todC2BA was PCR amplified from pDTG601A[4] using primers TODC2BAF (GTATAAGAAGGAGATATACAATGATTGATTCAGCCAAC) and TODC2BAR (TATCCAATTGAGATCTGCCATCACGTTAGGTCTCCTTC) and cloned into the NdeI site of pCDF-Duet-1 using NEBuilder® HiFi DNA Assembly (New England Biolabs) to yield plasmid pCP-01. todC1 (1.4-kbp) was then amplified from pDTG601A[4] using primers TODC1F (ACTTTAATAAGGAGATATACATGAATCAGACCGACACATCAC) and TODC1R (TGATGGTGATGGCTGCTGCCTCAGCGTGTCGCCTTCAG) and cloned into the NcoI site of pCP-01 using NEBuilder® HiFi DNA Assembly (New England Biolabs) to yield plasmid pCP-02.

Whole-cell fermentation 96 well-plate assay protocol[20,27]

E. coli BL21 (DE3) electrocompetent cells were transformed with isolated pCP-02 plasmids expressing toluene dioxygenase (parent and/or mutant libraries), and with isolated pCP-01 plasmids as negative controls. The transformation cultures were incubated on LB + streptomycin plates overnight. Single colonies were inoculated into 160 µL LB + streptomycin media with 0.3% glucose in a 96-well round bottom seed plate and incubated with shaking overnight. All plates included 3 or more wells containing *E. coli* BL21 (DE3) pCP-02 cells expressing the parent toluene dioxygenase enzyme (positive control), and 3 or more wells containing *E. coli* BL21 (DE3) pCP-01 (negative control). Seed plates were used to inoculate 5uL into 155 µL LB media containing streptomycin in a fresh 96-well round bottom assay plate, and the cultures were incubated with shaking 2.5 h. The assay plates were then pelleted, and the supernatant discarded. Cultures were resuspended in 160 uL minimal media (see above) containing streptomycin and incubated for a 1 h recovery period. Following this, the cultures were induced to a final concentration of 0.5 mM IPTG and the incubation temperature reduced to 30 °C. After a 2 h induction period, aromatic substrates were added as 68 mM stock solutions in DMSO to a final concentration of 2 mM. Cultures were incubated with aromatic substrates for 1 h at 30 °C, after which the cultures were pelleted. A 100 µL portion of supernatant from each well was transferred to 96-well black opaque assay plates. Reaction was initiated by adding a 50 µL of $NaIO_4$ stock solution to each well to a final concentration of 10 mM, and the assay plates were incubated with shaking at room temperature for 30 mins. Cleaved diols were detected by adding 50 µL of fluoresceinamine stock solution (prepared with 3 µL conc. HCl (11.65 M) / 1 mL fluoresceinamine solution) to each well to a final concentration of 0.1 mM. Assays plates were incubated with shaking at room temperature for 5 h. Fluorescence response from each well was analyzed at 485 nm (ex), 520 nm (em), and normalized to the mean fluorescence response of the negative controls ($[I - I_0] / I_0$)].

Mutant library generation

pCP-02 was used as the template for toluene dioxygenase mutant library generation. Saturation mutagenesis was performed using the QuikChange® XL II kit (Agilent Technologies). Amplification was performed using an ABI GeneAmp® 9700 Thermal Cycler. Mutagenic primers were designed with the aid of the QuikChange® primer design tool (Agilent Technologies) (toluene dioxygenase L272 forward primer – CTATGTCGGCGACCCCAATNNKATGCTTGCCATCATGGGG; reverse primer – CCCCATGATGGCAAGCATMNNATTGGGGTCGCCGACATAG).

Sequencing analyses were performed by GenHunter© corporation (Nashville, TN).

Acknowledgements

This work was made possible in part by the Ball State University Provost Start-Up program and by a Vanderbilt University Discovery Award grant to JF and BB. This material is based upon work supported by the National Science Foundation under Grant No. HRD 1618408, 2016-2021 (Indiana STEM Louis Stokes Alliance for Minority Participation) and by the Indiana Academy of Science Senior Research Grant Program (IASSG-F20-01). The authors thank undergraduate research assistants Phillip Betts, Olivia Susi and Kelsey Scott for their valuable contributions to this manuscript.

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 P. A. Williams and J. R. Sayers, *Biodegradation,* 1994, **5**, 195-217.

2 M. M. Abu-Omar, A. Loaiza and N. Hontzeas, *Chem. Rev*., 2005, **105**, 2227-2252.

3 (a) D. T. Gibson, J. R. Koch and R. E. Kallio, *Biochemistry,* 1968, **7**, 2653-2662; (b) D. T. Gibson, J. R. Koch, C. L. Schuld and R. E. Kallio, *Biochemistry,* 1968, **7**, 3795-3802; (c) D. T. Gibson, M. Hensley, H. Yoshioka and T. J. Mabry, *Biochemistry,* 1970, **9**, 1626-1630; (d) H. Ziffer, K. Kabuto, D. T. Gibson, V. M. Kobal and D. M. Jerina, *Tetrahedron,* 1977, ***33*****,** 2491-2496.

4 G. J. Zylstra and D. T. Gibson, *J. Biol. Chem.,* 1989, **264**, 14940-14946.

5 (a) Y. Feng, C. Y. Ke, G. Xue and L. Que. *Chem. Commun.,* 2009, 50-52; (b) E. H. Southgate, J. Pospech, J. Fu, D. R. Holycross and D. Sarlah, *Nat. Chem*., 2016, **8**, 922-928; (c) W. C. Wertjes, E. H Southgate and D. Sarlah, *Chem. Soc. Rev*., 2018, **47**, 7996-8017; (d) M. Okumura and D. Sarlah, *Synlett*, 2018, **29**, 845-855; (e) C. J. Huck and D. Sarlah, *Chem*., 2020, **6**, 1589-1603; (f) M. Okumura and D. Sarlah, *Chimia*, 2020, **74**, 577-583; (g) M. Okumura and D. Sarlah, *Eur. J. Org. Chem*., 2020, **10**, 1259-1273.

6 R. A. Johnson, *Org. React.,* 2004, **63**, 117-264.

7 Selected examples of the application of Rieske dioxygenase metabolites in synthesis: (a) A. M. Reiner, *J. Bacteriol*., 1971, **108**, 89-94; (b) W.-K. Yeh, D. T. Gibson and T.-N. Liu, *Biochem. Biophys. Res. Commun*., 1977, **78**, 401-410; (c) V. Subramanian, T.-N. Liu, W. K. Yeh and D. T. Gibson, *Biochem. Biophys. Res. Commun*., 1979, **91**, 1131-1139; (d) D. A. Widdowson, D. W. Ribbons and S. D. Thomas, *Janssen Chimica Acta.,* 1990, **8**, 3-9; (e) H. A. J. Carless, *Tetrahedron: Asymm*., 1992, **3**, 795-826; (f) T. Hudlicky and J. W. Reed, in *Advances in Asymmetric Synthesis,* ed. A. Hassner, JAI Press, Greenwich, 1995, pp. 271-312; (g) X. Tian, T. Hudlicky and K. Königsberger, *J. Am. Chem. Soc*., 1995, **117**, 3643-3644; (h) T. Hudlicky and A. J. Thorpe, *Chem. Commun*., 1996, 1993-2000; (i) T. Hudlicky, D. A. Entwistle, K. K. Pitzer and A. J. Thorpe, *Chem. Rev*., 1996, **96**, 1195-1220; (j) D. R. Boyd and G. N. Sheldrake, *Nat. Prod. Rep*., 1998, **15**, 309-324; (k) M. A. Endoma, V. P. Bui, J. Hansen and T. Hudlicky, *Org. Proc. Res. Dev*., 2002, **6**, 525-532; (l) M. G. Banwell, A. J. Edwards, G. J. Harfoot, K. A. Jolliffe, M. D. McLeod, K. J. McRae, S. G. Stewart and M. Vogtle, *Pure Appl. Chem*., 2003, **75**, 223-229; (m) D. R. Boyd and T. H. Bugg, *Org. Biomol. Chem*., 2006, **4**, 181-192; (n) A. T. Omori, K. J. Finn, H. Leisch, R. J. Carroll and T. Hudlicky, *Synlett,* 2007, 2859-2862; (o) T. Hudlicky and J. W. Reed, *Chem. Soc. Rev*., 2009, **38**, 3117-3132; (p) T. Hudlicky, J. W. Reed, *Synlett,* 2009, **5**, 685-703; (q) B. Sullivan, I. Carrera, M. Drouin and T. Hudlicky, *Angew. Chem. Int. Ed*., 2009, **48**, 4229-4231; (r) L. Werner, A. Machara and T. Hudlicky, *Adv. Synth. Catal*., 2010, **352,** 195-200; (s) J. F. Trant, J. Froese and T. Hudlicky, *Tetrahedron: Asymm*., 2013, **24**, 184-190; (t) S. E. Lewis, *Chem. Commun*., 2014, **50**, 2821-2830; (u) V. Varghese and T. Hudlicky, *Angew. Chem. Int. Ed*., 2014, **53**, 4355-4358; (v) S. Vshyvenko, M. R. Reisenauer; S. Rogelj, T. Hudlicky. *Bioorg. Med. Chem. Lett*., 2014, **24**, 4236-4238; (w) S. E. Lewis, in *Asymmetric Dearomatization Under Enzymatic Conditions,* ed. S.-L. You, Wiley, Chichester, 2016, p. 279-337; (x) J. Froese, C. Overbeeke and T. Hudlicky, *Chem. Eur. J*., 2016, **22**, 6180-6184; (y) L. V. White and M. G. Banwell, *J. Org. Chem*., 2016, **81**, 1617-1626; (z) D. Baidilov, L. Rycek, J. F. Trant, J. Froese, B. Murphy and T. Hudlicky, *Angew. Chem. Int. Ed.,* 2018, **57**, 10994-10998.

8 J. Froese, M.-A. Endoma-Arias and T. Hudlicky, *Org. Proc. Res. Dev.,* 2014, **18**, 801-809.

9 G. N. Sheldrake, in *Chirality in Industry: the commercial manufacture and application of optically active compounds*, ed. A. N. Collins, G. N. Sheldrake and J. Crosby, John Wiley & Sons Ltd, Chichester, 1992, pp. 127–166.

10 (a) B. D. Erickson and F. Mondello, *J. Appl. Environ. Microbiol*., 1993, **59**, 3858-3862; (b) N. Kimura, A. Nishi, M. Goto and K. Furukawa, *J. Bacteriol*., 1997, **179**, 3936-3943; (c) T. Kumamaru, H. Suenaga, M. Mitsuoka, T. Watanabe and K. Furukawa, *Nat. Biotechnol*., 1998, **16**, 663-666; (d) N. Zhang, B. G. Stewart, J. C. Moore, R. L. Greasham, D. K. Robinson, B. C. Buckland and C. Lee, *Metab. Engineering,* 2000, **2**, 339-348; (e) R. E. Parales, S. M. Resnick, C. L. Yu, D. R. Boyd, N. D. Sharma and D. T. Gibson, *J. Bacteriol*., 2000, **182**, 5495-5504; (f) C. L. Yu, R. E. Parales and D. T. Gibson, *J. Ind. Microbiol. Biotechnol*., 2001, **27**, 94-103; (g) T. Sakamoto, J. M. Joern, A. Arisawa and F. H. Arnold, *Appl. Env. Microbiol*., 2001, **67**, 3882-3887; (h) J. M. Joern, T. Sakamoto, A. Arisawa, F. H. Arnold, *J. Biomol. Screen*., 2001, **6**, 219–223; (i) H. Suenaga, T. Watanabe, M. Sato, Ngadiman and K. Furukawa, *J. Bacteriol*., 2002, **184**, 3682-3688; (j) R. E. Parales, *J. Ind. Microbiol. Biotechnol*., 2003, **30**, 271-278; (k) K. Pollmann, V. Wray, H. J. Hecht and D. H. Pieper, *Microbiology,* 2003, **149**, 903-913; (l) L. M. Newman, H. Garcia, T. Hudlicky and S. A. Selifonov, *Tetrahedron,* 2004, **60**, 729-734; (m) E. L. Ang, J. P. Obbard and H. M. Zhao, *Appl. Microbiol. Biotechnol*., 2009, **81**, 1063-1070; (n) C. Gally, B. M. Nestl and B. Hauer, *Angew. Chem. Int. Ed*., 2015, **54**, 12952-12956; (o) M. A. Vila, D. Umpierrez, N. Veiga, G. Seoane, I. Carrera and S. R. Giordano, *Adv. Synth. Catal*., 2017, **359**, 2149-2157; (p) J. M. Halder, B. M. Nestl, and B. Hauer, *ChemCatChem,* 2018, **10**, 178-182.

11 J. L. Wissner, W. Escobedo-Hinojosa, P. M. Heinemann, A. Hunold, and B. Hauer, in *Methods in Enzymology, Vol. 644*, ed. D. S. Tawfik, Academic Press, Cambridge, 2020, pp. 63-93.

12 G. L. Rosano and F. A.Ceccarelli, *Front. Microbiol*., 2014, **5**, 1-17.

13 (a) C. Engler, R. Kandzia and S. Marillonnet, *PLOS One,* 2008, **3**, e3647(1-7); (b) D. G. Gibson, L. Young, R.Y. Chuang, J. C. Venter, C. A. III Hutchison and H. O. Smith, *Nature Methods*, 2009, **6**, 343-345.

14 (a) M. T. Reetz, M. Bocola, J. D. Carballeira, D. X. Zha and A.Vogel, *Angew. Chem. Int. Ed.,* 2005, **44**, 4192-4196; (b) M.

T. Reetz, S. Prasad, J. D. Carballeira, Y. Gumulya and M. Bocola, *J. Am. Chem. Soc.,* 2010, **132**, 9144-1952.

15 (a) M. Olsen, B. Iverson and G. Georgiou, *Curr. Opin. Biotechnol*., 2000, **11**, 331-337; (b) D. Wahler and J. L. Reymond, *Curr. Opin. Chem. Biol*., 2001, **5**, 152-158; (c) S. V. Taylor, P. Kast and D. Hilvert, *Angew. Chem. Int. Ed*., 2001, **40**, 3310-3335; (d) Y. L. Boersma, M. J. Droge, and W. J. Quax, *FEBS J.,* 2007, **274**, 2181-2195; (e) J. A. Dietrich, A. E. McKee and J. D. Keasling, *Ann. Rev. Biochem.,* 2010, **79**, 563-590.

16 (a) H. S. Cao and M. D. Heagy, *J. Fluores*., 2004, **14**, 569-584; (b) F. E. Cappuccio, J. T. Suri, D. B. Cordes, R. A. Wessling and B. Singaram, *J. Fluoresc*., 2004, **14**, 521-533; (c) X. Wu, Z. Li, X. X. Chen, J. S. Fossey, T. D. James and Y. B. Jiang, *Chem. Soc. Rev*., 2013, **42**, 8032-3048; (d) Y. J. Huang, W. J. Ouyang, X. Wu, Z. Li, J. S. Fossey, T. D. James and Y. B. Jiang, *J. Am. Chem. Soc*., 2013, **135**, 1700-1703; (e) G. F. Whyte, R. Vilar and R. Woscholski, *J. Chem. Biol*., 2013, **6**, 161-174; (f) X. Sun and T. D. James, *Chem. Rev*., 2015, **115**, 8001-8037; (g) S. Xu, A. C. Sedgwick, S. A. Elfeky, W. Chen, A. S. Jones, G. T. Williams, T. A. Jenkins, S. D. Bull, J. S. Fossey and T. D. James, *Front. Chem. Sci. Eng*., 2020, **14**, 112-116.

17 M. G. Quintana, C. Didion and H. Dalton, *Biotechnol. Techniques*, 1997, **11**, 585-587.

18 (a) S. Uzu, S. Kanda, K. Imai, K. Nakashima and S. Akiyama, *Analyst,* 1990, **115**, 1477-1482; (b) L. Nondek, R. E. Milofsky and J. W. Birks, *Chromatographia,* 1991, **32**, 33-39; (c) M. Konarzycka-Bessler and U. T. Bornscheuer, *Angew. Chem. Int. Ed*., 2003, **42**, 1418-1420; (d) R. Yang, K. Li, N. Li, F. L. Zhao and W. H. Chan, *Anal. Chem*., 2003, **75**, 3908-3914; (e) Y. Xing, S. Wang, X. Mao, X. Zhao and D. Wei, *J. Fluores.,* 2011, **21**, 587-594; (f) A. K. Ressmann, D. Schwendenwein, S. Leonhartsberger, M. D. Mihovilovic, U. T. Bornscheuer, M. Winkler and F. Rudrof, *Adv. Synth. Catal*., 2019, **361**, 2538-2543.

19 (a) B. T. Golding, G. Kennedy and W. P. Watson, *Tetrahedron Lett*., 1988, **29**, 5991-5994; (b) E. Abu-Shqara and J. Blum, *J. Heterocyclic Chem*., 1991, **28**, 473-475; (c) J. A. Key, C. Li and C. W. Cairo, *Bioconjugate Chem.,* 2012, **23**, 363-371.

20 M. A. Endoma, V. P. Bui, J. Hansen and T. Hudlicky, *Org. Proc. Res. Dev.,* 2002, **6**, 525-532.

21 A. Baeyer, *Ber. Dtsch. Chem. Ges.,* 1871, **4**, 555-558.

22 C. Munkholm, D.-R. Parkinson and D. R. Walt, *J. Am. Chem. Soc*., 1990, **112**, 2608-2612.

23 (a) G. Gomori, *Methods Enzymol*., 1955, **1**, 138-146; (b) N. E. Good, G. D. Winget, W. Winter, T. N. Connolly, S. Izawa and R. M. M. Singh, *Biochemistry,* 1966, **5**, 467-477; (c) N. E. Good and S. Izawa, *Methods Enzymol*., 1972, **24**, 53-68; (d) W. J. Ferguson, K. I. Braunschweiger, W. R. Braunschweiger, J. R. Smith, J. J. McCormick, C. C. Wasmann, N. P. Jarvis, D. H. Bell and N. E. Good, *Anal. Biochem.,* 1980, **104**, 300-310.

24 (a) H. Leonhardt, L. Gordon and R. Livingston, *J. Phys. Chem*., 1971, **75**, 245-249; (b) M. M. Martin and L. Lindqvist, *J. Luminesc*., 1975, **10**, 381-390; (c) J. Yguerabide, E. Talavera, J. M. Alvarez and B. Quintero, *Photochem. Photobiol*., 1994, **60**, 435-441; (d) R. Sjöback, J. Nygren and M. Kubista, *Spectrochim. Acta A.,* 1995, **51**, 7-21; (e) N. Klonis and W. H. Sawyer, *J. Fluoresc.,* 1996, **6**, 147-157.

25 L. D. Lavis, T. J. Rutkoski and R. T. Raines, *Anal. Chem.,* 2007, **79**, 6775-6782.

26 (a) A. Kawski, *Photochem. Photobiol*. 1983, **38**, 487-508; (b) L.W. Runnels and S.F. Scarlata, *Biophys. J*. 1995, **69**, 1569-1583; (c) J. Karolin, M. Fa, M. Wilczynska, T. Ny and B.-A. Johansson, *Biophys. J*. 1998, **74**, 11-21; (d) X. Zhuang, T. Ha, H.D. Kim, T. Centner, S. Labeit and S. Chu, *Proc. Natl. Acad. Sci. USA*, 2000, **97**, 14241-14244; (e) J.R. Lakowicz, J. Malicka, S. D'Auria and I. Gryczynski, *Anal. Biochem*. 2003, 320, 13-20.

27 J. Marley, M. Lu and C. Bracken, *J. Biomol. NMR,* 2001, **20**, 71-75.

28 (a) G. J. Zylstra and E. J. Kim, *Ind. Microbiol. Biotechnol.*, 1997, **19**, 408-414; (b) J. Armengaud, B. Happe and K. N. Timmis, *J. Bacteriol.,* 1998, **180**, 3954-3966; (c) D. T. Gibson, *J. Ind. Microbiol. Biotechnol*., 1999, **23**, 284-293; (d) E. Kim and G. J. Zylstra, *J. Ind. Microbiol. Biotechnol.*, 1999, **23**, 294-302; (e) M. F. Romine, L. C. Stillwell, K.-K. Wong, S. J. Thurston, E. C. Sisk, C. Sensen, T. Gaasterland, J. K. Fredrickson and J. D. Saffer, *J. Bacteriol*., 1999, **181**, 1585-1602; (f) D. Pérez-Pantoja, R. Donoso, L. Agulló, M. Córdova, M. Seeger, D. H. Pieper and B. González, *Environ. Microbiol.*, 2012, **14**, 1091-1117; (g) N. Sathyanarayanan and H. G. Nagendra, *Bioinformation*, 2012, **8**, 1154-1161; (h) N. Sathyanarayanan and H. G. Nagendra, *Bioinformation*, 2014, **10**, 68-75.

29 M. T. Reetz, and J. D. Carballeira, *Nature Protocols*, 2007, **2**, 891-903.

30 R. Friemann, K. Lee, E. N. Brown, D. T. Gibson, H. Eklund and S. Ramaswany, *Acta Crys.,* 2009, **D65**, 24-33.

31 R. O. Burns and R. D. Demoss, *Biochim. Biophys. Acta*., 1962, **65**, 233-244.

32 T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris and T. E. Ferrin, *Protein Sci.,* 2018, **27**, 14-25.