A Labeled Substrate Approach to Discovery of Biocatalytic Reactions: A Proof of Concept Transformation with N-Methylindole

Jamie L. Rogers and John B. MacMillan*

Department of Biochemistry, The University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9038, United States

S *Supporting Information*

ABSTRACT: Biocatalysis has become an important method in the pharmaceutical industry for the incorporation of new functionality in small molecules. Currently this method is limited in the types of reactions that can be carried out and no strategy exists to systematically screen for new biocatalyzed reactions. This study involves the development of a medium throughput screen to identify and optimize new reactions using a series of marine-derived bacterial cell lines, which were screened against several ^{13}C labeled organic substrates. The reactions were analyzed using ^{13}C NMR as the primary screening tool. We describe the discovery of a bacterial catalyzed indole oxidation reaction in which complete conversion of ^{13}C labeled N-methyl indole to 3-hydroxyindole was observed. In addition, the sensitivity of this reaction to dO_2 levels can be exploited to oxidize to either 3-hydroxyindole or 2-oxoindole. This new platform sets up an important tool for the discovery of new organic transformations using an extensive library of marine bacteria.

Biocatalysis has become an important component in the toolbox of the pharmaceutical and fine chemical industries, being used as reagents in multistep synthetic processes.[1] The biggest role for biocatalysis in the pharmaceutical sector lies in the chemo-, regio-, and stereoselectivity properties that biological systems can achieve.[2] Additionally biocatalyzed reactions can be carried out in H_2O, without the use of protecting groups and under mild temperatures. Examples of processes that utilize biocatalysis in the manufacture of pharmaceuticals include Crixivan, where *Rhodococcus sp.* is used to perform a stereoselective epoxidation in the first step of the synthesis.[3] A second generation process route to Lyrica includes the use of a lipolase catalyzed resolution to generate a key (S)-mono acid enantiomer from a racemic cyanodiester.[4] An efficient enzymatic desymmetrization route was recently reported for the antidepression drug Paxil, where the key step involves a protease catalyzed desymmetrization of a meso diester, doubling the yield of the existing manufacturing process (Figure 1).[5]

As the pharmaceutical industry continues to expand their capabilities in biocatalysis, there is still a limitation in the strategies to identify new reactions and enzymes. The majority of biocatalyzed reactions focus on desymmetrizations and resolutions. The more widespread application of large-scale biocatalysis is currently restricted because of the limited
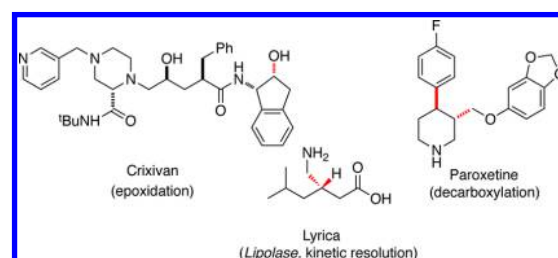

Figure 1. Pharmaceuticals produced by biocatalysis.

number of commercially available synthetically useful enzymes and limited development speed. Discovery methods that can tap into the full potential of enzyme catalyzed reactions, such as carbon bond-forming reactions (C−C, C−N, C−O, C−X) will expand the toolbox for this growing field.

We have developed a library of marine-derived bacteria (Actinomycetes, Bacilli, α-proteobacteria) that has been utilized in a number of drug discovery efforts.[6] We felt the library presented a unique opportunity to develop a platform for the exploration of new, whole-cell biocatalyzed organic transformations. By screening various substrates in a range of fermentation conditions we could search for interesting whole-cell biotransformations. There are many examples of screening microbial libraries for specific reactions, such as whole-cell biocatalyzed Baeyer−Villiger reactions, but we are unaware of unbiased approaches to biocatalysis discovery.[7] We desired a method to discover new or unexpected reactivity in an easy to detect method, such as that developed by the Liu lab on DNA-templated synthesis and *in vitro* selection to rapidly evaluate combinations of substrates for bond-forming reactions.[8] Additional approaches such as "accelerated serendipity" by the lab of David MacMillan for the development of new metal catalyzed reactions[9] and an innovative MS screening approach by Hartwig and co-workers have enhanced reaction discovery.[10]

The development of a robust, medium throughput method to accomplish this goal posed some unique challenges. Because we are using whole-cell fermentation with natural product producing microorganisms, there is significant background metabolite production, which interferes with analysis of reactions.[11] We desired a method to analyze a significant number (>100) of transformations and look specifically for changes to the substrate of choice. To this end, we selected substrates **1−5** for initial evaluation of the biotransformation
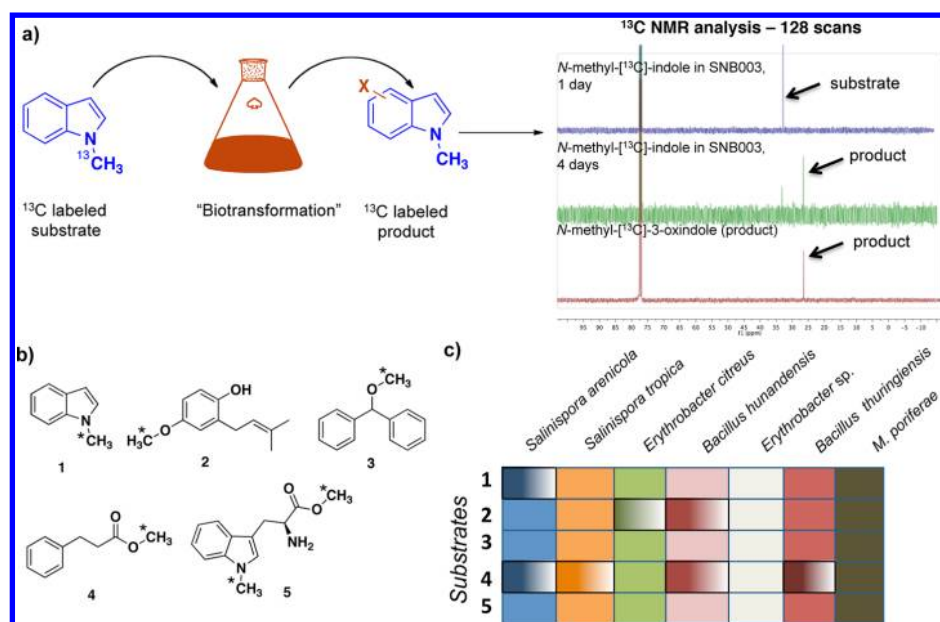
Figure 2. (a) Approach to discovery of biotransformations with readout of ^{13}C labeled substrates by ^{13}C NMR. (b) Labeled substrates used in proof-of-concept screen. (c) Matrix of bacteria and substrates used in initial screen. Boxes with gradient shading indicate bacteria/substrate combinations that resulted in a new product.

discovery platform looking for substrates that fit the following criteria: (1) substrates with motifs similar to those found in natural products and/or pharmaceuticals, (2) substrates wherein interesting biosynthetic reactions had been observed, and (3) the ability to incorporate a ^{13}C isotope label (Figure 2b).

The strategy that we developed utilizes the incorporation of a ^{13}C label on the exogenous substrate allowing for use of ^{13}C NMR spectroscopy as a primary screening tool (Figure 2a). Compared to the low natural abundance of ^{13}C, the use of labeled substrates provides a sensitive technique to distinguish products of biotransformation. As demonstrated in Figure 1a, the ^{13}C NMR of N-[^{13}C]methylindole only shows a single peak, which upon fermentation with the actinomycete *Salinispora arenicola* derives a new product with an upfield chemical shift.

An initial screen was performed in which five ^{13}C labeled substrates were added individually to seven unique strains of bacteria (Figure 2c). The reactions were monitored over the course of several days by removing a small aliquot of the culture and extracting with ethyl acetate. The extracts were analyzed via ^{13}C NMR and LC-MS. Shaded boxes in Figure 1c represent substrate/bacteria combinations where we saw conversion to one or more new products. We were able to find enzymatic conversion with four of the five substrates that were tested in this platform. Here, we will focus on the complete description of one particular biotransformation.

We were particularly drawn to one biotransformation; the conversion of N-[^{13}C]methylindole into an oxidized product by the bacterium *Salinispora tropica*, strain SNB-003. Analysis of the product revealed oxidation to N-[^{13}C]methyl-3-hydroxyindole (6) (Figure 3). This is an unusual oxidation of N-methylindole or indole, with the typical oxidation product under both synthetic and biocatalysis conditions being 2-oxindole.[12] Of the few reports of direct indole oxidation, they all suffer from poor yields and complex mixtures. Oxidation at the 3-position of an indole is a common structural feature in natural products; as such it is not surprising to find a bacterial
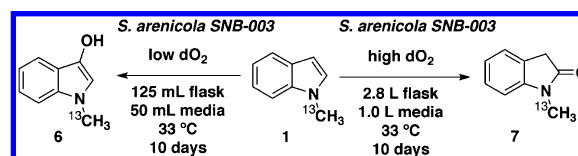
mediated oxidation in our screening approach. To verify the conversion was not a result of air oxidation or from the fermentation media, we carried out a series of control experiments with **1** in the absence of bacteria and saw no conversion to **6**. In addition to being a common structural motif in natural products, oxidized indole substrates are common in pharmaceuticals.[13] Typically the synthetic route to an oxidized indole is through oxidation of a simpler substrate followed by ring closing to get the indole. Direct oxidation on indole may allow for a more efficient synthesis of these molecules.[14]

Using the initial screening conditions, a 125 mL Erlenmeyer flask containing bacteria was charged with 10 mg of **1**, grown for 10 days and subsequently purified via HPLC to reveal complete conversion to N-[^{13}C]methyl-3-hydroxyindole, with no starting material or side products recovered. To follow up on these results, the initial conditions were scaled up into a 2.8 L Fernbach flask. After observing complete conversion of **1**, the resulting product was isolated and determined to be an alternative oxidation product, N-[^{13}C]methyl-2-oxindole (**7**). The only variable in this reaction was the size of the flask and the volume of media, which can change the amount of dissolved oxygen (dO$_2$) in the media (2.8 L Fernbach flasks are designed for aeration of cultures). This leads to the possibility that there are two enzymatic pathways at work, operating under different oxygen levels.

To probe this hypothesis we carried out a series of whole-cell biotransformations at various dO$_2$ levels. By altering the media volume in standard 500 mL Erlenmeyer flasks we were able to

Figure 3. Biocatalysis of **1** to products **6** and **7** under different oxygen concentrations.

vary the dO_2 concentration in each culture and cover a wide range of relative O_2 levels from 21 to 91% (Air = 100%) (Tables 1 and S1). We measured dO_2 concentration using an

Table 1. Role of % dO_2 on Conversion of 1 to 6 or 7 under Whole-Cell Biotransformation Conditions

entry	%**6**	%**7**	% dO_2 relative to air
1	0	100	91
2	0	100	89
3	0	100	81
4	100	0	77
5	63	27	73
6	100	0	67
7	70	30	61
8	80	0	61
9	75	0	58
10	100	0	58
11	60	0	50
12	100	0	46
13	40	0	44
14	40	0	39
15	30	0	21

Oakton submersible dissolved oxygen probe. At dO_2 levels lower than 45% there was little conversion of 1 to either of the oxidized indole products (Table 1, entries 13−15). In the dO_2 range of 45%−80% there was nearly quantitative conversion to the desired oxidized product 6 (Table 1, entries 4−12). There is a dramatic switch at 80% relative dO_2 where production of 6 is completely eliminated and uniform conversion to 7 is obtained (Table 1, entries 1−3). We can conclude that to obtain the 3-oxindole product most efficiently the oxygen level in the fermentation flask should be maintained at a relative dO_2 of 60%.

With these optimized dO_2 conditions we obtained a 35% overall isolated yield of 6. Interestingly, no other product was observed in this reaction and no starting material was recovered to account for the low yield. This could be due to challenges in extraction of 6 from media, loss of the ^{13}C label, or incorporation of the indole into the organism. To overcome this problem we looked at the possibility of moving the reaction from whole-cell biotransformation to an enzymatic preparation from the bacteria.

An enzyme preparation was performed with *S. arenicola* SNB003 in which the cells were grown in A1 media for 10 days and then lysed via repeated sonication. The large, insoluble particles were separated from the lysate via centrifugation at 5200 rpm (15 min, 4 °C), and the presence of proteins was established via Western blot. 1 was added to the flask maintained at 4 °C for four days. Multiple minor products were observed by LC-MS and ^{13}C NMR; however the two major products, following purification, were identified as the desired 3-hydroxyindole 6 and the higher molecular weight indole trimer 8 (Figure 4). Compounds 6 and 8 accounted for the 75% yield of this reaction in a 1:4 ratio of products. 8 is a known compound and is also structurally related to a previously reported natural product, which is lacking the *N*-methyl groups.[15] The trimer has been postulated to arise from a 3-hydroxyindole intermediate.[16] When the cell-free extract reaction is allowed to proceed for seven days, only 8 is observed, indicating that 6 is a substrate for further reactivity. Oxidation of indoles is a common transformation by a variety
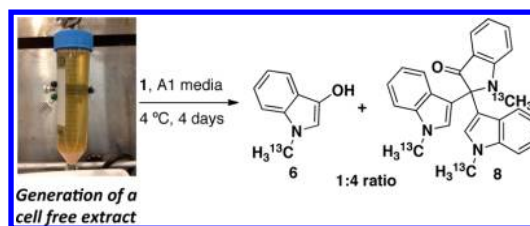

Figure 4. Cell-free extract conversion.

of P450s and monooxygenases; however oxidation results in the formation of multiple products and is driven toward formation of indigo or indirubin.[17] In the case of the whole-cell biotransformation of 1 with *S. arenicola*, the oxidation can be controlled to give 6, making this unique compared to previous examples. We are carrying out further studies to determine the oxidase responsible for the biotransformation.

We probed the substrate scope of the biotransformation using a variety of substituted *N*-methylindole and indole substituents (Table 1). The substrate scope is somewhat limited at this time, with more electron-rich substrates being readily oxidized with complete conversion to the 3-oxindole product (entries *a*−*d*). However, the electron-poor substrates we tested were completely unreactive and only starting material was recovered (entries *h* and *i*). Entries *f* and *g* represent substrates found in a number of natural products. 1,2,3,4-Tetrahydrocarbazole (*f*), lacking the *N*-methyl, was oxidized at the 3-position, which induced a rearrangement to the known spirocycle 20. The product 20 has been reported from the oxidation of tetrahydrocarbazole as early as 1951 by Witkop and co-workers.[18] A radical mechanism is proposed for this reaction in which homolytic cleavage of the NH bond initiates the oxidation. However, with *N*-methyl-1,2,3,4-tetrahydrocarbazole no reaction was observed and starting material was quantitatively recovered (entry *g*). We believe this is related to the fact that 21 is the least soluble substrate in Table 2 and is unlikely to be incorporated into the cells.

The development of additional "reagents" for biocatalysis is an important avenue of research as the pharmaceutical industry continues to build their capabilities for manufacturing scale biological approaches to chemistry. We have been able to use a biodiscovery platform to develop an enzyme catalyzed, regiospecific oxidation of a common aromatic moiety that is present in a number of biologically active natural products. Further efforts to isolate the enzyme responsible for the transformation would allow directed evolution approaches to optimize for a given substrate. Further efforts to look for stereospecific incorporation of oxygen in 3-alkylindoles would provide additional synthetic utility. However, this reaction establishes a robust discovery platform for whole-cell biocatalysis.

There are two technical challenges to be considered in using the approach: maintenance of the bacterial screening collection and detection. Advancements in microscale fermentation (12-well to 96-well plates)[19] make it simple to generate bacterial screening plates, and relatively little microbiology expertise is needed to carry out a screen, through the use of cryogenically stored screening plates. The only instrumentation needed for these studies is a small shaker table and well-plate compatible centrifuge. One challenge that we have encountered that has limited throughput is the differential growth rates of the bacteria in our collection. In order to utilize well plates efficiently it has been important to optimize growth conditions

Table 2. Substrate Scope of Whole-Cell Indole Oxidation

of each bacteria strain independently. From the analytical side, although ^{13}C NMR is not typically viewed as high-throughput, NMR systems with autosamplers and ^{13}C optimized probes allow for hundreds of samples to be run in a 12-h period. To enhance the throughput, we have begun looking at the possibility of adding multiple substrates with orthogonal ^{13}C labeling. The continual identification of ever more interesting enzymes in natural products biosynthesis justifies an unbiased approach to harness microbial capabilities for synthetically useful biotransformations.

ASSOCIATED CONTENT

Supporting Information

General procedures, chemical derivatization, compound characterization, and data table. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

john.macmillan@utsouthwestern.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

REFERENCES

(1) Schmid, A.; Dordick, J. S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. *Nature* **2001**, *409*, 258. (b) Straathof, A. J.; Panke, S.; Schmid, A. *Curr. Opin. Biotechnol.* **2002**, *13*, 548−. (c) Tao, J.; Xu, J. H. *Curr. Opin. Chem. Biol.* **2009**, *13*, 43.

(2) Pollard, D. J.; Woodley, J. M. *Trends Biotechnol.* **2006**, *25*, 66.

(3) Buckland, B. C.; Drew, S. C.; Connors, N. C.; Chartrain, M. M.; Lee, C.; Salmon, P. M.; Gbewonyo, K.; Zhou, W.; Gailliot, P.; Singhvi, R.; Olewinski, R. C.; Sun, W. J.; Reddy, J.; Zhang, J.; Jackey, B. A.; Taylor, C.; Goklen, K. E.; Junker, B.; Greasham, R. L. *Metab. Eng.* **1999**, *1*, 63.

(4) Martinez, C. A.; Hu, S.; Dumond, Y.; Tao, J.; Kelleher, P.; Tully, L. *Org. Process Res. Dev.* **2008**, *12*, 392.

(5) Glaser, V. *Gene Eng. Biotechnol. News* **2008**, *28*, 30.

(6) Hu, Y.; Legako, A. G.; Espindola, A.-P. D. M.; MacMillan, J. B. *J. Org. Chem.* **2012**, *77*, 3401. (b) Pan, E.; Jamison, M.; Yousufuddin, M.; MacMillan, J. B. *Org. Lett.* **2012**, *14*, 2390. (c) Hu, Y.; MacMillan, J. B. *Org. Lett.* **2011**, *13*, 6580. (d) Hu, Y.; Espindola, A.-P. D. M.; Steward, N. A.; Wei, S.; Posner, B. A.; MacMillan, J. B. *Bioorg. Med. Chem.* **2011**, *19*, 5183.

(7) Alphand, V.; Furstoss, R. *J. Mol. Catal. B: Enzym.* **2000**, *9*, 209.

(8) Kanan, M. W.; Rozenman, M. M.; Sakurai, K.; Syder, T. M.; Liu, D. R. *Nature* **2004**, *431*, 545.

(9) McNally, A.; Prier, C. K.; MacMillan, D. W. C. *Science* **2011**, *334*, 1114.

(10) Robbins, D. W.; Hartwig, J. H. *Science* **2011**, *333*, 1423.

(11) Although LC-MS profiling is an option, many low molecular weight compounds do not ionize well and have weak UV chromophores making detection of new molecules difficult.

(12) Medici, R.; Garaycoechea, J. I.; Dettorre, L. A.; Iribarren, A. M.; Lewkowicz, E. S. *Biotechnol. Lett.* **2011**, *33*, 1999. (b) Trost, B. M.; Kie, J.; Sieber, J. D. *J. Am. Chem. Soc.* **2011**, *133*, 20611. (c) Ling, K.; Li, W.; Sayre, L. M. *J. Am. Chem. Soc.* **2008**, *130*, 933.

(13) Jimenez, J. L.; Huber, U.; Moore, R. E.; Patterson, G. M. L. *J. Nat. Prod.* **1999**, *62*, 569. (b) Birch, A. L.; Wright, I. I. *J. Chem. Soc., Chem Commun.* **1969**, *12*, 644−. (c) Liu, J. R.; Jiang, Z. Y.; Wang, R. R.; Zheng, Y. T.; Chen, J. L.; Zhang, X. M.; Ma, Y. B. *Org. Lett.* **2007**, *9*, 4127.

(14) Onyeibor, O.; Croft, S. L.; Dodson, H. I.; Feiz-Haddad, M.; Kendrick, H.; Millington, N. J.; Parapini, S.; Phillips, R. M.; Seville, S.; Shnyder, S. D. *J. Med. Chem.* **2005**, *48*, 2701. (b) Gaywood, A. P.; McNab, H. *Synthesis* **2010**, *8*, 1361. (c) Azadi-Ardakani, M.; Alkhader, M. A.; Lipplatt, J. H.; Patel, D. I.; Smalley, R. R.; Higson, S. *J. Chem. Soc., Perkin Trans. 1* **1986**, *6*, 1107.

(15) Kiraz, C. I. A.; Enge, T. J.; Jimenez, L. S. *J. Org. Chem.* **2004**, *69*, 2200.

(16) Ganachaud, C.; Garfagnoli, V.; Tron, T.; Lacazio, G. *Tetrahedron Lett.* **2008**, *49*, 2476.

(17) Ensley, B. D.; Ratzkin, B. J.; Osslund, T. D.; Simon, M. J.; Wackett, L. P.; Gibson, D. T. *Science* **1983**, *222*, 167. (b) Murdock, D.; Ensley, B. D.; Serdar, C.; Thalen, M. *Bio/Technology* **1993**, *11*, 381. (c) Mermod, N.; Harayama, S.; Tinnis, K. N. *Bio/Technology* **1986**, *4*, 321. (d) Meyer, A.; Wursten, M.; Schmidt, A.; Kohler, H-P. E.; Witholt, B. *J. Biol. Chem.* **2002**, *277*, 34161.

(18) Wiktop, B.; Patrick, J. B. *J. Am. Chem. Soc.* **1951**, *73*, 2188.

(19) (a) Ferreira-Torres, C.; Michelleti, M.; Lye, G. J. *Biosyst. Eng.* **2005**, *28*, 83. (b) Minas, W.; Bailey, J. E.; Duetz, W. *Antonie Leeuwenhoek* **2000**, *78*, 297.

12381

dx.doi.org/10.1021/ja304767m | *J. Am. Chem. Soc.* 2012, 134, 12378−12381