

Understanding Flavin-Dependent Halogenase Reactivity via Substrate Activity Profiling

Mary C. Andorfer,[†] Jonathan E. Grob,[‡] Christine E. Hajdin,[‡] Julia R. Chael,[†] Piro Siuti,[‡] Jeremiah Lilly,[‡] Kian L. Tan,^{*,‡} and Jared C. Lewis^{*,†,ID}

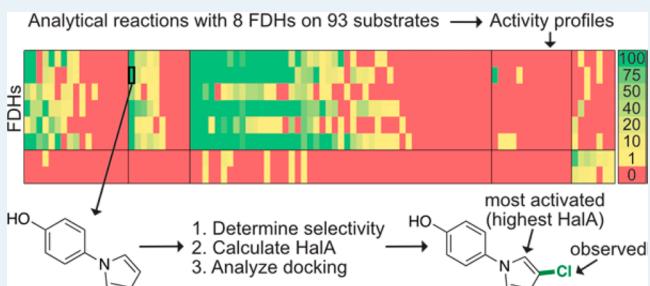
[†]Department of Chemistry, University of Chicago, Chicago, Illinois 60637, United States

[‡]Global Discovery Chemistry, Novartis Institutes for Biomedical Research, 250 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

Supporting Information

ABSTRACT: The activity of four native FDHs and four engineered FDH variants on 93 low-molecular-weight arenes was used to generate FDH substrate activity profiles. These profiles provided insights into how substrate class, functional group substitution, electronic activation, and binding affect FDH activity and selectivity. The enzymes studied could halogenate a far greater range of substrates than have been previously recognized, but significant differences in their substrate specificity and selectivity were observed. Trends between the electronic activation of each site on a substrate and halogenation conversion at that site were established, and these data, combined with docking simulations, suggest that substrate binding can override electronic activation even on compounds differing appreciably from native substrates. These findings provide a useful framework for understanding and exploiting FDH reactivity for organic synthesis.

KEYWORDS: halogenase, flavin, C–H functionalization, biocatalysis, site selective



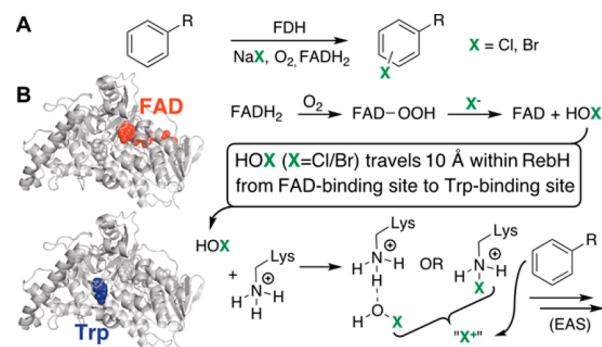
INTRODUCTION

The importance of halogenation for the synthesis and function of organic compounds has driven extensive efforts to identify^{1,2} and engineer³ halogenases to selectively install halogen substituents. Among the several classes of halogenases identified to date, flavin-dependent halogenases (FDHs) have proven particularly promising in this regard.⁴ FDHs for which in vitro activity has been established halogenate electron-rich arenes (**Scheme 1A**). While some FDHs require substrates linked to carrier proteins, many have activity on free substrates, which greatly simplifies biocatalysis and engineering efforts.¹

Tryptophan halogenases (Trp-FDHs), including RebH,⁵ PrnA,⁶ and PyrH,⁷ are perhaps the best studied FDHs with activity on free substrates. Mechanistic studies on RebH and PrnA suggest that these enzymes promote electrophilic aromatic substitution of enzyme-bound tryptophan by a formal halonium ion (X^+) donor, which is proposed to be either a lysine-derived haloamine (K79 in RebH)⁸ or bound HOX⁹ (**Scheme 1B**). This mechanism is consistent with the fact that relatively electron rich substrates are halogenated. Other factors are clearly important, however, as less electronically activated sites can be halogenated in the presence of more electronically activated sites.¹⁰

A number of fungal halogenases with activity on phenols and anisoles have also been identified.¹ For example, Rdc2, an FDH from *C. chiversii*, was found to halogenate a variety of resorcyclic

Scheme 1. (A) General Scheme for FDH-Catalyzed Halogenation and (B) Mechanism for Generation of Proposed Halonium for Electrophilic Aromatic Substitution (EAS) within FDHs



acid lactone substrates, the acyclic natural product curcumin, and two hydroxyisoquinolines.^{11,12} Gsfl, a homologue of Rdc2 from *P. aethiopicum*, catalyzes a selective chlorination in the biosynthesis of griseofulvin.¹³ No information on how the

Received: September 21, 2016

Revised: December 9, 2016

Published: January 31, 2017

substrate scope of these enzymes compares with that of Trp-FDHs has been reported.

Despite the sequence and structural homology of Trp-FDHs, we^{14,15} and others^{16,17} have observed significant differences in their activity toward non-native substrates. Likewise, the conservation of key sequence motifs in FDHs in general belies the apparent specificity of Trp, fungal, and likely other FDHs for distinct substrate classes on the basis of the limited exploration of substrates reported to date.¹ We therefore became interested in establishing detailed profiles^{18,19} of FDH activity toward a diverse panel of substrates (substrate activity profiles). These profiles provided insights into how substrate class, functional group substitution, electronic activation, and binding affect FDH activity and selectivity. These data, in turn, provide a useful framework for understanding and exploiting FDH reactivity in organic synthesis.

RESULTS

FDHs and Substrates Selected for Activity Profiles.

FDHs were selected to evaluate differences in the activity and selectivity between wild-type Trp-FDHs (RebH, Thal),^{20,21} engineered RebH variants (OK, 4V, 6TL, and 10S),^{10,15} and fungal halogenases (Rdc2 and Gsfl).^{11,13} The scope of RebH has been examined to a limited extent,^{15,20,22} but far less is known about the activity of Thal, a Trp-6-FDH, other than its ability to halogenate D- and L-tryptophan.²¹ Similarly, while Rdc2 activity has been confirmed on a handful of non-native phenol-containing substrates,^{11,12} there are no previous reports of Gsfl activity on non-native substrates. RebH variant OK was created by replacing active site residue Glu461 with Lys on the basis of the success of the analogous mutation in PrnA to increase activity on benzoic acids.¹⁶ Variant 4V was developed to accept larger, biologically active compounds,¹⁵ while variants 6TL and 10S were engineered during an effort¹⁰ to alter the selectivity of tryptamine halogenation.

When the sequence homology of these eight FDHs is assessed, they fall into two categories according to the organism of origin (Figure 1). RebH and Thal, both of which are

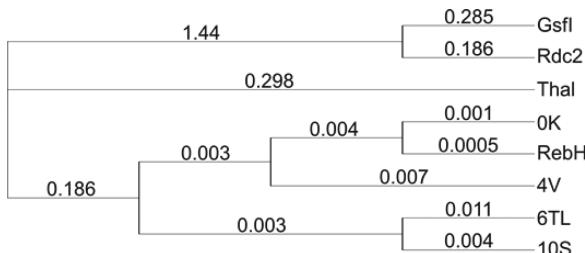


Figure 1. Phylogenetic tree of FDHs examined. Branch labels display amino acid substitutions per site.

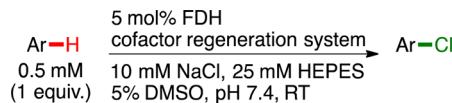
bacterial enzymes, have much higher sequence identities with respect to one another (61% identity, 99% query coverage) than to RebH and the fungal FDH Rdc2 (43% identity, 48% query coverage). We hypothesized this dramatic difference in primary sequence could lead to significant differences in scope, even beyond that previously observed between wild-type and engineered Trp-FDHs.

A substrate panel^{18,23} comprised of 86 N-containing compounds (panel 1, Table S1 in the Supporting Information) was used to evaluate FDH activity. Because halogenation with FDHs is believed to proceed via electrophilic aromatic

substitution,⁶ substrates in this panel contained at least one electronically activated site. In addition, substrates were selected to probe significant steric and electronic variation, as well as functional group substitution, across a range of substrate classes, including anilines, indoles, azoles, and pyrroles. A second panel (panel 2) of seven phenols and anisoles was also constructed, taking into consideration the native substrates of the fungal FDHs.¹ The 93 compounds contained within these two panels constitute the largest survey of FDH substrate scope to date. Moreover, these compounds are representative of those commonly used for fragment-based drug design;²⁴ therefore, understanding their reactivity toward FDHs could provide information on FDH activity toward motifs found in pharmaceuticals.

FDH Substrate Activity Profiles. The eight FDH genes were coexpressed with the GroEL/ES genes, and the resulting FDHs were purified as previously described for RebH.²⁰ This protocol produced high yields of all eight FDHs, including the fungal enzymes Gsfl (~35 mg/L) and Rdc2 (~60 mg/L). Purified enzyme was then used to conduct two analytical bioconversions on each substrate (Scheme 2). One set of

Scheme 2. General Scheme for FDH Bioconversions



reactions was allowed to proceed overnight, while the other was quenched after 1 h to ensure that monohalogenation could be observed on more reactive substrates for which dihalogenation is possible.^{15,25} Despite differences in optimal reaction conditions for the FDHs, conditions were standardized to directly compare conversions. Reactions were analyzed by LCMS, and conversion data were used to generate substrate activity profiles (Figure 2 and Table S1 in the Supporting Information).

Trp-FDH Activity on Anilines, Indoles, Pyrroles, and Azoles. The Trp-FDHs exhibited broad scope toward panel 1 substrates; 67% of these electron-rich arenes were halogenated to some extent (Figure 2). Notably, activity on a wide range of substituted anilines was observed, indicating that the narrow range of anilines evaluated to date¹⁶ could be significantly expanded. The Trp-FDHs also demonstrated remarkable functional group tolerance by accepting substrates containing amines, alcohols, esters, amides, sulfonamides, nitriles, thioethers, pyridines, quinolones, azoles, and pyrroles. Qualitatively similar scope was observed for Thal, 6TL, and RebH, which each halogenated 40–47% of substrates examined with >1% conversion. In contrast, the scope and activity for variants 10S and OK are more limited (>1% conversion for 23–24% of substrates). In the case of OK, this could result from the fact that this enzyme was designed for higher activity on benzoic acid substrates,¹⁶ but few panel 1 substrates contain an acidic functionality. Variant 10S was evolved to halogenate the 5-position of tryptamine with high selectivity, while 6TL, an intermediate in the 10S lineage, halogenated the 5-, 6-, and 7-positions of tryptamine with similar efficiency.¹⁰ The improved scope of 6TL relative to 10S is thus consistent with other examples in which evolutionary intermediates serve as more general catalysts in comparison to variants with high activity/selectivity on target substrates.^{26,27} Interestingly, the FDH with the broadest scope toward panel 1 substrates was 4V. This

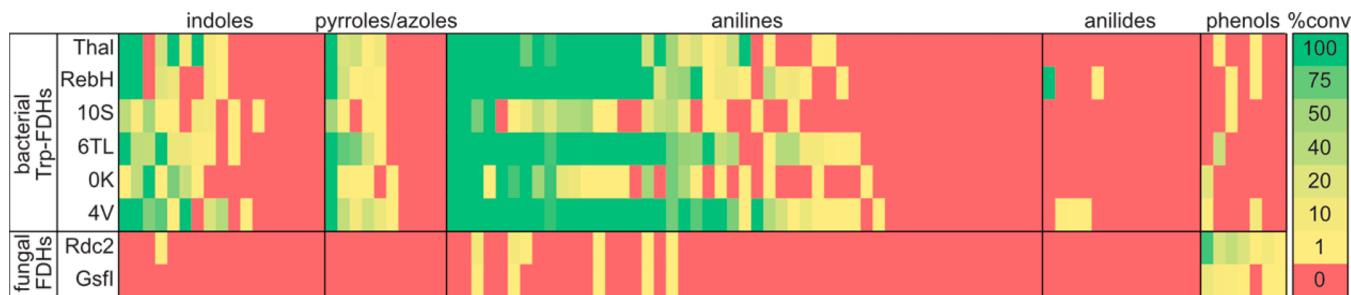


Figure 2. Substrate activity profiles in heat map form for eight FDHs on substrates in panels 1 (indoles, pyrroles, azoles, anilines, anilides) and 2 (phenols). Maximum conversion is shown for each enzyme–substrate pair (for complete data sets at both time points, see Table S1 in the Supporting Information).

variant was engineered to accept large indole-containing compounds,¹⁵ and this appears to have enabled activity on a range of substrates wider than that used in the original engineering effort.

Closer examination of the activity profiles revealed significant differences in the specific substrates halogenated by each Trp-FDH (Figure 3A). Many instances arose in which significant

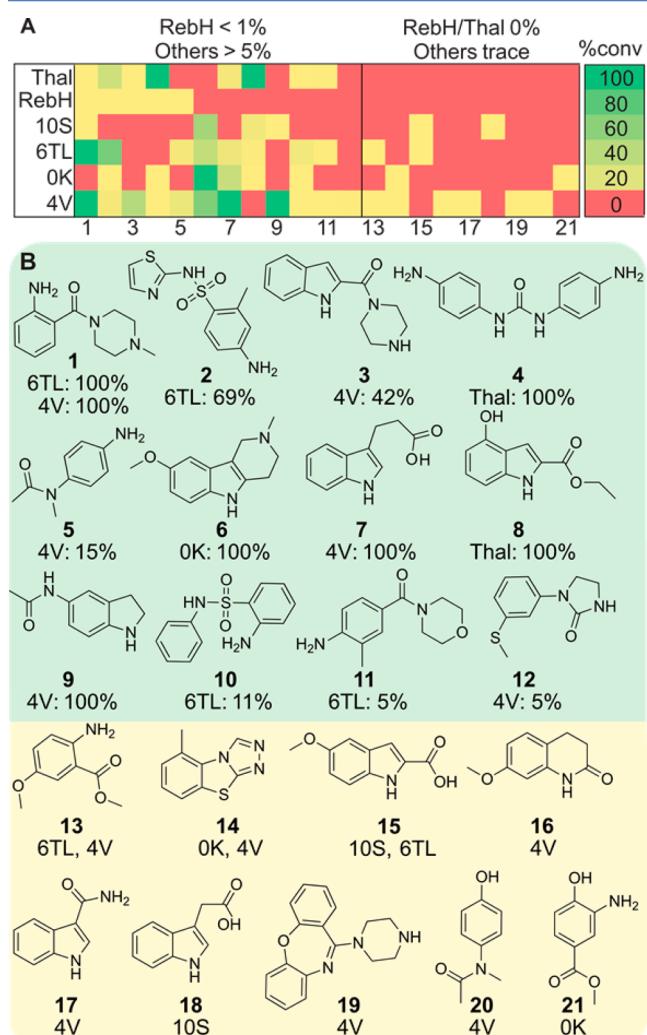


Figure 3. (A) Conversion data from initial activity profile for selected substrates. (B) Structures depicted in heat map. Compounds for which >5% conversion was observed are highlighted in green; those halogenated to a trace extent are highlighted in yellow.

conversion (>5%) was observed for Thal or the engineered variants on substrates where trace or no activity was observed for RebH (Figure 3B, 1–12). A number of cases were also observed in which neither RebH nor Thal displayed any conversion on a substrate, but trace activity was seen for the engineered halogenases (Figure 3B, 13–21). This trace activity, confirmed by LCMS using chlorine isotope patterns, provides a starting point for evolving enzymes with good activity on these different compound classes.¹⁵ We have previously demonstrated the feasibility of this with FDHs.^{10,15} Moreover, this finding clearly shows that even minor changes in halogenase sequence can significantly affect specificity and scope, which highlights the importance of evaluating multiple enzymes on a given substrate(s) to find active catalysts, be it for preparative reactions or for starting points for evolution.^{26,28,29}

Comparison of Trp-FDH and Fungal FDH Activity. While many panel 1 substrates were accepted as substrates by the Trp-FDHs, very few (7%) were halogenated by the fungal FDHs (Figure 2). For this reason, the activity of these enzymes toward panel 2 substrates was evaluated. Many of these less nucleophilic substrates³⁰ were not halogenated to a significant extent by the Trp-FDHs; however, all were halogenated by at least one of the fungal FDHs (Figure 4). Rdc2 and Gsfl, in

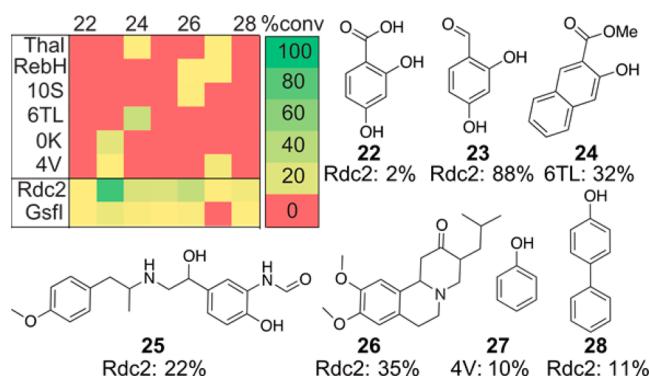


Figure 4. Conversion data from initial activity profile and structures for substrates 22–28.

contrast to Trp-FDHs, have a preference for the less electron rich substrates,³⁰ which indicates that the selectivity is more complex than a simple electronic activation. The practical consequence is that expanded substrate scope can be achieved by investigating diverse enzyme sets.

Selectivity of FDH Halogenation. The conversion data for substrates from panels 1 and 2 provide valuable insights into the scope and specificity of the eight FDHs in this study. To

determine whether site-selective halogenation occurred in reactions of **non-native substrates**, however, RebH- and Rdc2-catalyzed reactions that provided >5% conversion under high-throughput screening conditions (36 total) were conducted on a 1–10 mg scale. Products for 29 of these reactions were obtained from preparative LCMS. Conversion comparable to the initial substrate panel evaluation was generally observed (**Figure 2**). Characterization of these products by NMR spectroscopy and HRMS established that most of these substrates (22) were halogenated at a single position with >95% selectivity (**Figure 5**, representative examples; **Table S2** in the Supporting Information, all products).

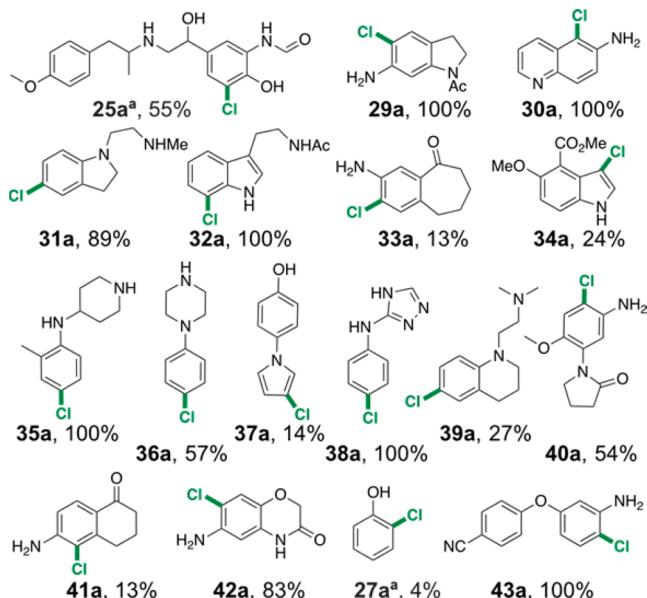


Figure 5. Conversions for substrates with multiple electronically activated sites that provided a single halogenated product. Reactions were conducted on 1–10 mg scale using either RebH (**29a–43a**) or Rdc2^a (**25a, 27a**). See **Table S2** in the Supporting Information for a full product list. Conversion and selectivity for **27a** were determined by comparison with authentic material.

Quantitative Evaluation of Substrate Activation. As previously noted, FDH-catalyzed halogenation is believed to proceed via electrophilic aromatic substitution.⁶ Substrates bound within the FDH active site are presumably bound such that a single site is located proximal to an electrophilic halogen species (a Lys79-chloramine and hypohalous acid have been proposed)^{8,9} that formally transfers a halonium ion (X^+) to the substrate. This enables Trp-FDHs to halogenate the benzo ring of tryptophan over the more reactive pyrrolo ring. The selectivity of the reactions whose products are shown in **Figure 5** are similarly remarkable, given that more than one site on each of the substrates might reasonably be considered sufficiently activated for electrophilic aromatic substitution.³⁰ Importantly, however, this level of catalyst control was achieved on diverse substrates without the need for multiple halogenase engineering efforts.¹⁰

We previously used calculated **halonium affinity (HalA)**³¹ values to quantitate the extent to which the observed selectivity of RebH variants overrides the electronic preference of tryptamine toward (hypothetical) halogenation by Cl^+ at different positions.¹⁰ This approach is analogous to several previous examples in which calculated electronic parameters

were used to rationalize the regioselectivity of electrophilic aromatic substitution reactions involving small-molecule reagents.^{30,32–34} HalA values for each sp^2 -hybridized carbon on each substrate for which halogenation selectivity was unambiguously assigned were calculated (29 substrates, **Table S2** in the Supporting Information), and a plot of conversion versus HalA at each site was constructed (**Figure 6**).

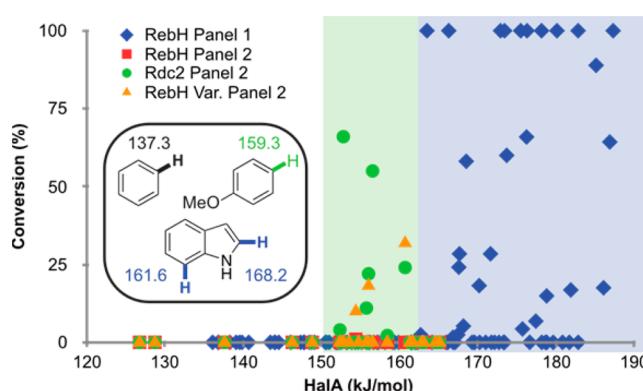


Figure 6. Conversion versus halonium affinity (HalA) for each sp^2 carbon on each substrate in **Table S2** in the Supporting Information. HalA ranges for panel 1 substrates halogenated by RebH and panel 2 substrates halogenated by Rdc2 are highlighted in blue and green, respectively. The inset shows representative HalA values for common arenes (not substrates).

Several useful trends can be obtained from this plot. Good conversion of panel 1 substrates by RebH (**Figure 6**, blue diamonds) occurs only at sites with HalA values >160 kcal/mol. Poor conversion of panel 2 substrates, which all possess HalA values <164 kcal/mol, was observed for RebH (**Figure 6**, red squares), but Rdc2 provides significant conversion of panel 2 substrates at sites with HalA values as low as 152 kcal/mol (**Figure 6**, green circles). Similarly, engineered RebH variants were able to halogenate sites on panel 2 substrates with HalA values as low as 154 kcal/mol (**Figure 6**, orange triangles). Sites on substrates with HalA values above the minimum for each enzyme, however, were often not halogenated; therefore, other factors clearly affect FDH selectivity. In short, our initial data suggest that there is a minimum electronic activation (HalA \approx 154 kJ/mol) necessary for halogenation by the FDHs examined, and this can serve as a coarse filter for identifying potential new substrates for these enzymes. To better understand the selectivity of the reactions, substrate binding must be considered (vide infra).

Substrate Docking. Deviation of FDH selectivity from outcomes expected on the basis of electronic effects alone likely results in part from the unique binding of different substrates within FDH active sites.^{7,10,16} Sixteen substrates for which the selectivity of RebH was established (**Figure 5**) were therefore docked into a multiconformer model of RebH to determine if computationally inexpensive methods^{35,36} could improve on a purely electronic view of RebH selectivity. Docking was performed using ICM Molsoft docking³⁷ and ROCS pharmacophore overlay.³⁸ The former employs a Monte Carlo minimization algorithm combined with empirical energy terms to model van der Waals forces, hydrophobic properties, and electrostatics.³⁹ This procedure returns multiple energetically favorable poses of a ligand within a target, but it is blind to known binding interactions. Given the existence of several

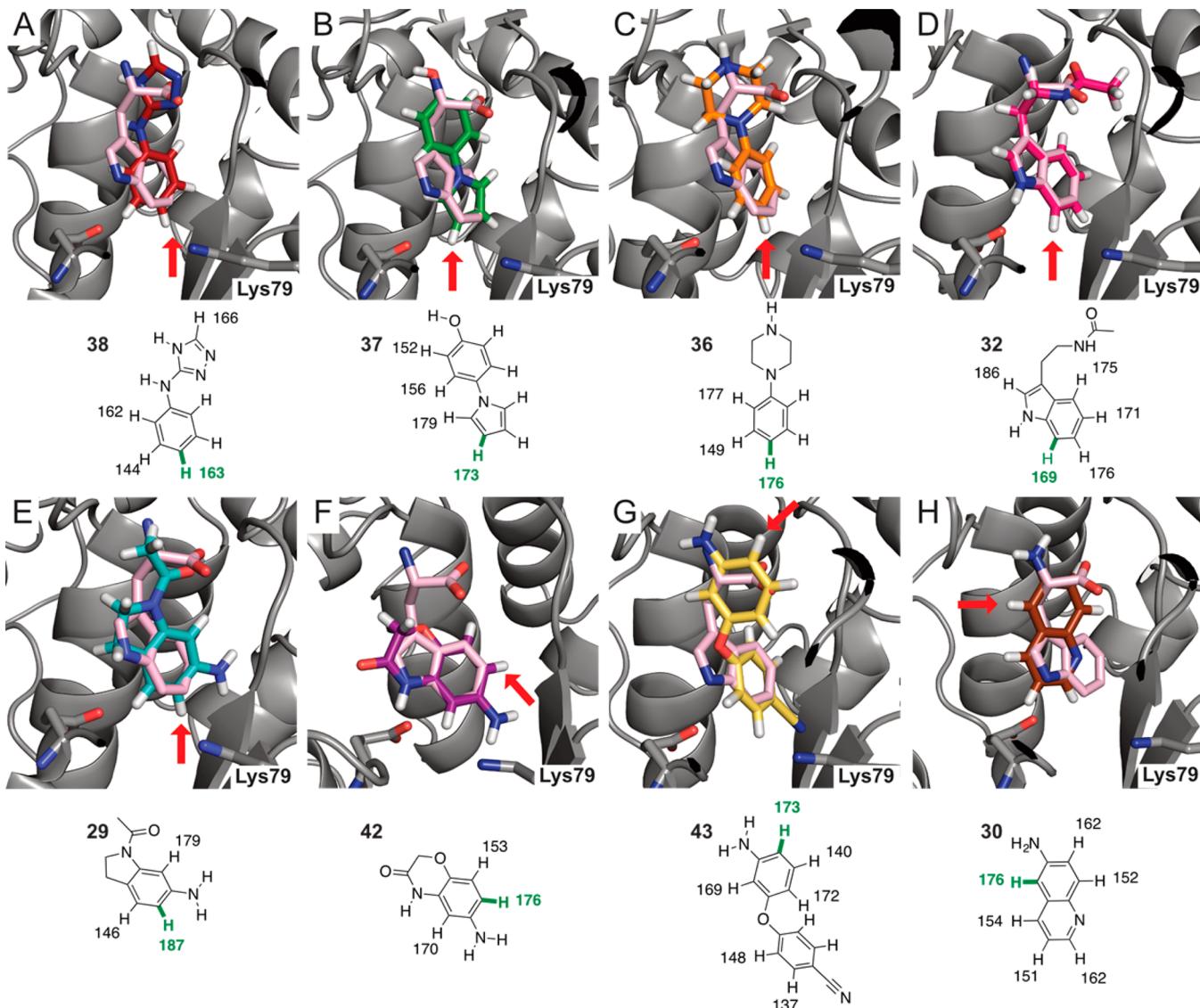


Figure 7. Poses observed for ROCS pharmacophore overlay. Tryptophan in the RebH–tryptophan complex (PDB 2OA1) is shown in pink. Lys79 is labeled in each model, and the site of halogenation for each substrate is indicated with a red arrow. Structures for each substrate are shown with HalA values for each site and the site halogenated in green.

crystal structures of RebH–tryptophan complexes, the ROCS pharmacophore overlay approach, which aligns query molecules to a given template on the basis of molecular shape and charge³⁸ was also used. As in our previous work, catalytically relevant poses were taken to be those in which an sp^2 -hybridized carbon was positioned proximal to Lys79 in the RebH active site.¹⁰

In general, neither method provided significant predictive ability regarding substrate specificity (i.e., whether or not a given substrate would be halogenated). Many unreactive substrates were nonetheless predicted to bind in the RebH active site. In addition, while ICM Molsoft docking returned several poses for each substrate within the RebH active site, very few of these appeared to be catalytically relevant. This was likely due to both the relatively large size of the binding pocket and the relatively small size of the compounds, allowing for multiple high-scoring (energetically feasible) docking poses.³⁶ ROCS pharmacophore overlay, on the other hand, provided a number of interesting results regarding the site selectivity of substrates that were halogenated (Figure 7, representative

examples). Perhaps most notably, Figure 7A–D shows cases where the site halogenated does not have the highest HalA but is predicted to bind proximal to Lys79. This selectivity is inconsistent with purely electronic effects (substrate control) but is consistent with predicted binding within the active site (catalyst control), suggesting that FDHs could be useful for obtaining novel selectivity on diverse structures. Moreover, the pharmacophore overlay predicts that this selectivity results from different polar functional groups distal to the site of halogenation binding to the pocket occupied by the substrate amino acid moiety in the crystallized RebH–tryptophan complex. Similar anchoring effects have been exploited for other enzymes,⁴⁰ but these results suggest that a range of different functional groups and even heterocycles can fulfill this role in RebH.

Several substrates, including the one shown in panel Figure 7E, were predicted to bind such that the most electronically activated site was proximal to Lys79, but cases in which electronic effects appear to outweigh any predicted binding effects were also found. In Figure 7F–H for example, substrates

were halogenated at the most electronically activated site even when this site was predicted to bind distal to Lys79. Presumably, in these cases, alternate substrate orientations within the RebH active site can be sampled, leading to halogenation at the most activated site(s). While ICM Molsoft docking tended to underrepresent binding modes similar to native Trp binding, pharmacophore overlay emphasized such binding modes, both of which reduce predictive utility. In combination with HalA values, however, computationally inexpensive docking simulations appear to provide insight into the observed selectivity of FDH-catalyzed halogenation on non-native substrates in many cases.

Controlling Site Selectivity Using Different FDHs. The catalyst control exhibited by RebH toward various substrates also suggests that different FDHs could be used to produce different halogenated products. To explore this possibility, halogenation of several substrates with multiple electronically activated sites that were selectively halogenated by RebH or Rdc2 was also examined using 4V, 6TL, and 10S (Table 1, entries 1–6). Two reactions in which RebH provided mixtures of compounds were also examined to determine if other enzymes could improve selectivity (Table 1, entries 7 and 8). Finally, reactions of two phenols were also examined (Table 1,

entries 9 and 10). To compare the product profiles of different halogenases with those produced by a common stoichiometric chlorinating reagent, these reactions were also conducted with N-chlorosuccinimide (NCS).⁴¹

Substrate selectivity profiles for the different halogenating species are presented in Table 1, with results for enzymes that provide similar profiles removed for clarity. Altered site selectivity was observed not only between FDHs but also between the FDHs and NCS. A complete switch in site selectivity is observed using different FDHs for substrates such as 27, 41, and 46. A remarkable example of this is phenol (27), for which Rdc2 halogenates exclusively ortho to the alcohol, while 4V only halogenates para. This particular example highlights how effective enzymatic catalyst control can be even on simple non-native substrates. For many substrates within this subset, an FDH gave a single isomer when mixtures were seen with NCS (Table 1; entries 2–6 and 10). In addition, NCS often begins to dihalogenate substrates to significant amounts before complete consumption of starting material. This problem is circumvented by using FDHs, which typically have decreased activity on halogenated products (Table 1; entries 2–4 and 7). FDHs are also catalytic and use chloride as a halide source and air as a terminal oxidant.²

For certain substrates, such as 46 (Table 1; entry 8), different ratios of two monochlorinated products can be observed for NCS, RebH and 4V. The use of FDHs presents the unique opportunity in such cases to tune existing selectivity through directed evolution.¹⁰ For example, RebH halogenates primarily para to the methyl group but is not highly selective for this site (70%). RebH could be tuned to give better selectivity at this position. On the other hand, 4V halogenates primarily ortho to the methyl group (77%) and could be used as a starting point to selectively halogenate this position if desired.

DISCUSSION

In this work, the substrate scope of several FDHs was probed by measuring their activity toward diverse panels of low-molecular-weight aromatic compounds. The resulting substrate activity profiles contain information that cannot be obtained by studying enzyme activity on native substrates or simple derivatives of these substrates.^{18,19,23} While substitution of compounds distal to sites of reaction is commonly used to probe electronic effects in organic reactions (e.g., LFERs),⁴² this can perturb substrate binding as well as electronic effects within an enzyme active site,⁴³ leading to spurious correlations between substitution and activity. The points in Figure 6 corresponding to substrate sites with sufficient electronic activation (as indicated by calculated HalA values) but for which halogenation was not observed likely exemplify this phenomenon. On the other hand, sampling a sufficient range of substrate classes, substituents, and substitution patterns provides general information regarding the halogenating ability of different enzymes. Analogous efforts on other enzymes could prove similarly enlightening.

The small set of FDHs examined provided remarkably broad substrate scope considering the substantial variation of the substrates evaluated relative to the native substrates of these enzymes. A wide range of functional groups were tolerated, but it is worth noting that this would not necessarily be apparent from evaluating individual substrates. For example, the simple ethylsulfonamide-substituted aniline 51 (Table S1 in the Supporting Information) was not halogenated, but sulfonamides 2 and 10 were both halogenated, showing that minor

Table 1. Relative Product Distributions for Halogenation of Representative Substrates using FDHs and NCS^a

| Entry | Substrate | Species | %1 | %2 | %3 | % di | % conv. |
|-------|-----------|---------|-----|-----|----|------|---------|
| 1 | | NCS | 100 | 0 | - | - | 97 |
| | | RebH | 0 | 100 | - | - | 55 |
| 2 | | NCS | 18 | 41 | 29 | 13 | 87 |
| | | RebH | 0 | 100 | 0 | 0 | 32 |
| | | 6TL | 17 | 83 | 0 | 0 | 51 |
| 3 | | NCS | 7 | 65 | - | 28 | 86 |
| | | RebH | 0 | 100 | - | 0 | 35 |
| 4 | | NCS | 41 | 20 | 30 | 9 | 80 |
| | | RebH | 2 | 0 | 98 | 0 | 94 |
| 5 | | NCS | 63 | 37 | - | - | 94 |
| | | 6TL | 3 | 97 | - | - | 100 |
| 6 | | NCS | 10 | 90 | - | - | 98 |
| | | RebH | 0 | 100 | - | - | 15 |
| | | 6TL | 93 | 7 | - | - | 81 |
| | | 4V | 70 | 30 | - | - | 22 |
| 7 | | NCS | 2 | 25 | - | 73 | 65 |
| | | RebH | 50 | 0 | - | 50 | 50 |
| | | 6TL | 58 | 42 | - | 0 | 44 |
| 8 | | NCS | 52 | 43 | - | 5 | 94 |
| | | RebH | 30 | 70 | - | 0 | 74 |
| | | 4V | 77 | 23 | - | 0 | 81 |
| 9 | | NCS | 37 | 52 | - | 11 | 62 |
| | | Rdc2 | 75 | 18 | - | 7 | 44 |
| | | Gsfl | 91 | 9 | - | 0 | 2 |
| 10 | | NCS | 23 | 75 | - | 2 | 94 |
| | | Rdc2 | 100 | 0 | - | - | 4 |
| | | 4V | 0 | 100 | - | - | 10 |

^aSites of halogenation (P1–3) are indicated for select substrates.

changes in substituents distal to reaction sites can significantly affect substrate reactivity. Similarly, compounds 21 (Figure 3) and 44 (Table 1, entry 5) are regioisomeric at methyl (aminohydroxybenzoates), but while 44 is halogenated in high conversion by multiple FDHs (100% with 4V, 56% with 6TL, 55% with RebH), 21 was halogenated only by OK to a trace amount. Examining activity on diverse substrates²³ can account for these differences and provide an accurate picture of the types of functional groups and fragments that can potentially be halogenated. The molecular context in which these moieties appear, however, will clearly influence their effect on FDH catalysis.

Subsequent analysis of substrates that undergo halogenation can then be used to rationalize how substrate properties affect FDH selectivity. Each sp^2 carbon on each substrate constitutes a potential reaction site; therefore, conversion values for each site can be plotted against steric or electronic descriptors. The use of calculated HalA values³¹ in this work indicated clear cutoffs in electronic activation that are required for substrate halogenation by FDHs, but electronics alone could not explain the observed selectivity in most cases. Docking simulations provided further insight into selectivity in combination with HalA values, and more rigorous computational methods (e.g., MD simulations)⁴⁴ could significantly improve the predictive models for FDH activity and selectivity.

These results constitute the largest examination of FDH substrate scope and selectivity to date and significantly improve our understanding of the types of substrates that can be halogenated using these enzymes. The range of substrates that could be halogenated clearly attests to the potential for halogenases evolved for one application to find utility toward another. Moreover, this broad substrate scope should facilitate the identification of novel FDHs, since native substrates are not necessarily required to confirm FDH activity.¹⁸ Given the unique scope and selectivity of the enzymes examined in this study relative to one another and to NCS, the use of probe substrates for functional characterization of putative FDHs holds great promise for expanding the utility of enzymatic halogenation.¹

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acscatal.6b02707](https://doi.org/10.1021/acscatal.6b02707).

Full description of materials, a complete substrate list, experimental details, and characterization of all compounds and enzymes (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail for K.L.T.: kian.tan@novartis.com.

*E-mail for J.C.L.: jaredlewis@uchicago.edu.

ORCID

Jared C. Lewis: [0000-0003-2800-8330](https://orcid.org/0000-0003-2800-8330)

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the NIH (1R01GM115665), The Novartis Institutes for Biomedical Research, and the NSF under the CCI Center for Selective C–H Functionalization (CHE-

1205646). M.C.A. was supported by an NIH Chemistry and Biology Interface training grant (T32 GM008720), an NSF predoctoral fellowship (DGE-1144082), a University of Chicago Department of Chemistry Helen Sellei-Beretvas Fellowship, and an ARCS Scholar Award. This work was completed in part with resources provided by the University of Chicago Research Computing Center with assistance from Dr. Jonathan Skone. We thank Prof. Karl-Heinz van Pée for providing a plasmid encoding Thal, Clayton Springer for building models of halonium adducts for all substrates studied, and Laurel Heckman and Randy Richter for housing M.C.A.

ABBREVIATIONS

FDH, flavin-dependent halogenase; Trp-FDH, tryptophan-flavin-dependent halogenase; EAS, electrophilic aromatic substitution; HalA, halonium affinity; NCS, *N*-chlorosuccinamide; LFER, linear free energy relationship

REFERENCES

- (1) Weichold, V.; Milbrecht, D.; van Pée, K.-H. *Angew. Chem., Int. Ed.* **2016**, *55* (22), 6374–6389.
- (2) Vaillancourt, F. H.; Yeh, E.; Vosburg, D. A.; Garneau-Tsodikova, S.; Walsh, C. T. *Chem. Rev.* **2006**, *106* (8), 3364–3378.
- (3) Payne, J. T.; Andorfer, M. C.; Lewis, J. C. *Engineering Flavin-Dependent Halogenases*, 1st ed.; Elsevier: Amsterdam, 2016; Methods in Enzymology **575**, pp 93–126.
- (4) van Pée, K.-H.; Patallo, E. *Appl. Microbiol. Biotechnol.* **2006**, *70* (6), 631–641.
- (5) Yeh, E.; Garneau, S.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102* (11), 3960–3965.
- (6) Dong, C.; Flecks, S.; Unversucht, S.; Haupt, C.; van Pée, K.-H.; Naismith, J. H. *Science* **2005**, *309* (5744), 2216–2219.
- (7) Zhu, X.; Laurentis, W. D.; Leang, K.; Herrmann, J.; Ihlefeld, K.; van Pée, K.-H.; Naismith, J. H. *J. Mol. Biol.* **2009**, *391* (1), 74–85.
- (8) Yeh, E.; Blasiak, L. C.; Koglin, A.; Drennan, C. L.; Walsh, C. T. *Biochemistry* **2007**, *46* (5), 1284–1292.
- (9) Flecks, S.; Patallo, E. P.; Zhu, X.; Ernyei, A. J.; Seifert, G.; Schneider, A.; Dong, C.; Naismith, J. H.; van Pée, K.-H. *Angew. Chem., Int. Ed.* **2008**, *47* (49), 9533–9536.
- (10) Andorfer, M. C.; Park, H.-J.; Vergara-Coll, J.; Lewis, J. C. *Chem. Sci.* **2016**, *7* (6), 3720–3729.
- (11) Zeng, J.; Zhan, J. *ChemBioChem* **2010**, *11* (15), 2119–2123.
- (12) Zeng, J.; Lytle, A. K.; Gage, D.; Johnson, S. J.; Zhan, J. *Bioorg. Med. Chem. Lett.* **2013**, *23* (4), 1001–1003.
- (13) Chooi, Y.-H.; Cacho, R.; Tang, Y. *Chem. Biol.* **2010**, *17* (5), 483–494.
- (14) Durak, L. J.; Payne, J. T.; Lewis, J. C. *ACS Catal.* **2016**, *6* (3), 1451–1454.
- (15) Payne, J. T.; Poor, C. B.; Lewis, J. C. *Angew. Chem., Int. Ed.* **2015**, *54* (14), 4226–4230.
- (16) Shepherd, S. A.; Karthikeyan, C.; Latham, J.; Struck, A.-W.; Thompson, M. L.; Menon, B. R. K.; Styles, M. Q.; Levy, C.; Leys, D.; Micklefield, J. *Chem. Sci.* **2015**, *6* (6), 3454–3460.
- (17) Latham, J.; Henry, J.-M.; Sharif, H. H.; Menon, B. R. K.; Shepherd, S. A.; Greaney, M. F.; Micklefield, J. *Nat. Commun.* **2016**, *7*, 11873.
- (18) Landwehr, M.; Carbone, M.; Otey, C. R.; Li, Y.; Arnold, F. H. *Chem. Biol.* **2007**, *14* (3), 269–278.
- (19) Zhang, K.; El Damaty, S.; Fasan, R. *J. Am. Chem. Soc.* **2011**, *133* (10), 3242–3245.
- (20) Payne, J. T.; Andorfer, M. C.; Lewis, J. C. *Angew. Chem., Int. Ed.* **2013**, *52* (20), 5271–5274.
- (21) Seibold, C.; Schnerr, H.; Rumpf, J.; Kunzendorf, A.; Hatscher, C.; Wage, T.; Ernyei, A. J.; Dong, C.; Naismith, J. H.; van Pée, K.-H. *BioCATAL. Biotransform.* **2006**, *24* (6), 401–408.
- (22) Frese, M.; Sewald, N. *Angew. Chem., Int. Ed.* **2015**, *54* (1), 298–301.

- (23) Carr, R.; Alexeeva, M.; Enright, A.; Eve, T. S. C.; Dawson, M. J.; Turner, N. J. *Angew. Chem., Int. Ed.* **2003**, *42* (39), 4807–4810.
- (24) Erlanson, D. A.; McDowell, R. S.; O'Brien, T. *J. Med. Chem.* **2004**, *47* (14), 3463–3482.
- (25) Heemstra, J. R.; Walsh, C. T. *J. Am. Chem. Soc.* **2008**, *130* (43), 14024–14025.
- (26) Lewis, J. C.; Arnold, F. H. *Chimia* **2009**, *63* (6), 309–312.
- (27) Matsumura, I.; Ellington, A. D. *J. Mol. Biol.* **2001**, *305* (2), 331–339.
- (28) Rentmeister, A.; Arnold, F. H.; Fasan, R. *Nat. Chem. Biol.* **2009**, *5* (1), 26–28.
- (29) Truppo, M. D.; Escalettes, F.; Turner, N. J. *Angew. Chem., Int. Ed.* **2008**, *47* (14), 2639–2641.
- (30) Pratihar, S.; Roy, S. *J. Org. Chem.* **2010**, *75* (15), 4957–4963.
- (31) Ashtekar, K. D.; Marzijarani, N. S.; Jagannathan, A.; Holmes, D.; Jackson, J. E.; Borhan, B. *J. Am. Chem. Soc.* **2014**, *136* (38), 13355–13362.
- (32) Tishchenko, O.; Pham-Tran, N.-N.; Kryachko, E. S.; Nguyen, M. T. *J. Phys. Chem. A* **2001**, *105* (38), 8709–8717.
- (33) Schwabe, T.; Grimme, S. *Eur. J. Org. Chem.* **2008**, *2008* (35), 5928–5935.
- (34) Aoki, S.; Matsuo, Y.; Ogura, S.; Ohwada, H.; Hisamatsu, Y.; Moromizato, S.; Shiro, M.; Kitamura, M. *Inorg. Chem.* **2011**, *50* (3), 806–818.
- (35) Stiefl, N.; Gedeck, P.; Chin, D.; Hunt, P.; Lindvall, M.; Spiegel, K.; Springer, C.; Biller, S.; Buenemann, C.; Kanazawa, T.; Kato, M.; Lewis, R.; Martin, E.; Polyakov, V.; Tommasi, R.; van Drie, J.; Vash, B.; Whitehead, L.; Xu, Y.; Abagyan, R.; Raush, E.; Totrov, M. *J. Chem. Inf. Model.* **2015**, *55* (4), 896–908.
- (36) Cross, J. B.; Thompson, D. C.; Rai, B. K.; Baber, J. C.; Fan, K. Y.; Hu, Y.; Humblet, C. *J. Chem. Inf. Model.* **2009**, *49* (6), 1455–1474.
- (37) Abagyan, R.; Totrov, M.; Kuznetsov, D. *J. Comput. Chem.* **1994**, *15* (5), 488–506.
- (38) Grant, J. A.; Gallardo, M. A.; Pickup, B. T. *J. Comput. Chem.* **1996**, *17* (14), 1653–1666.
- (39) Schapira, M.; Totrov, M.; Abagyan, R. *J. Mol. Recognit.* **1999**, *12* (3), 177–190.
- (40) Negretti, S.; Narayan, A. R. H.; Chiou, K. C.; Kells, P. M.; Stachowski, J. L.; Hansen, D. A.; Podust, L. M.; Montgomery, J.; Sherman, D. H. *J. Am. Chem. Soc.* **2014**, *136* (13), 4901–4904.
- (41) Ponce, M. A.; Tarzi, O. I.; Erra-Balsells, R. *J. Heterocycl. Chem.* **2003**, *40* (3), 419–426.
- (42) Anslyn, E. V.; Dougherty, D. A. *Modern Physical Organic Chemistry*; University Science Books: Mill Valley, CA, 2006; pp 441–471.
- (43) Weise, N. J.; Parmeggiani, F.; Ahmed, S. T.; Turner, N. J. *J. Am. Chem. Soc.* **2015**, *137* (40), 12977–12983.
- (44) Narayan, A. R. H.; Jiménez-Osés, G.; Liu, P.; Negretti, S.; Zhao, W.; Gilbert, M. M.; Ramabhadran, R. O.; Yang, Y.-F.; Furan, L. R.; Li, Z.; Podust, L. M.; Montgomery, J.; Houk, K. N.; Sherman, D. H. *Nat. Chem.* **2015**, *7* (8), 653–660.