

Data-Driven Construction of an Imine Reductase Library Capable of Broad-Scope Reductive Amination at Equimolar Substrate Concentrations

Sarah A. Berger^{1,5}, Christopher Grimm^{1,5}, Marco Cespuigi², Andreas Krassnigg², Tobias Schopper², Irene Marzuoli³, Isabel Oroz-Guinea^{1,4}, Stephan Vrabl¹, Lukas Roemer¹, Melanie A. Weber¹, Fabian M. Kulier¹, Yuliya Orel¹, Bettina M. Nestl², Christian C. Gruber², Georg Steinkellner², Serena Bisagni³, Francis Gosselin⁵, Hans Iding³, Kurt Puentener³, Dennis Wetzl³, Wolfgang Kroutil^{1,4,6}, Joerg H. Schrittwieser^{1,*}

¹ Department of Chemistry, University of Graz, NAWI Graz, Heinrichstrasse 28, 8010 Graz (Austria)

² Innophore GmbH, Am Eisernen Tor 3, 8010 Graz (Austria)

³ F. Hoffmann-La Roche Ltd., Process Chemistry and Catalysis, Grenzacherstrasse 124, 4070 Basel (Switzerland)

⁴ Field of Excellence "BioHealth", University of Graz, 8010 Graz (Austria)

⁵ Genentech, Inc., Department of Small Molecule Process Chemistry, 1 DNA Way, South San Francisco, California 94080 (United States)

⁶ BioTechMed Graz, 8010 Graz (Austria)

⁵ These authors contributed equally.

E-Mail: joerg.schrittwieser@uni-graz.at

Abstract: Reductive amination catalysed by imine reductase (IREd) and reductive aminase (RedAm) enzymes has recently been established as a powerful method for the asymmetric synthesis of chiral amines. While this biocatalytic technology has rapidly progressed from proof of concept to initial industrial applications, its scope and limitations remain to be fully explored. In this work, we report a broad and systematic profiling of reductive amination performance in the sequence space of IREds and RedAms. This investigation employed an iterative strategy for activity and stereoselectivity screening, guided by chemo- and bioinformatic modelling as well as machine learning. By evaluating the catalytic performance of 175 IREds against structurally diverse panels of 36 carbonyl compounds and 24 amines, we show that the majority of these enzymes is capable of asymmetric reductive amination at equimolar concentrations of the two substrates (50 mM each). The most effective enzymes identified in this study display sequence characteristics of RedAms, are active on 29–42% of the analysed substrate combinations, and combine high specific activities for the most favourable substrate pair (1.7–27.7 U/mg_{IREd}) with excellent stereoselectivity. Beyond assembling this high-performance enzyme panel, we demonstrate extrapolation from our collected screening data to new substrate combinations by deep learning and the scale-up of selected reactions to a preparative batch size (10 mmol substrate, 200 mL reaction volume), delivering gram amounts of reductive amination products in high yield (63–89%) and optical purity (98% to >99% *ee*).

Introduction

Chiral amines are a wide-spread structural motif in pharmaceuticals, agrochemicals, and natural products.¹ Nearly 80% of the 200 top-selling small-molecule drugs in 2023 contain at least one nitrogen atom, with 35% containing a non-racemic α -chiral amine or amino acid moiety.^{2,3} The most versatile method for synthesising chiral amines is asymmetric reductive amination, which proceeds through the condensation of a carbonyl compound and a primary or secondary amine to give an imine or iminium intermediate, followed by asymmetric reduction to the target chiral amine. Due to the two variable residues on each substrate, reductive amination offers a broad scope, theoretically allowing the convergent synthesis of almost any α -chiral amine (*cf.* Figure 1a).

Recent years have witnessed impressive advances in transition metal-catalysed, organocatalytic, and biocatalytic asymmetric reductive amination.^{4–9} Metal-catalysed approaches typically use chiral iridium(I), rhodium(I), or ruthenium(II) complexes at elevated hydrogen pressure (>10 bar) and have demonstrated broad ketone and amine scope. Organocatalytic methods predominantly rely on axially chiral phosphoric acids as catalysts, Hantzsch esters, silanes or boranes as reducing agents, and aniline derivatives as amine substrates. Enzymatic reductive amination has long been limited to the preparation of α -amino acids from α -keto acids using amino acid dehydrogenases and the synthesis of α -chiral primary amines from ketones using transaminases (the latter being a reversible group transfer process).^{9–16} Recently, however, the scope of biocatalytic reductive amination has been significantly expanded with the discov-

ery that imine reductases (IREds) – and in particular a sub-class of this enzyme family termed reductive aminases (RedAms) – are able to perform this reaction with a broad range of amine and carbonyl substrates using NADPH as the reducing agent.^{17–22}

Following the initial discovery of asymmetric reductive amination by IREds and RedAms,^{23–25} sequence-based and metagenomic enzyme discovery efforts have rapidly generated extensive libraries of these biocatalysts,^{26–30} and protein engineering has been used to tailor them to specific target reactions.^{31–41} Within a decade, IREd-catalysed reductive amination has progressed from proof of concept to industrial application,¹⁹ including the recent syntheses of a lysine-specific demethylase-1 (LSD1) inhibitor,³¹ a Janus kinase 1 (JAK1) inhibitor,³² and a cyclin-dependent kinase inhibitor on a kilogram-ton scale.^{36,37} Despite these advances, the scope and limitations of asymmetric reductive amination catalysed by IREds and RedAms need to be further explored. Many screening campaigns have focused on a narrow selection of model substrates, particularly in the amine dimension. Additionally, many studies use limited concentrations of the carbonyl compound and a large excess of the amine substrate (commonly 5–10 mM and 10–50 eq., respectively),^{25,26,28,30,34} which helps to achieve faster biotransformations but negatively impacts the resource efficiency and preparative applicability of these reactions.

Screening multiple-substrate catalytic reactions for diversity is intrinsically challenging, as multiplicative growth of the parameter space to be covered results in a large experimental workload, a problem known in data science as the “curse of dimensionality”.⁴² Literature-reported screenings of IREd-catalysed reductive amina-

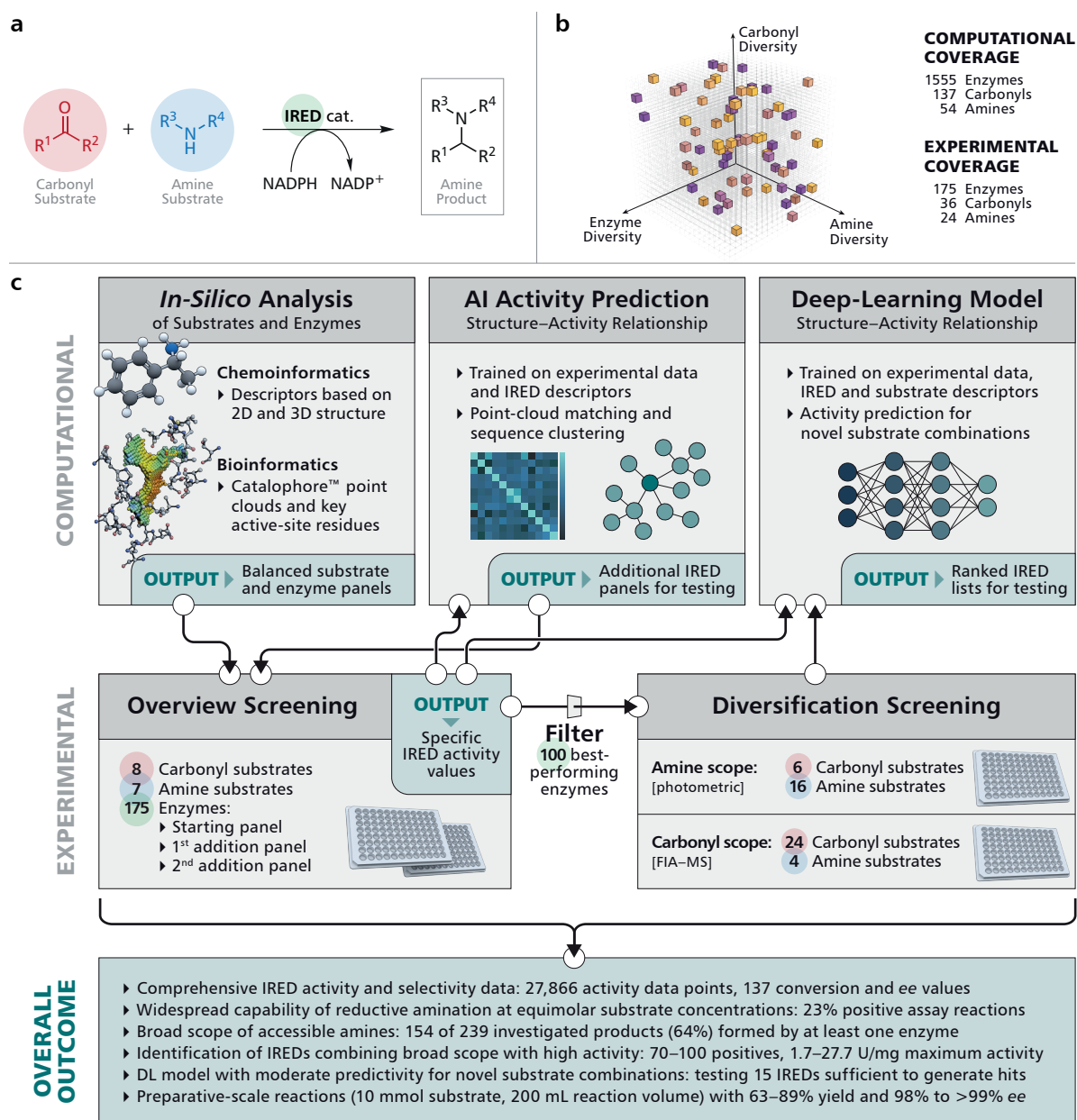


Figure 1 | Visual outline of the present study. **a**, General reaction scheme of IRED-catalysed reductive amination. Unless residues R¹ and R² are identical, the formed amine product is chiral. **b**, Schematic visualisation of the three-dimensional screening space. Representative coverage by non-exhaustive screening depends on the balanced selection of enzymes and substrates. **c**, Summary of computational and experimental workflows and overall outcome of the study. Substrate and enzyme panels were selected based on chemo- and bioinformatic analyses. Experimental screening was performed in two phases (Overview Screening and Diversification Screening) and the resulting data were used to train AI models with the aim to establish structure–activity relationships in the enzyme and substrate dimensions. The models were used to suggest additional IREDs for experimental testing, either on known substrate combinations (AI Activity Prediction) or on substrate combinations outside of the training data set (Deep-Learning Model). Arrows represent the flow of information between the different parts of the study.

tion have included over 150 carbonyl compounds and more than 50 amines, and the number of currently identified IRED sequences is in the thousands (the *Imine Reductase Engineering Database*, maintained by Pleiss and co-workers at the University of Stuttgart and available online at <https://ired.biocatnet.de/>, currently lists 1,409 sequences). Covering this search space exhaustively would hence require the analysis of more than 10 million individual substrate–enzyme combinations, a number and chemical diversity that appears overwhelming even for modern high-throughput experimentation.^{43–46} Non-exhaustive screening, however, depends on the selection of representative substrate and enzyme panels and still requires assay methods that offer high throughput and broad scope.

Even if these prerequisites are met, non-exhaustive screening leaves a significant part of the search space unexplored, which poses challenges for systematic data analysis and predictive modelling.⁴⁷ These fundamental problems have been addressed in various chemical contexts and in diverse ways, ranging from combinatorial screening strategies that increase experimental throughput,^{48–52} through predictive modelling in the substrate and catalyst dimensions,^{53–56} to fully autonomous robotic experimentation (“self-driving labs”).^{57–61}

In this work, we present a general strategy for the iterative, multi-dimensional screening of catalytic reactions, which we have used to broadly explore the IRED sequence space for reductive amination activity at equimolar substrate concentrations. Our approach

involves swift elimination of underperforming enzymes early in the screening campaign, with an emphasis on activity and substrate scope over stereoselectivity. Additionally, we leverage machine learning for performance prediction, and employ analysis methods of moderately high throughput that should be accessible to many catalysis laboratories. The screening data collected in this study provide a representative fingerprint of reductive amination performance in the IRED/RedAm enzyme family and have pinpointed IREDs that catalyse synthetically interesting reductive amination reactions with high activity and stereoselectivity.

Results and Discussion

Screening Strategy

Our screening strategy is iterative in all three dimensions (carbonyl, amine, and IRED; cf. Figure 1b) and comprises two phases: First, we rapidly evaluated the reductive amination performance of candidate enzymes using a limited substrate panel that is covered exhaustively (Overview Screening, 8 carbonyls \times 7 amines; Figure 1c). We eliminated poorly performing IREDs from the screening panel and replaced them with new candidates identified using a machine-learning model trained on the collected activity data. Next, we explored the substrate scope of the best-performing IREDs more broadly using a split screening that covers a diverse range of substrates non-exhaustively (Diversification Screening, 24 carbonyls \times 4 amines and 6 carbonyls \times 16 amines; Figure 1c). Although combinatorial reactions of multiple substrates or pooling of reactions before analysis can expedite screening,^{48–52,62,63} we chose to carry out and analyse reactions individually to generate a data set for machine learning that is as unaffected by inhibition phenomena and analytical artifacts as possible.

In-Silico Analysis

In the first stage of the screening process, we had to find representative panels of substrates and enzymes. To enable a rational and balanced selection, we compiled large collections of candidates, parametrised their structural properties, and applied dimensionality reduction by principal component analysis (PCA) and multidimensional scaling (MDS) to identify candidates that cover the search space relatively evenly.

Amines and carbonyl compounds included in the analysis comprised 136 literature-known IRED screening substrates and 55 additional structures relevant for medicinal chemistry applications (137 carbonyls and 54 amines in total; see Table S1). Ten descriptors based on two- and three-dimensional structure were computed for the amine and carbonyl substrates to represent their molecular geometry as well as their chemical characteristics (Table S21). Principal component analysis revealed clustering based mainly on the size and diameter of the compounds, the presence of halogens and, in case of carbonyls, of ester groups (Figures S2, S3). Multidimensional scaling confirmed these findings (Figures S4, S5). We selected eight carbonyl compounds and seven amines for experimental testing in the Overview Screening, and 30 carbonyl compounds and 19 amines for the Diversification Screening (see *Table of Screening Substrates* in the Supporting Information). These selections were made to thoroughly cover the majority of the principal component space, deliberately excluding regions of minor interest, such as small carbonyl compounds like acetone, 2-butanone, and butanal.

IRED sequences were gathered from various literature sources

(mainly refs. 26–28, 30, 64, and 65; see Table S23 for complete literature references) and from the BioCatNet IRED database (v. 3, accessed in January 2021). The initial dataset was expanded with additional sequences identified through BLAST searches⁶⁶ on the NCBI nr (non-redundant) database, using as queries nine IRED sequences reasonably distant from each other on a phylogenetic tree (see Figure S6). Redundancy removal at an identity threshold of 99% resulted in a dataset of 1,688 sequences. Further refinement was necessary to remove 68 β -hydroxyacid dehydrogenases (β -HADs), which share remarkable sequence similarity with IREDs (around 25% pairwise identity) but do not possess reductive amination activity. The filtering criterion was the presence of a catalytic lysine residue in the active site instead of the aspartic acid or tyrosine commonly found in IREDs.^{67–70} Homology models of the resulting 1,620 sequences were generated based on suitable templates (Figure S7, Table S22), and 65 further enzymes were removed based on model quality. Thus, the final *in-silico* panel comprised 1,555 curated sequences (available in CSV format at <https://doi.org/10.17632/jn4sfsrz8x.1>). From these, 115 IREDs (110 of which could be expressed in soluble form) were initially selected for the Overview Screening and later supplemented with batches of 45 and 20 additional enzymes based on machine-learning (ML) predictions of their reductive amination performance.

Overview Screening

The Overview Screening required the analysis of 9,800 individual biotransformations (8 carbonyls \times 7 amines \times 175 enzymes), a number that exceeds the capacities of standard chromatographic methods. As alternatives, we considered a colorimetric screen (called IREDy-to-go) developed by Turner and co-workers, which detects the formation of NADPH in the reverse reaction (*i.e.*, oxidative deamination),³⁰ and the photometric quantification of NADPH consumption in the reductive amination reaction, the feasibility of which has been demonstrated by Höhne and co-workers.⁷¹ We chose the latter option because it provides a more direct measure of the desired activity and implemented it in a 96-well microtiter plate format that required two plates per enzyme for the triplicate analysis of all 56 substrate combinations and a ‘no amine’ negative-control condition (see *Photometric IRED Activity Assay: Overview Screening*, Supporting Information). All carbonyl and amine substrates, except ammonia (500 mM, 10 eq.), were screened at equimolar concentrations of 50 mM each, and NADPH was supplied at 2 mM. Under these conditions, two substrate combinations (benzaldehyde–benzylamine and benzaldehyde–aniline) gave turbid assay mixtures, which prevented the photometric monitoring of NADPH consumption. All other combinations could be assayed without interfering absorption of light by substrates, imine, or product at a wavelength of 370 nm ($\epsilon_{\text{NADPH}} = 2,216 \text{ M}^{-1} \text{ cm}^{-1}$). Although all investigated IREDs were expressed with an N-terminal His-tag, we sought to avoid the effort of purifying all 175 proteins by establishing the assay using crude, lyophilised cell-free extracts. Previous studies suggest that the use of non-purified enzymes could result in pronounced background NADPH depletion; for instance, due to competing carbonyl reduction by endogenous *E. coli* ketoreductases.^{23,71} Gratifyingly, we found that our initially chosen assay conditions (bicine–NaOH buffer, 100 mM, pH 8.0, 10% v/v DMSO; 2 mg/mL crude IRED, corresponding to a substrate/enzyme ratio [w/w] of approx. 2–4) resulted in an acceptable background NADPH depletion rate (<5 mAU/min) in negative control reactions performed in the absence of amine, even for easily reduced carbonyl substrates such as benzaldehyde (c1) and cyclohexanone

(c2). This low background, combined with a statistical criterion for hit identification that takes the standard deviation of the triplicate biotransformations into account (Equation 3, Supporting Information), allowed us to reliably detect reductive amination activities as low as 5–10 mU/mg of IRED.

The results obtained in the photometric assay were validated by re-analysing selected substrate–enzyme combinations chromatographically. Biotransformations on a 500- μ L scale were set up using the appropriate substrates and IRED combined with phosphite dehydrogenase (PtDH) for cofactor regeneration, incubated for 24 h, and analysed by GC–MS or HPLC–MS. Three strategies guided the selection of substrate–enzyme combinations for re-analysis: (i) 50 combinations (25 positives, 25 negatives) were randomly selected to estimate the overall rate of false-positive and false-negative results, (ii) each positive substrate pair was re-analysed at least once, and (iii) low-level positives were given precedence for re-analysis, in particular if the respective substrate pair had yielded no high-level positives. These measures not only confirmed the overall validity of the photometric assay (Tables S8 and S9), but also contributed to the elimination of low-level false positives (Tables S10 and S11).

The consolidated activity data demonstrate that reductive am-

ination at equimolar substrate concentrations is more common across the entire IRED enzyme family than previously thought. Out of the 9,450 interpretable substrate–enzyme combinations, 1,437 (15%) were positives, and 36 of the 54 possible products (67%) could be accessed by at least one enzyme (Figure 2). Six of the 18 products that were inaccessible derive from α -ketoester c5, which was found to undergo competing ester hydrolysis under the assay conditions, a side reaction that severely interferes with the desired reductive amination process.

Aggregating the determined activities for each substrate across all coupling partners reveals general reactivity trends (Figure 3a,b): In line with previous findings, benzaldehyde (c1) and cyclohexanone (c2) were the most reactive carbonyl substrates, but different from earlier reports,^{25,28,29} the aromatic ketone acetophenone (c3) was accepted more readily than the aryl-aliphatic phenylacetone (c8). Among the tested amines, propargylamine (a4) and methylamine (a2) were accepted best, while the secondary amine piperidine (a5) proved to be a slightly more challenging substrate than the α -branched but primary isopropylamine (a3).

To compare the performance of different enzymes, we found it helpful to visualise the collected data along two aggregated di-

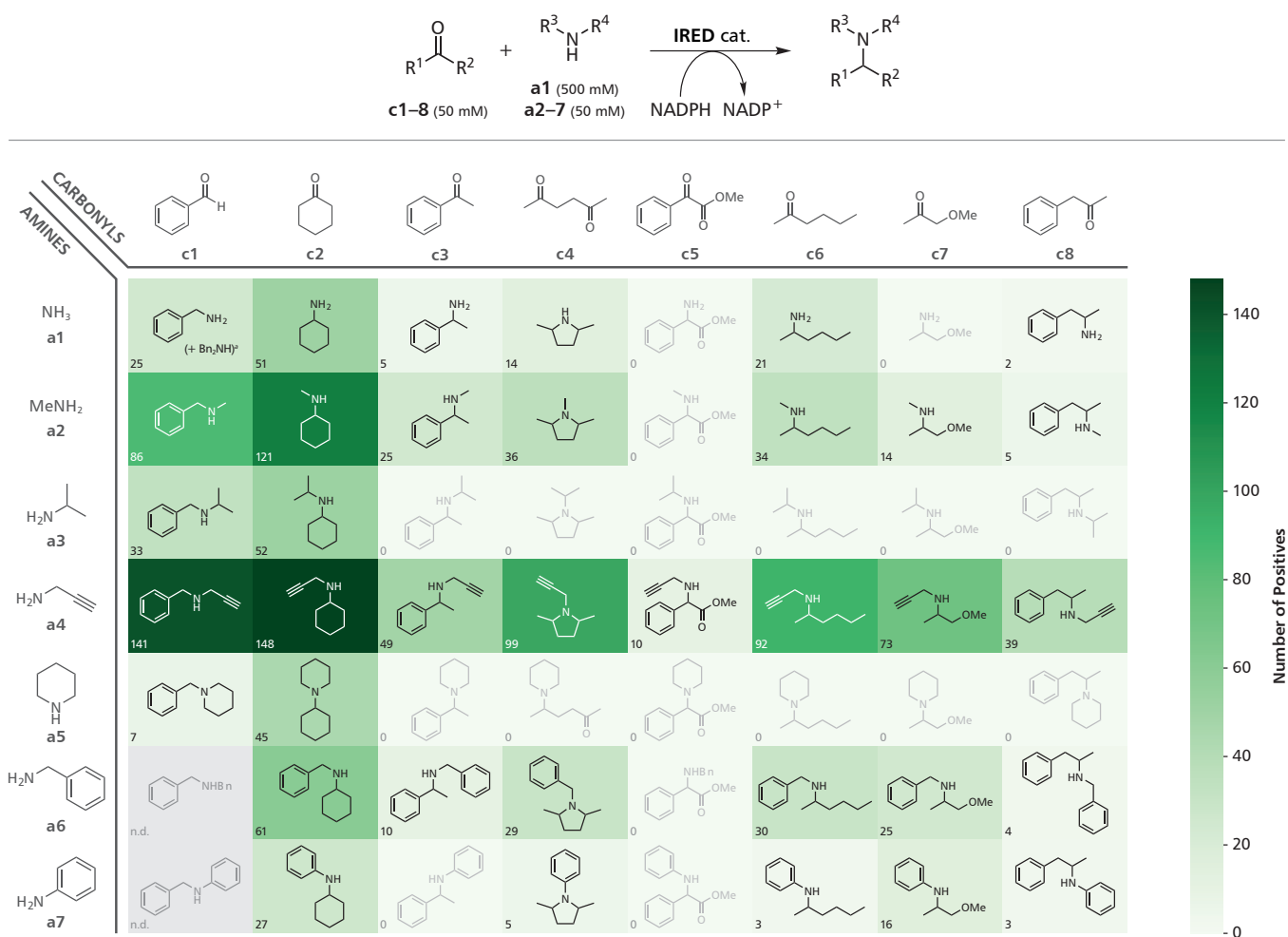


Figure 2 | Summary of positives identified in the Overview Screening. The coloured rectangles (tiles) show the products resulting from the 56 tested substrate combinations, along with the number of active IREDs for the respective combination given in the bottom-left corner of each tile. Activity was assayed spectrophotometrically by following the consumption of the IREDs' cofactor, NADPH, and confirmed by re-screening of selected positives using GC–MS or HPLC–MS analysis. The colour scale represents the number of confirmed hits found for each substrate combination. Two reactions that could not be assayed spectrophotometrically due to formation of turbidity (c1-a6, c1-a7) are shown as grey tiles. ^o Dibenzylamine (Bn₂NH) was formed alongside the expected benzylamine, and the latter was present in the EtOAc-extracted samples as the benzylidene imine (*N*-benzylidenebenzylamine).

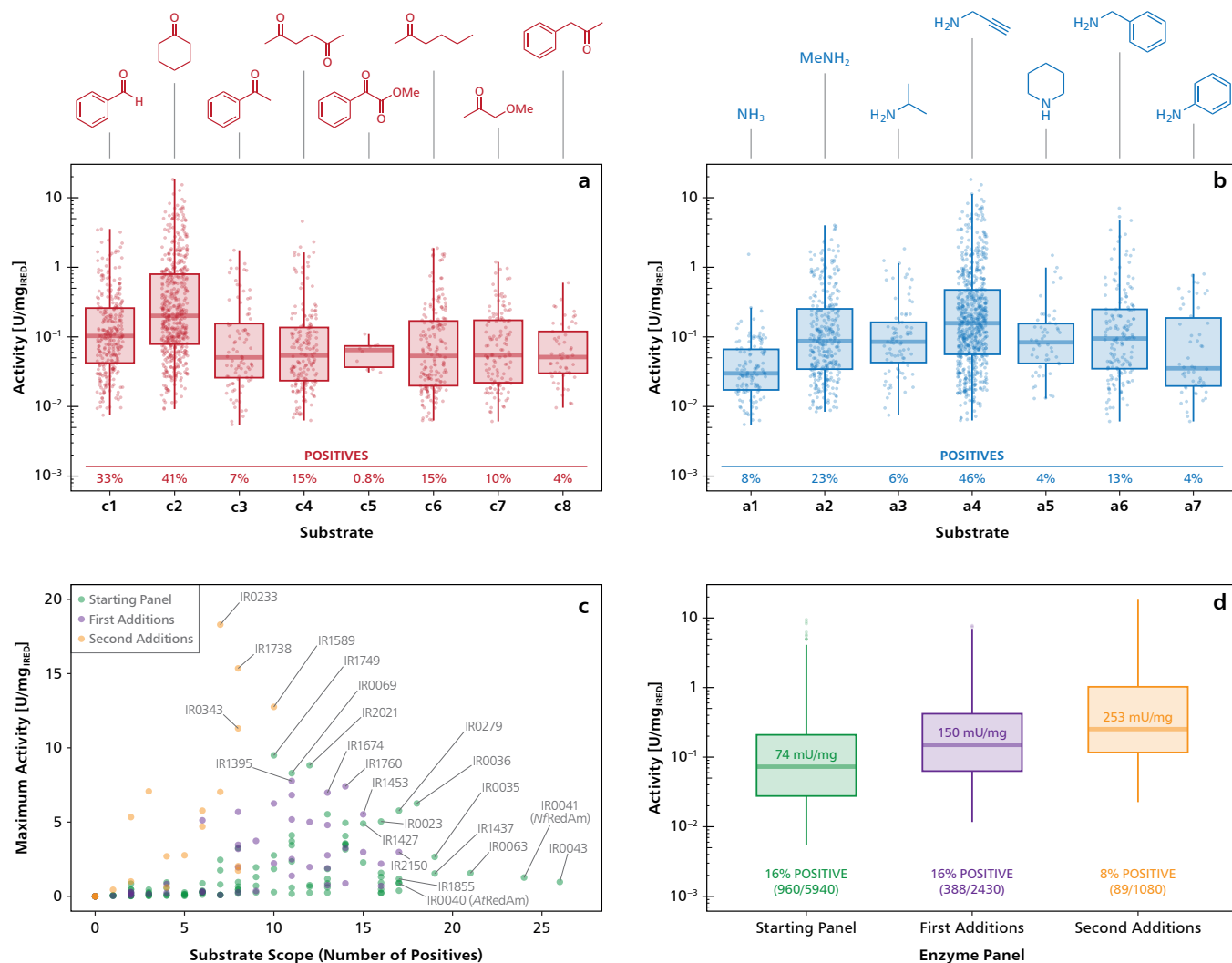


Figure 3 | Analysis of data collected in the Overview Screening. **a,b**, Box plots of non-zero activities (scaled logarithmically) of each carbonyl (**a**) and amine (**b**) substrate, showing median (centre band), 25th and 75th percentiles (lower and upper hinge of box), and data range (whiskers). Individual data points are shown as semi-transparent circles. **c**, Scatter plot of maximum activity vs. number of positive substrate combinations for each of the 175 tested IREDs. The labelled enzymes along the top-right front are compared to each other in Figure S15 and Table S4 (Supporting Information). **d**, Box plot of non-zero activities (scaled logarithmically) grouped by enzyme panel. Extreme values (data points farther than 1.5 times the inter-quartile range from the hinges) are shown as semi-transparent circles.

mensions: the maximum activity found across all tested substrate combinations (activity score), and the number of positive substrate combinations (generality score). This allows a rough categorisation of IREDs into generalists vs. specialists and fast vs. slow enzymes. As shown in Figure 3c, fast generalists (top-right quadrant) are notably absent, as the IREDs that reached the highest maximum activities are only moderately general, and *vice versa*, suggesting an inherent trade-off between substrate scope and catalytic efficiency in any single reaction. The enzymes located closest to the top-right quadrant (labelled in Figure 3c) share an average sequence identity of only 46% (with pairwise identities ranging from 26% to 96%) and originate from diverse phylogenetic backgrounds. However, most of them possess key active-site residues (Asn93, Asp169, Tyr177; AspRedAm numbering) that have been assigned a catalytic role in the proposed reaction mechanism for fungal RedAms (Figure S15, Table S4).^{18,25,72} In general, we found the presence of these key residues to be a useful predictor for reductive amination performance in the Overview Screening.

As discussed above (see *Screening Strategy, In-Silico Analysis*), the 175 tested IREDs were screened iteratively in three stages. Ini-

tially, a panel of 110 enzymes was tested, and the collected data were used to train a machine-learning (ML) model that correlated the determined activities to differences in physico-chemical descriptors of the enzymes' active sites, formulated as dissimilarity scores among the corresponding set of Catalophore™ point clouds (for details, see *Activity Prediction Using Machine-Learning Models*, Supporting Information). Results from this model were used to predict two additional panels of 45 and 20 IREDs expected to possess good reductive amination activity towards the substrate set of the Overview Screening. As shown in Figure 3c, the most general IREDs were found in the starting panel, while the enzymes recruited based on ML predictions tended towards moderate substrate scope but high maximum activity. This is reflected in an increase of the median non-zero activity across all substrate combinations from 74 mU/mg in the starting panel to 150 mU/mg (2-fold increase) in the first addition panel and to 253 mU/mg (3.4-fold increase) in the second addition panel (Figure 3d). Therefore, the ML model predicted high-activity enzyme candidates based on three-dimensional representations of the active sites and an experimental activity data set of 5,940 points (960 non-zero points).

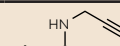
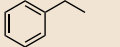
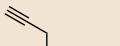
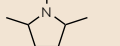

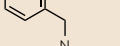
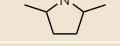
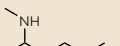
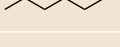
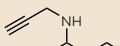
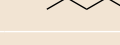
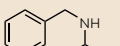

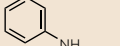
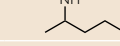
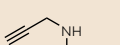
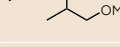
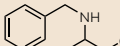
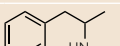
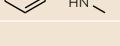
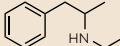
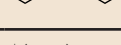
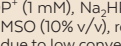
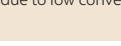




As most of the investigated reductive amination reactions lead to chiral amine products, stereoselectivity is another important measure of an enzyme's performance next to activity and substrate scope. To gather representative stereoselectivity data, we performed a total of 80 small-scale biotransformations covering 26 positive substrate combinations and determined conversion and enantiomeric excess of the products by GC and/or HPLC analysis. In agreement with previous studies on IRED-catalysed reductive amination,^{25,28,73} stereoselectivity varied widely across substrates even for the same enzyme, but among the investigated biotransformations were many that proceeded with high conversions and excellent enantio- and/or diastereoselectivity (Table 1; see Table S15 for complete data). Most reactions afforded the (*R*)-enantiomer of the product, but in some cases both enantiomers could be accessed (e.g., **c3-a4**, Table 1, entries 1 and 2; **c6-a4**, Table 1, entries 11–13). The reductive amination reactions of hexane-2,5-dione (**c4**) predominantly proceeded with high enantio- but moderate diastereoselectivity (Table 1, entries 3–7) and were accompanied by the formation of the corresponding pyrroles through a Paal–Knorr cyclisation that was also observed in negative control reactions performed in the absence of IRED. This

outcome is largely consistent with the findings of Cosgrove, Turner, and co-workers, who used similar 1,4-diketones as substrates for IRED-catalysed reductive amination.⁷⁴ The amphetamine derivative **c8-a4** was obtained with moderate conversion and high, but not exclusive, selectivity for the (*R*)-enantiomer. (*R*)-**c8-a4** is a precursor to selegiline, a pharmaceutical drug that can be obtained from **c8-a4** by *N*-alkylation and that is used for the treatment of Parkinson's disease and major depression.⁷⁵

Diversification Screening

We advanced the 100 best-performing IREDs (55 from the initial panel; 30 and 15 from the first and second addition panel, respectively) into a second screening phase to explore their reductive amination scope using a broader set of substrates, including larger and more functionalised compounds commonly used in medicinal chemistry. In this Diversification Screening, a total of 30 carbonyls and 19 amines were investigated, resulting in 57,000 individual substrate–enzyme combinations for exhaustive coverage. As this large number exceeded our screening capacity, we decided to explore

Table 1 | Conversions and stereoselectivities (enantiomeric excess, *ee*; diastereomeric ratio, *dr*) of selected IRED-catalysed reductive amination reactions of the Overview Screening.

Entry	Carbonyl	Amine	IRED	Product	Conv. [%]	<i>ee</i> [%]	<i>dr</i>
1	c3	a4	IRO059		1	98 (<i>S</i>)	
2	c3	a4	IRO041		10	85 (<i>R</i>)	
3	c4	a4	IR2021		24	60 (<i>A</i>) ^c	76:24
4	c4	a4	IRO036		15	78 (<i>A</i>) ^c	78:22
5	c4	a4	IRO068		12	81 (<i>A</i>) ^c	86:14
6	c4	a6	IRO464		14	>99 (<i>A</i>) ^c	15:85
7	c4	a6	IRO499		13	>99 (<i>A</i>) ^c	7:93
8	c6	a2	IR2021		59	98 (<i>R</i>) ^c	
9	c6	a2	IRO068		53	>99 (<i>R</i>) ^c	
10	c6	a2	IRO032		18	90 (<i>R</i>) ^c	
11	c6	a4	IR1749		97	>99 (<i>R</i>)	
12	c6	a4	IRO068		95	>99 (<i>R</i>)	
13	c6	a4	IRO395		61	99 (<i>S</i>)	
14	c6	a6	IR1453		87	99 (<i>R</i>)	
15	c6	a6	IR1460		73	>99 (<i>R</i>)	
16	c6	a6	IR2147		70	>99 (<i>R</i>)	
17	c6	a7	IRO041		6	>99 (<i>R</i>)	
18	c6	a7	IRO043		2	>99 (<i>R</i>)	
19	c7	a4	IRO068		>99	97 (<i>R</i>)	
20	c7	a4	IRO041		96	64 (<i>R</i>)	
21	c7	a4	IRO233		82	99 (<i>R</i>)	
22	c7	a6	IR2147		98	>99 (<i>R</i>)	
23	c7	a6	IR1453		94	>99 (<i>R</i>)	
24	c8	a2	IR1134		28	79 (<i>R</i>)	
25	c8	a2	IRO052		13	84 (<i>R</i>)	
26	c8	a4	IRO036		36	96 (<i>R</i>)	
27	c8	a4	IRO041		71	95 (<i>R</i>)	
28	c8	a4	IRO068		48	95 (<i>R</i>)	

Reaction conditions: Carbonyl compound (50 mM), amine (50 mM), NADP⁺ (1 mM), Na₂HPO₄ · 5 H₂O (100 mM), IRED preparation (4 mg/mL lyophilised cell lysate), PtDH preparation (4 mg/mL lyophilised cell lysate), bicine-NaOH buffer (100 mM, pH 8.0), DMSO (10% v/v), reaction volume 500 μL, 30 °C, 800 rpm, 24 h. ^a Conversion not determined due to hydrolytic instability of the carbonyl substrate. ^b Enantiomeric excess not determined due to low conversion. ^c (*A*) and (*B*) refer to the first-eluting and second-eluting enantiomer, respectively, where absolute configurations are unknown.

the substrate space non-exhaustively by splitting the screening into two parts: the carbonyl scope (24 carbonyls \times 4 amines) and amine scope (6 carbonyls \times 16 amines), each implemented in a 96-well plate format (Figures S25, S26). While this reduces the number of individual combinations to 19,200, we still considered triplicating every reaction (as was done in the Overview Screening) too experimentally costly. Therefore, only the negative controls – performed in the same plate layouts but using cell-free extract not containing any IRED (“empty-vector controls”) – were multiplied 9–10

times, while the enzymatic assays were carried out only once for each IRED. This makes the experimental replicates an additive rather than multiplicative effort, but still provides an estimate of global random error that allows the identification of hits based on a statistical criterion (Equations 5 and 7, Supporting Information).

The photometric NADPH consumption assay used in the Overview Screening could be readily adapted for the amine scope plate, with only six of 96 substrate combinations showing interfering turbidity. However, about a third of the combinations on the

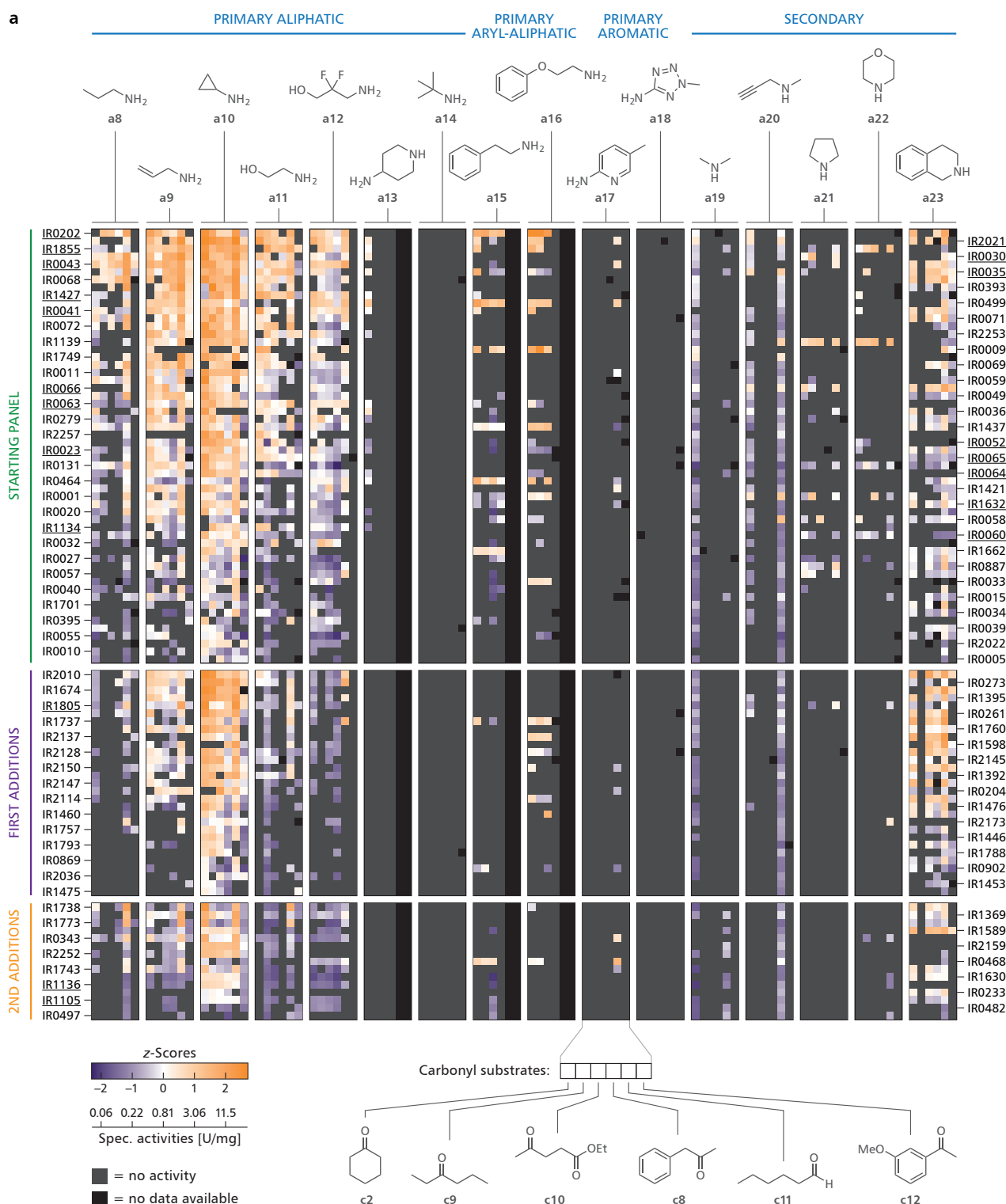


Figure 4 | Heatmap representation of all data collected in the Diversification Screening. (Continued on next page.)

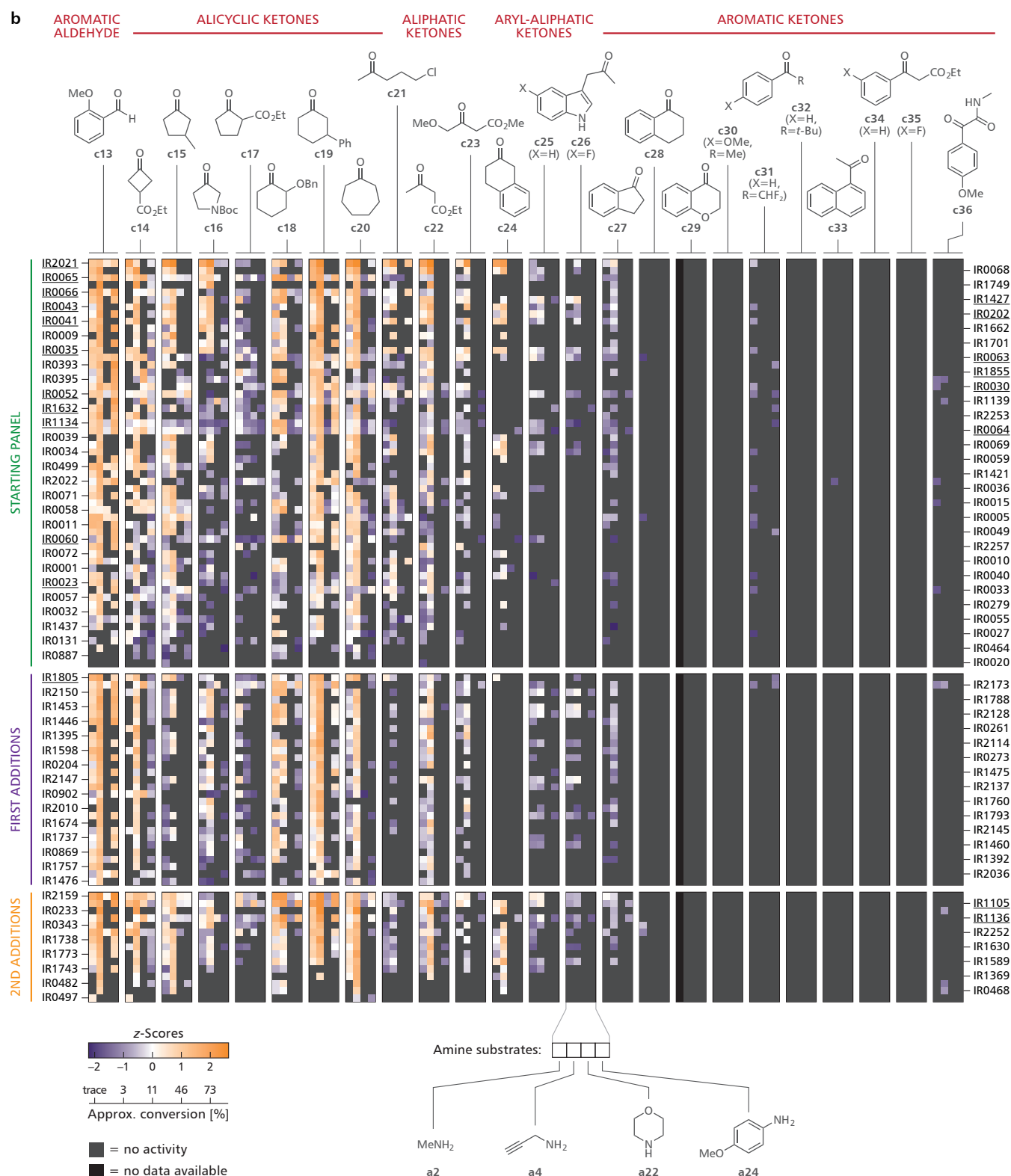


Figure 4 | Heatmap representation of all data collected in the Diversification Screening. Each of the 100 enzymes was screened against 6 carbonyl compounds \times 16 amines (amine scope, **a**) and 24 carbonyl compounds \times 4 amines (carbonyl scope, **b**). The collected specific activities (**a**) and target/buffer integral ratios (**b**) were logarithmised and converted to z-scores (distance from mean in units of standard deviation) as described in the Supporting Information. The z-scores are visualised using a diverging purple–orange colour scale with white as the centre point (mean, $z = 0$). Grey tiles represent inactive substrate–enzyme combinations, and black tiles indicate combinations for which no data are available. Enzymes are grouped according to the panel from which they were sourced, sorted within each panel by the sum of activities (highest on top), and labelled alternately on both sides for better readability. The twenty enzymes that showed the broadest substrate scope overall (amine and carbonyl scope combined) are underlined. (Continued from previous page.)

carbonyl scope plate proved unsuitable for photometric analysis due to turbidity or elevated background absorbance, making an alternative assay method necessary. Our recently developed flow-injection analysis mass spectrometry (FIA-MS) method,⁷⁶ which quantifies product formation by relating the mass signal ($M+H^+$) of the target amine to the mass signal of the bicine buffer, worked well for all combinations except one, in which the molecular masses of product and buffer coincide (**c29-a2**, $m/z = 164$ for $M+H^+$). FIA-MS thus allowed us to screen a diverse set of carbonyl substrates at a substantial concentration (50 mM), even when the reaction mixtures were not fully homogeneous.

While the Overview Screening had been set up manually using multi-channel pipettes, the two assay setups required for the Diversification Screening were automated using a liquid-handling robotic platform. This reduced the required hands-on time and improved throughput, as eight amine scope plates could be conveniently prepared and analysed in a single overnight run without any user interference. However, the robotic pipetting resulted in photometric reads of slightly lower quality compared to the manual preparation of plates, possibly due to changing the order of reagent addition to minimise the time required for the last pipetting operation and transfer of the plate into the reader (for details, see *Photometric IRED Activity Assay: Diversification Screening*, Supporting Information). Nevertheless, 8,916 of the 9,000 analysed assay reactions (99%) gave interpretable absorbance curves. For the carbonyl scope, samples were prepared autonomously by the robot and manually transferred to the autosampler for FIA-MS analysis. Hits were identified from the FIA-MS integral data by comparing the target/buffer ratio in enzymatic reactions to those in the multiplicate empty-vector controls, with no need for calibrating the target response (for details, see *FIA-MS Activity Assay: Diversification Screening*, Supporting Information).

As was the case for the Overview Screening (see above), we used chromatographic re-analysis of selected biotransformations to consolidate the results of the photometric and FIA-MS assays (Tables S12 and S13). This quality assurance step was particularly important for the carbonyl scope data, as the high sensitivity of single-ion monitoring MS resulted in product mass integrals significantly above the empty-vector background even for reactions that formed only minute levels of product. A cut-off for the target/buffer integral ratio was hence introduced to narrow down the identified hits to those that produced detectable amounts of product according to GC-MS or HPLC-MS (see *Validation Experiments for the Diversification Screening*, Supporting Information).

The complete results of the Diversification Screening are shown in Figure 4. To create this visualisation, the carbonyl scope and amine scope data, which are based on different performance metrics (specific activity in mU/mg_{IRED} for the amine scope, and target/buffer integral ratio as a dimensionless number for the carbonyl scope), had to be mapped onto a common scale. We found that both data sets, when plotted logarithmically, are reasonably well-described by Gaussian distributions (see *Data Transformation for Generating z-Score Heatmaps*, Supporting Information), enabling their transformation into z -scores (distance from mean in units of standard deviation). The aggregated data confirm our earlier conclusion that IRED-catalysed reductive amination at equimolar substrate concentrations is more common than previously believed. Overall, 27% of the interpretable assay reactions were positive (2,422 of 9,500 = 25% for the carbonyl scope, 2,529 of 8,916 = 28% for the amine scope; coloured tiles in Figure 4), and 118 out of 185 possible reductive amination products (64%) could be produced by at least one of the investigated IREDs (Figures S19 and

S20). Nevertheless, certain substrate motifs proved challenging for all tested enzymes, such as *tert*-butylamine (**a14**), which was not accepted at all, secondary amines, the heterocyclic primary amines **a17** and **a18**, and bulky, aromatic ketones (e.g., **c32-c35**; Figure 4). At the other end of the reactivity spectrum are amines like allylamine (**a9**) or cyclopropylamine (**a10**) and carbonyl compounds like 2-(benzyloxy)cyclohexanone (**c18**), 3-phenylcyclohexanone (**c19**), or *tert*-butyl 3-oxopyrrolidine-1-carboxylate (**c16**), which were accepted in 46–90% of the assayed reactions.

The data collected in the Diversification Screening do not indicate any trade-off between the substrate scope of an IRED and its maximum activity for any single substrate pair, in contrast to the findings of the Overview Screening (cf. Figure 3c). In fact, the enzymes with the broadest substrate acceptance are also among those with the highest individual activities, irrespective of whether the amine and carbonyl scope data are analysed separately or combined after z -score transformation (Figure S18, Table S6). The twenty top-performing IREDs (underlined in Figure 4 and highlighted in Figure S18) are active on more than one-third (34–44%) of the investigated substrate combinations and display maximum activities of 1.0–27.7 U/mg_{IRED} for the most favourable amine-carbonyl pair. Chromatographic re-investigation of representative positives with a focus on these 20 broad-scope enzymes revealed high conversions and excellent enantiomeric purities in many cases (Figure 5a; see Table S16 for complete data). Hence, our iterative screening approach has succeeded in identifying IREDs that catalyse reductive amination at equimolar substrate concentrations with broad scope, high activity, and high stereoselectivity.

Despite this gratifying outcome, it is essential to acknowledge the limitations of our approach to the Diversification Screening, which only covered 32% of all potential substrate combinations experimentally (18,416 of 57,000, considering only interpretable assay reactions). To address this limitation, we explored whether a machine-learning model trained on the available screening data could make useful activity predictions for amine-carbonyl pairs outside of the training set. We chose a gradient-boosted random forest for this task, and used physico-chemical descriptors of the substrates, extracted features of the IREDs' active-site point clouds, and z -transformed screening data as inputs (for details, see *Activity Prediction Using Machine-Learning Models*, Supporting Information). To assess the predictive performance of the model, we generated ranked lists of IREDs for four novel combinations of substrates (**c16-a6**, **c16-a15**, **c23-a10**, **c24-a6**) and picked the top 15 and bottom 5 candidates in each list for experimental testing. As shown in Tables S17–S20, the top-15 picks comprised positives with substantial conversion and *ee* in all cases, while the 5 lowest-ranked IREDs consistently showed little to no activity. Additionally, our model identified not only the 20 best-performing IREDs of the Diversification Screening (cf. Table S6), but also suggested less obvious candidates for experimental testing. In fact, in two cases these turned out to afford the best results in terms of conversion and enantioselectivity (IR0464 for **c16-a15**, IR2147 for **c24-a6**, see Figure 5b). This indicates that the DL model possesses predictive power for substrate pairs outside its training set that goes beyond pure face validity.

To demonstrate the synthetic applicability of the investigated biocatalytic reductive amination reactions, we scaled five of them to a preparative batch size (10 mmol carbonyl substrate, 200 mL reaction volume). Near-complete conversions (93% or higher) were achieved in all cases, and the desired products were isolated in 63–89% yield and high enantiomeric purity (98% to >99% *ee*, see Figure 5c).

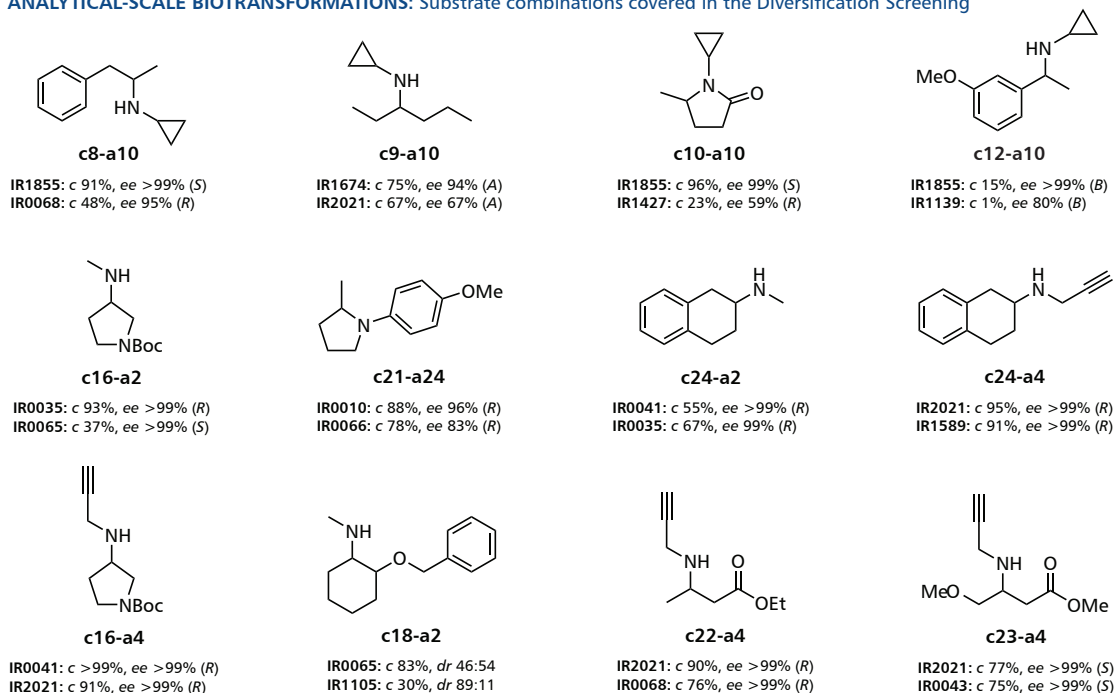
Conclusions

In summary, we have successfully used a data-driven, iterative screening approach to assemble a panel of imine reductases that perform asymmetric reductive amination at equimolar concentrations of the amine and carbonyl substrates with high activity, broad scope and good stereoselectivity. Instrumental in this process were the balanced, rational selection of enzymes and substrates for experimental testing based on systematic *in-silico* analysis, the use

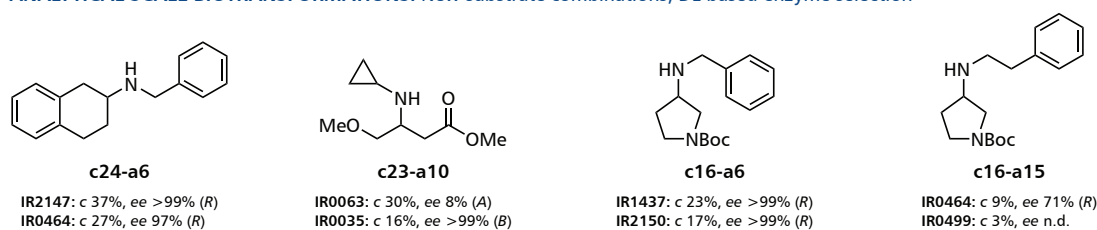
of photometric and mass spectrometry-based screening methods that combine sufficient throughput with broad applicability and a low rate of false positives and negatives, and the feedback of machine-learning models trained on the screening data, which guided our recruiting of additional IREDs and allowed us to extrapolate structure–activity relationships from the experimental data set to substrate combinations that were not covered by our screening.

The abundance of activity data this study has generated (27,866 total data points) is made openly accessible in a machine-readable

a ANALYTICAL-SCALE BIOTRANSFORMATIONS: Substrate combinations covered in the Diversification Screening



b ANALYTICAL-SCALE BIOTRANSFORMATIONS: New substrate combinations, DL-based enzyme selection



c PREPARATIVE-SCALE BIOTRANSFORMATIONS: 10 mmol substrates, 200 mL reaction volume

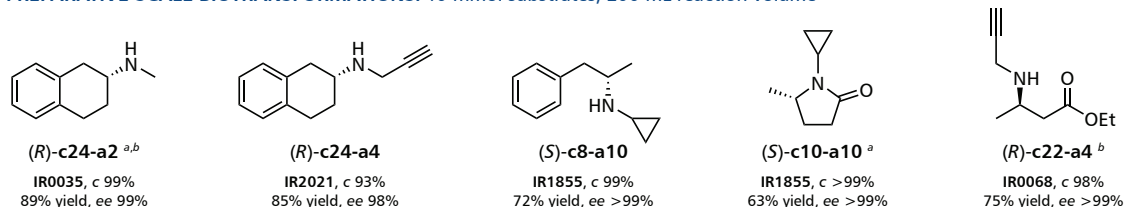


Figure 5 | Results of selected IRED-catalysed asymmetric reductive amination reactions. Conversion (c), enantiomeric excess (ee), and diastereomeric ratio (dr, where applicable) were determined by GC-MS, GC-FID, and HPLC-DAD analysis. Yield refers to isolated yield after purification by column chromatography. **a**, Biotransformations on analytical scale (0.5 mL) using substrate combinations covered in the Diversification Screening. **b**, Biotransformations on analytical scale (0.5 mL) using substrates of the Overview and Diversification Screenings in novel combinations. IREDs were selected for testing using a deep-learning (DL) model trained on the available screening data. **c**, Biotransformations on preparative scale (200 mL). *General reaction conditions:* Carbonyl compound (50 mM), amine (50 mM), NADP⁺ (1 mM), Na₂HPO₃ · 5 H₂O (100 mM), IRED preparation (4 mg/mL lyophilised cell lysate), PtDH preparation (4 mg/mL lyophilised cell lysate), bicine-NaOH buffer (100 mM, pH 8.0), DMSO (10% v/v), reaction volume 500 μL (**a**, **b**), 200 mL (**c**), 30 °C, 800 rpm, 24 h. ^a Increased amine substrate concentration (100 mM, 2 eq.) used. ^b Increased IRED loading (8 mg/mL lyophilised cell lysate) used.

format for re-analysis and re-use by other researchers. It represents a balanced fingerprint of reductive amination performance across the entire IRED/RedAm sequence space, without bias towards positive results⁷⁷ (77% of the tested substrate–enzyme combinations are non-hits). The data has been consolidated through experimental replicates, negative controls, and partial chromatographic re-analysis. It will serve as a valuable resource for the training and validation of machine-learning models that predict reductive amination outcome either from experimental data or from *in silico* analysis of substrate and enzyme structure directly (e.g., IREDFisher).⁵⁶

Overall, our screening demonstrates that reductive aminations at equimolar substrate concentrations are feasible with a substantial share of IRED enzymes and across a broad range of carbonyl and amine coupling partners. Out of the nearly 28,000 assay reactions conducted, a notable 23% (6,388 substrate–enzyme combinations) were positives, and 64% of the investigated amine products (154 out of 239) could be formed by at least one enzyme. The most proficient IREDs identified in our study combine a broad substrate scope (70–100 total positive combinations for Overview Screening and Diversification Screening) with high specific activities for the most favourable substrate pair (1.7–27.7 U/mg_{IRED}) and high stereoselectivity. This positions them as promising starting points for engineering towards a specific target reaction. While several of these top-performing biocatalysts are literature-known from earlier screening campaigns,^{26–28,30,73} others were experimentally characterised for the first time in the course of this work, highlighting both the remarkable proficiency of established IREDs and RedAms and the continued potential for discovery in this enzyme family.

Several of the reductive amination reactions investigated in this study were efficiently scaled up to gram scale. The desired amine products were isolated in good yields (63–89%) and with excellent enantiomeric purity (98% to >99% *ee*), demonstrating the preparative applicability of IRED-catalysed reductive amination at equimolar substrate concentrations. Therefore, we are confident that our results will actively encourage the wider use of imine reductases for biotransformations at low or no excess of amine, both in academic research and in industrial route development and production.

Methods

General Methods and Materials

¹H and ¹³C NMR spectra were recorded on 300 MHz and 500 MHz instruments (Bruker Avance III HD, Bruker Avance NEO 500). Chemical shifts are given in parts per million (ppm) relative to tetramethylsilane ($\delta = 0$ ppm) and coupling constants (*J*) are reported in Hertz (Hz). Thin layer chromatography (TLC) was used to follow the progress of synthetic reactions and to analyse fractions collected during preparative column chromatography. TLC was carried out on silica gel 60 F₂₅₄ plates (Merck) and compounds were visualised either by dipping into cerium ammonium molybdate (CAM) reagent [100 g/L (NH₄)₆Mo₇O₂₄ · 4 H₂O, 4 g/L Ce(SO₄)₂ · 4 H₂O, in 10% aq. H₂SO₄], by dipping into basic permanganate reagent (10 g/L KMnO₄, 50 g/L Na₂CO₃, 0.85 g/L NaOH, in H₂O), by spraying with ninhydrin solution (2 g/L ninhydrin in EtOH), or by UV light. Preparative chromatography was carried out on silica gel 60, particle size 40–63 μ m (230–400 mesh; Merck Millipore), either using traditional glassware or using an automated flash chromatograph (Biotage Selekt). Unit resolution GC–MS analyses were performed using electron impact (EI) ionisation at 70 eV and quadrupole mass selection. High-resolution MS analyses were performed using elec-

troscopy ionisation (ESI) and time-of-flight mass selection. Optical rotation values $[\alpha]_D^{20}$ were measured on an electronic polarimeter (Anton Paar MCP 5100) at 589 nm (Na D-line) and 20 °C using a cuvette of 1 dm path length.

Chemicals

Unless otherwise noted, all reagents and biotransformation substrates were obtained from commercial suppliers in reagent grade quality and used without further purification. Aniline used for reference synthesis or as biotransformation substrate was purified by short-path distillation (Kugelrohr) before use. Product reference compounds were either obtained from commercial suppliers in reagent-grade quality (c1-a1, c1-a2, c1-a3, c1-a5, c1-a6, c1-a7, c2-a1, c2-a2, c2-a5, c2-a6, c2-a7, c2-a9, c2-a11, c2-a15, c2-a19, c2-a21, c2-a22, c3-a1, c3-a2, c6-a1, c7-a1, c11-a19, c13-a2, c16-a2, c20-a2) or synthesised as described in the Supporting Information.

Enzymes

The imine reductases and phosphite dehydrogenase used in this study were produced by heterologous expression in *E. coli* as described in the Supporting Information. Whole-cell preparations of transaminases used for the synthesis of non-racemic reference compounds were prepared as described in previous publications from our group.^{78–80} Novozym® 435 (lipase B from *Candida antarctica* immobilised on acrylic resin; Novozymes A/S, Bagsvaerd, Denmark) was obtained from Pointner & Rothschädl GmbH (Baden, Austria).

Biotransformations

Reductive amination on analytical scale (general procedure): Stock solutions of the amine substrates (100 mM) in bicine–NaOH buffer (100 mM, pH 8.0) and of the carbonyl substrates (500 mM) in DMSO were prepared and stored as described in the Supporting Information. A stock solution (2.5 \times) of NADP⁺ sodium salt (1.9 mg/mL, 2.5 mM), disodium hydrogen phosphite pentahydrate (54 mg/mL, 250 mM) and PtDH preparation (10 mg/mL) was freshly prepared for each series of biotransformations in the required quantity. IRED preparations (2.0 \pm 0.2 mg) were weighed into microcentrifuge tubes (2.0 mL) and dissolved in NADP–phosphite–PtDH stock (200 μ L), followed by addition of amine stock (250 μ L) and carbonyl stock (50 μ L) to give final concentrations of 50 mM carbonyl compound, 50 mM amine, 1 mM NADP⁺, 100 mM phosphite, 4 mg/mL lyophilised IRED preparation (cell-free extract), 4 mg/mL lyophilised PtDH preparation (cell-free extract), and 10% (v/v) DMSO in a total volume of 500 μ L. Alternatively, individual stock solutions of NADP⁺ sodium salt (15.3 mg/mL, 20 mM), disodium hydrogen phosphite pentahydrate (216 mg/mL, 1.0 M), and PtDH preparation (40 mg/mL) could be prepared and used to achieve the same final concentrations and total reaction volume. The biotransformations were incubated at 30 °C and 120 rpm (horizontal placement of microcentrifuge tubes) in an incubator shaker or at 30 °C and 800 rpm (vertical placement of microcentrifuge tubes) in a benchtop thermoshaker for 24 h. Afterwards, the samples were subjected to one of the three work-up procedures described below.

Work-up A (extraction for GC): The reaction mixtures were basified (pH \geq 12) by addition of saturated aqueous Na₂CO₃ solution (200 μ L) and extracted with EtOAc (2 \times 500 μ L, cont. 10 mM *n*-decane or *n*-dodecane as internal standard), whereby phase separation was accelerated by centrifugation (13,000 rpm, 1 min). The organic phases were combined in another microcentrifuge tube (1.5 mL),

dried over Na_2SO_4 , and centrifuged again to pellet the solid. The supernatant was transferred to a crimp-cap glass vial for GC-MS or GC-FID analysis.

Work-up B (extraction and acetylation for GC): The reaction mixtures were basified ($\text{pH} \geq 12$) by addition of saturated aqueous Na_2CO_3 solution (200 μL), and EtOAc (1.0 mL, cont. 10 mM *n*-decane or *n*-dodecane as internal standard) as well as acetic anhydride (10 μL) were added. The samples were shaken by hand for 1 min, followed by centrifugation (13,000 rpm, 1 min) to separate the liquid phases. The organic phase was transferred to another microcentrifuge tube (1.5 mL), dried over Na_2SO_4 and K_2CO_3 , and centrifuged again to pellet the solids. The supernatant was transferred to a crimp-cap glass vial for GC-MS or GC-FID analysis.

Work-up C (dilution and protein removal for HPLC-MS): Aliquots (100 μL) of the reaction mixtures were transferred to microcentrifuge tubes (1.5 mL) and diluted with HPLC-grade acetonitrile (900 μL). Precipitated protein was pelleted by centrifugation (13,000 rpm, 1 min), and the supernatant was filtered through a layer of cotton in a pipette tip and transferred to a crimp-cap glass vial for HPLC-MS analysis.

Reductive amination on preparative scale (general procedure): A screw-cap Erlenmeyer flask (250 mL) was charged with a solution of the carbonyl substrate (10 mmol, 1 eq.; final concentration: 50 mM) in DMSO (20 mL) and a solution of the amine substrate (10–20 mmol, 1–2 eq.; final concentration: 50–100 mM) in bicine-NaOH buffer (180 mL; 100 mM, pH 8.0). Then, NADP⁺ sodium salt (153 mg, 0.2 mmol; final concentration: 1 mM), disodium hydrogen phosphite pentahydrate (4.32 g, 20 mmol, 2 eq.; final concentration: 100 mM), IRED preparation (0.8–1.6 g lyophilised cell lysate, final concentration 4–8 mg/mL), and PtDH preparation (0.8 g lyophilised cell lysate, final concentration 4 mg/mL) were added, and the flask was closed and placed in an incubator shaker at 30 °C and 120 rpm for 24 h. After this time, a sample (500 μL) was taken, worked up as described above (see *Reductive amination on analytical scale, Work-up A*), and analysed by GC-MS, indicating near-complete conversion in each case. The reaction mixture was basified ($\text{pH} \geq 12$) by addition of half-saturated Na_2CO_3 solution (50 mL), transferred to polypropylene tubes (8 × 50 mL), and extracted with EtOAc (4 × 20 mL per tube), whereby phase separation was accelerated by centrifugation (4,500 rpm, 5 min). The combined organic phases were dried over Na_2SO_4 , filtered, and the solvent was removed under reduced pressure to give the crude amine products, which were purified by column chromatography.

(S)-N-(1-Phenylpropan-2-yl)cyclopropanamine, (S)-c8-a10. Biotransformation using phenylacetone (**c8**; 1.36 g, 10.1 mmol), cyclopropylamine (**a10**; 0.73 mL, 0.60 g, 10.5 mmol), and IR1855 (0.8 g), followed by the general work-up, gave 4.2 g of a yellow liquid. Column chromatography using silica gel 60 (100 g, 15 × 5 cm) and cyclohexane/EtOAc = 1:2 afforded the title compound (1.28 g, 7.3 mmol, 72%) as a pale-yellow liquid. TLC (silica gel 60, cyclohexane/EtOAc = 1:2, KMnO_4 staining): R_f = 0.35. *ee* >99% (HPLC). $[\alpha]_{\text{D}}^{20}$ = +17.9 (*c* 1.00, CHCl_3), +15.9 (*c* 1.11, MeOH); lit.⁸¹ $[\alpha]_{\text{D}}^{26}$ = +16.1 (*c* = 0.56 mM, MeOH). ¹H NMR (300 MHz, CDCl_3): δ [ppm] = 7.33–7.25 (2H, m, Ar-H), 7.25–7.16 (3H, m, Ar-H), 3.04 (1H, sext, *J* = 6.5 Hz, $\text{CH}_2\text{-CH-CH}_3$), 2.80 (1H, dd, *J* = 13.3, 6.9 Hz, Ar- $\text{CH}_2\text{-CH}$), 2.59 (1H, dd, *J* = 13.3, 6.7 Hz, Ar- $\text{CH}_2\text{-CH}$), 2.06 (1H, tt, *J* = 6.6, 3.7 Hz, cyclopropyl CH), 1.82 (1H, br s, NH), 1.10 (3H, d, *J* = 6.3 Hz, CH_3), 0.53–0.25 (4H, m, cyclopropyl CH_2). ¹³C NMR (75 MHz, CDCl_3): δ [ppm] = 139.7, 129.4, 128.5, 126.2, 55.6, 43.8, 28.9, 20.7, 7.3, 6.0. GC-MS (EI, 70 eV): *m/z* = 175 (M^+ , 2), 160 (6), 146 (7), 118 (15), 115 (5), 91 (38), 84 (100), 65 (8), 56 (6), 41 (15). HRMS calcd for $\text{C}_{12}\text{H}_{18}\text{N}^+$ [$\text{M}+\text{H}^+$]: 176.1434, found 176.1445.

(R)-N-(Prop-2-yn-1-yl)-1,2,3,4-tetrahydronaphthalen-2-amine, (R)-c24-a4. Biotransformation using 2-tetralone (**c24**; 1.49 g, 10.2 mmol), propargylamine (**a4**; 0.67 mL, 0.58 g, 10.4 mmol), and IR2021 (0.8 g), followed by the general work-up, gave 3.6 g of a brown liquid. Column chromatography using silica gel 60 (70 g, 11 × 5 cm) and cyclohexane/EtOAc = 1:1 afforded the title compound (1.61 g, 8.7 mmol, 85%) as a dark-brown liquid. TLC (silica gel 60, cyclohexane/EtOAc = 1:1, KMnO_4 staining): R_f = 0.38. *ee* 98% (HPLC). $[\alpha]_{\text{D}}^{20}$ = +85.6 (*c* 1.03, CHCl_3). ¹H NMR (300 MHz, CDCl_3): δ [ppm] = 7.16–7.02 (4H, m, Ar-H), 3.55 (2H, dd, *J* = 2.5, 0.8 Hz, $\text{NH-CH}_2\text{-C}\equiv\text{CH}$), 3.28–3.13 (1H, m, NH-CH), 3.01 (1H, ddd, *J* = 16.0, 5.0, 1.6 Hz, CH_2), 2.96–2.76 (2H, m, CH_2), 2.60 (1H, dd, *J* = 16.0, 9.2 Hz, CH_2), 2.21 (1H, td, *J* = 2.4, 0.8 Hz, $\text{C}\equiv\text{CH}$), 2.11–1.95 (1H, m, CH_2), 1.72–1.53 (1H, m, CH_2), 1.48 (1H, br s, NH). ¹³C NMR (75 MHz, CDCl_3): δ [ppm] = 136.3, 135.0, 129.5, 128.8, 126.0, 125.8, 82.3, 71.5, 51.8, 36.3, 35.5, 29.2, 27.9. GC-MS (EI, 70 eV): *m/z* = 185 (M^+ , 28), 184 (96), 170 (73), 155 (23), 144 (42), 130 (100), 115 (68), 105 (24), 104 (85), 91 (43), 80 (88), 78 (37), 65 (14), 51 (12), 39 (29). HRMS calcd for $\text{C}_{13}\text{H}_{16}\text{N}^+$ [$\text{M}+\text{H}^+$]: 186.1277, found 186.1298.

(R)-N-Methyl-1,2,3,4-tetrahydronaphthalen-2-amine, (R)-c24-a2. Biotransformation using 2-tetralone (**c24**; 1.49 g, 10.2 mmol), methylamine (**a2**; 1.73 mL of a 40 wt-% solution in water, 20.0 mmol, 2 eq.), and IR0035 (1.6 g), followed by the general work-up, gave 4.4 g of a brown liquid. Column chromatography using silica gel 60 (100 g, 15 × 5 cm) and MTBE/MeOH/ NH_4OH = 90:9:1 afforded product that was still contaminated with residual DMSO. This material was taken up in MTBE (50 mL), washed with a 10% (w/v) aq. LiCl solution (20 mL), and the organic phase was dried over Na_2SO_4 and concentrated under reduced pressure to give the title compound (1.46 g, 9.1 mmol, 89%) as a brown liquid. TLC (silica gel 60, MeOH, KMnO_4 staining): R_f = 0.17. *ee* 99% (HPLC). $[\alpha]_{\text{D}}^{20}$ = +51.5 (*c* 1.18, CHCl_3). ¹H NMR (300 MHz, CDCl_3): δ [ppm] = 7.15–7.02 (4H, m, Ar-H), 3.03 (1H, ddd, *J* = 15.8, 4.9, 1.7 Hz, CH_2), 2.97–2.73 (3H, m, CH_2 and NH overlap), 2.60 (1H, ddd, *J* = 16.0, 9.3, 1.2 Hz, CH_2), 2.52 (3H, s, CH_3), 2.15–2.00 (1H, m, CH_2), 1.69–1.50 (2H, m, CH_2). ¹³C NMR (75 MHz, CDCl_3): δ [ppm] = 136.3, 135.3, 129.5, 128.8, 125.9, 125.8, 55.4, 36.4, 33.9, 29.2, 28.0. GC-MS (EI, 70 eV): *m/z* = 161 (M^+ , 100), 146 (29), 130 (56), 115 (21), 104 (36), 91 (11), 78 (13), 70 (11), 57 (10), 42 (9). HRMS calcd for $\text{C}_{11}\text{H}_{16}\text{N}^+$ [$\text{M}+\text{H}^+$]: 162.1277, found 162.1286.

(R)-Ethyl 3-(prop-2-yn-1-ylamino)butanoate, (R)-c22-a4. Biotransformation using ethyl acetoacetate (**c22**; 1.31 g, 10.1 mmol), propargylamine (**a4**; 0.67 mL, 0.58 g, 10.4 mmol), and IR0068 (1.6 g), followed by the general work-up, gave 3.3 g of a yellow liquid. Column chromatography using silica gel 60 (100 g, 15 × 5 cm) and cyclohexane/EtOAc = 1:1 afforded the title compound (1.28 g, 7.6 mmol, 75%) as a pale-yellow liquid. TLC (silica gel 60, cyclohexane/EtOAc = 1:1, KMnO_4 staining): R_f = 0.33. *ee* >99% (GC). $[\alpha]_{\text{D}}^{20}$ = -2.2 (*c* 1.93, CHCl_3). ¹H NMR (300 MHz, CDCl_3): δ [ppm] = 4.14 (2H, q, *J* = 7.2 Hz, $\text{CH}_2\text{-CH}_3$), 3.53 (1H, dd, *J* = 17.1, 2.5 Hz, $\text{NH-CH}_2\text{-C}\equiv\text{CH}$), 3.45 (1H, dd, *J* = 17.1, 2.5 Hz, $\text{NH-CH}_2\text{-C}\equiv\text{CH}$), 3.37 (sext, *J* = 6.4 Hz, NH-CH), 2.94 (1H, br s, NH), 2.52 (1H, dd, *J* = 15.6, 6.8 Hz, $\text{CH}_2\text{-CO}_2\text{Et}$), 2.40 (1H, dd, *J* = 15.6, 6.0 Hz, $\text{CH}_2\text{-CO}_2\text{Et}$), 2.24 (1H, t, *J* = 2.5 Hz, $\text{C}\equiv\text{CH}$), 1.25 (3H, t, *J* = 7.1 Hz, $\text{CH}_2\text{-CH}_3$), 1.16 (3H, d, *J* = 6.4 Hz, 3H, CH-CH_3). ¹³C NMR (75 MHz, CDCl_3): δ [ppm] = 172.0, 81.0, 72.3, 60.7, 48.9, 41.1, 35.5, 19.6, 14.3. GC-MS (EI, 70 eV): *m/z* = 169 (M^+ , <1), 154 (11), 130 (11), 108 (12), 96 (9), 82 (100), 66 (8), 54 (16), 42 (11). HRMS calcd for $\text{C}_9\text{H}_{16}\text{NO}_2^+$ [$\text{M}+\text{H}^+$]: 170.1176, found 170.1175.

(S)-1-Cyclopropyl-5-methylpyrrolidin-2-one, (**S**)-**c10-a10**. Biotransformation using ethyl levulinate (**c10**; 1.46 g, 10.1 mmol), cyclopropylamine (**a10**; 1.46 mL, 1.20 g, 21.1 mmol), and IR1855 (0.8 g), followed by the general work-up, afforded a mixture of the desired lactam along with the open-chain amino ester in a ratio of approx. 10:90. This material was taken up in MTBE (100 mL), washed with a 10% (w/v) aq. LiCl solution (20 mL) to remove residual DMSO, and the organic phase was dried over Na₂SO₄ and concentrated under reduced pressure to give 2.1 g of a yellow liquid. The residue was dissolved in cyclohexane (100 mL) and the solution transferred to a screw-cap Erlenmeyer flask (250 mL), Novozym® 435 (100 mg) was added, and the flask was closed and placed in an incubator shaker at 30 °C and 120 rpm for 16 h. Afterwards, the immobilised enzyme was removed by filtration through a glass frit and the filtrate was concentrated under reduced pressure to give the crude lactam (1.2 g) as a yellow liquid. Column chromatography using silica gel 60 (70 g, 10 × 5 cm) and EtOAc/MeOH = 5:1 afforded the title compound (0.89 g, 6.4 mmol, 63%) as a yellow liquid. TLC (silica gel 60, EtOAc/MeOH = 5:1, KMnO₄ staining): *R*_f = 0.55. *ee* >99% (GC). [α]_D²⁰ = -90.4 (c 1.07, CHCl₃). ¹H NMR (300 MHz, CDCl₃): δ [ppm] = 3.57 (1H, dqd, *J* = 7.7, 6.3, 4.6 Hz, CH-CH₃), 2.50–2.23 (3H, m, CH₂), 2.09 (1H, dddd, *J* = 12.7, 9.6, 7.8, 7.1 Hz, CH₂), 1.57 (1H, dddd, *J* = 12.7, 9.5, 6.0, 4.5 Hz, cyclopropyl CH), 1.24 (3H, d, *J* = 6.3 Hz, CH₃), 0.93 (1H, dtd, *J* = 9.7, 7.0, 5.4 Hz, cyclopropyl CH₂), 0.79 (1H, dddd, *J* = 10.4, 6.7, 5.4, 3.9 Hz, cyclopropyl CH₂), 0.65 (1H, dtd, *J* = 9.8, 7.0, 5.2 Hz, cyclopropyl CH₂), 0.51 (1H, dddd, *J* = 10.3, 7.0, 5.2, 4.0 Hz, cyclopropyl CH₂). ¹³C NMR (75 MHz, CDCl₃): δ [ppm] = 176.1, 55.2, 30.8, 26.4, 23.2, 20.1, 7.5, 4.3. GC-MS (EI, 70 eV): *m/z* = 139 (M⁺, 53), 124 (100), 96 (14), 84 (36), 68 (17), 57 (5), 55 (20), 41 (25), 39 (12). HRMS calcd for C₈H₁₄NO⁺ [M+H⁺]: 140.1070, found 140.1072.

Supporting Information

Additional experimental and computational data, a full description of the experimental, computational, and analytical methods, as well as compound characterisation data can be found in the Supporting Information.

Data Availability

The data that support the findings of this study (consolidated screening data, NMR spectra, protein and DNA sequences, chemoinformatic descriptors of the substrates) are openly available in the Mendeley Data repository (<https://data.mendeley.com/>) at <https://doi.org/10.17632/jn4sfsrz8x.1>. The Python scripts used for data collection and processing are available on GitHub (https://github.com/saberger/Screening_for_Generality).

Acknowledgements

This work was funded by F. Hoffmann–La Roche Ltd. (Basel, Switzerland) and Genentech Inc. (San Francisco, USA) via the companies' Technology Innovation & Science programme. Financial support by the University of Graz and the Field of Excellence 'Bio-Health' is gratefully acknowledged. The authors would like to thank Markus Köck (University of Graz) for HRMS measurements, Bernd Werner (University of Graz) for recording NMR spectra, and Lisa M. Berger (University of Graz) for help with data analysis.

Author Contributions

Conception of the project: J.H.S., D.W., H.I., C.C.G., G.S.

Acquisition of funding: J.H.S., D.W., H.I., C.C.G., G.S., W.K.

Experiments: S.A.B. (enzyme expression, screening, method development on robotic platform, synthesis of reference compounds, biotransformations on analytical scale, biotransformations on preparative scale, development of analytical methods); C.G. (enzyme expression, screening, method development on robotic platform, synthesis of reference compounds, biotransformations on analytical scale, development of analytical methods); I.O.-G. (method development on robotic platform); J.H.S. (synthesis of reference compounds, biotransformations on analytical scale, development of analytical methods); S.V., L.R., M.W., F.M.K., Y.O. (enzyme expression, synthesis of reference compounds, biotransformations on analytical scale)

Computation: S.A.B. (script-based data processing); M.C., B.M.N. (bioinformatic analysis and selection of enzyme panel); I.M. (chemoinformatic analysis of substrate panel); A.K., T.S. (AI analysis of screening results); J.H.S. (explorative data analysis and *z*-score transformation)

Supervision of the project: J.H.S., B.M.N., S.B., H.I., F.G., K.P., D.W.

Preparation of manuscript: S.A.B., J.H.S., I.M., A.K., M.C. (first draft, preparation of tables and figures); S.A.B., S.V., C.G., B.M.N., S.B., H.I., J.H.S., W.K., F.G., K.P. (editing and proof-reading)

References

- Nugent, T. C. *Chiral Amine Synthesis: Methods, Developments and Applications*. (Wiley-VCH, 2010).
- Top 200 Small Molecule Drugs by Sales in 2023 Poster*. (Njardarson Laboratory, University of Arizona, 2024). Available online at: <https://sites.arizona.edu/njardarson-lab/top200-posters/>
- McGrath, N. A., Brichacek, M. & Njardarson, J. T. A Graphical Journey of Innovative Organic Architectures That Have Improved Our Lives. *J. Chem. Educ.* **87**, 1348–1349 (2010).
- Cunillera, A., Claver, C., Godard, C., Urrutigoity, M. & Kalck, P. in *Methologies in Amine Synthesis*. (eds Alfredo Ricci & Luca Bernardi), 155–185 (Wiley-VCH, 2021).
- Li, T. et al. Asymmetric Reductive Amination in Organocatalysis and Biocatalysis. *Eur. J. Org. Chem.* **26**, e202300507 (2023).
- Reshi, N. U. D., Saptal, V. B., Beller, M. & Bera, J. K. Recent Progress in Transition-Metal-Catalyzed Asymmetric Reductive Amination. *ACS Catal.* **11**, 13809–13837 (2021).
- Shi, Y., Rong, N., Zhang, X. & Yin, Q. Synthesis of Chiral Primary Amines via Enantioselective Reductive Amination: From Academia to Industry. *Synthesis* **55**, 1053–1068 (2023).
- Tian, Y., Hu, L., Wang, Y.-Z., Zhang, X. & Yin, Q. Recent advances on transition-metal-catalysed asymmetric reductive amination. *Org. Chem. Front.* **8**, 2328–2342 (2021).
- Hollmann, F., Opperman, D. J. & Paul, C. E. Biocatalytic Reduction Reactions from a Chemist's Perspective. *Angew. Chem. Int. Ed.* **60**, 5644–5665 (2021).
- Gomm, A. & O'Reilly, E. Transaminases for chiral amine synthesis. *Curr. Opin. Chem. Biol.* **43**, 106–112 (2018).
- Hummel, W. & Gröger, H. in *Enzyme Catalysis in Organic Synthesis*. 3rd edn, Vol. 1, 1165–1203 (Wiley-VCH, 2012).
- Kelly, S. A., Mix, S., Moody, T. S. & Gilmore, B. F. Transaminases for industrial biocatalysis: novel enzyme discovery. *Appl. Microbiol. Biotechnol.* **104**, 4781–4794 (2020).
- Kelly, S. A. et al. Application of ω -Transaminases in the Pharmaceutical Industry. *Chem. Rev.* **118**, 349–367 (2018).
- Patil, M. D., Grogan, G., Bommarius, A. & Yun, H. Recent Advances in ω -Transaminase-Mediated Biocatalysis for the Enantioselective Synthesis of Chiral Amines. *Catalysts* **8**, 254 (2018).
- Seah, S. Y. K. in *Industrial Enzymes: Structure, Function and Applications*. (eds Julio Polaina & Andrew P. MacCabe), 489–504 (Springer, 2007).

16. Xue, Y.-P., Cao, C.-H. & Zheng, Y.-G. Enzymatic asymmetric synthesis of chiral amino acids. *Chem. Soc. Rev.* **47**, 1516–1561 (2018).
17. Cosgrove, S. C. et al. in *Methods Enzymol.*, Vol. 608 (ed Nigel Scrutton) 131–149 (Academic Press, 2018).
18. Ducrot, L., Bennett, M., Grogan, G. & Vergne-Vaxelaire, C. NAD(P) H-Dependent Enzymes for Reductive Amination: Active Site Description and Carbonyl-Containing Compound Spectrum. *Adv. Synth. Catal.* **363**, 328–351 (2021).
19. Gilio, A. K., Thorpe, T. W., Turner, N. & Grogan, G. Reductive aminations by imine reductases: from milligrams to tons. *Chem. Sci.* **13**, 4697–4713 (2022).
20. Grogan, G. Synthesis of chiral amines using redox biocatalysis. *Curr. Opin. Chem. Biol.* **43**, 15–22 (2018).
21. Wu, K., Huang, J. & Shao, L. Imine Reductases: Multifunctional Biocatalysts with Varying Active Sites and Catalytic Mechanisms. *ChemCatChem* **14**, e202200921 (2022).
22. Aleku, G. A. Imine Reductases and Reductive Aminases in Organic Synthesis. *ACS Catal.* **14**, 14308–14329 (2024).
23. Huber, T. et al. Direct Reductive Amination of Ketones: Structure and Activity of S-Selective Imine Reductases from *Streptomyces*. *ChemCatChem* **6**, 2248–2252 (2014).
24. Scheller, P. N., Lenz, M., Hammer, S. C., Hauer, B. & Nestl, B. M. Imine Reductase-Catalyzed Intermolecular Reductive Amination of Aldehydes and Ketones. *ChemCatChem* **7**, 3239–3242 (2015).
25. Aleku, G. A. et al. A reductive aminase from *Aspergillus oryzae*. *Nat. Chem.* **9**, 961–969 (2017).
26. Wetzl, D. et al. Asymmetric Reductive Amination of Ketones Catalyzed by Imine Reductases. *ChemCatChem* **8**, 2023–2026 (2016).
27. Roiban, G.-D. et al. Efficient Biocatalytic Reductive Aminations by Extending the Imine Reductase Toolbox. *ChemCatChem* **9**, 4475–4479 (2017).
28. France, S. P. et al. Identification of Novel Bacterial Members of the Imine Reductase Enzyme Family that Perform Reductive Amination. *ChemCatChem* **10**, 510–514 (2018).
29. Montgomery, S. L. et al. Characterization of imine reductases in reductive amination for the exploration of structure-activity relationships. *Sci. Adv.* **6**, eaay9320 (2020).
30. Marshall, J. R. et al. Screening and characterization of a diverse panel of metagenomic imine reductases for biocatalytic reductive amination. *Nat. Chem.* **13**, 140–148 (2021).
31. Schober, M. et al. Chiral synthesis of LSD1 inhibitor GSK2879552 enabled by directed evolution of an imine reductase. *Nat. Catal.* **2**, 909–915 (2019).
32. Kumar, R. et al. Biocatalytic reductive amination from discovery to commercial manufacturing applied to abrocitinib JAK1 inhibitor. *Nat. Catal.* **4**, 775–782 (2021).
33. Ma, E. J. et al. Machine-Directed Evolution of an Imine Reductase for Activity and Stereoselectivity. *ACS Catal.* **11**, 12433–12445 (2021).
34. Yang, Z.-Y. et al. Direct Reductive Amination of Biobased Furans to *N*-Substituted Furfurylamines by Engineered Reductive Aminase. *Adv. Synth. Catal.* **363**, 1033–1037 (2021).
35. Zhang, J. et al. Tuning an Imine Reductase for the Asymmetric Synthesis of Azacycloalkylamines by Concise Structure-Guided Engineering. *Angew. Chem. Int. Ed.* **61**, e202201908 (2022).
36. Duan, S. et al. Application of Biocatalytic Reductive Amination for the Synthesis of a Key Intermediate to a CDK 2/4/6 Inhibitor. *Org. Process Res. Dev.* **26**, 879–890 (2022).
37. Steflik, J. et al. Engineering of a Reductive Aminase to Enable the Synthesis of a Key Intermediate to a CDK 2/4/6 Inhibitor. *ACS Catal.* **13**, 10065–10075 (2023).
38. Liu, T. et al. Enantiodivergent Synthesis of 1-Heteroaryl Tetrahydroisoquinolines Catalyzed by Imine Reductases. *Org. Lett.* **25**, 2438–2443 (2023).
39. Zhang, J. et al. Structure-guided semi-rational design of an imine reductase for enantio-complementary synthesis of pyrrolidinamine. *Chem. Sci.* **14**, 4265–4272 (2023).
40. Zhu, F. et al. Semi-rational design of an imine reductase for asymmetric synthesis of alkylated S-4-azepanamines. *Org. Biomol. Chem.* **21**, 4181–4184 (2023).
41. Casamajo, A. R. et al. Biocatalysis in Drug Design: Engineered Reductive Aminases (RedAms) Are Used to Access Chiral Building Blocks with Multiple Stereocenters. *J. Am. Chem. Soc.* **145**, 22041–22046 (2023).
42. Bellman, R. E. *Dynamic Programming*. (Princeton University Press, 1957).
43. Longwell, C. K., Labanieh, L. & Cochran, J. R. High-throughput screening technologies for enzyme engineering. *Curr. Opin. Biotechnol.* **48**, 196–202 (2017).
44. Markel, U. et al. Advances in ultrahigh-throughput screening for directed enzyme evolution. *Chem. Soc. Rev.* **49**, 233–262 (2020).
45. Payne, E. M., Holland-Moritz, D. A., Sun, S. & Kennedy, R. T. High-throughput screening by droplet microfluidics: perspective into key challenges and future prospects. *Lab Chip* **20**, 2247–2262 (2020).
46. Vervoort, N., Goossens, K., Baeten, M. & Chen, Q. Recent advances in analytical techniques for high throughput experimentation. *Anal. Sci. Adv.* **2**, 109–127 (2021).
47. Bansal, A., Sharma, R. & Kathuria, M. A Systematic Review on Data Scarcity Problem in Deep Learning: Solution and Applications. *ACM Computing Surveys* **54**, 208 (2022).
48. Goddard, J.-P. & Reymond, J.-L. Enzyme Activity Fingerprinting with Substrate Cocktails. *J. Am. Chem. Soc.* **126**, 11116–11117 (2004).
49. Kim, H. et al. A multi-substrate screening approach for the identification of a broadly applicable Diels–Alder catalyst. *Nat. Commun.* **10**, 770 (2019).
50. Knorrscheidt, A. et al. Simultaneous screening of multiple substrates with an unspecific peroxxygenase enabled modified alkane and alkene oxyfunctionalisations. *Catal. Sci. Technol.* **11**, 6058–6064 (2021).
51. McDonald, A. D., Higgins, P. M. & Buller, A. R. Substrate multiplexed protein engineering facilitates promiscuous biocatalytic synthesis. *Nat. Commun.* **13**, 5242 (2022).
52. Wagen, C. C., McMinn, S. E., Kwan, E. E. & Jacobsen, E. N. Screening for generality in asymmetric catalysis. *Nature* **610**, 680–686 (2022).
53. Clements, H. D. et al. Using Data Science for Mechanistic Insights and Selectivity Predictions in a Non-Natural Biocatalytic Reaction. *J. Am. Chem. Soc.* **145**, 17656–17664 (2023).
54. Haas, B. C., Goetz, A. E., Bahamonde, A., McWilliams, J. C. & Sigman, M. S. Predicting relative efficiency of amide bond formation using multivariate linear regression. *Proc. Natl. Acad. Sci. USA* **119**, e2118451119 (2022).
55. Xu, J. et al. Atroposelective Negishi Coupling Optimization Guided by Multivariate Linear Regression Analysis: Asymmetric Synthesis of KRAS G12C Covalent Inhibitor GDC-6036. *J. Am. Chem. Soc.* **144**, 20955–20963 (2022).
56. Yu, Y. et al. Structure-Based Design of Small Imine Reductase Panels for Target Substrates. *ACS Catal.* **13**, 12310–12321 (2023).
57. Abolhasani, M. & Kumacheva, E. The rise of self-driving labs in chemical and materials sciences. *Nat. Synth.* **2**, 483–492 (2023).
58. Bennett, J. A. & Abolhasani, M. Autonomous chemical science and engineering enabled by self-driving laboratories. *Curr. Opin. Chem. Eng.* **36**, 100831 (2022).
59. Häse, F., Roch, L. M. & Aspuru-Guzik, A. Next-Generation Experimentation with Self-Driving Laboratories. *Trends Chem.* **1**, 282–291 (2019).
60. Seifrid, M. et al. Autonomous Chemical Experiments: Challenges and Perspectives on Establishing a Self-Driving Lab. *Acc. Chem. Res.* **55**, 2454–2466 (2022).
61. Dai, T. et al. Autonomous mobile robots for exploratory synthetic chemistry. *Nature* (2024). DOI: 10.1038/s41586-024-08173-7
62. Satyanarayana, T. & Kagan, H. B. The Multi-Substrate Screening of Asymmetric Catalysts. *Adv. Synth. Catal.* **347**, 737–748 (2005).
63. Stanišić, A., Hüsken, A. & Kries, H. HAMA: a multiplexed LC-MS/MS assay for specificity profiling of adenylate-forming enzymes. *Chem. Sci.* **10**, 10395–10399 (2019).
64. Velikogne, S., Resch, V., Dertnig, C., Schrittwieser, J. H. & Kroutil, W. Sequence-Based *In-silico* Discovery, Characterisation, and Biocatalytic Application of a Set of Imine Reductases. *ChemCatChem* **10**, 3236–3246 (2018).
65. Wetzl, D. et al. Expanding the Imine Reductase Toolbox by Exploring the Bacterial Protein-Sequence Space. *ChemBioChem* **16**, 1749–1756 (2015).
66. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).
67. Fademrecht, S., Scheller, P. N., Nestl, B. M., Hauer, B. & Pleiss, J. Identification of imine reductase-specific sequence motifs. *Proteins Struct. Funct. Bioinf.* **84**, 600–610 (2016).
68. Lenz, M. et al. New imine-reducing enzymes from β -hydroxyacid dehydrogenases by single amino acid substitutions. *Protein Eng. Des. Sel.* **31**, 109–120 (2018).

69. Stockinger, P., Roth, S., Müller, M. & Pleiss, J. Systematic Evaluation of Imine-Reducing Enzymes: Common Principles in Imine Reductases, β -Hydroxy Acid Dehydrogenases, and Short-Chain Dehydrogenases/Reductases. *ChemBioChem* **21**, 2689–2695 (2020).
70. Stockinger, P. et al. Engineering of Thermostable β -Hydroxyacid Dehydrogenase for the Asymmetric Reduction of Imines. *ChemBioChem* **21**, 3511–3514 (2020).
71. Matzel, P., Krautschick, L. & Höhne, M. Photometric Characterization of the Reductive Amination Scope of the Imine Reductases from *Streptomyces tsukubaensis* and *Streptomyces ipomoeae*. *ChemBioChem* **18**, 2022–2027 (2017).
72. Sharma, M. et al. A Mechanism for Reductive Amination Catalyzed by Fungal Reductive Aminases. *ACS Catal.* **8**, 11534–11541 (2018).
73. Mangas-Sanchez, J. et al. Asymmetric synthesis of primary amines catalyzed by thermotolerant fungal reductive aminases. *Chem. Sci.* **11**, 5052–5057 (2020).
74. Ramsden, J. I. et al. Bifunctional Imine Reductase Cascades for the Synthesis of Saturated N-Heterocycles. *ACS Catal.* **14**, 14703–14710 (2024).
75. Hu, Y. et al. Chemoenzymatic Synthesis of Selegiline: An Imine Reductase-Catalyzed Approach. *Molecules* **29**, 1328 (2024).
76. Berger, S. A. et al. Rapid, Label-Free Screening of Diverse Biotransformations by Flow-Injection Mass Spectrometry. *ChemBioChem* **24**, e202300170 (2023).
77. Taniike, T. & Takahashi, K. The value of negative results in data-driven catalysis research. *Nat. Catal.* **6**, 108–111 (2023).
78. Koszelewski, D., Pressnitz, D., Clay, D. & Kroutil, W. Deracemization of Mexiletine Biocatalyzed by ω -Transaminases. *Org. Lett.* **11**, 4810–4812 (2009).
79. Fuchs, C. S. et al. Asymmetric Amination of α -Chiral Aliphatic Aldehydes via Dynamic Kinetic Resolution to Access Stereocomplementary Brivaracetam and Pregabalin Precursors. *Adv. Synth. Catal.* **360**, 768–778 (2018).
80. Alvarenga, N. et al. Asymmetric Synthesis of Dihydropinidine Enabled by Concurrent Multienzyme Catalysis and a Biocatalytic Alternative to Krapcho Dealkoxycarbonylation. *ACS Catal.* **10**, 1607–1620 (2020).
81. Gergel, S. et al. Engineered cytochrome P450 for direct arylalkene-to-ketone oxidation via highly reactive carbocation intermediates. *Nat. Catal.* **6**, 606–617 (2023).