Biocatalytic reductive amination from discovery to commercial manufacturing applied to abrocitinib JAK1 inhibitor

Rajesh Kumar1, Michael J. Karmilowicz1, Dylan Burke2, Michael P. Burns1, Leslie A. Clark2, Christina G. Connor1, Eric Cordi1, Nga M. Do1, Kevin M. Doyle3, Steve Hoagland3, Chad A. Lewis1, David Mangan2, Carlos A. Martínez2, Emma L. McInturff1, Kevin Meldrum4, Robert Pearson1, Jeremy Steflik1, Anil Rane1 and John Weaver1

Enzymatic reductive amination, being a direct, selective and green methodology, has attracted significant interest in a short period of time and is emerging as a powerful tool for the synthesis of chiral alkylated amines. The discovery of an increasing number of imine reductases with reductive aminase (RedAm) activity has enabled mechanistic and substrate profiling studies. However, their potential for commercial applications has not been realized. Here, we report the discovery of RedAm activity in an imine reductase enzyme for the direct reductive amination of a cyclic ketone with methylamine. We also investigate engineering the enzyme to access a cis-cyclobutyl-N-methylamine for the manufacturing of a late-stage drug candidate, Janus kinase 1 (JAK1) inhibitor abrocitinib. The engineered enzyme, SpRedAm-R3-V6, showed >200-fold improvement in performance over the wild-type enzyme and was successfully used to develop a commercial manufacturing process with 73% isolated yield at 99.5% purity and high selectivity (>99:1 cis:trans). This process has been successfully used to manufacture multi-metric tons of the amine, demonstrating the potential of RedAm technology for commercial manufacturing.

Efficient and sustainable synthesis of chiral amines, a prominent synthetic motif in drug molecules1, has spurred the recent advances in the development of innovative and sustainable synthetic methods, including both traditional chemical2–4 and enzymatic methods5–7. Reductive amination, being highly versatile, is one of the most frequently used transformations for the synthesis of a variety of amines. Chemical methods for reductive amination commonly use stoichiometric amounts of reducing agents at low temperature, or transition metal catalysis, making them unsafe at scale and environmentally unsustainable2,8. Enzymatic methods are more desirable9–11 for organic synthesis and manufacturing due to their potential for high selectivity, coupled with safety and environmental benefits. Enzymatic synthesis of primary amines has been well established12 using transaminases13,14 or amino acid and amine dehydrogenases15,16. Recently, considerable progress has been made to evolve amino acid and amine dehydrogenases to perform reductive amination of ketones with ammonia and in some cases with methylamine17 for preparative scale applications. Although significant advances have been made for the enzymatic synthesis of primary amines, reductive amination of ketones with alkylamines remains a challenge. One of the recent innovations in enzyme catalysis is the discovery of imine reductases (IREDs)18–22 for catalysing the reduction of C=N bonds to give amines. The ability of a subfamily of IREDS classified as reductive aminases (RedAms) to perform reductive amination of a ketone or an aldehyde with an alkylamine to give alkylated amines is transformative, as this requires concurrent binding of both substrates in the active site with proper orientation23–26. Enzymatic reductive amination has attracted significant interest in a short period of time, with several studies reporting on the potential reaction mechanism for in situ imine formation followed by reduction and identification of new enzymes with broad tolerance for carbonyl substrates and amine nucleophiles27–30. More recently, an IRED enzyme was successfully engineered and was applied to greater than 100 g scale for the reductive amination of an aldehyde to resolve an amine26, further underscoring the potential value of RedAm technology. A versatile biocatalytic reductive amination could revolutionize the synthesis of chiral alkylamines, by providing a direct, selective, safe, green and sustainable alternative to traditional methods.

Abrocitinib belongs to a group of Janus kinase (JAK) inhibitors and is in late stage development as a JAK1 inhibitor for the treatment of atopic dermatitis31. Abrocitinib is structurally related to tofacitinib (US Food and Drug Administration approved for the treatment of rheumatoid arthritis) and contains a unique N-methylene-substituted cis-cyclobutane headpiece (Fig. 1). In early synthetic approaches to abrocitinib, the amine was synthesized by a chemical reductive amination at low temperatures, providing an approximately 80:20 mixture of cis/trans isomers. The desired cis isomer 1 could be obtained after multiple crystallizations to purge the undesired diastereomer. While effective for synthesis of smaller amounts of abrocitinib to support clinical studies, a more efficient synthetic route, including access to the key cis-cyclobutyl-N-methylamine 1 is required for commercial scale manufacturing. We envisioned the application of RedAm technology to address the structural complexity of the N-methylene headpiece 1 to support the commercial manufacturing of abrocitinib.
Enzymatic reductive amination reactions are reported to be highly stereoselective in creating new chiral centres\(^ {30} \) and in some cases are able to control the stereoselectivity at adjacent chiral centres\(^ {31} \). However, this study entails introduction of cis-selectivity during enzymatic reductive amination, which requires controlling the configuration of remotely lying substituents during reduction of an imine intermediate. Also, four-membered cyclic ketones are reported to have much lower activity compared to five-membered and six-membered cyclic ketones\(^ {32} \). Therefore this study was a significant undertaking, and in the absence of any prior reports on the scalability of RedAm activity at the outset of this work, was an even larger challenge under an accelerated development timeline driven by the ‘Breakthrough Therapy Designation’ from the US Food and Drug Administration for abrocitinib.

In this Article, we report on the successful transition of RedAm technology from the initial identification of enzymatic activity to kilogram scale production and eventually to commercial manufacturing for the synthesis of key intermediate 1 for the new drug, abrocitinib. Our production of the intermediate cis-cyclobutyl-N-methylamine (1) is a successful demonstration of RedAm technology for the commercial scale synthesis of substituted amines by the reductive amination of a ketone with methylamine.

**Results**

**Route selection and identification of RedAm enzyme.** Reductive amination of ketones with alkylamines remains a challenge. Several drawbacks include incomplete reaction, over-reaction to give bis-alkylated by-products and low selectivity leading to challenges in isolation and loss of yield. Transaminases have proven to efficiently transfer an ammonia equivalent, but require the subsequent alkylation of the primary amine, which often results in over-alkylation and the use of potentially genotoxic alkylating reagents\(^ {33,34} \). A direct reductive amination was identified as the most efficient method to introduce a methylamine moiety to synthesize the N-methylamine-substituted cis-cyclobutane headpiece 1 (Fig. 1).

A screen of the Pfizer in-house enzyme panel, consisting of over 80 wild-type IRED enzymes from various sources, was performed to identify an enzyme capable of performing reductive amination of ketone 2 with methylamine to give the desired N-methylamine 1. Several enzymes were identified with reductive amination activity that resulted in the formation of the desired cis isomer 1. A few enzymes also showed activity for the undesired trans isomer (Supplementary Table 1). The three best enzyme hits from initial screening were re-tested to confirm their performance, resulting in the selection of SpRedAm from *Streptomyces purpurascens*\(^ {35} \) as the best candidate with both reductive aminase activity and high selectivity (diastereometric ratio (d.r.) > 99:1) for the desired cis isomer 1. The reaction with SpRedAm was scaled to 7.5 g using a substrate loading of 20 g l\(^ {-1} \) of 2 and 8 g l\(^ {-1} \) of enzyme as a lysate, to give 27% isolated yield of amine 1 with high selectivity for the desired cis isomer (d.r. > 99:1), confirming the performance seen in initial screening. Further testing of SpRedAm wild-type enzyme under more reasonable process conditions (100 g l\(^ {-1} \) ketone 2 and 1.5 g l\(^ {-1} \) enzyme loading) showed only 0.75% conversion (24 h) to product 1, but retained high selectivity. The significant drop in percentage conversion with increase in substrate ketone 2 loading from 20 to 100 g l\(^ {-1} \), suggested
a low substrate tolerance for wild-type SpRedAm enzyme and it was not suitable for targeted 100 g⁻¹ substrate loading for a commercial process. In addition, the wild-type enzyme SpRedAm showed a narrow range of pH (7–8) and temperature (25–30°C) for optimal activity, which could be challenging to manage in commercial manufacturing facilities.

**Enzyme engineering.** After initial evaluation, over 200-aggregate fold (~127x) in activity and 2x methylamine tolerance, see Table 1) improvement in enzyme performance over the parent wild-type (WT) SpRedAm was needed to enable a commercial manufacturing process. In addition, to ensure robust performance at scale a broader window for operational stability (pH and temperature) was highly desirable. Therefore, we pursued enzyme engineering to improve the performance of the SpRedAm wild-type enzyme, targeting increased substrate tolerance and activity while retaining high selectivity (see the estimated process targets for commercial manufacturing, Table 1).

A multi-pronged approach for enzyme engineering was designed and applied in view of an accelerated development timeline. This included a computational-based and a bioinformatics-based approach for initial site selection and library design, coupled with a data-driven approach to identify hot spots and their synergistic recombination to achieve the desired performance.

Initial library design and site selection were done using structural homology models as there was limited mechanistic and structural information available for this emerging class of enzyme. A total of 93 sites (out of 296) were selected for single site saturation mutagenesis (SSM), covering both binding site and secondary shell residues. Overall, 34 binding site residues, 55 secondary shell residues and 4 additional sites identified by bioinformatics, were selected for SSM library synthesis. SSM libraries from the first round of enzyme engineering were initially screened at a low substrate concentration (20 g l⁻¹ of ketone 2) and were progressively increased in subsequent rounds of screening. This resulted in the rapid identification of over 20 improved variants with single amino acid substitution, representing 12 amino acid residues as hot spots for improved activity (Fig. 2). The most active variants exhibited up to fivefold improvement over the parent (FIOP) and included the most active variants with four to sixfold improvement in percentage conversion over their single mutation parents with retention of high selectivity (Fig. 3a).

Double mutant Q13R/F214I showed a threefold improvement in percentage conversion. However, it resulted in a drop in selectivity and gave a cis/trans isomer ratio of 98:5:1:5. Addition of F214I/N to A170C or F180M, individually, showed a twofold improvement. However, their combination as a triple mutant was non-synergistic and resulted in only a modest increase in activity (Fig. 3a).

Computational analysis showed amino acid residue 214 is part of the active site (chain B), which is formed at the interface between two chains and is adjacent to other binding site residues 180 and 176 (from chain A, Fig. 4). This provided a useful insight for future recombination of hot spots and residue 214 was evaluated by both ISM (ref. 35) and CASTing (ref. 36), using multiple parental templates. Double variants N131H/A170C and A170C/F180M showed the highest performance and were selected to progress further by targeted mutagenesis. Based on the crystal structure of SpRedAm, both the 170 and 180 amino acid residues are present on the same α-helix and residue 180 is part of the binding site. On the contrary, amino acid residue 170 is a secondary shell residue which is 8 Å away from the binding site and is oriented parallel or slightly away from the site. As the number of positive mutations increased, a multi-site random recombination approach was also introduced. Several of the most active double and triple mutation variants were used as parental templates to create diverse variant libraries, in reactions including up to 12 additional mutations.

Screening of several hundred variants from the resultant libraries identified improved variants with four to six mutations. Screening of all the libraries from round 2 was performed at 75 g l⁻¹ and any identified hits were retested under multiple screening conditions (increasing substrate and co-substrate loadings).

In round 3, hits obtained in round 1 were recombined with multiple approaches to generate various combinations. In this round, an additional 20 plates containing various combinations of mutations were screened and analysed. This resulted in identification of multiple variants with an additional fourfold to fivefold improvement in enzyme performance. The key variants from each round were further characterized to calculate kinetic parameters $K_m$ (Michaelis constant), $V_m$ (Michaelis constant) and turnover numbers (TONs) (Fig. 5). The final variant R3-V6 from round 3 with substitutions N131H, A170C, F180M and G217D was selected for reaction and process.

**Table 1 | Process targets for commercial scale manufacturing in comparison to initial performance of wild-type SpRedAm**

| Process parameter                   | Targets for commercial process | Initial performance SpRedAm wild-type | Estimated performance improvement targets |
|-------------------------------------|--------------------------------|---------------------------------------|------------------------------------------|
| Substrate loading (g 1⁻¹)           | 100                           | 100                                   |                                          |
| Enzyme loading (g 1⁻¹)              | 1.5                           | 1.5                                   |                                          |
| % Conversion (24 h)                 | >95                           | 0.75                                  | 127x                                     |
| Selectivity (cis/trans)             | >99:1                         | >99:1                                 |                                          |
| Methylamine tolerance (mM)          | 1,000                         | 500                                   | 2x                                       |
| Background activity (ketone reduction) | <0.5%                        | 1.5%                                  |                                          |

F180M and F214N retained their improved activity (fourfold to fivefold over WT), suggesting increased tolerance for higher substrate and methylamine concentrations. Variants A170C/M, F180M and F214N showed the highest increase in activity and substrate tolerance, resulting in an aggregated improvement of approximately 20-fold.

The top five variants from round 1 were tested at multi-gram scale at 50 g 1⁻¹ loading of 2 with 8 g 1⁻¹ enzyme lysate and gave >75% conversion over 96 h with >99:1 cis/trans isomers. This represented a significant improvement in enzyme performance for substrate tolerance, given the final substrate loading target of 100 g 1⁻¹. In parallel, we also started to evaluate various process parameters to provide insights into the design of the next set of screening conditions. Computational analysis of the top hot spots from round 1 of enzyme engineering showed that most were associated with either the enzyme active site or cofactor nicotinamide adenine dinucleotide phosphate reduced (NADPH) binding pocket (Fig. 2b,c).

In round 2 of enzyme engineering, recombination of beneficial mutations was performed and they were screened under more stringent conditions at 75 g 1⁻¹ (480 mM) of substrate loading using 2 equiv. of methylamine, close to the desired process targets for a scalable process (Fig. 3b). Initially, targeted recombination was performed to produce double and triple mutation variants. Several double mutants (Q13R/A170M, Q13R/N131H and A170C/F180M) showed up to threefold improvement in percentage conversion over their single mutation parents with retention of high selectivity (Fig. 3a). Double mutant Q13R/F214I showed a threefold improvement in percentage conversion. However, it resulted in a drop in selectivity and gave a cis/trans isomer ratio of 98:5:1:5. Addition of F214I/N to A170C or F180M, individually, showed a twofold improvement. However, their combination as a triple mutant was non-synergistic and resulted in only a modest increase in activity (Fig. 3a).
Out of four mutations in SpRedAm-R3-V6, only F180M is part of the enzyme active site, directly interacting with the substrate. The other three mutations N131H, A170C and G217D are part of the secondary shell and have close interactions with active site residues. Also, the mutations A170C and F180M are part of the α-helix which undergoes a change in conformation on substrate binding26.

**Process development.** Once the enzyme variant SpRedAm-R3-V6 from round 3 was identified as having improved performance, we focused on process design and development for commercial manufacturing. Multiple cofactor recycling systems were screened and the glucose/glucose dehydrogenase (GDH) cofactor recycling system was selected for process development, due to higher catalytic efficiency and clean reaction profile without any side products from reduction of 2 to the corresponding alcohol. Cofactor recycling systems Lactobacillus brevis alcohol dehydrogenase (ADH)/isopropanol and formate dehydrogenase/amino formate were also evaluated, but both resulted in multiple side products. In addition, the effects of reaction temperature, pH, substrate loading, enzyme
loading, glucose loading, cofactor NADP+ loading and methylamine concentration were screened to identify optimum reaction conditions to achieve high conversion and selectivity. The engineered enzyme showed high selectivity and specificity for the desired amine 1, controlling the formation of the only observed by-product, the trans isomer of 1, to <0.5%. An increase in temperature and pH resulted in higher conversion. However, reaction temperature above 30 °C and maintaining pH > 8 resulted in an increased rate of hydrolysis of the isopropyl ester of both ketone 2 and product amine 1. Hydrolysis was controlled by maintaining the reaction temperature between 20 and 30 °C and the pH between 6 and 8. Enzyme loading studies showed an increase in percentage conversion to the product with higher enzyme loadings and the loading was optimized to 1.5 wt% to enable efficient enzyme removal downstream for product isolation. The optimized laboratory process was further refined and engineered to fit the scale and manufacturing equipment train and was successfully scaled from the laboratory (gram scale) to kilolaboratory (1–10 kg of 2 per batch), pilot plant (50–100 kg of 2 per batch) and finally commercial manufacturing plant (>200 kg of 2 per batch). Consistent reaction performance of >91% conversion in 48 h was observed irrespective of scale (Fig. 6a).

Under optimized process conditions, the engineered enzyme SpRedAm-R3-V6 showed on average ~77% conversion (versus 0.75% with WT) after 24 h (see Fig. 6a), resulting in a 103× improvement in enzyme performance. In addition, it showed 2× improvement in tolerance for methylamine since the final process was run with >1 M of methylamine (versus 0.5 M for wild-type SpRedAm), leading to an overall improvement of 206-fold (103×2) over wild type. The product amine 1 was isolated by removal of the enzyme and extraction at pH >11.5 using methyl tert-butyl ether (MTBE) and then crystallized as the succinate salt. A total of >3.5 MT of amine 1 as the succinate salt was manufactured in >99% purity and >99:1 cis:trans selectivity, representing the successful implementation of RedAm technology on a commercial manufacturing scale (Fig. 6b).

Conclusions
RedAm technology was successfully applied for the commercial scale manufacturing of a secondary amine via direct reductive amination of a ketone with methylamine. This was accomplished by discovering an IRED enzyme with the desired RedAm activity coupled with state-of-the-art enzyme engineering and high-throughput

---

**Fig. 4** | Active-site view of the cofactor NADPH-bound SpRedAm-R3-V6. **a**, All the mutations in the final variant are shown with the product amine docked into the site. **b**, The key active site residues (residues highlighted in brown are from chain A and residues highlighted in cyan are from chain B) showing interaction with the product are shown.

**Fig. 5** | Performance of best variants from different rounds of enzyme engineering. **a**, Kinetic characterization of advanced variants from each round. **b**, Performance improvement with each round of enzyme engineering for key variants from each round with fold improvement in TONs over wild-type SpRedAm.

---

| Variant #     | TON  | K<sub>m</sub> (mM) | K<sub>cat</sub> (s<sup>-1</sup>) | Specific activity (mM g<sup>-1</sup>) |
|---------------|------|-------------------|--------------------------|----------------------------------|
| SpRedAm-WT    | 288  | 2.5               | 2.68                     | 2.64                             |
| SpRedAm-R1-V1 | 7,904| 5.58              | 14.58                    | 10.96                            |
| SpRedAm-R1-V2 | 9,954| 14.22             | 28.02                    | 15.37                            |
| SpRedAm-R2-V3 | 17,308| 14.97             | 17.92                    | 8.65                             |
| SpRedAm-R2-V4 | 21,154| 8.94              | 15.97                    | 8.33                             |
| SpRedAm-R3-V5 | 33,000| 8.46              | 22.82                    | 16.53                            |
| SpRedAm-R3-V6 | 36,538| 13.80             | 51.22                    | 29.42                            |
screening to generate an engineered enzyme SpRedAm-R3-V6 to provide an overall performance improvement of >200-fold over the wild type. The engineered enzyme SpRedAm-R3-V6 was successfully implemented in the commercial process to give a space–time yield of 60 g per litre per day with high purity (>99.5%) and selectivity (>99:1 cis:trans) to access the amine 1 required to synthesize abrocurtinib. The successful transition of initial laboratory activity to commercial industrial scale manufacturing was demonstrated under significantly accelerated timelines, highlighting the potential of rapid development of enzyme catalysis as a competitive green alternative to traditional chemical methods. This work contributes to the rapidly developing field of enzyme catalysis that is emerging as a critical strategy for pharmaceutical and fine chemical manufacturing as a sustainable alternative to existing methods.

Methods

All reagents and solvents used in this study were purchased from commercial suppliers and were used as received, unless specified. NADP+ was purchased from Oriental Yeast Co., GDH (CDX 901) enzyme was purchased from Codexis Inc. Engineered enzyme SpRedAm-R3-V6 at large scale was custom produced by commercial enzyme producers.

Identification and cloning of enzymes in screening panel. The Pfizer IRED screening panel included various wild-type IRED enzymes from multiple sources4,5,18,19,25,27. Enzyme identification and cloning of various wild-type enzymes included in the Pfizer IRED screening panel was previously published18,25. Identification and cloning of enzymes in screening panel. The Pfizer IRED screening panel included various wild-type IRED enzymes from multiple sources4,5,18,19,25,27. Enzyme identification and cloning of various wild-type enzymes included in the Pfizer IRED screening panel was previously published18,25.

Expression and lysis of enzymes for screening panel. Plasmids containing IRED genes were transformed into chemically competent BL21 Gold (DE3) cells (Agilent 2301132) following standard protocols, plated on LB + kanamycin agar plates and incubated at 37 °C overnight. Next, 800 μl LB + kanamycin+ seed cultures were inoculated with single colonies from transformation plates in 96-well deep-well plates (VWR P9636) and incubated at 37 °C for 20 h with shaking (210 rpm, 2 °C). Frozen cell pellets were thawed on ice and then fully suspended at 120 mg ml−1 in Bug Buster Master Mix (Millipore Sigma 71491) expression cultures were inoculated with 80 μl overnight seed cultures in fresh deep-well plates and incubated for 24 h at 32 °C with shaking (1,000 rpm, 3 mm orbit). Cells were pelleted by centrifugation at 4,500g for 15 min and frozen at −80 °C.

Bioinformatics. Template identification and homology model construction. At the start of this engineering project no experimentally determined structures for SpRedAm were available. The Chemical Computing Group’s Molecular Operating Environment (MOE) suite of tools was used to calculate amino acid conservation rates at each position along the protein backbone using an alignment file containing the closest 250 non-redundant sequences identified through a BLAST search. A template search for homology modelling was performed in MOE as well as a BLAST search of Protein Data Bank proteins. The selected template, ZHIB from Streptomyces kanamyceticus, resulted in an alignment with 52% identity to SpRedAm and a pairwise percentage positive (BLSM62) of 61%. A homology model was constructed using MOE. For this model, the conserved cofactor binding motif GxGxxG was constrained between the template and SpRedAm sequences, as were the highly conserved active site residues corresponding to Asp169 and Thr177.

Identification of positions for site saturation mutagenesis. Binding site residues in the homology model were identified using the Site Finder function in MOE. Secondary shell residues were identified by selecting residues within 4.5 Å of binding site residues. All residues were sorted and scored by calculated conservation rates, variability of amino acids seen in the alignment mentioned previously (number and chemical characteristics) and distance from the binding site. Ninety-three positions were selected for site saturation mutagenesis library creation, including all 34 of the identified binding site residues. The remaining positions comprised non-conserved secondary shell and protein surface residues.

Substrate docking. Once an experimentally determined structure of SpRedAm-R3-V6 was determined, MOE was used to first identify the binding site using the Site Finder function and then the product amine was docked using the General Docking function with an Induced Fit refinement.

Enzyme engineering (library synthesis, assembly and expression). Single site saturation variant libraries (93 residues), consisting of 34 active site residues, 55 secondary shell residues and 4 additional residues were identified by bioinformatics. SpRedAm gene variant libraries were synthesized by Twist Bioscience. Twist Bioscience was supplied with a codon-optimized gene sequence which also included 30 base pairs of 5′ and 3′ flanking DNA sequences to allow cloning of the linear DNA libraries, into an expression plasmid, using Gibson Assembly. The single site saturation variant libraries were delivered/received as lyophilized linear DNA, in a 96-well plate (one targeted position per residue per well). The library pools were rehydrated with 80 μl TE buffer (pH 7.0).

Library assembly. The individual library pools were amplified by polymerase chain reaction (PCR) with Phusion 2x Hot Start DNA polymerase mix (NEB), using flanking primers (IDT) designed to accommodate cloning into an expression vector.

Table 6 | Scale-up performance of engineered enzyme SpRedAm-R3-V6. a, Reaction profile for conversion versus time. b, Isolated yield and purity for succinate salt (reaction was stopped after 60 h) from laboratory up to commercial manufacturing on multi 100s of kg scale.
plasmid by Gibson Assembly. The PCR-amplified variant library DNA was purified using Qiagen PCR purification kit reagents and the DNA was eluted with water. The library pools were cloned into the pET28b vector (EMD Biosciences), which had been linearized with restriction enzymes (NEB) and purified with Qiagen PCR purification kit reagents. Hi-fi DNA assembly reagents (NEB) were used to perform the variant library cloning step.

**Library expression.** Transformation of the pET28 variant libraries was performed using electroporation BL21 DE3 Eschericia coli (Lucigen) by electroporation using GENE PULSER (BioRad) with 1 mm gap cuvettes (BTX). Recovery media Eschericia coli (Lucigen) by electroporation perform the variant library cloning step. The library pools were cloned into the pET28b vector (EMD Biosciences), which had been linearized with restriction enzymes (NEB) and purified with Qiagen.

Screening protocol B (substrate loading 20 mg ml\(^{-1}\)) as a solution in DMSO (500 mg dissolved in 1 ml DMSO) to each μ (128 mM) was added 190 μ (101 mg), glucose (160 mM, 1.25 equiv.) and methylamine (260 mM, −20 min, 4 °C) and the clarified supernatant was transferred to 96-well deep-well plates which were sealed with aluminium film and frozen at −20 °C, prior to screening.

Screening of enzyme libraries for round 1. The SSN libraries were synthesized as a pooled library (one targeted site per residue per well) and 46–92 colonies were picked from each pool of the library for screening to cover for all the possible 19 variants. After the mixture was sterilized for 30 min, carbon D3CO60 powder (66 kg, 20 wt% relative to 2) was charged to the reactor and the suspension was stirred for an additional 30 min. The carbon was filtered through a layer of Celite and rinsed with water (2301, 11 kg, LR). MTBE (4,1001, 201 kg, LR) was charged to the reactor and the mixture was cooled to 5 °C, after which 20 wt% aqueous NaaOH (551 kg, 1 equiv.) was added. The phases were split and the organic layer was collected and concentrated under reduced pressure (500 mbar) to a final volume of 1,150 Fresh MTBE (1,150) was added. In a separate vessel, MTBE (2,3001, 101 kg, LR) and succinic acid (139 kg, 0.90 equiv.) were charged to the reactor at 20 °C. Seed crystals of succinate salt of 1 (2.5 kg, 1 wt% relative to 2) were added, followed by the transfer of the MTBE solution to the succinyl chloride slurry. The resulting slurry was granulated for 1 h at 20 °C. The solids were filtered off, rinsed with MTBE, (1,150, 51 kg, LR) and dried in a vacuum oven at 40 °C to afford the desired amine succinate salt 1 (311 kg, 73% yield) as a white solid.

**H NMR (400 MHz, DMSO-d6), δ 10.66 (br s, 2H), 4.88 (hept, J = 6.2 Hz, 1H), 3.30 (tt, J = 8.8, 7.3 Hz, 1H), 2.82 (tt, J = 9.8, 8.2 Hz, 1H), 2.43–2.34 (m, 2H), 2.31 (d, J = 2.7 Hz, 2H), 2.13–2.00 (m, 2H), 1.18 (d, J = 6.3 Hz, 6H), 1.4 C NMR (100 MHz, DMSO-d6), δ 174.9, 173.1, 67.3, 48.8, 31.1, 30.6, 30.4, 21.5; IR (cast film, cm⁻¹): 2987, 2745, 2502, 1702, 1637, 1597, 1351, 1254, 1191, 1098, 1070, 1017, 875, 801, 755, 689, 585; m.p. 88–89 °C; HRMS (TOF) m/z calculated for C₂H₁₄N₂O₂ ([M + H]⁺): 272.1338, found: 272.1329.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

Additional data supporting the findings reported in this paper are available as Supplementary Information. All other data are available from the authors upon reasonable request.

Received: 15 February 2021; Accepted: 5 August 2021;
Published online: 21 September 2021

**References**

1. Nijger, T. C. E. & El-Shazly, M. Chiral amine synthesis–recent developments and trends for enamide reduction, reductiveamination, and imine reduction. Adv. Synth. Catal. 352, 753–819 (2010).
2. Afanas’ev, O. I., Kuchuk, E., Usanov, D. L. & Chusov, D. Reductive amiation in the synthesis of pharmaceuticals. Chem. Rev. 119, 11857–11911 (2019).
3. Topczewski, J. J., Cabrera, P. J., Saper, N. I. & Sanford, M. S. Palladium-catalysed transannular C–H functionalization of aliphatic amines. Nature 542, 220–224 (2017).
4. Verzil, G. K. et al. Asymmetric synthesis of a key intermediate for tacofentin via a dynamic kinetic resolution-reductive amination protocol. Org. Process Res. Dev. 22, 1817–1822 (2018).
5. Koszelewski, D., Tauber, K., Faber, K. & Kroutil, W. a-Transaminases for the synthesis of non-racemic a-chiral primary amines. Trends Biotechnol. 28, 324–332 (2010).
6. Gomm, A. & O’Reilly, E. Transaminases for chiral amine synthesis. Curr. Opin. Chem. Biol. 43, 106–112 (2018).
7. Hughes, D. L. Biocatalysis in drug development—highlights of the recent patent literature. Org. Process Res. Dev. 22, 1063–1080 (2018).
8. Abrahamsson, M. J., Wong, J. W. & Bommarius, A. S. The evolution of an amine dehydrogenase biocatalyst for the asymmetric production of chiral amides. Adv. Synth. Catal. 355, 1780–1786 (2013).
9. Abrahamsson, M. J., Vázquez-Figueroa, E., Woodall, N. B., Moore, J. C. & Bommarius, A. S. Development of an amine dehydrogenase for synthesis of chiral amine. Angew. Chem. Int. Ed. 51, 3969–3972 (2012).
10. Bornscheuer, U. et al. Engineering the third wave of biocatalysis. *Nature* **485**, 185–194 (2012).

11. Wu, S. et al. Biocatalytic enzymatic synthesis for industrial applications. *Angew. Chem. Int. Ed.* **60**, 88–119 (2021).

12. Bornscheuer, U. T. The fourth wave of biocatalysis is approaching. *Philos. Trans. R. Soc. A* **376**, 20170063 (2018).

13. Devine, P. N. et al. Extending the application of biocatalysis to meet the challenges of drug development. *Nat. Rev. 2*, 409–421 (2018).

14. Midelfort, K. S. et al. Redesigning and characterizing the substrate specificity and activity of *Vibrio fluvialis* aminotransferase for the synthesis of imagabalin. *Protein Eng. Des. Sel.* **26**, 25–33 (2013).

15. Savile, C. K. et al. Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **329**, 303–309 (2010).

16. Chen, H. et al. Engineered imine reductases and methods for the reductive amination of ketone and amine compounds. US patent 9803224B2 (2017).

17. Scheller, P. N. et al. Enzyme toolbox: novel enantioselective amination and reductive amination enzymes. *ChemBioChem* **15**, 2201–2204 (2014).

18. Li, H., Luan, Z. J., Zheng, G. W. & Xu, J. H. J. A. S. Efficient synthesis of chiral indolines using an imine reductase from *Pannibacillus lactis*. *Adv. Synth. Catal.* **357**, 1692–1696 (2015).

19. Wetzl, D. et al. Expanding the imine reductase toolbox by exploring the bacterial protein-sequence space. *ChemBioChem* **16**, 1749–1756 (2015).

20. France, S. P. et al. Biocatalytic routes to enantiomerically enriched dibenz[c,e]azepines. *Angew. Chem. Int. Ed.* **56**, 15589–15593 (2017).

21. Lenz, M., Borlinghaus, N., Weimann, L. & Nestl, B. M. Recent advances in imine reductase-catalyzed reactions. *World J. Microbiol. Biotechnol.* **33**, 199 (2017).

22. Mangas-Sanchez, J. et al. Imine reductases (IREDs). *Curr. Opin. Chem. Biol.* **19**, 267–273 (2017).

23. Wetzl, D. et al. Asymmetric reductive amination of ketones catalyzed by imine reductases. *ChemCatChem* **8**, 2023–2026 (2016).

24. Aleku, G. A. et al. A reductive amination from *Aspergillus oryzae*. *Nat. Chem.* **9**, 961–969 (2017).

25. 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).

26. Sharma, M. et al. A mechanism for reductive amination catalyzed by fungal reductive aminases. *ACS Catal.* **8**, 11534–11541 (2018).

27. Roiban, G. D. et al. Efficient biocatalytic reductive amination by extending the imine reductase toolbox. *ChemCatChem* **9**, 4475–4479 (2017).

28. Schober, M. et al. Chiral synthesis of LSD1 inhibitor GSK2879552 enabled by directed evolution of an imine reductase. *Nat. Catal.* **2**, 909–915 (2019).

29. Vazquez, M. L. et al. Identification of N-[cis-3-[methyl(7H-pyrrolo[2,3-d]pyrimidin-4-yl)amino]cyclobutyl]propane-1-sulfonamide (PF-04965842): a selective JAK1 inhibitor from Streptomyces tsukubensis. *Chembiochem*. **18**, 2022–2027 (2017).

30. Marshall, J. R. et al. Screening and characterization of a diverse panel of metagenomic imine reductases for biocatalytic reductive amination. *Nat. Chem.* **13**, 140–140 (2021).

31. Matzel, P. et al. Synthesis of β-chiral amines by dynamic kinetic resolution of α-branched aldehydes applying imine reductases. *ChemCatChem* **11**, 4281–4285 (2019).

32. Matzel, P. et al. Photometric characterization of the reductive amination scope of the imine reductases from *Streptomyces tsukubensis* and *Streptomyces ipomoaeus*. *ChemBioChem* **18**, 2022–2027 (2017).

33. Xia, G. et al. Novel process for preparing Pramipexole and its optical isometric mixture by reduction with sodium triacetoxyborohydride. *WO patent WO2006070349 (2006).*

34. Patil, P., Pansare, P., Jagtap, A. & Krishnamurthy, D. An improved process for the preparation of pramipexole dihydrochloride monohydrate. *WO patent WO 2015155704A1 (2015).*

35. Reetz, M. T. & Carballaie, J. D. Iterative saturation mutagenesis (ISM) for rapid directed evolution of functional enzymes. *Nat. Protoc.* **2**, 891 (2007).

36. Bonmarius, A. S., Blum, J. K. & Abrahamson, M. J. Status of protein engineering for biocatalysts: how to design an industrially useful biocatalyst. *Curr. Opin. Chem. Biol.* **15**, 194–200 (2011).

37. Lu, X., Hora, B., Cai, F. & Gao, F. Generation of random mutant libraries with multiple primers in a single reaction. *J. Virol. Methods* **167**, 146–151 (2010).

**Acknowledgements**

We would like to thank Pfizer leadership L. Handanyan, S. Caron, N. Thomson, J. Nelson and C. McWilliams for their support for the work and publication. We would also like to thank R. Lewis and S. France for proof-reading the manuscript.

**Author contributions**

R.K., M.J.K., J.S., M.P.B. and C.A.M. contributed to enzyme engineering and enzyme manufacturing scale reactions. C.A.L., K.M.D., R.P. and J.W. provided analysis, sourcing and C. McWilliams for their support for the work and publication. We would also like to thank R. Lewis and S. France for proof-reading the manuscript.

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41929-021-00671-5.

**Correspondence and requests for materials** should be addressed to Rajesh Kumar.

**Peer review information** *Nature Catalysis* thanks John Woodley and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

**Reprints and permissions information** is available at www.nature.com/reprints.

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2021
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

- [X] The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- [X] An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- [ ] Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- [ ] A description of all covariates tested
- [X] A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- [X] A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- [X] For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- [X] For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- [X] For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- [ ] Clearly defined error bars
- [X] State explicitly what error bars represent (e.g. SD, SE, CI)

Software and code

Policy information about availability of computer code

Data collection

MOE software from Chemical Computing group

Data analysis

GraphPad Prism 8.4.2, GraphPad Software Inc and Microsoft Excel

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Additional data is available in supplementary information and all other data is available from authors upon reasonable request.
Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

[X] Life sciences  [ ] Behavioural & social sciences  [ ] Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Sample size         | Not relevant to the work reported |
|---------------------|-----------------------------------|
| Data exclusions     | No data relevant to the work described in the manuscript are excluded |
| Replication         | The experimental findings were reliably reproduced. |
| Randomization       | Not relevant to the work reported |
| Blinding            | Not relevant to the work reported |

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Study description | Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study). |
|-------------------|----------------------------------------------------------------------------------------------------------------------------------|
| Research sample   | State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source. |
| Sampling strategy | Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed. |
| Data collection   | Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection. |
| Timing            | Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort. |
| Data exclusions   | If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established. |
| Non-participation | State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation. |
| Randomization     | If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled. |

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Study description | Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates. |
Research sample
Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.

Sampling strategy
Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

Data collection
Describe the data collection procedure, including who recorded the data and how.

Timing and spatial scale
Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken.

Data exclusions
If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

Reproducibility
Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.

Randomization
Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.

Blinding
Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Did the study involve field work? Yes No

Field work, collection and transport

Field conditions
Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).

Location
State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).

Access and import/export
Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).

Disturbance
Describe any disturbance caused by the study and how it was minimized.

Reporting for specific materials, systems and methods

Materials & experimental systems
| n/a | Involved in the study |
|-----|-----------------------|
| X   | Unique biological materials |
| X   | Antibodies |
| X   | Eukaryotic cell lines |
| X   | Palaeontology |
| X   | Animals and other organisms |
| X   | Human research participants |

Methods
| n/a | Involved in the study |
|-----|-----------------------|
| X   | ChIP-seq |
| X   | Flow cytometry |
| X   | MRI-based neuroimaging |

Unique biological materials
Policy information about availability of materials

Obtaining unique materials
Describe any restrictions on the availability of unique materials OR confirm that all unique materials used are readily available from the authors or from standard commercial sources (and specify these sources).

Antibodies

Antibodies used
Describe all antibodies used in the study; as applicable, provide supplier name, catalog number, clone name, and lot number.
Validation

Describe the validation of each primary antibody for the species and application, noting any validation statements on the manufacturer’s website, relevant citations, antibody profiles in online databases, or data provided in the manuscript.

Eukaryotic cell lines

Policy information about: **cell lines**

**Cell line source(s)**
State the source of each cell line used.

**Authentication**
Describe the authentication procedures for each cell line used OR declare that none of the cell lines used were authenticated.

**Mycoplasma contamination**
Confirm that all cell lines tested negative for mycoplasma contamination OR describe the results of the testing for mycoplasma contamination OR declare that the cell lines were not tested for mycoplasma contamination.

**Commonly misidentified lines**
(See **ERAC** register)
Name any commonly misidentified cell lines used in the study and provide a rationale for their use.

Palaeontology

**Specimen provenance**
Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information).

**Specimen deposition**
Indicate where the specimens have been deposited to permit free access by other researchers.

**Dating methods**
If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.

Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Animals and other organisms

Policy information about: **studies involving animals**: ARRIVE guidelines recommended for reporting animal research

**Laboratory animals**
For laboratory animals, report species, strain, sex and age OR state that the study did not involve laboratory animals.

**Wild animals**
Provide details on animals observed in or captured in the field, report species, sex and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.

**Field-collected samples**
For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.

Human research participants

Policy information about: **studies involving human research participants**

**Population characteristics**
Describe the covariate-relevant population characteristics of the human research participants (e.g. age, gender, genotypic information, past and current diagnosis and treatment categories). If you filled out the behavioural & social sciences study design questions and have nothing to add here, write “See above.”

**Recruitment**
Describe how participants were recruited. Outline any potential self-selection bias or other biases that may be present and how these are likely to impact results.

ChIP-seq

**Data deposition**

- Confirm that both raw and final processed data have been deposited in a public database such as **GEO**.
- Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

**Data access links**
May remain private before publication.

For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.

**Files in database submission**

Provide a list of all files available in the database submission.

**Genome browser session**
(see **IUCDS**)

Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.
Methodology

Replicates
Describe the experimental replicates, specifying number, type and replicate agreement.

Sequencing depth
Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.

Antibodies
Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.

Peak calling parameters
Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.

Data quality
Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.

Software
Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.

Flow Cytometry

Plots
Confirm that:

- The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- All plots are contour plots with outliers or pseudocolor plots.
- A numerical value for number of cells or percentage [with statistics] is provided.

Methodology

Sample preparation
Describe the sample preparation, detailing the biological source of the cells and any tissue processing steps used.

Instrument
Identify the instrument used for data collection, specifying make and model number.

Software
Describe the software used to collect and analyze the flow cytometry data. For custom code that has been deposited into a community repository, provide accession details.

Cell population abundance
Describe the abundance of the relevant cell populations within post-sort fractions, providing details on the purity of the samples and how it was determined.

Gating strategy
Describe the gating strategy used for all relevant experiments, specifying the preliminary FSC/SSC gates of the starting cell population, indicating where boundaries between "positive" and "negative" staining cell populations are defined.

☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type
Indicate task or resting state; event-related or block design.

Design specifications
Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.

Behavioral performance measures
State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).

Acquisition

Imaging type(s)
Specify: functional, structural, diffusion, perfusion.

Field strength
Specify in Tesla

Sequence & imaging parameters
Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.

Area of acquisition
State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.
**Preprocessing**

| Preprocessing software | Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.). |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Normalization           | If data were normalized/standardized, describe the approach(es); specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization. |
| Normalization template   | Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MN1305, ICBM152) OR indicate that the data were not normalized. |
| Noise and artifact removal | Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration). |
| Volume censoring         | Define your software and/or method and criteria for volume censoring, and state the extent of such censoring. |

**Statistical modeling & inference**

| Model type and settings | Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation). |
|-------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Effect(s) tested        | Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used. |
| Specify type of analysis: | Whole brain | ROI-based | Both |
| Statistic type for inference | (See [Edlund et al., 2016](#)) | Specify voxel wise or cluster wise and report all relevant parameters for cluster wise methods. |
| Correction               | Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo). |

**Models & analysis**

| n/a | Involved in the study |
|-----|-----------------------|
|     | Functional and/or effective connectivity |
|     | Graph analysis |
|     | Multivariate modeling or predictive analysis |

**Functional and/or effective connectivity**

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

**Graph analysis**

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

**Multivariate modeling and predictive analysis**

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.