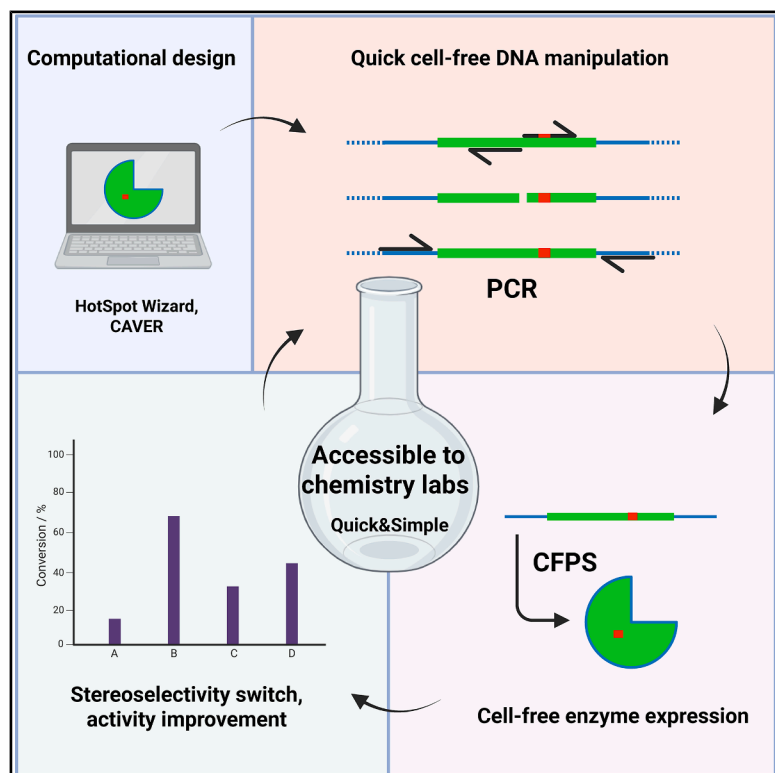


# Accessible biocatalyst development by rapid *in vitro* semi-rational engineering (RISE) of enzymes

## Graphical abstract



## Authors

András Telek, Viktor Zsolt Gaják, Gergő Dargó, Jan Mičan, David Bednár, Beáta G. Vértessy, Gábor Tasnádi

## Correspondence

vertessy.beata@ttk.hu (B.G.V.), gabor.tasnadi@servier.com (G.T.)

## In brief

Biotechnology; Enzyme engineering; Biocatalysis

## Highlights

- Innovative cell-free DNA manipulation enables enzyme engineering in chemistry labs
- Fast and accessible biocatalyst development is demonstrated on a ketimine reductase
- Focused engineering results in activity enhancement and stereoselectivity switch
- Engineered variants are used in gram-scale synthesis of drug intermediates



## Article

# Accessible biocatalyst development by rapid *in vitro* semi-rational engineering (RISE) of enzymes

András Telek,<sup>1,2</sup> Viktor Zsolt Gaják,<sup>2,3</sup> Gergő Dargó,<sup>2</sup> Jan Mičan,<sup>4,5</sup> David Bednář,<sup>4,5</sup> Beáta G. Vértessy,<sup>1,6,7,\*</sup> and Gábor Tasnádi<sup>2,\*</sup>

<sup>1</sup>Department of Applied Biotechnology and Food Science, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, Budapest 1111, Hungary

<sup>2</sup>Servier Research Institute of Medicinal Chemistry, Budapest 1037, Hungary

<sup>3</sup>Eötvös Loránd University, Institute of Chemistry, Budapest 1117, Hungary

<sup>4</sup>Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Brno 625 00, Czech Republic

<sup>5</sup>International Clinical Research Center, St. Anne's University Hospital, Brno 625 00, Czech Republic

<sup>6</sup>Institute of Molecular Life Sciences, Research Centre for Natural Sciences, HUN-REN, Budapest 1117, Hungary

<sup>7</sup>Lead contact

\*Correspondence: [vertessy.beata@ttk.hu](mailto:vertessy.beata@ttk.hu) (B.G.V.), [gabor.tasnadi@servier.com](mailto:gabor.tasnadi@servier.com) (G.T.)

<https://doi.org/10.1016/j.isci.2025.114257>

## SUMMARY

Tailoring natural enzymes to synthetic needs is often associated with high costs and long timelines, hindering the broader adoption of biocatalysis in the chemical and pharmaceutical industries. To address this, we developed the RISE (rapid *in vitro* semi-rational engineering) workflow that makes enzyme engineering accessible to chemistry laboratories. RISE integrates four key concepts: computational design of focused variant libraries, rapid generation of linear mutant DNA libraries via PCR, cell-free protein synthesis from linear template DNA, and iterative cycles of mutagenesis, expression, and testing to accumulate beneficial mutations. In a proof-of-concept study, we engineered a ketimine reductase from *Rattus norvegicus* (*RnKIRE*D), achieving stereoselectivity inversion in one reductive amination reaction and a 400-fold activity improvement in another. These engineered variants enabled the gram-scale synthesis of key intermediates for ACE2 inhibitor drugs. RISE bridges the gap between inefficient wild-type enzymes and expensive directed evolution, promoting biocatalysis implementation in early chemical development.

## INTRODUCTION

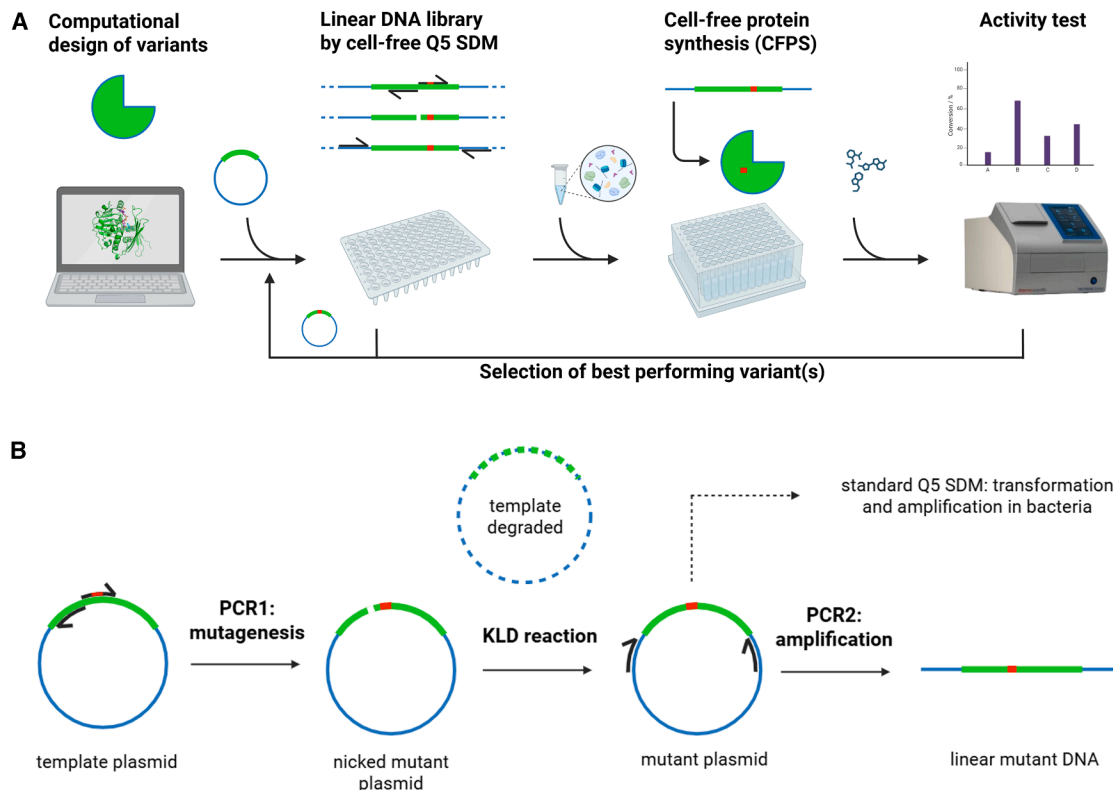
Biocatalysis is often pictured as a mature technology that has reached industrial viability through advances in molecular biology techniques, bioinformatic tools, and directed protein evolution.<sup>1–3</sup> This is well exemplified by numerous industrial enzymatic processes reported in the literature or filed as patents.<sup>4,5</sup> However, developing an industrial biocatalyst still requires significant time and financial investment, which limits the adoption of biocatalysis by a broader chemistry community.<sup>6,7</sup> As an example, in the early chemical development of active pharmaceutical ingredients a biocatalytic route can prove inferior to chemical methods due to the low activity or insufficient stability of the available wild-type enzymes.<sup>8,9</sup> Directed evolution could be used to improve those properties, but its high cost and time demands are often not justified at that stage causing the enzymatic solution to be discarded. As an alternative, focused rational or semi-rational engineering of enzymes can also provide improved variants, but the laborious DNA manipulation and protein expression involved in this process are often still prohibitive for non-specialized laboratories. To make this option more accessible, we developed RISE (rapid *in vitro* semi-rational

engineering), a simple workflow enabling quick development of biocatalysts.

Central to RISE is cell-free protein synthesis (CFPS) that simplifies the expression of enzyme variants by avoiding DNA transformation, culture growth, and cell lysis to obtain the expressed protein.<sup>10,11</sup> *E. coli* extract based *in vitro* transcription-translation systems are available from multiple vendors or can be produced in-house as well.<sup>12</sup> Such systems are increasingly used in ultra-high-throughput microfluidic screenings<sup>13,14</sup> but there are also several reports of medium-throughput microtiter plate-based applications.<sup>15,16</sup> A major advantage of CFPS is its ability to express proteins from linear template DNA. This opens the possibility to directly use PCR products for expression instead of plasmids that are much more laborious to prepare.

Semi-rational enzyme engineering<sup>17,18</sup> involves the creation of focused mutant libraries of the parent enzyme by site-directed mutagenesis (SDM). Well-established PCR-based methods like QuikChange, Gibson-assembly, or Q5 SDM are used to create plasmids containing the prespecified mutation. To date, there have only been scattered reports of methods creating linear template DNA for single-point mutant proteins. RAPPER<sup>19</sup> involves a primer overlap extension PCR strategy while DiRect<sup>20</sup> uses both





**Figure 1. The concept of RISE**

(A) General workflow of RISE.

(B) Workflow of the cell-free Q5 SDM protocol.

overlap extension and subsequent nested PCR reactions. Both methods suffer from contamination from the original template DNA, which is mitigated by either the use of biotinylated primers or multiple amplification steps. In RISE, we use a self-developed cell-free Q5 SDM protocol that involves a second PCR step instead of transforming the KLD (kinase, ligase, DpnI) reaction into bacteria. This step results in quick and cell-free amplification of the mutated coding sequence without contamination of the wild-type sequence which is digested by DpnI (Figure 1B). Very recently, Landwehr et al. followed a similar strategy to create a single-point mutant linear DNA library using the Gibson assembly.<sup>21</sup>

A key step in every semi-rational enzyme engineering effort is the design of the mutations to test. Computational methods have proven efficient in selecting mutational hotspots for engineering.<sup>22</sup> Several tools are available to predict activity- or stability-enhancing mutations.<sup>23</sup> Since epistatic effects arising from the combination of mutations are still hard to predict, often experimental strategies need to be used to accumulate beneficial mutations. In the context of low- or medium-throughput enzyme engineering, FRISM<sup>24</sup> (focused rational iterative site-specific mutagenesis) has been a widely used strategy. In FRISM, a reduced alphabet of amino acids is introduced in selected positions and the best-performing variant is used as a starting point for iteration.

Putting these considerations together, we have defined the workflow of RISE (Figure 1A) as follows. First, a focused mutant

library is designed using computational tools. Then, a linear DNA library encoding for each designed variant is produced by a cell-free Q5 SDM protocol and the variants are expressed by CFPS. The activity of variants is tested in small-scale biocatalytic reactions and the best-performing variants are selected and used as templates for an iterative mutagenesis, expression, and testing cycle.

## RESULTS

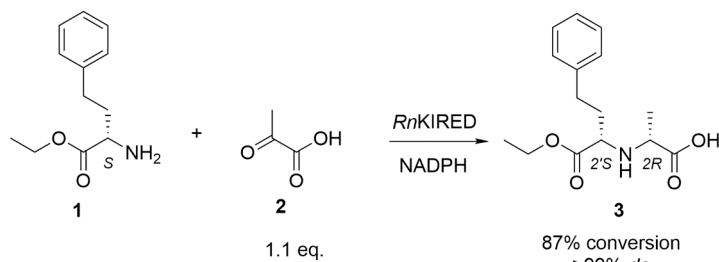
### Validation of the cell-free Q5 SDM protocol

In a proof-of-concept study, first, the cell-free Q5 SDM protocol (Figure 1B) had to be validated since all other steps are well-documented in the literature. We have selected two unrelated enzymes to perform the validation: a ketimine reductase from *Rattus norvegicus* (*RnKIRE*)<sup>25,26</sup> and a 2'-deoxyribosyltransferase from *Lactobacillus leichmannii* (*LINDT*).<sup>27,28</sup> HotSpot Wizard<sup>29</sup> webserver and CAVER analyst<sup>30,31</sup> were used on snapshots from molecular dynamics simulations to select positions for mutagenesis based on their predicted mutability and location within the active site and access tunnels. The process identified ten positions in *RnKIRE* and 15 in *LINDT* that were selected for mutation. Six amino acid changes were proposed in each, considering their size, predicted stability, and chemical diversity (Figures S1 and S2; Tables S1–S4). This smart library design resulted in a 60-member variant library for *RnKIRE* and a 90-member variant library for *LINDT*. The primers for each

### Reaction 1

Ref. 26

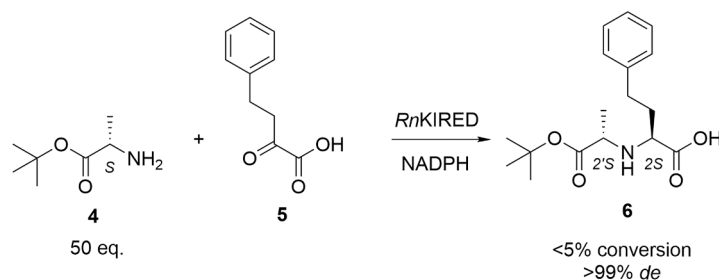
Aim: stereoselectivity  $R \rightarrow S$



### Reaction 2

Ref. 26

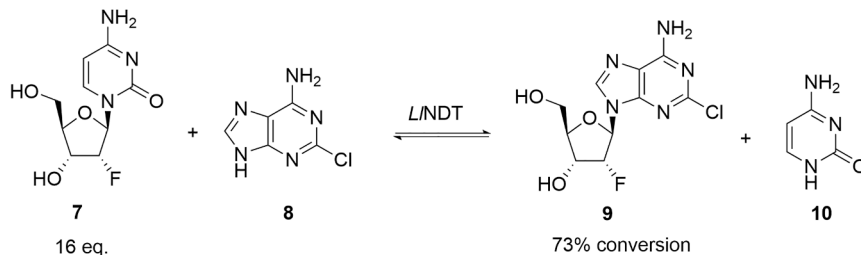
Aim: conversion  $\uparrow$  amine excess  $\downarrow$



### Reaction 3

Ref. 28

Aim: conversion  $\uparrow$  donor excess  $\downarrow$



mutation have been designed with the widely used NEBaseChanger and NEB  $T_m$  calculator tools to allow the use of a single annealing temperature so that the mutagenic PCR can be performed in a single experiment (Table S5). For the second PCR reaction, a general primer pair was designed 100 base-pairs upstream and downstream of the T7 promoter and T7 terminator regions on the plasmid backbone. Then, we could perform the whole cell-free Q5 SDM protocol (Figure 1B) and obtain the linear mutant DNA libraries for both *RnKIREd* and *LINDT* (Figures S3 and S5). Notably, the library preparation process could be completed within 5–8 h. The presence of the designed mutations has been confirmed by sequencing. This demonstrates that our cell-free Q5 SDM protocol can be used to rapidly generate linear single-point mutant DNA libraries.

#### Validation of the RISE workflow

Next, all subsequent steps of RISE have also been performed on these two libraries to complete the validation of all experimental protocols. Using the commercial NEBExpress kit, the expression of all 60 *RnKIREd* and 90 *LINDT* variants has been achieved (Figures S4 and S6).

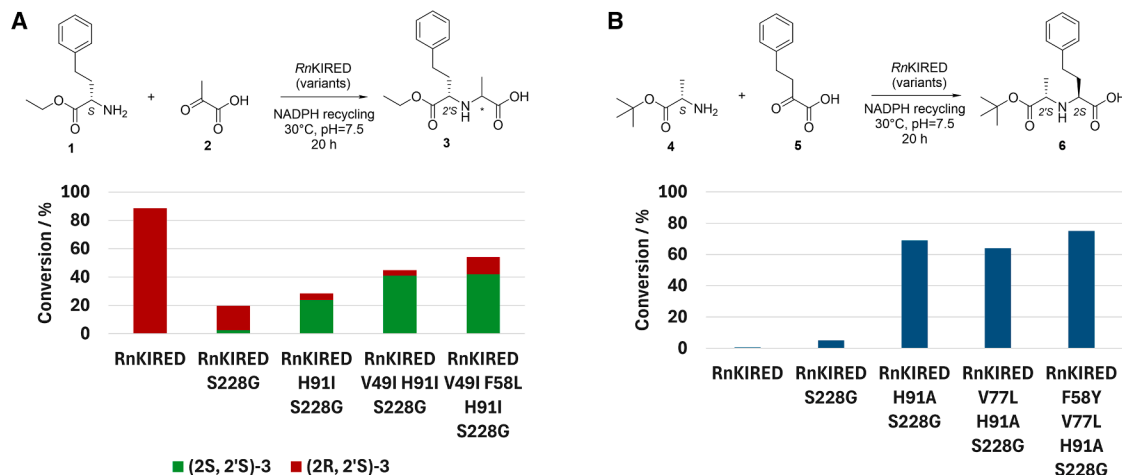
#### Scheme 1. Test reactions and conditions to optimize by enzyme engineering with RISE

To assess the activity of the produced variants, we have selected test reactions where the activity or selectivity of *RnKIREd* and *LINDT* has to be improved to reach synthetic applicability. *RnKIREd* shows high activity and *R*-selectivity in the reaction of L-homophenylalanine ethyl ester (1) and pyruvic acid (2) (Reaction 1, Scheme 1), and *S*-selectivity but low activity in the reaction of L-alanine *tert*-butyl ester (4) and benzylpyruvic acid (5) (Reaction 2, Scheme 1).<sup>26</sup> Since (2*S*,2'*S*)-products are key building blocks of several ACE2 inhibitor drugs, both reactions were selected, targeting modified stereoselectivity in the former and improved activity in the latter. For *LINDT*, 2'-modified nucleoside analogs are pharmaceutically relevant synthetic targets since they are important building blocks of therapeutic oligonucleotides.<sup>32</sup> The wild-type *LINDT* has shown only modest activity on 2'-fluoro derivatives,<sup>28</sup> thus, we selected the sugar transfer reaction from 2'-fluoro-cytidine (7) to 2-chloro-adenine (8) (Reaction 3, Scheme 1), targeting improvement in conversion to nucleoside product 9.

In Reaction 1, we have seen significant increase in (2*S*,2'*S*)-product formation from <1% to 15% for one of 60 single-point variants, *RnKIREd* S228G (Figure S7A). In Reaction 2, we have observed a detectable conversion to the product for five of 60 single-point variants with *RnKIREd* S228G standing out reaching 37% conversion to 6 in two days (Figure S8A). In Reaction 3, only marginally increased conversion was observed for the best four of 90 single-point variants *LINDT* F8A, A10G, Y21R, and V41I (Figure S9). The findings in each case have been validated by performing the reactions with bacterially produced and purified enzymes (Figures 2 and S10). These results confirm that the experimental protocols of RISE are suitable for the rapid production of focused enzyme variant libraries in appropriate amounts for activity screening.

#### Engineering *RnKIREd* with RISE

The contrast between improvements observed in the *RnKIREd* and *LINDT* libraries prompted us to perform the iterative cycles of RISE only on the former. Using the validated protocols, we combined further mutations on *RnKIREd* S228G to further modify stereoselectivity in Reaction 1 and improve activity in Reaction 2. A second mutation has been introduced in eight positions (six amino acids in each), resulting in a 48-member double mutant library.



**Figure 2. Validation of results with purified variants from RISE**

(A) Conversion and product diastereomer distribution achieved with the top hits from four rounds of RISE. Conditions, 25 mM **1**, 75 mM **2**, 0.4 mM NADP<sup>+</sup>, 62.5 mM glucose, 6 U/mL GDH, 100 mM Na-phosphate, pH = 7.5, 20 v/v% DMF, 0.1 mg/mL enzyme.

(B) Conversions achieved with the top hits from four rounds of RISE. Conditions, 100 mM **4**, 50 mM **5**, 0.4 mM NADP<sup>+</sup>, 125 mM glucose, 6 U/mL GDH, 100 mM Na-phosphate, pH = 7.5, 20 v/v% DMF, 0.1 mg/mL enzyme.

Testing this library in Reaction 1 has revealed 13 of 48 mutations increasing the ratio of the (2S,2'S)-**3**, but *RnKIRED* H91I/S228G stood out with inverted stereoselectivity of 35% *de* (2S,2'S)-**3** (Figure S7B). Notably, the activity of both *RnKIRED* S228G and *RnKIRED* H91I/S228G is significantly decreased in Reaction 1 compared to the wild-type *RnKIRED*. Next, we produced a 42-membered triple-mutant library with *RnKIRED* H91I/S228G as the template. From this, we identified two of 42 improved variants, among which *RnKIRED* V49I/H91I/S228G was superior, displaying further increased stereoselectivity of 80% *de* (2S,2'S)-**3** while also improving activity in Reaction 1 (Figure S7C). In a fourth round of mutagenesis, a 36-membered quadruple-mutant library was produced based on *RnKIRED* V49I/H91I/S228G. From that, no variant has shown further improvement in stereoselectivity, the activity could be further improved only at the expense of decreased selectivity toward (2S,2'S)-**3** (Figure S7D).

In Reaction 2, we found 14 of 48 activity increasing mutations in the initial double mutant library, the two best double mutants *RnKIRED* V77G/S228G and *RnKIRED* H91A/S228G displayed over 90% conversion to **6** in one day (Figure S8B). Therefore, we produced a 78-membered triple-mutant library using both *RnKIRED* V77G/S228G and *RnKIRED* H91A/S228G as templates. The high activity of the double mutants in reaction 2 allowed us to lower the excess of **4** used in the screening from 50 to 2 equivalents while increasing the substrate loading as well. The ~20% conversion achieved by the template double mutants under these conditions enabled capturing further increase of activity. With that we have detected 12 of 78 activity-improving mutations and identified *RnKIRED* V77G/H91A/S228G and *RnKIRED* V77L/H91A/S228G as top performers showing over 30% conversion to **6** (Figure S8C). In the fourth round, a 60-membered quadruple-mutant library was produced based on both *RnKIRED* V77G/H91A/S228G and *RnKIRED* V77L/H91A/S228G. From that, we have identified 3 of 60 improved variants, among which *RnKIRED*

F58Y/V77L/H91A/S228G achieved close to 50% conversion to **6** (Figure S8D).

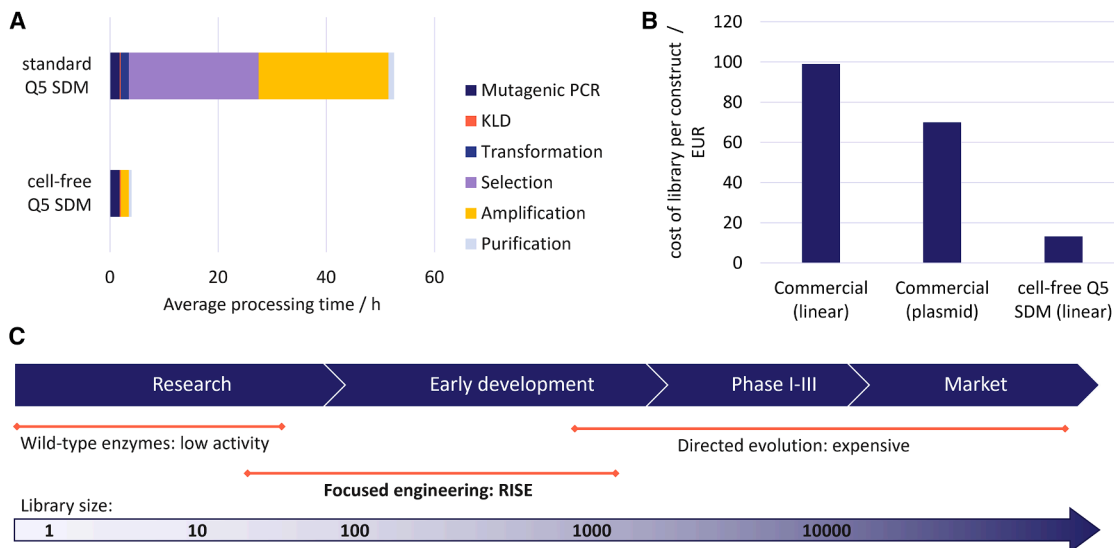
### Synthetic application of *RnKIRED* variants

While performing the iterative mutagenesis rounds, we have also started the synthetic application of the improved variants. In Reaction 1, *RnKIRED* V49I/H91I/S228G enabled the synthesis of (2S,2'S)-**3** with >70% conversion on preparative scale (543 mg, 39% isolated yield, >99% *de*). In Reaction 2, *RnKIRED* H91A/S228G has been used to produce **6** on gram scale with >97% conversion and 63% isolated yield (967 mg, >99% *de*).

### DISCUSSION

Taking these results together, we have validated the whole workflow of RISE by performing consecutive rounds of semi-rational engineering on *RnKIRED* with two separate aims. When we aimed at changing stereoselectivity, we achieved inversion of >99% *R*-selectivity to 80% *S*-selectivity by testing 186 variants in four rounds. The variant *RnKIRED* V49I/H91I/S228G has also sufficient activity to reach synthetic applicability (Figure 2A). When we aimed at increasing activity, we achieved a 400-fold increase in specific activity (Figure S11) and high conversions with low aminoester excess by testing 258 variants in four rounds (Figure 2B). The variant *RnKIRED* H91A/S228G reached >97% conversion in a gram-scale transformation.

The effect of mutations on the stereoselectivity and activity of *RnKIRED* cannot be easily rationalized by structural or mechanistic intuitions. This might be subject to further studies to inform future evolution campaigns. Most importantly, S228G appeared to be a key mutation modifying stereoselectivity and decreasing activity in Reaction 1 while improving activity without changing stereoselectivity in Reaction 2. In the second round, we have observed strong epistatic effects with position 91 in both reactions.



**Figure 3. Advantages of RISE**

(A) Comparison of timelines for DNA manipulation by the standard and cell-free Q5 SDM (for supporting data see [Table S6](#)).

(B) Comparison of the cost of a 60-member single-point mutant DNA library obtained from commercial sourcing and in-house cell-free Q5 SDM (for supporting data see [Table S6](#)).

(C) Proposed role of focused enzyme development by RISE in chemical development in the pharmaceutical industry.

Our results demonstrate that RISE has the potential to enhance the adoption of biocatalysis in a wider chemistry community by lowering both technical and financial barriers for enzyme engineering. On the technical side, the cell-free nature of the workflow comes with two distinct advantages: it significantly reduces the time required for DNA manipulation ([Figure 3A](#)) and it also makes RISE accessible to non-specialized laboratories lacking cell manipulation capabilities. It requires only a PCR machine and basic equipment for DNA and protein analysis, while all reagents used are commercially available. Moreover, the cell-free Q5 SDM protocol also enables direct coupling with large-scale expression of top performing variants by transforming KLD reactions into bacteria. On the financial side, in-house production by cell-free Q5 SDM reduces the cost of focused mutant DNA libraries compared to commercial sourcing ([Figure 3B](#)). Furthermore, for chemistry laboratories relying on external partners for enzyme production, expression of single hits instead of whole libraries also results in cost reduction.

The significance of RISE is also underlined by a recent pioneering work in the field of cell-free enzyme engineering that was developed in parallel to our efforts. The paper of Landwehr et al. describes a workflow similar to our cell-free Q5 SDM for the production of linear mutant DNA libraries. In contrast to RISE, they use automation to generate large libraries to inform machine learning algorithms.<sup>21</sup> Our approach aims to provide an enzyme engineering tool for medium and small-scale experiments accessible to chemistry laboratories due to its simplicity and low technical demand.

In summary, our proof-of-concept study demonstrates that synthetically applicable enzyme variants can be obtained from only a few rounds of laboratory evolution by RISE starting from an unfit wild-type enzyme (unsuitable stereoselectivity or low activity).

While such a focused engineering campaign might not provide an industrial biocatalyst, the improved variants can help bridge the gap between the use of inefficient wild-type enzymes and the ultimate biocatalyst obtained by full-scale directed evolution ([Figure 3C](#)). The simplicity of the protocol and the short turnaround time enable the demonstration of the synthetic utility of an enzyme in early phases of chemical development helping enzymatic processes to progress further toward greener manufacturing routes.

### Limitations of the study

Notably, there are some limitations of RISE. Computational design of mutations requires detailed understanding of the structure and mechanism of the biocatalyst of interest and the focused library will ultimately constrain the achievable fitness improvement. The enzyme should be well-expressed in CFPS, which might exclude proteins having more complex folds, requiring prosthetic groups, metal ions, or cofactors for folding. The full workflow cannot be self-sufficiently performed in chemistry laboratories; bacterial expression is still required for hit validation and synthetic application. The FRISM principle applied for the accumulation of beneficial mutations introduces strict path dependency in the sequence space that might lead to evolutionary dead ends. In the future, a more advanced version of RISE might include data-driven prediction of higher order mutations by machine learning or artificial intelligence tools.

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Beáta G. Vértessy ([vertessy.beata@ttk.hu](mailto:vertessy.beata@ttk.hu)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- All data reported in this paper are available within the paper and the supplemental information files. Data will be shared by the [lead contact](#) upon request.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

### ACKNOWLEDGMENTS

Project no. C1580174 (AT) and C2266727 (VZG) have been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the NVKDP-2021 and KDP-2023 funding scheme, respectively.

This work was supported by the European Commission, under Horizon Europe's Research and Innovation Programme (Bluetools project, grant agreement no 101081957). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

This article is based upon work from COST Action COZYME CA21162, supported by COST (European Cooperation in Science and Technology).

BGV was supported by the National Research, Development and Innovation Fund of Hungary (K135231, K146890, NKP-2018-1.2.1-NKP-2018-00005, 2022-1.2.2-TÉT-IPARI-UZ-2022-00003), the TKP2021-EGA-02 grant, implemented with support provided by the Ministry for Innovation and Technology of Hungary from the National Research, Development and Innovation Fund, and the ICGEB Research Grants Programme 2023 (CRP/HUN23-02).

J.M. is supported by the scholarship Erno Ph.D. Talent. The project was supported by the Ministry of Education, Youth and Sports of the Czech Republic - RECETOX RI (LM2023069), INFRA CZ (90254), and Elixir CZ (LM2023055), European Union's Horizon 2020 Research and Innovation Programme projects CETOCOEN (857560) and CLARA (1011366070), and the Grant Agency of the Czech Republic (25-18233M). The article reflects the author's view, and the Agency and European Commission are not responsible for any use that may be made of the information it contains.

The authors thank the Analytical Department of SRIMC for assisting in the characterization of the compounds.

Graphical figures were created using [BioRender.com](#). Graphical abstract, <https://BioRender.com/nrq6oq0> Figure 1A: <https://BioRender.com/s75hva4> Figure 1B: <https://BioRender.com/vsch3ev>.

### AUTHOR CONTRIBUTIONS

The manuscript was written through contributions of all authors. A.T.: conceptualization, methodology, investigation, validation, and writing – original draft. V.Z.G.: investigation and validation. G.D.: investigation. J.M.: software and formal analysis. D.B.: software and formal analysis. B.G.V.: conceptualization, methodology, and writing – review & editing. G.T.: conceptualization, investigation, writing – review & editing, and supervision.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

### DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this work, the authors used Servier Secure GPT—internal service based on Microsoft's Azure OpenAI service—in order to rewrite the summary section to fit within the required word limit. After using this tool or service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

### STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
- [METHOD DETAILS](#)
  - Computational analysis and library design
  - Site-directed mutagenesis with the cell-free Q5 SDM protocol
  - Plasmid production for hit variants
  - Cell-free protein synthesis
  - Bacterial enzyme production and purification
  - RnKlRED test reactions
  - LINDT test reaction
  - Measurement of specific activity
  - Preparative-scale enzymatic reactions
  - Analytical methods
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
- [ADDITIONAL RESOURCES](#)

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at.

Received: September 8, 2025

Revised: October 29, 2025

Accepted: November 25, 2025

Published: November 27, 2025

### REFERENCES

1. Buller, R., Lutz, S., Kazlauskas, R.J., Snajdrova, R., Moore, J.C., and Bornscheuer, U.T. (2023). From nature to industry: Harnessing enzymes for biocatalysis. *Science* 382, eadh8615. <https://doi.org/10.1126/science.adh8615>.
2. Bell, E.L., Finnigan, W., France, S.P., Green, A.P., Hayes, M.A., Hepworth, L.J., Lovelock, S.L., Niikura, H., Osuna, S., Romero, E., et al. (2021). Biocatalysis. *Nat. Rev. Methods Primers* 1, 46. <https://doi.org/10.1038/s43586-021-00044-z>.
3. Wu, S., Snajdrova, R., Moore, J.C., Baldenius, K., and Bornscheuer, U.T. (2021). Biocatalysis: Enzymatic Synthesis for Industrial Applications. *Angew. Chem. Int. Ed.* 60, 88–119. <https://doi.org/10.1002/anie.202006648>.
4. Choi, J.M., Han, S.S., and Kim, H.S. (2015). Industrial applications of enzyme biocatalysis: Current status and future aspects. *Biotechnol. Adv.* 33, 1443–1454. <https://doi.org/10.1016/j.biotechadv.2015.02.014>.
5. Hughes, D.L. (2022). Highlights of the Recent Patent Literature—Focus on Biocatalysis Innovation. *Org. Process. Res. Dev.* 26, 1878–1899. <https://doi.org/10.1021/acs.oprd.1c00417>.
6. Woodley, J.M. (2019). Accelerating the implementation of biocatalysis in industry. *Appl. Microbiol. Biotechnol.* 103, 4733–4739. <https://doi.org/10.1007/s00253-019-09796-x>.
7. Romero, E.O., Saucedo, A.T., Hernández-Meléndez, J.R., Yang, D., Chakrabarty, S., and Narayan, A.R.H. (2023). Enabling Broader Adoption of Biocatalysis in Organic Chemistry. *JACS Au* 3, 2073–2085. <https://doi.org/10.1021/jacsau.3c00263>.
8. Truppo, M.D. (2017). Biocatalysis in the Pharmaceutical Industry: The Need for Speed. *ACS Med. Chem. Lett.* 8, 476–480. <https://doi.org/10.1021/acsmedchemlett.7b00114>.
9. Goodwin, N.C., Morrison, J.P., Fuerst, D.E., and Hadi, T. (2019). Biocatalysis in Medicinal Chemistry: Challenges to Access and Drivers for Adoption. *ACS Med. Chem. Lett.* 10, 1363–1366. <https://doi.org/10.1021/acsmedchemlett.9b00410>.

10. Gregorio, N.E., Levine, M.Z., and Oza, J.P. (2019). A User's Guide to Cell-Free Protein Synthesis. *Methods Protoc.* 2, 24. <https://doi.org/10.3390/mps2010024>.
11. Silverman, A.D., Karim, A.S., and Jewett, M.C. (2020). Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21, 151–170. <https://doi.org/10.1038/s41576-019-0186-3>.
12. Levine, M.Z., Gregorio, N.E., Jewett, M.C., Watts, K.R., and Oza, J.P. (2019). *Escherichia coli*-Based Cell-Free Protein Synthesis: Protocols for a robust, flexible, and accessible platform technology. *J. Vis. Exp.* 144, e58882. <https://doi.org/10.3791/58882>.
13. Contreras-Llano, L.E., and Tan, C. (2018). High-throughput screening of biomolecules using cell-free gene expression systems. *Synth. Biol.* 3, ysy012. <https://doi.org/10.1093/synbio/ysy012>.
14. Vasina, M., Kovar, D., Damborsky, J., Ding, Y., Yang, T., Demello, A., Mazurenko, S., Stavakis, S., and Prokop, Z. (2023). In-depth analysis of biocatalysts by microfluidics: An emerging source of data for machine learning. *Biotechnol. Adv.* 66, 108171. <https://doi.org/10.1016/j.biotechadv.2023.108171>.
15. Hadi, T., Nozzi, N., Melby, J.O., Gao, W., Fuerst, D.E., and Kvam, E. (2020). Rolling circle amplification of synthetic DNA accelerates biocatalytic determination of enzyme activity relative to conventional methods. *Sci. Rep.* 10, 10279. <https://doi.org/10.1038/s41598-020-67307-9>.
16. Madani, A., Krause, B., Greene, E.R., Subramanian, S., Mohr, B.P., Holton, J.M., Olmos, J.L., Xiong, C., Sun, Z.Z., Socher, R., et al. (2023). Large language models generate functional protein sequences across diverse families. *Nat. Biotechnol.* 41, 1099–1106. <https://doi.org/10.1038/s41587-022-01618-2>.
17. Qu, G., Li, A., Acevedo-Rocha, C.G., Sun, Z., and Reetz, M.T. (2020). The Crucial Role of Methodology Development in Directed Evolution of Selective Enzymes. *Angew. Chem. Int. Ed.* 59, 13204–13231. <https://doi.org/10.1002/anie.201901491>.
18. Qin, Z., Yuan, B., Qu, G., and Sun, Z. (2024). Rational enzyme design by reducing the number of hotspots and library size. *Chem. Commun.* 60, 10451–10463. <https://doi.org/10.1039/D4CC01394H>.
19. Quertinmont, L.T., Orru, R., and Lutz, S. (2015). RAPid Parallel Protein EvaluatoR (RAPPER), from gene to enzyme function in one day. *Chem. Commun.* 51, 122–124. <https://doi.org/10.1039/C4CC08240K>.
20. Watanabe, S., Ito, M., and Kigawa, T. (2021). DiRect: Site-directed mutagenesis method for protein engineering by rational design. *Biochem. Biophys. Res. Commun.* 551, 107–113. <https://doi.org/10.1016/j.bbrc.2021.03.021>.
21. Landwehr, G.M., Bogart, J.W., Magalhaes, C., Hammarlund, E.G., Karim, A.S., and Jewett, M.C. (2025). Accelerated enzyme engineering by machine-learning guided cell-free expression. *Nat. Commun.* 16, 865. <https://doi.org/10.1038/s41467-024-55399-0>.
22. Planas-Iglesias, J., Marques, S.M., Pinto, G.P., Musil, M., Stourac, J., Damborsky, J., and Bednar, D. (2021). Computational design of enzymes for biotechnological applications. *Biotechnol. Adv.* 47, 107696. <https://doi.org/10.1016/j.biotechadv.2021.107696>.
23. Marques, S.M., Planas-Iglesias, J., and Damborsky, J. (2021). Web-based tools for computational enzyme design. *Curr. Opin. Struct. Biol.* 69, 19–34. <https://doi.org/10.1016/j.sbi.2021.01.010>.
24. Bao, Y., Xu, Y., and Huang, X. (2024). Focused rational iterative site-specific mutagenesis (FRISM): A powerful method for enzyme engineering. *Mol. Catal.* 553, 113755. <https://doi.org/10.1016/j.mcat.2023.113755>.
25. Hyslop, J.F., Lovelock, S.L., Sutton, P.W., Brown, K.K., Watson, A.J.B., and Roiban, G.D. (2018). Biocatalytic Synthesis of Chiral N-Functionalized Amino Acids. *Angew. Chem. Int. Ed.* 57, 13821–13824. <https://doi.org/10.1002/anie.201806893>.
26. Telek, A., Dargó, G., Kovács, R., Molnár, Z., Vértessy, B.G., and Tasnádi, G. (2025). Enzymatic Production of Opine-Type Chiral Amines with Controlled Stereoselectivity. *ChemCatChem* 17, e202402066. <https://doi.org/10.1002/cctc.202402066>.
27. Salihovic, A., Ascham, A., Taladriz-Sender, A., Bryson, S., Withers, J.M., McKean, I.J.W., Hoskisson, P.A., Grogan, G., and Burley, G.A. (2024). Gram-scale enzymatic synthesis of 2'-deoxyribonucleoside analogues using nucleoside transglycosylase-2. *Chem. Sci.* 15, 15399–15407. <https://doi.org/10.1039/D4SC04938A>.
28. Salihovic, A., Ascham, A., Rosenqvist, P.S., Taladriz-Sender, A., Hoskisson, P.A., Hodgson, D.R.W., Grogan, G., and Burley, G.A. (2025). Biocatalytic synthesis of ribonucleoside analogues using nucleoside transglycosylase-2. *Chem. Sci.* 16, 1302–1307. <https://doi.org/10.1039/D4SC07521H>.
29. Sumbalova, L., Stourac, J., Martinek, T., Bednar, D., and Damborsky, J. (2018). HotSpot Wizard 3.0: web server for automated design of mutations and smart libraries based on sequence input information. *Nucleic Acids Res.* 46, W356–W362. <https://doi.org/10.1093/NAR/GKY417>.
30. Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., et al. (2012). CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* 8, e1002708. <https://doi.org/10.1371/JOURNAL.PCBI.1002708>.
31. Jurcik, A., Bednar, D., Byska, J., Marques, S.M., Furmanova, K., Daniel, L., Kokkonen, P., Brezovsky, J., Strnad, O., Stourac, J., et al. (2018). CAVER Analyst 2.0: analysis and visualization of channels and tunnels in protein structures and molecular dynamics trajectories. *Bioinformatics* 34, 3586–3588. <https://doi.org/10.1093/bioinformatics/bty386>.
32. Smith, C.I.E., and Zain, R. (2019). Therapeutic Oligonucleotides: State of the Art. *Annu. Rev. Pharmacol. Toxicol.* 59, 605–630. <https://doi.org/10.1146/annurev-pharmtox-010818-021050>.
33. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., De Beer, T.A.P., Rempfer, C., Bordoli, L., et al. (2018). SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46, W296–W303. <https://doi.org/10.1093/NAR/GKY427>.
34. Borel, F., Hachi, I., Palencia, A., Gaillard, M.C., and Ferrer, J.L. (2014). Crystal structure of mouse mu-crystallin complexed with NADPH and the T3 thyroid hormone. *FEBS J.* 281, 1598–1612. <https://doi.org/10.1111/febs.12726>.
35. Anandkrishnan, R., Aguilar, B., and Onufriev, A.V. (2012). H++ 3.0: Automating pK prediction and the preparation of biomolecular structures for atomistic molecular modeling and simulations. *Nucleic Acids Res.* 40, W537–W541. <https://doi.org/10.1093/nar/gks375>.
36. Maier, J.A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K.E., and Simmerling, C. (2015). ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. *J. Chem. Theory Comput.* 11, 3696–3713. <https://doi.org/10.1021/acs.jctc.5b00255>.
37. Wang, J., Wang, W., Kollman, P.A., and Case, D.A. (2006). Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* 25, 247–260. <https://doi.org/10.1016/j.jmgm.2005.12.005>.
38. Holmberg, N., Ryde, U., and Bülow, L. (1999). Redesign of the coenzyme specificity in L-Lactate dehydrogenase from *Bacillus stearothermophilus* using site-directed mutagenesis and media engineering. *Protein Eng.* 12, 851–856. <https://doi.org/10.1093/protein/12.10.851>.
39. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. <https://doi.org/10.1063/1.445869>.
40. Hopkins, C.W., Le Grand, S., Walker, R.C., and Roitberg, A.E. (2015). Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* 11, 1864–1874. <https://doi.org/10.1021/ct5010406>.

41. Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* *23*, 327–341. [https://doi.org/10.1016/0021-9991\(77\)90098-5](https://doi.org/10.1016/0021-9991(77)90098-5).
42. Davidchack, R.L., Handel, R., and Tretyakov, M.V. (2009). Langevin thermostat for rigid body dynamics. *J. Chem. Phys.* *130*, 234101. <https://doi.org/10.1063/1.3149788>.
43. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A., and Haak, J.R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* *81*, 3684–3690. <https://doi.org/10.1063/1.448118>.
44. Armstrong, S.R., Cook, W.J., Short, S.A., and Ealick, S.E. (1996). Crystal structures of nucleoside 2-deoxyribosyltransferase in native and ligand-bound forms reveal architecture of the active site. *Structure* *4*, 97–107. [https://doi.org/10.1016/S0969-2126\(96\)00013-5](https://doi.org/10.1016/S0969-2126(96)00013-5).

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Bacterial and virus strains</b>		
<i>E. coli</i> XL1Blue	Agilent Technologies	50-125-058
<i>E. coli</i> BL21(DE3)	Sigma-Aldrich	69450-M
<i>E. coli</i> BL21(DE3)	Thermo Scientific	EC0114
<b>Chemicals, peptides, and recombinant proteins</b>		
Chemical 1	Combi-Blocks	OR-1599
Chemical 1	Angene Chemical	AG003Q00
Chemical 2	Fluorochem	169520
Chemical 4	BLD Pharmatech	BD8618
Chemical 5	Combi-Blocks	QC-2820
Chemical 7	Angene Chemical	AG0006OP
Chemical 8	Combi-Blocks	QK-5329
RnKIREd and variants (Table S2)	This study	N/A
LINDT and variants (Table S4)	This study	N/A
<b>Critical commercial assays</b>		
NEBExpress® Cell-free <i>E. coli</i> Protein Synthesis System	NEB	E5360L
<b>Oligonucleotides</b>		
Primers (Table S5)	Microsynth	N/A
<b>Recombinant DNA</b>		
pET14b-RnKIREd	Genscript	N/A
pET14b-RnKIREd variants (Table S2)	This study	N/A
pET14b-LINDT	Genscript	N/A
pET14b-LINDT variants (Table S4)	This study	N/A
<b>Software and algorithms</b>		
CAVER Analyst	Loschmidt Laboratories	<a href="https://caver.cz/index.php">https://caver.cz/index.php</a>
Hotspot Wizard	Loschmidt Laboratories	<a href="https://loschmidt.chemi.muni.cz/hotspotwizard/">https://loschmidt.chemi.muni.cz/hotspotwizard/</a>
Amber	N/A	<a href="https://ambermd.org/index.php">https://ambermd.org/index.php</a>

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

*Escherichia coli* XL1Blue strain was used to prepare the plasmids. The cells were incubated in Luria-Bertani (LB) medium containing 50-50 mg/mL tetracycline and carbenicillin at 37°C for 16 h. *Escherichia coli* BL21(DE3) strain was used to express the target proteins. The cells were grown in Luria-Bertani (LB) medium supplemented with carbenicillin (50 mg/mL). The culture was grown to OD = 0.6 and induced by addition of 0.5 mM IPTG and then incubated in at 16°C for 16 h.

## METHOD DETAILS

## Computational analysis and library design

**Acquisition and preparation of structure:** RnKIREd (UniProt ID: Q9QYU4) homology model structure was generated by the SWISS-MODEL<sup>33</sup> server using default settings. The template for modeling was chain A of the crystal structure of the ketimine reductase from *Mus musculus* (PDB: 4BVA)<sup>34</sup>, having 97.5% sequence identity with RnKIREd and 1.75 Å resolution. Titratable residues were protonated, and orientations of Asn, Gln, and His side chains were optimized using the H<sup>++</sup><sup>35</sup> server at 0.15 M salinity and pH 7.5, with the rest of the parameters left at default values. After processing with H<sup>++</sup>, coordinated crystallographic waters were added from the structure of 4BVA.

**Parametrization:** Input topologies and coordinates were prepared using the Tleap module of AMBER 14. The protein atoms of the system were parametrized using the ff14SB<sup>36</sup> force field. The NADPH cofactor was parametrized using the GAFF<sup>37</sup> force field with partial charges and force field parameter modifications taken from the AMBER parameter database.<sup>38</sup> The system was solvated in a truncated octahedron of water molecules so that all protein atoms were at least 9 Å from the water box's surface, so the system is well contained within the periodic box. The TIP3P<sup>39</sup> water model was used. The charge of the system was neutralized using 4 Na<sup>+</sup> ions. The number of ions added for the production simulations was determined using the average volume in the last stage of the equilibration simulation to achieve a final salinity of 0.15 M NaCl. Lastly, masses of protein groups containing hydrogen atoms were re-partitioned using the AMBER14 program ParmEd by its command HMassRepartition.<sup>40</sup>

**Molecular Dynamics Simulations:** Energy minimizations and MD simulations were performed using the PMEMD.CUDA module of AMBER16. Initially, the system was minimized by 10 steps of steepest descent followed by 9990 steps of conjugate gradient with 500 kcal mol<sup>-1</sup> × Å<sup>-2</sup> restraints on all atoms of the protein. The system was further minimized in four more rounds, each consisting of 2500 steps of steepest descent and 7500 steps of conjugate gradient minimization with decreasing harmonic restraints. The restraints were applied as follows: 500, 125, and 25 kcal mol<sup>-1</sup> × Å<sup>-2</sup> on all backbone atoms of the protein. Finally, the system was minimized with 5000 steps of steepest descent and 15000 steps of conjugate gradient minimization without any restraints. The subsequent MD simulations employed periodic boundary conditions, the Particle Mesh Ewald method to treat electrostatic interactions with the long-range cut-off of 10 Å and the same cut-off for nonbonded interactions, the SHAKE<sup>41</sup> algorithm to fix all bonds containing hydrogens, and a 4 fs time step. Equilibration simulations consisted of two steps: (I) 500 ps of gradual heating from 0 to 300 K using the Langevin<sup>42</sup> thermostat with a collision frequency of 5.0 ps<sup>-1</sup> and constant volume, (II) 400 ps in the NPT ensemble using the Berendsen barostat<sup>43</sup> at constant pressure of 1.0 bar and pressure relaxation time of 1.0 ps<sup>-1</sup> and harmonic restraints of 150.0 kcal × mol<sup>-1</sup> × Å<sup>-2</sup> on the positions of all protein backbone atoms. Then, the system was further equilibrated during 8800 ps at 310 K in the NPT ensemble as in step (II) in 10 rounds of 400 ps each with decreasing harmonic restraints. The restraints were applied as follows: 100.0, 75.0, 50.0, 25, 15.0, 10.0, 5.0, 1.0, 0.5, and 0.0 kcal × mol<sup>-1</sup> × Å<sup>-2</sup> on the backbone atoms of the protein. After equilibration with the final NaCl concentration, 35 ns long production MD simulations were run using the same settings as the last equilibration step. Coordinates were saved at intervals of 4 ps. The simulation was run in 4 replicas, producing 140 ns of aggregated simulation time.

**Analysis of molecular dynamics simulations using CAVER Analyst 2:** The production MD simulations were visualized in CAVER Analyst 2.0 beta. Tunnels were calculated from the production MD simulation trajectories. The tunnel origin was set on the center point between residues R48 and K75 C $\alpha$  atoms. The calculation was run using default settings except probe\_radius = 1.4 Å and exclude\_end\_zone = 2 Å.

**Assessment using Hotspot Wizard:** The structure of RnKIREd was submitted to the Hotspot Wizard web server and was processed using default settings. The conservation and correlation of the individual positions, as well as their mutational landscape and mutational effect predictions, were used to construct the mutational libraries. For RnKIREd, residues labeled as hotspots, both reliably and with unreliable mutational data, were considered for selection into the mutant library (Figure S1 and Table S1). In the case of LINDT (UniProt ID: Q9R5V5), we focused on residues lining the ligand in the structure (PDB: 1F8X)<sup>44</sup> which included hotspots, tentative hotspots as well as non-hotspots. Hotspot Wizard was used to select amino acid substitutions based on stability score and chemical diversity (Figure S2 and Table S3).

### Site-directed mutagenesis with the cell-free Q5 SDM protocol

Mutagenic PCR reactions were conducted on 96-well PCR plate according to the Q5 Site-Directed Mutagenesis Kit Quick Protocol provided by the manufacturer (12.5  $\mu$ L reaction volume). A single annealing temperature was used (65°C for RnKIREd libraries, 64°C for the LINDT library) for 30 cycles. KLD reactions were assembled according to the manufacturer's protocol and incubated at room temperature for 1h. 2  $\mu$ L from the KLD reactions was transferred to a new 50  $\mu$ L Q5 Hot Start PCR reaction containing the primers for amplification of the coding sequence at 0.5  $\mu$ M concentration. 72°C was used for both annealing and elongation (50s, 35 cycles) to minimize crossreactions with contaminating primers from the first PCR step. KLD reactions were stored at -20°C until eventual use for plasmid production. PCR products were purified using the Mag-Bind TotalPure NGS kit according to the manufacturers protocol and sent for sequencing. DNA concentration was determined by measuring absorbance at 260 nm using Thermo Scientific Multiskan SkyHigh Microplate Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) with  $\mu$ DropDuo plate.

### Plasmid production for hit variants

KLD reactions of selected mutants were transformed into XL1Blue chemically competent cells and were selected on LB-TC/CAR agarose plates. Single colonies were picked and used to inoculate 5 mL overnight cultures. From these, plasmids were purified according to the NucleoSpin Plasmid Prep protocol. The purified plasmids were sent to sequencing to verify mutations, and then used for bacterial protein expression.

### Cell-free protein synthesis

CFPS was carried out using the NEBExpress Cell-free *E. coli* Protein Synthesis System according to manufacturer's protocol (12.5–50  $\mu$ L reaction volumes, 16h incubation at 29°C). The synthesized variants were eventually immobilized on NEBExpress Ni-NTA Magnetic Beads applying 10  $\mu$ L slurry per reaction (see Figure S8).

### Bacterial enzyme production and purification

*RnKIRED* and *LINDT* genes were codon optimized for *E. coli*, synthesized by Genescript and cloned into pET14b vector using NdeI and BamHI cloning sites resulting in N-terminal His-tag proteins (see [supplemental information](#) for protein sequences).

pET14b-*RnKIRED* or its variants were transformed into *E. coli* BL21(DE3) chemically competent cells, and 5 mL overnight cultures were grown in Luria-Bertani (LB) medium supplemented with carbenicillin. Then, 4 mL was added to 500 mL fresh LB medium. The culture was grown at 37°C to OD = 0.6 and induced by addition of 0.5 mM IPTG. The flasks were then incubated in a shaker at 16°C overnight. The cells were harvested by centrifugation at 4000 rpm, 4°C for 30 min. The pellets were resuspended in lysis buffer (50 mM Tris pH = 8.5, 300 mM NaCl, 1 mM benzamidine, 2 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM Tris(2-carboxyethyl)-phosphine (TCEP), DNase) and lysed by ultrasonication. The suspension was centrifuged at 11000 rpm, 4°C for 30 min. The supernatant was loaded onto equilibrated Ni-NTA column, washed with 15 mM imidazole in lysis buffer. The bound protein was eluted with 250 mM imidazole in lysis buffer and then dialyzed in 50 mM Tris pH = 8.5. The resulting solution was aliquoted after addition of 10% glycerol and then stored at –20°C until further use. Total protein concentration was determined based on absorption at 280 nm measured on a NanoDrop spectrophotometer.

The transformation of pET14b-*LINDT* or its variants and subsequent selection and preparation of the master cell bank (glycerol stock) was carried out according to the protocols provided by the manufacturer of *E. coli* BL21(DE3) competent cells (Thermo Scientific). Then, seed culture was prepared by inoculating 100 mL of LB medium containing ampicillin (100 mg L<sup>-1</sup>) with 1000 mL glycerol stock of the producer strains. The flasks were incubated at 28°C, 250 rpm overnight. OD<sub>600</sub> reached 6.0–6.5 and the homogeneity of the culture was confirmed by microscopic examination. 20 mL of *E. coli* seed culture was used to inoculate 2000 mL of 2 YT medium containing ampicillin (100 mg L<sup>-1</sup>) and incubated at 37°C, 250 rpm agitation until the culture reached OD<sub>600</sub> 0.4–0.5. The expression was induced by addition of IPTG (1 mM final concentration). The fermentation was continued for 3 h at 37°C. The biomass was harvested by centrifugation at 5500g and 4°C for 20 min and cell pellets were stored at –20°C. The cell pellet was resuspended (100 mL end volume) in chromatography load buffer (20 mM NaP<sub>i</sub>, 500 mM NaCl, 20 mM imidazole, pH 7.4) and homogenized twice (1000 bar, 4°C–8°C), centrifuged twice (8000 g, 4°C, 10 min) and the supernatant was used for protein purification. The purification of the enzymes was carried out on an AKTA Purifier (UPC100) and a Cytiva's HisTrap HP 5 mL column according to the manufacturer's recommendations using load buffer for washing and elution buffer containing 20 mM NaP<sub>i</sub>, 500 mM NaCl, 500 mM imidazole, pH: 7.4. Then the purified solutions were dialyzed in a dialysis buffer composed of 50 mM Tris×HCl, 250 mM NaCl, pH:7.0 twice for 12 h using 50× volume dialysis buffer at 4°C with gentle stirring. Then, the protein solution was concentrated using Amicon Ultra Centrifugal Filter, 10 kDa MWCO combined with further buffer exchange (4×). After the purification process, protein samples were stable in the elution buffer at 4°C but during the dialysis, significant precipitation was observed in case of all the enzymes. Precipitated protein was removed by centrifugation, but further precipitation was observed. After the second removal of the precipitated protein, the solutions seemed stable at 4°C (5 days of observation). There was at least 50–75% of protein loss was confirmed by Bradford-assay. The proteins were stored at –20°C.

### RnKIRED test reactions

For activity screening of variant libraries, the reaction mixtures (100 µL) contained 0.4 mM NADP<sup>+</sup>, 25 mM **1** and 1.1–3 equivalents of **2** with 2.5 equivalents of d-glucose and 6 U/mL GDH (Codexis CDX-901, 50 U/mg) in 100 mM Na-phosphate buffer containing 20% v/v DMF, adjusted to pH 7.5 (Reaction 1) or 0.4 mM NADP<sup>+</sup>, 10–50 mM **4** and 2–50 equivalents of **3** with 2.5 equivalents of d-glucose and 6 U/mL GDH (Codexis CDX-901, 50 U/mg) in 100 mM Na-phosphate buffer containing 10% v/v DMSO, adjusted to pH 7.5 (Reaction 2). The blank mix was added directly to the CFPS reactions or to the magnetic beads containing the immobilized variants (for details see Screening results). The reaction mixtures were incubated at 30°C with shaking at 600 rpm for 16–48 h. Analysis was done on HPLC-UV and HPLC-MS. Conversions (%) were calculated by using the HPLC area of limiting reactant (including hydrolyzed amino acid form when applicable) and expected product, and were given as area percentage of expected product at 210 nm, assuming similar UV responses of substrate and product.

### LINDT test reaction

For activity screening of the variant library, the reaction mixtures (150 µL) contained 5 mM 2-chloroadenine (**8**) and 50 mM 2'-deoxy-2'-fluorocytidine (**7**) in 10 mM NaP<sub>i</sub> buffer, adjusted to pH 7.0. The reaction mix was added to the magnetic beads containing the immobilized variants or the wild type enzyme and to three empty wells as well. The reaction mixture was incubated at 40°C with shaking at 400 rpm for 18–22 h. Analysis was done on HPLC-UV and HPLC-MS. Transfer rate (%) was calculated as the ratio of HPLC area under the curve, at 254 ± 4 nm, of nucleoside product (2-chloroadenosine, **9**) and the sum of 2-chloroadenine and 2-chloroadenosine given as a percentage, assuming similar UV response of substrate and product.

### Measurement of specific activity

Enzyme activities were measured spectrophotometrically at 30°C in 96-well half-area plates in a Thermo Scientific Multiskan SkyHigh Microplate Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA). Standard assays were carried out by using 20 mM **5** and 0.4 mM NADPH in 100 mM sodium phosphate buffer containing 10% v/v DMSO (pH = 7.5). Depending on the activities of each enzyme, enzyme concentrations were fixed between 0.05 and 0.5 mg/mL. The reaction was started by the addition of 40 mM **4**. The

decrease in absorbance was monitored at 340 nm for 2.5 h. Initial velocities were calculated from the linear section of the plots using an NADPH calibration and specific activities calculated based on enzyme amounts used in each reaction.

### Preparative-scale enzymatic reactions

**Scale-up of (2S,2'S)-3.** 100 mL reaction mixture contained 50 mM **1**, 250 mM (5 eq.) **2**, 0.25 mM (0.25 eq.) NADP<sup>+</sup> and 2.5 equivalents of d-glucose (125 mM) and 6 U/mL GDH (Codexis CDX-901, 50 U/mg) in 100 mM Na-phosphate buffer containing 20% v/v DMF, adjusted to pH 7.5. The reaction mixture was incubated at 30°C with stirring at 200 rpm in a 250-mL Optimax glass reactor. *Rn*KIRED V49I/H91I/S228G enzyme was added to the reaction in two aliquots ( $t_0$ : 10 mg, 25 h: 1 mg). pH was monitored and kept at 7.5 with 2M NaOH solution using an SP-50 Dosing Unit. The reaction was stopped after 69 h, when conversion reached 90%+, using 150 mL MeCN. The 5 G Celite was added to the solution, stirred for another 15 min, then filtered. The filtrate was concentrated in vacuo, then lyophilized. Purification was carried out using preparative HPLC. Fractions containing the (2S,2'S)-**3** compound were merged, concentrated and lyophilized resulting in 543 mg pure product (yield = 39%, >99% *de*).

**Scale-up of (2S,2'S)-6.** 100 mL reaction mixture contained 50 mM **4**, 200 mM (4 eq.) **2**, 0.25 mM (0.25 eq.) NADP<sup>+</sup> and 2.5 equivalents of d-glucose (125 mM) and 6 U/mL GDH (Codexis CDX-901, 50 U/mg) in 100 mM Na-phosphate buffer containing 10% v/v DMSO, adjusted to pH 7.5. The reaction mixture was incubated at 30°C with stirring at 200 rpm in a 250-mL Optimax glass reactor. *Rn*KIRED H91A/S228G enzyme was added to the reaction in two aliquots ( $t_0$ : 10 mg, 24 h: 0.9 mg). pH was monitored and kept at 7.5 with 2M NaOH solution using an SP-50 Dosing Unit. The reaction was stopped after 44 h, when conversion reached 97%+. From the mother liquor 523 mg pure (2S,2'S)-**6** could be isolated by filtration. Then, the filtrate was extracted using 2 x 200 mL DCM. The organic phases were merged and dried on MgSO<sub>4</sub>, followed by filtration of the drying agent and removal of the organic solvent in vacuo. The crude product was redissolved in water, cooled down to 4°C resulting in precipitation of the product. By filtration another 444 mg pure (2S,2'S)-**6** could be isolated. (total yield = 63%, >99% *de*).

### Analytical methods

**HPLC methods.** HPLC measurements were performed on Agilent Technologies 1260 LC system equipped with a DAD detector using Gemini 3  $\mu$ m NX-C18, 50 mm  $\times$  3.00 mm i.d. 110 Å column and 5 mM aqueous NH<sub>4</sub>HCO<sub>3</sub> solution and MeCN as eluents in gradient mode. Analytical LC-MS was performed on an Agilent Technologies 1200 LC system equipped with Agilent 6140 quadrupole MS, operating in positive or negative ion electrospray ionization mode (molecular weight scan range was 100–1350 m/z) with parallel UV detection using Gemini 3  $\mu$ m NX-C18, 50 mm  $\times$  3.00 mm i.d. 110 Å column and 5 mM aqueous NH<sub>4</sub>HCO<sub>3</sub> solution and MeCN as eluents in gradient mode.

**Compound purification.** Purifications using preparative HPLC were carried out with Teledyne Isco preparative HPLC using a Gemini 5  $\mu$ m NX-C18, 250 mm  $\times$  50 mm i.d. 110 Å column and 5 mM aqueous NH<sub>4</sub>HCO<sub>3</sub> solution and MeCN as eluents in gradient mode. Collected fractions were analyzed using HPLC, the fractions containing the desired product were merged, concentrated by rotavap, then lyophilized.

**NMR measurements.** <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra were recorded on a Bruker Avance Ultrashield 400 (100 MHz <sup>13</sup>C) instrument with Bruker Prodigy Cryo Probe and are internally referenced to residual protium solvent signals (note: DMSO referenced at 2.50 and 39.52 ppm in <sup>1</sup>H and <sup>13</sup>C NMR measurements, respectively). Samples (3–8 mg) were dissolved in 0.5 mL DMSO-*d*<sub>6</sub>. Data for <sup>1</sup>H NMR are reported as follows: chemical shift ( $\delta$  ppm), multiplicity (s = singlet, d = doublet, dd = doublet of a doublet, t = triplet, dt = doublet of a triplet, m = multiplet, br is used for broad signals), integration, coupling constant (Hz). Assignments of protons are listed on the individual spectra. Data for <sup>13</sup>C NMR are reported in terms of chemical shift and no special nomenclature is used for equivalent carbons.

**High resolution mass spectrometry.** HRMS measurements were carried out on an Agilent 6545 Q-TOF mass spectrometer system, mass resolution: 45,000 FWHM @ m/z 2,722 Da, ion source: AJS-ESI, sheath gas temperature: 300°C, drying gas temperature: 300°C, ionizing voltage: 2500V, nozzle voltage: 1000V.

### QUANTIFICATION AND STATISTICAL ANALYSIS

DNA concentration was determined by measuring absorbance at 260 nm using Thermo Scientific Multiskan SkyHigh Microplate Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) with  $\mu$ DropDuo plate. Protein concentration was determined based on absorption at 280 nm measured on a NanoDrop spectrophotometer. For *Rn*KIRED, conversions (%) were calculated by using the HPLC area of limiting reactant (including hydrolyzed amino acid form when applicable) and expected product, and were given as area percentage of expected product at 210 nm, assuming similar UV responses of substrate and product. For *L*INDT, transfer rate (%) was calculated as the ratio of HPLC area under the curve, at 254  $\pm$  4 nm, of nucleoside product (2-chloroadenosine, **9**) and the sum of 2-chloroadenosine and 2-chloroadenosine given as a percentage, assuming similar UV response of substrate and product. Enzyme activities were measured spectrophotometrically at 30°C in 96-well half-area plates in a Thermo Scientific Multiskan SkyHigh Microplate Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) in triplicates. The decrease in absorbance was monitored at 340 nm. Initial velocities were calculated from the linear section of the plots using an NADPH calibration and specific activities calculated based on enzyme amounts used in each reaction.

**ADDITIONAL RESOURCES**

[Quick Protocol for Q5 Site-Directed Mutagenesis Kit \(E0554\) | NEB.](#)  
[Mag-Bind TotalPure NGS – Omega Bio-tek.](#)  
[NEBExpress Cell-free Ecoli Protein Synthesis System | NEB.](#)  
[NEBExpress Ni-NTA Magnetic Beads | NEB.](#)