UC Irvine
UC Irvine Previously Published Works

Title
A sequence motif enables widespread use of noncanonical redox cofactors in natural enzymes

Permalink
https://escholarship.org/uc/item/3vw2w5g8

Journal
Nature Chemical Biology

ISSN
1552-4450

Authors
Saleh, Samer
Hsu, Ning-Hsiang
Luu, Emma
et al.

Publication Date
2026-09-01

DOI
10.1038/s41589-026-02315-w

Copyright Information
This work is made available under the terms of a Creative Commons Attribution-NonCommercial-ShareAlike License, available at https://creativecommons.org/licenses/by-nc-sa/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

nature chemical biology

Article

https://doi.org/10.1038/s41589-026-02315-w

A sequence motif enables widespread use of noncanonical redox cofactors in natural enzymes

Received: 18 August 2025

Accepted: 31 July 2026

Published online: 09 September 2026

Check for updates

Samer Saleh[1,10], Ning-Hsiang Hsu[2,10], Emma Luu[3], Vincent C. Martin[1], Hans Jefferson C. Ng[1], William B. Black[1], Shiding Zhang[1], Jin-Kwang Kim[2], Banumathi Sankaran[4], Andre H. T. Tran[2], Ryan L. Hayes[1,5,6,7], Justin B. Siegel[3,8,9], Feng Qiao[2] ✉ & Han Li[1,2,5] ✉

Noncanonical redox cofactors (NRCs) are low-cost alternatives to the natural redox cofactors nicotinamide adenine dinucleotide (NAD^+) and nicotinamide adenine dinucleotide phosphate ($NADP^+$) for biomanufacturing, offering exquisite electron-delivery control, yet their adoption is limited by the scarcity of compatible enzymes. Screening the aldehyde dehydrogenase (ALDH) family, we identified a conserved RH/QxxR motif that enables widespread NRC activity among natural enzymes. *Bos taurus* ALDH3a1 exhibits unprecedented turnover with nicotinamide mononucleotide (NMN^+), with k_{cat} values exceeding NAD^+ and surpassing most engineered NRC-active enzymes by 10–10^5-fold. Structural analyses reveal that this motif reinforces cofactor positioning and preorganizes the active site independently of the NAD^+ adenosine monophosphate moiety. This motif supports activity across simple-synthetic NRCs such as 1-(2-carbamoylmethyl)nicotinamide and, when introduced into diverse ALDH scaffolds, enhances NMN^+ activity up to 60-fold. These findings elucidate nature's solution to engineering NRC-active enzymes and offer a blueprint to mine latent evolutionary plasticity in natural enzymes that serve as superior engineering starting points.

Biomanufacturing, a paradigm where cell and enzyme catalysts facilitate the production of commodities, materials and medicines, is a key component in human society's sustainable future. While biomanufacturing has been successfully applied toward a variety of products, many of these applications depend on the natural cofactors nicotinamide adenine dinucleotide (phosphate) ($NAD(P)^+$) to mediate redox reactions. In whole-cell systems, dependence on natural redox cofactors poses many challenges, including undesirable side reactions, disrupted cell fitness and insufficient driving force for thermodynamically challenging reactions[1–3]. In cell-free systems, $NAD(P)^+$ also introduces a prohibitive cost, even with recycling[4,5].

Noncanonical redox cofactors (NRCs) are biomimetic analogs of $NAD(P)^+$ that can replace $NAD(P)^+$ in oxidoreductases[6–13]. These cofactors have attracted interest for biotechnological purposes because they can overcome the common drawbacks of natural cofactors. Nonetheless, NRCs have not yet seen wide adoption because of the

[1]Department of Chemical and Biomolecular Engineering, University of California, Irvine, Irvine, CA, USA. [2]Department of Biological Chemistry, University of California, Irvine, Irvine, CA, USA. [3]Genome Center, University of California, Davis, Davis, CA, USA. [4]Molecular Biophysics and Integrated Bioimaging, Berkeley Center for Structural Biology, Lawrence Berkeley National Laboratory, Berkeley, CA, USA. [5]UCI Center for Synthetic Biology, University of California, Irvine, Irvine, CA, USA. [6]Department of Pharmaceutical Sciences, University of California, Irvine, Irvine, CA, USA. [7]Department of Chemistry, University of California, Irvine, Irvine, CA, USA. [8]Department of Chemistry, University of California, Davis, Davis, CA, USA. [9]Department of Biochemistry and Molecular Medicine, University of California, Davis, Davis, CA, USA. [10]These authors contributed equally: Samer Saleh, Ning-Hsiang Hsu.
✉e-mail: qiao@uci.edu; han.li@uci.edu

scarcity of enzymes that can accept them effectively[14]. Engineering NRC-dependent enzymes has proven difficult, usually yielding variants with catalytic efficiencies toward NRCs many orders of magnitude below the parent enzyme's native NAD(P)$^+$ activity, as discussed below.

This gap in performance between natural and engineered enzymes is well documented, underscoring the power of natural evolution in shaping enzymes[15]. Designing the complex conformational dynamics required for optimal enzyme catalysis is a long-standing challenge, as evident from the severely suboptimal catalytic rates of engineered enzymes[15]. In this study, we took a systems approach to survey a broad natural sequence space, identified naturally occurring NRC-using enzymes, distilled a sequence motif that dictates this activity and elucidate its molecular mechanism. These enzymes not only offer ready-to-optimize scaffolds for engineering but also teach guiding principles that accelerate the development of efficient NRC-dependent enzymes.

We chose nicotinamide mononucleotide (NMN$^+$) as the NRC for this study because of its previously demonstrated success in model biomanufacturing systems, both in vivo and in vitro[9,16–20]. NMN$^+$ is a biomimetic cofactor lacking the adenosine monophosphate (AMP) moiety of NAD(P)$^+$ (Fig. 1). Aside from flavin-dependent enzymes, which are known to accept divergent electron donors and acceptors[21–24], examples of NMN$^+$-using natural enzymes are very rare[24–26]. To investigate this phenomenon, we selected the aldehyde dehydrogenase (ALDH) protein family (pFAM) because it is diverse in sequence and function while maintaining high structural homology[27]. This makes it easier to deconvolute sequence–function relationships and allows NRC-enabling features to be translated across scaffolds, as demonstrated here. ALDHs have also seen use in the production of bioplastics, fragrances and biofuels, thus representing an interesting target for enhanced industrial biocatalysis[28–32].

Briefly, we generated a sequence similarity network (SSN) encompassing the entire ALDH pFAM and tested representatives for activity towards NAD$^+$ and NMN$^+$ in high throughput. Surprisingly, more than one third of the ALDHs tested showed activity towards NMN$^+$. Using a design, build, test and learn (DBTL) approach, a signature RH/QxxR motif was discovered to be conserved in the sequence subnetwork housing the two most efficient NRC-using enzymes: *Bos taurus* ALDH3a1 (*Bt*ALDH3) and a putative ALDH from *Pseudanabaena biceps* (*Pb*ALDH). *Bt*ALDH3 shows 1.5-fold higher k_{cat} toward NMN$^+$ ($2.1 \pm 0.1\ s^{-1}$) than NAD$^+$ ($1.4 \pm 0.1\ s^{-1}$). *Pb*ALDH also exhibits a high k_{cat} toward NMN$^+$ ($3.02 \pm 0.01\ s^{-1}$). To provide a reference point, the ALDH functions belong to secondary metabolism. Among diverse secondary metabolism enzymes, the median value of k_{cat} for natural reactions is ~$2.5\ s^{-1}$ (ref. 33). These ALDHs are $10–10^5$-fold better than all other natural or engineered enzymes reported so far, according to the ratio of NMN$^+$ catalytic efficiency (k_{cat}/K_M) to the wild-type (WT) enzyme's native NAD(P)$^+$ activity, with a single exception[19].

This unusually high level of activity calls for mechanistic study. We revealed the mechanism of the RH/QxxR motif as an active site organization center using mutational studies, crystallography and molecular dynamics modeling. The three key residues in this motif are proposed to control the conformational flexibility of the active site and reinforce cofactor positioning during catalysis by supporting interactions with an 'aromatic lid'. Importantly, NMN$^+$ catalytic efficiency was increased up to 60-fold when RH/QxxR was translated onto eight unrelated ALDHs. Moreover, the RH/QxxR motif-containing ALDHs have been shown to use a broad range of structurally divergent, nonnucleotide, simple-synthetic NRCs such as 1-benzylnicotinamide (BNA$^+$), 1-(4-carboxybenzyl)nicotinamide (BANA$^+$), 1-[3-(4-methoxyphenyl) propyl]nicotinamide (P3NA-OMe$^+$) and 1-(carbamoylmethyl)nicotinamide (AmNA$^+$), opening new avenues to use these cost-effective NAD(P)$^+$ alternatives in biotransformation.

NAD(P)$^+$ serve as the universal redox cofactors across all domains of life. They are widely regarded as molecular fossils of the ancient RNA world and are proposed to predate the emergence of enzymes. Consequently, one may posit that redox enzymes have been evolutionarily optimized to bind NAD(P)$^+$ with exquisite specificity. This work unveils that there remains considerable plasticity in the cofactor recognition landscape, likely within enzymes that more firmly contact the electron transfer nicotinamide 'head' on the NMN$^+$ moiety of the cofactor, while relying less on interactions with the rest of the cofactor, which would be the AMP moiety 'tail' of the cofactor that engineers seek to diversify[34]. A similar approach can be taken for other enzyme classes to understand the generalizability of this observation. Natural evolution, with its immense combinatorial depth, has generated a rich library of enzyme scaffolds that could be mined to advance NRC enzyme engineering.

Results

High-throughput screening of ALDHs for activity toward NMN$^+$

To guide our efforts in identifying NMN$^+$ active ALDHs, we followed a DBTL cycle (Fig. 1). The ALDH pFAM represents over 508,000 sequences, covering divergent functions such as detoxification of aldehydes, synthesizing cell signaling messengers and mitigating ocular damage from ultraviolet irradiation[27,35]. To efficiently explore the natural sequence diversity of ALDHs with experimentally tractable throughput, we reduced the redundancy of available sequences and identified representatives from each region of sequence space. To achieve this, we reduced the number of available sequences down to ~10,000 by using UniRef50 cluster representatives rather than all available sequences[36]. Next, we constructed an SSN to visualize the relationships between ALDHs in sequence space (Fig. 1a)[37].

From the ALDH SSN, we chose 42 candidate sequences, sampling from all the largest subnetworks, as well as a set from the isolated nodes. Sequences were primarily chosen from each subnetwork at random; however, a small subset of sequences were chosen on the basis of literature reports of attractive properties such as thermostability, high turnover or broad substrate scope. Subnetworks containing sequences encoding for bifunctional or CoA-acylating ALDHs were avoided. The 42 chosen ALDHs originate from a wide range of organisms and share an average of 18% pairwise sequence identity. The chosen ALDHs were cloned, expressed in *Escherichia coli* and purified (Supplementary Tables 1 and 2). The purified enzymes were then subjected to a 96-well plate colorimetric cycling assay (Fig. 1b), in which a diaphorase from *Geobacillus* sp. (*Gs*DI) rapidly oxidizes the reduced cofactor (NADH or NMNH) generated by ALDH and concomitantly reduces the WST-1 tetrazolium dye, producing a sensitive absorbance readout[13].

As most of the chosen ALDH sequences are only inferred to belong to the ALDH pFAM by sequence homology and are otherwise uncharacterized, we performed preliminary screening to determine their preferred aldehyde substrate when coupled to NAD$^+$. We selected various aliphatic aldehyde substrates (acetaldehyde, butyraldehyde and hexanal) and coupled each ALDH with NAD$^+$ to determine their substrate preference (Supplementary Table 3). Next, each ALDH was assayed with its preferred aldehyde substrate and NMN$^+$ as the cofactor (Fig. 1b). Surprisingly, we identified that more than one third of the ALDHs tested had activity with NMN$^+$ (Fig. 1c and Supplementary Table 3). The ALDH showing the highest activity toward NMN$^+$ was *Bt*ALDH3. *Bt*ALDH3 rapidly achieved 100% conversion using NMN$^+$ and its NMN$^+$ activity was robust across substrates (Supplementary Fig. 1).

Characterization of NMN$^+$-active ALDHs

We characterized the top two performing ALDHs, *Bt*ALDH3 and *Pb*ALDH, by measuring their apparent kinetic parameters with NAD$^+$ and NMN$^+$ (Table 1). To contextualize the unusual NMN$^+$ activity these enzymes exhibit, we compared them to natural and engineered enzymes previously characterized for NMN$^+$ use[9,13,16,17,19,20,38–40]. From the existing literature, all other WT enzymes show NMN$^+$ activities that are $10^3–10^6$ fold lower than their native activities with their preferred

https://doi.org/10.1038/s41589-026-02315-w

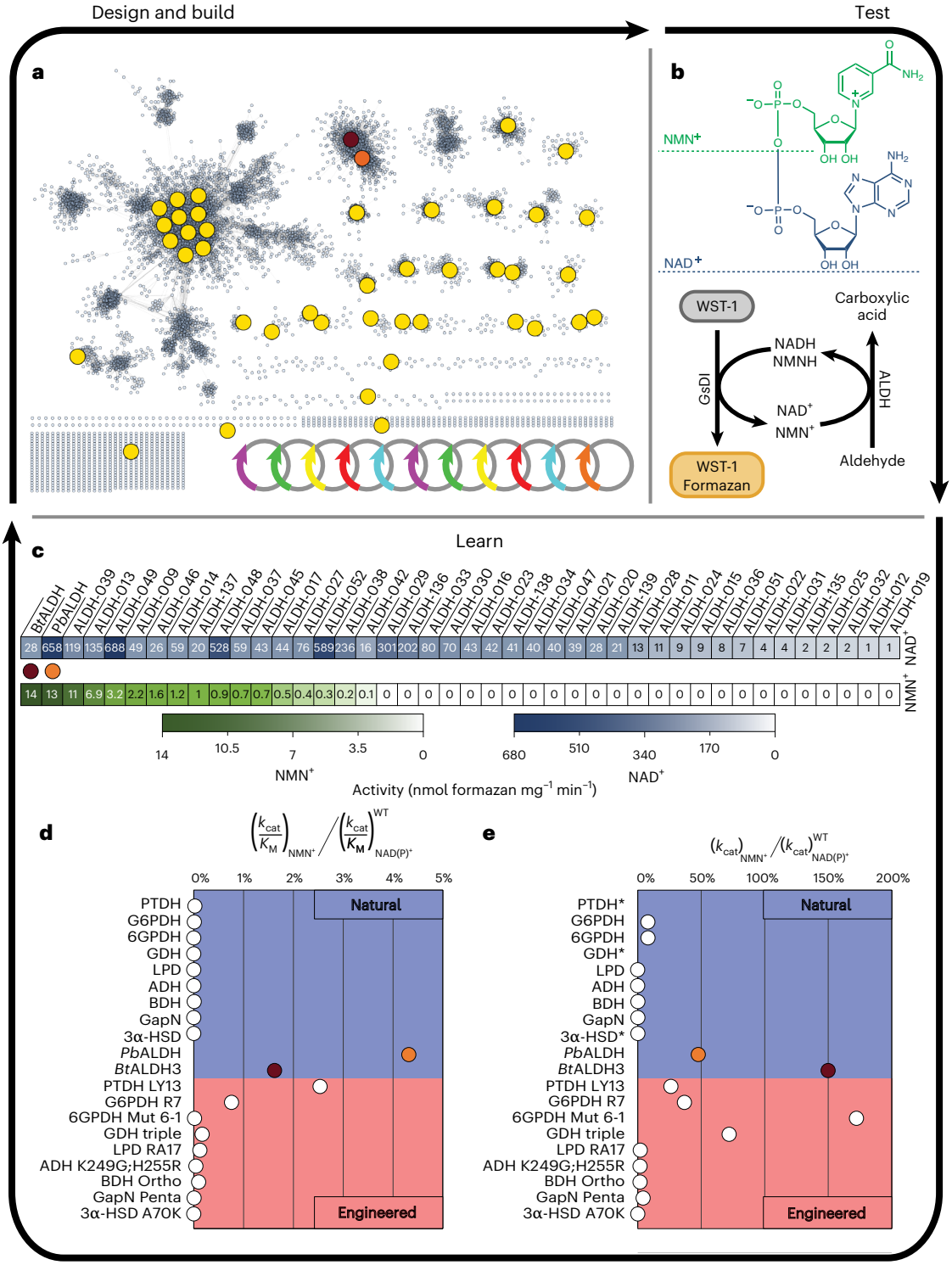

Fig. 1 | DBTL cycle for bioprospecting NMN$^+$-using ALDHs. a, The ALDH SSN is used to efficiently screen ALDHs. Each node represents a UniRef50 cluster of related ALDHs and edges represent shared sequence similarity between each node. A total of 42 nodes (yellow, burgundy, and orange) were chosen for characterization. **b**, Once designed and built, each ALDH scaffold was tested in a high-throughput colorimetric cycling assay toward NAD$^+$ and NMN$^+$. Cofactor recycling and WST-1 dye enable screening of many conditions and generates a sensitive readout for NMN$^+$ activity with a wide dynamic range[13]. **c**, Results of colorimetric cycling assay. Each ALDH was assayed with its preferred substrate (acetaldehyde, butyraldehyde or hexanal) for activity with NAD$^+$ (blue) and NMN$^+$ (green). The highest NMN$^+$-using ALDHs, *Bt*ALDH3 (burgundy node) and *Pb*ALDH (orange node) show remarkably high levels of NRC activity for naturally occurring enzymes. **d**, Comparing their ratio of catalytic efficiency (k_{cat}/K_M) with NMN$^+$ to that with of the WT with NAD$^+$, these ALDHs outperform all other naturally occurring enzymes similarly characterized in the literature and all but one engineered enzyme (PTDH LY13)[19], according to the ratio relative to native cofactor (NAD(P)$^+$) activity, by several orders of magnitude. **e**, Similarly, comparing their k_{cat} with NMN$^+$ to their native cofactor k_{cat}, *Bt*ALDH3 and *Pb*ALDH outperform all other naturally occurring enzymes and most engineered enzymes in the literature. All colorimetric assay measurements were carried out with $n = 2$ biologically independent replicates. *The k_{cat} values for PTDH and GDH with NMN$^+$ are not reported in the literature.

https://doi.org/10.1038/s41589-026-02315-w

Table 1 | Kinetic parameters of ALDHs and NMN$^+$-using dehydrogenases

	NAD$^+$			NMN$^+$		
This work[a]	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($s^{-1}mM^{-1}$)	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($s^{-1}mM^{-1}$)
*Bt*ALDH3	1.4±0.1	0.040±0.001	34.0±1.5	2.1±0.1	3.7±0.2	0.56±0.03
*Pb*ALDH	6.32±0.60	0.92±0.19	6.96±0.86	3.02±0.01	10.2±0.2	0.29±0.01
ALDH-013	0.84±0.04	0.59±0.06	1.4±0.2	ND[b]	ND[b]	(4.8±0.8)×10^{-3}
ALDH-013 RHR (S342R, I343H, G346R)	0.39±0.05	1.20±0.10	0.33±0.04	0.31±0.01	37.5±2.6	(8.4±0.6)×10^{-3}
ALDH-037	0.50±0.04	0.23±0.02	2.2±0.1	ND[b]	ND[b]	(2.0±0.1)×10^{-4}
ALDH-037 RHR (T365R, Q366H, K369R)	0.13±0.01	0.42±0.03	0.32±0.02	0.38±0.02	34.6±4.7	0.012±0.001
ALDH-137	0.44±0.03	0.029±0.001	14.9±1.4	0.059±0.01	37.6±3.6	(1.6±0.1)×10^{-3}
ALDH-137 RHR (S312R, I313H, Y316R)	0.035±0.001	0.028±0.006	1.3±0.3	0.11±0.01	6.7±0.2	0.017±0.001
	NAD(P)$^+$[c]			NMN$^+$		
Previous studies	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($s^{-1}mM^{-1}$)	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($s^{-1}mM^{-1}$)
PTDH[19]	1.03±0.01	0.06±0.01	18±1	ND[b]	ND[b]	(4±1)×10^{-3}
PTDH LY13 (ref. 19)	2.06±0.11	0.18±0.03	12±2	0.27±0.01	0.62±0.05	0.44±0.03
G6PDH (ref. 20)[c]	16±1[c]	0.037±0.00[c]	440[c]	1.3±0.1	28±7	0.046
G6PDH R7 (ref. 20)[c]	1.1±0.2[c]	7.1±0.0[c]	0.15[c]	5.9±0.2	1.8±0.1	3.3
6PGDH (ref. 13)[c]	15.9±0.2[c]	(1.2±0.1)×10^{-3}[c]	13,394[c]	1.3±0.1	30.6±1.7	0.04
6PGDH Mut 6-1 (ref. 13)[c]	28.9±0.3[c]	0.19±0.01[c]	148.6[c]	27.4±0.5	13.5±0.5	2.04
GDH (ref. 16)[c]	4.3±0.1[c]	0.015±0.001[c]	280±10[c]	ND[b]	ND[b]	(4.7±0.5)×10^{-4}
GDH Triple (ref. 16)[c]	4.4±0.1[c]	0.61±0.15[c]	7.5±1.8[c]	3.1±0.1	6.4±0.8	0.51±0.06
LPD[38]	150±10	1.1±0.1	130±10	(1.7±0.1)×10^{-3}	8.3±0.3	(2.1±0.1)×10^{-4}
LPD RA17 (ref. 38)	33±2	37±3	0.9±0.1	3.2±0.1	19±2	0.17±0.02
ADH[9]	1.42	0.06	23.6	5×10^{-4}	2.50	2×10^{-4}
ADH K249G;H255R (ref. 9)	3.00	0.46	6.5	0.03	2.6	0.01
BDH[17]	20±1	0.22±0.03	96±20	0.043±0.004	5.6±0.3	(7.9±1)×10^{-3}
BDH Ortho[17]	0.019±0.001	1.6±0.1	0.012±0.001	0.38±0.04	4.3±0.4	0.086±0.002
GapN (ref. 39)[c]	19±0.7[c]	0.020±0.001[c]	930[c]	0.012±0.001	8.1±0.8	1.5×10^{-3}
GapN Penta (ref. 39)[c]	0.26±0.01[c]	2.6±0.4[c]	0.10[c]	0.82±0.1	12±4	0.067
3α-HSD[40]	135±5	0.0414±0.0052	3,200±300	0.025±0.003	17.1±2.6	(1.46±0.06)×10^{-3}
3α-HSD A70K (ref. 40)	ND[b]	ND[b]	0.29±0.01	0.19±0.01	14±2	0.0127±0.0007

[a]To determine apparent kinetic parameters, reactions were carried out in 100 mM potassium phosphate buffer pH 7.4, 5 mM hexanal, 5 mM dithiothreitol and varying concentrations of NAD$^+$ or NMN$^+$ (0.005–100 mM) and varying concentrations of purified ALDH (0.01–0.5 mg ml^{-1}). All reactions were carried out with $n=3$ biologically independent replicates. Data are reported as the mean±s.d. [b]The enzyme could not be saturated with the cofactor concentrations tested. [c]Kinetic parameters are reported for NADP$^+$. All others are reported for NAD$^+$ or NMN$^+$.

natural cofactors (NAD(P)$^+$) as measured by the ratio of catalytic efficiency $(k_{cat}/K_M)_{NMN^+}$ to $(k_{cat}/K_M)_{NAD(P)^+}$ (Fig. 1d and Table 1). After engineering, many of the reported enzymes showed significantly increased catalytic efficiency for NMN$^+$, albeit still $10–10^5$ fold lower than their natural activities as calculated by the ratio between $(k_{cat}/K_M)_{NMN^+}$ of engineered enzymes to $(k_{cat}/K_M)_{NAD(P)^+}$ of the WT parent (Fig. 1d and Table 1). Remarkably, *Bt*ALDH3 and *Pb*ALDH, two WT enzymes without any engineering, already have NMN$^+$ catalytic efficiencies that are ~2% and ~4% of NAD$^+$, respectively (Fig. 1d). This may seem small; however, with a single exception (engineered phosphite dehydrogenase (PTDH) LY13, which has NMN$^+$ catalytic efficiency at ~2.5% of the WT parent's natural NAD$^+$ catalytic efficiency[19]), *Bt*ALDH3 and *Pb*ALDH are $10–10^5$ fold better than all other enzymes, both before and after engineering, in terms of their $(k_{cat}/K_M)_{NMN^+}$ compared to their evolutionarily honed native activity (Fig. 1d).

While informative, using catalytic efficiency as a metric might unjustly exaggerate the gap between NMN$^+$ and NAD(P)$^+$ activity because K_M values strongly depend on ligand sizes[33]. Below 350 Da (molecular weight of NMN$^+$ = 334 Da, NAD$^+$ = 663 Da), K_M exponentially increases as the ligand molecular weight decreases, suggesting that enzymes face intrinsic limitations in recognizing ligands < 350 Da (ref. 33). These limits may be insurmountable even with natural evolution's unparalleled depth in searching sequence space. In this respect, k_{cat} alone provides a more direct measure of intrinsic catalytic compatibility under saturating conditions. Indeed, when we compared k_{cat}, the gap between NMN$^+$ and NAD(P)$^+$ activity narrowed (Fig. 1e and Table 1). With a single exception (engineered 6-phosphogluconate dehydrogenase Mut 6-1)[13], WT *Bt*ALDH3 was better than all other enzymes, both before and after engineering, in terms of the ratio between $(k_{cat})_{NMN^+}$ and the corresponding WT parent's natural $(k_{cat})_{NAD(P)^+}$ for its preferred cofactor. *Pb*ALDH also ranked high, surpassing all but two engineered enzymes in this analysis (Fig. 1e and Table 1), and outperformed WT enzymes for which this metric has been published. The absolute k_{cat} values of *Bt*ALDH3 (2.1 ± 0.1 s^{-1}) and *Pb*ALDH (3.02 ± 0.1 s^{-1}) with NMN$^+$ are in line with the expected k_{cat} observed in secondary metabolism enzymes[33]. Notwithstanding the absence of standardized NRC assay

https://doi.org/10.1038/s41589-026-02315-w

conditions and the difficulty of comparing enzymes across functional classes, the available literature supports our discovery of natural enzymes that rival purpose-engineered enzymes for NMN^+ use.

Elucidating the sequence determinant of NMN^+ activity

Both *Bt*ALDH3 and *Pb*ALDH originate from the same subnetwork in the ALDH SSN (Fig. 1a). The colocalization of NMN^+-using ALDHs indicated that this corresponding subnetwork may be carrying a common sequence feature that enables activity toward NMN^+. To investigate this hypothesis, we iterated on our DBTL approach and generated a refined SSN by isolating the nearest neighbors of *Bt*ALDH3 and *Pb*ALDH (Fig. 2a). From this refined SSN, *Bt*ALDH3 and *Pb*ALDH were found to be related by three degrees of separation, representing five ALDH sequences. These five ALDHs include *Homo sapiens* ALDH3a1 (*Hs*ALDH3), as well as four putative ALDHs (ALDH-053, ALDH-054, ALDH-140, and ALDH-141), and share an average of ~50% pairwise sequence identity. In addition to *Bt*ALDH and *Pb*ALDH, these five ALDHs share edges to ~80% of the nodes within their respective subnetwork, forming a representative group of ALDHs. All five ALDHs were cloned, expressed and tested toward various aliphatic aldehydes using NAD^+ in a 340-nm spectrophotometric assay and their preferred substrate was identified. Next, the specific activity of *Bt*ALDH3, *Pb*ALDH and their five related ALDHs were similarly measured for both NAD^+ and NMN^+ (Fig. 2b and Supplementary Fig. 2). All five enzymes demonstrated activity toward NMN^+ and, in four of five cases, this activity was comparable to or exceeded *Bt*ALDH3 and *Pb*ALDH. These results further suggested that the sequence space containing *Bt*ALDH3 and *Pb*ALDH represents a group of NMN^+-using ALDHs.

To elucidate what sequence features facilitate activity toward NMN^+, we investigated the sequences of *Bt*ALDH3 and its related ALDHs, focusing only on regions structurally relevant to the cofactor-binding domain. Multiple-sequence alignment of these ALDHs (Fig. 2c) uncovered a unique R289, Q/H290, R293 (*Bt*ALDH3 numbering) motif (RH/QxxR motif), which is conserved in the *Bt*ALDH3 sequence subnetwork but not commonly observed outside this subnetwork of ALDH sequences. In fact, a comparison of the consensus sequences generated from each subnetwork in the ALDH pFAM SSN shows that the *Bt*ALDH3 subnetwork is the only one carrying this motif as the consensus sequence (Supplementary Table 4). These results indicated that the RH/QxxR motif could be the unique feature of the *Bt*ALDH3 subnetwork enabling NMN^+ activity. However, to ensure that the observed activity toward NMN^+ is indeed enabled by the RH/QxxR motif and not some other sequence feature associated with the *Bt*ALDH3 subnetwork, we identified two additional ALDHs containing the RH/QxxR motif (ALDH-155 and ALDH-238) from outside the *Bt*ALDH3 subnetwork (Supplementary Fig. 3). Both ALDHs showed activity toward NMN^+, further supporting our hypothesis. Similarly, we reexamined the ALDHs from our colorimetric screen that showed detectable levels of NMN^+ activity to identify whether they shared any similar sequence features to the RH/QxxR motif. Analysis of the top performers shows an enrichment for similar sequences to the RH/QxxR motif, including ALDH-009 (I343, Q344, R347; IQxxR), ALDH-039 (A326, H327, R330; AHxxR) and ALDH-046 (K346, Q347, R350; KQxxR).

Structural mechanism of NMN^+ use by RH/QxxR motif

Having identified a potential sequence determinant of NMN^+ use in ALDHs, we performed alanine scanning at the RH/QxxR motif of *Bt*ALDH3 to assess its importance (Fig. 3a). The R289A, H290A and R293A single mutants showed a 47%, 89% and 98% reduction in activity toward NMN^+, respectively, relative to the WT. The *Bt*ALDH3-R289A;H290A;R293A triple mutant showed 99% reduction in relative activity toward both NAD^+ and NMN^+. In all cases, except for the R289A;H290A;R293A triple mutant, the reduction in activity observed for NMN^+ was always significantly greater than that observed for NAD^+ (Supplementary Fig. 4). These results indicate that, while

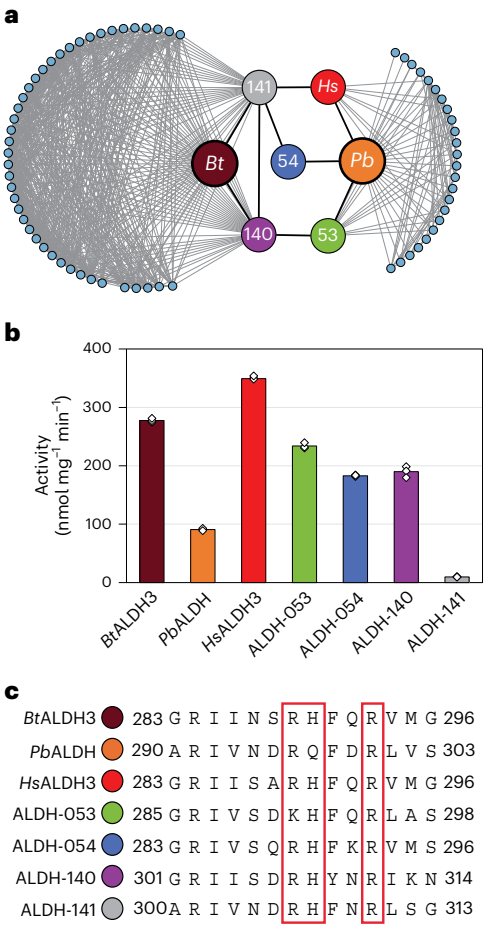

Fig. 2 | Refinement and investigation of the *Bt*ALDH3/*Pb*ALDH subnetwork. **a**, *Bt*ALDH3a1 (*Bt*ALDH3, burgundy) and *Pb*ALDH (orange) show the highest level of NMN^+ activity in our initial screen and appear in a shared region of the ALDH SSN. To investigate the sequence–function relationship of enzyme NRC activity, this region was isolated to generate a refined subnetwork of their interrelated sequences. *Bt*ALDH3 and *Pb*ALDH are related by three degrees of separation encompassed by five ALDHs sharing an average pairwise sequence identity of ~50%. **b**, All ALDHs interrelated between *Bt*ALDH3 and *Pb*ALDH showed activity toward NMN^+ when assayed with their preferred substrate (acetaldehyde for ALDH-054, hexanal for the rest). **c**, Multiple-sequence alignment of *Bt*ALDH3, *Pb*ALDH and their related ALDHs uncover a highly conserved RH/QxxR motif unique to this subnetwork of ALDH sequences. All specific activity data were generated by monitoring the formation of reduced cofactor (NMNH) at 340 nm and are reported as the mean of $n = 3$ independent biological replicates. Error bars correspond to 1 s.d.

NAD^+ can partially compensate for the loss of the RH/QxxR motif, it is more critical for NMN^+ use. Next, we sought to elucidate the structural mechanism of NMN^+ use by determining the crystal structure of *Bt*ALDH3 complexed with NAD^+ or NMN^+.

We solved the crystal structures of *Bt*ALDH3 in complex with NAD^+ (1.46-Å resolution) and NMN^+ (1.77-Å resolution), as well as an apo structure of *Bt*ALDH3-H290L (1.40-Å resolution) (Fig. 3b–e and Supplementary Table 5). Like the alanine scanning variants, *Bt*ALDH3-H290L displays a substantial reduction in NMN^+ activity relative to WT, which is significantly greater than its reduction in NAD^+ activity (Fig. 3a and Supplementary Fig. 4). This impaired variant is, thus, a suitable candidate for probing the structural role of the RH/QxxR motif. *Bt*ALDH3 shares the structural architecture of class 3 dimeric ALDHs from *Rattus norvegicus*[41], *H. sapiens*[42] and *Pseudomonas putida*[43]. The binary complexes with NAD^+ and NMN^+ are defined by

https://doi.org/10.1038/s41589-026-02315-w

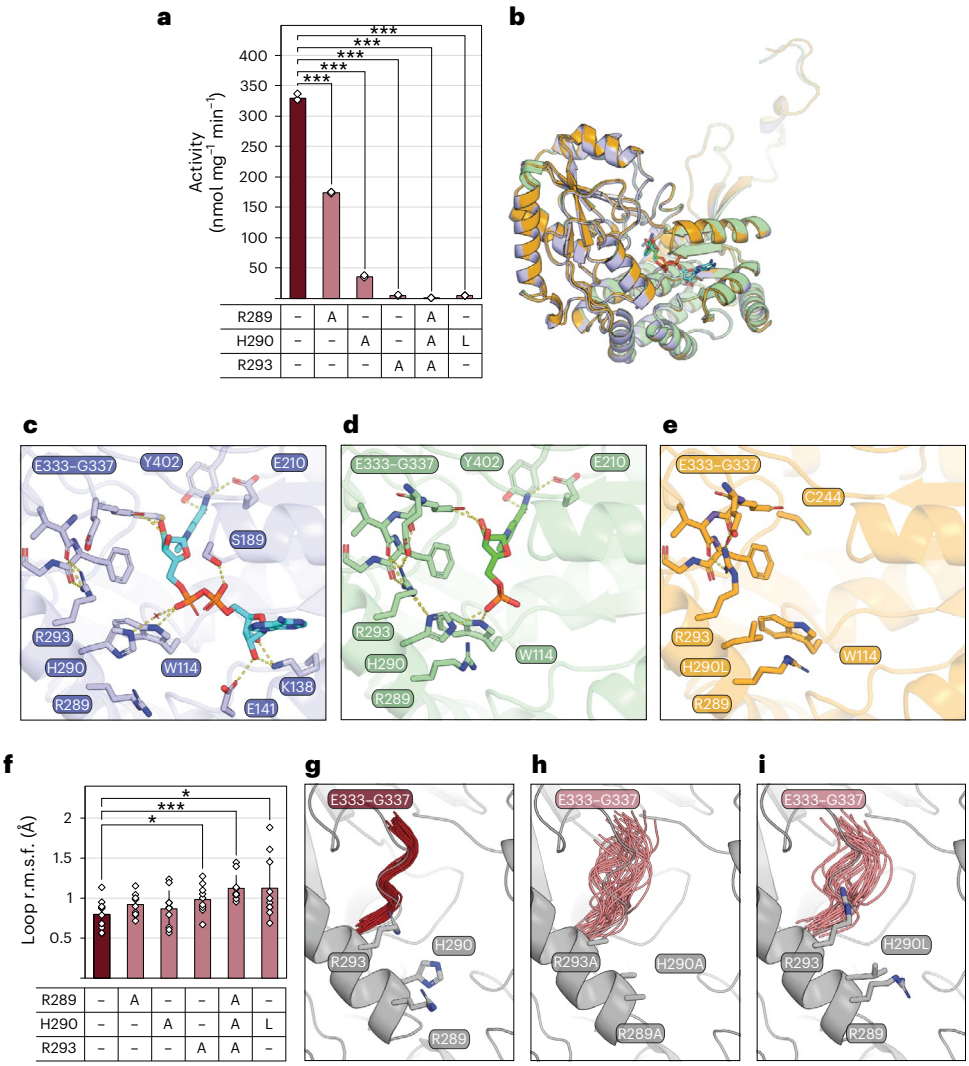

Fig. 3 | Structural analysis of *Bt*ALDH3. a, *Bt*ALDH3 RH/QxxR motif is essential for NMN^+ use (R289A, $P = 4 \times 10^{-4}$; H290A $P = 4 \times 10^{-5}$; R293A, $P = 1 \times 10^{-4}$; R289A;H290A;R293A, $P = 1 \times 10^{-4}$; H290L, $P = 1 \times 10^{-4}$). **b**, Superimposed crystal structures of *Bt*ALDH3 in complex with NAD^+ (blue) or NMN^+ (green) and *Bt*ALDH3-H290L variant (orange). **c**,**d**, *Bt*ALDH3 WT cofactor-binding site with NAD^+ (**c**) or NMN^+ (**d**) bound. The RH/QxxR motif (R289, H290 and R293) supports cofactor binding by stabilizing the active site. Hydrogen bonds to E333–G337 through R293 stabilize this important loop. H290 coordinates the NMN^+ phosphate through direct interaction with the cofactor and supporting W114. **e**, *Bt*ALDH3-H290L crystallized under identical conditions to WT cocrystal, in the presence of NAD^+, shows no NAD^+ electron density. Aside from the alternative side-chain conformation of R293 and lack of observed NAD^+, the key difference between WT and H290L is a greater *B* factor in the H290L active site, notably at E333–G337 (Supplementary Fig. 5). **f**, Molecular dynamics simulations of apo *Bt*ALDH3 variants shows a notable increase in the r.m.s.f. of the E333–G337 Cα, between the WT (0.80 ± 0.19 Å) and the R289A;H290A;R293A triple mutant (1.12 ± 0.17 Å, $P = 6 \times 10^{-4}$). Similar r.m.s.f. increases are seen for other variants (R289A, $P = 0.11$; H290A, $P = 0.45$; R293A, $P = 0.03$; H290L, $P = 0.03$) **g**–**i**, Representative simulation frames illustrate that WT (**g**) maintains an ordered loop, unlike the R289A;H290A;R293A triple mutant (**h**) or the H290L mutant (**i**). Specific activity data are reported as the mean of $n = 3$ independent biological replicates. Simulations are reported as the mean of $n = 10$ trajectories. Error bars correspond to 1 s.d. Pairwise statistical comparisons were performed using a two-tailed Welch's *t*-test (*$P < 0.05$ and ***$P < 0.001$).

clear electron densities (Supplementary Fig. 5). The primary interactions with the cofactors are formed at the carboxamide, NMN^+ ribose and NMN^+ phosphate and mediated by E210, E344, S189 and W114 (Fig. 3c,d). In addition to these interactions, the NAD^+ complex shows additional interactions with the AMP ribose through K138 and E141 and to adenosine through V170, T173, F186, V192 and V195.

To our surprise, NAD^+ and NMN^+ bind nearly identically to *Bt*ALDH3 with a root-mean-square deviation (r.m.s.d.) of 0.193 Å for NMN^+ portion of both cofactors. Like the NAD^+ conformation captured in the crystal structures of other class 3 ALDHs, *Bt*ALDH3 cocrystallizes with NAD^+ and NMN^+ in their hydride transfer poses[42–46]. The structural architecture of *Bt*ALDH3 shows no notable differences in main-chain conformation upon NAD^+ or NMN^+ binding (Fig. 3b). In contrast, cofactor binding has been shown to modulate conformational dynamics in numerous enzymes[47–51]. In particular, the AMP moiety of the NAD^+ can exert long-range allosteric effects to preorganize the enzyme active site[25,52]. In fact, buttressing the void in an enzyme's AMP-binding region has been proven as a generalizable design principle in engineering NMN(H)-dependent enzymes[17,19,53]. As discussed below, our results indicate that the RH/QxxR motif contributes to the high rigidity of *Bt*ALDH3 active site, such that catalysis is well ordered without relying on the AMP moiety of NAD^+ to orchestrate active site organization. Both the *Bt*ALDH3 WT and the H290L mutant structures show almost identical main chains (r.m.s.d. = 0.164 Å; Fig. 3b). Although the H290L crystal was obtained in the presence of fivefold molar excess of NAD^+, no electron density was observed for the cofactor, although an unassignable electron density was observed, which was hypothesized to be an unknown adduct to the catalytic cysteine (Supplementary Fig. 5).

https://doi.org/10.1038/s41589-026-02315-w

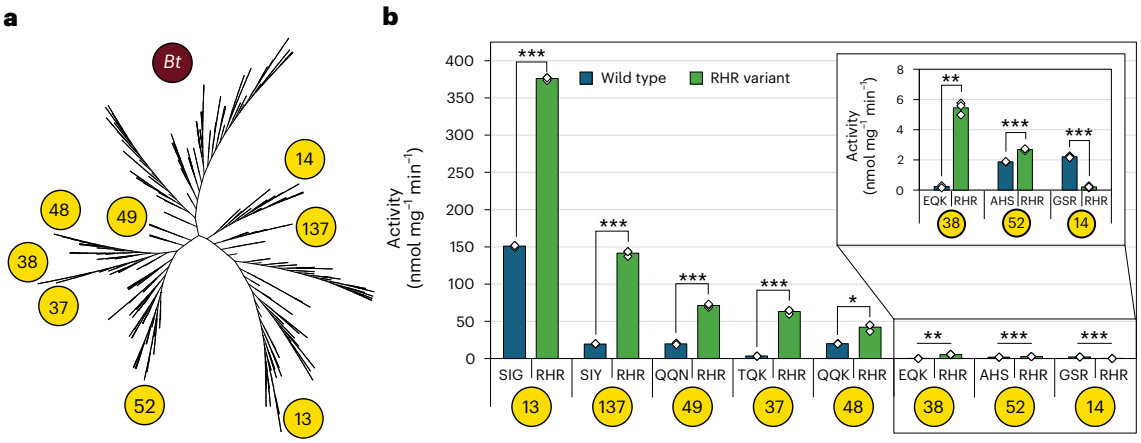

Fig. 4 | Transferring the RH/QxxR motif onto ALDHs enhances NMN^+ activity. **a**, To investigate the transferability of the RH/QxxR motif, we selected all the ALDH scaffolds from our colorimetric screening assay that showed greater than 50 nmol formazan mg^{-1} min^{-1} activity with NAD^+ and any detectable activity toward NMN^+ (yellow nodes). Onto these ALDHs, we incorporated the RH/QxxR motif. These ALDHs show varying starting levels of activity toward NMN^+ and are phylogenetically distinct. ALDH identifiers are shown in yellow circles. **b**, All the RH/QxxR variants of these ALDHs were cloned and all were expressed; however, the ALDH-027 K337R;F338H;W341R mutant could not be purified and was not characterized. All but one showed significantly improved activity toward NMN^+, up to ~23 fold improvement. While ALDH-014 showed a decrease in activity, this may be attributed to its highly unconserved fold. All specific activity data were generated by monitoring the formation of reduced cofactor (NMNH) at 340 nm and are reported as the mean of $n = 3$ independent biological replicates. Error bars correspond to 1 s.d. Statistical comparisons between WT and RHR variants were performed using a two-tailed Welch's t-test (*$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$). ALDH-013, $P = 2 \times 10^{-7}$; ALDH-137, $P = 2 \times 10^{-4}$; ALDH-049, $P = 4 \times 10^{-5}$; ALDH-037, $P = 8 \times 10^{-4}$; ALDH-048, $P = 0.02$; ALDH-038, $P = 0.002$; ALDH-052, $P = 4 \times 10^{-4}$; ALDH-014, $P = 4 \times 10^{-6}$.

This result is consistent with the variant's impaired NAD^+ and NMN^+ activity (Fig. 3a). The catalytic cysteine in enzymes[54,55], including class 3 ALDH[56], has been shown to be sensitive to oxidation or inhibitory modifications. The presence of cofactors has been shown to protect this highly reactive residue[55]. It is intriguing whether the firm binding of cofactor supported by RH/QxxR motif functions to shield the catalytic cysteine, whereas mutating this motif causes cofactor disassociation, exposing the cysteine to nonspecific interactions. Further biochemical and biophysical studies are needed to validate this proposed function.

A key feature observed in both *Bt*ALDH3-NAD^+ and *Bt*ALDH3-NMN^+ crystal structures is that F336 and W114 form an aromatic lid (Fig. 3c,d) that encloses the NMN^+ moiety of the cofactor in the binding pocket. This potentially supports NRCs such as NMN^+ through catalysis without depending on the AMP moiety of NAD^+ for anchoring. Notably, the RH/QxxR motif is critical for positioning this aromatic lid. R293 forms a series of hydrogen bonds to the backbones of E334, V335 and F336, as well as a salt bridge with E333 (Fig. 3c,d and Supplementary Fig. 6). This network of interactions reinforces the loop spanning E333–G337, which holds F336 in place. H290 coordinates the NMN^+ phosphate through a direct hydrogen-bond interaction, in addition to a hydrogen bond to W114. Additionally, H290 forms a supporting hydrogen bond that coordinates R293 (Fig. 3d and Supplementary Fig. 6). In the H290L variant, the R293 guanidinium shifts ~2 Å and establishes alternative interactions with the E333–G337 loop (Fig. 3e). The lack of interactions with W114 and inability of H290L to correctly position R293 is hypothesized to result in a more disordered cofactor-binding site that eliminates activity with NMN^+.

R289 is hypothesized to modulate the electrostatic environment and support the hydrogen-bond network mediated through H290 and R293. This mechanism, albeit not captured in the static crystal structure, was inferred from molecular dynamic studies, as discussed below. The spacer residues (xx) in the RH/QxxR motif are required to maintain the α-helix between H290 and R293, positioning R293 to face the active site and interact directly with the E333–G337 loop. While their side-chain identities are relatively conserved among the natural NRC-using ALDHs tested above, those side chains are solvent exposed and do not appear to have any notable synergy with the rest of the motif. Ultimately, while R289, H290 and R293 are required for enabling NRC activity, the identities of the spacer residues (xx) can vary (see below).

Despite significantly reducing activity for NMN^+, *Bt*ALDH3-H290L shows no major structural difference compared to *Bt*ALDH3 WT. To better understand the effect of this substitution, we examined the normalized temperature factors of H290L in comparison to WT. Congruent with the hypothesis that the RH/QxxR motif rigidifies the active site, H290L shows a higher normalized temperature factor at the E333–G337 loop compared to the WT (Supplementary Fig. 5). To further probe this hypothesis, we performed molecular dynamics analysis to pinpoint the effect of the RH/QxxR motif in stabilizing the catalytically important E333–G337 loop. In addition to housing the aromatic lid residue F336, the E333–G337 loop may have a broader impact on active site dynamics. F336 is responsible for modulating the conformational dynamics of the nicotinamide portion of the cofactor during catalysis[44]. Moreover, the E333–G337 loop is adjacent to a flexible loop containing the central catalytic residue of *Bt*ALDH3, C244. Therefore, fastening down the E333–G337 loop may prevent excessive fluctuations that disrupt the catalytic cysteine.

Molecular dynamics simulations on apo *Bt*ALDH3 WT, the R289A, H290A and R293A single mutants, the corresponding triple mutant and the H290L mutant were conducted. The flexibility of E333–G337 was quantified by determining the average root-mean-square fluctuation (r.m.s.f.) across the Cα of E333–G337, after alignment to the average trajectory structure (Fig. 3f). The triple alanine substitution significantly increased the r.m.s.f. of the E333–F336 loop from 0.80 ± 0.19 Å to 1.12 ± 0.17 Å ($P = 0.0006$). This can be visually understood by observing that the E333–G337 loop adopts an ordered conformation in *Bt*ALDH3 WT (Fig. 3g) but adopts a wider range of catalytically irrelevant poses upon mutating the RH/QxxR motif (Fig. 3h,i). Similar increases in r.m.s.f. were observed for the single mutants R289A (0.92 ± 0.14 Å, $P = 0.11$) and H290A (0.87 ± 0.23 Å, $P = 0.45$) and were significant for R293A (0.98 ± 0.18 Å, $P = 0.03$) and H290L (1.12 ± 0.39 Å, $P = 0.03$). These results are consistent with the specific activity measurements of these variants, where significant increases in r.m.s.f. are associated with greater reductions in activity toward NMN^+.

https://doi.org/10.1038/s41589-026-02315-w

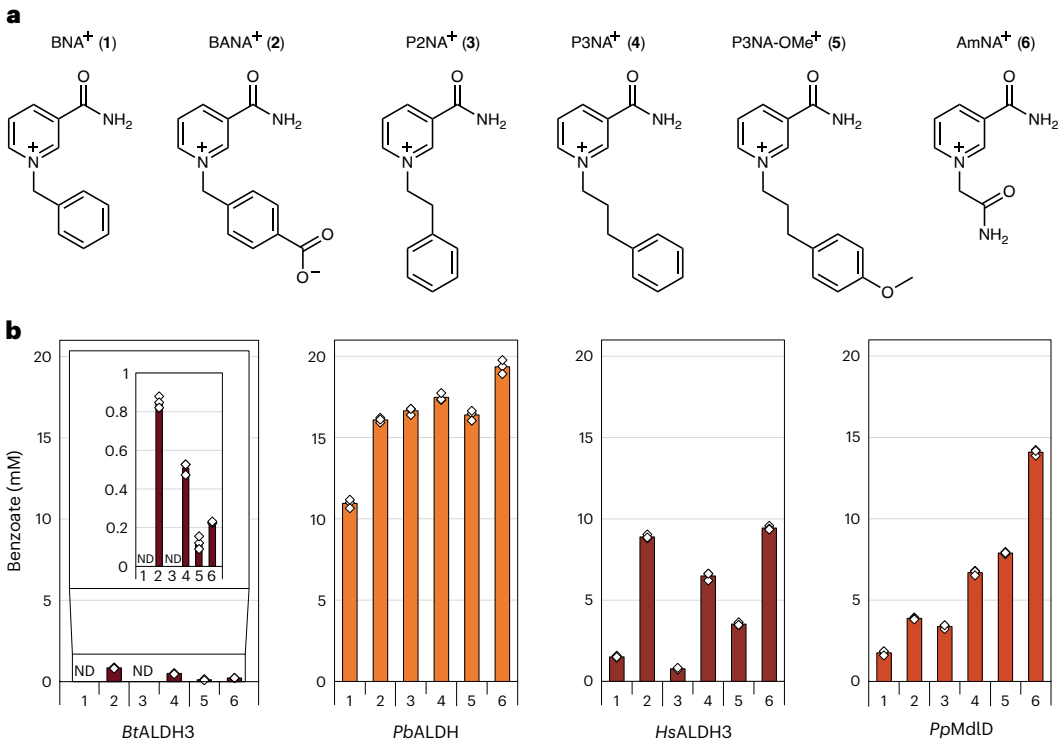

Fig. 5 | RH/QxxR motif enables NRC activity with nonnucleotide, simple-synthetic NRCs. a, The simple NRCs BNA$^+$, BANA$^+$, P2NA$^+$, P3NA$^+$, P3NA-OMe$^+$ and AmNA$^+$ are readily synthesized and allow functional probing of cofactor promiscuity in ALDHs. **b**, In a conversion assay, *Bt*ALDH3a1 (*Bt*ALDH3) only shows less than 1 mM conversion of benzaldehyde to benzoate with simple-synthetic NRCs over 24 h. However, *Pb*ALDH shows much higher levels of benzaldehyde conversion with simple-synthetic NRCs, achieving greater than 50% conversion with all simple-synthetic NRCs tested and near-complete conversion with AmNA$^+$ (19.4 ± 0.4 mM, 97% conversion). *Hs*ALDH3 and *Pp*MdlD, which also carry the RH/QxxR motif, showed intermediate levels of conversion, generally preferring the more polar cofactors BANA$^+$, P3NA-OMe$^+$ and AmNA$^+$. All reactions were carried out and reported as the mean of $n = 3$ biologically independent replicates. ND, no benzoate detected.

RH/QxxR motif is transferable to diverse ALDHs and NRCs

To test the extent to which the RH/QxxR motif is self-sufficient in enabling NMN$^+$ activity, we transferred RH/QxxR onto ALDHs lacking the motif. For this investigation, we selected all the ALDHs from our initial colorimetric screen that showed greater than 50 nmol formazan mg^{-1} min^{-1} of activity toward NAD$^+$ and any detectable level of activity with NMN$^+$. These ALDHs originate from diverse regions of sequence space within the ALDH SSN and are phylogenetically distinct (Fig. 4a). The corresponding RH/QxxR variants of these ALDHs were cloned, expressed and purified, apart from the ALDH-027 K337R;F338H;W341R mutant, which could not be purified. All but one (ALDH-014) of the RH/QxxR variants of these ALDHs showed significant improvements in activity toward NMN$^+$, achieving up to ~23 fold improvement in specific activity (Fig. 4b). These improvements do not appear to correlate with phylogenetic proximity or sequence identity to *Bt*ALDH3, indicating that the role of this motif in improving NMN$^+$ activity is not constrained to the evolutionary trajectory of *Bt*ALDH3. The broad success in translating the RH/QxxR motif demonstrates the generalizable and central role of the motif for NMN$^+$ use in ALDHs. The lack of improvement observed in ALDH-014 may be attributed to its unconserved structure. AlphaFold models of ALDH-014 indicate it lacks the highly conserved oligomerization domain widely observed in ALDHs and carries a unique N-terminal tail. These deviances in its structure may make the enzyme unamenable to mutagenesis. Moreover, given the nonuniform improvement in activity observed across ALDHs, there may be additional undiscovered features that act alongside the RH/QxxR motif to enhance NRC activity. Nonetheless, transferring the RH/QxxR motif onto ALDHs appears to be a facile approach for enhancing NRC activity in these enzymes.

From these ALDH mutants, we chose ALDH-013 RHR, ALDH-137 RHR and ALDH-037 RHR for further kinetic characterization. ALDH-013 RHR, ALDH-037 RHR and ALDH-137 RHR showed 2-fold, 60-fold and 11-fold increases in catalytic efficiencies toward NMN$^+$, respectively, compared to their corresponding WT (Table 1). These improvements primarily stem from reductions in their K_M, which supports that the RH/QxxR motif enhances NMN$^+$ binding by rigidifying and preorganizing the cofactor-binding pocket. Accordingly, the RHR mutants did not enhance catalytic efficiency toward NAD$^+$.

While NMN$^+$ has many advantages as an NRC, other nonnucleotide, simple-synthetic cofactors have also garnered interest because of their facile synthesis, reduced cost and enhanced redox potential[7,10–12,21–24,53]. We theorized that the RH/QxxR motif may also stabilize the binding of other NRCs, which vary in size and polarity, but all share the nicotinamide head with NMN$^+$ that the aromatic lid packs against (Fig. 5a). In addition to testing *Bt*ALDH3, *Pb*ALDH and *Hs*ALDH3 we chose to test the well-characterized class 3 ALDH from *P. putida* (*Pp*MdlD), which also contains the RH/QxxR motif (R290, Q291 and R294). We synthesized six simple-synthetic cofactors: BNA$^+$, BANA$^+$, 1-phenethylnicotinamide (P2NA$^+$), 1-(3-phenylpropyl)nicotinamide(P3NA$^+$), P3NA-OMe$^+$ and AmNA$^+$. Next, we performed a biotransformation assay over the course of 24 h and measured the cofactor-dependent conversion of benzaldehyde to benzoate mediated by the four chosen ALDHs. Excitingly, these ALDHs showed conversion using simple-synthetic NRCs (Fig. 5b). *Pb*ALDH showed the highest level of conversion, achieving >10 mM (>50%) conversion with all cofactors and 19.4 ± 0.4 mM (97%) conversion with AmNA$^+$. Similarly, *Pp*MdlD showed conversion with all cofactors tested, achieving up to 14.1 ± 0.2 mM (70%) conversion with AmNA$^+$. Under these assay conditions, all four enzymes showed complete conversion of benzaldehyde in under 10 min with NAD$^+$, or NMN$^+$.

*Bt*ALDH3 did not show detectable conversion with BNA^+ or $P2NA^+$ and only achieved a maximum of 4% conversion with $BANA^+$. On the other hand, *Hs*ALDH3, which shares 83% sequence identity with *Bt*ALDH3, displayed substantially higher conversion with a broader scope of cofactors, reaching 8.9 ± 0.1 mM (45%) conversion with $BANA^+$ and up to 9.4 ± 0.1 mM (47%) conversion with $AmNA^+$ (Fig. 5b). This is likely to be understood in the context of the relatively inferior catalytic rate constant of *Bt*ALDH3 ($k_{cat,\ NAD+}$ = 1.4 ± 0.1 s^{-1}) compared to *Hs*ALDH3 ($k_{cat,\ NAD+}$ = 14.1 ± 0.4 s^{-1}) (Supplementary Table 6). This trend persists for other high-conversion ALDHs *Pb*ALDH and *Pp*MdlD, both having higher $k_{cat,\ NAD+}$ than *Bt*ALDH3 (Table 1 and Supplementary Table 6). While the RH/QxxR motif is sufficient to permit the use of NRCs, overall conversion efficiency is expected to depend on the intrinsic rate at which the enzyme samples catalytically competent states[57–59]. Therefore, k_{cat} is influenced not only by the active site but also by the global conformational dynamics of the enzyme, which differs even among closely related homologs. Consistent with this view, the RH/QxxR motif-containing ALDHs exhibiting higher k_{cat} values also display higher conversion with NRCs. Nonetheless, our results suggest that the RH/QxxR motif can serve as a sequence indicator for searching NRC-dependent activity in ALDH enzymes beyond NMN^+. However, the magnitude of this activity, particularly the conversion rate in biotransformation under relatively high cofactor concentrations, is governed by catalytic rate and the innate conformational dynamics of enzymes.

During the preparation of this manuscript, a parallel effort was published demonstrating that an ALDH from *Sphingobium* sp. can also use simple-synthetic cofactors (Supplementary Fig. 7)[60]. *Sp*ALDH2 also contains the RH/QxxR motif (R343, Q344, R347), which represents an independent corroboration of our findings.

Discussion

In this work, we used a DBTL approach, iteratively leveraging the natural sequence space to discover and engineer ALDHs toward NMN^+ and simple-synthetic NRCs. While previous studies have focused on engineering activity toward NRCs from the ground up on the basis of a particular scaffold[8,9,11–14,16–20,23,24,38–40], we systematically screened through the ALDH pFAM and identified naturally occurring NMN^+-using ALDHs. Among these, *Bt*ALDH3 and *Pb*ALDH are ~10^1–10^5-fold better than all other natural or engineered enzymes previously reported, as measured by the ratio between NMN^+ catalytic efficiency and WT parent's native $NAD(P)^+$ catalytic efficiency. A single exception to this is PTDH LY13, which reaches similar level of relative catalytic proficiency to *Bt*ALDH3 and *Pb*ALDH for NMN^+ (Fig. 1d). Interestingly, PTDH LY13 was obtained through a cell-fitness-based in vivo directed evolution campaign[19]. This selection parallels our search through natural evolutionary sequence space, albeit at a much smaller scale and depth.

Using a systems approach, we identified a distinct sequence motif that determines NRC use in natural enzymes. Although *Bt*ALDH3 and *Pb*ALDH only share 46% sequence identity and are phylogenetically distant, we were able to use SSNs to isolate their relationship in sequence space and distill the RH/QxxR sequence motif. This motif is not only validated to be essential in *Bt*ALDH3 for NMN^+ activity but can also be translated to other unrelated ALDHs to impart NMN^+ activity. Biochemical, structural and computational studies elucidated that RH/QxxR motif functions through rigidifying a loop in the active site to more firmly support the hydride transfer portion of cofactors, without discriminating different 'recognition handles' of NRCs. In line with these findings, we also successfully mined nonnucleotide, simple-synthetic NRC activities in natural ALDHs guided by the presence of RH/QxxR.

The RH/QxxR motif's mechanism can be classified as active site preorganization, which is more optimized in natural enzymes than engineered enzymes; natural enzymes have active site residues pointing in catalytically competent directions more frequently than engineered enzymes[15,61]. Active site preorganization requires more complex and integrated interaction networks, as exemplified here by *Bt*ALDH3, which is challenging to rationally design, although the latest ensemble modeling and deep learning methods may contribute new solutions to this challenge[61,62]. Directed evolution, which emulates natural evolution in microcosm, has been shown to improve enzyme function by narrowing enzyme's conformational ensemble to enrich for catalytically competent states[63]. Through studying the mechanisms of the RH/QxxR motif, it appears that NRC-using ALDHs feature a narrowly defined active site shaped by evolution.

The prevalence of notable NMN^+ activity in natural enzymes raises a fundamental question. Is this activity a byproduct of nonphysiological promiscuity or does it indicate some physiological relevance? One can find partial support for either hypothesis. The RH/QxxR motif appears to be a highly conserved feature of dimeric ALDHs[35,42,43]. It is possible that the motif evolved to compensate for stabilizing interactions normally mediated by additional monomers in higher-order oligomeric assemblies. As a side product, the stabilized active site gained promiscuous activity toward NRCs.

On the other hand, although adenosine is suggested to be the product of an RNA world that was fossilized in contemporary cofactors, Nature has exploited both adenylated and nonadenylated versions of cofactors such as flavin adenine dinucleotide and flavin mononucleotide for flavins or adenosine triphosphate (ATP) and pyrophosphate in energy metabolism[64,65]. As the AMP moiety of the cofactor is added at the cost of ATP, evidence suggests that the simpler, nonadenylated cofactors may function under stress or energy-constrained conditions[66,67]. The RH/QxxR motif is enriched in class 3 ALDHs in higher organisms. Members of class 3 ALDHs such as *Bt*ALDH3 and *Hs*ALDH3 are primarily located in the eyes, an organ that is particularly vulnerable to oxidative stress because of constant irradiation[35,68]. Indeed, high oxidative stress can impair NAD^+ biosynthesis in ocular tissues through aging and disease, which leads to the accumulation of NAD^+ precursors such as NMN^+, creating a biochemical context in which cofactor availability rather than enzyme abundance may become rate limiting[69]. In this context, the ability of a dominant oxidoreductase to tolerate or use a nonadenylated NAD^+ precursor may confer a functional advantage. While NMN^+ is present at notable levels in humans and other organisms as an NAD^+ biosynthesis precursor, we emphasize that our results do not establish in vivo use of NMN^+ as a dominant physiological pathway; more studies are certainly required to draw any conclusion on whether this nonadenylated version of NAD^+ is actually used as a redox cofactor in vivo by naturally occurring enzymes such as those identified in this work[70–72].

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-026-02315-w.

References

1. Butler, N. D., Anderson, S. R., Dickey, R. M., Nain, P. & Kunjapur, A. M. Combinatorial gene inactivation of aldehyde dehydrogenases mitigates aldehyde oxidation catalyzed by E. coli resting cells. *Metab. Eng.* **77**, 294–305 (2023).
2. Montaño López, J., Duran, L. & Avalos, J. L. Physiological limitations and opportunities in microbial metabolic engineering. *Nat. Rev. Microbiol.* **20**, 35–48 (2022).
3. Bekiaris, P. S. & Klamt, S. Network-wide thermodynamic constraints shape NAD(P)H cofactor specificity of biochemical reactions. *Nat. Commun.* **14**, 4660 (2023).
4. Rollin, J. A., Tam, T. K. & Percival Zhang, Y.-H. New biotechnology paradigm: cell-free biosystems for biomanufacturing. *Green Chem.* **15**, 1708–1719 (2013).

5. Chánique, A. M. & Parra, L. P. Protein engineering for nicotinamide coenzyme specificity in oxidoreductases: attempts and challenges. *Front. Microbiol.* **9**, 194 (2018).
6. Weusthuis, R. A., Folch, P. L., Pozo-Rodríguez, A. & Paul, C. E. Applying non-canonical redox cofactors in fermentation processes. *iScience* **23**, 101471 (2020).
7. Knaus, T. et al. Better than nature: nicotinamide biomimetics that outperform natural coenzymes. *J. Am. Chem. Soc.* **138**, 1033–1039 (2016).
8. Wang, X. et al. Creating enzymes and self-sufficient cells for biosynthesis of the non-natural cofactor nicotinamide cytosine dinucleotide. *Nat. Commun.* **12**, 2116 (2021).
9. Campbell, E., Meredith, M., Minteer, S. D. & Banta, S. Enzymatic biofuel cells utilizing a biomimetic cofactor. *Chem. Commun.* **48**, 1898–1900 (2012).
10. Platt, A. P., Klem, H., Mallinson, S. J., Bomble, Y. J. & Paton, R. S. Computer-aided design of stability enhanced nicotinamide cofactor biomimetics for cell-free biocatalysis. *Green Chem.* **27**, 6831–6844 (2025).
11. Gu, X., Zhou, J., Wang, H., Zhu, Y. & Ni, Y. Engineering glucose dehydrogenase from *Sulfolobus sulfataricus* toward utilizing diverse nicotinamide cofactor biomimetics. *Green Synth. Catal.* https://doi.org/10.1016/j.gresc.2025.03.008 (2025).
12. Zhu, Y. et al. Engineering a newly identified alcohol dehydrogenase from *Sphingobium* sp. for efficient utilization of nicotinamide cofactors biomimetics. *Bioresour. Bioprocess.* **12**, 41 (2025).
13. Huang, R., Chen, H., Upp, D. M., Lewis, J. C. & Zhang, Y. H. P. J. A high-throughput method for directed evolution of NAD(P)$^+$-dependent dehydrogenases for the reduction of biomimetic nicotinamide analogues. *ACS Catal.* **9**, 11709–11719 (2019).
14. King, E., Maxel, S. & Li, H. Engineering natural and noncanonical nicotinamide cofactor-dependent enzymes: design principles and technology development. *Curr. Opin. Biotechnol.* **66**, 217–226 (2020).
15. Mak, W. S. & Siegel, J. B. Computational enzyme design: transitioning from catalytic proteins to enzymes. *Curr. Opin. Struct. Biol.* **27**, 87–94 (2014).
16. Black, W. B. et al. Engineering a nicotinamide mononucleotide redox cofactor system for biocatalysis. *Nat. Chem. Biol.* **16**, 87–94 (2020).
17. Aspacio, D. et al. Shifting redox reaction equilibria on demand using an orthogonal redox cofactor. *Nat. Chem. Biol.* **20**, 1535–1546 (2024).
18. Richardson, K. N., Black, W. B. & Li, H. Aldehyde production in crude lysate-and whole cell-based biotransformation using a noncanonical redox cofactor system. *ACS Catal.* **10**, 8898–8903 (2020).
19. Zhang, L. et al. Directed evolution of phosphite dehydrogenase to cycle noncanonical redox cofactors via universal growth selection platform. *Nat. Commun.* **13**, 5021 (2022).
20. Li, Q. et al. Stoichiometric regeneration of biomimetic nicotinamide coenzyme powered by biomass sugars via in vitro synthetic enzymatic biosystems. *ChemSusChem* **18**, e202401263 (2025).
21. Nowak, C. et al. Characterization of biomimetic cofactors according to stability, redox potentials, and enzymatic conversion by NADH oxidase from *Lactobacillus pentosus*. *ChemBioChem* **18**, 1944–1949 (2017).
22. Paul, C. E. et al. Nonenzymatic regeneration of styrene monooxygenase for catalysis. *ACS Catal.* **5**, 2961–2965 (2015).
23. Ryan, J. D., Fish, R. H. & Clark, D. S. Engineering cytochrome P450 enzymes for improved activity towards biomimetic 1, 4-NADH cofactors. *ChemBioChem* **9**, 2579–2582 (2008).
24. Drenth, J., Yang, G., Paul, C. E. & Fraaije, M. W. A tailor-made deazaflavin-mediated recycling system for artificial nicotinamide cofactor biomimetics. *ACS Catal.* **11**, 11561–11569 (2021).
25. Sicsic, S., Durand, P., Langrene, S. & le Goffic, F. A new approach for using cofactor dependent enzymes: example of alcohol dehydrogenase. *FEBS Lett.* **176**, 321–324 (1984).
26. Hu, X., Zhang, J. & Rydström, J. Interactions of reduced and oxidized nicotinamide mononucleotide with wild-type and αD195E mutant proton-pumping nicotinamide nucleotide transhydrogenases from *Escherichia coli*. *Biochim. Biophys. Acta* **1367**, 134–138 (1998).
27. Shortall, K., Djeghader, A., Magner, E. & Soulimane, T. Insights into aldehyde dehydrogenase enzymes: a structural perspective. *Front. Mol. Biosci.* **8**, 659550 (2021).
28. Steffler, F., Guterl, J. K. & Sieber, V. Improvement of thermostable aldehyde dehydrogenase by directed evolution for application in synthetic cascade biomanufacturing. *Enzyme Microb. Technol.* **53**, 307–314 (2013).
29. Park, Y. S. et al. Engineering an aldehyde dehydrogenase toward its substrates, 3-hydroxypropanal and NAD$^+$, for enhancing the production of 3-hydroxypropionic acid. *Sci. Rep.* **7**, 17155 (2017).
30. Chen, Z., Wang, C., Wang, R., Li, L. & Du, J. Engineering of an aldehyde dehydrogenase from *Neurospora crassa* for crocetin production. *J. Agric. Food Chem.* **73**, 13710–13721 (2025).
31. Kanter, J. P. et al. An enzymatic tandem reaction to produce odor-active fatty aldehydes. *Appl. Microbiol. Biotechnol.* **106**, 6095–6107 (2022).
32. Zhou, J. et al. Three multi-enzyme cascade pathways for conversion of C1 to C2/C4 compounds. *Chem. Catalysis* **2**, 2675–2690 (2022).
33. Bar-Even, A. et al. The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* **50**, 4402–4410 (2011).
34. Black, W. B., Perea, S. & Li, H. Design, construction, and application of noncanonical redox cofactor infrastructures. *Curr. Opin. Biotechnol.* **84**, 103019 (2023).
35. Jackson, B. et al. Update on the aldehyde dehydrogenase gene (ALDH) superfamily. *Hum. Genomics* **5**, 283 (2011).
36. Suzek, B. E., Wang, Y., Huang, H., McGarvey, P. B., Wu, C. H. & The UniProt Consortium. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31**, 926–932 (2015).
37. Zallot, R., Oberg, N. & Gerlt, J. A. The EFI web resource for genomic enzymology tools: leveraging protein, genome, and metagenome databases to discover novel enzymes and metabolic pathways. *Biochemistry* **58**, 4169–4182 (2019).
38. Aspacio, D. et al. Engineering *Escherichia coli* pyruvate metabolism to generate noncanonical reducing power. *ACS Catal.* **14**, 9776–9784 (2024).
39. King, E. et al. Engineering Embden–Meyerhof–Parnas glycolysis to generate noncanonical reducing power. *ACS Catal.* **12**, 8582–8592 (2022).
40. Chen, Y. L. et al. Rational engineering of 3α-hydroxysteroid dehydrogenase/carbonyl reductase for a biomimetic nicotinamide mononucleotide cofactor. *Catalysts* **12**, 1094 (2022).
41. Liu, Z. J. et al. The first structure of an aldehyde dehydrogenase reveals novel interactions between NAD and the Rossmann fold. *Nat. Struct. Biol.* **4**, 317–326 (1997).
42. Khanna, M. et al. Discovery of a novel class of covalent inhibitor for aldehyde dehydrogenases. *J. Biol. Chem.* **286**, 43486–43494 (2011).
43. Zahniser, M. P. et al. Structure and mechanism of benzaldehyde dehydrogenase from *Pseudomonas putida* ATCC 12633, a member of the class 3 aldehyde dehydrogenase superfamily. *Protein Eng. Des. Sel.* **30**, 273–280 (2017).

44. Perez-Miller, S. J. & Hurley, T. D. Coenzyme isomerization is integral to catalysis in aldehyde dehydrogenase. *Biochemistry* **42**, 7100–7109 (2003).
45. Huo, L. et al. Crystallographic and spectroscopic snapshots reveal a dehydrogenase in action. *Nat. Commun.* **6**, 5935 (2015).
46. d'Ambrosio, K. et al. The first crystal structure of a thioacylenzyme intermediate in the ALDH family: new coenzyme conformation and relevance to catalysis. *Biochemistry* **45**, 2978–2986 (2006).
47. Boehr, D. D. et al. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642 (2006).
48. Zou, Y. et al. Crystal structures of phosphite dehydrogenase provide insights into nicotinamide cofactor regeneration. *Biochemistry* **51**, 4263–4270 (2012).
49. Cristobal, J. R., Nagorski, R. W. & Richard, J. P. Utilization of cofactor binding energy for enzyme catalysis: formate dehydrogenase-catalyzed reactions of the whole NAD cofactor and cofactor pieces. *Biochemistry* **62**, 2314–2324 (2023).
50. Colonna-Cesari, F. et al. Interdomain motion in liver alcohol dehydrogenase. Structural and energetic analysis of the hinge bending mode. *J. Biol. Chem.* **261**, 15273–15280 (1986).
51. Price, A. C., Zhang, Y. M., Rock, C. O. & White, S. W. Cofactor-induced conformational rearrangements establish a catalytically competent active site and a proton relay conduit in FabG. *Structure* **12**, 417–428 (2004).
52. Skarżyński, T. & Wonacott, A. J. Coenzyme-induced conformational changes in glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *J. Mol. Biol.* **203**, 1097–1118 (1988).
53. King, E. et al. Orthogonal glycolytic pathway enables directed evolution of noncanonical cofactor oxidase. *Nat. Commun.* **13**, 7282 (2022).
54. Zhu, J. Y. et al. Functional consequence of covalent reaction of phosphoenolpyruvate with UDP-*N*-acetylglucosamine 1-carboxyvinyltransferase (MurA). *J. Biol. Chem.* **287**, 12657–12667 (2012).
55. Park, J. & Rhee, S. Structural basis for a cofactor-dependent oxidation protection and catalysis of cyanobacterial succinic semialdehyde dehydrogenase. *J. Biol. Chem.* **288**, 15760–15770 (2013).
56. Kimble-Hill, A. C. et al. Development of selective inhibitors for aldehyde dehydrogenases based on substituted indole-2,3-diones. *J. Med. Chem.* **57**, 714–722 (2014).
57. Henzler-Wildman, K. A. et al. Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838–844 (2007).
58. Schwartz, S. D. Protein dynamics and enzymatic catalysis. *J. Phys. Chem. B* **127**, 2649–2660 (2023).
59. Boehr, D. D., Nussinov, R. & Wright, P. E. The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* **5**, 789–796 (2009).
60. Zhou, J., Gu, X., Chen, S., Zhu, Y. & Ni, Y. Biooxidation of aromatic aldehydes by aldehyde dehydrogenase from *Sphingobium* sp. using synthetic nicotinamide cofactor biomimetics. *Green Chem. Eng.* **7**, 245–250 (2026).
61. Lauko, A. et al. Computational design of serine hydrolases. *Science* **388**, eadu2454 (2025).
62. Guo, A. B. et al. Deep learning-guided design of dynamic proteins. *Science* **388**, eadr7094 (2025).
63. Otten, R. et al. How directed evolution reshapes the energy landscape in an enzyme to boost catalysis. *Science* **370**, 1442–1446 (2020).
64. Macheroux, P., Kappes, B. & Ealick, S. E. Flavogenomics—a genomic and structural view of flavin-dependent proteins. *FEBS J.* **278**, 2625–2634 (2011).
65. Heinonen, J. K. *Biological Role of Inorganic Pyrophosphate* 1st edn (Springer, 2012).
66. Igamberdiev, A. U. & Kleczkowski, L. A. Pyrophosphate as an alternative energy currency in plants. *Biochem. J* **478**, 1515–1524 (2021).
67. Baykov, A. A., Malinen, A. M., Luoto, H. H. & Lahti, R. Pyrophosphate-fueled Na^+ and H^+ transport in prokaryotes. *Microbiol. Mol. Biol. Rev.* **77**, 267–276 (2013).
68. Lindahl, R. & Petersen, D. R. Lipid aldehyde oxidation as a physiological role for class 3 aldehyde dehydrogenases. *Biochem. Pharmacol.* **41**, 1583–1587 (1991).
69. Liu, S. & Zhang, W. NAD^+ metabolism and eye diseases: current status and future directions. *Mol. Biol, Rep.* **50**, 8653–8663 (2023).
70. Imai, S. I. NAD World 3.0: the importance of the NMN transporter and eNAMPT in mammalian aging and longevity control. *npj Aging* **11**, 4 (2025).
71. Sorci, L. et al. Nicotinamide mononucleotide synthetase is the key enzyme for an alternative route of NAD biosynthesis in *Francisella tularensis*. *Proc. Natl Acad. Sci. USA* **106**, 3083–3088 (2009).
72. Gazzaniga, F., Stebbins, R., Chang, S. Z., McPeek, M. A. & Brenner, C. Microbial NAD metabolism: lessons from comparative genomics. *Microbiol. Mol. Biol. Rev.* **73**, 529–541 https://doi.org/10.1128/mmbr.00042-08 (2009).

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

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

Methods

Generation of ALDH SSN

SSNs were generated using the EFI-EST webtool supplied with the ALDH pFAM (PF00171) as the input query, filtered by UniRef50 sequences[37]. UniRef50 clusters are clustered groups of unique sequences from the UniProtKB database sharing 50% sequence identity and 80% sequence length overlap, which are assigned a representative sequence by choosing the highest quality UniProt entry in the cluster, sequences deriving from model organisms or by assigning the longest sequence as the representative[36]. Generated networks were visualized with Cytoscape. The SSN presented in Fig. 1a is arranged using the edge-weighted spring-embedded layout, where the distance between connected nodes in a given subnetwork is approximately inversely correlated with their alignment score. The distance between unconnected nodes and subnetworks is arbitrary. The refined *Bt*ALDH3 subnetwork (Fig. 2a) was generated in Cytoscape by selecting *Bt*ALDH3 and *Pb*ALDH, extracting their nearest neighbors using the built-in selection tool and exporting the connected subset as a new network. In the refined SSN presented in Fig. 2a, nodes are arranged using a circular layout for visual ease; the spatial positions of nodes and edge lengths do not correspond to sequence similarity.

Plasmid construction

All plasmids used in this study are listed in Supplementary Table 1, with corresponding nucleotide and amino acid sequences in Supplementary Table 2. To clone the gene encoding *Geobacillus* sp. diaphorase (*Gs*DI) and genes encoding ALDHs used in this study, their DNA sequences were first codon-optimized for expression in *E. coli* and their corresponding gene fragments were obtained as synthesized oligonucleotides (Integrated DNA Technologies). Synthesized gene fragments were inserted into the pQElac vector (PLlacO1, ColE1 ori, AmpR, N-terminal 6×His tag) using the Gibson isothermal DNA assembly method. To isolate plasmids for sequence verification, XL1-Blue competent cells (Agilent Technologies) were transformed with the Gibson assembly products, plated on 2×YT agar plates (16 g L^{-1} tryptone, 10 g L^{-1} yeast extract and 5 g L^{-1} NaCl) containing 100 mg L^{-1} ampicillin and incubated at 37 °C overnight. Single colonies were chosen from each construct, inoculated into 5 ml of 2×YT liquid medium with 100 mg L^{-1} ampicillin and shaken overnight at 37 °C. Plasmid DNA was then extracted from each culture using the QIAprep Spin miniprep kit (Qiagen) and sequence-verified by Sanger sequencing. pET28a-His-Smt3-*Bt*ALDH3 was similarly constructed using pET28a-His-Smt3 (T7:: BtALDH3, pBR322 ori, KanR, N-terminal 10×His–Smt3 tag) vector in place of pQElac.

Variants of *Bt*ALDH3, ALDH-013, ALDH-037 and ALDH-137 were generated through site-directed mutagenesis by PCR using the KOD1 PCR master mix (Toyobo). Mutated gene fragments were amplified using primers containing point mutations (Integrated DNA Technologies). Gene fragments were assembled into pQElac, isolated and sequence-verified as described above. All unique plasmid constructs presented in this study are available from the corresponding authors upon request.

Protein expression and purification

The ALDHs characterized in this study were expressed in *E. coli* BL21(DE3) (Invitrogen) transformed with the respective pQElac ALDH plasmid. To begin, single colonies of *E. coli* BL21(DE3), each carrying an ALDH expression plasmid, were inoculated into individual 10-ml culture tubes containing 5 ml of 2×YT liquid medium containing 100 mg L^{-1} ampicillin and shaken overnight at 30 °C. Next, 50 ml of 2×YT medium supplemented with 200 mg L^{-1} ampicillin was inoculated with 0.5 ml of overnight culture in a 250-ml baffled flask. Cultures were grown at 37 °C, 250 rpm until reaching an optical density of ~0.6–1, then induced with 0.5 mM IPTG and incubated at room temperature. After 24 h of expression, the cultures were transferred to 50-ml Falcon tubes and centrifuged at 4,000*g*, 4 °C. The supernatant was discarded and the cell pellet was resuspended in 1 ml of ice-cold equilibration buffer (50 mM sodium phosphate buffer pH 7.7, 300 mM sodium chloride, 10 mM imidazole and 0.03% Triton X-100) and transferred to a prechilled 2-ml bead-beating tube containing 0.5 ml of 0.1-mm-diameter soda lime glass beads (BioSpec Products) and stored on ice. Each resuspended cell pellet was then homogenized on a benchtop homogenizer (FastPrep-24, MP Biomedicals) at 4 m s^{-1} for 35 s and cooled on ice for 5 min. This process was repeated twice for three total homogenization steps. The crude lysate was separated from the cell debris by centrifugation at 21,000*g*, 4 °C and purified using HisPur Ni-NTA protein miniprep (Thermo Fisher Scientific) according to the manufacturer's instructions. All enzymes were eluted in protein elution buffer (50 mM sodium phosphate buffer pH 7.7, 300 mM sodium chloride and 250 mM imidazole) and stored in 20% (v/v) glycerol at −80 °C. *Gs*DI was similarly expressed and purified, with the only deviation being postinduction incubation at 30 °C rather than room temperature. Purified protein yields for ALDHs are presented in Supplementary Table 3 and representative SDS–PAGE gels of *Bt*ALDH3, *Pb*ALDH and *Pp*MdlD are presented in Supplementary Fig. 8.

High-throughput colorimetric cycling assay

The high-throughput colorimetric cycling assay (Fig. 1c and Supplementary Table 3) was carried out as follows. A 100-µl reaction was prepared containing 100 mM potassium phosphate buffer pH 7.4, 5 mM aldehyde (acetaldehyde, butyraldehyde or hexanal), 0.5 mM WST-1 tetrazolium dye, 3 mM cofactor (NAD^+, NMN^+ or no cofactor) and 1 µM of purified *Gs*DI, along with varying concentrations of purified ALDH (0.01–0.5 mg ml^{-1}). Reactions were initiated by the addition of ALDH to reaction buffer and carried out in 96-well plates at 30 °C. The formation of WST-1 formazan was monitored kinetically at 438 nm ($\varepsilon_{\text{WST-1 formazan}} = 37 \times 10^3\ M^{-1}\ cm^{-1}$) and activity is reported as nmol of WST-1 formazan generated per mg of ALDH per min. To ensure that the NMN^+ activity measured was always cofactor dependent and not as a consequence of native cofactor contamination, colorimetric assays were carried out in parallel where cofactor was substituted with deionized (DI) water in the reaction mixture and the background activity was measured. Enzyme concentrations were titrated for each ALDH to achieve a signal-to-noise ratio of at least 10:1 with NAD^+. The limit of detection was determined to be 0.1 nmol formazan formed per min. All colorimetric activity data were controlled by subtracting any activity observed under no-cofactor conditions. All reactions were carried out with $n = 2$ biologically independent replicates.

Specific activity assays

The specific activity assays (Figs. 2b, 3a and 4b) were carried out as follows. A 200-µl reaction was prepared containing 100 mM potassium phosphate buffer pH 7.4, 5 mM aldehyde (acetaldehyde, butyraldehyde or hexanal), 5 mM dithiothreitol and 3 mM cofactor (NAD^+, NMN^+ or no cofactor), along with varying concentrations of purified ALDH (0.01–0.5 mg ml^{-1}). Reactions were initiated by the addition of ALDH to reaction buffer and carried out in 96-well plates at 30 °C. The formation of reduced cofactor was monitored kinetically at 340 nm ($\varepsilon_{\text{NADH/NMNH}} = 6.22 \times 10^3\ M^{-1}\ cm^{-1}$). To ensure that the NMN^+ activity measured was always cofactor dependent and not as a consequence of native cofactor contamination, specific activity assays were carried out in parallel where cofactor was substituted with DI water in the reaction mixture and the background activity was measured. All specific activity data were controlled by subtracting any activity observed under no-cofactor conditions. All reactions were carried out with $n = 3$ biologically independent replicates.

Steady-state kinetic characterization

The steady-state kinetic characterization of ALDHs was performed as follows. A 200-µl reaction was prepared containing 100 mM potassium phosphate buffer pH 7.4, 5 mM hexanal, 5 mM dithiothreitol and

varying concentrations of NAD^+ or NMN^+ (0.005–100 mM), along with varying concentrations of purified ALDH (0.01–0.5 mg ml^{-1}). Reactions were initiated by the addition of ALDH to reaction buffer and carried out in 96-well plates at 30 °C. The formation of reduced cofactor was monitored kinetically at 340 nm ($\varepsilon_{NADH/NMNH} = 6.22 \times 10^3\ M^{-1}\ cm^{-1}$). The resulting specific activity data were plotted and fitted to the standard Michaelis–Menten equation. In cases where the enzyme could not be saturated with 100 mM cofactor, it was assumed that $K_m > 100$ mM; therefore, the kinetic parameters were determined using the simplified Michaelis–Menten equation ($V_o \approx k_{cat}[E][S]/K_m$). All reactions were carried out with $n = 3$ biologically independent replicates.

Protein expression and purification for crystallography of *Bt*ALDH3

The plasmids pET28a-His-smt3-*Bt*ALDH3 and pET28a-His-smt3-*Bt*ALDH3-H290L were transformed into *E. coli* BL21(DE3) cells. Protein expression was induced in log-phase cultures with 0.2 mM IPTG for 16 h at 16 °C.

Cells were harvested at 6,000 rpm for 10 min and resuspended in lysis buffer containing 25 mM Tris-HCl pH 8.0, 350 mM NaCl, 30 mM imidazole, 5 mM β-mercaptoethanol, 1 mM PMSF and 2 mM benzamidine. Cells were lysed by sonication and clarified by centrifugation. The supernatant was incubated with Ni-NTA resin (Qiagen) for 30 min at 4 °C. After washing, bound protein was eluted with buffer containing 300 mM imidazole. The His–Smt3 tag was removed by Ulp1 protease cleavage and the cleaved protein was further purified by ion-exchange chromatography.

Removal of carryover NAD^+ from recombinant *Bt*ALDH3 before crystallography

In our initial attempts to determine the crystal structure of *Bt*ALDH3 in complex with NMN^+, we observed minor electron density of NAD^+'s AMP moiety. To ensure no NAD^+ was carried into the NMN^+ cocrystal we performed the following procedure. To remove carryover NAD^+ while preserving overall fold integrity, purified *Bt*ALDH3 was diluted to 0.5–3.0 mg ml^{-1} (~10–60 μM) and subjected to mild denaturation by dialysis against 500 mM NaCl supplemented with 1.0–5.0 M urea for 6 h to overnight. This treatment disrupts cofactor interactions with the protein, thus facilitating dissociation of carryover NAD^+ without complete unfolding of the protein.

NMN^+ incorporation with recombinant *Bt*ALDH3 for crystallography

NMN^+ was introduced during refolding to occupy the cofactor-binding site. Protein samples were dialyzed against 500 mM NaCl containing 20–100 μM NMN^+ at 4 °C. Dialysis proceeded for 1–3 h with periodic mixing. The external buffer was then replaced with fresh 500 mM NaCl containing 20–100 μM NMN^+ and dialysis was continued overnight to complete refolding and cofactor incorporation. Refolded *Bt*ALDH3–NMN complexes were concentrated to ~10–15 mg ml^{-1}. The resulting sample was used directly for crystallization trials.

Crystallography of *Bt*ALDH3

Crystallization of *Bt*ALDH3 WT and H290L was performed by hanging-drop vapor diffusion at 22 °C. Crystals were obtained in two conditions: 190 mM magnesium chloride, 100 mM Tris-HCl pH 8.5 and 21% (w/v) PEG8000 or 200 mM magnesium formate, 100 mM Tris-HCl (pH 8.5) and 20% PEG3350. Both *Bt*ALDH3 WT and *Bt*ALDH3-H290L were cocrystallized with an approximately five fold molar excess of NAD^+ (760 μM) to enzyme (150 μM) or 500-fold molar excess of NMN^+ (73 mM) to enzyme.

Crystals were isolated, cryoprotected in mother liquor supplemented with glycerol and flash-frozen in liquid nitrogen for data collection. Diffraction data were collected to 1.4-Å resolution at NSLS II and processed with Mosflm[73]. Crystals of *Bt*ALDH3 (NAD^+) belonged to space group P1, with unit-cell parameters $a = 46.68$ Å, $b = 61.53$ Å, $c = 87.95$ Å and $\alpha = 94.01°$, $\beta = 100.94°$, $\gamma = 115.36°$. *Bt*ALDH3 (NMN^+) crystallized in space group P1 21 1, with $a = 94.17$ Å, $b = 59.28$ Å, $c = 158.64$ Å and $\alpha = 90°$, $\beta = 98.43°$, $\gamma = 90°$. *Bt*ALDH3-H290L crystallized in space group P1, with $a = 46.63$ Å, $b = 59.35$ Å, $c = 87.70$ Å and $\alpha = 77.22°$, $\beta = 79.81°$, $\gamma = 69.31°$. The structure was solved by molecular replacement using an AlphaFold-predicted *Bt*ALDH3 model. Three-dimensional structural figures were generated in PyMOL. Final data collection and refinement statistics are presented in Supplementary Table 5.

Molecular dynamics

Molecular dynamics simulations of apo *Bt*ALDH3 WT, single-mutant variants R289A, H290A, H290L and R293A and the AAA triple mutant were performed as follows. Coordinates for the WT and H290L variants were obtained from their crystal structures. The remaining variants were modeled from the WT *Bt*ALDH3 crystal structure using PyMOL. The apo structure was first generated by removing all ligands from its crystal structure. The system was solvated using in-house scripts in a rhombic dodecahedron with a 10-Å buffer on all sides for a periodic image distance of 120 Å. Simulations used the CHARMM36 force field for proteins and the TIP3P model for water. The box was solvated with 100 mM NaCl and protonation states were assigned at a pH of 7 by H++ and visual inspection, resulting in protonation of glutamate and histidine in different systems, as shown in Supplementary Table 7. Molecular dynamics simulations were conducted with GROMACS 2025.3 software using particle mesh Ewald for the treatment of long-range electrostatics with fourth-order interpolation, a grid length of 128 and $\kappa = 0.32\ Å^{-1}$. Lennard–Jones interactions used force switching with a switch radius of 9 Å and a cutoff of 10 Å. After minimization, both position-restrained simulations imposed constraints on the positions of heavy atoms in the system (under NPT and NVT ensembles). The integration time step was set to 2 fs. The LINCS algorithm was used for bond constraints and the leapfrog algorithm was used for the integration of motion equations. In the NVT ensemble, each system was gradually heated from 0 K to 298.15 K over 200 ps, followed by a 400-ps MD simulation under the NPT ensemble ($T = 298.15$ K, $P = 1$ bar). During both of these periods, the protein was restrained with a force constant of 1,000 kJ mol^{-1} nm^{-2}. The V-rescale thermostat was used for temperature control, while the Parrinello–Rahman barostat was used to control pressure. Subsequently, production simulations of each system were run in quintuplicate for 250 ns, which gave ten trajectories of cofactor pocket dynamics when considering the two halves of the dimer. The postsimulation data were extracted and analyzed using GROMACS analysis tools. For each trajectory, monomers were aligned to an average trajectory structure and r.m.s.f. values were computed for Cα atoms across residues 333–337. The r.m.s.f. values were first averaged across residues for each simulation and then across all ten trajectories to quantify residue-level flexibility. Pairwise statistical comparisons between r.m.s.f. values of ALDH variants were performed using a two-tailed Welch's *t*-test. Trajectories used for structural visualization were subsampled using GROMACS. A total of 30 random frames were sampled from the trajectory. Results were visualized with PyMOL.

Multiple-sequence alignment and phylogenetic analysis

Multiple-sequence alignments (Fig. 2c) and phylogenetic trees (Fig. 4a) were generated using the EMBL-EBI Clustal Omega webserver tool[74]. The phylogenetic tree shown in Fig. 4a was visualized using the Interactive Tree of Life phylogenetic tree visualizer web tool[75]. Sequence logo plots (Supplementary Table 4) were generated from the multiple-sequence alignments using the WebLogo sequence logo plot generator[76].

Synthesis of nonnucleotide NRCs

Synthesis of nonnucleotide NRCs was carried out using established methods[53]. Briefly, one equivalent of nicotinamide was dissolved with one equivalent of alkyl bromide in acetonitrile and heated under reflux overnight to yield a white precipitate. The precipitate was

filtered and dried under vacuum and analyzed by nuclear magnetic resonance (NMR). The alkyl bromides used for the synthesis of BNA^+, $BANA^+$, $P2NA^+$, $P3NA^+$, $P3NA\text{-}OMe^+$ and $AmNA^+$ were (bromomethyl) benzene, 4-(bromomethyl)benzoic acid, (2-bromoethyl)benzene, (3-bromopropyl)benzene, 1-(3-bromopropyl)-4-methoxybenzene and 2-bromoacetamide, respectively. NMR spectra and detailed reaction conditions are provided in the Supplementary Methods.

Nonnucleotide cofactor conversion assay

To assess the activity of ALDHs toward nonnucleotide cofactors, we initially attempted to measure activity in a spectrophotometric assay, measuring either WST-1 formazan in a colorimetric cycling assay or reduced cofactor formation in a specific activity assay (as described above). However, the activity of the tested ALDHs toward these simple-synthetic cofactors was too low to accurately quantify using these methods. Instead, we performed a simple conversion assay, measuring the conversion of benzaldehyde to benzoate (Fig. 5b). Benzaldehyde was chosen as the substrate as it is used by the tested ALDHs and its product is readily detectable by high-performance liquid chromatography (HPLC). The reactions were carried out as follows. A 20-µl reaction was prepared containing 100 mM Tris-Cl buffer pH 7, 20 mM benzaldehyde, 20 mM cofactor (NAD^+, BNA^+, $BANA^+$, $P2NA^+$, $P3NA^+$, $P3NA\text{-}OMe^+$, $AmNA^+$ or no cofactor) and 0.25 g L^{-1} purified ALDH. Reactions were initiated by the addition of ALDH to reaction buffer and carried out in PCR tubes at 37 °C for 24 h. After 24 h, reactions were arrested by the addition of 40 µl of a 2.5% formic acid and 50% methanol solution. Each sample was then filtered and analyzed by HPLC under the following chromatographic conditions. Separation was achieved on an Agilent 1200 system equipped with a diode array detector and Luna Omega PS C18 column (5 µm, 100 Å, 150 × 4.6 mm; Phenomenex) maintained at 50 °C. The mobile phase consisted of solvent A (water with 0.1% formic acid) and solvent B (50% methanol in 0.1% formic acid), run under isocratic conditions at 50% B. The flow rate was 2.0 ml min^{-1}, with an injection volume of 5 µl. Total run time was 8 min per sample. Detection of benzoate and benzaldehyde was carried out at 230 nm. Concentration of benzoate formed was determined by comparison to authentic standards. To ensure that the NRC activity measured was always cofactor dependent and not as a consequence of native cofactor contamination, assays were carried out in parallel where cofactor was substituted with DI water in the reaction mixture and the background conversion was measured. All conversion data were controlled by subtracting the conversion achieved under no-cofactor conditions. All reactions were carried out with $n = 3$ biologically independent replicates.

Material availability

All unique plasmid constructs presented in this study are available from the corresponding authors upon request.

Reporting summary

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

Data availability

All data supporting the findings of this work are available within the paper and its Supplementary Information. Crystal structures of *Bt*ALDH3 WT in complex with NAD^+ (PDB 9PXQ), *Bt*ALDH3 WT in complex with NMN^+ (PDB 9PXH) and *Bt*ALDH3 H290L (PDB 10PQ) were deposited to the Protein Data Bank. Molecular dynamics trajectories were deposited to Zenodo (https://doi.org/10.5281/zenodo.21229755)[77]. Source data are provided with this paper.

Code availability

The source code for molecular dynamics simulations of *Bt*ALDH3 WT and variants are available from GitHub (https://github.com/hanli-lab/BtALDH3_MD.git).

References

73. Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R. & Leslie, A. G. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 271–281 (2011).
74. Madeira, F. et al. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Res.* **50**, W276–W279 (2022).
75. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **49**, W293–W296 (2021).
76. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
77. Saleh, S. et al. A sequence motif enables widespread use of noncanonical redox cofactors in natural enzymes. *Zenodo* https://doi.org/10.5281/zenodo.21229755 (2026).

Acknowledgements

S.S., V.M., W.B.B, H.N., S.Z. and H.L. acknowledge funding from the Advanced Research Projects Agency–Energy EcoSynBio program (DE-AR0001508), the National Science Foundation (NSF) (2328145), the National Institutes of Health (NIH) (1R35GM153401-01), Department of Energy (DOE) (EE0008923), University of California Irvine Engineering+ Health Institute and Sloan Research Fellowship. S.S. acknowledges support from the NSF Graduate Research Fellowship Program (DGE-1839285). E.L. and J.B.S. acknowledge funding from Advanced Research Projects Agency–Energy EcoSynBio program (DE-AR0001508), the National Institute of Environmental Health Sciences (P42ES004699), the NIH (R01GM076324-11) and the NSF (2328145, 1627539, 1805510 and 1827246). N.-H.H., J.-K.K. and F.Q. acknowledge funding from the NSF (2328145) and NIH (R35GM149572). This research used resources at AMX beamline of the National Synchrotron Light Source II, a US DOE Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory (DE-SC0012704). The Center for BioMolecular Structure is primarily supported by the NIH, National Institute of General Medical Sciences through a Center Core P30 grant (P30GM133893) and by the DOE Office of Biological and Environmental Research (KP1605010). The ALS-ENABLE beamlines are supported in part by the NIH, National Institute of General Medical Sciences (P30GM124169-01). The Advanced Light Source is a DOE Office of Science User Facility (DE-AC02-05CH11231). This work used the high-performance computing cluster maintained by the Research Cyberinfrastructure Center at the University of California, Irvine.

Author contributions

S.S., N.-H.H., F.Q. and H.L. designed all experiments. E.L., S.S. and J.B.S. performed the bioinformatics analysis and SSN construction. E.L., S.S., W.B.B., J.B.S. and H.L. selected the ALDH candidates. W.B.B. and S.S. constructed the ALDH expression vectors. S.S. performed the colorimetric assays, 340-nm specific activity assays and benzaldehyde conversion assays of ALDHs. S.S. and V.C.M. performed steady-state kinetic characterization of ALDHs. N.-H.H. and F.Q. expressed, purified, reconstituted and crystallized *Bt*ALDH3 and its variants. B.S. collected diffraction data. N.-H.H., J.-K.K., A.H.T.T. and F.Q. solved and analyzed the crystal structures. R.L.H., S.Z. and S.S. performed and analyzed the molecular dynamics simulations. H.J.C.N. and S.S. synthesized simple-synthetic cofactors. All authors analyzed the data and wrote the manuscript.

Funding

This work was supported by the Advanced Research Projects Agency–Energy EcoSynBio program (DE-AR0001508), the NSF (DGE-1839285, 2328145, 1627539, 1805510 and 1827246), the NIH (1R35GM153401-01, R35GM149572, R01GM076324-11, P42ES004699, P30GM133893 and P30GM124169-01), the US DOE (EE0008923,

KP1605010, DE-SC0012704 and DE-AC02-05CH11231) and a Sloan Research Fellowship.

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/s41589-026-02315-w.

Correspondence and requests for materials should be addressed to Feng Qiao or Han Li.

Peer review information *Nature Chemical Biology* thanks the anonymous reviewers for their contribution to the peer review of this work.

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

nature portfolio

Corresponding author(s): Han Li, Feng Qiao

Last updated by author(s): Apr 21, 2026

Reporting Summary

Nature Portfolio 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 Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☐	☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters 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)
☐	☒ 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.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Specific activity and kinetic characterization parameters were collected using SoftMax Pro 7.0 Software. LC data were collected and analyzed using ChemStation B.04.03-SP1. Molecular dynamics data were collected using GROMACS 2025.3. All custom software used is deposited at [https://github.com/hanli-lab/BtALDH3_MD.git].
Data analysis	Crystallography data were analyzed with iMosFlm version 7.4.0., autoPROC 1.0.5, and PHENIX 1.20.1_4487. Molecular dynamics data were analyzed using Python 3.0. Multiple sequence alignments and phylogentic trees were generated with EMBL-EBI Clustal Omega webserver tool and the Interactive Tree of Life phylogenetic tree visualizer web tool, respectively. All other data were analyzed with Microsoft Excel 16

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

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 description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All data supporting the findings of this work are available within the paper and its Supplementary Information files. Source Data files are also provided with this paper. Crystal structures of BtALDH3 WT in complex with NAD+ (9PXQ), BtALDH3 WT in complex with NMN+ (9PXH), and BtALDH3 (10PQ) are deposited in the Protein Data Bank. Molecular dynamics trajectories have been deposited at 10.5281/zenodo.21229755.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Not applicable
Reporting on race, ethnicity, or other socially relevant groupings	Not applicable
Population characteristics	Not applicable
Recruitment	Not applicable
Ethics oversight	Not applicable

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

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

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

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

Sample size	Colorimetric assays were performed in duplicate (n=2) as high-throughput primary screens intended to flag candidates for subsequent validation rather than to provide definitive quantification. Molecular dynamics simulations were run with ten independent replicates (n=10), and all other experiments were performed in triplicate (n=3). Sample sizes follow established practice in the field: duplicate screening is sufficient for a first-pass filter given downstream confirmation, while triplicate measurement supports determination of the mean and standard deviation for the reported parameters.
Data exclusions	No data were excluded.
Replication	All replication attempts were successful.
Randomization	All samples in the same experiment were treated identically, therefore randomization was not applicable.
Blinding	This study reports objective measurements of samples derived from purified proteins or chemical reactions. In each experiment, the samples were treated identically, therefore blinding was not applicable.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Plants

Seed stocks	Not applicable
Novel plant genotypes	Not applicable
Authentication	Not applicable