This is a repository copy of *Expanding the Substrate Scope of Native Amine Dehydrogenases through In Silico Structural Exploration and Targeted Protein Engineering*.

**Article:**

**ChemCatChem**

Research Article
doi.org/10.1002/cctc.202200880

Europe
European Chemical
Societies Publishing

www.chemcatchem.org

# Expanding the Substrate Scope of Native Amine Dehydrogenases through *In Silico* Structural Exploration and Targeted Protein Engineering

Laurine Ducrot,[a] Megan Bennett,[b] Gwenaëlle André-Leroux,[c] Eddy Elisée,[a] Sacha Marynberg,[a] Aurélie Fossey-Jouenne,[a] Anne Zaparucha,[a] Gideon Grogan,[b] and Carine Vergne-Vaxelaire*[a]

Native Amine Dehydrogenases (nat-AmDHs) are NAD(P)H-enzymes performing reductive amination, mainly active towards small aliphatic aldehydes and cyclic ketones, due to active site volumes limited by the presence of several bulky amino acids. Herein, inspired by the diversity of residues at these positions among the family, we report the implementation of mutations F140A and W145A in *Cfus*AmDH and their transposition in nine other members. Moderate to high conversions were obtained with substrates not accepted by the native enzymes, notably *n*-alkylaldehydes (44.6%–99.5% for hexanal to nonanal) and *n*-alkylketones (16.0%–53.7% for hexan-2-one to nonan-2-one) with retention of excellent (*S*)-enantioselectivity (>99% *ee*). Complementary to the reported (*R*)-selective AmDHs, the promising mutant *Cfus*AmDH–W145A was further characterized for its synthetic potential. Crystal structure resolution and molecular dynamics gave insights into the cofactor and substrate specificity and the whole structural dynamics, thus providing keys for mutagenesis work on this enzyme family.

## Introduction

The amine function is ubiquitous among biologically active compounds[1] and reductive amination is one of the top 10 reactions used for the industrial synthesis of pharmaceuticals, representing 5.3% of all the reactions in the chemists' toolbox.[2] Therefore, facing significant economic and environmental challenges, society needs to intensify the search for sustainable ways to perform this reaction. Among catalytic technologies, biocatalysis stands out as the most environmentally friendly approach.[3] ω-Transaminases (ω-TAs), which can reversibly catalyze the formation of amine groups by transferring the amine function from a sacrificial amino donor to a ketone or aldehyde,[4] are already widely used even on an industrial scale,

the greatest example being the production of the antidiabetic drug, sitagliptin.[5] However, alternative enzymes that can perform direct reductive amination in atom-efficient processes are required. For example, native or first generation mutants of oxidoreductases (ORs), such as amine dehydrogenases (AmDHs), imine reductases (IREDs) and reductive aminases (RedAms), catalyze the direct reductive amination of carbonyl substrates. Nevertheless, they mostly display a narrow substrate spectrum generally restricted to small molecules due to the limited size of the active site pockets.[6] To extend the applicability of these ORs in synthetic processes, further protein engineering work has been reported to design enzymes that offer higher conversions of longer or more hindered carbonyl compounds. Key examples are the improved L-AmDH-TV L39A/A112G,[7] *Gk*AmDH–M0/M3/M8[8] and L-*Bc*AmDH–M5 which displayed activity towards nonan-2-one, dioxolo-benzylacetone and the even more bulky ketones 6-phenylpentan-2-one and 6-phenylhexan-2-one,[9] respectively. As the main effort was directed towards the production of chiral amines, only very few enzyme activities towards long aliphatic aldehydes have been reported thus far. *Es*LeuDH-DM, obtained through a double mutation of a LeuDH from *Exiguobacterium sibiricum*, displayed reductive amination activities with *n*-alkyl C3-C8 aldehydes in the range 0.09 to 1.16 U mg$^{-1}$.[10] This class of primary amines, such as hexan-1-amine, is widely used either for direct application or as building blocks for the synthesis of commodity chemicals along with high value-added pharmaceuticals, agro-chemicals or fine chemicals.[11] The nat-AmDH family, first discovered in our group in 2019 and more recently expanded with the use of metagenomic data, comprises NAD(P)H-dependent enzymes performing direct reductive amination, using mainly ammonia. They largely exhibit activities towards short aliphatic aldehydes, particularly MATOUAmDH2, which displays

[a] L. Ducrot, Dr. E. Elisée, S. Marynberg, A. Fossey-Jouenne, Prof. A. Zaparucha, Dr. C. Vergne-Vaxelaire
Génomique Métabolique, Genoscope
Institut François Jacob
CEA, CNRS, Univ Evry, Université Paris-Saclay
2 rue Gaston Crémieux 91057 Evry (France)
E-mail: carine.vergne@genoscope.cns.fr
[b] M. Bennett, Prof. G. Grogan
York Structural Biology Laboratory
Department of Chemistry
University of York
Heslington, York, YO10 5DD (UK)
[c] Dr. G. André-Leroux
Université Paris-Saclay
INRAE, MaIAGE
78352 Jouy-en-Josas (France)

**ChemCatChem**

Research Article
doi.org/10.1002/cctc.202200880

**Chemistry
Europe**
European Chemical
Societies Publishing

the highest specific activities with up to 11 U mg$^{-1}$ for isobutyraldehyde at 50 °C.[12] However, either no, or only traces of activity, have been detected towards linear aldehydes or ketones above C5 or aromatic ketones such as 4-phenylbutan-2-one.[6d,12b,13] Up to now, only few protein engineering experiments were performed on this nat-AmDH family. Using the structure of AmDH4 in complex with NAD$^+$ (PDB: 6G1M), the enzyme was engineered to convert pentan-2-one to (2S)-pentan-2-amine by mutating the carboxylate binding site of its natural substrate.[13c,14] Additionally, its catalytic efficiency towards the bio-based levulinic acid was improved by 18-fold through directed evolution carried out by Cai et al.[15] More recently, MATOUAmDH2 was engineered to give active site mutants with greater activity for the amination of cyclohexanone with both ammonia and methylamine.[12a]

In the present work, with the aim at gaining activity towards larger aliphatic aldehydes and ketones by enlargement of the active site volumes, we performed rational substitutions on ten nat-AmDHs from a structure-based strategy using two structures of nat-AmDHs (CfusAmDH, PDB: 6IAU and MsmeAmDH, PDB: 6IAQ) (Scheme 1). X-ray crystallographic structures and Molecular Dynamics (MD) studies are described for one of the most promising mutants, to understand the effects of the mutation and to gain insights into the conformational dynamics in nat-AmDHs. Eventually, we aim at providing sequence/structure/function correlation to inform further improvements.

## Results and Discussion

### *In silico* analysis of biodiversity and determination of the mutated sites

The structural resolution of AmDH4 in complex with NAD$^+$ (Uniprot: A8BHL2, PDB: 6G1M) previously enabled the identification of the residue positions defining the active site, named P1-P20. This allowed us to compare the modeled active sites of the whole available nat-AmDH family, defined thanks to an ASMC (Active Site Modeling and Clustering) analysis.[12b,13c]
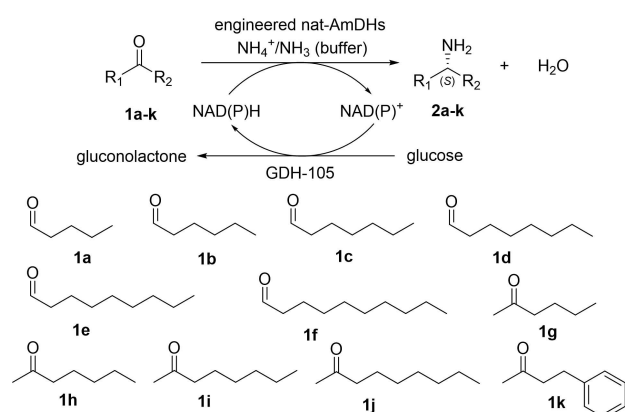


**Scheme 1.** Reductive amination of selected ketones and aldehydes **1 a**–**1 k** into the corresponding amines **2 a**–**2 k** catalyzed by engineered nat-AmDHs in this work.

Based on the structure of CfusAmDH, MsmeAmDH and homologs, the list of residues constituting the active site was reconsidered by adding a new position between P4 and P5. Unlike in AmDH4, this position is occupied by bulky amino acids. The active sites are now defined by residues P1-P21. In contrast to many studies that generally base their protein engineering strategy on only a few structures or models, we decided to exploit the substantial resource of data from biodiversity to target and mutate candidate residues. Therefore, we looked at the structures of CfusAmDH (Uniprot: S9Q235, PDB: 6IAU), MsmeAmDH (UniprotKB entry: A0A8B4R7U8 formerly A0A0D6I8P6, PDB: 6IAQ) and the model of their 291 homologs already reported in the ASMC.[13c] The key hypothesis to be made was that the restricted substrate scope of nat-AmDHs may be due to a very small active site, mainly formed by bulky aromatic amino acids located on the first layer, including Y or F at P5, P12 and P17, Y, F or W at P10 and W at P8 (Figure S1–S3).

The docking of n-alkylketones such as pentan-2-one already reported,[16] revealed an accomodation of the n-alkyl chain towards the positions P5 (F140 in CfusAmDH) and P8 (W145 in CfusAmDH). These two positions were selected for mutations to enlarge the pocket and so to extend the substrate scope to longer n-alkyl ketones and aldehydes (Figure 1). As depicted in Figure S3, discerning examples from biodiversity (UniprotKB entries are used to name them) display smaller residues at these positions, such as A at P5 in A0A1N7N2Q9 and A0A1X4NNW3, C at P8 in MATOUAmDH1 and MATOUAmDH2,[12a] G at P8 in OM-RGCAmDH2, A0A078BAL0, A0A078AU13 and A0A077ZQK8, and L at P8 for example in A0A076NBW6. The occurrence in close AmDHs homologs of smaller amino acids at positions P5 and P8, while harboring identical or similar neighbors, suggests that mutations at these positions into smaller residues should not be detrimental neither for enzyme folding nor for activity. The effect of a small residue at P8 for broader substrate scope was supported by activity results in presence of bulkier substrates such as benzaldehyde or norcamphor and structural analysis of the structure of MATOUAmDH2 (PDB: 7R09, 7ZBO).[12] The RX structure of CfusAmDH (PDB: 6IAU), displaying F140 and W145 respectively at P5 and P8, was used as a reference enzyme for *in silico* mutations and subsequent analysis of the extended pocket size through docking experiments. As the native enzyme showed a preference for aldehydes over ketones,[13c] the analysis was carried out with the docking of n-terminal primary amine products, from pentan-1-amine (**2 a**) to decan-1-amine (**2 f**). As depicted in Figure 1, the separate mutations F140A and W145A increased the active site volume from 957 Å$^3$ in wild-type (WT) CfusAmDH to 1194 Å$^3$ and 1176 Å$^3$, respectively. For WT CfusAmDH, the small chains **2 a** and the longest amine formed experimentally **2 b**, were obtained in the highest number of energetically viable poses inside the active site with top pose binding energy of −5.5 and −5.9 kJ mol$^{-1}$, respectively. Fewer than three poses of the ten were obtained with larger ligands up to nonan-1-amine (**2 e**) and in a single conformation with higher energy (−5.0 kJ mol$^{-1}$) (Figure 1A and Figure S4) that appeared highly constrained within the active site. In CfusAmDH–F140A and CfusAmDH–W145A mutants, the pocket
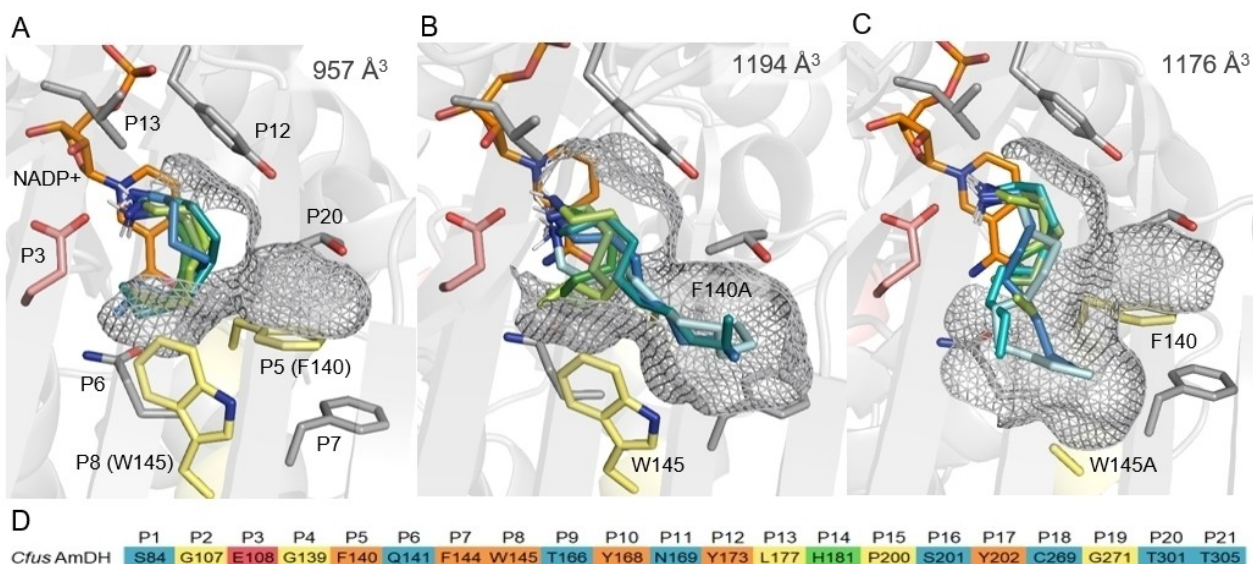
**Figure 1.** Active site of *Cfus*AmDH (A) (RX structure; PDB: 6IAU), *Cfus*AmDH–F140A (B) and *Cfus*AmDH–W145A (C) with docked structures of pentan-1-amine (**2a**) to heptan-1-amine (**2c**) in shades from dark to light green and octan-1-amine (**2d**) to decan-1-amine (**2f**) (only up to nonan-1-amine (**2e**) for *Cfus*AmDH) in shades from dark to light blue. NADP$^+$ is represented in orange, the catalytic residue E108 (P3) is in pink, F/A140 (P5) and W/A145 (P8) are in yellow and some residues that form the first layer of the pocket are in grey. For sake of clarity, only the most energetically favorable pose is represented. The pocket enlargement in the two mutants is shown using the grey surface representation. The active site volumes are given in Å³. (D) P1-P21 residues in *Cfus*AmDH. The color code used refers to the polarity and charge of the corresponding residue [blue: polar residues, yellow: hydrophobic residues, orange: aromatic residues, red: negatively charged residues, and green: positively charged residues (charges at physiological pH)].

could easily accommodate longer *n*-alkylamines up to **2f** (longer chains have not been tested) with nine and ten poses inside the pocket out of ten generated poses, and with binding energies even lower than for *Cfus*AmDH/**2a** (−6.8 and −6.5 kJ mol$^{-1}$ for *Cfus*AmDH–F140A and *Cfus*AmDH–W145A respectively) (Figure S4). Docked amines occupied the space freed up by the small A residue and led to linear conformations directed towards the bottom of the pocket in the case of *Cfus*AmDH–F140A (Figure 1B) and to slightly curved conformations for *Cfus*AmDH–W145A (Figure 1C). The structure of nat-AmDHs has been observed in both open and closed conformations, the open conformation maintaining the active site pocket accessible to the solvent.[12a,13c] Thus we can hypothesize that the accessibility of larger compounds such as heptanal (**1c**) to decanal (**1f**) should not be a limitation compared to enzymes that display a restrictive access tunnel.[17] Further *in silico* experiments such as MD simulations studying how the substrates access the pocket could complement this hypothesis.

### *In vitro* mutations of *Cfus*AmDH

Based on *in silico* results, positions F140 and W145 in *Cfus*AmDH were mutated into A (Table S1–S2 and Figure S5 and S6). Due to slight protein degradation of *Cfus*AmDH–F140A at 30°C, only the specific activity of *Cfus*AmDH–W145A was measured towards 10 mM of pentanal (**1a**) to octanal (**1d**) in presence of 5% DMSO (Table 1). As expected from the docking experiments, WT *Cfus*AmDH displayed very low or no activity towards substrates larger than **1a** (<20 mU mg$^{-1}$). On the contrary,

| Table 1. Specific activities of WT *Cfus*AmDH and *Cfus*AmDH–W145A. |
|---|

| | Specific activities [mU mg$^{-1}$][a] | | | |
|---|---|---|---|---|
| | *Cfus*AmDH | | *Cfus*AmDH–W145A | |
| | NADH | NADPH | NADH | NADPH |
| **1a** | 73.3 ± 6.8 | 187.8 ± 1.8 | 15.5 ± 1.5 | 61.9 ± 0.9 |
| **1b** | 11.5 ± 1.0 | 18.9 ± 0.3 | 195.2 ± 7.4 | 400.9 ± 1.5 |
| **1c** | 16.6 ± 5.1 | *n.d* | 404.6 ± 9.8 | 371.1 ± 10.7 |
| **1d** | *n.d* | *n.d* | 196.7 ± 11.5 | 483.4 ± 22.8 |

[a] Reaction conditions: final volume 100 μL, 10 mM aldehyde substrate, 2 M NH₄HCO₂ buffer, pH 8.5, 0.2 mM NAD(P)H, 0.03–0.25 mg mL$^{-1}$ purified enzyme, 30°C. Errors represent standard deviations of two or three independent experiments *n.d* not detected

*Cfus*AmDH–W145A displayed activities of 401, 405 and 483 mU mg$^{-1}$ respectively towards **1b**, **1c** and **1d**. Interestingly, this mutant was less active than the WT enzyme towards the smaller substrate **1a**. This tendency to gain activity for longer substrates, while losing activity towards shorter ones, was already reported for some engineered Amino Acids Dehydrogenases (AADHs).[7,18] This might be due to an unsuitable positioning of the shorter substrates in the pocket, induced by an active spatial arrangement of the substrate mainly by steric hindrance along with the hypothesized weak stabilization by interaction with Y168 (P10). The specific activities of the W145A mutant towards aldehydes **1b**–**1d** were significantly higher than the only one reported for the same substrates with *Es*LeuDH-DM (320 *vs* 401 mU mg$^{-1}$ for **1b**, 190 *vs* 371 mU mg$^{-1}$ for **1c** and 90 *vs* 483 mU mg$^{-1}$ for **1d**).[10]

Another interesting observation concerns the cofactor preference of the mutant, which appeared to depend on the

substrate as seen preliminary for the WT enzyme.[13c] With **1a**, both WT *Cfus*AmDH and *Cfus*AmDH–W145A seemed to favor NADPH over NADH with ratios of 2.5:1 and 4:1, respectively. However, this ratio decreased to 2:1 with **1b** and was even reversed with **1c** with a slight preference for NADH (specific activities of 405 and 371 mU mg$^{-1}$ with NADH and NADPH, respectively). For the longer substrate **1d**, NADPH appeared to be preferred. This phenomenon was more pronounced for the W145A mutant while activities remained very low for the WT enzyme (Table 1).

Mutation of P5 and P8 to G, also found at P8 in some native AmDHs (Figure S3), did not lead to notably better activities towards longer substrates **1d**–**1e** compared to F140A or W145A (Table S1–S2 and Figure S5 for production, activity data not shown).

The potential of mutants F140A and W145A for biocatalytic reactions was confirmed with conversions of 39.6 and 68.0% of **1c**, respectively, in contrast to the WT enzyme which gave conversions <10%. Mutant F140A appeared to be more sensitive to the co-solvent DMSO, as the conversion dropped from 39.6% at 1% DMSO to 2.2% at 5% DMSO, unlike mutant W145A still leading to 45.0% conversion at 20% v/v DMSO (Figure S7 and S8).

## Transposing mutations that confer specificity for larger substrates to other AmDHs

In order to investigate the transposability of the mutational pattern defined above, the mutations corresponding to F140A and W145A in *Cfus*AmDH were replicated in nine other nat-AmDHs (*Msme*AmDH; *Micro*AmDH; *Apau*AmDH; MATOUAmDH1; MATOUAmDH2; *Porti*AmDH; *Rgna*AmDH; *Chat*AmDH; IGCAmDH1) (Figure S3). Mutations into A were preferred over G to avoid unstable and/or inactive protein as observed in *Cfus*AmDH–F140G, although it could have helped reaching activity towards even larger substrates in other templates. Some of the selected nat-AmDHs were already reported active towards a broader range of substrates compared to the rest of the family, namely MATOUAmDH1, MATOUAmDH2[12] and *Micro*AmDH, the latter displaying low activity towards C3-keto substrates.[13c] *Porti*AmDH and *Rgna*AmDH were part of the nat-AmDHs found among metagenomic databases but were not extensively characterized as purified enzymes.[12b] Among the nine selected WT enzymes and their 18 mutants overexpressed in *E. coli* and purified on Ni-NTA columns, only MATOUAmDH1 and its two mutants could not be produced with a sufficient yield to be used in activity screening. All variants (Table S1 and S2 and Figure S5 and S6) were assayed in 5% v/v of DMSO against a range of carbonyl substrates, including aldehydes from hexanal (**1b**) to decanal (**1f**), and ketones from hexan-2-one (**1g**) to nonan-2-one (**1j**), as well as the commonly tested 4-phenylbutan-2-one (**1k**) (Figure 2 and Table S4 and S5). The most notable improvements were seen for the aldehydes **1b**–**1e** with moderate to high conversions (20–>99%) by mutant enzymes whereas most of the WT enzymes, such as *Chat*AmDH, IGCAmDH1, *Rgna*AmDH, *Cfus*AmDH and *Apau*AmDH, could not
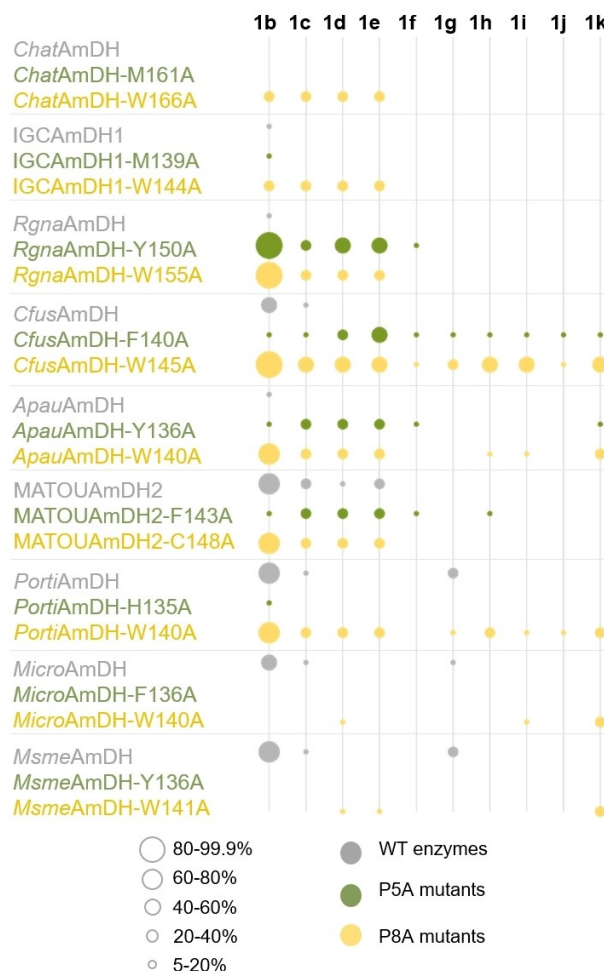


**Figure 2.** Conversions obtained for WT enzymes and mutants towards substrates **1b**–**1k**. Reaction conditions: final volume 100 μL, 10 mM substrate (DMSO 5% (v/v)), 2 M NH$_4$HCO$_2$ buffer, pH 8.5, 0.4 mM NAD(P)$^+$, 1.2 eq. D-glucose, 3 U mL$^{-1}$ GDH-105, 0.5 mg mL$^{-1}$ purified enzyme, 25 °C, 24 h. The monitoring of the remaining aldehydes or ketones **1b**–**1j**, the amines **2b**–**2j** and the corresponding alcohols formed was performed by GC-FID after ethyl acetate extraction. The monitoring of the amine **2k** formed was carried out by UHPLC-UV after derivatization with Marfey's reagent (FDAA) (Table S3-S5).

convert substrates larger than **1b**. For example, *Rgna*AmDH converted 12.5% of **1b** whereas its mutants Y150A and W155A enabled conversions of 95.2% and 85.6%, respectively. *Rgna*AmDH–Y150A even converted **1e** with 44.6% conversion. *Chat*AmDH and IGCAmDH1 were not active as WT enzymes towards aldehydes **1b**–**1e**, but the P8 mutations W166A and W144A achieved moderate conversions of 26.2% and 34.6%, respectively. Similarly, for *Porti*AmDH, for which the WT enzyme converted 60.8% of **1b**, only the mutant W140A exhibited significant conversions (22.0–25%) of longer chain substrates up to nonanal (**1e**). An even higher conversion of >99.5% was obtained with **1b** and *Cfus*AmDH–W145A. As expected, the enlarged active site of WT MATOUAmDH2 favored the conversion of **1b** with 72.7% and larger substrates up to **1e** with 20.5%. In this case, the mutations of P5 and P8 residues (F143A and C148A, respectively) did not noticeably increase the

**ChemCatChem**

Research Article
doi.org/10.1002/cctc.202200880

**Chemistry
Europe**
European Chemical
Societies Publishing

conversion for all the other aldehydes. Interestingly, only *Rgna*AmDH–Y150A, MATOUAmDH2-F143A, *Cfus*AmDH–F140A and *Apau*AmDH–Y136A mutated in P5 displayed notable conversions of **1 f** (7.8–11.5 %), suggesting that this mutation led to improved accommodation of the longest carbon chain tested. Such reported conversion data towards aldehydes **1 b**–**1 f** with the direct use of ammonia as the amine donor are of importance, as the only other example in the literature is that of *Es*LeuDH-DM.[10] Some conversion data on **1 b** and **1 c** by enzymes performing reductive amination have been reported, but only with methylamine and propargylamine as the amine donor.[6c,19]

For the *n*-alkylketones from **1 g** to **1 j**, most of the WT enzymes displayed very low or no conversion, except for *Porti*AmDH and *Micro*AmDH, which converted 26.5 % and 22.6 % of **1 g** with 79.4 % and 94.6 % *ee*, respectively, in favor of the (*S*)-amine. Among the mutants, *Cfus*AmDH–F140A, *Cfus*AmDH–W145A, *Apau*AmDH–W145A and *Porti*AmDH–W140A gained activity towards larger ketones from **1 g** to **1 j**. *Cfus*AmDH–W145A gave the highest conversions with 39.7 % of **2 g**, (99.3 % *ee*), 53.7 % of **2 h** (99.7 % *ee*), 40.3 % of **2 i** (99.5 % *ee*) and 16.0 % of **2 j** (> 99.9 % *ee*). In most cases, the (*S*)-enantioselectivity of the WT enzymes was conserved. Notably, the P8 mutation of *Porti*AmDH (*Porti*AmDH–W140A) enabled slightly lower conversion (26.5 % for WT against 18.2 % for the mutant) but with a noticeable improved stereoselectivity with *ee* from 79.4 % to 99.5 %. Thanks to these mutations, we have now in hand alternative AmDH templates for the reductive amination of aliphatic ketones, complementary to both AmDHs engineered from α- and ε-AADHs and RedAms, which already displayed high efficiency and stereoselectivity towards such kind of carbonyl substrates, but which gave the opposite (*R*)-enantiomer. For instance, with *Gk*AmDH from *Geobacillus kaustophilus*, Liu *et al.* reported 99 % and 98 % conversion (> 99 % *ee*) for the formation of (2*R*)-**2 g** and (2*R*)-**2 h** respectively, but the conversion dropped to < 5 % with **1 i** despite high enzyme loadings (1–3 mg mL$^{-1}$).[20] Importantly, *Cfus*AmDH–W145A, and to a lesser extent *Porti*AmDH–W140A, constitute interesting alternatives to *Nf*RedAm, the only enzyme reported thus far that gives the (*S*)-alkylamines with 41 % formation of (2*S*)-decan-2-amine with 84 % *ee*.[6e] Finally, *Cfus*AmDH–F140A, *Cfus*AmDH–W145A, *Apau*AmDH–W141A, *Porti*AmDH–W140A and *Micro*AmDH–W141A were shown to be active towards 4-phenylbutan-2-one (**1 k**), a substrate not observed previously to be (*S*)-aminated by nat-AmDHs and converted to (*R*)-**2 k** by reported engineered enzymes.[6c,h,7–9,19,21] Higher conversions were again obtained with *Cfus*AmDH–W145A, with 52.5 % conversion into (2*S*)-**2 k** (99.6 % *ee*). It is worth noting that for *Micro*AmDH and *Msme*AmDH, the P8 mutation of the residue was only beneficial for **1 k** transformation. For all assays, the alcohol formation did not exceed 4.0 %, reached with IGCAmDH1-W144A for the reduction of octanal **1 d** into octan-1-ol (Table S5). These results confirmed, as previously reported, that WT nat-AmDHs do not significantly catalyze the reduction of carbonyls to alcohols in presence of excess of ammonia and that this behavior was not altered by the P5 and P8 mutations.[13c,22] All the mutants were tested

towards other shorter but bulky substrates such as benzylamine or acetophenone but the mutations did not enable to increase the conversion and even annihilated the activity towards these substrates in some cases. The positive results obtained towards the bulky and non aliphatic linear substrate **1 k** suggest the potential for binding even larger substrates, perhaps with substitution on the aromatic ring or functionalization of the alliphatic chain

### Kinetic parameters of *Cfus*AmDH-W145A

Kinetic parameters were determined to further characterize the most promising mutant *Cfus*AmDH–W145A arising from the conversion screening, and more specifically to gain insights on the observed change in cofactor specificity dependent upon the aldehyde substrate. The following four operating conditions were assayed: **1 c**/NADPH/NH$_3$, **1 c**/NADH/NH$_3$, **1 a**/NADPH/NH$_3$, **1 a**/NADH/NH$_3$ (Table 2). Plots are reported in Figure S9–S12.

As expected, whether using NADPH or NADH, the W145A mutant displayed 57 to 128-fold higher (apparent) catalytic efficiency for heptanal (**1 c**) than for pentanal (**1 a**), with $k_{cat}/K_M$ with NADPH cofactor of 2.56 $10^3$ and 1.58 $10^1$ s$^{-1}$ M$^{-1}$, respectively. This is mainly due to the $K_M$ being 43 and 4-fold higher for pentanal **1 c** than heptanal **1 a** with NADPH and NADH, respectively, that can attest a difference of protein affinity for the two substrates. Surprisingly, compared to the results from previous specific activity assays, the $K_M$ of the cofactor were 11–14-fold lower for NADPH over NADH with **1 a** and **1 c**. Interestingly, the poor affinity of W145A mutant for **1 a** occurs with an inhibition effect with a $K_i$ estimated at 41.34 mM with NADPH (Figure S11 and S12). This also led to an inhibition over 1.5 M NH$_4$HCO$_2$ buffer, which was not observed with **1 c** at even higher concentrations of this amine (Figure S11 and S12). No explanation was found for this behavior but complementary analyses confirmed that it was neither due to pH, ionic strength of the reaction mixture nor product inhibition (data not shown).

**Table 2.** Kinetic parameters of *Cfus*AmDH–W145A.

|  | $k_{cat}$ [a] | $K_M$ [a] | $K_i$ [a] | $k_{cat}/K_M$ [a] |
|---|---|---|---|---|
| **1 c** [d] | 0.74 ± 0.03 | 0.29 ± 0.04 |  | 2.56 $10^3$ |
| NADPH[d] | 0.78 ± 0.02 | 0.21 ± 0.02 |  | 3.71 $10^3$ |
| NH$_3$[d] | 0.72 ± 0.02 | 196.20 ± 25.50 |  | 3.62 |
| **1 c**[e] | 0.59 ± 0.02[c] | 1.03 ± 0.15[c] |  | 5.72 $10^2$ [c] |
| NADH[e] | 1.78 ± 0.17 | 3.00 ± 0.77[b] |  | 5.94 $10^2$ |
| NH$_3$[e] | 0.52 ± 0.01[c] | 126.40 ± 13.30[c] |  | 3.98[c] |
| **1 a**[f] | 0.20 ± 0.05 | 12.51 ± 4.60 | 41.34 ± 15.73 | 1.58 $10^1$ |
| NADPH[f] | 0.10 ± 0.01 | 0.07 ± 0.01 |  | 1.40 $10^3$ |
| NH$_3$[f] | 0.10 ± 0.01[c] | 319.80 ± 52.00[c] |  | 0.31[d] |
| **1 a**[g] | 0.03 ± 0.01[c] | 3.95 ± 0.43[c] |  | 7.35 [c] |
| NADH[g] | 0.07 ± 0.01 | 0.74 ± 0.05[b] |  | 9.20 $10^1$ |
| NH$_3$[g] | 0.03 ± 0.01[c] | 191.00 ± 33.90[c] |  | 0.17[c] |

[a] Kinetic parameters were measured in a 100 μL final volume in NH$_4$HCO$_2$ buffer (0.05–4 M, pH 8.5) at concentrations of 0.1–12 mM of heptanal (**1 c**), 0.5–120 mM of pentanal (**1 a**), 0.01–0.9 mM of NADPH and 0.05–0.9 mM of NADH. $k_{cat}$ are given in s$^{-1}$, $K_M$ and $K_i$ in mM and $k_{cat}/K_M$ in s$^{-1}$ M$^{-1}$. [b] $K_M$ values calculated with Lineweaver-Burk model. [c] Apparent kinetic parameters. [d] with **1 c**/NADPH/ NH$_3$. [e] with **1 c**/NADH/ NH$_3$. [f] with **1 a**/ NADPH/ NH$_3$. [g] with **1 a**/NADH/ NH$_3$.

ChemCatChem

Research Article
doi.org/10.1002/cctc.202200880

Chemistry
Europe
European Chemical
Societies Publishing

It is worth noting that the mutation did not impact significantly the affinity for $NH_3$ in any of the assayed conditions; the $K_M$ of 98.9 mM being previously reported for the WT enzyme[13c] and $K_M$ of 100–300 mM being common for AmDHs. Only the $k_{cat}$ and affinity of longer aldehydes such as **1c** were beneficially impacted; no data could be obtained with the WT enzyme as it does not accept such an aldehyde. To our knowledge, no kinetic parameter has been reported for such aldehydes and enzymes performing reductive amination that would enable any comparison.

## Thermostability

Despite the occurrence of small amino acid residues at P8 in native homologs, one cannot exclude that W145A mutation could have an impact on the structure flexibility and stability of *Cfus*AmDH. Thus, we determined the thermostability of *Cfus*AmDH–W145A at 25, 30 and 40 °C and compared the data with those of *Cfus*AmDH. As shown in Figure 3, the mutant appeared to be slightly less thermostable than the WT enzyme but still displayed good stability at 25 °C and 30 °C with half-life around 350 h and 200 h, respectively, compared to >600 h and 480 h for *Cfus*AmDH. The features of *Cfus*AmDH–W145A (stability at the optimum temperature range, DMSO tolerance) make this mutant a good candidate with respect to envisaged process applications (Figure S7).

## Substrate loading and semi-preparative scale reaction on octan-2-one 1i and heptanal 1c

For biocatalytic synthetic purposes, the substrate tolerance of *Cfus*AmDH–W145A was studied with increasing loadings of heptanal (**1c**) and octan-2-one (**1i**) up to 200 mM and various enzyme loadings for reaction times of 24 h and 48 h. The results
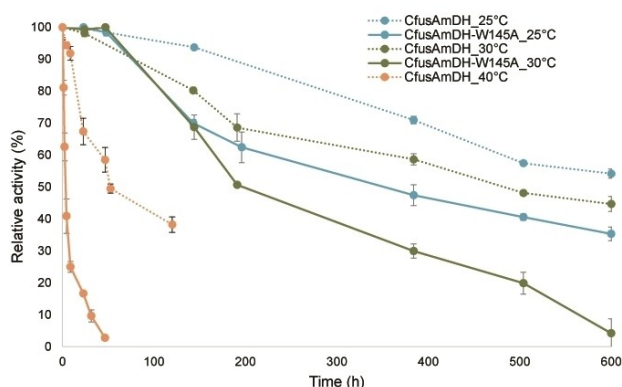


**Figure 3.** Thermostability of *Cfus*AmDH and *Cfus*AmDH–W145A at 25, 30 and 40 °C. The remaining specific activities were measured as described in the Experimental Section with 0.05–0.12 mg mL$^{-1}$ of purified enzyme, 30 °C. The maximum activity of *Cfus*AmDH towards **1a** was 219.2 mU mg$^{-1}$ and the maximum activity of *Cfus*AmDH–W145A towards **1c** was 256.3 mU mg$^{-1}$. Error bars represent standard deviations of two or three independent experiments.
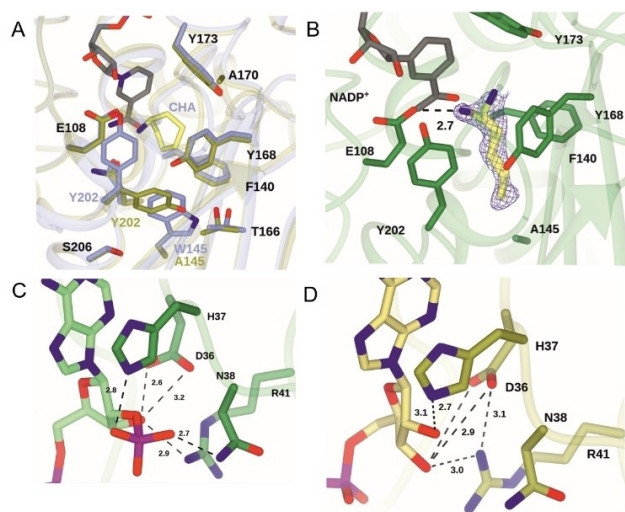
obtained after UHPLC-UV analysis showed that *Cfus*AmDH–W145A still catalyzed the formation of more amine at 100 and 200 mM of **1c** and **1i**, respectively, than at 10 mM. Up to these respective concentrations, the same trend was observed for both substrates; a decrease in conversion but still 20.6% and 15.0% at 48 h with 0.5 mg mL$^{-1}$ (Figure S13). The enzyme loading did not drastically enhance the performance. The semi-preparative reaction was performed on 50 mL-scale with 50 mM of **1i** and 0.5 mg mL$^{-1}$ enzyme loading for 24 h to afford 116 mg of the hydrochloride salt (2S)-**2i** (32.2% conversion, 28.3% isolated yield, *ee* 98.4%) (Figure S14–S16).

## Structures of *Cfus*AmDH-W145A with NAD$^+$ and NADP$^+$

In order to explore the reasons for the altered activity of the *Cfus*AmDH–W145A mutant and its cofactor specificity, its structure was solved using X-ray crystallography. Two structures were determined; one in complex with NAD$^+$ (PDB: 7QZN) and the other with NADP$^+$ (PDB: 7QZL) and pentan-1-amine (**2a**), to resolutions of 1.64 Å and 1.50 Å, respectively (Table S6). Overall, the structures superimposed well with the WT structure in complex with NADP$^+$ and cyclohexylamine (6IAU[13c]), with a RMS deviation <0.361 Å over >277 C-alpha carbons, each representing 'closed' conformations of the AmDH as previously defined.[13c] The structures verify that the W145A mutation has created the extra active site space for the accommodation of longer aliphatic aldehyde and ketone substrates as predicted. Notable differences were observed between the NADP$^+$- and NAD$^+$-mutant structures both in the active site and also in the cofactor phosphate binding site. In the NAD$^+$ complex, the most striking difference locates in the active site with the side-chain rotation of 110–115° for Y202. In subunit A, this side chain could be modeled in two conformational rotamers at 50% occupancy each. In subunit B, only the rotated form is observed. This rotation led to the loss of the hydrogen bond between E108 carboxylate moiety and Y202 hydroxyl group. The W145A mutation has left space for Y202 side chain to rotate and take the space previously occupied by the indole ring of W145 (Figure 4A). In the NADP$^+$ complex with **2a**, the Y202 side chain is held in the WT position in both subunits, but the alkyl chain of the amine extends into the space that was made by the W145A mutation (Figure 4B). Each of these results suggests that the space made by the mutation can result in flexible substrate specificity of this mutant through the combined effects of the mutation and the rotation of Y202. Compared to *in silico* docking experiments, **2a** harbors a similar orientation of the carbon chain in the active site but with a different positioning in the pocket. Notably, a shift of the nitrogen atom position from 0.7 to 2.2 Å can be observed (Figure S17).

In terms of cofactor binding, the NADP$^+$ ribose 2′ phosphate in the W145A NADP$^+$ complex makes interactions with the side chains of H37, N38 and a number of water molecules, while the 3′ OH interacts with D36 and R41, as observed for the WT *Cfus*AmDH (Figure 4C).[13c] When the NADP$^+$ phosphate is replaced by the simple 2′ hydroxyl in the NAD$^+$ complex, this group only makes direct contact with the

**ChemCatChem**

Research Article
doi.org/10.1002/cctc.202200880

Chemistry
Europe
European Chemical
Societies Publishing

**Figure 4.** (A) Active sites of *Cfus*AmDH/NADP$^+$ complex with cyclohexyl-amine (CHA, 6IAU, carbon atoms in blue) and *Cfus*AmDH–W145A/NAD$^+$ (carbon atoms in gold), showing rotation of Y202 side chain to fill space vacated by W154A mutation; (B) Active site of *Cfus*AmDH–W145A/NADP$^+$ complex with pentan-1-amine (**2a**), modeled in two conformations, each at 50% occupancy. (C) Cofactor binding pocket of *Cfus*AmDH–W145A/NADP$^+$ complex (D) Cofactor binding pocket of *Cfus*AmDH–W145A/NAD$^+$. The ionic interaction between D36 and R41 side chains in the NAD$^+$ complex is absent in the NADP$^+$ complex. Electron density corresponds to the $F_o$-$F_c$ (omit) map at a level of 3σ obtained prior to building and refinement of the ligand. Interaction between the ligand nitrogen atom and E108 side chain is shown as a black dashed line with the distance in Å.

side chain of H37, which moves slightly towards the cofactor (Figure 4D). In addition, the side chain of R41 displays a different orientation that introduces an ionic interaction of the NH1 atom with the side chain of D36 that is absent in the NADP$^+$ complex. The reduced number of interactions in the NAD$^+$ complex can perhaps reflect the overall preference of *Cfus*AmDH–W145A for NADP$^+$ over NAD$^+$ of 11:1. A comparison of the active sites of NADP$^+$ and NAD$^+$ complexes with W145A does not offer a structural explanation for the apparent different cofactor specificities observed when using different substrates. However, the flexibility of the active site, enhanced by W145A substitution and evidenced by Y202 rotation, suggests a plasticity in substrate recognition that may give rise to variable results with different substrates under different conditions.

**Molecular Dynamics**

In addition to X-ray crystallography, MD simulations were run to understand the influence of W145A mutation on protein flexibility and specificity towards cofactors and substrates.[8,23] Accordingly, the study was carried out on an extensive set of Enzyme-Cofactor-Ligand systems where both *Cfus*AmDH–W145A (M) or *Cfus*AmDH (WT) were considered, in complex with NAD$^+$ (NAD) or NADP$^+$ (NADP) cofactors, and **2a** (PEN) or **2c** (HEP) ligands.
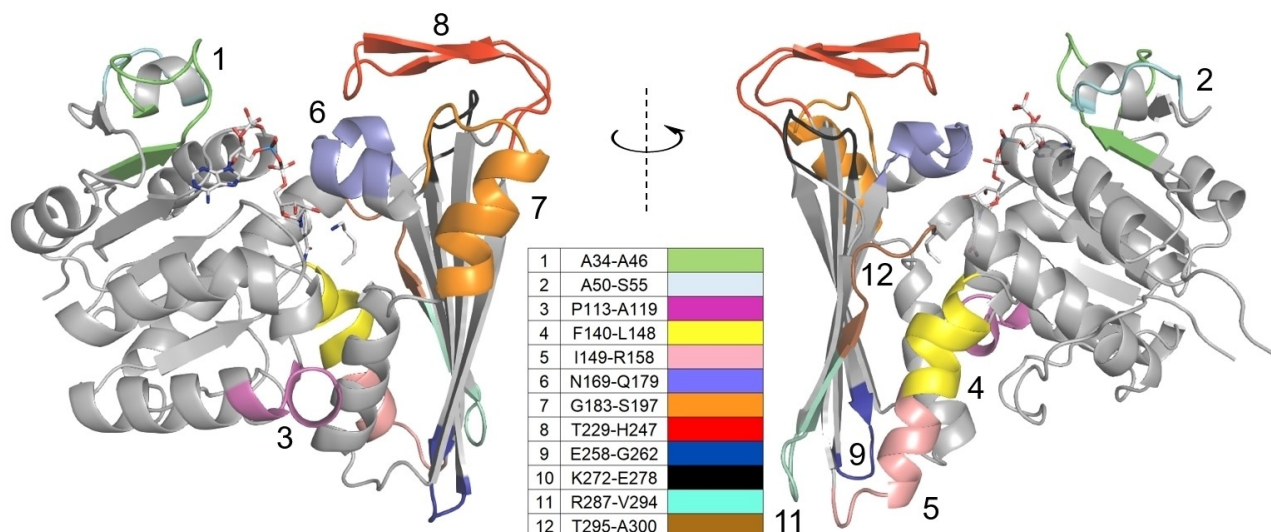
The analysis of all MD simulations led to the following observations that deal with (1) the overall dynamics of the protein, (2) the affinity of M for larger substrates, (3) the mutated site W145A and its coordination sphere and (4) the adenosine binding site.

*(1) Overall Dynamic.* M did not suffer from any major unfolding over the time of simulation, which confirms that W145A did not impact the global 3D-fold of the protein as already evidenced by both the experimental measurements and the structural studies. The closed-to-open event, previously observed in AmDH4 (PDB: 6G1M),[13c] and more recently for MATOUAmDH2,[12a] was observed for all the systems (Figure S18A). The Root Mean Square Fluctuation (RMSF), measured in Å for each full length protein, identified segments N169-Q179 (6), G183-S197 (7) and T229-H247 (8) as highly flexible (Figure S19). Monitoring the distance between G14 and G174, which are facing each other and are located on each side of the aperture, reveals that all the systems converge towards a so-called "relaxed" conformation opened at 14 Å (Figure S18B and Figure S20). Interestingly, in M systems, a higher velocity of opening is observed when the protein is complexed to NADP$^+$ as compared to NAD$^+$. Therefore, in contrast to WT complexes, the nature of the cofactor could influence the velocity of the loop opening in M complexes and so the $k_{cat}$ of the system.

*(2) The affinity of M for heptan-1-amine* **2c**. The Enzyme-Ligand interaction energies confirmed a better adaptation of HEP ligand for the mutated pocket, with a global 1.2-fold lower energy values as compared to WT, for both NADP$^+$ and NAD$^+$ (Figure S21). However, the calculated interaction energies for M_PEN systems, which were similar or lower when compared to those for WT_PEN, do not correlate with experimental data as M displays a loss of specific activity towards PEN.

*(3) Mutated site W145A and its coordination sphere.* A 14-residue long helix, located between F140 and G154 of the β-sheet domain, and harboring the mutation site, is encompassed within the flexible segments 4 and 5 (Figure 5, Figure S19 and S22). Notably, this helix is positioned within the core of the protein, where it could act as a structural spine, and is largely involved in the homodimerization (Figure S18). MD suggests that this spine helix remains rather unchanged over the opening-closing cycles despite a particularly tensed twist at F140-G146, mainly due to the presence of W145 in WT. Consistently, W145A substitution could relax this tension and provide elasticity to this helix, as suggested by the RMSF around D142-G146 in the systems M_NAD_HEP and M_NADP_PEN, respectively calculated on average 0.21 Å and 0.39 Å higher than in the corresponding WT systems (Figure S22). In the other M systems, this parameter is more similar to the corresponding WT systems.

Also in the coordination sphere of the mutated site, and particularly in the catalytic area, Q141 plays a key role in maintaining the nicotinamide moiety of the cofactor through hydrogen bonds between its oxygen and Q141 amide moiety. In WT, when Q141 loses its interaction with NAD(P), it remains quite stable in its initial position while in M, Q141 becomes more mobile (Figure 6 and Figure S23). Therefore, as the mutation has an impact on the fixation sphere of the cofactor

ChemCatChem

Research Article
doi.org/10.1002/cctc.202200880

Chemistry
Europe
European Chemical
Societies Publishing

**Figure 5.** 3D representation at 0° and 180° of WT monomer in complex with NAD and PEN. Segments 1–12, displaying a backbone fluctuation beyond 1.5 Å in Root Mean Square Fluctuation (RMSF) analysis (Figure S18),are highlighted with a gradient of colors reported in the inserted table.
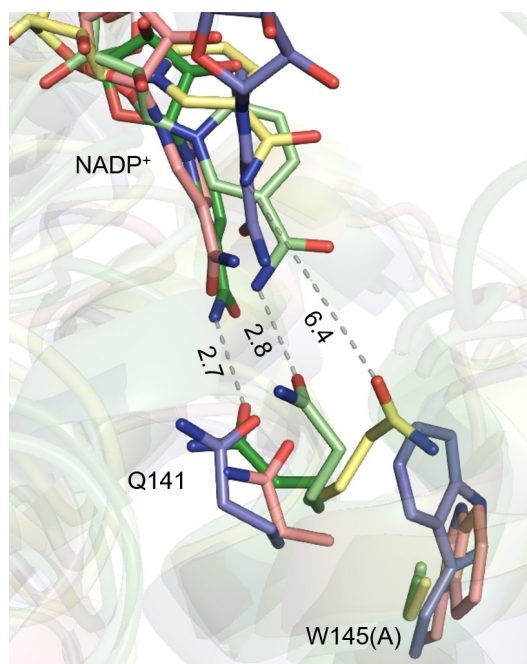


**Figure 6.** Close-view of Q141 in WT and M at different time of the dynamic. M_NAD_HEP_2 frames at 0.1, 16.2 and 18.5 ns are represented in dark green, light green and yellow, respectively. WT_NAD_HEP_4 at 29.0 ns and WT_NAD_HEP_1 at 53.2 ns are represented in purple and pink, respectively.

*Rgna*AmDH (P) and IGCAmDH6 (E), which suggests different dynamic behaviors depending both on cofactors and ligands.

*(4) The adenosine binding site.* Considering the equilibrated system and frames at 0 ns, the distances registered between H37 and adenine, as well as between D36 and ribose revealed a closer proximity of the binding site with the non-phosphory-lated cofactor (Figure S25 and Figure S26). The steric hindrance brought about by $PO_3^-$ towards N38 seems to induce a shift of 0.25 Å in the ribose positioning. In the case of the shorter substrate PEN, which does not occupy the entire mutated catalytic pocket, the global larger space left over might play a role. Either way, the additional $PO_3^-$-N38 interaction appeared to increase the anchoring of the phosphorylated cofactor in the binding site, hence the affinity of the cofactor for the protein. This correlates with the average of 1.73 and 2.28-fold better interaction energies at 0 ns in NADPH systems compared to NAD systems, between the adenosine moiety and the protein in M and WT systems, respectively. This tendency remained over the whole course of the MD (Figure S27). This is also in accordance with the *in vitro* experiments displaying 10-fold lower $K_M$ for M_NADP as compared to M_NAD systems. Reversely, when tracing these distances along the simulation path, molecular explanations were difficult to rationalize for a higher mobility of the adenosine ring calculated in the case of M_PEN systems as compared to M_HEP systems (Figure S26).

through its alteration at Q141, it may affect the affinity of the protein towards its cofactor. This correlates with the average of 1.1-fold higher Enzyme-Cofactor interaction energies between WT and M, whichever cofactor and ligand considered (Figure S24). Interestingly, the spatial position equivalent to Q141 (P6) is quite conserved among the nat-AmDH family, except in some cases such as MATOUAmDH1 and 2 (I and L respectively),

## Conclusion

Supported by *in silico* analysis of active site diversity within the nat-AmDH family and using *Cfus*AmDH as a model, we successfully predicted two significant mutation sites in nat-AmDHs. The transposition of the mutation pattern to nine other AmDHs drastically relaxed the nat-AmDHs substrate scope from four-to-five-membered carbon chains up to eight and nine-

**ChemCatChem**

Research Article
doi.org/10.1002/cctc.202200880

Chemistry
Europe

European Chemical
Societies Publishing

membered chains. The biocatalytic efficient formation of *n*-alkyl terminal primary amines from C5 to C9 by these mutants can complement the only reported enzyme *Es*LeuDH-DM active towards corresponding aldehyde substrates.[10] The mutants remained more active towards these aldehydes, which supplements well the substrate scope of engineered AmDHs from α- and ε-AADHs, which have been mainly described for the transformation of prochiral ketones into the corresponding (*R*)-amines. Most of all, the stereoselectivity of the mutants was maintained for the formation of the (*S*)-enantiomers with high *ee*. Bulky amines such as 4-phenylbutan-2-amine (**2k**) and heptan-2-amine (**2h**) can now be obtained with good conversions (52.5% and 53.7% for *Cfus*AmDH–W145A, respectively) and excellent enantiomeric excess (99.6% and 99.7% *ee*, respectively). Crystal structures and MD simulations enabled to gain insights into the overall dynamics of these enzymes and more especially of *Cfus*AmDH–W145A and to postulate the role of some specific residues of the cofactor binding site or mutant active site. This work should greatly facilitate subsequent mutagenesis efforts of this family of enzymes. It also emphasizes that coupling biodiversity to targeted protein engineering can provide efficient biocatalysts with desired properties, being thus complementary to mutagenesis libraries.

## Experimental Section

### General

All the chemicals were purchased from commercial sources (see Supporting Information) and used without additional purification. NMR spectra were recorded on a Bruker (Bruker) 600 MHz spectrometer (Evry University, France) for ¹H and ¹³C experiments. Chemical shifts (expressed in ppm) of ¹H spectra were referenced to the solvent peaks δ(H) = 4.65 for D$_2$O. UHPLC analyses were performed on a UHPLC U3000 RS 1034 bar system (Thermo Fisher Scientific) equipped with a DAD3000 diode array detector and a MSQ Plus™ Single Quadrupole Mass Spectrometer using a Kinetex® F5 (Phenomenex) column (100×2.1 mm; 1.7 μm). GC-FID analyses were performed on a Gas Chromatograph Trace 1300 (Thermo Fisher Scientific) equipped with an autosampler injector AI/AS1310, a Flame Ionization Detector (FID) and a H$_2$ Generator Alliance (FDGSi). Samples were injected on a TG-35MS AMINE (Thermo Fischer Scientific) column (30 m; 0,25 mm; 1 μm). Injector temperature 220°C, split flow 20, split ratio 10, flame 250°C, H$_2$ flow 30 mL min⁻¹. Spectrophotometric assays were recorded on a Safas UVMC2 (Safas) thermostated with a refrigerated/heating circulator Polystat36 (Fisher Scientific) using microcells high-precision cell quartz with 6 mm light path (Hellma Analytics).

### Docking analysis

The mutants *Cfus*AmDH–F140A and *Cfus*AmDH–W145A were modeled from the solved RX structure of *Cfus*AmDH (PDB: 6iau, chain B) using the Wizard mutagenesis tool on PyMOL and rotamer selection. The products pentan-1-amine (**2a**), hexan-1-amine (**2b**), heptan-1-amine (**2c**), octan-1-amine (**2d**), nonan-1-amine (**2e**) and decan-1-amine (**2f**) were docked into the resolved crystal structures of *Cfus*AmDH and into the models *Cfus*AmDH–F140A and *Cfus*AmDH–W145A. The ligand PDB files were generated using the CORINA Molecular Online Tool. With AutoDockTool,[24] the docking

simulations were performed on rigid structures, with no flexibility given to any catalytic pocket residue. The number of Genetic Algorithm (GA) runs was fixed at 10 using the Lamarckian GA (4.2). The 10 ligand conformations obtained were then analyzed in PyMOL. The cavity volume of *Cfus*AmDH, *Cfus*AmDH–F140A and *Cfus*AmDH–W145A were calculated using CAVER Web Tool v1.1, considering the result with 100% relevance score.[25]

### Mutagenesis

QuikChange Site-Directed Mutagenesis Kit® (Agilent technology) (forward primers were used for each mutation in this case) was used according to manufacturer's instructions for the production of enzyme variants. All the primers used are reported in Table S1. Successful cloning and creation of the target mutations were confirmed by sequencing. The original sequence of the enzymes mutated are already reported in Mayol *et al.* 2019 and Caparco *et al.* 2020.[12b,13c] The proteins were produced by heterologous expression in *E. coli* BL-21C+ and purified by nickel affinity chromatography with Ni-NTA column (QIAGEN) as already described.[13c] Protein concentrations were determined by the Bradford method with bovine serum albumin as the standard, and the samples were analyzed by SDS-PAGEs using the Invitrogen NuPAGE system. All the final enzyme concentrations and NuPAGE gels obtained are reported in Table S2 and Figure S5 and S6. The purified proteins were stored at −80°C.

### Large-Scale purification

Large-scale purification of *Cfus*AmDH and *Cfus*AmDH–W145A were conducted from a 500-mL culture by nickel affinity chromatography in tandem with gel filtration (Hi Load 16/600 Superdex 200 pg) as described elsewhere.[26] The storage buffer was 50 mM phosphate pH 7.5, 50 mM NaCl, 10% glycerol and 1 mM DTT. Large-scale purification of *Cfus*AmDH–F140A was conducted using the same protocol except for mechanical sonication instead of using Bug Buster® Protein Extraction Reagent. The final ratio of glycerol has also been increased to 30% due to inactivity and/or instability of this mutant with the usual method used for *Cfus*AmDH and *Cfus*AmDH–W145A (Figure S6).

### Specific activity measurements

All the reactions were conducted in duplicates at 30°C in spectrophotometric cell (6 mm light path) in a final reaction volume of 100 μL. To a mixture of ammonium formate buffer (2 M NH$_4$HCO$_2$/NH$_4$OH, pH 8.5), carbonyl-containing compound (10 mM) was added, the mixture was vortexed and incubated at 30°C for 1 min, then NADH or NADPH (0.2 mM) was added, the mixture was vortexed again and introduced in the pre-heated cell. The reaction was initiated with addition of an appropriate amount of purified enzyme (0.03–0.25 mg mL⁻¹). The initial slope measured at 340 nm determined the specific activity of the enzyme according to Beer–Lambert's law and the molar absorptivity of β-NAD(P)H (ε = 6220 M⁻¹cm⁻¹) after subtraction of the slope obtained under the same conditions without enzyme.

### Biocatalytical reactions monitored by UHPLC-UV

*Cfus*AmDH, *Cfus*AmDH–F140A and *Cfus*AmDH–W145A biocatalytic potential and their tolerance to DMSO was conducted through conversion assays. To a reaction mixture (100 μL in 500 μL Eppendorf tubes) containing 10 mM carbonyl-containing substrate (prepared as solutions in DMSO at different initial concentration for

**ChemCatChem**

Research Article
doi.org/10.1002/cctc.202200880

**Chemistry
Europe**

European Chemical
Societies Publishing

a 1–20% final ratio of DMSO, for 0% the solution was prepared in water at 25 mM but the solubility was still not complete), 0.2 mM NAD$^+$, 0.2 mM NADP$^+$, 3 U mL$^{-1}$ Glucose Dehydrogenase (GDH-105), 1.2 eq. D-glucose in 2 M NH$_4$HCO$_2$/NH$_4$OH buffer (pH 8.5) was added 0.5 mU mg$^{-1}$ of purified enzymes. The calibration points were prepared using various standard amine concentrations (0, 3, 6 and 10 mM) in mixture containing 2 M NH$_4$HCO$_2$/NH$_4$OH buffer (pH 8.5) and an appropriate volume of desalting buffer (average of the volume of enzyme solution used in the reaction mixture). The reaction mixtures and the calibration points were let at 25 °C for 24 h under agitation at 400 rpm. All the reaction mixtures were analyzed after derivatization with benzoyl chloride (BzCl) according to the already described protocol.[13c]

### Screening of nat-AmDHs WT and mutants (GC-FID monitoring)

All the mutants produced were screened for conversion of a range of carbonyl-containing substrates as follows: to a reaction mixture (100 μL in 500 μL Eppendorf tubes) containing 10 mM carbonyl-containing substrate (prepared as solutions in DMSO at 200 mM for a 5% final ratio of DMSO), 0.2 mM NAD$^+$, 0.2 mM NADP$^+$, 3 U mL$^{-1}$ GDH-105, 1.2 eq. D-glucose in 2 M NH$_4$HCO$_2$/NH$_4$OH buffer (pH 8.5) was added 0.5 mU mg$^{-1}$ of purified enzymes. For each substrate, a blank mixture was prepared in the same manner but lacking the enzyme replaced by an appropriate volume of desalting buffer. Calibrations points were prepared as above. The reaction mixtures and the calibration points were let at 25 °C for 24 h under agitation at 400 rpm. The mixtures were extracted in ethyl acetate as follows: to a mix of 20 μL NaOH 10 M and 80 μL of the sample was added 100 μL ethyl acetate, after 5 s vortexing, a volume of 70 μL of the organic layer was recovered (×2). The combined organic layers were then analyzed by GC-FID. The gradients used for each substrate are given in Table S3. A summary of the substrates and mutants tested in this screening is given in Figure S28 and chromatograms are given in Figure S29–S37.

### Enantiomeric Excess and 1k ketone conversion monitored by UHPLC-UV

The conversion of **1k** and the hits obtained with prochiral ketones from the screening in GC-FID were analyzed for enantiomeric excess of the amine formed by derivatization with Marfey's reagent in a 96-well plate as already described,[16] followed by UHPLC-UV analysis (eluent MeOH/H$_2$O 0,1% formic acid with a linear gradient 40/60 during 1 min, then 40/60 to 85/15 in 2 min (hold 3 min), then 85/15 to 40/60 in 1 min and a re-equilibration time of 2 min; flow 0.3 mL min$^{-1}$; temperature 25 °C; injection volume 3 μL; UV detection at λ = 340 nm). A summary of the substrates and mutants tested in this screening is given in Figure S28 and chromatograms are given in Figure S38–S42.

### Kinetic parameters of CfusAmDH-W145A

Kinetic parameters of *Cfus*AmDH–W145A for the carbonyl-containing compound, the cofactor and ammonia were determined by spectrophotometry through NAD(P)H-monitoring at 340 nm. All the reactions were performed in ammonium formate buffer pH 8.5 in a final volume of 100 μL at 30 °C, in spectrophotometer cell with optical paths of 0.6 cm and the procedure was as described in the Specific activities measurements paragraph. Initial rates of the reaction were measured with various concentrations of substrate and saturated concentrations of the other substrates. Data were fitted to the Michaelis and Menten equation or to the equation provided in case of inhibition by excess of substrate, using Sigma

Plot software. The uncertainties are those generated out of two to three experiments (Figure S9–S12).

### Semi-preparative scale reaction

In a 50 mL-Greiner tube equipped with a screw cap was poured **1i** or **1c** (2.5 mL of a 1 M stock solution in DMSO, final ratio of DMSO 5%, 2.5 mmol), distilled water (28.6 mL), ammonium formate pH 8.5 (2 mL of a 10 M stock solution), D-glucose (3 mL of a 1 M stock solution, 3 mmol), NADP$^+$ (1 mL of a 10 mM, 10 μmol), NAD$^+$ (1 mL of a 10 mM, 10 μmol), GDH (Codexis GDH-105, 1.5 mL of a 100 U mL$^{-1}$ stock solution, 150 U) and purified *Cfus*AmDH–W145A (2.4 mL of a 10.5 mg mL$^{-1}$ stock solution). The reaction was shaken at 25 °C, 350 rpm. After 24 h, the conversion was monitored with UHPLC-UV after BzCl derivatization of 20 μL sample. The reaction from **1i** was acidified with HCl > 37% solution to pH < 2 and the product was washed twice with methyl *tert*-butyl ether (MTBE) to remove the remaining ketone (2×15 mL). The aqueous layers were then basified to pH 12 with 10 M KOH solution and the product extracted with MTBE (5×20 mL). The combined layers were then washed with 10 mL 5% m/v LiCl solution to remove the remaining DMSO and ammonia. The organic layers were dried (MgSO$_4$) and filtrated before the addition of a solution of 2 M HCl in diethylether (720 μL, 1.4 mmol, 1.8 eq. of the estimated **2i** produced). Solvent was removed by evaporation under reduced pressure to afford (2S)-octan-2-amine (**2i**) as monohydrochloric salt (116 mg, 28% isolated yield, white solid). Enantiomeric excess 98.4% were obtained by UHPLC-UV after FDAA derivatization of the obtained amine and commercial (S)-**2i** at 0.025–0.05 mM (Figure S14). NMR analysis were in accordance with the commercial standard. $^1$H NMR (600 MHz) δ 3.22 (m, 1H), 1.48 (m, 2H), 1.27–1.14 (m, 8H; d, J = 6.24 Hz, 3H), 0.74 (t, J = 6.90 Hz, 3H). $^{13}$C NMR (150 MHz, D$_2$O) δ 47.9, 34.0, 30.8, 28.0, 24.5, 21.9, 17.7, 13.4 (Figure S15 and S16).

### Crystallization of CfusAmDH-W145A

Pure *Cfus*AmDH–W145A was concentrated to 10 mg mL$^{-1}$ and complexed with either 10 mM NAD$^+$ or NADP$^+$. Protein was subjected to crystallization trials using commercially available screens in 96-well plate format in which 300 nL drops, comprising 150 nL protein solution and 150 nL of precipitant solution were employed. The best NAD$^+$ crystals were obtained using 0.1 M Tris-HCl buffer 8.5 with 0.2 M MgCl$_2$.6H$_2$O, 8% (w/v) PEG 20,000 and 8% (v/v) PEG 500 MME. The best NADP$^+$ crystals were obtained using 0.1 M Bis-Tris buffer pH 5.5 with 25% (w/v) PEG 3350. These crystals were incubated with a solution of the mother liquor containing 10 mM pentan-1-amine (**2a**) and incubated for 30 min prior to fishing. Crystals were flash-cooled from the crystallization drops in liquid nitrogen and tested for diffraction using a Rigaku Micromax-007HF fitted with Osmic multilayer optics and a MARRESEARCH MAR345 imaging plate detector. Those crystals displaying diffraction of > 3 Å resolution were retained for data collection at the synchrotron.

### Data collection and processing, structure solution and refinement

Datasets for crystals of *Cfus*AmDH–W145A in complex with NAD$^+$ or NADP$^+$ and **2a** were collected on beamline I03 at the Diamond Light Source Synchrotron in Oxford, U.K. Data, which were collected to 1.64 and 1.50 Å respectively, were processed and integrated using XDS[27] and scaled using SCALA[28] as part of the Xia2 processing system.[29] Data collection statistics are given in Table S6. The structures were solved with the program MOLREP,[30] using a monomer of the WT *Cfus*AmDH (PDB code 6IAU[13c]) as a model.

Each solution contained two molecules in the asymmetric unit. The structures were built and refined using iterative cycles within the programs COOT[31] and REFMAC,[32] respectively. Following building of the protein and water molecules in the NADP$^+$ structure, clear density was observed adjacent to the nicotinamide ring of NADP$^+$. This was successfully modeled and refined as the added ligand **2a**. The NAD$^+$ and NADP$^+$ complexes were refined to $R_{cryst}/R_{free}$ values of 17.8/21.4 and 16.4/18.8%, respectively. The structures were validated upon deposition within the Protein DataBank (PDB). Refinement statistics are presented in Table S6. Coordinates for the *Cfus*AmDH–W145A structures in complex with NAD$^+$ or NADP$^+$ and **2a** have been deposited in the PDB with the accession codes 7QZN and 7QZL, respectively.

### Molecular dynamics

MD simulations were executed using the GROMACS software (version 2021. 3), with CHARMM36 force field, following the GROMACS tutorial on protein/ligand complex.[33] In order to generate MD simulations starting from fully equilibrated structures, we first considered every system composed of Enzyme-Cofactor-Ligand, where Enzyme is either *Cfus*AmDH (WT) or mutant *Cfus*AmDH–W145A (M), Cofactor is either NAD$^+$ (NAD) or NADP$^+$ (NADP) and eventually, Ligand is either pentan-1-amine (**2a**) (PEN) or heptan-1-amine (**2c**) (HEP). Each system was placed in a dodecahedron box, filled with TIP3P waters to model water solvatation, then counter-ions were added to reach system neutralization. The borders of the simulation box were cut off at 10 Å from the surface of the protein. A first step of energy minimization was carried out through steepest descent algorithm, to relax the eventual steric constraints. Then, each system followed a two-step protocol equilibration, with position restraints applied to the protein, co-factor and substrate heavy atoms. The first phase ran a 100 ps NVT (constant volume) simulation followed by a second 100 ps NPT (constant pressure) simulation to maintain pressure isotropically at 1.0 bar. All production steps were carried out under NPT conditions with 2 fs step size. The LINCS algorithm was used to constrain the lengths of hydrogen containing bonds and the waters were restrained using the SETTLE algorithm. Van der Waals forces were treated using a 12 Å cut-off. Long-range electrostatic forces were treated using the particle mesh Ewald method (PME). A number of four independent MD simulations (with random initial velocities at NVT equilibration) were executed per system. Each production contains a trajectory of 100 ns, with a snapshot saved every 100 ps, thus resulting in 1000 frames (counting the starting structure) per simulation. These frames were submitted to analysis and the systems were compared using the average of all the four simulations performed. Visualization and figures of protein structures were generated using PyMOL Schrödinger LLC.[34]

## Conflict of Interest

The authors declare no conflict of interest.

## Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

[1] O. I. Afanasyev, E. Kuchuk, D. L. Usanov, D. Chusov, *Chem. Rev.* **2019**, *119*, 11857–11911.

[2] S. D. Roughley, A. M. Jordan, *J. Med. Chem.* **2011**, *54*, 3451–3479.

[3] F. G. Mutti, T. Knaus, *Enzymes Applied to the Synthesis of Amines.in Biocatalysis for Practitioners* (Eds.: G. De Gonzalo, I. Lavandera), Wiley-VCH, Weinheim, Germany **2021**, Ch. 6, pp 143–180.

[4] a) S. A. Kelly, S. Mix, T. S. Moody, B. F. Gilmore, *Appl. Microbiol. Biotechnol.* **2020**, *104*, 4781–4794; b) I. Slabu, J. L. Galman, R. C. Lloyd, N. J. Turner, *ACS Catal.* **2017**, *7*, 8263–8284; c) F. Guo, P. Berglund, *Green Chem.* **2017**, *19*, 333–360; d) M. Fuchs, J. E. Farnberger, W. Kroutil, *Eur. J. Org. Chem.* **2015**, 6965–6982; e) A. Gomm, E. O'Reilly, *Curr. Opin. Chem. Biol.* **2018**, *43*, 106–112; f) M. D. Patil, G. Grogan, A. Bommarius, H. Yun, *Catalysts* **2018**, *8*, 254.

[5] a) S. A. Kelly, S. Pohle, S. Wharry, S. Mix, C. C. R. Allen, T. S. Moody, B. F. Gilmore, *Chem. Rev.* **2018**, *118*, 349–367; b) M. D. Patil, G. Grogan, A. Bommarius, H. Yun, *Catalyst* **2018**, *8*, 254; c) K. Savile, J. M. Janey, E. C. Mundorff, J. C. Moore, S. Tam, W. R. Jarvis, J. C. Colbeck, A. Krebber, F. J. Fleitz, J. Brands, P. N. Devine, G. W. Huisman, G. J. Hughes, *Science* **2010**, *329*, 305–309.

[6] a) M. J. Abrahamson, E. Vazquez-Figueroa, N. B. Woodall, J. C. Moore, A. S. Bommarius, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 3969–3972; b) M. J. Abrahamson, J. W. Wong, A. S. Bommarius, *Adv. Synth. Catal.* **2013**, *355*, 1780–1786; c) G. A. Aleku, S. P. France, H. Man, J. Mangas-Sanchez, S. L. Montgomery, M. Sharma, F. Leipold, S. Hussain, G. Grogan, N. J. Turner, *Nat. Chem.* **2017**, *9*, 961–969; d) L. Ducrot, M. Bennett, G. Grogan, C. Vergne-Vaxelaire, *Adv. Synth. Catal.* **2021**, *363*, 328–351; e) J. Mangas-Sanchez, M. Sharma, S. C. Cosgrove, J. I. Ramsden, J. R. Marshall, T. W. Thorpe, R. B. Palmer, G. Grogan, N. J. Turner, *Chem. Sci.* **2020**, *11*, 5052–5057; f) J. R. Marshall, P. Yao, S. L. Montgomery, J. D. Finnigan, T. W. Thorpe, R. B. Palmer, J. Mangas-Sanchez, R. A. M. Duncan, R. S. Heath, K. M. Graham, D. J. Cook, S. J. Charnock, N. J. Turner, *Nat. Chem.* **2021**, *13*, 140–148; g) S. L. Montgomery, A. Pushpanath, R. S. Heath, J. R. Marshall, U. Klemstein, J. L. Galman, D. Woodlock, S. Bisagni, C. J. Taylor, J. Mangas-Sanchez, J. I. Ramsden, B. Dominguez, N. J. Turner, *Sci. Adv.* **2020**, *6*, eaay9320; h) V. Tseliou, T. Knaus, M. F. Masman, M. L. Corrado, F. G. Mutti, *Nat. Commun.* **2019**, *10*, 3717.

[7] R. D. Franklin, C. J. Mount, B. R. Bommarius, A. S. Bommarius, *Chem-CatChem* **2020**, *12*, 2436–2439.

[8] D.-H. Wang, Q. Chen, S.-N. Yin, X.-W. Ding, Y.-C. Zheng, Z. Zhang, Y.-H. Zhang, F.-F. Chen, J.-H. Xu, G.-W. Zheng, *ACS Catal.* **2021**, *11*, 14274–14283.

[9] X. Mu, T. Wu, Y. Mao, Y. Zhao, Y. Xu, Y. Nie, *ChemCatChem* **2021**, *13*, 5243–5253.

[10] J. Löwe, A. Siewert, A.-C. Scholpp, L. Wobbe, H. Gröger, *Sci. Rep.* **2018**, *8*, 10436.

[11] H. Groger, *Appl. Microbiol. Biotechnol.* **2019**, *103*, 83–95.

[12] a) M. Bennett, L. Ducrot, C. Vergne-Vaxelaire, G. Grogan, *ChemBioChem* **2022**, *23*, e202200136; b) A. A. Caparco, E. Pelletier, J. L. Petit, A. Jouenne, B. R. Bommarius, V. de Berardinis, A. Zaparucha, J. A. Champion, A. S. Bommarius, C. Vergne-Vaxelaire, *Adv. Synth. Catal.* **2020**, *362*, 2427–2436.

[13] a) S. Lee, H. Jeon, P. Giri, U.-J. Lee, H. Jung, S. Lim, S. Sarak, T. P. Khobragade, B.-G. Kim, H. Yun, *Biotechnol. Bioprocess Eng.* **2021**, *26*, 384–391; b) J. Liu, W. Kong, J. Bai, Y. Li, L. Dong, L. Zhou, Y. Liu, J. Gao, R. T. Bradshaw Allen, N. J. Turner, Y. Jiang, *Chem. Catal.* **2022**, *2*, 1494–1498; c) O. Mayol, K. Bastard, L. Beloti, A. Frese, J. P. Turkenburg, J.-L. Petit, A. Mariage, A. Debard, V. Pellouin, A. Perret, V. de Berardinis, A. Zaparucha, G. Grogan, C. Vergne-Vaxelaire, *Nat. Catal.* **2019**, *2*, 324–333.

[14] O. Mayol, S. David, E. Darii, A. Debard, A. Mariage, V. Pellouin, J.-L. Petit, M. Salanoubat, V. de Berardinis, A. Zaparucha, C. Vergne-Vaxelaire, *Catal. Sci. Technol.* **2016**, *6*, 7421–7428.

[15] R.-F. Cai, L. Liu, F.-F. Chen, A. Li, J.-H. Xu, G.-W. Zheng, *ACS Sustainable Chem. Eng.* **2020**, *8*, 17054–17061.

[16] L. Ducrot, M. Bennett, A. A. Caparco, J. A. Champion, A. S. Bommarius, A. Zaparucha, G. Grogan, C. Vergne-Vaxelaire, *Front. Catal.* **2021**, *1*, 781284.

[17] I. G. Riziotis, A. J. M. Ribeiro, N. Borkakoti, J. M. Thornton, *J. Mol. Biol.* **2022**, *434*, 167517.

[18] a) F. Chen, S. C. Cosgrove, W. R. Birmingham, J. Mangas-Sanchez, J. Citoler, M. Thompson, G.-W. Zheng, J.-H. Xu, N. J. Turner, *ACS Catal.* **2019**, *9*, 11813–11818; b) F. Chen, G.-W. Zheng, L. Liu, H. Li, Q. Chen, F.-L. Li, C.-X. Li, J.-H. Xu, *ACS Catal.* **2018**, *8*, 2622–2628.

[19] V. Tseliou, M. F. Masman, W. Böhmer, T. Knaus, F. G. Mutti, *ChemBioChem* **2019**, *20*, 800–812.

[20] L. Liu, D.-H. Wang, F.-F. Chen, Z.-J. Zhang, Q. Chen, J.-H. Xu, Z.-L. Wang, G.-W. Zheng, *Catal. Sci. Technol.* **2020**, *10*, 2353–2358.

[21] a) T. Knaus, W. Böhmer, F. G. Mutti, *Green Chem.* **2017**, *19*, 453–463; b) A. Pushpanath, E. Siirola, A. Bornadel, D. Woodlock, U. Schell, *ACS Catal.* **2017**, 3204–3209; c) M. Sharma, J. Mangas-Sanchez, S. P. France, G. A. Aleku, S. L. Montgomery, J. I. Ramsden, N. J. Turner, G. Grogan, *ACS Catal.* **2018**, *8*, 11534–11541; d) L. J. Ye, H. H. Toh, Y. Yang, J. P. Adams, R. Snajdrova, Z. Li, *ACS Catal.* **2015**, *5*, 1119–1122.

[22] E. P. J. Jongkind, A. Fossey-Jouenne, O. Mayol, A. Zaparucha, C. Vergne-Vaxelaire, C. E. Paul, *ChemCatChem* **2022**, *14*, e202101576.

[23] a) E. Rajakumara, S. Abhishek, K. Nitin, D. Saniya, P. Bajaj, U. Schwaneberg, M. D. Davari, *ACS Chem. Biol.* **2022**, *17*, 266–280; b) D. D. Wang, L. Ou-Yang, H. Xie, M. Zhu, H. Yan, *Comput. Struct. Biotechnol. J.* **2020**, *18*, 439–454.

[24] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, A. J. Olson, *J. Comput. Chem.* **2009**, *30*, 2785–2791.

[25] J. Stourac, O. Vavra, P. Kokkonen, J. Filipovic, G. Pinto, J. Brezovsky, J. Damborsky, D. Bednar, *Nucleic Acids Res.* **2019**, *47*, W414-W422.

[26] N. Perchat, P.-L. Saaidi, E. Darii, C. Pellé, J.-L. Petit, M. Besnard-Gonnet, V. de Berardinis, M. Dupont, A. Gimbernat, M. Salanoubat, C. Fischer, A. Perret, *Proc. Nat. Acad. Sci.* **2018**, *115*, E4358-E4367.

[27] W. Kabsch, *Acta Crystallogr. Sect. D* **2010**, *66*, 125–132.

[28] P. Evans, *Acta Crystallogr. Sect. D* **2011**, *67*, 282–292.

[29] G. Winter, *J. Appl. Crystallogr.* **2010**, *43*, 186–190.

[30] A. Vagin, A. Teplyakov, *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.

[31] P. Emsley, K. Cowtan, *Acta Crystallogr. Sect. D* **2004**, *60*, 2126–2132.

[32] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallogr. Sect. D* **1997**, *53*, 240–255.

[33] Lemkul, J. A.; "GROMACS tutorial, Protein-Ligand Complex", can be found under http://www.mdtutorials.com/gmx/complex/index.html.

[34] a) M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, *SoftwareX* **2015**, *1–2*, 19–25; b) H. J. C. Berendsen, D. van der Spoel, R. van Drunen, *Comput. Phys. Commun.* **1995**, *91*, 43–56; c) B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435–447; d) E. Lindahl, B. Hess, D. van der Spoel, *J. Mol. Model.* **2001**, *7*, 306–317; e) S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, E. Lindahl, *Bioinformatics* **2013**, *29*, 845–854; f) D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. C. Berendsen, *J. Comput. Chem.* **2005**, *26*, 1701–1718; g) H. Bekker, H. J. C. Berendsen, E. J. Dijkstra, S. Archterop, R. Van Drunen, D. Van Der Spoel, A. Sijbers, H. Keegstra, *Phys. Comput.* **1993**, *92*, 252–256.
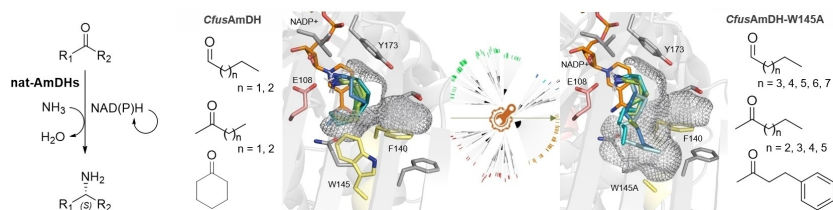
# RESEARCH ARTICLE



Biodiversity is a good inspiring source for protein engineering. A targeted mutation in the active site of diverse native Amine Dehydrogenases allowed to extend their substrate scope to linear aliphatic aldehydes and ketones up to C8. The structure and dynamics of these enzymes are now better understood thanks to crystal structure of the mutant and molecular dynamics studies.

L. Ducrot, M. Bennett, Dr. G. André-Leroux, Dr. E. Elisée, S. Marynberg, A. Fossey-Jouenne, Prof. A. Zaparucha, Prof. G. Grogan, Dr. C. Vergne-Vaxelaire*

**Expanding the Substrate Scope of Native Amine Dehydrogenases through *In Silico* Structural Exploration and Targeted Protein Engineering**