Main Manuscript for

An adenine dinucleotide variant of the nickel-pincer nucleotide cofactor in ornithine cyclodeaminases.

Benoît Desguin.

Dr. Benoît Desguin,
Louvain Institute of Biomolecular Science and Technology (LIBST)
Université catholique de Louvain, B-1348 Louvain-La-Neuve, Belgium.

*B. Desguin.

Email: benoit.desguin@uclouvain.be

Author Contributions: B.D conceived the project, carried out the experiments and wrote the article.

Competing Interest Statement: Disclose any competing interests here.

Classification: Biochemistry (major).

Keywords: NPN cofactor • ornithine cyclodeaminase • adenine dinucleotide

This PDF file includes:

Main Text
Figures 1 to 5
Tables 1 to 2

Abstract

The nickel-pincer nucleotide (NPN), which was first identified in lactate racemase, is a widespread cofactor present in around one-fifth of bacterial species and more than half of archaeal species. However, only a fraction of the enzymes utilizing this cofactor have been identified. In this study, we identified several novel enzyme families that are potentially dependent on NPN, including one that is usually linked to arginine catabolism. Four enzymes from this family, which were isolated from various genetic sources, exhibited ornithine cyclodeaminase (OCD) activity when co-expressed with the NPN biosynthetic enzymes. Biochemical analysis revealed that these enzymes utilize a unique cofactor: a nickel-pincer adenine dinucleotide (NPAD), a novel derivative of NPN. Cofactor exchange experiments and mass spectrometry confirmed NPAD as the active cofactor, which fits better in the structure of OCDs than NAD^+. This finding expands the known roles of nickel-pincer cofactors, introducing NPAD as a critical cofactor in bacteria, where it supports nitrogen remobilization and amino acid fermentation, and in archaea, where it is involved in L-proline biosynthesis. This discovery not only highlights the versatility of the nickel-pincer cofactor and suggests the presence of other undiscovered enzyme families that may rely on NPAD.

Introduction

Tridentate planar metal complexes, also known as pincer complexes, have long been recognized by chemists for decades as useful catalysts (1), yet the first enzyme harboring a natural nickel-pincer complex, the lactate racemase (LarA) and its nickel-pincer nucleotide (NPN) cofactor, was only characterized recently (2). This cofactor combines the properties of a nicotinic acid-derived scaffold, enabling it to accept a hydride, with a nickel ion coordinated by two sulfur atoms and forming a nickel-carbon bond with the C4 position of the pyridinium ring. It has been hypothesized that the NPN cofactor functions in the reversible capture of a hydride, in a manner similar to that of nicotinamide adenine dinucleotide (NAD^+) (2). In addition, the presence of the nickel bond implies that the hydride can be transferred to the nickel (3, 4), with the nickel ion exhibiting catalytic properties of its own (5). The NPN cofactor biosynthetic

pathway starts with nicotinic acid adenine dinucleotide (NaAD) and involves three biosynthetic enzymes: LarB, LarE, and LarC (6-8). A bioinformatics analysis revealed that *larA*, *larB*, *larC*, and *larE* are present in 10% of the genomes, indicating their widespread distribution among prokaryotes (9). LarA homologs form an enzyme superfamily whose identified members catalyze racemization or C2-epimerization of α-hydroxyacids (10, 11). Furthermore, NPN biosynthesis enzyme genes were identified without an accompanying *larA* homolog in another 15% of the studied genomes, suggesting that NPN is also used by other enzymes that are not homologous to LarA (9). These enzymes could be involved in very different types of biochemical reactions that remain to be explored. In this study, I identified and characterized a new enzyme family that catalyzes ornithine cyclodeamination and depends on the presence of a dinucleotidic variant of the NPN cofactor.

Materials and Methods

Materials and growth conditions

Bacterial strains and plasmids used in the present study are listed in Table S1. Chemicals were purchased from Merck. *L. lactis* NZ3900 (12) was grown in M17 broth supplemented with 0.5% glucose at 28 °C with chloramphenicol (10 $mg \cdot L^{-1}$). To induce genes under the control of the *nisA* expression signals, nisin A was added during the early exponential phase (OD_{600} = 0.3-0.4) at a concentration of 1 $\mu g \cdot L^{-1}$ and the cells were collected after 4 h. *E. coli* DH10B cells were grown with agitation at 37 °C in lysogeny broth with ampicillin (200 $mg \cdot L^{-1}$). When expressing pBADHisA derivatives, induction was initiated by addition of L-arabinose (0.2%) during the early exponential phase and cells were grown at 37 °C for 4 h with agitation.

DNA techniques

DNA was introduced into *E. coli* and *L. lactis* cells by electrotransformation (13, 14). PCR amplifications used Q5 high-fidelity DNA polymerase (New England Biolabs). The primers used in this study were purchased from Eurogentec and are listed in Table S1. For construction of plasmid pGIR330, *agrE* was amplified from *Synechocystis* sp. PCC 6803 chromosome with

the primers Sll_NheI_A and SllCOG_Pst1_B, digested by NheI and PstI and cloned into a similarly digested pBADHisA. For construction of plasmid pGIR331, *nadD* was amplified from *L. lactis* chromosome with the primers NadDa_A and NadDa_B, digested by NheI and HindIII and cloned into similarly digested pBADHisA. For plasmids pGIR220, pGIR221, pGIR222, and pGIR223 the *agrE*, *mxOCD*, *auOCD*, and *eaOCD* genes were amplified from *Synechocystis* sp. PCC 6803, *M. xanthus* DSM 16526, *A. urinae* CCUG 36881, and *E. asparagiformis* DSM 15981 chromosomes with the primer pairs SLL1336_A + SLL1336_B, MxCOG1915_A + MxCOG1915_B2, AuCOG1915A_A3 + AuCOG1915B_B2, and ClasCOG1915a_A + ClasCOG1915a_B, respectively, digested by PciI and NheI and cloned into a similarly digested pGIR210. For construction of plasmid pGIR083, *mxOCD* was amplified from the *M. xanthus* DSM 16526 chromosome with the primers MxCOG1915_A + MxCOG1915_B2, digested by PciI and NheI and cloned into a NcoI- and NheI-digested pGIR082. For construction of plasmid pGIR224, the synthetic gene of *mmOCD* was purchased from IDT, digested with PciI and NheI and cloned into a similarly digested pGIR210. For construction verification by PCR and sequencing (realized by Microsynth), primers pNZ8048_UP' and LarZ-B_B were used. For construction of plasmid pGIR222_M4, pGIR222 plasmid was amplified with primers AuCOG_M4_A + AuCOG_M4_B, digested with DpnI and transformed into *L. lactis* cells, following QuickChange protocol for site-directed mutagenesis (15). The presence of the mutation was screened with primers AuCOG_M4_V and LarZ-B_B.

Protein purification and in vitro *NPN biosynthesis*

Strep-tagged proteins were purified following published protocols on a Strep-tactin XT column (7). His-tagged proteins (NMAT and AgrE) were purified by following published protocols on a Ni-NTA column (16). Protein concentrations were estimated with the Bradford assay. The nickel content was estimated using 4-(2-pyridylazo)resorcinol, as described previously (9).

Enzymatic assays

The general protocol for assaying OCD activity was as follows: 5 µL of OCD sample was added to 95 µL reaction mix from the Megazyme Ammonia Assay Kit (Rapid) containing 50 mM L-

ornithine (if not otherwise stated), and the absorbance at 340 nm was directly measured every min for 30 min with an Infinite 200 PRO plate reader (Tecan). The linear rate of absorbance decrease was used for measuring the enzyme rate. The arginine dihydrolase activity was measured similarly, with L-arginine as substrate.

Concentrations of the substrate were varied in order to measure the kinetic parameters. For pH and temperature dependent studies, the OCDs were incubated with 50 mM L-ornithine in 20 µL reactions with the appropriate buffers (see below) and the reaction was stopped after 30 min by incubation at 80°C for 5 min, then the product concentration was assayed with the Megazyme Ammonia Assay Kit (Rapid) without L-ornithine. The final difference in absorbance was used for measuring the enzyme rate. For pH dependence studies, citric acid buffer (pH 2 to pH 4), acetic acid buffer (pH 4 to pH 6), phosphate buffer (pH 6 to pH 8) and borate buffer (pH 8 to pH 10) were used at room temperature. For temperature studies, the enzymes were incubated at the indicated temperatures in Tris-HCl buffer pH 8. For stability studies, 50 µL samples of purified OCDs were incubated at room temperature and 5 µL samples were retrieved at the indicated times for measurement of OCD activity as described above. For the stability of MmOCD in presence of substrate, the activity of 5 µL of MmOCD sample was followed every minute for 30 minutes as described above with 20 mM L-ornithine. For the inhibition of OCDs by L-proline; 0, 1, or 10 mM of L-proline were added to general protocol for assaying OCD activity and a concentration of 10 mM L-ornithine was used.

For cofactor transfer experiments, a 20 µL reaction volume was used. AgrE, $LarA_{Tt}$, or NMAT were used at 1 µM concentration; AuCOG, $AuCOG^{H209A}$, or LarC at 5 µM concentration, denaturation was performed by 1-min incubation at 80 °C. The reaction mixture was incubated for 1 min at room temperature (for the NPAD biosynthetic reactions, the incubation was 1 hour instead of 1 minute) and 5 µL of OCD samples were added to 95 µL reaction mix from the Megazyme Ammonia Assay Kit (Rapid) containing 50 mM L-ornithine, and the absorbance at 340 nm directly measured as described above. Biosynthetic NPN was synthesized in a mix of purified LarB, LarE, and LarC as described before (17) and used at an approximative NPN

final concentration of 1 µM. The Lar assay used the D-/L-Lactic Acid Assay Kit from Megazyme as described (17) with L-lactate as substrate.

For the purification of NPAD, 1 mL of biosynthetic NPN was incubated for 1 hour in presence of 2 µM NMAT. Then, 10 mM β-mercaptoethanol (BME) was added, the proteins were denatured by 1-min incubation at 80 °C, the supernatant was collected by centrifugation at 20,000g for 2 min, loaded onto a 5-ml HiScreen Q HP column (GE Healthcare) using an Akta Purifier (GE Healthcare), washed with Tris-HCl buffer (30mM, pH 8) containing 1 mM BME, and eluted with a gradient of NaCl (0 to 1 M) in Tris-HCl buffer (30 mM, pH 8) containing 1 mM BME, as previously described for NPN (7). The NPAD peak was identified by absorbance at 430 nm and confirmed by OCD activity tests as described above with 1 µM purified AgrE.

Mass spectrometry analysis of intact protein

Protein samples (around 100 µg) were concentrated using Amicon® Ultra Centrifugal Filters (10 kDa molecular weight cut-off) and diluted with 2.5 mM ammonium acetate buffer three times, then diluted two times in 2.5 mM ammonium acetate – 2% DMSO. Intact protein analyses were done on a SYNAPT G2-Si high definition mass spectrometer (Waters) equipped with a NanoLockSpray dual electrospray ion source (Waters). Coated fused silica PicoTipTM Econo 12 Emitters for nanoelectrospray, outer diameters: 1 µm Tip (New Objective, USA) were filled with 10 µL of samples and placed on the Universal NanoFlow™ Sprayer (Waters). The eluent was sprayed at a spray voltage of 2.2 KV. The source temperature was set to 80°C. The cone gas flow was 20 liters/h with a nano flow gas pressure of 0.5 bar. MS spectra were acquired and processed with MassLynx software (Waters). MaxEnt1 deconvolution algorithm was used to extract the exact masses. Range from 38000 Da to 42000 and from 78000 to 82000 Da, 1 Da resolution and 0.4 Da width at half height Uniform Gaussian method were used.

Mass spectrometry analysis of NPAD

100 µL of solution containing cofactor was acidified with 10 µL of 1% TFA, desalted and concentrated with C18 ZipTip® (Millipore) according to protocol manufacturer.

Cofactor solution was separated by reverse phase chromatography on a NanoACQUITY UPLC MClass system (Waters) working with MassLynx V4.1 (Waters) software. 5 µL of sample was injected on a trap C18, 100Å 5 µm, 180 µm x 20 mm column (Waters) and desalted using a short isocratic condition with a flow rate of 8 µL/min using a 99% formic acid and 1% (v/v) ACN buffer for 0.5 min to avoid to lose the cofactor in the waste. Sample was subjected to reverse phase chromatography on a C18, 100Å 1.8 µm, 75 µm x 150 mm column (Waters) PepMap for 35 min at 35°C at a flow rate of 300 nL/min using a two part linear gradient from 1% (v/v) ACN, 0.1 % formic acid to 40 % (v/v) ACN, 0.1 % formic acid for 15 min and from 40% (v/v) ACN, 0.1 % formic acid to 85 % (v/v) ACN, 0.1 % formic acid for 10 min. The column was re-equilibrated at initial conditions after washing 30 min at 85% (v/v) ACN, 0.1 % formic acid at a flow rate of 300 nL/min. For online LC-MS analysis, the nanoUPLC was coupled to the mass spectrometer through a nano-electrospray ionization (nanoESI) source emitter.

DDA (Data Dependent Analysis) analysis were performed on an SYNAPT G2-Si high definition mass spectrometer (Waters) equipped with a NanoLockSpray dual electrospray ion source (Waters). Precut fused silica PicoTipR Emitters for nanoelectrospray, outer diameters: 360 µm; inner diameter: 20 µm; 10 µm tip; 2.5” length (Waters) were used for samples and Precut fused silica TicoTipR Emitters for nanoelectrospray, outer diameter: 360 µm; inner diameter: 20 µm; 2.5” length (Waters) were used for the lock mass solution. The eluent was sprayed in negative mode at a spray voltage of 2.8 kV with a sampling cone voltage of 25 V and a source offset of 30 V. The source temperature was set to 80°C. The cone gas flow was 20 liters/h with a nano flow gas pressure of 0.4 bar and the purge gas was turned off. The SYNAPT G2Si instrument was operated in DDA (data-dependent mode), automatically switching between MS and MS2. Full scan MS and MS2 spectra (m/z 50- 2000) were acquired directly after injection to 35 min in resolution mode (20,000 resolution FWHM at m/z 400) with a scan time of 0.1 sec. Tandem mass spectra of up to 10 precursors were generated in the trapping region of the ion mobility cell by using a collision energy ramp from 17/19 V (low mass, start/end) to up to 65/75 V (high mass, start/end). Charged ions (-1, -2,) are selected to be submitted to the MSMS

fragmentation over the m/z range from 50 to 2000 with a scan time of 0.25 sec. For the post-acquisition lock mass correction of the data in the MS method, the doubly charged monoisotopic ion of [Glu1]-fibrinopeptide B was used at 100 fmol/µl using the reference sprayer of the nanoESI source with a frequency of 30 s at 0.5 µl/min into the mass spectrometer.

Data were processed with MassLynx V4.1 (Waters). Spectrum were extracted from the TIC (Total Ion Chromatogram) and submitted to a search of masses of interest.

In silico analyses

Kinetic parameters and 95% confidence interval were estimated using Prism (GraphPad). The COG co-occurrence study utilized Excel (Microsoft). The genomic context study was done with the online tool GeConT 2 (18). The sequences selected for the multiple alignment were the COG1915 genes from the GeConT 2 database, with the addition of *eaOCD*. The multiple alignment was performed using ClustalX (19) and the phylogenetic tree was visualized using Dendroscope 3 (20). The model of MxOCD with L-ornithine and NPAD was obtained from the Chai Discovery server (21). Figures were made and the catalytic site residues were selected based on visual inspection of the structures of AgrE (PDB code 6LRG) and EaOCD model with The PyMOL Molecular Graphics System, Version 3.0 Schrödinger, LLC.

Results and discussion

Three enzyme families other than the LarA superfamily are potentially NPN-dependent.

I hypothesized that novel families of NPN-dependent enzymes would be present in genomes harboring the genes encoding the NPN biosynthetic enzymes and would be absent from those lacking these genes. To identify these enzyme families, the Cluster of Orthologous Groups (COG) database was searched (22) for COGs that co-occur with the NPN biosynthetic enzyme LarE. Each COG was assigned a score equal to the number of co-occurrences minus the number of absences in the presence of LarE (Fig. 1*A*). As expected, the COGs of the other NPN biosynthetic enzymes (LarB and LarC) showed the highest scores, followed by four other COGs: COG3875, which corresponds to the enzymes of the LarA superfamily, COG1453, which corresponds to a predicted oxidoreductase of the aldo/keto reductase family, and

COG2006 along with COG1915 which correspond to uncharacterized conserved proteins. The observation that the LarB, LarC, and LarA families showed the highest scores confirmed the reliability of this *in silico* analysis.

Assuming the identified enzyme families are NPN-dependent, I tested whether they could explain the presence of the NPN-biosynthetic genes in so many genomes. The COG3875 group alone can explain the presence of *larE* in only 54% of all *larE*-containing genomes. However, the addition of the other three potential NPN-dependent enzyme families increases this percentage to 94% (Fig. 1*A*, black line). This observation shows that by taking these three families of enzymes into account makes it now possible to explain the presence of the NPN-biosynthetic genes in most genomes where they are found.

Next, I sought to identify the function of these three families of potential NPN-dependent enzymes by investigating the genomic context of their corresponding genes, in case these genes are transcriptionally linked in operons with other genes participating to the same function (23). The GeConT tool (18) was searched for the genes that occur most frequently in their genomic context (three genes before and three genes after) (Fig. 1*B*). The genomic context analysis of COG1453 is not informative as the most frequently observed COGs are only observed for 5% of the genes. COG2006 genes are often found alongside Fe-S cluster proteins (17% of genes) and glycosyltransferases (8% of genes), which suggests that these enzymes may exchange electrons with Fe-S cluster proteins and indicates a potential link with carbohydrate metabolism. Finally, the genomic context of COG1915 enzymes shows an association with arginine aminohydrolases or arginases (31% of genes), indicating a clear link with arginine catabolism. I thus focused on the COG1915 family of enzymes. One member of this family, AgrE, is a bifunctional enzyme exhibiting arginine dihydrolase activity of the N-terminal domain and ornithine cyclodeaminase (OCD) activity of the C-terminal domain (Fig. 2*A*) (24), although this latter domain does not belong to the μ-crystallin family, to which NAD^+-dependent OCDs belong (25). OCD activity was observed when the enzyme was purified from

Anabaena cells, but not when purified from *Escherichia coli* cells (24). The enzyme's structure was obtained in the presence of NAD^+, which the authors believed to be necessary for OCD activity (26), but it was suggested that a cyanobacterial factor was missing from the enzyme (27).

I therefore hypothesized that the AgrE C-terminal domain, and by extension the enzymes of the COG1915 family, are NPN-dependent OCDs. Indeed, *Cyanobacteria* and other microorganisms which harbor COG1915 genes typically harbor the NPN-biosynthetic genes as well (Fig. 1*A*), whereas *E. coli* does not (9).

Enzymes of the COG1915 family are ornithine cyclodeaminases requiring the *in vivo* expression of LarBCE.

AgrE was overexpressed and purified from *E. coli* cells (Fig. S1*A*) and its activity was tested in the presence of both L-arginine and of L-ornithine, to confirm its arginine dihydrolase and OCD activities, with and without *in vitro*-synthesized NPN. The results showed that AgrE hydrolyzes L-arginine, but no OCD activity was observed, even with added NPN (Fig. 2*B*). I hypothesized that NPN might be required during *in vivo* production of the protein rather than after its isolation. Consequently, the AgrE-encoding gene was cloned into the pGIR210 plasmid (10), enabling AgrE to be overexpressed and purified from *Lactococcus lactis* cells co-expressing LarBCE *in vivo*. Unfortunately, it was not possible to purify AgrE under these conditions, possibly due to protein misfolding or enzyme toxicity in *L. lactis* cells (Fig. S1*A*).

The OCD homolog of the COG1915 family from *Myxococcus xanthus* (hereafter named MxOCD) was overexpressed and purified using *L. lactis* cells that co-expressed LarBCE *in vivo* (Fig. S1*B*) and OCD activity was finally observed (Fig. 2*C*). The level of activity was strongly reduced when no nickel was added to the growth medium and no activity was observed when MxOCD was purified from cells that did not co-express LarBCE, as expected for a NPN-dependent enzyme. However, activity was still absent when *in vitro*-synthesized NPN was added, as observed for AgrE (Fig. 2*B and C*). This suggests that the NPN needs to be produced and incorporated into the enzyme *in vivo*, and not *in vitro*, for OCD activity. The

absorbance spectrum of the purified MxOCD showed a peak of absorbance around 450 nm (Fig. 2*D*), which is typical of the NPN cofactor (2), further confirming that a NPN-like cofactor is present in MxOCD.

Similar to MxOCD, three other members of the COG1915 family from the following species were purified: *Aerococcus urinae*, (hereafter named AuOCD), *Enterocloster asparagiformis* (hereafter named EaOCD), and *Methanococcus maripaludis* (hereafter named MmOCD) (Fig. S1*B*). Two bands are present for the purified AuOCD, as two paralogs (AuOCD1 and AuOCD2) are present in this species, with their corresponding genes located next to each other within the same operon. When both paralogs were overexpressed simultaneously in *L. lactis*, they copurified (Fig. S1*B*), even when only one paralog was strep-tagged, demonstrating that they form heterodimers. Purification yields were typically less than 0.5 mg/L of culture for all proteins, except for EaOCD, which was purified in much greater quantities (typically more than 5 mg/L of culture). All purified enzymes exhibited OCD activity when purified from *L. lactis* cells co-expressing LarBCE. The kinetic parameters of all four OCDs were measured (Table 1, Figs S2, S3 and S4). The observed k_{cat} values were highly variable, ranging from 0.085 s^{-1} for EaOCD to 12.3 s^{-1} for MxOCD, as were the observed K_m values, ranging from 0.19 mM for AuOCD to 34 mM for EaOCD. Due to this variability, the enzymatic efficiencies varied by four orders of magnitude between the different OCDs (Table 1, Fig. S2). These variabilities may be partially explained by the differences observed in nickel content, which is used as a proxy for cofactor content (Table 1). Due to the small quantity of purified enzymes, the measured nickel contents were imprecise; however, they were always equal to or less than 50% for all purified OCDs, suggesting that the observed k_{cat} values are most probably underestimated, particularly for EaOCD, which exhibited the lowest nickel content.

The apparent optimal temperatures for most OCDs range from 45 to 55 °C, which is within the expected range for enzymes from mesophilic bacteria. However, AuOCD exhibited an optimal temperature of 25-30 °C, which is below the optimal growth temperature of *A. urinae*, around 37 °C (Table 1 and Fig. S3). The optimal pH was generally around pH 8.0, except for EaOCD,

which showed an optimal pH of 6.5 (Table 1 and Fig. S4). The lower optimal pH for EaOCD may be related to the enzyme's high K_m and low k_{cat} values (Table 1), which may indicate that the catalytic activity is not optimal in the *in vitro* assay conditions.

The time-dependency of the enzymes' OCD activity was also assessed, given that NPN-dependent enzymes are generally unstable over time (2). Most OCDs were relatively stable over time, with a half-life ranging from around 1 day for EaOCD to 3 days for MxOCD. However, AuOCD showed a half-life of around 30 minutes (Fig. S5*A*). AuOCD rapid loss of OCD activity can be attributed to its low optimal temperature, both are probably the result of its instability. The stability of MmOCD was greatly reduced in the presence of substrate, with its half-life of activity reduced from around 1.5 days in the absence of substrate (Fig. S5*A*) to 10 minutes in its presence (Fig. S5*B*).

AuOCD and MxOCD did not seem to be inhibited by their product, L-proline, unlike EaOCD and especially MmOCD, which were inhibited by L-proline (Fig. S5*C*). A lower affinity for the substrate thus seems to correlate with a higher affinity for the reaction product, resulting in product inhibition.

Overall, these results demonstrate that all investigated OCDs exhibit good stability, except MmOCD in the presence of its substrate and AuOCD both with and without substrate. The investigated OCDs possess similar biochemical properties to other NPN-dependent enzymes, albeit with lower catalytic efficiencies (10, 11), except EaCOD, which exhibited suboptimal activity probably due to a low cofactor loading.

OCDs of the COG1915 family depend on a nickel-pincer adenine dinucleotide cofactor.

The binding of NAD^+ to the AgrE OCD domain (26) and the requirement for NPN biosynthesis *in vivo* for OCD activity (Fig. 2*C*) suggest the presence of a new cofactor in OCDs of the COG1915 family: a nickel-pincer adenine dinucleotide (NPAD). Mass spectrometry analysis of several purified OCDs revealed an intact mass that corresponded exactly to the predicted mass of the apoenzyme minus its N-terminal methionine (Fig. S6), indicating that the cofactor is not covalently bound to the protein, like NPN to LarA (2).

To explore the cofactor content of these OCDs, several cofactor transfer experiments were conducted. First, the cofactor present in an OCD holoenzyme was transferred to an OCD apoprotein to demonstrate the ability to exchange the cofactor between the enzymes. AgrE, when purified from *E. coli*, is inactive, as is AuOCD when heat-denatured (Fig. 3*A*). However, when AgrE was incubated in the presence of denatured AuOCD, OCD activity was restored, suggesting that the cofactor released from denatured AuOCD was taken up by AgrE. Furthermore, the observed activity was three times greater than the original activity of AuOCD, suggesting that AgrE has a higher k_{cat} than AuOCD. To confirm that the observed OCD activity observed is not due to the refolding of AuOCD in the presence of AgrE, the inactive H209A variant of AuOCD was used. The H209 residue, which is in AuOCD1, was chosen because it is a conserved residue that is probably important for catalysis or cofactor binding. Denatured AuOCD^{H209A} was able to activate AgrE, demonstrating that the cofactor present in AuOCD could be transferred to another OCD upon denaturation of the enzyme.

In a second cofactor transfer experiment, denatured OCD holoenzyme was incubated with LarA apoprotein, which could be activated by NPN synthesized *in vitro*. No Lar activity was detected following this incubation unless LarC and manganese ions were added to the reaction (Fig. 3*B*). As LarC has been shown to hydrolyze pyridinium-3,5-bisthiocarboxylic acid adenine dinucleotide (P2TAD) into its corresponding mononucleotide (28), it is most likely that LarC is hydrolyzing NPAD into NPN, thereby activating LarA apoprotein. The addition of CTP was not required, demonstrating that the hydrolytic activity of LarC is indeed necessary, rather than its nickel insertion activity. In fact, adding CTP reduced this activation (Fig. 3*B*), suggesting that CTP inhibits the binding and/or hydrolysis of NPAD by LarC.

In the final cofactor transfer experiment, NPAD was synthesized from NPN *in vitro*. Nicotinate mononucleotide adenylyltransferases (NMATs) are present in most living organisms as part of the NAD^+ biosynthesis pathway, using Mg·ATP to convert nicotinate mononucleotide into NaAD, the precursor of NAD^+. As NPN appears to be converted into NPAD *in vivo*, it is likely that endogenous NMAT is involved in this conversion. NMAT from *L. lactis* was purified and

incubated with *in vitro* synthesized NPN, Mg·ATP, and AgrE. OCD activity was observed when all the elements were present, and absent when NPN or NMAT was missing. The residual activity in absence of added ATP is likely due to residual Mg·ATP from the NPN biosynthetic reaction (Fig. 3*C*). This final experiment demonstrates that NPN can be converted into NPAD by NMAT.

Overall, the cofactor transfer experiments suggest that NPAD can be converted into NPN by the hydrolytic activity of LarC and that NPN can be converted into NPAD by NMAT in the presence of Mg·ATP (Fig. 3*D*).

In vitro synthesized NPAD was then subjected to anion exchange chromatography. Two peaks which exhibited absorbance at 430 nm were observed, but only the fractions corresponding to the first peak were able to generate OCD activity in presence of AgrE (Fig. S7). These fractions were analyzed by liquid chromatography–mass spectrometry (LC-MS) and the observed monoisotopic mass, isotopic distribution, and fragmentation pattern corresponded to NPAD (Fig. S8). The other peak observed is the mononucleotide form of NPN which has not been converted into NPAD and which is unable to activate AgrE.

The ability to activate AgrE apoprotein with NPAD enabled the kinetic properties of the AgrE C-terminal domain to be studied (Table 2, Figs S2-S5), which are relatively similar to those of previously identified OCDs (Table 1). However, AgrE exhibited substrate inhibition with a K_i of 320 mM, a characteristic not observed in other OCDs.

In conclusion, in vitro cofactor transfer experiments and LC-MS analyses definitely confirms that NPAD is the cofactor used by OCDs of the COG1915 family.

Structural analysis of OCDs of the COG1915 family.

As the structure of AgrE with NAD^+ had previously been obtained (26), the AgrE OCD domain structure was compared with the MxOCD model structure with docked NPAD and L-ornithine, which was obtained using Chai-1 (Fig. 4) (21). This comparison highlights a difference in the domain organization of the two OCDs: AgrE OCD domain comprises two small N-terminal domains: one domain composed of two α-helices and three β-sheets (hereafter named αβ

domain), and another domain composed of seven α-helices (hereafter named the 7α domain), followed by the C-terminal Rossman fold domain that binds the cofactor. In contrast, MxOCD only possesses the small N-terminal 7α domain and the C-terminal Rossman fold domain. Furthermore, in AgrE the αβ and 7α domains are situated on one side of the Rossman fold domain, whereas in MxOCD, the 7α domain is situated on the other side of the Rossman fold domain (Fig. 4*A* and *B*). This additional domain, as well as the organization of domains, could explain why AgrE exhibits substrate inhibition, whereas other OCDs do not (Table 1 and 2). However, the active site is very similar for both OCDs and lies at the interface between the two monomers. The residues forming the catalytic site, which are probably involved in catalysis, have been identified (Fig. 4*C*). Most of these residues belong to one monomer, with four residues (M59, D60, N296 and K326 in MxOCD) being provided by the other monomer. In AgrE, only two residues are provided by the other monomer due to the different domain organization. One histidine residue (H182 in MxOCD structure) is located 2.3 Å from the nickel ion, consistent with its role in coordinating it, as observed in LarA (2). One arginine residue (R253) is located 2.4 Å from the second thiocarboxylic acid of NPAD, consistent with the formation of a salt bridge. These interactions would be absent if NPAD was replaced by NAD^+, thereby confirming that NPAD fits better into the structure.

The position of L-ornithine, as modelled by Chai-1, is perfectly positioned for a hydride abstraction by NPAD and proton abstraction by an aspartate residue (D255 in MxOCD structure). This would facilitate a proton-coupled hydride transfer reaction mechanism (Fig. 4*C* and *D*), analogous to that hypothesized for NAD^+-dependent OCDs (25). Another aspartate residue is also well positioned to abstract a proton from the terminal amine of L-ornithine, facilitating the cyclization of the substrate (Fig. 4*D*). When MxOCD structure is modelled with NPN or NAD^+, the angle formed by the α-carbon of L-ornithine, the C4 and the N of the pyridinium ring appears to be less favorable for hydride transfer. This angle indeed increases from 107.1° with NPAD (Fig. 4*C*), similarly to D-lactate in LarA (29), to 117.7° and 123.3° with

NPN and NAD^+, respectively (Fig. S10). This observation could be one reason for the absence of reaction with NPN or NAD^+.

Furthermore, an alignment of the sequences of the purified OCDs was realized and shows that most of the identified active site residues are conserved among all OCDs, with the notable exceptions of AuOCD2 and the residues from AuOCD1 which participate in the active site of AuOCD2 (Fig. S10). The presence of a Glu residue instead of an Ala residue at the position corresponding to A178 in MxOCD prevents AuOCD2 from binding to any nucleotidic cofactor (Fig. 4*C*), suggesting that it is catalytically inactive. This hypothesis is consistent with the observation that the $AuOCD^{H209A}$ variant completely abolishes the OCD activity of AuOCD (Fig. 3*A*). In conclusion, the sequences and structural models of the OCDs of the COG1915 family are consistent with the presence of NPAD in their active site.

Roles of OCDs of the COG1915 family

A phylogenetic tree was created with 309 members of the COG1915 family to investigate the distribution of NPAD-dependent OCDs in bacteria and archaea, (Fig. 5 and Dataset 1). This phylogenetic tree revealed the widespread distribution of OCD homologs in bacteria and archaea. Most of the identified OCDs were monofunctional; the only bifunctional OCDs observed were in *Cyanobacteria*. Some *Bacillota* also exhibited the presence of two heterodimer-forming OCD paralogs, one of each inactive, similarly to AuOCD. The role of the inactive paralog requires further investigation.

The genomic context of these OCDs was also analyzed (Fig. 5). Arginine dihydrolase genes were frequently found in association with OCDs in many bacteria, primarily in *Actinobacteria* and *Planctomycetota*, and were fused to OCDs in *Cyanobacteria*. This congruence indicates that all OCDs in these species probably have a similar function, most likely for nitrogen remobilization, as has been suggested for *Anabaena* (24).

In archaea, OCDs are sometimes found in association with arginases, confirming that they function as OCDs (Fig. 5). *Methanocaldococcus jannaschii* has been shown to synthesize L-proline through OCD activity, but the enzyme responsible for this activity has not yet been

identified (30). Potential NAD^+-dependent OCDs of the μ-crystallin family exhibit either alanine dehydrogenase activity or pyrroline-2-carboxylate reductase activity in archaea (31, 32). This suggests that L-proline biosynthesis in many archaea is achieved via NPAD-dependent OCDs of the COG1915 family (33, 34).

In *Bacillota* and *Proteobacteria*, OCDs are frequently found in association with carbamate kinases (ArcC) (Fig. 5). In this case, arginine deiminases (ArcA) and L-ornithine carbamoyltransferases (ArgF) are often associated, too, similarly to what is observed in *E. asparagiformis* and *A. urinae* (Fig. S11*A*). These enzymes form the arginine deiminase pathway, which generates ATP through the deamination of arginine into citrulline and the production of carbamoyl-phosphate (35), and is widely used in bacteria to generate metabolic energy. The addition of an OCD would enable L-ornithine to be utilized as a carbon and/or electron source instead of being exported as a waste product (36) (Fig. S11*B*). OCD activity is also required for the fermentation of L-arginine, L-citrulline, and L-ornithine in the Stickland metabolism of proteolytic *Clostridia*. In *Clostridium sporogenes*, an NAD^+-dependent OCD is thought to be involved in this pathway (37). However, a NPAD-dependent OCD is also present (Dataset 1), indicating that both types of OCD might catalyze the same reaction, highlighting the importance of this reaction in proteolytic *Clostridia*.

In summary, the functions of the COG1915 family of OCDs depend on the organism in which they are expressed: nitrogen remobilization in *Actinobacteria*, *Planctomycetota*, and *Cyanobacteria*; L-arginine, L-citrulline, and L-ornithine fermentation in *Bacillota* and *Proteobacteria*; and L-proline biosynthesis in archaea.

Conclusion and perspectives

This study describes a new family of OCDs present in 11% of all bacterial and archaeal genomes included in the COG database (22). This family of enzymes has various functions, including nitrogen remobilization, L-arginine, L-citrulline and L-ornithine fermentation, and L-proline biosynthesis. L-proline plays a significant role in the pharmaceutical and chemical industries, as it is involved in the formation of several short peptide drugs and catalytic

processes, and OCD activity is one pathway leading to L-proline (39). Furthermore, reversing the OCD activity could enable the non-reductive amination of α-hydroxy acids. This would eliminate the need for a strong reducing agent in these reactions and could lead to industrial applications in amino acid synthesis.

This enzyme family requires a dinucleotide variant of the NPN cofactor: the NPAD cofactor. Similar to the flavin cofactors FMN and FAD, it shows that the NPN cofactor also exists in different nucleotidic forms. To improve clarity, I propose to rename the NPN nickel-pincer mononucleotide (NPMN) and to use the term 'NPN' for both NPMN and NPAD. The catalytic properties of NPAD and NPMN are probably highly similar and they are both used for the transient oxidation of the substrate. This enables isomerization in the case of LarA homologs (3) and deamination in the case of OCDs, similarly to NAD^+-dependent OCDs (25). This functional similarity in function raises the question of why two different enzyme families catalyze the same reaction using two different cofactors. Citing Prof. Zamble in (38): "why an *organism would go to* the *trouble* of *making such* a *fancy cofactor*?", a question all the more intriguing given that NAD^+ can catalyze the same reaction. Using NPAD instead of NAD^+ might give NPAD-dependent OCDs a higher catalytic efficiency, as suggested by the k_{cat}/K_m of 44 $mM^{-1}s^{-1}$ observed for MxOCD (Table 1), which is nine times greater than that observed for the NAD^+-dependent OCD from *Pseudomonas putida* (39).

This new family of OCDs is the first family of enzymes described featuring the NPAD cofactor, a dinucleotide variant of NPN cofactor. This confirms my initial hypothesis claiming that other NPN-dependent enzyme families remain to be discovered, and extends the role of NPN-dependent enzymes beyond racemization and epimerization. Furthermore, identifying NPAD-dependent OCDs now opens the door to the investigation of other unknown NPN-dependent enzyme families, which were identified by *in silico* analysis (Fig. 1). This research lays the groundwork for a new field of research on NPN-dependent enzymes, with implications for biochemistry and microbial physiology, as well as numerous potential applications in biocatalysis and metabolic engineering.

Acknowledgments

Funding was provided by the UCLouvain Fond Special de Recherche and by the FNRS (Belgium) (PDR 40027734). BD is Research Associate at FNRS (Belgium). I thank Kevin Lafontaine, Thomas Hanquet, David Baland, and Satyam Kumar Yadav for their help in carrying out this research and Hervé Degand for the mass spectrometry analyses.

References

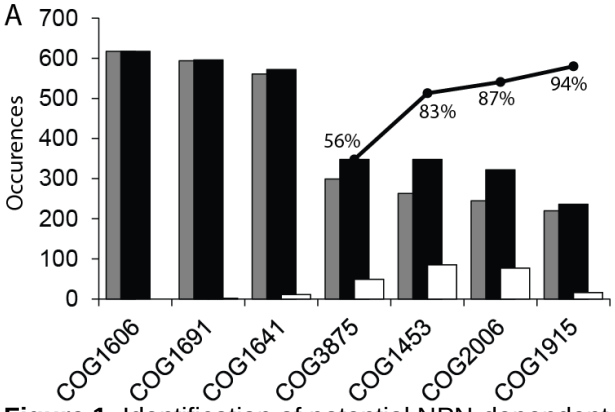
1. K. J. Szabó, "Pincer Complexes as Catalysts in Organic Chemistry" in Organometallic Pincer Chemistry, G. van Koten, D. Milstein, Eds. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2013), pp. 203-241.
2. B. Desguin *et al*. A tethered niacin-derived pincer complex with a nickel-carbon bond in lactate racemase. *Science* **349**, 66-69 (2015).
3. J. A. Rankin *et al.*, Lactate Racemase Nickel-Pincer Cofactor Operates by a Proton-Coupled Hydride Transfer Mechanism. *Biochemistry* **57**, 3244-3251 (2018).
4. S. Gatreddi, D. Sui, R. P. Hausinger, J. Hu, Irreversible inactivation of lactate racemase by sodium borohydride reveals reactivity of the nickel-pincer nucleotide cofactor. *ACS Catal* **13**, 1441-1448 (2023).
5. J. L. Boer, S. B. Mulrooney, R. P. Hausinger, Nickel-dependent metalloenzymes. *Arch. Biochem. Biophys.* **544**, 142-152 (2014).
6. B. Desguin, P. Soumillion, P. Hols, R. P. Hausinger, Nickel-pincer cofactor biosynthesis involves LarB-catalyzed pyridinium carboxylation and LarE-dependent sacrificial sulfur insertion. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 5598-5603 (2016).
7. B. Desguin *et al.*, Biosynthesis of the nickel-pincer nucleotide cofactor of lactate racemase requires a CTP-dependent cyclometallase. *J. Biol. Chem.* **293**, 12303-12317 (2018).
8. M. Fellner, J. A. Rankin, B. Desguin, J. Hu, R. P. Hausinger, Analysis of the Active Site Cysteine Residue of the Sacrificial Sulfur Insertase LarE from *Lactobacillus plantarum*. *Biochemistry* **57**, 5513-5523 (2018).
9. B. Desguin *et al.*, Lactate racemase is a nickel-dependent enzyme activated by a widespread maturation system. *Nat. Commun.* **5**, 3615 (2014).
10. B. Desguin *et al.*, Uncovering a superfamily of nickel-dependent hydroxyacid racemases and epimerases. *Sci Rep* **10**, 18123 (2020).
11. J. Urdiain-Arraiza, A. Vandenberghe, G. Dimitrova, B. Desguin, Unveiling 14 novel 2-hydroxy acid racemization and epimerization reactions in the lactate racemase superfamily. *J. Biol. Chem.* **301**, 108069 (2025).
12. P. G. de Ruyter, O. P. Kuipers, W. M. de Vos, Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. *Appl. Environ. Microbiol.* **62**, 3662-3667 (1996).
13. G. Zhang, D. A. Mills, D. E. Block, Development of chemically defined media supporting high-cell-density growth of lactococci, enterococci, and streptococci. *Appl. Environ. Microbiol.* **75**, 1080-1087 (2009).
14. S. J, F. E, M. T, Molecular cloning: a laboratory manual, second edition. *New York, NY, USA: Cold Spring Harbour Laboratory Press* (1989).
15. L. Zheng, U. Baumann, J. L. Reymond, An efficient one-step site-directed and site-saturation mutagenesis protocol. *Nucleic Acids Res.* **32**, e115 (2004).
16. C. L. Kielkopf, W. Bauer, I. L. Urbatsch, Purification of Polyhistidine-Tagged Proteins by Immobilized Metal Affinity Chromatography. *Cold Spring Harb. Protoc.* **2020**, 102194 (2020).

17. R. P. Hausinger, J. Hu, B. Desguin, The nickel-pincer coenzyme of lactate racemase: A case study of uncovering cofactor structure and biosynthesis. *Methods Enzymol.* **685**, 341-371 (2023).
18. R. Ciria, C. Abreu-Goodger, E. Morett, E. Merino, GeConT: gene context analysis. *Bioinformatics* **20**, 2307-2308 (2004).
19. M. A. Larkin *et al.*, Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947-2948 (2007).
20. D. H. Huson, C. Scornavacca, Dendroscope 3: an interactive tool for rooted phylogenetic trees and networks. *Syst. Biol.* **61**, 1061-1067 (2012).
21. C. Discovery *et al.*, Chai-1: Decoding the molecular interactions of life, bioRxiv 2024.10.10.615955 (2024).
22. M. Y. Galperin *et al.*, COG database update: focus on microbial diversity, model organisms, and widespread pathogens. *Nucleic Acids Res.* **49**, D274-D281 (2021).
23. M. D. Ermolaeva, O. White, S. L. Salzberg, Prediction of operons in microbial genomes. *Nucleic Acids Res.* **29**, 1216-1221 (2001).
24. M. Burnat, S. Picossi, A. Valladares, A. Herrero, E. Flores, Catabolic pathway of arginine in *Anabaena* involves a novel bifunctional enzyme that produces proline from arginine. *Mol. Microbiol.* **111**, 883-897 (2019).
25. J. L. Goodman *et al.*, L-ornithine cyclodeaminase: structure, mechanism of action, and implications for the mu-crystallin family. *Biochemistry* **43**, 13883-13891 (2004).
26. H. Lee, S. Rhee, Structural and mutational analyses of the bifunctional arginine dihydrolase and L-ornithine cyclodeaminase AgrE from the cyanobacterium *Anabaena*. *J. Biol. Chem.* **295**, 5751-5760 (2020).
27. E. Flores, Arginine catabolism enzyme AgrE/ArgZ likely involves a cyanobacterial specific factor. *J. Biol. Chem.* **295**, 2915 (2020).
28. T. Vucko, D. Strilets, P. Soumillion, B. Desguin, S. P. Vincent, Chemo-enzymatic synthesis of NPN cofactor taking advantage of ADP-ribosyl cyclase and LarC cyclometallase promiscuous activities. *Bioorganic chemistry* **153**, 107879 (2024).
29. Gatreddi S, Urdiain-Arraiza J, Desguin B, Hausinger RP, Hu J. Structural Basis for the Catalysis and Substrate Specificity of a LarA Racemase with a Broad Substrate Spectrum. *ACS Catal*. **15**, 2857-2866 (2025).
30. M. Graupner, R. H. White, *Methanococcus jannaschii* generates L-proline by cyclization of L-ornithine. *J. Bacteriol.* **183**, 5203-5205 (2001).
31. I. Schroder, A. Vadas, E. Johnson, S. Lim, H. G. Monbouquette, A novel archaeal alanine dehydrogenase homologous to L-ornithine cyclodeaminase and mu-crystallin. *J. Bacteriol.* **186**, 7680-7689 (2004).
32. S. Watanabe, Y. Tozawa, Y. Watanabe, L-ornithine cyclodeaminase/mu-crystallin homolog from the hyperthermophilic archaeon *Thermococcus litoralis* functions as a novel Delta(1)-pyrroline-2-carboxylate reductase involved in putative trans-3-hydroxy-l-proline metabolism. *FEBS Open Bio* **4**, 617-626 (2014).
33. Y. Fukuyama *et al.*, Development of a rapid and highly accurate method for (13)C tracer-based metabolomics and its application on a hydrogenotrophic methanogen. *ISME Commun* **4**, ycad006 (2024).
34. N. Goyal, H. Widiastuti, I. A. Karimi, Z. Zhou, A genome-scale metabolic model of *Methanococcus maripaludis* S2 for CO2 capture and conversion to methane. *Mol. Biosyst.* **10**, 1043-1054 (2014).
35. M. H. Beck, M. Flaiz, F. R. Bengelsdorf, P. Durre, Induced heterologous expression of the arginine deiminase pathway promotes growth advantages in the strict anaerobe *Acetobacterium woodii*. *Appl. Microbiol. Biotechnol.* **104**, 687-699 (2020).
36. T. Pols, S. Singh, C. Deelman-Driessen, B. F. Gaastra, B. Poolman, Enzymology of the pathway for ATP production by arginine breakdown. *FEBS J.* **288**, 293-309 (2021).

37. Y. Liu *et al.*, *Clostridium sporogenes* uses reductive Stickland metabolism in the gut to generate ATP and produce circulating metabolites. *Nat Microbiol* **7**, 695-706 (2022).
38. D. Zamble, It costs more than a nickel. *Science* **349**, 35-36 (2015).
39. M. Long *et al.*, Directed Evolution of L-ornithine Cyclodeaminase Using an EvolvR-Based Growth-Coupling Strategy for Efficient Biosynthesis of l-Proline. *ACS Synth Biol* **9**, 1855-1863 (2020).

Figures and Tables

A



B

Genomic context of COG1453 genes
COG1259: bifunctional DNase/RNase (5%)
COG0307: riboflavin synthase alpha chain (5%)
COG1309: DNA-binding protein, AcrR family (5%)
Genomic context of COG2006 genes
COG2768: uncharacterized Fe-S cluster protein (9%)
COG0348: polyferredoxin NapH (8%)
COG0438: glycosyltransferase (8%)
Genomic context of COG1915 genes
COG1834: arginine aminohydrolase (24%)
COG0730: permease (7%)
COG0010: arginase/agmatinase family enzyme (7%)

Figure 1. Identification of potential NPN-dependent enzyme families. (*A*) COGs co-occurrence score with LarE (COG1606) (grey bar), corresponding to the total co-occurrences (black bar) minus the number of presences of the COG in the absence of LarE (white bar). COG1691, COG1641, and COG3875 are positive controls corresponding to LarB, LarC, and LarA. The black line indicates the percentage of *larE*-containing genomes harboring at least one NPN-dependent enzyme, when incorporating these additional COG families. (*B*) Three most often observed COGs and their typical genomic contexts for the three probable NPN-dependent COGs. The numbers in parentheses indicate the percentages of genes with this COG in their genomic context.

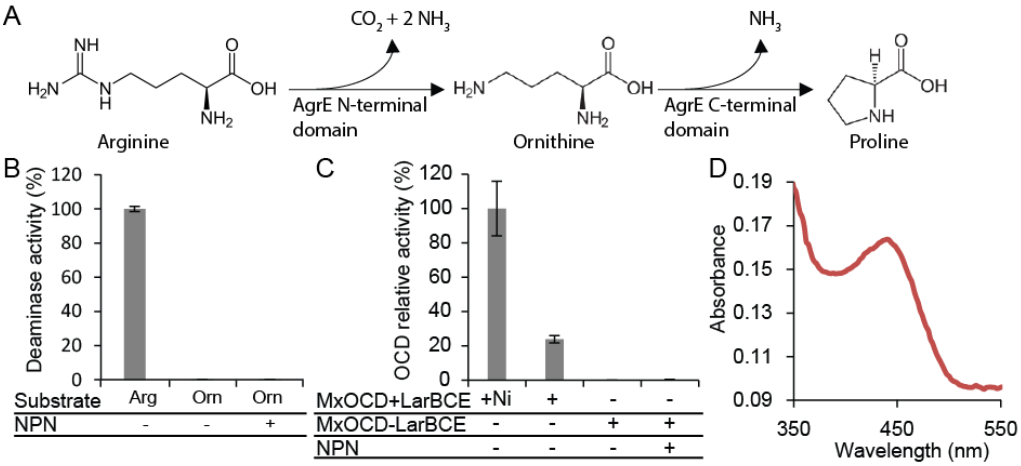


Figure 2. OCDs of the COG1915 family are NPN-dependent enzymes. (*A*) Reactions catalyzed by AgrE. (*B*) Deaminase activity of purified AgrE in the presence of different substrates and *in vitro* synthesized NPN. The deaminase activity with L-arginine was set to 100%. (*C*) OCD relative activity of MxOCD purified from *L. lactis* cells co-expressing LarBCE (+LarBCE) or not co-expressing LarBCE (-LarBCE). "+Ni" indicates that 1 mM $NiSO_4$ was added to the growth medium. The activity of MxOCD+LarBCE +Ni was set to 100%. (*D*) Absorbance scan of purified MxOCD at 1 mg/mL concentration. The error bars represent the standard deviation (N=4) in panels *B* and *C*.

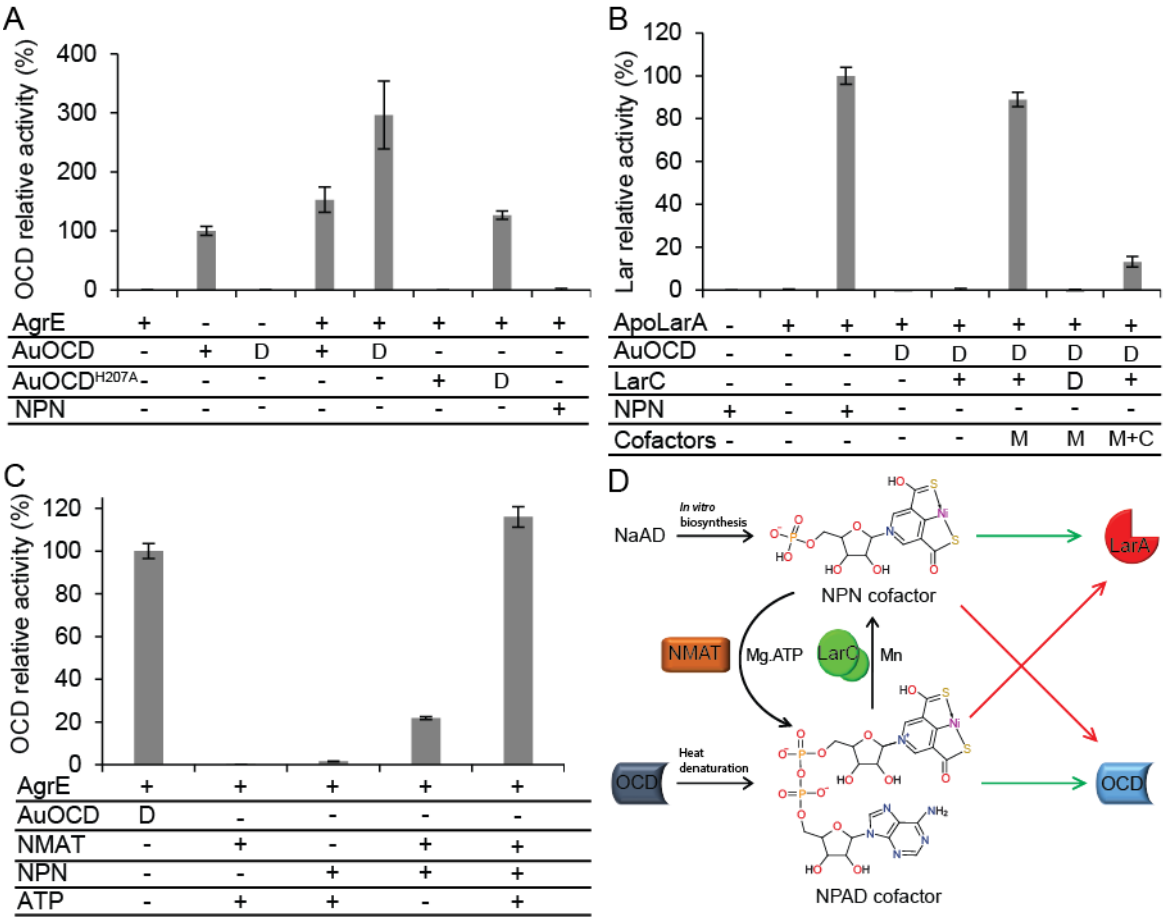


Figure 3. Cofactor transfer experiments. (*A*) OCD→OCD cofactor transfer. AgrE is purified from *E. coli* and lacks the cofactor, whereas AuOCD and AuOCD^{H209A} were purified from *L. lactis* cells coexpressing LarBCE and contain the cofactor. "D" signifies that the protein was heat denatured. NPN is *in vitro* synthesized NPN. OCD activity of AuOCD was set to 100%. (*B*) OCD→Lar cofactor transfer. Same legend as for *A* and "ApoLarA" is LarA apoprotein from *Thermoanaerobacterium thermosaccharolyticum*, "M" means 1 mM $MnCl_2$ and "C" indicates 1 mM CTP. Lar activity of ApoLarA with NPN was set to 100%. (*C*) NPN→NPAD cofactor conversion. Same legend as in *A* with NMAT being nicotinate mononucleotide adenylyltransferase. OCD activity of AgrE with denatured AuOCD was set to 100%. (*D*) Summary of cofactor transfer experiments. Red arrows indicate no activation, green arrows indicate activation observed. NaAD: nicotinic acid adenine dinucleotide. The error bars represent the standard deviation (N=4) in panels *A*, *B*, and *C*.

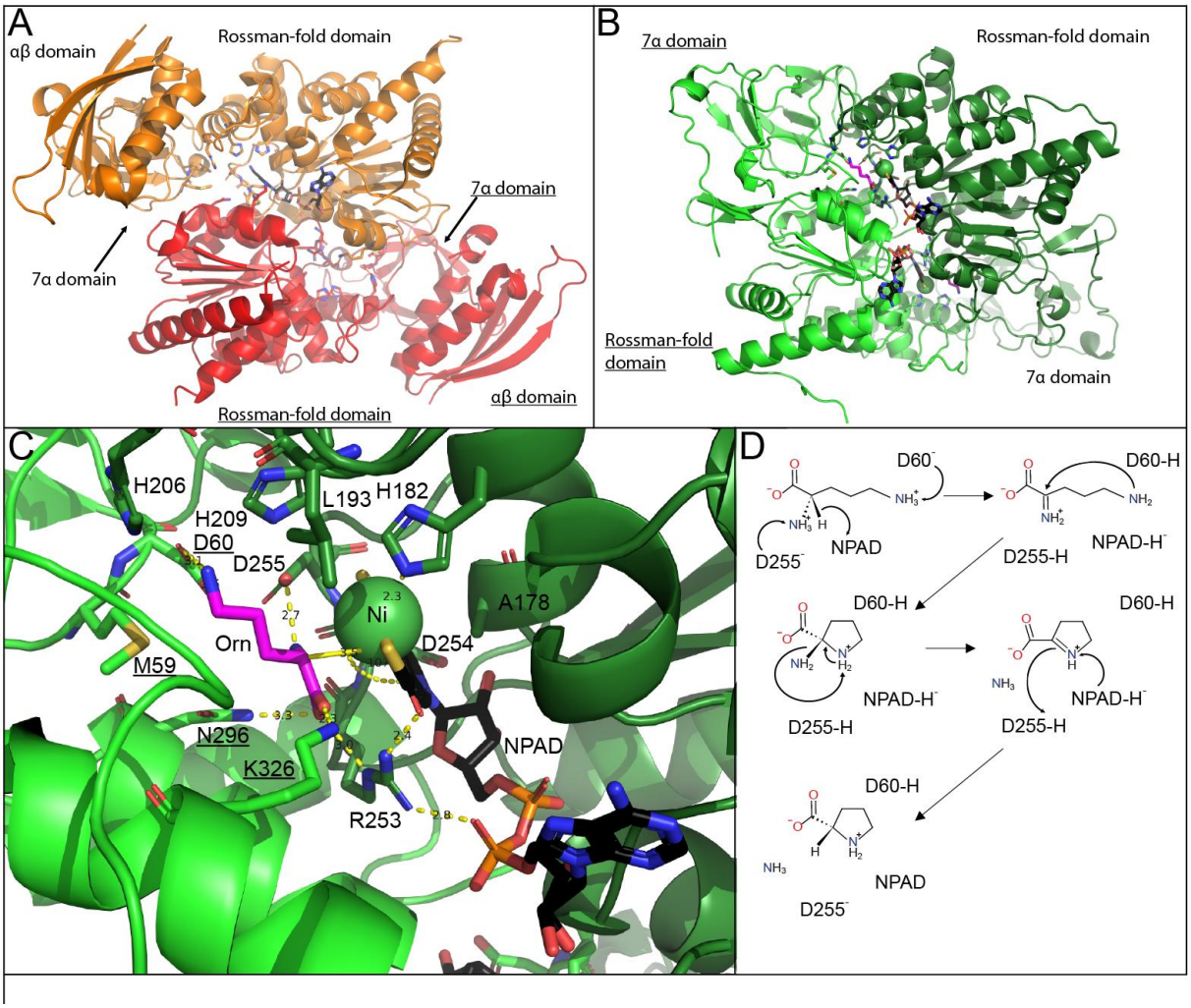


Figure 4. Structure of NPAD-dependent OCDs. (*A*) and (*B*) Ribbon structures of a dimer of the OCD domain of AgrE (PDB code 6LRG) and of a dimer of MxOCD with docked NPAD and L-ornithine (obtained with Chai-1 webserver), respectively. Stick representation is used for the residues likely to be involved in catalysis (with O atoms in red, N in blue, and S in yellow), for L-ornithine (C atoms in magenta), and for NPAD (C atoms in black, P atoms in orange, and Ni as a green sphere). The labels for the second monomer domains are underlined. (*C*) Expanded view of the catalytic site of *B*. Distances are indicated in Å and angles in degrees. The labels for residues from the second dimer are underlined. (*D*) Proposed catalytic mechanism.

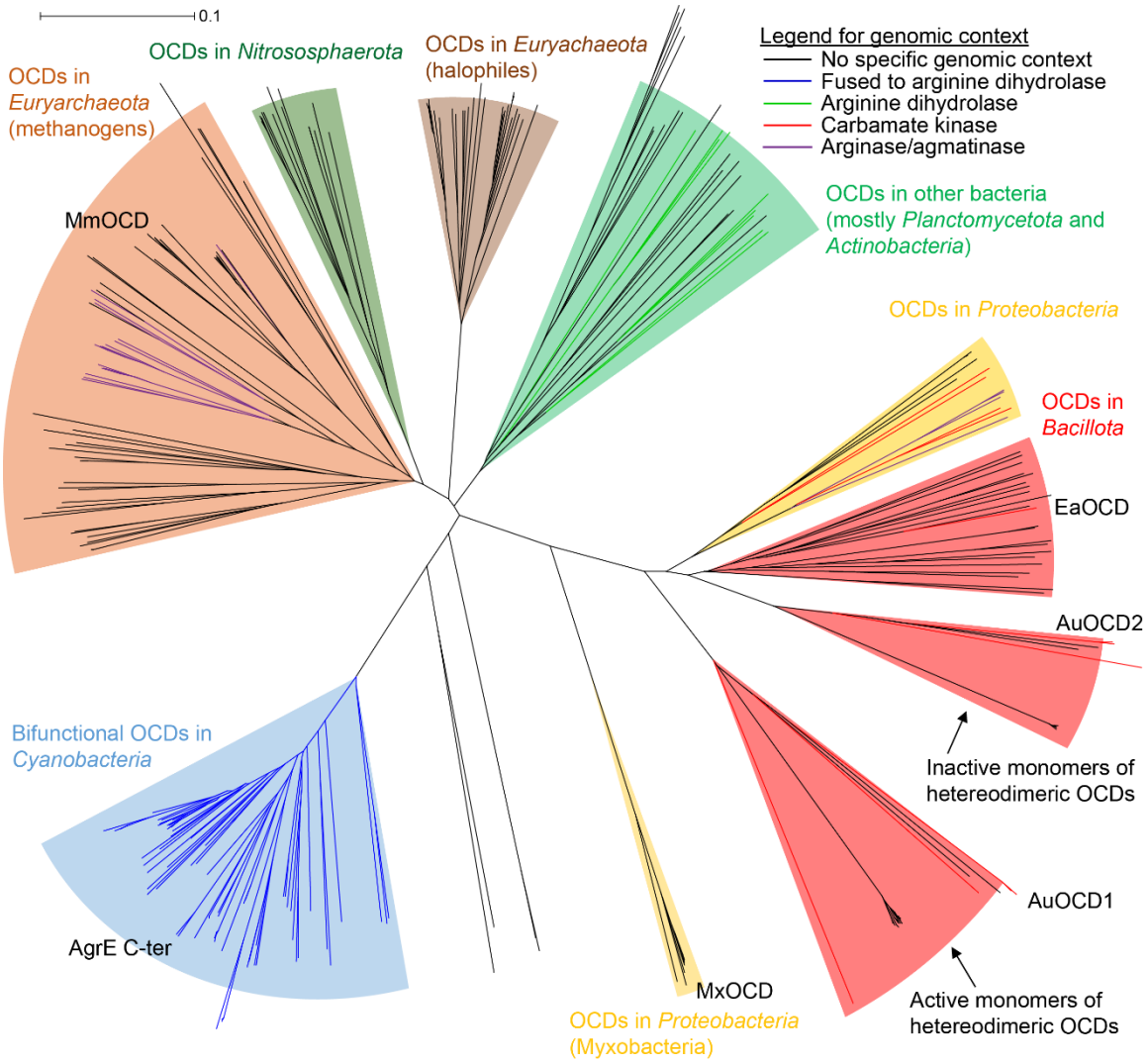


Figure 5. Phylogenetic tree and genomic context analysis of 309 OCD homologs. The different subfamilies of OCD are highlighted in different colors depending on the phyla in which they are observed. The lines are colored depending on the observed genomic context.

Table 1. Properties of purified OCDs of the COG1915 family.

Enzyme	k_{cat} (s^{-1})	95% CI (s^{-1})	K_m (mM)	95% CI (mM)	k_{cat}/K_m ($s^{-1}.mM^{-1}$)	95% CI ($s^{-1}.mM^{-1}$)	Temp (°C)	pH	Ni (%)	95% CI (%)
AuOCD	1.18	1.13-1.23	0.19	0.15-0.23	6.2	4.9-8.2	25-30	7.5-8	6	4-8
MxOCD	12.3	11.9-12.8	0.28	0.22-0.35	44	34-51	50-55	8	50	13-87
EaOCD	0.085	0.07-0.12	34	18-63	0.0025	0.001-006	40	6.5	3.4	1.2-5.7
MmOCD	3.1	2.9-3.4	4.1	3.2-5.2	0.8	0.5-1.1	ND	7.5-8	33	8-58

95% CI is the 95% confidence interval.
ND: not determined due to insufficient activity.

Table 2. Properties of purified AgrE activated by cofactor transfer.

Enzyme	k_{cat} (s^{-1})	95% CI (s^{-1})	K_m (mM)	95% CI (mM)	k_{cat}/K_m ($s^{-1}.mM^{-1}$)	95% CI ($s^{-1}.mM^{-1}$)	Temp (°C)	pH	K_i[a] (mM)	95 % CI (mM)
AgrE	32	26-39	8.9	5.5-14.2	3,6	1.8-7.1	45	8-9	320	180-600

95% CI is the 95% confidence interval.

[a] K_i was estimated based on the following equation $V_0 = \frac{V_{max}.[S]}{K_M+[S]+\frac{[S]^2}{K_i}}$

Supplemental material for:

An adenine dinucleotide variant of the nickel-pincer nucleotide cofactor in ornithine cyclodeaminases.

Benoît Desguin.

Figures S1-S11

Table S1

Legend for Dataset 1

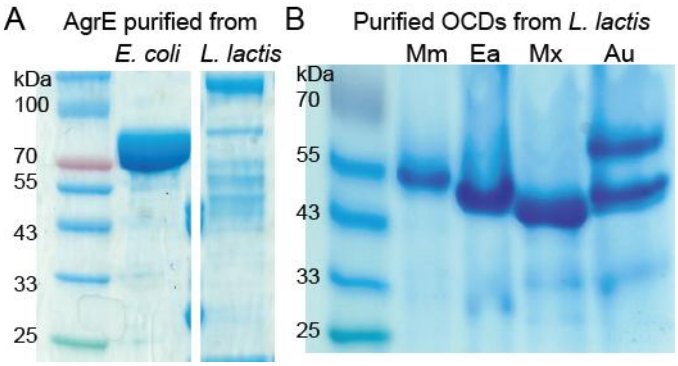


Figure S1. SDS-PAGE analysis of purified OCDs. (*A*) AgrE purified from *E. coli* harboring pGIR330 or from *L. lactis* containing pGIR220. (*B*) Purified OCDs from *L. lactis* harboring pGIR221-224.

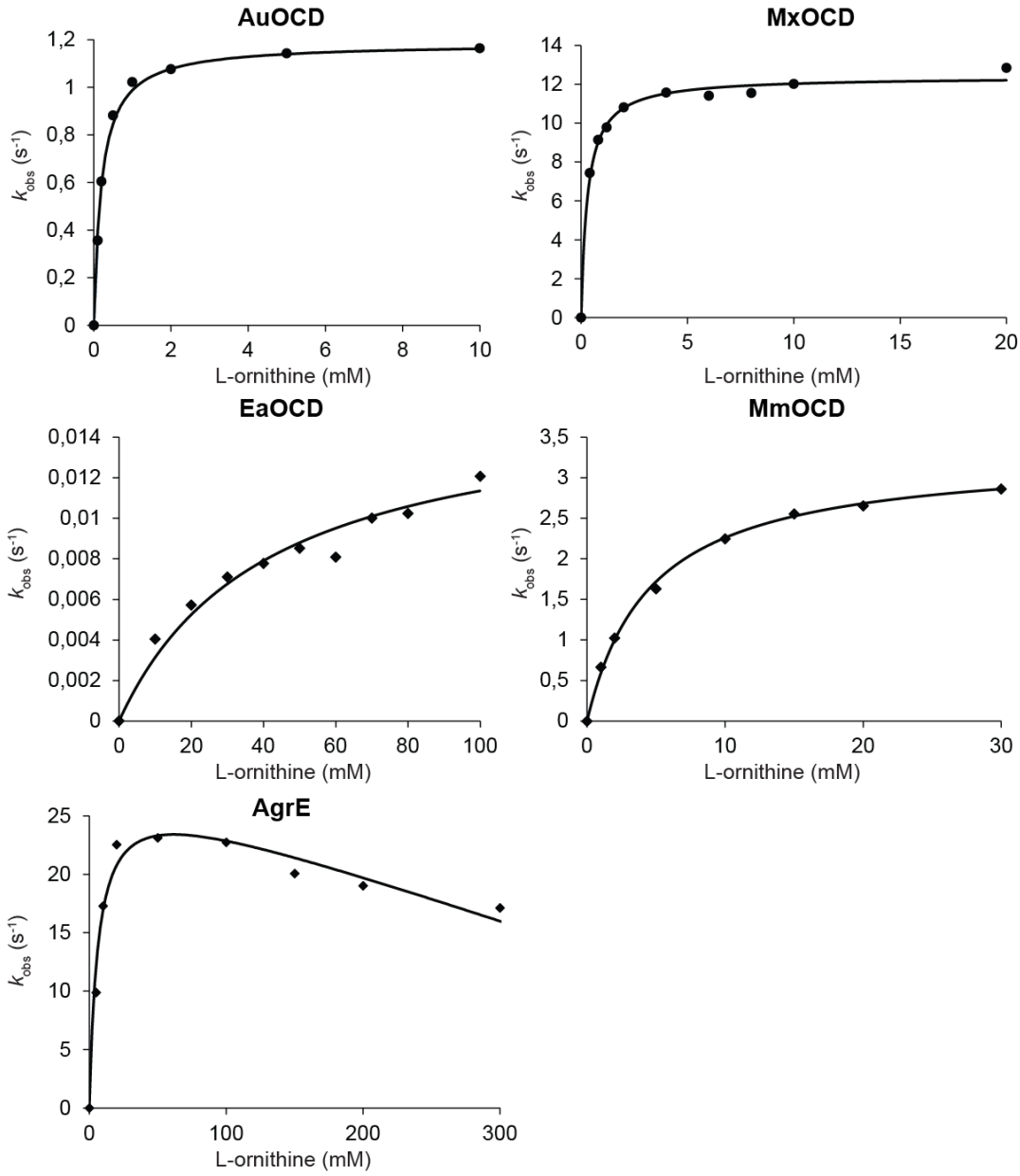


Figure S2. Kinetic parameters of purified OCDs. The points are the experimental results (means of three repetitions), the lines are the simulated values based on the calculated kinetic parameters.

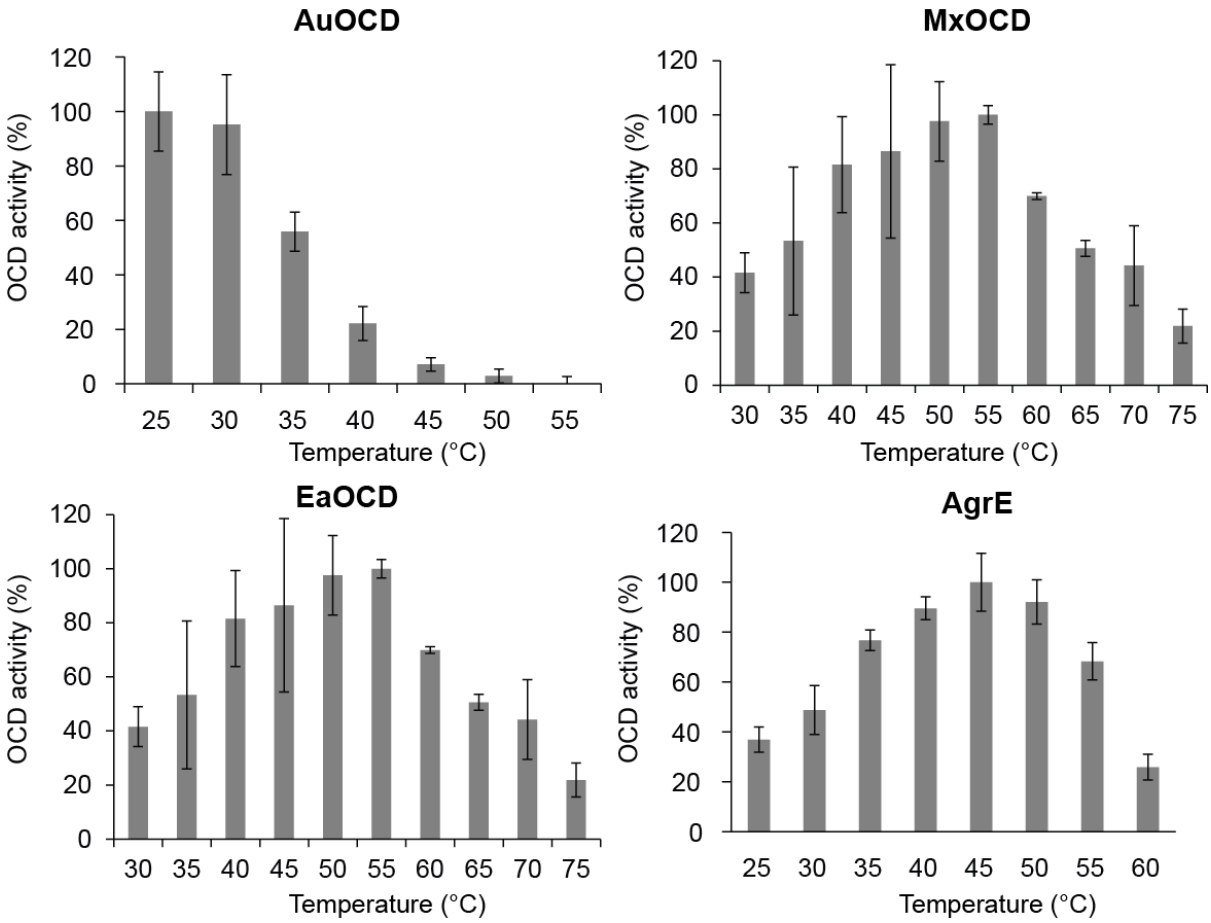


Figure S3. Optimal temperature of purified OCDs. The error bars represent the standard deviation (N=3). Maximal activity was set to 100%.

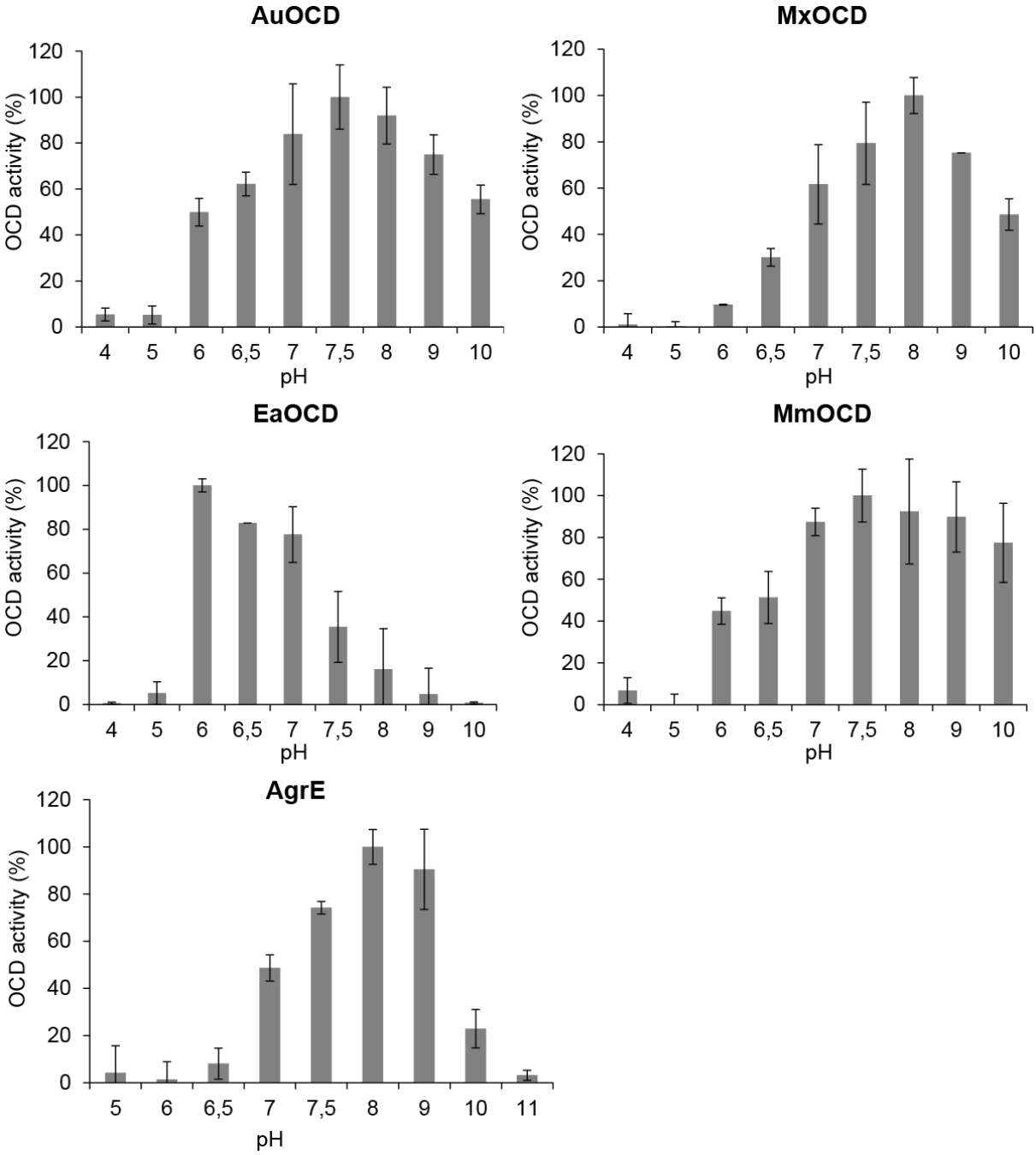


Figure S4. Optimal pH of purified OCDs. The error bars represent the standard deviation (N=3). Maximal activity was set to 100%.

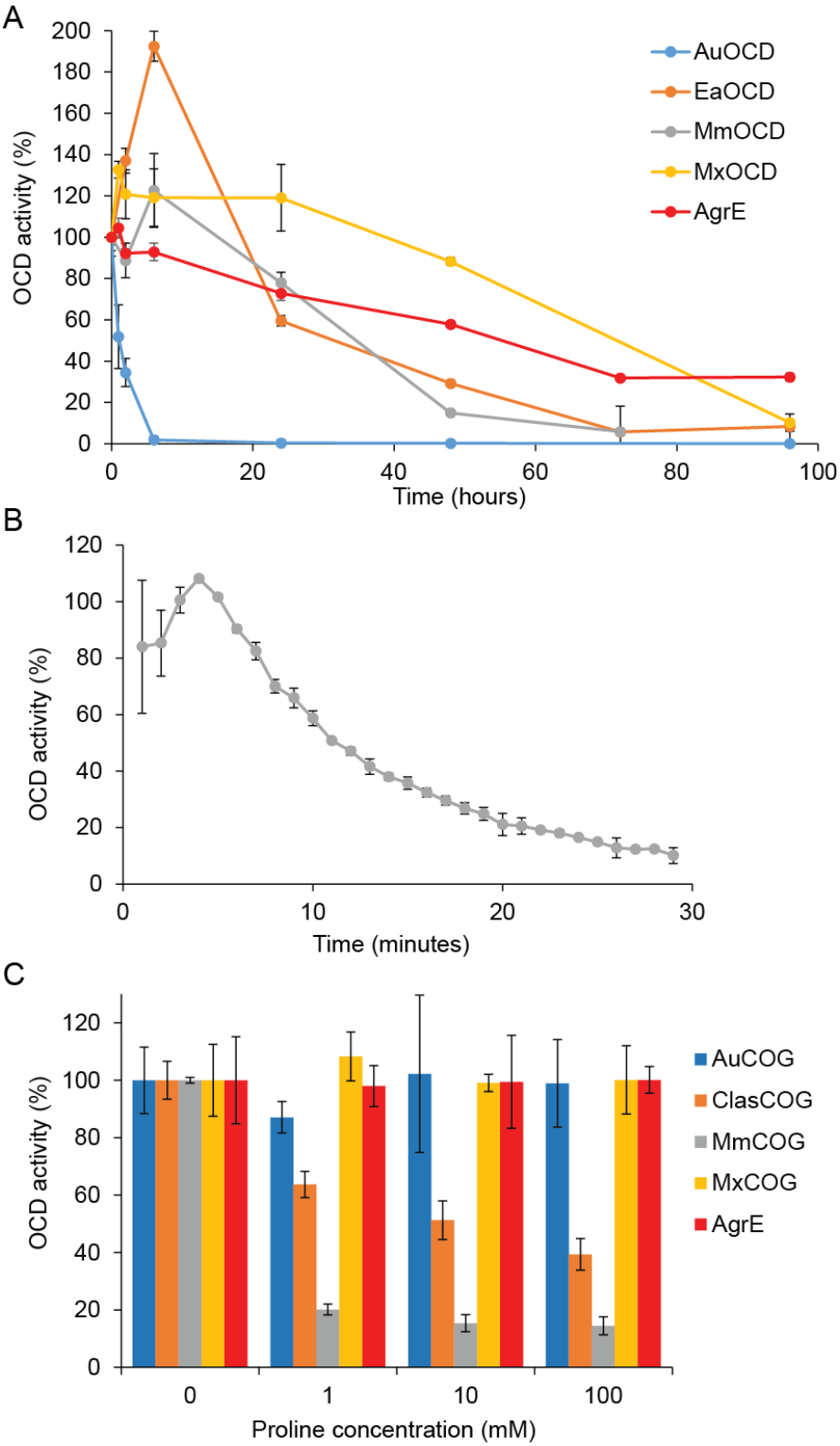


Figure S5. (*A*) Stability of purified OCDs. The activity at time 0 was set to 100% for each OCD. (*B*) Stability of purified MmOCD in the presence of 20 mM ornithine. Maximal activity was set to 100%. (*C*) L-Proline inhibition of purified OCDs. Activity without L-proline was set to 100%. The error bars represent the standard deviation (N=3) in all panels.

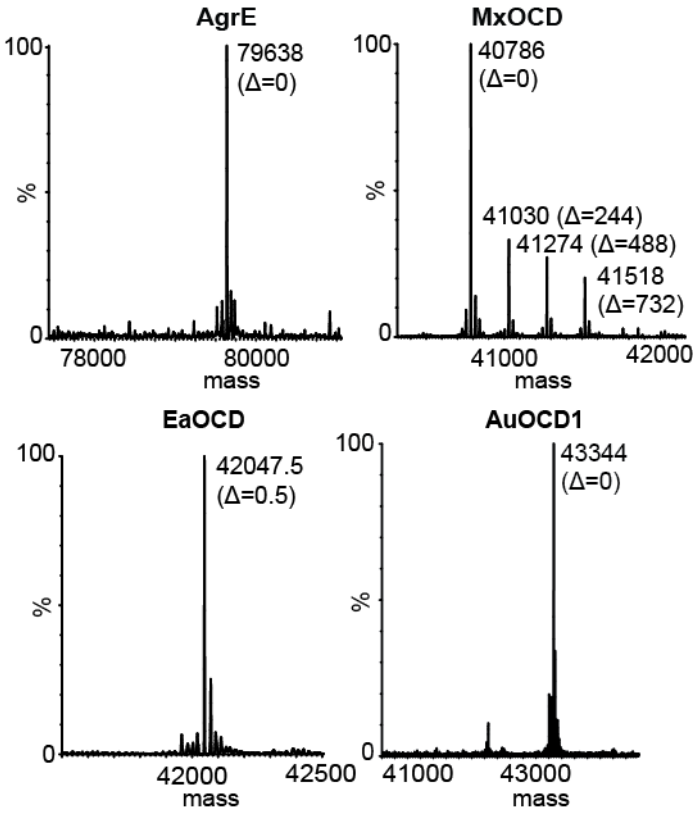


Figure S6. Masses of purified OCDs obtained by mass spectrometry analysis of intact proteins. The number in parentheses indicates the difference of mass compared to the expected mass of the apoprotein minus the N-terminal methionine.

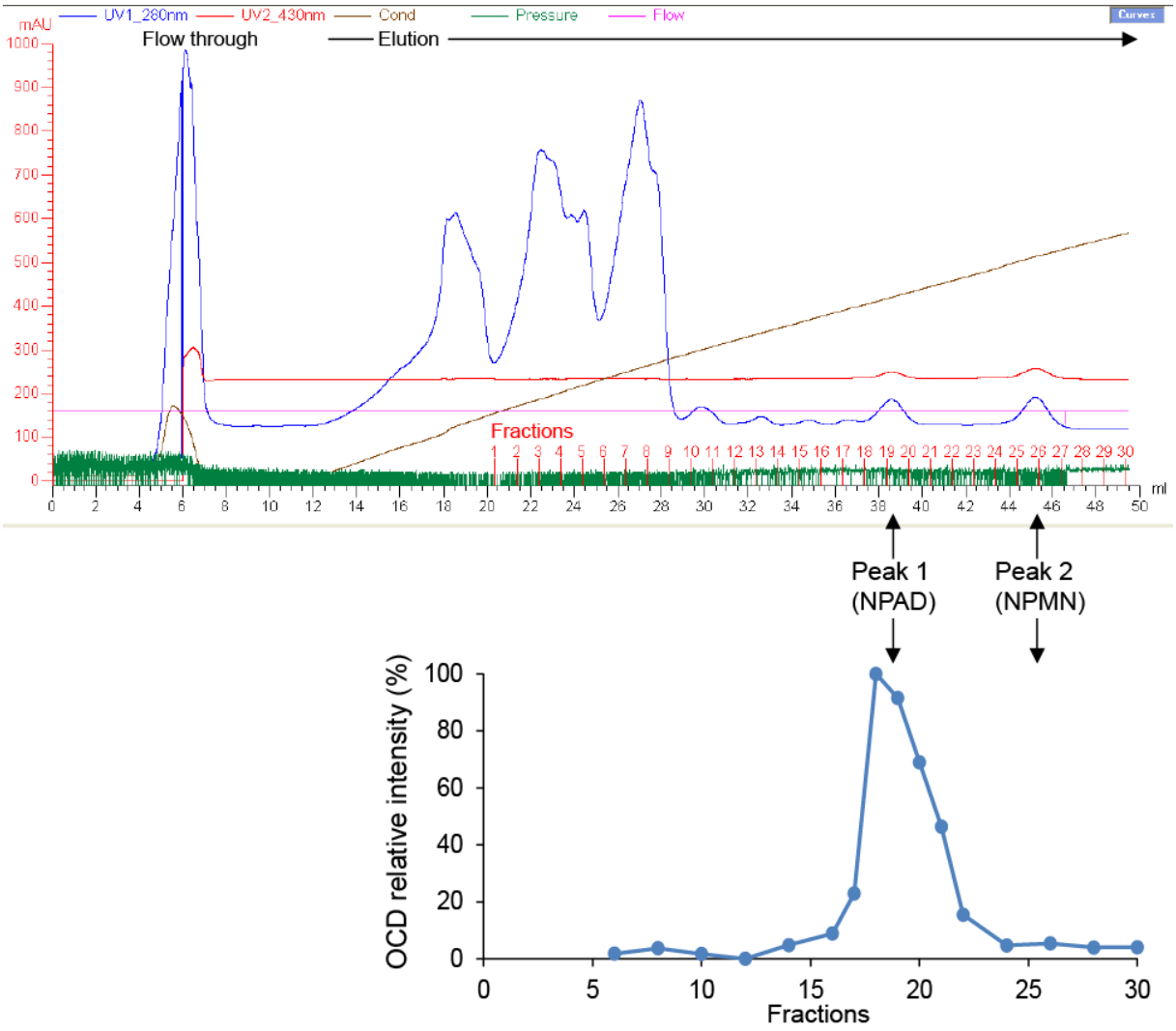


Figure S7. Purification of NPAD by anion exchange chromatography. Top: chromatogram of the purification of NPAD showing the absorbance at 280 nm (in blue), the absorbance at 430 nm (in red), the conductometry (in brown), the pressure (in green), the flow (in pink), and the fractions collected (in red). Bottom: relative OCD activity of the fractions in presence of added AgrE. The OCD activity of fraction 19 was set to 100%.

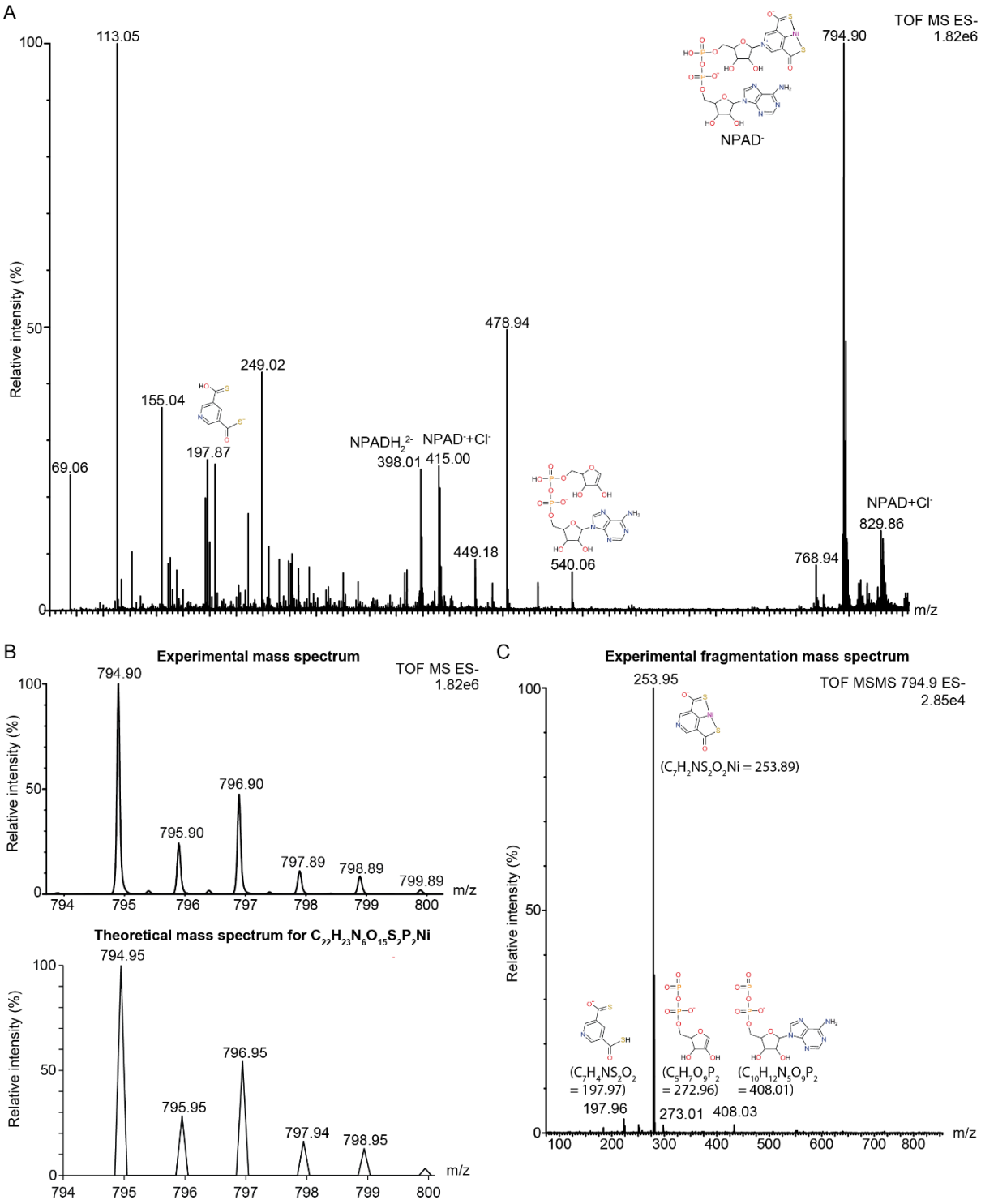


Figure S8. Identification of NPAD by mass spectrometry (MS). (*A*) Full MS spectrum (from 50 to 850 m/z) recorded between minute 33 and minute 34.5 of the liquid chromatography run. The identified peaks are labelled with the probable structure and/or name of molecule. (*B*) Expanded view of the peak corresponding to NPAD in *A* (top) and theoretical mass spectrum of NPAD (bottom). (*C*) Fragmentation mass spectrum of NPAD with the identified peaks labelled with the probable structure of the fragments and their chemical composition and monoisotopic m/z indicated in parentheses.

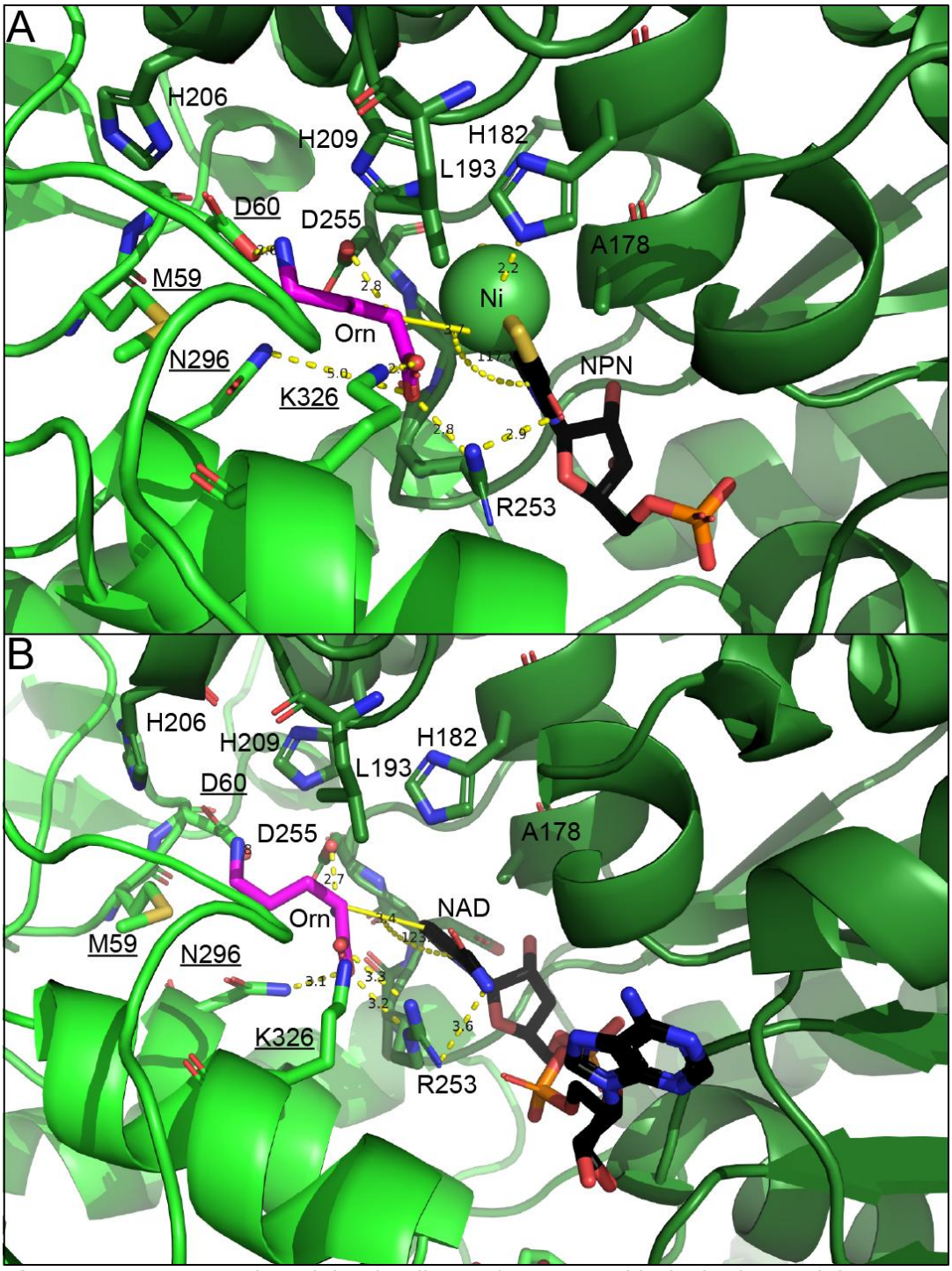


Figure S10. Structural models of a dimer of MxOCD with docked NPN (*A*) or NAD (*B*) and L-ornithine (obtained with Chai-1 webserver), respectively. Stick representation is used for the residues likely to be involved in catalysis (with O atoms in red, N in blue, and S in yellow), for L-ornithine (C atoms in magenta), and for NAD^+ or NPN (C atoms in black, P atoms in orange, and Ni as a green sphere). The labels for residues from the second dimer are underlined. Distances are indicated in Å and angles in degrees.

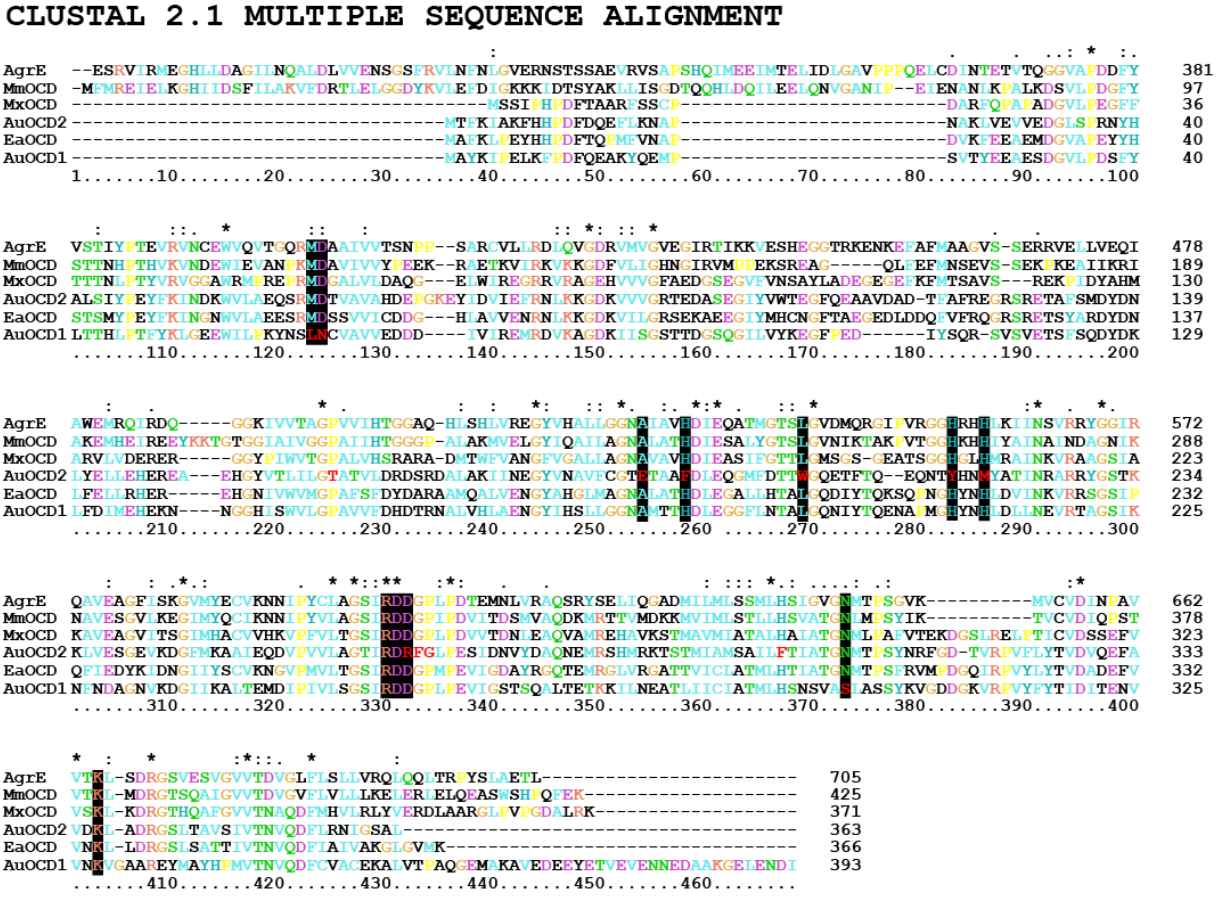


Figure S10. Sequence alignment of six OCD sequences. The standard Clustal 2.1 color code was used except that residues likely to be involved in catalysis are highlighted in black and colored in in red when they differ from the typical residues at this position. Only the C-terminal part of AgrE is shown (from residue 283 on).

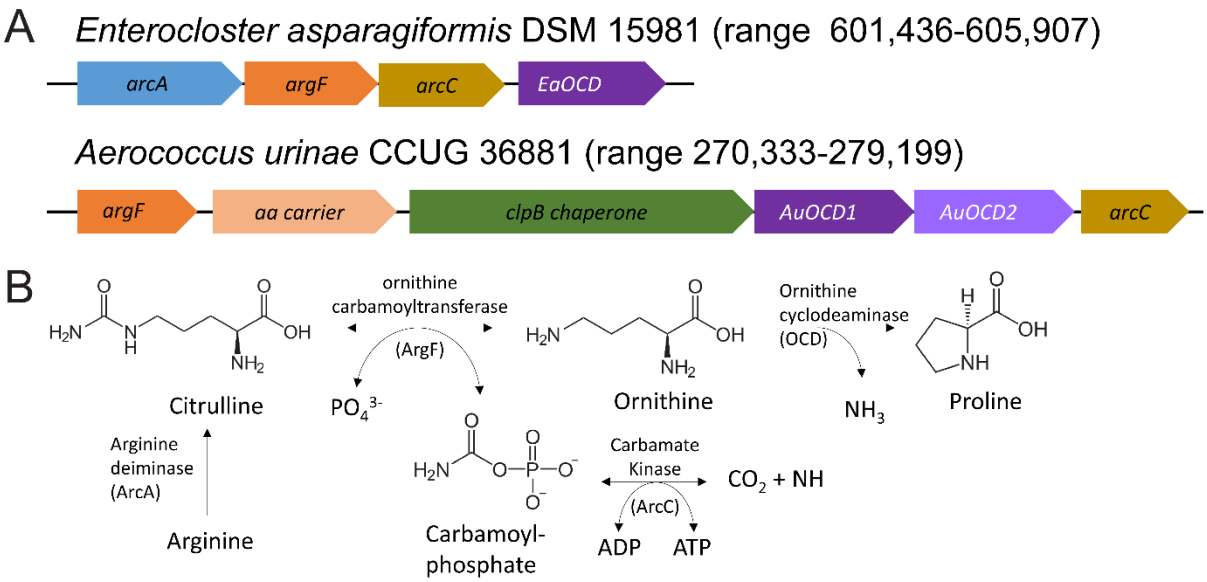


Figure S11. (*A*) Genomic context of EaOCD and AuOCD. (*B*) Arginine deiminase pathway with the addition of OCD.

Table S1. Strains, plasmids and primers.

Strains	Characteristic(s)		Source or reference
L. Lactis NZ3900	MG1363 derivative		(12)
E. coli DH10B	F- *endA*1 *recA*1 *galE*15 *galK*16 *nupG rpsL* Δ*lacX*74 Φ80*lacZ*ΔM15 *araD*139 Δ(*ara*,*leu*)7697 *mcrA* Δ(*mrr-hsdRMS-mcrBC*) λ-		Invitrogen
Plasmids	**Characteristic(s)**	**Usage**	**Source or reference**
pGIR082	Cm^r; pNZ8048 with DNA encoding *larA*$_{Tt}$ translationally fused to the StrepII-tag sequence.	LarA$_{Tt}$ purification	(9)
pGIR026	Em^r Amp^r; pGIR660 with DNA encoding *larB* translationally fused to the StrepII-tag sequence.	LarB purification	(6)
pGIR031	Cm^r; pNZ8048 with DNA encoding the StrepII-tag sequence translationally fused to *larC*.	LarC purification	(9)
pGIR076	Amp^r; pBADHisA with DNA encoding *larE* translationally fused to the StrepII-tag sequence.	LarE purification	(6)
pGIR330	Amp^r; pBADHisA with DNA encoding *agrE*.	AgrE purification	This study
pGIR331	Amp^r; pBADHisA with DNA encoding *L. lactis nadD*.	NMAT purification	This study
pGIR210	Cm^r; pNZ8048 with DNA encoding the StrepII-tag sequence in operon with *larBCDE* from *Lactiplantibacillus plantarum.*	Contruction of pGIR22X plasmids	(10)
pGIR220	Cmr; pGIR210 with DNA encoding *agrE* translationally fused to the StrepII-tag sequence.	AgrE purification from *L. lactis* cells overexpressing LarBCE.	This study
pGIR221	Cmr; pGIR210 with DNA encoding *mxCOG* translationally fused to the StrepII-tag sequence.	MxOCD purification from *L. lactis* cells overexpressing LarBCE.	This study
pGIR083	Cm^r; pNZ8048 with DNA encoding *mxCOG* translationally fused to the StrepII-tag sequence.	MxOCD purification from *L. lactis* cells without LarBCE.	This study
pGIR222	Cmr; pGIR210 with DNA encoding *auCOG1* and *auCOG2* translationally fused to the StrepII-tag sequence.	AuCOG purification from *L. lactis* cells overexpressing LarBCE.	This study

pGIR222_M4	Cmr; pGIR210 with DNA encoding *auCOG1*^{H209A} and *auCOG2* translationally fused to the StrepII-tag sequence.	AuCOG^{H209A} purification from *L. lactis* cells overexpressing LarBCE.	This study
pGIR223	Cmr; pGIR210 with DNA encoding *eaOCD* translationally fused to the StrepII-tag sequence.	EaOCD purification from *L. lactis* cells overexpressing LarBCE.	This study
pGIR224	Cmr; pGIR210 with DNA encoding *mmOCD* translationally fused to the StrepII-tag sequence.	MmOCD purification from *L. lactis* cells overexpressing LarBCE.	This study

Primers	**Sequence**	**Usage**	**Source or reference**
Sll_NheI_A	AAAGCTAGCGAAAGTCGGGTGATTCGCATGGAG	Amplication of *agrE*	This study
Sll_Pst1_B	TTTCTGCAGCTAAGCAAGTGTTTCCGCCAAGC		This study
MxCOG1915_AC	TTTACATGTCGTCCATTCCTCATCCTGACTTCAC	Amplification of *mxOCD*	This study
MxCOG1915_B2	TTTGCTAGCCTTCCTCAGCGCGTCCCC		This study
AuCOG1915A_A3	TTTACATGTCGGCTTATAAGATACCAGAATTTAAATTTCCTG	Amplification of *auOCD*	This study
AuCOG1915B_B2	AAAGCTAGCAAGGGCGCTACCAATATTCCGTAAG		This study
ClasCOG1915a_A	TTTACATGTCTGCTTTTAAATTACCTGAATACCATCATCCGG	Amplification of *eaOCD*	This study
ClasCOG1915a_B	TTTGCTAGCTTTCATCACGCCCAGCCCCTTG		This study
AuCOG_M4_A	TGGGCGCATACAACCACTTGGACCTCTTAAATG	Construction of pGIR222_M4	This study
AuCOG_M4_B	TCGGTGCATTCTCTTGAGTATAAATATTTTG		This study
AuCOG_M4_V	ATGCACCGATGGGCGC		This study
pNZ8048_UP'	ACAATGATTTCGTTCGAAGGAACTAC	Verification	This study
LarZ-B_B	GCATCGATTGCCGCCACTTGTTGTAATATTTC		(9)
NadDa_A	GGGGCTAGCGATAAAGGGAACAGAAAAAAAATAGGTTTGTTG	Amplification of *nadD*	This study
NadDa_B	CCCAAGCTTATTGGTCATTTTTTTTATATAATTTTTCAGTTTCTATATAGC		This study

Dataset 1. List of 309 OCD homologs for the construction of the phylogenetic tree (Fig. 5). “Name” refers to the name used in this study. “COGs” refers to the enzyme family or families encoded by each gene. “Context” refers to one typical genomic context identified by the GeCOnT tool (18).

See Excel file “Dataset1”.