CHEMBIOCHEM

Supporting Information

Production of Propane and Other Short-Chain Alkanes by Structure-Based Engineering of Ligand Specificity in Aldehyde-Deformylating Oxygenase

Basile Khara,[a] Navya Menon,[a] Colin Levy,[a] David Mansell,[a] Debasis Das,[b] E. Neil G. Marsh,[b] David Leys,[a] and Nigel S. Scrutton*[a]

cbic_201300307_sm_miscellaneous_information.pdf

On-line Supporting Information

1. Supplementary Methods.

Expression and purification of cADO

The V41 and A134 variant forms of cADO were isolated by site-directed mutagenesis (PfuUltra™ II Fusion HS DNA Polymerase, Stratagene). The primers (MWG Eurofins) used for the mutagenesis were as follows (base changes are indicated by underlining):
V41Y Forward primer: 5'- CATCAACGCCATTTATATCGAAGGCGAAC -3',
V41Y Reverse primer: 5'- GTTCGCCTTCGATATAAATGGCGTTGATG -3',
A134F Forward primer: 5'- ATTTGCGATTTCTTTTTATCATACGTACA -3',
A134F Reverse primer: 5'- TGTACGTATGATAAAAAGAAATCGCAAAT -3'.
PCRs were performed using the following conditions; 95 °C for 2 min; 30 cycles of 95 °C for 30 s, 70°C (V41Y) / 62 °C (A134F) for 20 s and 72 °C for 1 min 40 s; and 72 °C for 3 min. V41Y and A134F mutation were verified by DNA sequencing (MWG Eurofins).

Protein expression and purification

The plasmid cADO-pET28 [1] was transformed into *E. coli* BL21 (DE3) competent cells. The transformed cells were grown in 500 mL 2YT medium containing kanamycin (40 μg/mL) at 37 ˚C, 220 rpm. When the cells reached an optical density at 600 nm (OD_{600}) of 0.6, the expression of cADO was induced by the addition of IPTG to a concentration of 0.3 mM. The cells were further grown for 10 h before being harvested by centrifugation. The cell pellets were then resuspended in chilled binding buffer (50 mM HEPES, 0.4 M NaCl, 30 mM imidazole, pH 7.4) containing protease inhibitors (Roche) and disrupted by sonication. Cell debris was removed by centrifugation (35,000 rpm for 30 min) and the supernatant was loaded onto a 2.0 x 6.0 cm column packed with Ni^{2+} Sepharose fast-flow resin, pre-equilibrated with the same binding buffer. After this point all the purification steps were carried out at 4 ˚C in a cold cabinet. The column was then washed with 10 column volumes of binding buffer and 5 column volumes of binding buffer containing 140 mM imidazole, to remove any loosely bound contaminants. The recombinant protein was eluted using binding buffer containing 500 mM imidazole. The concentrated protein was buffer-exchanged using a PD10 column equilibrated with 50mM HEPES buffer, pH 7.4, and then loaded onto a 20 mL Q-Sepharose column equilibrated with the same buffer. The pure fractions were eluted with a gradient (0 mM to 500 mM NaCl) in elution buffer (50 mM HEPES buffer, pH 7.4). The fractions were collected after examining the purity of eluted factions by SDS-PAGE.

Whole cell assays of enzyme activity.

Cell cultures (10 mL; with/without the aprpopriate expression construct for cADO) were grown in LB media in 20 mL screw cap headspace vial (Fisher Scientific) to reach an optical density at 600 nm of 0.7. At this point, and where appropriate, expression of cADO was induced by the addition of IPTG to a concentration of 0.3 mM. All cell cultures (induced and non-induced) were then grown for a further 1 h. Butanal was then added to the cultures at a concentration of 10 mM. All cultures were then incubated for a further 24 h at 37 ˚C, 220 rpm using a shaking incubator to facilitate the conversion of butanal to propane. Stock solutions of butanal were prepared freshly before each experiment in DMSO. A 1 mL headspace sample was injected manually with a gas-tight syringe (Hamilton) into a Varian 3800 GC. All whole-cell culture assays were performed in triplicate.

1. Supplementary Figures

Determination of unknown ligand(s) using GC-MS

Purified cADO protein (500 mL, 0.8 mM) was acidified using 1 M HCl and extracted with ethyl acetate. The organic layer was collected and dried by passing through $MgSO_4$. The solvent was evaporated under vacuum under reduced pressure and the residue re-dissolved in ethyl acetate (250 μL) and methanol (50 μL). The sample was derivatised to the methyl ester using trimethylsilyl diazomethane (20 mL, 2 M solution in hexane). After standing for 30 minutes at room temperature the excess reagent was destroyed by the addition of acetic acid (5 mL) and the sample allowed to stand for 20 min prior to analysis. GC-MS analysis was performed using an Agilent 7890A GC equipped with an Agilent 7683B series autosampler injector coupled with an Agilent 5975C inert XL MSD with triple axis detector. The column employed was a DB-5ms (30 m x 250 μm x 0.25 μm): oven temperature was held at 60 ˚C for 2 min and then increased to 280 ˚C at 30˚C/min and finally maintained at 280 ˚C for 3 min. Helium was used as the carrier gas with a flow rate of 1.0 mL/min and the inlet temperature was maintained at 250 ˚C with a split ratio of 10:1. Chromatographic data were analysed using AMIDS-32 software.

A

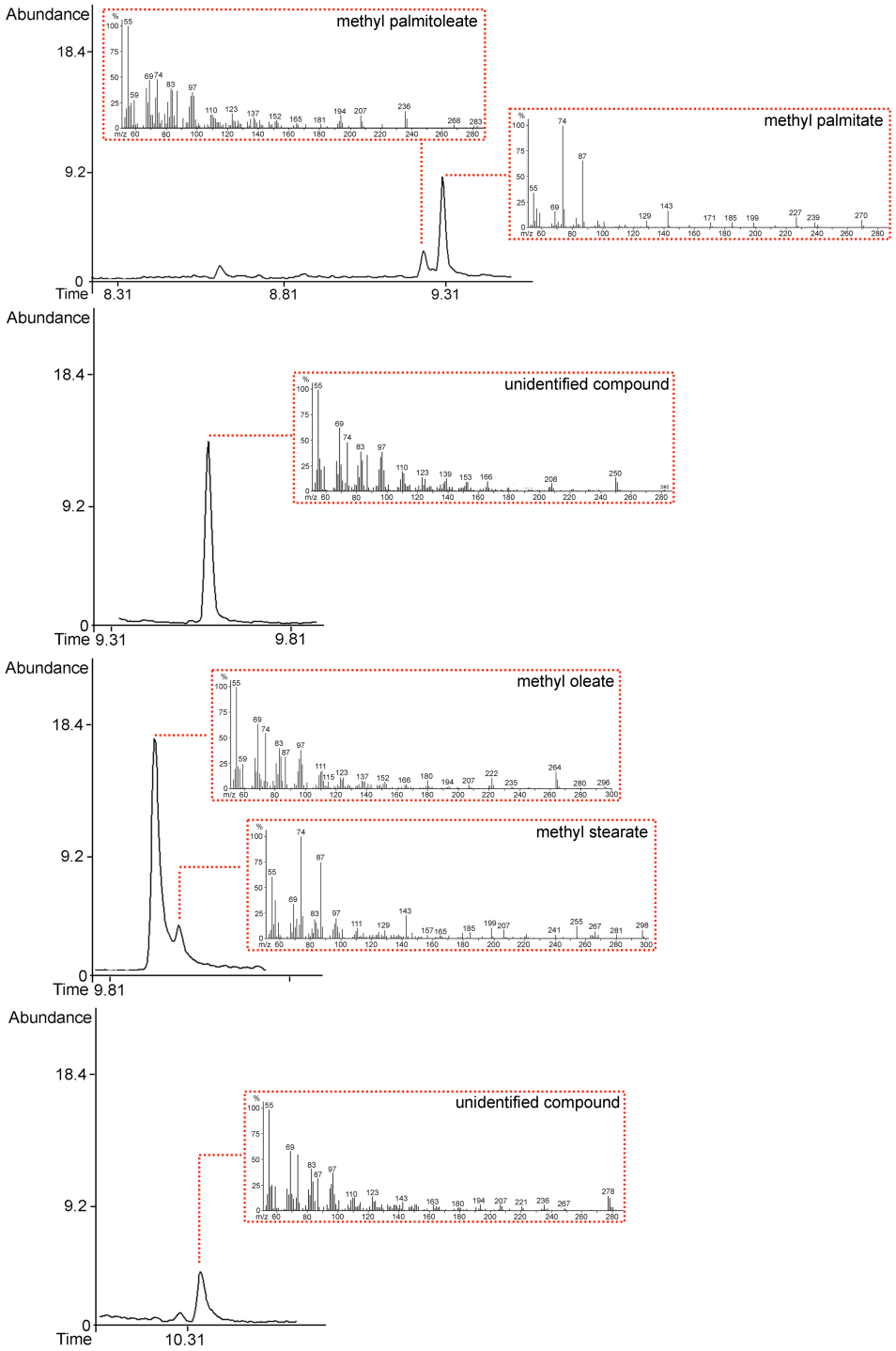

Figure S1. A) GC-MS analysis and verification of unknown ligands from wild-type cADO. A sample of purified wild-type cADO protein (0.8 mM) was acidified (using 1M HCl) and extracted with ethyl acetate. The organic was dried over $MgSO_4$ and the solvent evaporated under reduced pressure. The residue was re-dissolved in ethyl acetate containing methanol and derivatisation to the methyl ester was performed using trimethylsilyl diazomethane (2 M solution in hexane) prior to analysis by GC-MS (see methods). Ligand identification was performed by comparison of the mass spectrum of the analyte with the NIST library database. Two analytes gave insufficient match factors to achieve unambiguous identification.

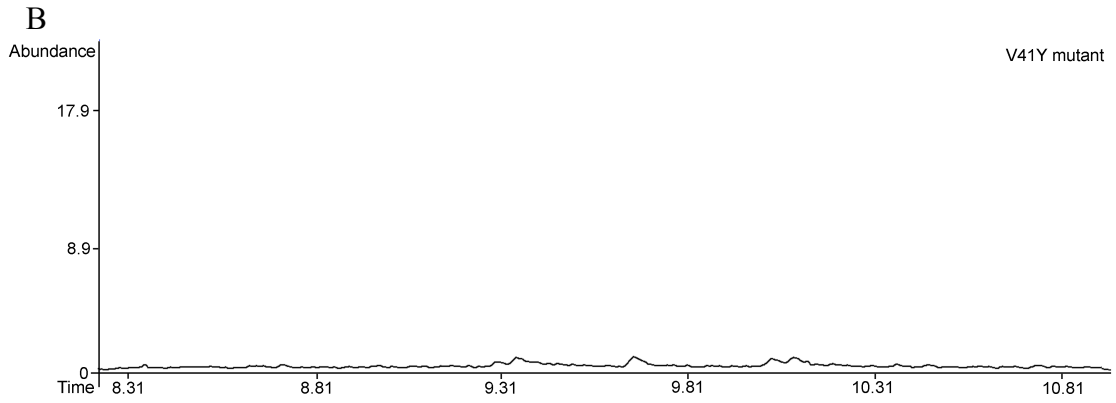

Figure S1. B) GC-MS analysis of V41Y variant. A sample of purified V41Y variant was extracted and analysed as described above. In this case, the GC trace shows absence of endogenous ligands.

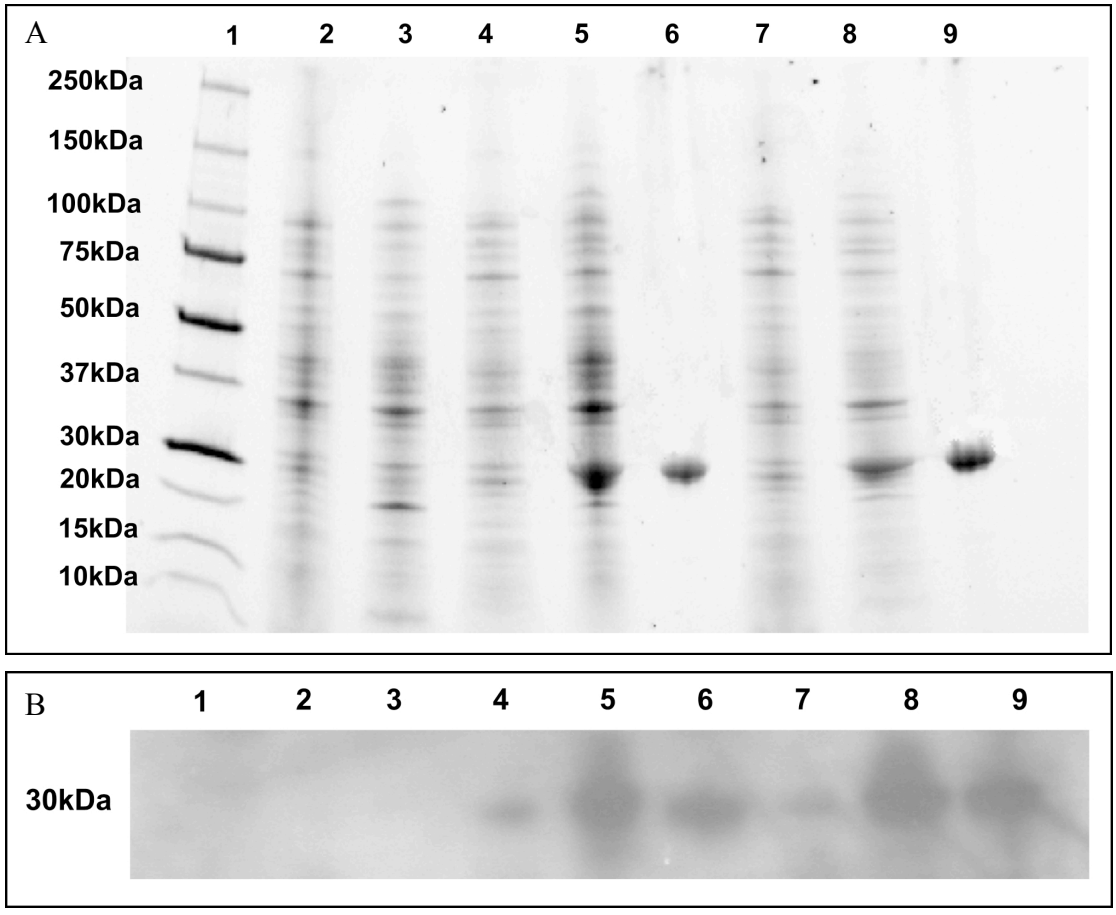

Figure S2. A) Cell lysates on reducing SDS-PAGE gels. 2: untransformed cell without IPTG; 3: untransformed cell with IPTG; 4: wild-type cADO uninduced fraction; 5: wild-type cADO induced fraction; 6: purified wild-type cADO fraction; 7: A134F cADO uninduced fraction; 5: A134F cADO induced fraction; 6: purified A134F cADO fraction. The same dilution was used for the cell lysate samples. **B)** Western blot analysis of corresponding cell lysates and purified cADO proteins. The blot shows expressed cADO protein from uninduced and induced cells after 24 hours of induction.

Crystallisation.

All crystals of cADO were obtained using the sitting drop vapour diffusion technique and 200 nL plus 200 nL drops, trays were incubated at 277 K with crystals forming over a period of ~96 h. Crystallisation details are shown in Table S1.

Table S1. Crystallisation conditions.

Sample	Concentration mg/mL	Buffer	Reservoir Solution	Cryoprotectant
Wild-type cADO	10	50mM Hepes	0.3M sodium acetate, 0.1M Tris, pH 8.5, 10% w/v PEG 8K,10% w/v PEG 1K	Paratone-N
A134F	12	50mM Hepes	0.1 M SPG buffer, pH 8, and 25 % w/v PEG 1500	Paratone-N
V41Y	12	50mM Hepes	0.1 M sodium HEPES; MOPS buffer, pH 7.5, 0.1M carboxylic acids*, ethylene glycol / PEG 8K 30%	NA

* 0.1M carboxylic acids [Na-formate; NH_4-acetate; Na_3-citrate; Na K-tartrate (racemic); Na-oxamate]

Diffraction data processing, structure determination and refinement.

X-ray diffraction data were collected at Diamond (Oxford, UK) beamlines and were integrated and scaled using the program XDS[2]. Structures were solved by molecular replacement in PHASER[3] using a template derived from structure 2OC5. The resulting models were completed through iterative rounds of rebuilding in COOT[4] and refinement in Phenix[5]. Structure validation with MOLPROBITY[6] was integrated as part of the iterative rebuild and refinement procedure.

Table S2. Data collection and refinement statistics. Figures in parentheses relate to the highest resolution shell.

Data Collection	*WT cADO (Palmitic acid)*	*A134F-hexanoic acid*	*V41Y-hexanoic acid*
Space Group	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell Dimensions			
a=b,c (Å)	77.0, 115.9	77.1, 115.9	78.2, 117.8
Resolution (Å)	1.89 (1.84-1.89)	1.67 (1.73-1.67)	1.88 (1.95-1.88)
Wavelength (Å)	0.9762	0.9795	0.9795
R_{meas} (%)	7.1 (90.4)	5.8 (87.0)	5.4 (79.0)
I/σI	22.3 (3.9)	21.8 (3.2)	21.4 (3.2)
Completeness (%)	100 (100)	99.97 (99.98)	98.4 (99.4)
Redundancy	12.8	9.8	7.7
Unique reflections	30929 (2245)	41207 (4059)	29998 (2971)
R_{work}	18.0 (21.8)	17.9 (22.2)	18.9 (26.0)
R_{free}	22.4 (33.0)	18.6 (23.4)	21.8 (28.5)
No. Atoms	2052	2085	2013
Protein	1826	1787	1775
Ligands	20	10	10
Water	195	288	228
Protein residues	223	221	221
RMS (bonds Å)	0.026	0.006	0.009
RMS (angles °)	1.99	0.98	1.14
Ramachandran favoured (%)	99.1	99.6	99.5
Ramachandran outliers (%)	0	0	0
Average B-factor ($Å^2$)	29.6	28.4	35
Protein ($Å^2$)	28.3	26.4	33.7
Solvent ($Å^2$)	40.4	40.8	44.8

Table S3. Kinetic parameter of wild-type and corresponding cADO variants.[a,b]

		WT cADO	V41Y cADO	A134F cADO	V41YA134F cADO
		k_{app} (min^{-1})	k_{app} (min^{-1})	k_{app} (min^{-1})	k_{app} (min^{-1})
Carbon length	Aldehyde				
4	Butanal	0.00065±0.0002	0.00068±0.00004	0.0029±0.0005	0.00064±0.0001
5	Pentanal	0.0039±0.001	0.0069±0.002	0.023±0.0006	0.0052±0.0001
6	Hexanal	0.064±0.002	0.05±0.002	0.215±0.0002	0.032±0.001
7	Heptanal	0.057±0.0001	0.03±0.001	0.11±0.0007	0.015±0.002
8	Octanal	0.12±0.002	0.047±0.0004	0.12±0.0007	0.02±0.0009
9	Nonanal	0.046±0.0002	0.051±0.0003	0.091±0.0009	0.024±0.002
10	Decanal	0.065±0.0004	0.071±0.001	0.042±0.0001	0.017±0.002
11	Undecanal	0.0044±0.0004	0.0093±0.00009	0.0084±0.0002	
12	Dodecanal	0.007±0.003	0.0078±0.0007	0.0097±0.0001	
13	Tridecanal	0.011±0.0006	0.0035±0.0005	0.011±0.0003	
14	Tetradecanal	0.011±0.001	0.005±0.001	0.0078±0.0007	
15	Pentadecanal	0.04±0.002	0.01±0.00005	0.011±0.001	
16	Hexadecanal	0.043±0.0007	0.015±0.0001	0.0094±0.0003	
17	Heptadecanal	0.037±0.004	0.015±0.002	0.0083±0.0006	
18	Octadecanal	0.063±0.001	0.0092±0.001	0.0072±0.001	

[a] Assays were perfomed in microaerobic condition and contained 10 μM cAD, 20 μM ferrous ammonium sulfate, 75 μM phenazine methosulfate (PMS) and 1 mM NADH ("chemical" reducing system) with varying concentrations of aldehydes in 100 mM potassium phosphate buffer, pH 7.2 containing 100 mM KCl and 10% glycerol.

[b] C4-C10 aldehydes were screened at 2 mM and due to low substrate solubility for longer chain substrates C11-C18 aldehydes were screened at 300 μM in a total volume of 0.5 mL, respectively.

2. Supplementary references

[1] D. Das, B. E. Eser, J. Han, A. Sciore, E. N. Marsh, *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 7148-7152.

[2] W. Kabsch, *Acta Crystall. Sect. D-Biol. Crystall.* **2010**, *66*, 125-132.

[3] A. J. Mccoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystall.* **2007**, *40*, 658-674.

[4] P. Emsley, K. Cowtan, *Acta Crystall. Sect. D-Biol. Crystall.* **2004**, *60*, 2126-2132.

[5] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystall. Sect. D-Biol. Crystall.* **2010**, *66*, 213-221.

[6] V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystall. Sect. D-Biol. Crystall.* **2010**, *66*, 12-21.