Structural insights into hydrolytic defluorination of difluoroacetate by microbial fluoroacetate dehalogenases

Khusnutdinova, Anna N.; Batyrova, Khorcheska; Brown, Greg; Fedorchuck, Tatiana; Chai, Yao Sheng; Skarina, Tatiana; Flick, Robert; Petit, Alain-Pierre; Savchenko, Alexei; Stogios, Peter; Yakunin, Alexander

Febs Journal

DOI:
10.1111/febs.16903

Published: 01/10/2023

Peer reviewed version

Supporting information for the manuscript

"Structural insights into hydrolytic defluorination of difluoroacetate by microbial fluoroacetate dehalogenases"

Anna N. Khusnutdinova, Khorcheska A. Batyrova, Greg Brown, Tatiana Fedorchuk, Yao Sheng Chai, Tatiana Skarina, Robert Flick, Alain-Pierre Petit, Peter Stogios, Alexei Savchenko, and Alexander F. Yakunin

Supplementary Figures

Supplementary Tables

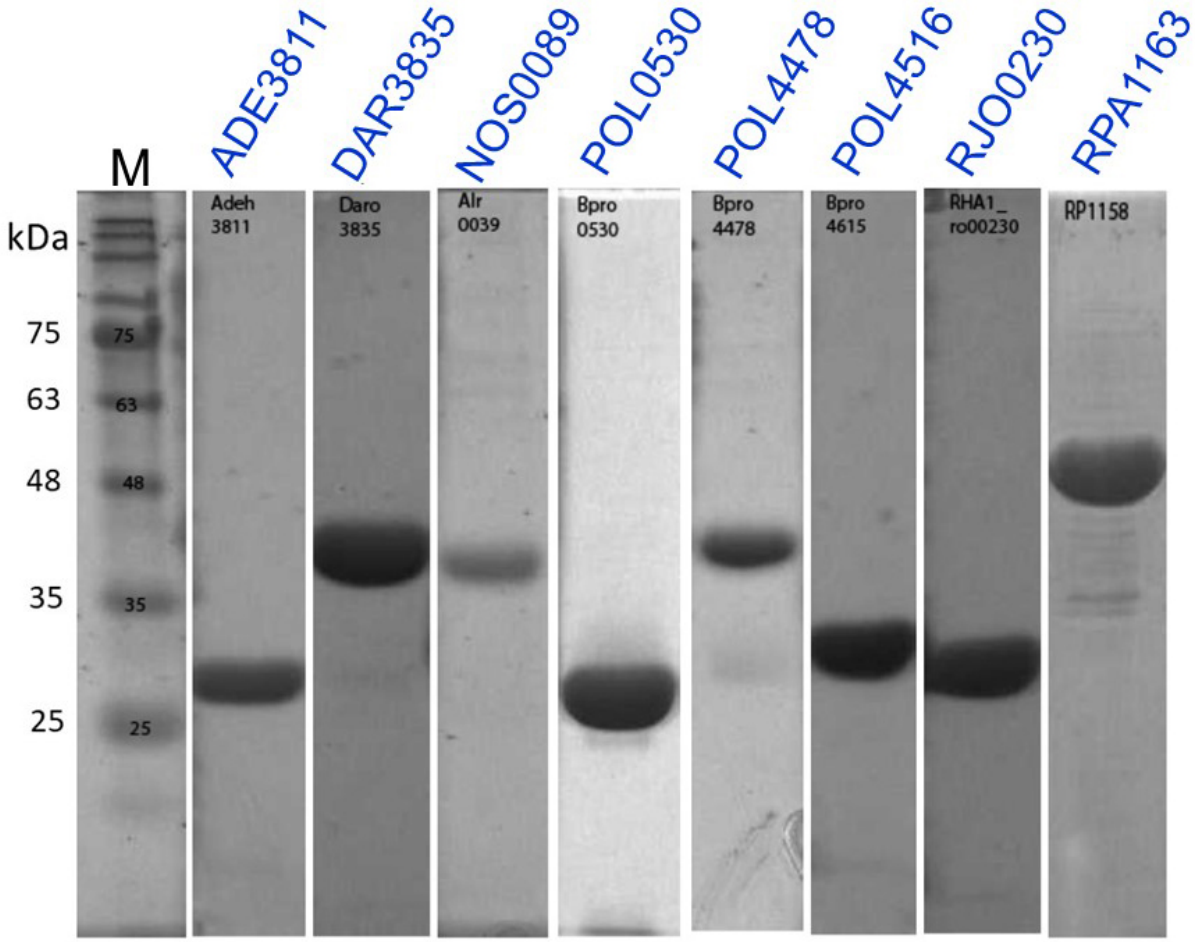

Fig. S1. SDS-PAGE analysis of purified FA dehalogenases used in this study. Coomassie-stained SDS gels showing purified proteins and molecular weight markers (M, kDa).

Fig. S2. Multiple sequence alignment of hydrolytic dehalogenases used in this study. (A), α/β-hydrolases: DAR3835 from *Dechloromonas aromatica* (Uniprot ID Q479B8), ADE3811 from *Anaeromyxobacter dehalogenans* (Q2IG66), NOS0089 from *Nostoc* sp. (Q8Z0Q1), POL0530 from *Polaromonas* sp. (Q12G50). (B), HADs: RPA1163 from *Rhodopseudomonas palustris* (Q6NAM1), POL4478 from *Polaromonas* sp. (Q123C8), POL4516 from *Polaromonas* sp. (Q122Z0), RJO0230 from *Rhodococcus jostii* (Q0SK70). The catalytic residues are designated by red boxes (Asp107, Asp131, His274 in DAR3835), fluoride coordinating residues are labelled by blue rectangles (His152, Trp153, Tyr215 in DAR3835), and carboxylate binding residues (Arg108 and Arg111 in DAR3835) are indicated by green rectangles, and conserved residues are shown in bold. The secondary structural elements of DAR3835 and RJO0230 (PDB code 3UMG).

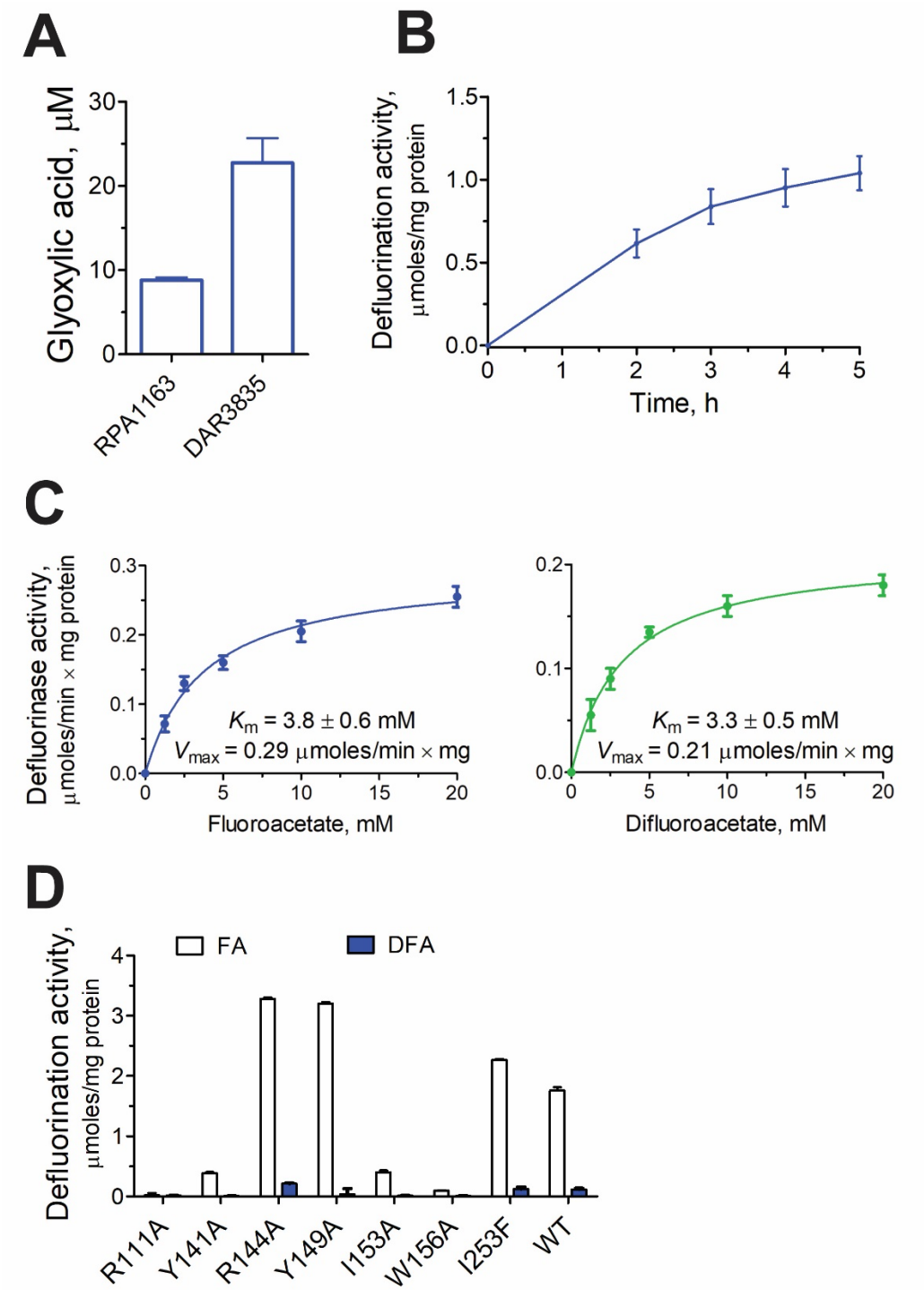

Fig. S3. Defluorination activity of purified DAR3835 and RPA1163. (A), Formation of glyoxylic acid during DFA defluorination by DAR3835 and RPA1163. The reaction mixtures (200 µL) contained 10 mM DFA and 20 µg of protein (overnight incubation at 30 ºC), and glyoxylic acid was measured using LC-MS under negative ionization (see Materials and Methods for experimental details). (B, C), Defluorinase activity of purified DAR3835 with FA and DFA as a function of time and substrate concentration (20 µg of enzyme/assay). (D), Site-directed mutagenesis of RPA1163: defluorination activity with FA or DFA as substrates. The reaction mixtures contained 10 mM substrate, 2 µg/ml of phenol red, and 20 µg of purified enzyme. All assays were carried out in triplicate, and results are means ± SD from at least two independent determinations.

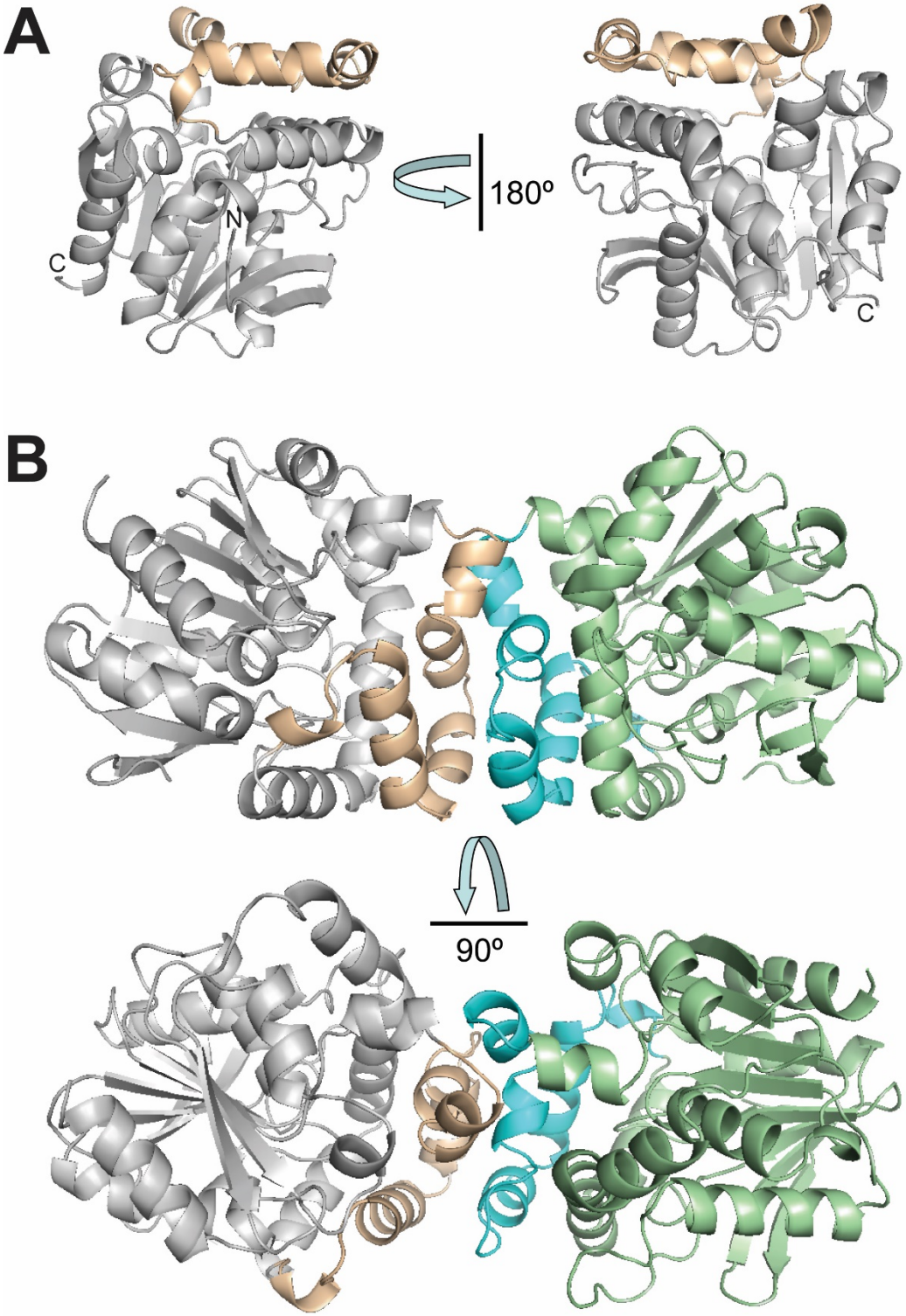

Fig. S4. Crystal structure of NOS0089: overall fold. (A), Overall fold of the protomer: two views related by a 180º rotation. The protein core domain is coloured in grey, whereas the lid domain is represented in light orange. The protein N- and C-termini are labelled (N and C). (B), Two views of the NOS0089 dimer shown in two views related by a 90º rotation. The protein subunits shown as ribbon diagrams, whereas the core and lid domains of two protomers are coloured in different colours.

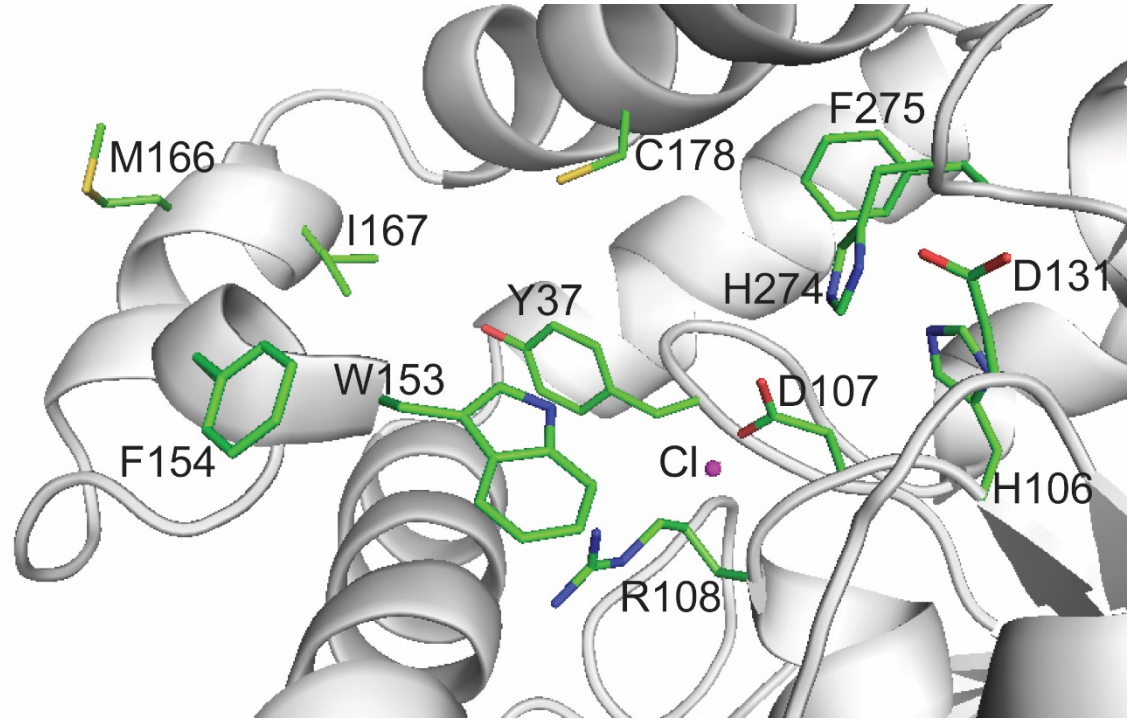

Fig. S5. DAR3835 active site: amino acid residues selected for site-directed mutagenesis. The protein ribbon is coloured in grey with residue side chains shown as sticks and carbon atoms coloured green, whereas the bound Cl^- ion is indicated as a magenta-coloured sphere.

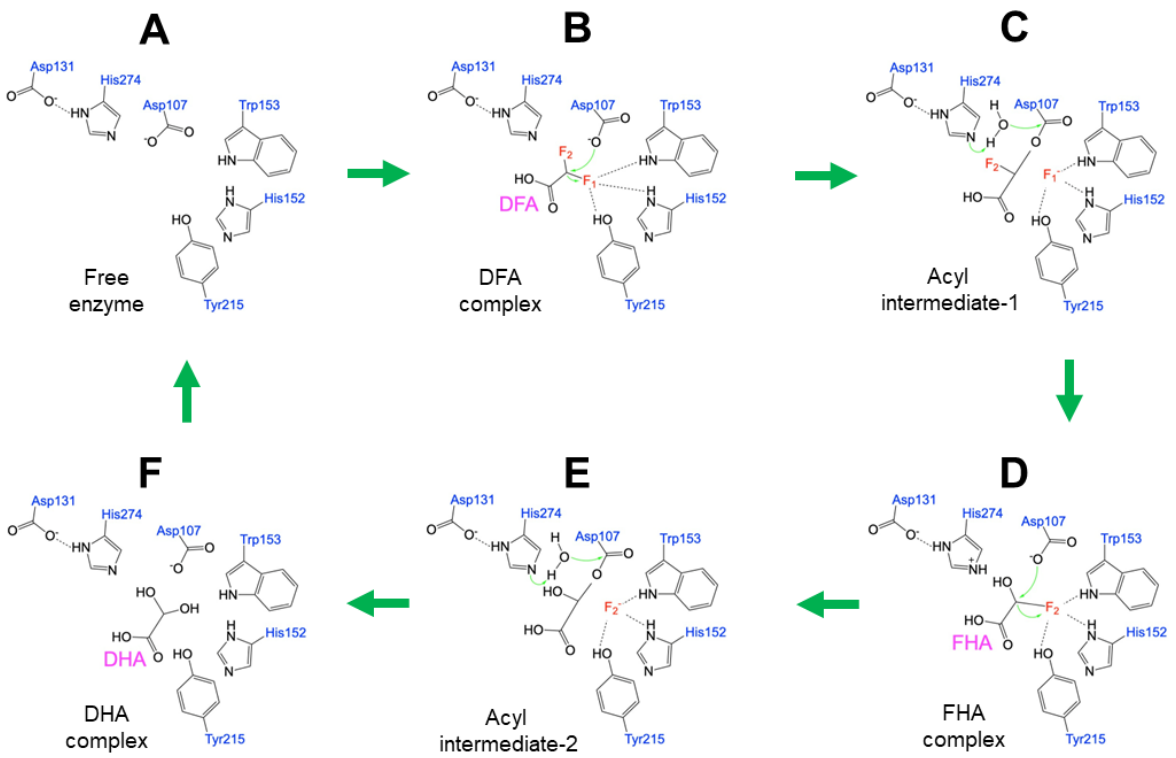

Fig. S6. Potential catalytic mechanism of DFA defluorination by DAR3835 without tetrahedral intermediates. The reaction involves the DAR3835 catalytic triad (Asp107, His274, and Asp131) and includes the following steps: (A), free enzyme; (B), the DAR3835-DFA (Michaelis) complex; (C), 2-fluoroglycolyl intermediate (acyl intermediate-1); (D), the DAR3835 complex with 2-fluoro-2-hydroxy-acetate (FHA); (E), 2-hydroxyglycolyl intermediate (acyl intermediate-2); (F), the DAR3835 complex with dihydroxyacetate. In organic solvents, DHA is dehydrated to glyoxylic acid, which was detected using LC-MS.

Table S1. Bacterial hydrolytic dehalogenases purified and screened in this study.

Protein name	Protein length	Gene name	Protein superfamily	Uniprot ID	Microorganism
1. ADE3811	226 aa	Adeh3811	HAD[a]	Q2IG66	*Anaeromyxobacter dehalogenans*
2. DAR3835	303 aa	Daro3835	α/β hydrolase	Q479B8	*Dechloromonas aromatica*
3. NOS0089	291 aa	Alr0039	α/β hydrolase	Q8Z0Q1	*Nostoc sp.*
4. POL0530	229 aa	Bpro0530	HAD	Q12G50	*Polaromonas sp.*
5. POL4478	317 aa	Bpro4478	α/β hydrolase	Q123C8	*Polaromonas sp.*
6. POL4516	242 aa	Bpro4516	HAD	Q122Z0	*Polaromonas sp.*
7. RJO0230	254 aa	RHA0230	HAD	Q0SK70	*Rhodococcus jostii*
8. RPA1163	302 aa	RPA1163	α/β hydrolase	Q6NAM1	*Rhodopseudomonas palustris*

[a] HAD, haloacid dehalogenase

Table S2. Amino acid sequence identity of defluorinases used in this study.

Enzyme	DAR3835	NOS0089	POL4478	RPA1163	POL0530	POL4516	ADE3811	RJO0230
DAR3835[a]		53.6	41.8	41.5	12.2	11.4	13.5	8.0
NOS0089[a]	53.6		46.1	45.3	7.8	8.6	10.8	6.4
POL4478[a]	41.8	46.1		38.9	8.2	9.4	9.9	6.5
RPA1163[a]	41.5	45.3	38.9		7.0	9.7	12.3	7.9
POL0530[b]	12.2	7.8	8.2	7.0		39.3	37.1	17.0
POL4516[b]	11.4	8.6	9.4	9.7	39.3		39.1	23.4
ADE3811[b]	13.5	10.8	9.9	12.3	37.1	39.1		20.6
RJO0230[b]	8.0	6.4	6.5	7.9	17.0	23.4	20.6	

[a] α/β hydrolase superfamily

[b] HAD-like superfamily

Table S3. X-ray crystallographic statistics for the structures of DAR3835 and NOS0089.

Structure	**DAR3835**	**DAR3835 H274N** (D107-glycolyl intermediate)	**NOS0089** (ALR0039)
PDB code	8SDC	8SDD	3QYJ
Data collection			
Space group	P1	$P2_1$	$P2_1$
Unit cell			
a, *b, c* (Å)	45.2, 45.2, 75.2	42.94, 58.49, 138.19	44.54, 88.20, 83.59
α, β, γ, (°)	87.6, 75.6, 63.0	90, 93.0, 90	90, 93.1, 90
Resolution, Å	19.66 – 1.86	46.00 – 2.00	20.00 – 1.78
R_{merge}[a]	0.049 (0.111)[c]	0.049 (0.098)	0.037 (0.258)
R_{pim}[b]	0.030 (0.061)	0.025 (0.075)	NC[d]
$CC_{1/2}$	0.997 (0.987)	0.998 (0.986)	NC
$I / \sigma(I)$	17.7 (9.4)	23.3 (11.3)	42.8 (6.3)
Completeness, %	94.3 (90.6)	94.6 (85.5)	97.4 (94.5)
Redundancy	3.6 (3.3)	4.8 (4.2)	3.8 (3.6)
Refinement			
Resolution, Å	19.66 – 1.86	30.40 – 2.00	19.86 – 1.78
No. unique reflections: working, test	40567, 2127	43847, 1995	58321, 2021
R-factor/free *R*-factor[e]	12.5/17.4 (13.5/19.0)	21.9/26.2 (24.4/30.8)	14.9/17.9 (23.4/NC)
No. refined atoms, chains			
Protein	4673,2	4652, 2	4732, 2
Solvent	2	N/A	N/A
Water	1018	667	634
B-factors			
Protein	15.2	32.1	20.7
Solvent	20.4	N/A	N/A
Water	33.0	38.7	30.4
r.m.s.d.			
Bond lengths, Å	0.008	0.005	0.025
Bond angles, °	0.897	0.736	1.84
Ramachandran			
favoured	97.1%	97.1%	99.8%
allowed	2.9%	2.8%	0.2%
outliers	0%	0.1%	0%

[a] $R_{merge} = \Sigma_{hkl}\Sigma_j|I_{hkl,j} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_j I_{hk,j}$, where $I_{hkl,j}$ and $\langle I_{hkl}\rangle$ are the *j*th and mean measurement of the intensity of reflection *j*.

[b] $R_{pim} = \Sigma_{hkl}\sqrt{(n/n-1)}\ \Sigma^n_{j=1}|I_{hkl,j} - \langle I_{hkl}\rangle|/\Sigma_{hkl}\Sigma_j I_{hk,j}$

[c] all values in brackets and $CC_{1/2}$ refer to the highest resolution shell.

[d] NC = not calculated.

[e] $R = \Sigma|F_p^{obs} - F_p^{calc}|/\Sigma F_p^{obs}$, where F_p^{obs} and F_p^{calc} are the observed and calculated structure factor amplitudes, respectively.

Table S4. Structure-based computational analysis of protein tunnels in DAR3835, NOS0089, and RPA1163 using MOLE 2.5[a].

Protein (subunit/tunnel)	Lining amino acids	Length, Å	Charge[b]	Hydrophobicity[c]	Polarity[d]
DAR3835 (A/1)	D107, R108, R111, S149, Y150, H152, W153, R181, W182, M184, F247, R250, H274	21.2 ± 0.3	2	0.43 ± 0.02	17.6 ± 2.4
DAR3835 (B/1)	D107, R111, D131, I132, S149, Y150, H152, W153, R181, W182, M184, F247, R250, H274	21.4 ± 1.2	2	0.34 ± 0.04	18.5 ± 1.4
NOS0089 (A/1)	D104, R108, D128, I129, Y147, K178, W179, K181, G242, I243, I244, H270	19.9 ± 0.7	0.5	0.30 ± 0.06	22.6 ± 3.4
NOS0089 (B/1)	D104, R105, R108, D128, A146, Y147, W150, K178, W179, G242, I243, H270	17.1 ± 1.0	0	0.20 ± 0.08	21.3 ± 0.3
RPA1163 (A/1)	D110, R114, D134, I135, Y141, Y149, K152, I153, H155, W156, I253, H280	18.7 ± 0.2	1	0.53 ± 0.19	15.2 ± 2.6
RPA1163 (B/1)	D110, R114, I135, Y141, K152, I153, H155, W156, W185, H280	15.1 ± 0.5	1	0.58 ± 0.13	28.9 ± 2.1
RPA1163 (B/2)	D110, R114, D134, I135, L136, Y141, R144, Y149, W156, G252, P260, H280	16.7 ± 1.1	0	0.09 ± 0.01	21.9 ± 1.7

[a] Charge, hydrophobicity, and polarity of active site residues were calculated as indicated in Materials and Methods.

[b] Charge: calculated as a sum of charges of amino acid residues along the tunnel.

[c] Hydrophobicity: average hydrophobicity of side chains lining the tunnel (the most hydrophobic amino acid Ile = 1.81, the most hydrophilic amino acid Glu = -1.14).

[d] Polarity: average polarity of amino acids lining the tunnel ranging from nonpolar residues (Ala, Gly = 0) through polar (Ser = 1.67) to charged amino acids (Glu = 49.9, Arg = 52.0) (according to Zimmerman et al., 1968) (83).

Table S5. Structure-based computational analysis of the active site cavities of DAR3835, NOS0089, and RPA1163 using MOLE 2.5.

Protein (chain)	**Amino acid residues lining the cavity**	**Volume**, $Å^3$
DAR3835 (A)	D107, R108, R111, D131, I132, A133, T138, L146, Y150, H152, T153, W182, Y215, G243, G246, F247, V248, G249, Y252, G273, H274	733.6
DAR3835 (B)	D107, R108, R111, D131, I132, A133, T138, L146, Y150, H152, W153, W182, Y215, G243, G246, F247, V248, G249, Y252, G273, H274	724.9
NOS0089 (A)	Y34, D104, R105, R108, D128, I129, A130, M135, Y147, W150, L153, Y171, C175, W179, Y211, G239, G242, I244, Y248, V250, T253, G269, H270	1043.3
NOS0089 (B)	Y34, D104, R105, R108, L126, D128, I129, A130, M135, F143, Y147, H149, W150, W179, Y211, V236, L237, W238, I244, Y248, D249, V250, T253, W254, H270	909.3
RPA1163 (A)	D110, R111, R114, D134, I135, L136, P137, T138, Y141, R144, M145, A150, I153, H155, W156, W185, Y219, A254, V263, H280	668.3
RPA1163 (B)	F40, D110, R111, R114, V132, D134, L135, L136, Y141, I153, Y154, H155, W156, Y219, A246, L247, L261, W264, H280	1062.3

Table S6. Binding energies for ligand binding in the active sites of DAR3835 and RPA1163 calculated by docking simulations [a].

RPA1163			
FA	**Free energy of binding (kcal/mol)**	**DFA**	**Free energy of binding (kcal/mol)**
Model 1	-3.28	Model 1	-2.93
Model 2	-3.28	Model 2	-2.82
Model 3	-3.27	Model 3	-2.81
Model 4	-3.27	Model 4	-2.80
Model 5	-3.27	Model 5	-2.80
Model 6	-3.27	Model 6	-2.78
Model 7	-3.27	Model 7	-2.78
Model 8	-3.27	Model 8	-2.78
Model 9	-3.27	Model 9	-2.77
Model 10	-3.25	Model 10	-2.77
DAR3835			
FA	**Free energy of binding (kcal/mol)**	**DFA**	**Free energy of binding (kcal/mol)**
Model 1	-1.73	Model 1	-1.95
Model 2	-1.73	Model 2	-1.94
Model 3	-1.72	Model 3	-1.94
Model 4	-1.72	Model 4	-1.94
Model 5	-1.70	Model 5	-1.93
Model 6	-1.70	Model 6	-1.91
Model 7	-1.70	Model 7	-1.90
Model 8	-1.66	Model 8	-1.89
Model 9	-1.64	Model 9	-1.88
Model 10	-1.63	Model 10	-1.88

[a] FA- fluoroacetate, DFA, difluoroacetate. The active site models with minimal binding energies are shown in Fig. 8.