SUPPLEMENTARY MATERIALS

Protein purification and crystallization artifacts: the tale usually not told

Ewa Niedzialkowska*, Olga Gasiorowska*, Katarzyna B. Handing, Karolina A. Majorek, Przemyslaw J. Porebski, Ivan G. Shabalin, Ewelina Zasadzinska, Marcin Cymborowski, Wladek Minor‡

* contributed equally, ‡ corresponding author

Inventory of supplementary materials

Table S1.
Summary of data collection and refinement statistics for the deposited YodA, YadF, and trimethylamine methyltransferase structures

Table S2.

The list of deposits used to identify crystallization artifacts by MR which is carried out using predefined set of structures of possible artifacts as templates together with the resulting *R*-factor values.

Table S1. Summary of data collection and refinement statistics for the deposited YodA, YadF, and trimethylamine methyltransferase structures.

Protein	**YodA**	**YadF**	**Trimethylamine methyltransferase**	
PDB accession code	4TNN	4ZNZ	Anomalous dataset	4YYC
Data processing				
Wavelength	0.9786	1.27822	0.9789	1.0781
Space group	$P3_121$	$P4_22_12$	$P2_12_12$	$P2_12_12$
Unit cell	Length [Å] a = 76.58 b = 76.58 c = 61.94 Angles [°] α = 90.00 β = 90.00 γ = 120.00	Length [Å] a = 67.91 b = 67.91 c = 84.93 Angles [°] α = 90.00 β = 90.00 γ = 90.00	Length [Å] a = 89.16 b = 60.03 c = 88.35 Angles [°] α = 90.00 β = 90.00 γ = 90.00	Length [Å] a = 89.16 b = 60.03 c = 88.35 Angles [°] α = 90.00 β = 90.00 γ = 90.00
Resolution (Å)	50.00 – 1.95 (1.95 - 1.98) [+]	50.00 – 2.70 (2.70 - 2.75) [+]	50.00 – 1.80 (1.80 - 1.83) [+]	50.00 – 1.56 (1.56 - 1.59) [+]
R_{merge}[#] (%)	8.0 (76.7)	12.2 (92.6)	9.2 (95.3)	5.2 (61.5)
No. of reflections (Bijvoet merged)	360750/15647	44824/5527	403463/42748	315705/ 67995
$<I/\sigma I>$	30.5 (2.5)	18.9 (1.8)	16.3 (1.3)	28.1 (1.9)
Completeness (%)	99.9 (100)	93.6 (95.8)	100 (100)	99.3 (97.4)
Multiplicity	7.3 (7.4)	8.1 (8.1)	4.7 (4.6)	4.6 (3.8)
Refinement				
No. of reflections(refinement/R_{free})	15616/782	5247/277	-	64442/3363
R_{work} / R_{free}[‡]	0.167/0.215	0.149/0.258	-	0.139/0.165
No. Atoms				
Protein	1551	1674	-	3741
Heterogens	31	1	-	8
Water	137	26	-	530
Ramachandran plot statistics[§] (%)				
Favored/outliers	95.79/0	97.18/0	-	98.58/0.22
R.m.s. deviations from target values				
Bond lengths (Å)	0.013	0.012	-	0.013
Bond angles (°)	1.5	1.6	-	1.6

[+] Values in parentheses are for the highest resolution shell.

[#] $R_{merge} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where ‹I(hkl)› is the mean of I observations $I_i(hkl)$ of reflection hkl.

[‡] $R = \sum ||F_o| - |F_c|| / \sum |F_o|$, where F_o and F_c are the observed and calculated structure factors, respectively, calculated for working set and R_{free} reflections.

[§] Ramachandran plot statistics are calculated by Coot.

Table S2. The list of deposits used to identify crystallization artifacts by MR which is carried out using predefined set of structures of possible artifacts as templates together with the resulting *R*-factor values.

Resulting overall R factor <= 0.30 colored in green. Missing values indicate that MOLREP was not able to find any solution.

			R-factor			
Name of protein	**PDB**	**MW**	**Sigma70**	**GNAT**	**xSix**	**021790**
Metal-binding lipocalin (YodA)	1OEJ	24.5	0.28	0.24	0.54	0.55
	1OEE	22.3	0.29	0.26	0.56	0.55
	4TNN	24.5	0.24	0.30	0.56	-
	1S7D	24.8	0.28	0.24	0.55	0.54
	1TXL	24.6	0.29	0.25	0.55	0.54
Carbonic dehydratase (YadF)	1T75	25.1	0.54	0.52	0.21	0.55
	3QY1	25.1	0.54	0.52	0.22	0.55
	1I6O	27.2	0.54	0.51	0.22	0.54
	1I6P	27.2	0.54	0.52	0.2	0.54
Ferric uptake regulator (Fur)	2FU4	9.33	0.53	0.51	0.55	0.55
cAMP-regulatory protein (CRP)	2GZW	23.51	0.55	0.51	0.52	0.54
Glucosamine-6-phosphate synthase	1MOQ	40.31	-	0.53	-	0.54
	2J6H	66.76	-	0.55	-	0.55
	4AMV	66.86	-	0.55	-	0.55
Glycogen synthase (GlgA)	2QZS	53.89	-	0.54	-	0.55
Component 1 of the 2-oxoglutarate	2JGD	105.03	-	-	-	0.56
Component E2 of the dihydrolipoamide	1E2O	26.07	0.53	0.52	0.55	0.54
Formyl transferase (YfbG)	2BLN	33.21	0.54	0.52	0.56	0.55
	1YRW	32.83	0.54	0.53	0.56	-
	2BLL	40	0.54	0.53	-	0.55
Cu/Zn-superoxide dismutase (Cu/Zn-	1ESO	15.74	0.53	0.51	-	0.54

Chloramphenicol-O-acetyl transferase	3U9F	25.66	0.52	0.52	0.55	0.54
	1PD5	25.66	0.53	0.51	0.56	0.54
Host factor-I protein (Hfq)	4PNO	8.12	0.54	0.52	0.54	0.55
	2YHT	8.25	0.53	0.52	-	0.55
Maltose-binding protein (MBP)	1PEB	40.06	-	0.54	-	0.54
	1NL5	40.06	0.54	0.52	-	0.54
	1ZJL	40.06	0.55	0.52	-	0.55
	1ZIU	40.06	-	0.53	-	0.55
N-Utilization substance (NusA)	2CXC	16.04	0.53	0.51	-	0.54
	1U9L	7.77	0.53	0.52	0.53	0.55
	1L2F	40.9	0.54	0.53	-	0.54
	1K0R	39.81	0.55	0.51	0.55	0.54