wwPDB X-ray Structure Validation Summary Report (i)

Nov 21, 2023 – 12:45 AM JST

PDB ID	:	7E6O
Title	:	Crystal structure of polyol dehydrogenase from Paracoccus denitrificans
Authors	:	Sha, F.; Zheng, Y.C.
Deposited on	:	2021-02-22
Resolution	:	2.10 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (i)) were used in the production of this report:

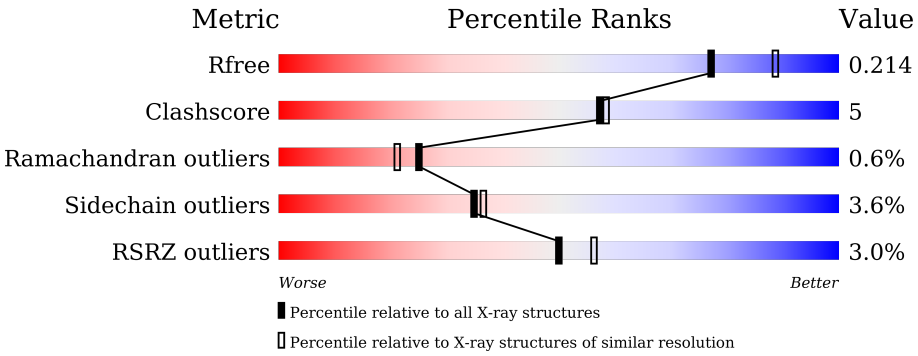
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	264	4% 81% 10% 8%
1	B	264	3% 80% 9% 9%
1	C	264	3% 84% 11%
1	D	264	% 85% 9% 5%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7684 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Short-chain dehydrogenase/reductase SDR.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	256	Total	C	N	O	S	8	1	0
			1885	1186	339	347	13			
1	A	242	Total	C	N	O	S	8	0	0
			1752	1101	315	324	12			
1	B	239	Total	C	N	O	S	8	0	0
			1723	1083	309	319	12			
1	D	250	Total	C	N	O	S	4	0	0
			1821	1147	326	336	12			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-7	MET	-	initiating methionine	UNP A1BBK7
C	-6	GLY	-	expression tag	UNP A1BBK7
C	-5	HIS	-	expression tag	UNP A1BBK7
C	-4	HIS	-	expression tag	UNP A1BBK7
C	-3	HIS	-	expression tag	UNP A1BBK7
C	-2	HIS	-	expression tag	UNP A1BBK7
C	-1	HIS	-	expression tag	UNP A1BBK7
C	0	HIS	-	expression tag	UNP A1BBK7
A	-7	MET	-	initiating methionine	UNP A1BBK7
A	-6	GLY	-	expression tag	UNP A1BBK7
A	-5	HIS	-	expression tag	UNP A1BBK7
A	-4	HIS	-	expression tag	UNP A1BBK7
A	-3	HIS	-	expression tag	UNP A1BBK7
A	-2	HIS	-	expression tag	UNP A1BBK7
A	-1	HIS	-	expression tag	UNP A1BBK7
A	0	HIS	-	expression tag	UNP A1BBK7
B	-7	MET	-	initiating methionine	UNP A1BBK7
B	-6	GLY	-	expression tag	UNP A1BBK7
B	-5	HIS	-	expression tag	UNP A1BBK7
B	-4	HIS	-	expression tag	UNP A1BBK7
B	-3	HIS	-	expression tag	UNP A1BBK7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	HIS	-	expression tag	UNP A1BBK7
B	-1	HIS	-	expression tag	UNP A1BBK7
B	0	HIS	-	expression tag	UNP A1BBK7
D	-7	MET	-	initiating methionine	UNP A1BBK7
D	-6	GLY	-	expression tag	UNP A1BBK7
D	-5	HIS	-	expression tag	UNP A1BBK7
D	-4	HIS	-	expression tag	UNP A1BBK7
D	-3	HIS	-	expression tag	UNP A1BBK7
D	-2	HIS	-	expression tag	UNP A1BBK7
D	-1	HIS	-	expression tag	UNP A1BBK7
D	0	HIS	-	expression tag	UNP A1BBK7

- Molecule 2 is water.

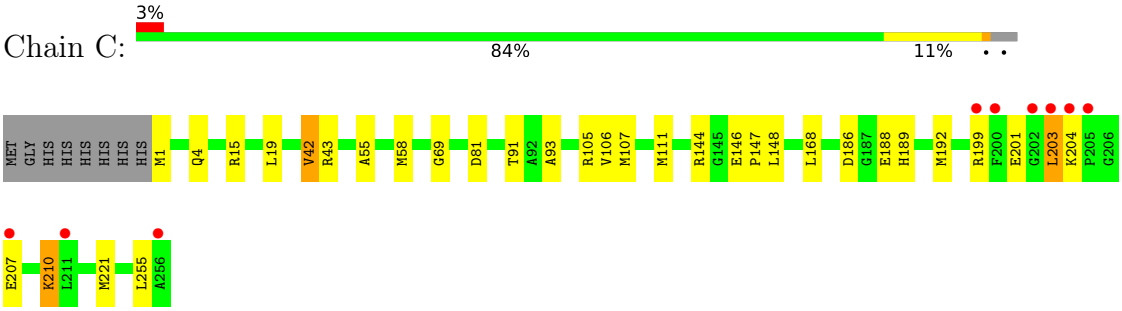
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	C	131	Total O 131 131	0	0
2	A	109	Total O 109 109	0	0
2	B	124	Total O 124 124	0	0
2	D	139	Total O 139 139	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

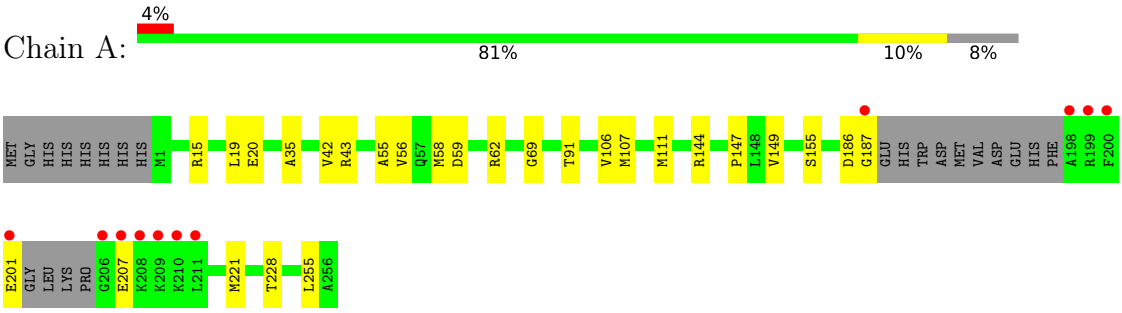
• Molecule 1: Short-chain dehydrogenase/reductase SDR

Chain C:



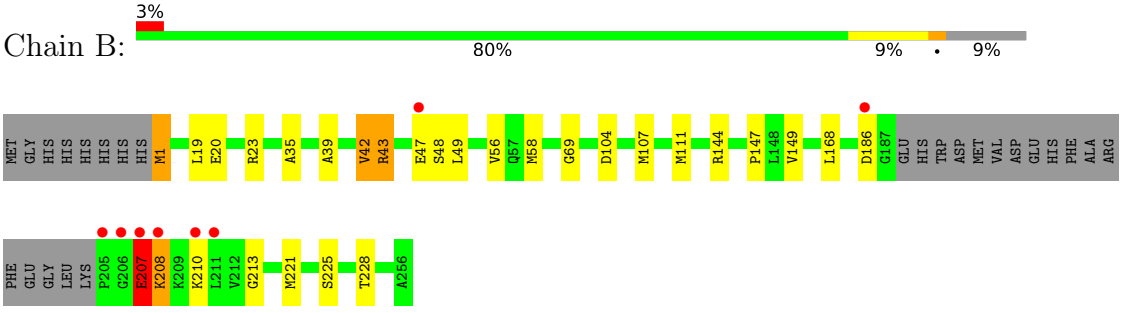
• Molecule 1: Short-chain dehydrogenase/reductase SDR

Chain A:



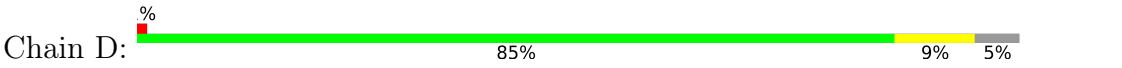
• Molecule 1: Short-chain dehydrogenase/reductase SDR

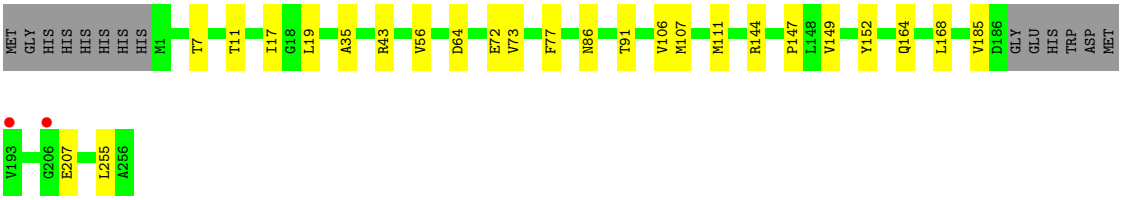
Chain B:



• Molecule 1: Short-chain dehydrogenase/reductase SDR

Chain D:




MET
GLY
HIS
HIS
HIS
HIS
HIS
M1
T7
T11
I17
G18
L19
A35
R43
V56
D64
E72
V73
F77
N86
T91
V106
M107
M111
R144
P147
L148
V149
Y152
Q164
L168
V185
D186
GLY
GLU
HIS
TRP
ASP
MET
V193
G206
E207
L255
A256

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	77.71Å 86.89Å 134.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.97 – 2.10 43.93 – 2.10	Depositor EDS
% Data completeness (in resolution range)	94.4 (43.97-2.10) 94.4 (43.93-2.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	5.39 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.160 , 0.206 0.171 , 0.214	Depositor DCC
R_{free} test set	2523 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	0.101	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 49.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	7684	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.16% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	#$\|Z\|$ >5	RMSZ	#$\|Z\|$ >5
1	A	0.77	1/1773 (0.1%)	0.85	1/2392 (0.0%)
1	B	0.77	1/1745 (0.1%)	0.89	1/2357 (0.0%)
1	C	0.86	1/1914 (0.1%)	0.90	3/2586 (0.1%)
1	D	0.95	1/1846 (0.1%)	0.86	3/2493 (0.1%)
All	All	0.84	4/7278 (0.1%)	0.87	8/9828 (0.1%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	43	ARG	CD-NE	-25.83	1.02	1.46
1	C	43	ARG	CD-NE	-19.65	1.13	1.46
1	A	43	ARG	CD-NE	-10.22	1.29	1.46
1	B	43	ARG	CD-NE	6.96	1.58	1.46

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	43	ARG	CG-CD-NE	12.31	137.66	111.80
1	D	43	ARG	CG-CD-NE	8.73	130.13	111.80
1	D	64	ASP	CB-CG-OD1	8.30	125.77	118.30
1	C	15	ARG	CD-NE-CZ	6.60	132.84	123.60
1	C	43	ARG	CD-NE-CZ	6.48	132.67	123.60

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1752	0	1777	19	0
1	B	1723	0	1753	19	0
1	C	1885	0	1896	20	0
1	D	1821	0	1844	17	0
2	A	109	0	0	1	0
2	B	124	0	0	4	0
2	C	131	0	0	2	0
2	D	139	0	0	2	0
All	All	7684	0	7270	70	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 5.

The worst 5 of 70 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:186:ASP:HA	1:C:221:MET:HE3	1.48	0.95
1:A:147:PRO:O	2:A:301:HOH:O	1.98	0.80
1:B:107:MET:SD	1:B:111:MET:CE	2.70	0.80
1:A:107:MET:SD	1:A:111:MET:HE3	2.22	0.80
1:C:186:ASP:HA	1:C:221:MET:CE	2.14	0.78

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	A	236/264 (89%)	228 (97%)	8 (3%)	0	100 100
1	B	235/264 (89%)	222 (94%)	10 (4%)	3 (1%)	12 7
1	C	255/264 (97%)	242 (95%)	10 (4%)	3 (1%)	13 8
1	D	246/264 (93%)	239 (97%)	7 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	972/1056 (92%)	931 (96%)	35 (4%)	6 (1%)	25 21

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	208	LYS
1	C	203	LEU
1	B	186	ASP
1	C	204	LYS
1	C	93	ALA

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	167/187 (89%)	163 (98%)	4 (2%)	49 53
1	B	165/187 (88%)	157 (95%)	8 (5%)	25 24
1	C	181/187 (97%)	171 (94%)	10 (6%)	21 19
1	D	175/187 (94%)	172 (98%)	3 (2%)	60 67
All	All	688/748 (92%)	663 (96%)	25 (4%)	35 36

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1	MET
1	B	43	ARG
1	D	207	GLU
1	B	42	VAL
1	B	47	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no monosaccharides in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	242/264 (91%)	-0.11	11 (4%)	33	38	18, 27, 53, 133	2 (0%)
1	B	239/264 (90%)	-0.16	8 (3%)	46	53	17, 25, 51, 82	2 (0%)
1	C	256/264 (96%)	-0.09	9 (3%)	44	50	16, 26, 62, 120	2 (0%)
1	D	250/264 (94%)	-0.23	2 (0%)	86	88	18, 26, 45, 71	1 (0%)
All	All	987/1056 (93%)	-0.15	30 (3%)	50	56	16, 26, 52, 133	7 (0%)

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	203	LEU	11.9
1	A	201	GLU	9.7
1	A	206	GLY	8.1
1	C	202	GLY	8.0
1	A	207	GLU	8.0

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.