



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 03:30 AM EDT

PDB ID : 5MVI
Title : Crystal structure of 2-methylcitrate dehydratase (PrpD) from *Salmonella enterica*
Authors : Race, P.R.; Baker, G.E.
Deposited on : 2017-01-16
Resolution : 3.05 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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.37.1
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) : 2.37.1

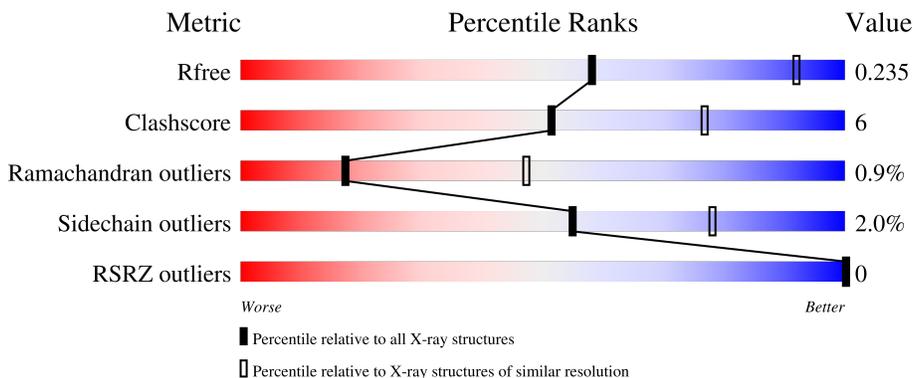
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.05 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	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)

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 $\leq 5\%$. 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	483	76% (green), 18% (yellow), 6% (orange), 0% (red), 0% (grey)
1	B	483	63% (green), 6% (yellow), 31% (grey)
1	C	483	62% (green), 6% (yellow), 31% (grey)
1	D	483	64% (green), 6% (yellow), 30% (grey)
1	E	483	63% (green), 5% (yellow), 31% (grey)

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	483	 63% 5% 31%

2 Entry composition [i](#)

There is only 1 type of molecule in this entry. The entry contains 15106 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 2-methylcitrate dehydratase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	E	334	2341	1491	410	428	12	0	0	0
1	D	337	2406	1534	417	444	11	0	0	0
1	A	463	3283	2100	566	604	13	0	0	0
1	B	334	2359	1508	409	433	9	0	0	0
1	C	332	2375	1517	409	437	12	0	0	0
1	F	334	2342	1493	407	429	13	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	8	VAL	ILE	conflict	UNP A0A0L5RC52
E	123	THR	ILE	conflict	UNP A0A0L5RC52
E	418	GLY	ASP	conflict	UNP A0A0L5RC52
D	8	VAL	ILE	conflict	UNP A0A0L5RC52
D	123	THR	ILE	conflict	UNP A0A0L5RC52
D	418	GLY	ASP	conflict	UNP A0A0L5RC52
A	8	VAL	ILE	conflict	UNP A0A0L5RC52
A	123	THR	ILE	conflict	UNP A0A0L5RC52
A	418	GLY	ASP	conflict	UNP A0A0L5RC52
B	8	VAL	ILE	conflict	UNP A0A0L5RC52
B	123	THR	ILE	conflict	UNP A0A0L5RC52
B	418	GLY	ASP	conflict	UNP A0A0L5RC52
C	8	VAL	ILE	conflict	UNP A0A0L5RC52
C	123	THR	ILE	conflict	UNP A0A0L5RC52
C	418	GLY	ASP	conflict	UNP A0A0L5RC52
F	8	VAL	ILE	conflict	UNP A0A0L5RC52
F	123	THR	ILE	conflict	UNP A0A0L5RC52

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	418	GLY	ASP	conflict	UNP A0A0L5RC52

ALA
VAL
PRO
LEU
LEU
PHE
GLY
ARG
LEU
THR
ALA
ALA
ASP
TYR
GLU
ASP
GLU
VAL
ALA
GLN
ASP
LYS
ARG
ILE
ASP
ALA
LEU
ARG
GLU
LYS
ILE
VAL
CYS
TYR
GLU
ASP
PRO
ALA
PHE
THR
ALA
ASP
TYR
HIS
ASP
PRO
GLU
LYS
ARG
ALA
ILE
GLY
ASN
ALA
ILE
SER
VAL
GLU
PHE
THR

ASP
GLY
SER
ARG
PHE
GLY
VAL
VAL
VAL
GLU
TYR
P425
M446
R449
A467
R468
V474
V482
ILE

4 Data and refinement statistics i

Property	Value	Source
Space group	H 3	Depositor
Cell constants a, b, c, α , β , γ	139.66Å 139.66Å 513.48Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	39.14 – 3.05 39.14 – 3.05	Depositor EDS
% Data completeness (in resolution range)	99.5 (39.14-3.05) 99.6 (39.14-3.05)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.71 (at 3.06Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.213 , 0.239 0.211 , 0.235	Depositor DCC
R_{free} test set	3483 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å ²)	25.5	Xtriage
Anisotropy	0.373	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 13.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.36$, $\langle L^2 \rangle = 0.19$	Xtriage
Estimated twinning fraction	0.387 for -h-k,k,-l	Xtriage
Reported twinning fraction	0.532 for H, K, L 0.468 for K, H, -L	Depositor
Outliers	0 of 70839 reflections	Xtriage
F_o, F_c correlation	0.80	EDS
Total number of atoms	15106	wwPDB-VP
Average B, all atoms (Å ²)	9.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality [i](#)

5.1 Standard geometry [i](#)

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	0.79	1/3352 (0.0%)	0.76	1/4559 (0.0%)
1	B	0.73	0/2406	0.71	2/3272 (0.1%)
1	C	0.66	0/2424	0.71	3/3293 (0.1%)
1	D	0.68	1/2454 (0.0%)	0.72	3/3333 (0.1%)
1	E	0.67	0/2385	0.71	2/3239 (0.1%)
1	F	0.59	1/2385 (0.0%)	0.66	2/3238 (0.1%)
All	All	0.70	3/15406 (0.0%)	0.72	13/20934 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	4
1	B	0	1
1	D	0	1
1	E	0	2
1	F	0	1
All	All	0	9

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	481	TYR	CB-CG	11.14	1.68	1.51
1	D	202	TRP	CB-CG	-5.81	1.39	1.50
1	F	126	TRP	CB-CG	5.12	1.59	1.50

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	28	ILE	CG1-CB-CG2	-8.72	92.22	111.40
1	D	49	LEU	CA-CB-CG	6.50	130.26	115.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	478	LEU	CA-CB-CG	6.50	130.26	115.30
1	C	480	LEU	CA-CB-CG	6.38	129.98	115.30
1	C	127	LEU	CA-CB-CG	6.07	129.26	115.30

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	317	ILE	Peptide
1	A	318	GLU	Peptide
1	D	424	TYR	Peptide
1	E	129	ARG	Peptide
1	E	285	ILE	Peptide

5.2 Too-close contacts [i](#)

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	3283	0	2890	69	0
1	B	2359	0	2114	24	0
1	C	2375	0	2152	24	0
1	D	2406	0	2172	17	0
1	E	2341	0	2098	20	0
1	F	2342	0	2113	18	0
All	All	15106	0	13539	160	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 6.

The worst 5 of 160 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:B:460:LEU:O	1:B:461:ASP:O	1.81	0.98
1:E:24:MET:CE	1:E:191:ARG:HA	2.02	0.89
1:A:38:HIS:ND1	1:A:466:ARG:CG	2.45	0.80

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:283:PHE:O	1:A:427:GLY:HA3	1.83	0.79
1:E:24:MET:HE3	1:E:191:ARG:HA	1.63	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	459/483 (95%)	431 (94%)	23 (5%)	5 (1%)	14	42
1	B	330/483 (68%)	319 (97%)	8 (2%)	3 (1%)	17	47
1	C	328/483 (68%)	319 (97%)	7 (2%)	2 (1%)	25	55
1	D	333/483 (69%)	323 (97%)	7 (2%)	3 (1%)	17	47
1	E	330/483 (68%)	317 (96%)	10 (3%)	3 (1%)	17	47
1	F	330/483 (68%)	320 (97%)	6 (2%)	4 (1%)	13	40
All	All	2110/2898 (73%)	2029 (96%)	61 (3%)	20 (1%)	17	47

5 of 20 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	21	ASP
1	A	204	ASP
1	A	415	SER
1	B	461	ASP
1	C	204	ASP

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	270/395 (68%)	261 (97%)	9 (3%)	38	67
1	B	196/395 (50%)	194 (99%)	2 (1%)	76	89
1	C	207/395 (52%)	201 (97%)	6 (3%)	42	70
1	D	207/395 (52%)	203 (98%)	4 (2%)	57	79
1	E	194/395 (49%)	191 (98%)	3 (2%)	65	83
1	F	196/395 (50%)	194 (99%)	2 (1%)	76	89
All	All	1270/2370 (54%)	1244 (98%)	26 (2%)	55	78

5 of 26 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	475	ASN
1	B	266	THR
1	F	262	PHE
1	B	262	PHE
1	C	204	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	280	ASN
1	A	325	HIS
1	A	455	GLN
1	C	280	ASN
1	F	280	ASN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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	463/483 (95%)	-0.79	0 100 100	2, 2, 24, 41	0
1	B	334/483 (69%)	-0.78	0 100 100	2, 2, 23, 43	0
1	C	332/483 (68%)	-0.71	0 100 100	2, 4, 26, 53	0
1	D	337/483 (69%)	-0.72	0 100 100	2, 2, 37, 69	0
1	E	334/483 (69%)	-0.73	0 100 100	2, 2, 31, 52	0
1	F	334/483 (69%)	-0.67	0 100 100	2, 11, 32, 53	0
All	All	2134/2898 (73%)	-0.74	0 100 100	2, 3, 29, 69	0

There are no RSRZ outliers to report.

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.