Full wwPDB X-ray Structure Validation Report ⓘ

Nov 14, 2023 – 11:31 AM JST

PDB ID	:	5Z5M
Title	:	Crystal structure of (S)-allantoin synthase
Authors	:	Oh, J.; Percudani, R.; Rhee, S.
Deposited on	:	2018-01-18
Resolution	:	1.85 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

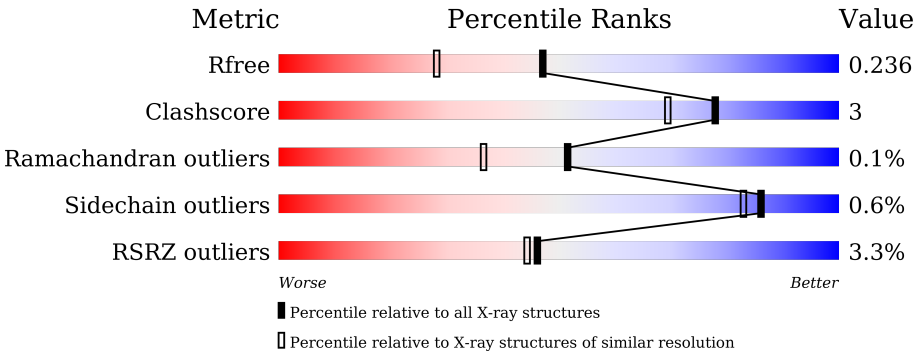
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 1.85 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	2469 (1.86-1.86)
Clashscore	141614	2625 (1.86-1.86)
Ramachandran outliers	138981	2592 (1.86-1.86)
Sidechain outliers	138945	2592 (1.86-1.86)
RSRZ outliers	127900	2436 (1.86-1.86)

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	299	2% 86% 6% 9%
1	B	299	% 83% 8% 9%
1	C	299	6% 83% 8% 9%
1	D	299	3% 85% 6% 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9263 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 Predicted protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	273	Total	C	N	O	S	0	0	0
			2101	1340	357	394	10			
1	B	272	Total	C	N	O	S	0	0	0
			2099	1338	357	394	10			
1	C	272	Total	C	N	O	S	0	0	0
			2096	1337	356	393	10			
1	D	272	Total	C	N	O	S	0	0	0
			2069	1323	355	381	10			

- Molecule 2 is water.

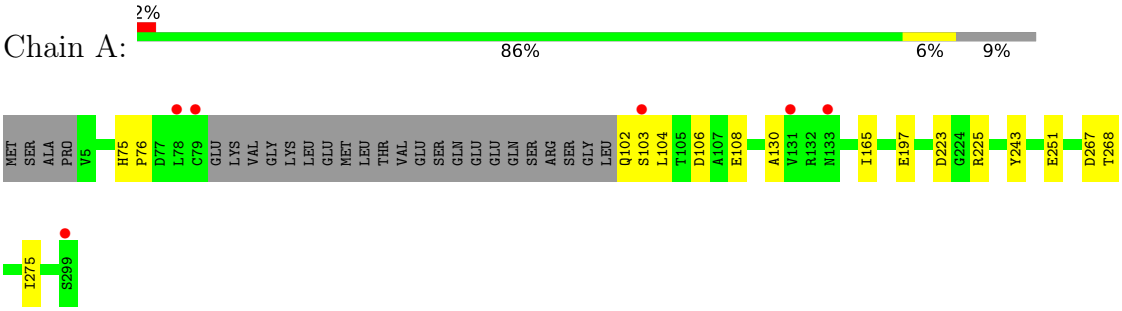
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	228	Total	O	0	0
			228	228		
2	B	251	Total	O	0	0
			251	251		
2	C	217	Total	O	0	0
			217	217		
2	D	202	Total	O	0	0
			202	202		

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.

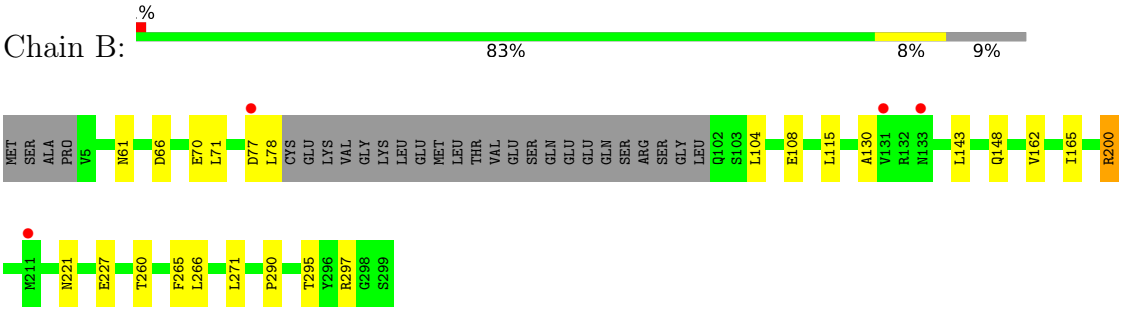
• Molecule 1: Predicted protein

Chain A:



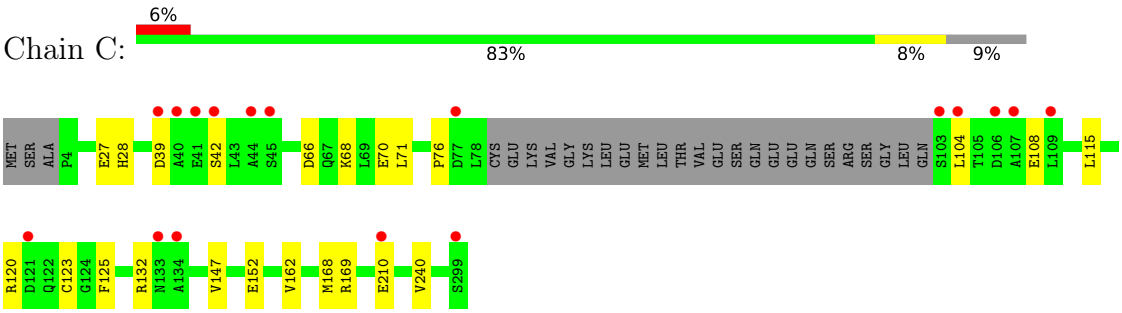
• Molecule 1: Predicted protein

Chain B:



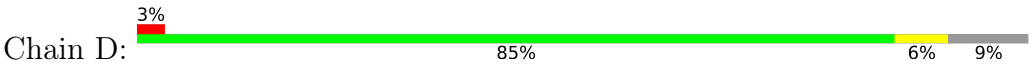
• Molecule 1: Predicted protein

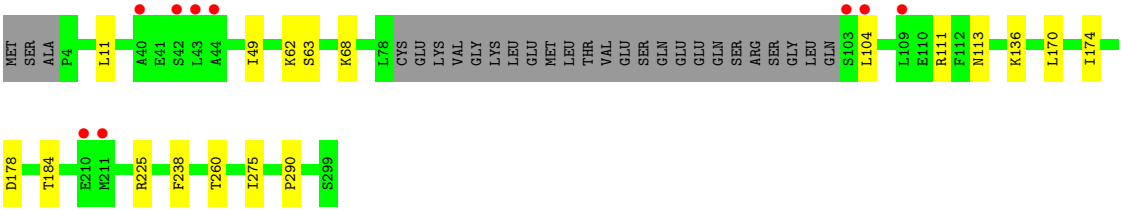
Chain C:



• Molecule 1: Predicted protein

Chain D:




MET
SER
ALA
P4
L11
A40
E41
S42
L43
A44
I49
K62
S63
K68
L78
CYS
GLU
LYS
VAL
GLY
LYS
LEU
GLU
MET
LEU
THR
VAL
GLU
SER
GLN
GLU
GLU
GLN
SER
ARG
SER
GLY
LEU
GLN
S103
L104
L109
E110
R111
F112
N113
K136
L170
I174
D178
T184
E210
M211
R225
F238
T260
I275
P290
S299

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	65.55Å 99.92Å 90.18Å 90.00° 109.35° 90.00°	Depositor
Resolution (Å)	30.92 – 1.85 30.92 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.5 (30.92-1.85) 99.6 (30.92-1.85)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.54 (at 1.85Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.199 , 0.237 0.199 , 0.236	Depositor DCC
R_{free} test set	4615 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å^2)	15.6	Xtriage
Anisotropy	0.696	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 57.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.034 for h,-k,-h-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9263	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% 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	0.27	0/2151	0.45	0/2916
1	B	0.28	0/2149	0.46	0/2913
1	C	0.27	0/2147	0.45	0/2911
1	D	0.27	0/2120	0.46	0/2877
All	All	0.27	0/8567	0.45	0/11617

There are no bond length outliers.

There are no bond angle outliers.

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	2101	0	2051	9	0
1	B	2099	0	2053	13	0
1	C	2096	0	2051	16	0
1	D	2069	0	2015	11	0
2	A	228	0	0	2	0
2	B	251	0	0	2	0
2	C	217	0	0	2	0
2	D	202	0	0	2	0
All	All	9263	0	8170	48	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 3.

All (48) 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:D:184:THR:HB	1:D:225:ARG:HH12	1.47	0.79
1:C:152:GLU:OE2	1:D:68:LYS:NZ	2.12	0.79
1:C:132:ARG:NH1	1:C:168:MET:SD	2.61	0.73
1:C:71:LEU:HD21	1:C:162:VAL:HG11	1.78	0.65
1:B:221:ASN:HB3	1:B:227:GLU:HG2	1.80	0.62
1:A:104:LEU:HD22	1:A:108:GLU:HB3	1.82	0.61
1:A:223:ASP:OD2	1:A:225:ARG:NH1	2.34	0.58
1:B:71:LEU:HD21	1:B:162:VAL:HG11	1.85	0.57
1:A:267:ASP:OD1	1:A:268:THR:N	2.32	0.57
1:D:104:LEU:HD23	1:D:136:LYS:HE3	1.87	0.57
1:D:178:ASP:OD1	1:D:178:ASP:N	2.41	0.52
1:C:68:LYS:NZ	2:C:803:HOH:O	2.42	0.52
1:B:130:ALA:HB2	1:B:165:ILE:HG13	1.92	0.52
1:B:200:ARG:NH1	2:B:309:HOH:O	2.42	0.51
1:A:130:ALA:HB2	1:A:165:ILE:HG13	1.93	0.51
1:B:104:LEU:HD22	1:B:108:GLU:HB3	1.94	0.50
1:B:297:ARG:HH21	1:B:297:ARG:HG2	1.77	0.49
1:A:197:GLU:HG2	2:A:421:HOH:O	2.13	0.49
1:C:66:ASP:O	1:C:70:GLU:HG3	2.12	0.49
1:C:123:CYS:HB3	1:C:125:PHE:CE1	2.48	0.48
1:D:63:SER:O	1:D:68:LYS:HE3	2.13	0.47
1:B:115:LEU:HB3	1:B:143:LEU:HG	1.97	0.47
1:C:115:LEU:HD22	1:C:147:VAL:HG11	1.97	0.46
1:B:61:ASN:ND2	2:B:307:HOH:O	2.40	0.46
1:C:120:ARG:HH11	1:C:120:ARG:HG3	1.81	0.46
1:D:111:ARG:NE	2:D:310:HOH:O	2.49	0.45
1:C:169:ARG:NH1	2:C:813:HOH:O	2.50	0.45
1:B:265:PHE:CD2	1:B:266:LEU:HG	2.52	0.44
1:D:62:LYS:NZ	2:D:303:HOH:O	2.37	0.44
1:A:75:HIS:HA	1:A:76:PRO:HD3	1.86	0.44
1:C:27:GLU:HG2	1:C:76:PRO:HD3	2.00	0.44
1:D:49:ILE:HG13	1:D:174:ILE:HD12	2.00	0.43
1:A:102:GLN:HB2	1:A:103:SER:H	1.63	0.43
1:A:251:GLU:OE2	2:A:301:HOH:O	2.21	0.43
1:C:104:LEU:HD22	1:C:108:GLU:HB3	2.01	0.43
1:B:77:ASP:OD2	1:B:78:LEU:N	2.52	0.43
1:C:132:ARG:NH2	1:C:240:VAL:HG11	2.34	0.42
1:A:243:TYR:CE1	1:A:275:ILE:HD12	2.53	0.42
1:C:152:GLU:H	1:C:152:GLU:CD	2.23	0.42
1:B:66:ASP:O	1:B:70:GLU:HG3	2.19	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:170:LEU:HG	1:D:174:ILE:HD12	2.02	0.42
1:B:260:THR:HG22	1:B:290:PRO:HB3	2.02	0.41
1:D:238:PHE:CE1	1:D:275:ILE:HD13	2.55	0.41
1:B:271:LEU:HD21	1:B:295:THR:HG21	2.02	0.41
1:C:27:GLU:HG3	1:C:28:HIS:ND1	2.35	0.41
1:C:123:CYS:HB3	1:C:125:PHE:CZ	2.56	0.41
1:D:260:THR:HG22	1:D:290:PRO:HB3	2.03	0.41
1:C:39:ASP:HB2	1:C:42:SER:HB3	2.04	0.40

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	269/299 (90%)	266 (99%)	3 (1%)	0	100	100
1	B	268/299 (90%)	264 (98%)	4 (2%)	0	100	100
1	C	268/299 (90%)	263 (98%)	5 (2%)	0	100	100
1	D	268/299 (90%)	262 (98%)	5 (2%)	1 (0%)	34	19
All	All	1073/1196 (90%)	1055 (98%)	17 (2%)	1 (0%)	51	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	113	ASN

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	222/249 (89%)	221 (100%)	1 (0%)	88 86
1	B	223/249 (90%)	221 (99%)	2 (1%)	78 72
1	C	223/249 (90%)	222 (100%)	1 (0%)	91 89
1	D	215/249 (86%)	214 (100%)	1 (0%)	88 86
All	All	883/996 (89%)	878 (99%)	5 (1%)	86 83

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	106	ASP
1	B	148	GLN
1	B	200	ARG
1	C	210	GLU
1	D	11	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	148	GLN
1	D	113	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

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	273/299 (91%)	-0.10	6 (2%)	62 61	8, 19, 40, 51	0
1	B	272/299 (90%)	-0.09	4 (1%)	73 74	8, 19, 43, 63	0
1	C	272/299 (90%)	0.19	17 (6%)	20 19	7, 22, 52, 79	0
1	D	272/299 (90%)	0.06	9 (3%)	46 44	7, 21, 46, 65	0
All	All	1089/1196 (91%)	0.02	36 (3%)	46 44	7, 20, 46, 79	0

All (36) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	41	GLU	6.0
1	D	42	SER	5.9
1	C	134	ALA	5.6
1	D	103	SER	5.5
1	C	133	ASN	5.5
1	C	40	ALA	5.3
1	C	210	GLU	4.9
1	D	40	ALA	4.2
1	C	103	SER	3.9
1	B	211	MET	3.8
1	D	104	LEU	3.8
1	C	42	SER	3.8
1	C	104	LEU	3.6
1	A	133	ASN	3.5
1	C	39	ASP	3.5
1	B	133	ASN	3.5
1	C	121	ASP	3.4
1	D	44	ALA	3.1
1	A	79	CYS	3.1
1	C	45	SER	3.1
1	D	109	LEU	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	131	VAL	2.8
1	C	106	ASP	2.8
1	C	44	ALA	2.7
1	D	211	MET	2.7
1	C	107	ALA	2.6
1	D	210	GLU	2.5
1	A	103	SER	2.5
1	C	77	ASP	2.4
1	A	78	LEU	2.3
1	B	131	VAL	2.2
1	A	299	SER	2.2
1	B	77	ASP	2.2
1	C	109	LEU	2.1
1	C	299	SER	2.1
1	D	43	LEU	2.1

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.