WORLDWIDE wwPDB

PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jun 27, 2022 – 06:12 PM EDT

PDB ID	:	8D27
Title	:	Arginase Domain of Ornithine Decarboxylase/Arginase from Fusobacterium nucleatum
Authors	:	Chan, A.C.; Kolesnikov, M.; Murphy, M.E.
Deposited on	:	2022-05-27
Resolution	:	2.25 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:

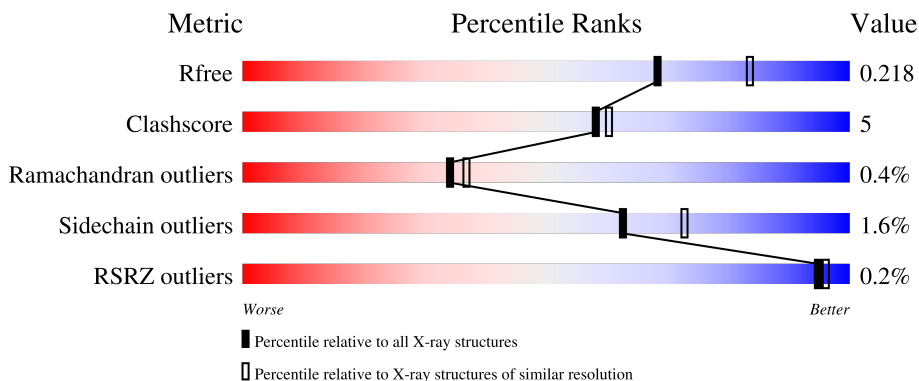|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.29
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

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 2.25 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	Percentile Ranks	Value
Rfree		0.218
Clashscore		5
Ramachandran outliers		0.4%
Sidechain outliers		1.6%
RSRZ outliers		0.2%

Worse *Better*

▌ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

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	287	84% 13% ··
1	B	287	89% 8% ·

2 Entry composition ⓘ

There are 2 unique types of molecules in this entry. The entry contains 8800 atoms, of which 4353 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 Arginase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	281	Total	C	H	N	O	S	0	0	0
			4372	1387	2184	361	430	10			
1	B	281	Total	C	H	N	O	S	0	0	0
			4351	1382	2169	360	430	10			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	497	GLY	-	expression tag	UNP Q8RG18
A	498	SER	-	expression tag	UNP Q8RG18
A	499	HIS	-	expression tag	UNP Q8RG18
B	497	GLY	-	expression tag	UNP Q8RG18
B	498	SER	-	expression tag	UNP Q8RG18
B	499	HIS	-	expression tag	UNP Q8RG18

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	47	Total	O	0	0
			47	47		
2	B	30	Total	O	0	0
			30	30		

3 Residue-property plots ⓘ

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: Arginase

Chain A:

• Molecule 1: Arginase

Chain B:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	44.82Å 78.22Å 73.53Å 90.00° 90.03° 90.00°	Depositor
Resolution (Å)	44.82 – 2.25 44.82 – 2.25	Depositor EDS
% Data completeness (in resolution range)	97.0 (44.82-2.25) 96.9 (44.82-2.25)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.97 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.19.2-4158	Depositor
R, R_{free}	0.180 , 0.212 0.183 , 0.218	Depositor DCC
R_{free} test set	1057 reflections (4.53%)	wwPDB-VP
Wilson B-factor (Å^2)	27.9	Xtriage
Anisotropy	0.831	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 20.1	EDS
L-test for twinning[2]	$< |L| > = 0.38, < L^2 > = 0.20$	Xtriage
Estimated twinning fraction	0.418 for h,-k,-l	Xtriage
Reported twinning fraction	0.450 for h,-k,-l	Depositor
Outliers	0 of 23341 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	8800	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.16% 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	0.28	0/2223	0.51	0/3006
1	B	0.34	0/2217	0.53	0/2999
All	All	0.31	0/4440	0.52	0/6005

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	2188	2184	2178	26	0
1	B	2182	2169	2160	17	0
2	A	47	0	0	0	0
2	B	30	0	0	0	0
All	All	4447	4353	4338	42	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 42 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:664:ILE:HG23	1:B:669:ARG:NH2	2.06	0.71

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:625:ILE:HG23	1:B:628:MET:HE2	1.77	0.66
1:B:638:ASP:HB3	1:B:641:LEU:HD12	1.78	0.65
1:B:574:VAL:HG21	1:B:597:VAL:HG22	1.79	0.64
1:A:628:MET:N	1:A:629:PRO:HD3	2.16	0.59

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	279/287 (97%)	266 (95%)	12 (4%)	1 (0%)	34	37
1	B	279/287 (97%)	266 (95%)	12 (4%)	1 (0%)	34	37
All	All	558/574 (97%)	532 (95%)	24 (4%)	2 (0%)	34	37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	628	MET
1	B	628	MET

5.3.2 Protein sidechains ⓘ

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.

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	244/252 (97%)	240 (98%)	4 (2%)	62	73
1	B	242/252 (96%)	238 (98%)	4 (2%)	60	71
All	All	486/504 (96%)	478 (98%)	8 (2%)	62	73

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	724	SER
1	B	693	VAL
1	B	523	ASP
1	A	732	SER
1	B	626	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	655	ASN
1	A	780	ASN
1	B	626	HIS

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, 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(Å²)	Q<0.9
1	A	281/287 (97%)	-0.36	1 (0%)	92	93	23, 31, 39, 60	0
1	B	281/287 (97%)	-0.32	0	100	100	24, 31, 40, 47	0
All	All	562/574 (97%)	-0.34	1 (0%)	95	96	23, 31, 40, 60	0

All (1) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	704	ASP	2.0

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 ⓘ

There are no such residues in this entry.