



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 11, 2024 – 10:19 PM EDT

PDB ID : 6NBK
Title : Crystal structure of Arginase from *Bacillus cereus*
Authors : Chang, C.; Evdokimova, E.; Mcchesney, M.; Joachimiak, A.; Savchenko, A.;
Center for Structural Genomics of Infectious Diseases (CSGID)
Deposited on : 2018-12-07
Resolution : 1.91 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	:	2.36.2
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)	:	2.36.2

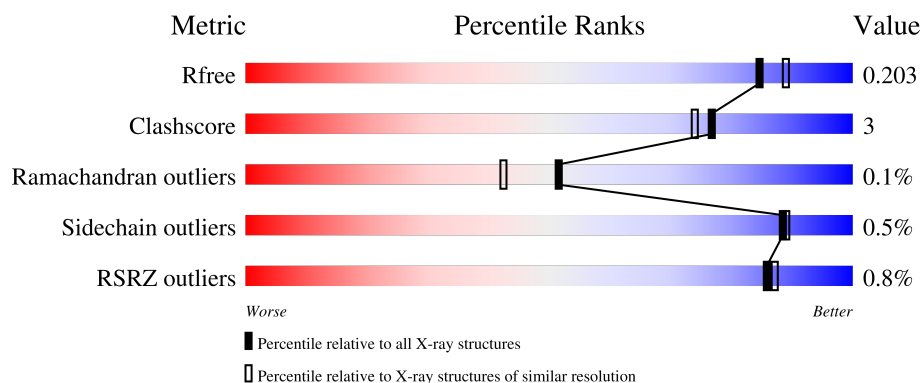
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.91 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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	297	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 89% 8% . </div> </div>
1	B	297	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 89% 7% . </div> </div>
1	C	297	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 91% 6% . </div> </div>
1	D	297	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 88% 9% . </div> </div>
1	E	297	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; width: 100%;"> % 91% 6% . </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	297	<div><div><div>%</div><div><div></div></div><div>91%</div><div>6%</div><div></div></div></div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 14725 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 Arginase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	289	Total	C	N	O	S	0	1	0
			2197	1375	376	432	14			
1	B	287	Total	C	N	O	S	0	0	0
			2173	1361	371	427	14			
1	C	286	Total	C	N	O	S	0	1	0
			2173	1361	372	426	14			
1	D	287	Total	C	N	O	S	0	1	0
			2176	1363	371	427	15			
1	E	288	Total	C	N	O	S	0	0	0
			2181	1366	372	429	14			
1	F	289	Total	C	N	O	S	0	1	0
			2193	1374	374	431	14			

- Molecule 2 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	3	Total	Mn	0	0
			3	3		
2	B	3	Total	Mn	0	0
			3	3		
2	C	2	Total	Mn	0	0
			2	2		
2	D	2	Total	Mn	0	0
			2	2		
2	E	2	Total	Mn	0	0
			2	2		
2	F	2	Total	Mn	0	0
			2	2		

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Ca 1 1	0	0
3	B	1	Total Ca 1 1	0	0
3	C	1	Total Ca 1 1	0	0

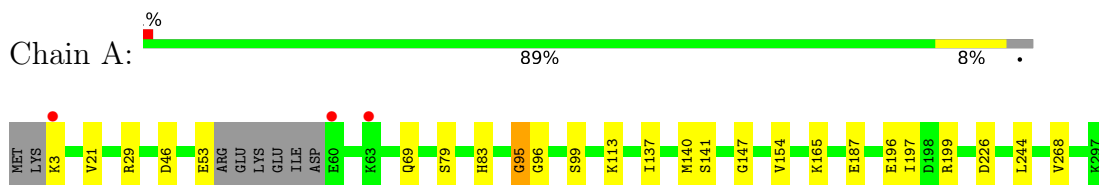
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	307	Total O 307 307	0	0
4	B	267	Total O 267 267	0	0
4	C	238	Total O 238 238	0	0
4	D	262	Total O 262 262	0	0
4	E	239	Total O 239 239	0	0
4	F	302	Total O 302 302	0	0

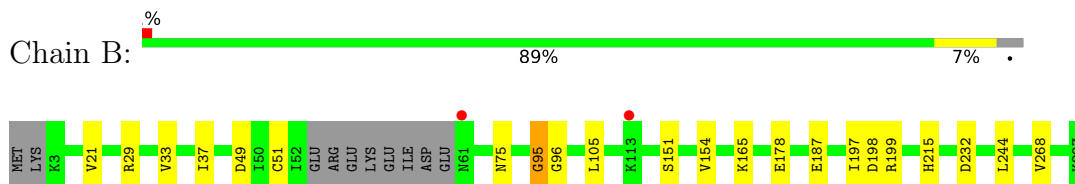
3 Residue-property plots [i](#)

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.

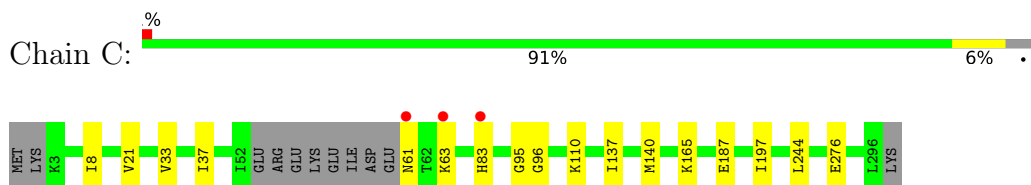
- Molecule 1: Arginase



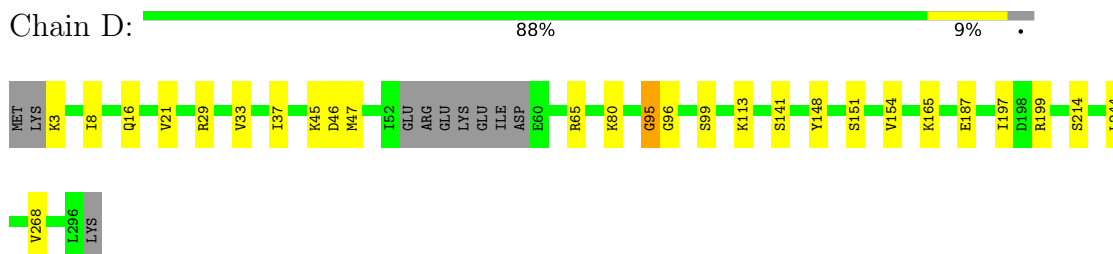
- Molecule 1: Arginase



- Molecule 1: Arginase



- Molecule 1: Arginase

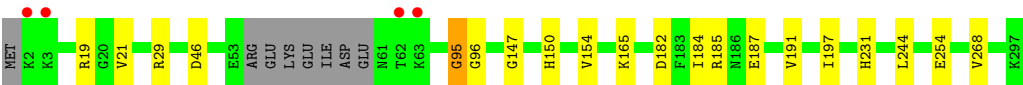
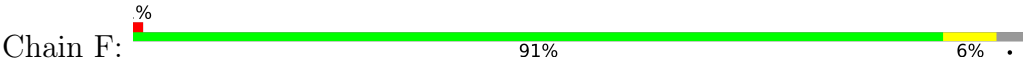


- Molecule 1: Arginase





● Molecule 1: Arginase



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	148.02Å 153.80Å 83.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.44 – 1.91 48.44 – 1.91	Depositor EDS
% Data completeness (in resolution range)	92.7 (48.44-1.91) 92.7 (48.44-1.91)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.31 (at 1.91Å)	Xtriage
Refinement program	PHENIX 1.12_2829	Depositor
R, R_{free}	0.161 , 0.203 0.161 , 0.203	Depositor DCC
R_{free} test set	6774 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	18.5	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 46.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.010 for k,h,-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	14725	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 3.82% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, MN

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	0.37	0/2231	0.57	1/3013 (0.0%)
1	B	0.34	0/2206	0.54	0/2979
1	C	0.33	0/2207	0.53	0/2983
1	D	0.35	0/2209	0.54	0/2985
1	E	0.34	0/2214	0.55	0/2991
1	F	0.36	0/2226	0.57	1/3008 (0.0%)
All	All	0.35	0/13293	0.55	2/17959 (0.0%)

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	1
1	B	0	1
1	C	0	1
1	D	0	1
1	E	0	1
1	F	0	1
All	All	0	6

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	F	19	ARG	NE-CZ-NH2	-6.72	116.94	120.30
1	A	226	ASP	CB-CG-OD1	5.87	123.58	118.30

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	95	GLY	Peptide
1	B	95	GLY	Peptide
1	C	95	GLY	Peptide
1	D	95	GLY	Peptide
1	E	95	GLY	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	2197	0	2172	16	0
1	B	2173	0	2158	13	0
1	C	2173	0	2151	9	0
1	D	2176	0	2155	16	0
1	E	2181	0	2164	12	0
1	F	2193	0	2174	12	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
2	E	2	0	0	0	0
2	F	2	0	0	0	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
4	A	307	0	0	3	2
4	B	267	0	0	5	0
4	C	238	0	0	1	0
4	D	262	0	0	2	1
4	E	239	0	0	3	1
4	F	302	0	0	3	2
All	All	14725	0	12974	78	3

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.

The worst 5 of 78 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:A:140:MET:SD	4:A:613:HOH:O	2.34	0.86
1:A:165:LYS:HD2	1:A:187:GLU:HB3	1.68	0.74
1:A:29:ARG:NH1	1:A:46:ASP:OD2	2.21	0.72
1:E:29:ARG:NH1	1:E:46:ASP:OD2	2.22	0.72
1:C:165:LYS:HD2	1:C:187:GLU:HB3	1.74	0.67

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:600:HOH:O	4:F:625:HOH:O[3_754]	2.14	0.06
4:A:483:HOH:O	4:F:548:HOH:O[3_754]	2.19	0.01
4:D:506:HOH:O	4:E:409:HOH:O[1_554]	2.19	0.01

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	286/297 (96%)	279 (98%)	7 (2%)	0	100	100
1	B	283/297 (95%)	273 (96%)	10 (4%)	0	100	100
1	C	283/297 (95%)	274 (97%)	9 (3%)	0	100	100
1	D	284/297 (96%)	277 (98%)	7 (2%)	0	100	100
1	E	284/297 (96%)	277 (98%)	6 (2%)	1 (0%)	34	24
1	F	286/297 (96%)	278 (97%)	8 (3%)	0	100	100
All	All	1706/1782 (96%)	1658 (97%)	47 (3%)	1 (0%)	51	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	62	THR

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	235/243 (97%)	235 (100%)	0	100	100
1	B	233/243 (96%)	231 (99%)	2 (1%)	78	78
1	C	233/243 (96%)	231 (99%)	2 (1%)	78	78
1	D	233/243 (96%)	232 (100%)	1 (0%)	91	91
1	E	234/243 (96%)	232 (99%)	2 (1%)	78	78
1	F	235/243 (97%)	235 (100%)	0	100	100
All	All	1403/1458 (96%)	1396 (100%)	7 (0%)	88	89

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	83	HIS
1	D	214	SER
1	E	190	LYS
1	E	64	LEU
1	C	8	ILE

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

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

Of 17 ligands modelled in this entry, 17 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 [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	289/297 (97%)	-0.40	3 (1%) 82 84	5, 15, 39, 93	0
1	B	287/297 (96%)	-0.36	2 (0%) 87 89	8, 18, 43, 84	0
1	C	286/297 (96%)	-0.19	3 (1%) 82 84	8, 21, 45, 96	0
1	D	287/297 (96%)	-0.28	0 100 100	7, 18, 41, 73	0
1	E	288/297 (96%)	-0.21	2 (0%) 87 89	7, 20, 45, 81	0
1	F	289/297 (97%)	-0.28	4 (1%) 75 77	7, 16, 40, 81	0
All	All	1726/1782 (96%)	-0.29	14 (0%) 86 87	5, 18, 43, 96	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	61	ASN	3.8
1	E	61	ASN	3.5
1	F	63	LYS	3.4
1	E	62	THR	3.2
1	A	60	GLU	3.1

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	MN	B	303	1/1	0.92	0.08	35,35,35,35	0
2	MN	B	301	1/1	0.98	0.07	15,15,15,15	0
2	MN	A	303	1/1	0.98	0.09	27,27,27,27	0
2	MN	D	301	1/1	0.98	0.07	14,14,14,14	0
2	MN	C	301	1/1	0.99	0.04	17,17,17,17	0
2	MN	C	302	1/1	0.99	0.04	21,21,21,21	0
2	MN	B	302	1/1	0.99	0.04	20,20,20,20	0
2	MN	D	302	1/1	0.99	0.04	20,20,20,20	0
2	MN	F	301	1/1	0.99	0.05	13,13,13,13	0
3	CA	B	304	1/1	0.99	0.16	28,28,28,28	0
3	CA	C	303	1/1	0.99	0.07	26,26,26,26	0
2	MN	E	302	1/1	1.00	0.04	17,17,17,17	0
2	MN	A	301	1/1	1.00	0.03	17,17,17,17	0
2	MN	F	302	1/1	1.00	0.03	17,17,17,17	0
3	CA	A	304	1/1	1.00	0.11	26,26,26,26	0
2	MN	A	302	1/1	1.00	0.04	12,12,12,12	0
2	MN	E	301	1/1	1.00	0.04	21,21,21,21	0

6.5 Other polymers [i](#)

There are no such residues in this entry.