wwPDB X-ray Structure Validation Summary Report (i)

Apr 27, 2023 – 07:42 pm BST

PDB ID	:	8BKF
Title	:	Structure of E. coli Class 2 L-asparaginase EcAIII, mutant M200T (crystal M200T#o)
Authors	:	Sciuk, A.; Ruszkowski, M.; Jaskolski, M.; Loch, J.I.
Deposited on	:	2022-11-09
Resolution	:	1.22 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:

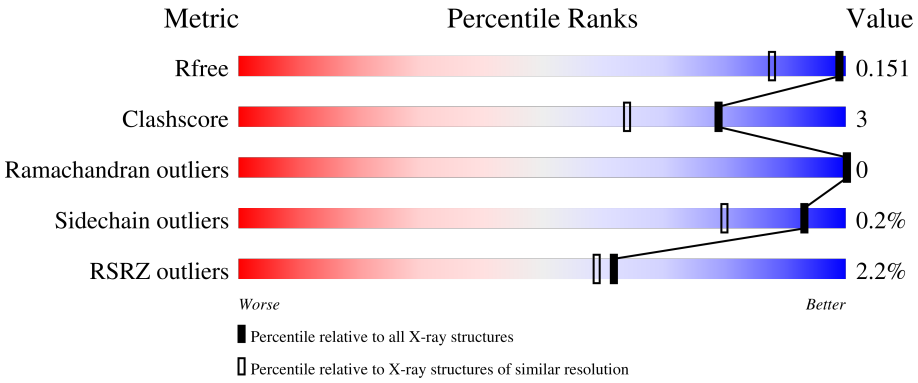
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.32.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.32.2

1 Overall quality at a glance (i)

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

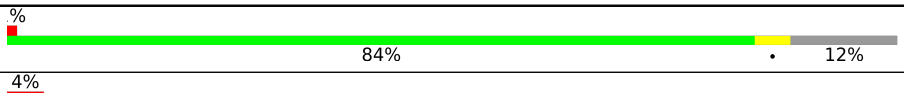
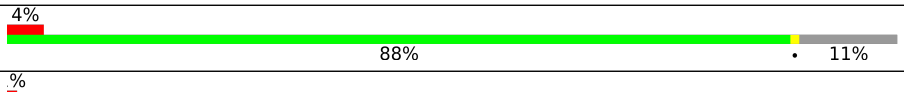
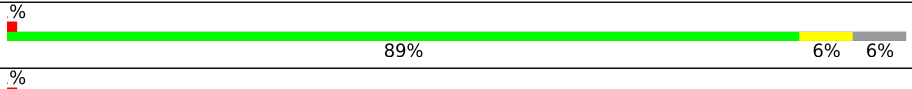
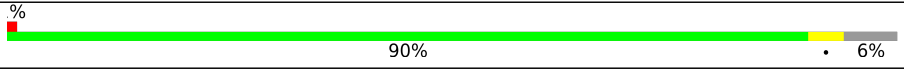
The reported resolution of this entry is 1.22 Å.

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	1232 (1.24-1.20)
Clashscore	141614	1294 (1.24-1.20)
Ramachandran outliers	138981	1251 (1.24-1.20)
Sidechain outliers	138945	1250 (1.24-1.20)
RSRZ outliers	127900	1209 (1.24-1.20)

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	AAA	178	
1	CCC	178	
2	BBB	143	
2	DDD	143	

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 5078 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 Isoaspartyl peptidase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	157	Total	C	N	O	S	0	9	0
			1242	772	218	241	11			
1	CCC	158	Total	C	N	O	S	0	5	0
			1213	756	214	233	10			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP P37595
CCC	1	MET	-	initiating methionine	UNP P37595

- Molecule 2 is a protein called Isoaspartyl peptidase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	BBB	135	Total	C	N	O	S	0	6	0
			995	623	166	200	6			
2	DDD	135	Total	C	N	O	S	0	9	0
			1026	645	170	205	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	200	THR	MET	engineered mutation	UNP P37595
DDD	200	THR	MET	engineered mutation	UNP P37595

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	AAA	1	Total	Na	0	0
			1	1		
3	CCC	1	Total	Na	0	0
			1	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	BBB	1	Total Cl 1 1	0	0
4	DDD	3	Total Cl 3 3	0	0

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	BBB	1	Total Mg 1 1	0	0

- Molecule 6 is water.

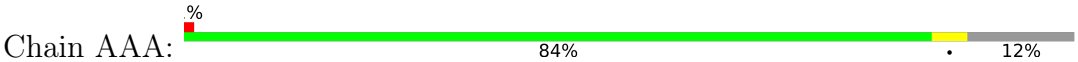
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	AAA	188	Total O 188 188	0	0
6	BBB	126	Total O 126 126	0	0
6	CCC	174	Total O 174 174	0	0
6	DDD	106	Total O 107 107	0	1

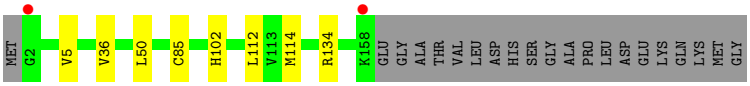
3 Residue-property plots (i)

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: Isoaspartyl peptidase subunit alpha

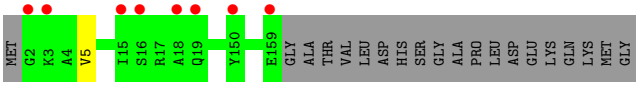
Chain AAA:





• Molecule 1: Isoaspartyl peptidase subunit alpha

Chain CCC:

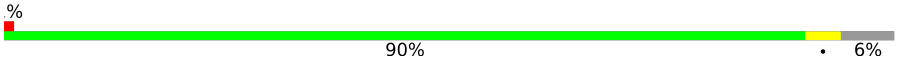


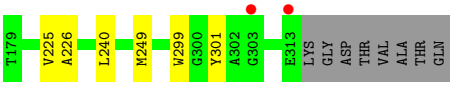
• Molecule 2: Isoaspartyl peptidase subunit beta

Chain BBB:

• Molecule 2: Isoaspartyl peptidase subunit beta

Chain DDD:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	49.85Å 75.89Å 147.76Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	73.89 – 1.22 67.50 – 1.22	Depositor EDS
% Data completeness (in resolution range)	99.9 (73.89-1.22) 99.9 (67.50-1.22)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.43 (at 1.22Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.128 , 0.150 0.129 , 0.151	Depositor DCC
R_{free} test set	1018 reflections (0.61%)	wwPDB-VP
Wilson B-factor (Å^2)	13.2	Xtriage
Anisotropy	0.266	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 44.3	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	0.98	EDS
Total number of atoms	5078	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 26.22 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.7423e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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

Bond lengths and bond angles in the following residue types are not validated in this section: MG, CL, NA

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	AAA	0.59	0/1260	0.65	0/1697
1	CCC	0.60	0/1231	0.66	0/1659
2	BBB	0.56	0/1015	0.66	0/1385
2	DDD	0.56	0/1047	0.66	0/1431
All	All	0.58	0/4553	0.66	0/6172

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	AAA	1242	0	1240	7	0
1	CCC	1213	0	1219	1	0
2	BBB	995	0	969	14	0
2	DDD	1026	0	996	17	0
3	AAA	1	0	0	0	0
3	CCC	1	0	0	0	0
4	BBB	1	0	0	0	0
4	DDD	3	0	0	0	0
5	BBB	1	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	AAA	188	0	0	1	0
6	BBB	126	0	0	0	0
6	CCC	174	0	0	0	0
6	DDD	107	0	0	0	0
All	All	5078	0	4424	25	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.

The worst 5 of 25 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)
2:BBB:240[B]:LEU:CD2	2:DDD:240[B]:LEU:HD21	1.26	1.66
2:BBB:240[B]:LEU:HD21	2:DDD:240[B]:LEU:CD2	1.31	1.55
2:BBB:240[B]:LEU:CD2	2:DDD:240[B]:LEU:CD2	2.08	0.96
2:DDD:299[A]:TRP:HZ3	2:DDD:301[A]:TYR:CE2	1.94	0.86
2:BBB:240[B]:LEU:HD21	2:DDD:240[B]:LEU:CG	2.11	0.80

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	AAA	165/178 (93%)	163 (99%)	2 (1%)	0	100	100
1	CCC	161/178 (90%)	159 (99%)	2 (1%)	0	100	100
2	BBB	139/143 (97%)	133 (96%)	6 (4%)	0	100	100
2	DDD	142/143 (99%)	134 (94%)	8 (6%)	0	100	100
All	All	607/642 (94%)	589 (97%)	18 (3%)	0	100	100

There are no Ramachandran outliers to report.

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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	AAA	130/136 (96%)	130 (100%)	0	100 100
1	CCC	126/136 (93%)	126 (100%)	0	100 100
2	BBB	99/99 (100%)	98 (99%)	1 (1%)	76 47
2	DDD	102/99 (103%)	102 (100%)	0	100 100
All	All	457/470 (97%)	456 (100%)	1 (0%)	93 80

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	BBB	230	THR

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

Of 7 ligands modelled in this entry, 7 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

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	AAA	157/178 (88%)	-0.42	2 (1%)	77	75	10, 14, 28, 50	0
1	CCC	158/178 (88%)	-0.31	8 (5%)	28	26	11, 16, 34, 56	0
2	BBB	135/143 (94%)	-0.45	1 (0%)	87	86	10, 14, 24, 49	0
2	DDD	135/143 (94%)	-0.61	2 (1%)	73	71	10, 14, 25, 53	0
All	All	585/642 (91%)	-0.44	13 (2%)	62	59	10, 15, 28, 56	0

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	CCC	2	GLY	5.9
1	AAA	158	LYS	3.8
1	AAA	2	GLY	3.3
1	CCC	159	GLU	3.3
2	DDD	313	GLU	3.3

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

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, 95^{th} 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
4	CL	DDD	402	1/1	0.99	0.06	33,33,33,33	0
4	CL	DDD	403	1/1	0.99	0.11	31,31,31,31	0
5	MG	BBB	402	1/1	0.99	0.25	38,38,38,38	0
4	CL	DDD	401	1/1	1.00	0.03	18,18,18,18	0
3	NA	AAA	201	1/1	1.00	0.03	12,12,12,12	0
3	NA	CCC	201	1/1	1.00	0.03	13,13,13,13	0
4	CL	BBB	401	1/1	1.00	0.03	18,18,18,18	0

6.5 Other polymers (i)

There are no such residues in this entry.