



# wwPDB X-ray Structure Validation Summary Report ⓘ

Apr 27, 2023 – 07:36 pm BST

PDB ID : 8C0I  
Title : Structure of E. coli Class 2 L-asparaginase EcAIII, mutant M200L (acyl-enzyme intermediate)  
Authors : Sciuk, A.; Jaskolski, M.; Loch, J.I.  
Deposited on : 2022-12-16  
Resolution : 1.90 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](mailto: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  
Mogul : 1.8.4, CSD as541be (2020)  
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

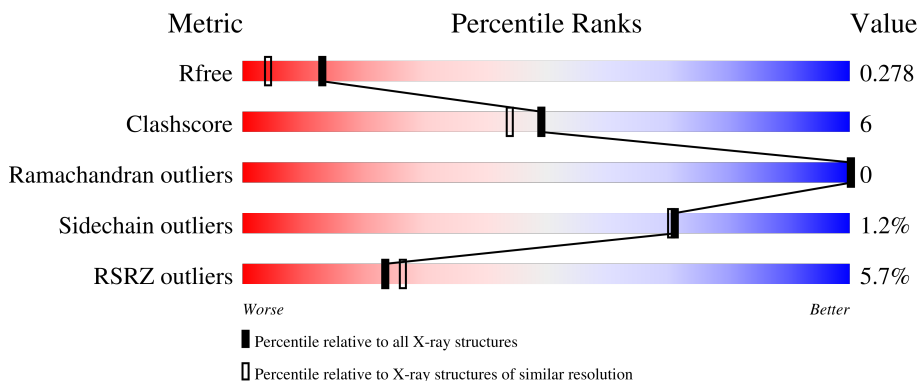
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.90 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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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  $\geq 3$ , 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	AAA	178	 5% 72% 15% 12%
1	CCC	178	 8% 76% 9% 15%
2	BBB	143	 2% 83% 12% 6%
3	DDD	143	 4% 78% 16% 6%

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 4432 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
			Total	C	N	O	S			
1	AAA	156	1169	729	207	223	10	0	1	0
1	CCC	152	1143	715	200	218	10	0	1	0

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
			Total	C	N	O	S			
2	BBB	135	958	597	162	193	6	0	1	0

There is a discrepancy between the modelled and reference sequences:

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

- Molecule 3 is a protein called Isoaspartyl peptidase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	DDD	135	972	608	163	195	6	0	2	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
DDD	179	AEI	THR	modified residue	UNP P37595

*Continued on next page...*

*Continued from previous page...*

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

- Molecule 4 is SODIUM ION (three-letter code: NA) (formula: Na).

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

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

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	AAA	2	Total Mg 2 2	0	0

- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	DDD	1	Total Cl 1 1	0	0

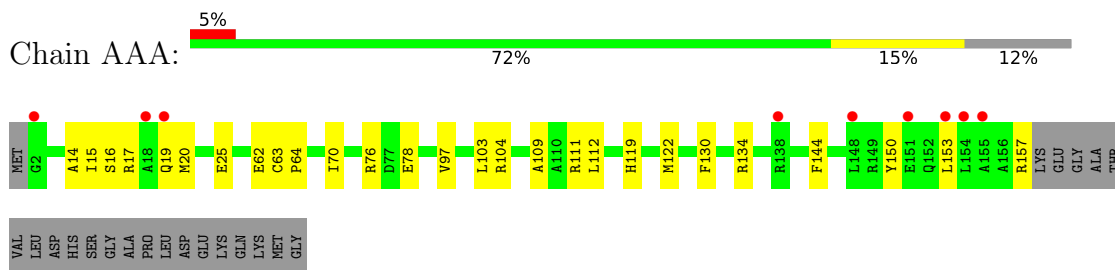
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	AAA	59	Total O 59 59	0	0
7	BBB	40	Total O 40 40	0	0
7	CCC	48	Total O 48 48	0	0
7	DDD	38	Total O 38 38	0	0

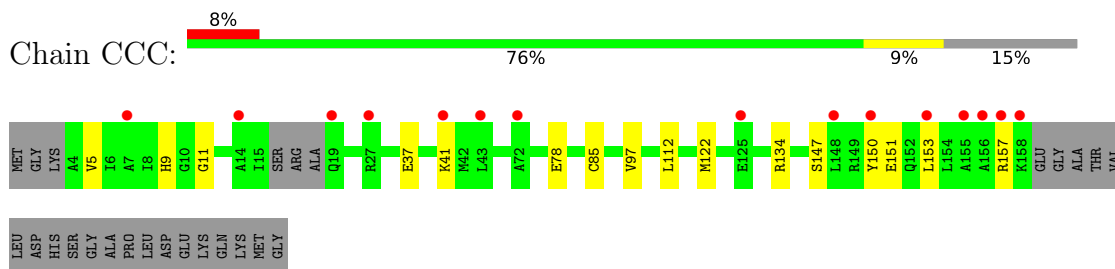
### 3 Residue-property plots [i](#)

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.

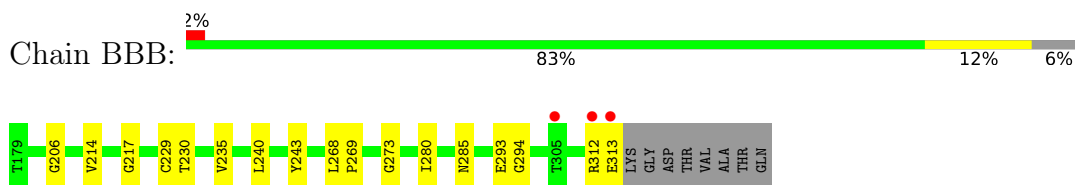
- Molecule 1: Isoaspartyl peptidase subunit alpha



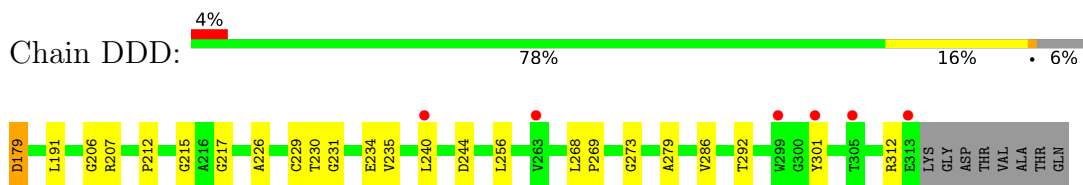
- Molecule 1: Isoaspartyl peptidase subunit alpha



- Molecule 2: Isoaspartyl peptidase subunit beta



- Molecule 3: Isoaspartyl peptidase subunit beta



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.71Å 74.85Å 146.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.30 – 1.90 19.30 – 1.90	Depositor EDS
% Data completeness (in resolution range)	97.1 (19.30-1.90) 97.2 (19.30-1.90)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.17 (at 1.90Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.232 , 0.275 0.235 , 0.278	Depositor DCC
$R_{free}$ test set	1022 reflections (2.39%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	19.4	Xtrriage
Anisotropy	0.123	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.41 , 52.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.42$ , $\langle L^2 \rangle = 0.24$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	4432	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	23.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 11.78% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, AEI, 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	0.79	0/1186	0.86	0/1599
1	CCC	0.72	0/1159	0.84	0/1562
2	BBB	0.75	0/973	0.91	0/1325
3	DDD	0.71	0/975	0.89	0/1328
All	All	0.75	0/4293	0.87	0/5814

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 [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	AAA	1169	0	1178	20	0
1	CCC	1143	0	1151	11	0
2	BBB	958	0	933	14	0
3	DDD	972	0	954	21	0
4	AAA	1	0	0	0	0
4	CCC	1	0	0	0	0
5	AAA	2	0	0	0	0
6	DDD	1	0	0	0	0
7	AAA	59	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	BBB	40	0	0	0	0
7	CCC	48	0	0	0	0
7	DDD	38	0	0	0	0
All	All	4432	0	4216	50	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 50 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:LEU:HD11	3:DDD:215:GLY:HA3	1.66	0.78
2:BBB:243:TYR:CG	3:DDD:240[B]:LEU:HD12	2.23	0.72
1:AAA:16:SER:H	1:AAA:19:GLN:HE21	1.41	0.67
1:AAA:16:SER:H	1:AAA:19:GLN:NE2	1.98	0.60
2:BBB:243:TYR:CB	3:DDD:240[B]:LEU:HD12	2.31	0.60

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	155/178 (87%)	149 (96%)	6 (4%)	0	100	100
1	CCC	149/178 (84%)	144 (97%)	5 (3%)	0	100	100
2	BBB	134/143 (94%)	129 (96%)	5 (4%)	0	100	100
3	DDD	135/143 (94%)	129 (96%)	6 (4%)	0	100	100
All	All	573/642 (89%)	551 (96%)	22 (4%)	0	100	100

There are no Ramachandran outliers to report.



### 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	AAA	120/136 (88%)	117 (98%)	3 (2%)	47	41
1	CCC	118/136 (87%)	118 (100%)	0	100	100
2	BBB	94/99 (95%)	93 (99%)	1 (1%)	73	73
3	DDD	94/98 (96%)	93 (99%)	1 (1%)	73	73
All	All	426/469 (91%)	421 (99%)	5 (1%)	71	70

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

Mol	Chain	Res	Type
1	AAA	20	MET
1	AAA	76	ARG
1	AAA	111	ARG
2	BBB	230	THR
3	DDD	230	THR

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](#)

1 non-standard protein/DNA/RNA residue is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	AEI	DDD	179	3	11,14,15	1.38	2 (18%)	14,18,20	1.97	4 (28%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	AEI	DDD	179	3	-	4/16/18/20	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	DDD	179	AEI	OG1-CD	3.35	1.43	1.34
3	DDD	179	AEI	OG1-CB	-2.00	1.43	1.46

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	DDD	179	AEI	CZ-CE2-CD	-4.95	102.38	112.85
3	DDD	179	AEI	OT1-CH2-CZ	-2.31	113.97	122.14
3	DDD	179	AEI	OG1-CD-CE2	2.28	115.66	111.46
3	DDD	179	AEI	OE1-CD-CE2	-2.23	119.80	124.73

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	DDD	179	AEI	O-C-CA-CB
3	DDD	179	AEI	OT2-CH2-CZ-NH1
3	DDD	179	AEI	OT1-CH2-CZ-NH1
3	DDD	179	AEI	OG1-CD-CE2-CZ

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	DDD	179	AEI	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 5 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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	AAA	156/178 (87%)	0.51	9 (5%) 23 25	10, 19, 41, 47	0
1	CCC	152/178 (85%)	0.67	15 (9%) 7 8	13, 24, 44, 58	0
2	BBB	135/143 (94%)	0.38	3 (2%) 62 64	11, 19, 28, 49	0
3	DDD	134/143 (93%)	0.47	6 (4%) 33 36	13, 22, 35, 51	0
All	All	577/642 (89%)	0.51	33 (5%) 23 26	10, 21, 38, 58	0

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	CCC	155	ALA	4.1
2	BBB	313	GLU	3.8
3	DDD	313	GLU	3.5
1	CCC	148	LEU	3.5
3	DDD	305	THR	3.4

### 6.2 Non-standard residues in protein, DNA, RNA chains [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	AEI	DDD	179	15/16	0.87	0.14	21,26,30,30	0

### 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	NA	CCC	201	1/1	0.96	0.06	17,17,17,17	0
5	MG	AAA	202	1/1	0.96	0.07	24,24,24,24	0
6	CL	DDD	401	1/1	0.98	0.06	23,23,23,23	0
5	MG	AAA	203	1/1	0.99	0.10	20,20,20,20	0
4	NA	AAA	201	1/1	0.99	0.05	9,9,9,9	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.