Full wwPDB X-ray Structure Validation Report ⓘ

Oct 14, 2021 – 11:05 am BST

PDB ID	:	7P9C
Title	:	Escherichia coli type II L-asparaginase
Authors	:	Maggi, M.; Scotti, C.
Deposited on	:	2021-07-27
Resolution	:	1.60 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 following versions of software and data (see references ⓘ) were used in the production of this report:

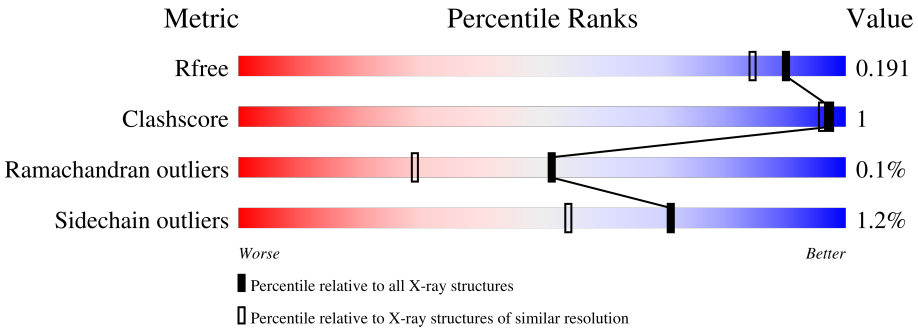
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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%

Mol	Chain	Length	Quality of chain
1	A	333	93%
1	B	333	90% 7%
1	C	333	93% 5%
1	D	333	89% 5% 7%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 20121 atoms, of which 9365 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 L-asparaginase 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	321	Total	C	H	N	O	S	0	0	0
			4728	1485	2347	405	483	8			
1	B	311	Total	C	H	N	O	S	0	0	0
			4616	1449	2297	395	467	8			
1	C	326	Total	C	H	N	O	S	14	0	0
			4840	1515	2416	415	486	8			
1	D	311	Total	C	H	N	O	S	0	0	0
			4627	1454	2305	395	465	8			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP P00805
A	-5	HIS	-	expression tag	UNP P00805
A	-4	HIS	-	expression tag	UNP P00805
A	-3	HIS	-	expression tag	UNP P00805
A	-2	HIS	-	expression tag	UNP P00805
A	-1	HIS	-	expression tag	UNP P00805
A	0	HIS	-	expression tag	UNP P00805
B	-6	MET	-	initiating methionine	UNP P00805
B	-5	HIS	-	expression tag	UNP P00805
B	-4	HIS	-	expression tag	UNP P00805
B	-3	HIS	-	expression tag	UNP P00805
B	-2	HIS	-	expression tag	UNP P00805
B	-1	HIS	-	expression tag	UNP P00805
B	0	HIS	-	expression tag	UNP P00805
C	-6	MET	-	initiating methionine	UNP P00805
C	-5	HIS	-	expression tag	UNP P00805
C	-4	HIS	-	expression tag	UNP P00805
C	-3	HIS	-	expression tag	UNP P00805
C	-2	HIS	-	expression tag	UNP P00805
C	-1	HIS	-	expression tag	UNP P00805
C	0	HIS	-	expression tag	UNP P00805

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	-6	MET	-	initiating methionine	UNP P00805
D	-5	HIS	-	expression tag	UNP P00805
D	-4	HIS	-	expression tag	UNP P00805
D	-3	HIS	-	expression tag	UNP P00805
D	-2	HIS	-	expression tag	UNP P00805
D	-1	HIS	-	expression tag	UNP P00805
D	0	HIS	-	expression tag	UNP P00805

- Molecule 2 is water.

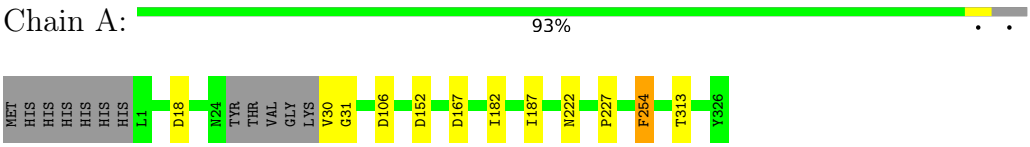
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	355	Total O 355 355	0	0
2	B	372	Total O 372 372	0	0
2	C	320	Total O 320 320	0	0
2	D	263	Total O 263 263	0	0

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. 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. 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.

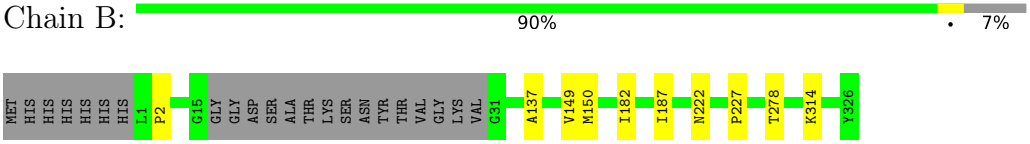
• Molecule 1: L-asparaginase 2

Chain A:



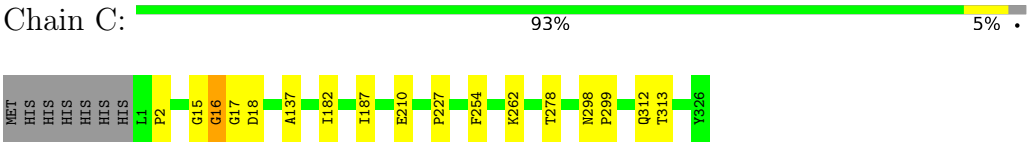
• Molecule 1: L-asparaginase 2

Chain B:



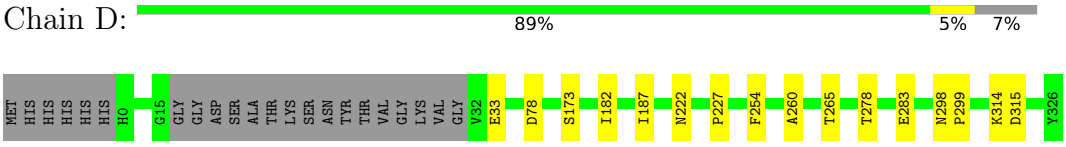
• Molecule 1: L-asparaginase 2

Chain C:



• Molecule 1: L-asparaginase 2

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	152.24Å 62.62Å 143.62Å 90.00° 118.10° 90.00°	Depositor
Resolution (Å)	39.22 – 1.60 39.22 – 1.15	Depositor EDS
% Data completeness (in resolution range)	98.8 (39.22-1.60) 97.2 (39.22-1.15)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.65 (at 1.15Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158, PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.164 , 0.190 0.164 , 0.191	Depositor DCC
R_{free} test set	20991 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	5.2	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	20121	wwPDB-VP
Average B, all atoms (Å^2)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.60% 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	0.53	1/2416 (0.0%)	0.72	3/3294 (0.1%)
1	B	0.53	1/2354 (0.0%)	0.71	4/3210 (0.1%)
1	C	0.55	4/2461 (0.2%)	0.72	6/3353 (0.2%)
1	D	0.42	0/2357	0.65	4/3212 (0.1%)
All	All	0.51	6/9588 (0.1%)	0.70	17/13069 (0.1%)

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	C	0	1

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	17	GLY	C-N	9.78	1.56	1.34
1	C	16	GLY	C-O	-6.89	1.12	1.23
1	A	254	PHE	C-N	-5.60	1.21	1.34
1	B	314	LYS	C-N	-5.54	1.21	1.34
1	C	18	ASP	C-N	-5.51	1.21	1.34
1	C	313	THR	C-N	5.50	1.46	1.34

All (17) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	17	GLY	O-C-N	10.75	139.90	122.70
1	C	17	GLY	CA-C-N	-8.06	99.48	117.20
1	C	16	GLY	C-N-CA	-7.92	105.67	122.30
1	A	167	ASP	CB-CG-OD1	7.85	125.37	118.30
1	B	314	LYS	O-C-N	-7.69	110.40	122.70
1	C	17	GLY	C-N-CA	-6.98	104.24	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	313	THR	C-N-CA	6.18	137.14	121.70
1	B	314	LYS	C-N-CA	6.14	137.05	121.70
1	D	173	SER	N-CA-C	-6.03	94.73	111.00
1	D	315	ASP	O-C-N	5.89	132.28	121.10
1	B	314	LYS	CA-C-N	5.86	130.08	117.20
1	D	278	THR	N-CA-C	-5.61	95.86	111.00
1	C	312	GLN	O-C-N	5.45	131.41	122.70
1	A	18	ASP	CB-CG-OD1	5.23	123.00	118.30
1	C	278	THR	N-CA-C	-5.19	96.99	111.00
1	D	78	ASP	CB-CG-OD2	5.11	122.90	118.30
1	B	278	THR	N-CA-C	-5.11	97.22	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	15	GLY	Mainchain

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	2381	2347	2348	4	0
1	B	2319	2297	2299	4	0
1	C	2424	2416	2419	4	0
1	D	2322	2305	2313	5	0
2	A	355	0	0	0	0
2	B	372	0	0	0	0
2	C	320	0	0	0	0
2	D	263	0	0	0	0
All	All	10756	9365	9379	15	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 1.

All (15) 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:283:GLU:HG2	1:D:283:GLU:O	2.04	0.57
1:A:30:VAL:HG23	1:A:31:GLY:N	2.22	0.54
1:A:30:VAL:HG23	1:A:31:GLY:H	1.73	0.53
1:B:182:ILE:HG12	1:B:187:ILE:HG12	1.97	0.45
1:D:298:ASN:HB2	1:D:299:PRO:CD	2.45	0.45
1:A:182:ILE:HG12	1:A:187:ILE:HG12	1.98	0.45
1:A:227:PRO:HB3	1:B:227:PRO:HB3	1.99	0.43
1:B:2:PRO:HG2	1:B:137:ALA:HB1	2.01	0.43
1:C:298:ASN:HB2	1:C:299:PRO:CD	2.49	0.43
1:D:182:ILE:HG12	1:D:187:ILE:HG12	2.01	0.43
1:C:182:ILE:HG12	1:C:187:ILE:HG12	2.00	0.42
1:D:260:ALA:HB1	1:D:265:THR:HB	2.01	0.42
1:C:2:PRO:HG2	1:C:137:ALA:HB1	2.01	0.42
1:C:227:PRO:HB3	1:D:227:PRO:HB3	2.02	0.41
1:B:149:VAL:C	1:B:150:MET:HG2	2.42	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	317/333 (95%)	310 (98%)	7 (2%)	0	100	100
1	B	307/333 (92%)	304 (99%)	3 (1%)	0	100	100
1	C	324/333 (97%)	313 (97%)	10 (3%)	1 (0%)	41	21
1	D	307/333 (92%)	302 (98%)	5 (2%)	0	100	100
All	All	1255/1332 (94%)	1229 (98%)	25 (2%)	1 (0%)	51	29

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	16	GLY

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	258/273 (94%)	254 (98%)	4 (2%)	62	41
1	B	253/273 (93%)	252 (100%)	1 (0%)	91	84
1	C	264/273 (97%)	261 (99%)	3 (1%)	73	57
1	D	253/273 (93%)	249 (98%)	4 (2%)	62	41
All	All	1028/1092 (94%)	1016 (99%)	12 (1%)	71	54

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	106	ASP
1	A	152	ASP
1	A	222	ASN
1	A	254	PHE
1	B	222	ASN
1	C	210	GLU
1	C	254	PHE
1	C	262	LYS
1	D	33	GLU
1	D	222	ASN
1	D	254	PHE
1	D	314	LYS

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 (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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.