



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 09:24 pm BST

PDB ID : 2JK0  
Title : Structural and functional insights into Erwinia carotovora L- asparaginase  
Authors : Papageorgiou, A.C.; Posypanova, G.A.; Andersson, C.S.; Sokolov, N.N.; Krasotkina, J.  
Deposited on : 2008-05-23  
Resolution : 2.50 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 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	:	2.11
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)	:	2.11

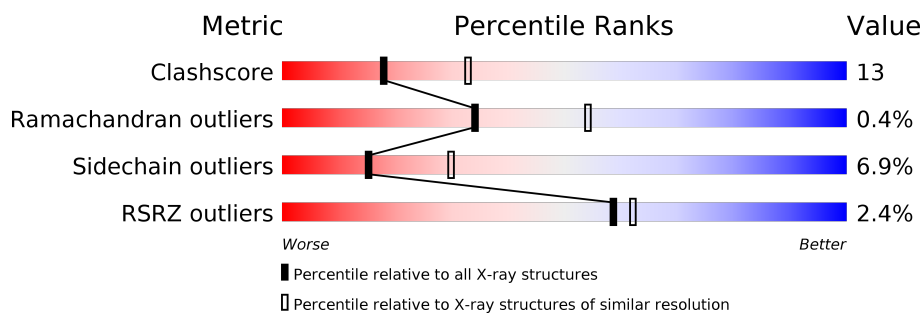
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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))
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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 on 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	A	325	<div> <div>3%</div> <div>77%</div> <div>21%</div> <div>.</div> </div>
1	B	325	<div> <div>2%</div> <div>77%</div> <div>18%</div> <div>..</div> </div>
1	C	325	<div> <div>2%</div> <div>71%</div> <div>22%</div> <div>..</div> </div>
1	D	325	<div> <div>%</div> <div>75%</div> <div>21%</div> <div>..</div> </div>
1	E	325	<div> <div>3%</div> <div>70%</div> <div>22%</div> <div>..</div> </div>
1	F	325	<div> <div>3%</div> <div>68%</div> <div>26%</div> <div>..</div> </div>
1	G	325	<div> <div>2%</div> <div>73%</div> <div>20%</div> <div>5%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	H	325	<div><div></div><div>4%</div><div>71%</div><div>22%</div><div></div><div></div></div>

## 2 Entry composition

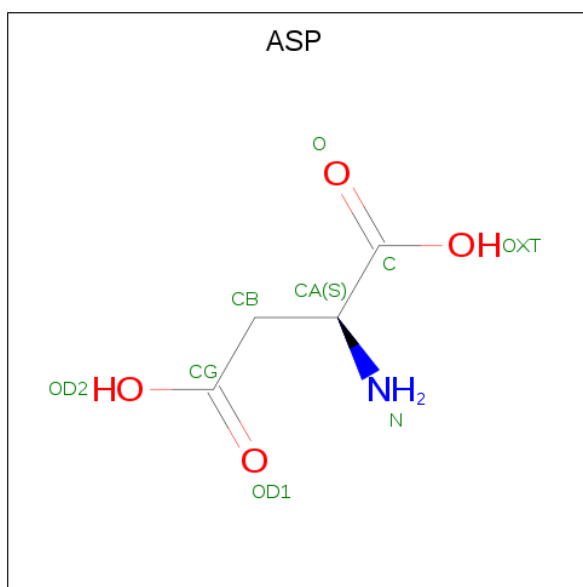
There are 3 unique types of molecules in this entry. The entry contains 19499 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 L-ASPARAGINASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	325	Total	C	N	O	S	0	0	0
			2386	1502	413	465	6			
1	B	314	Total	C	N	O	S	0	0	0
			2331	1472	401	452	6			
1	C	315	Total	C	N	O	S	0	0	0
			2335	1473	402	454	6			
1	D	320	Total	C	N	O	S	0	0	0
			2364	1489	408	461	6			
1	E	312	Total	C	N	O	S	0	0	0
			2320	1464	399	451	6			
1	F	317	Total	C	N	O	S	0	0	0
			2350	1481	405	458	6			
1	G	309	Total	C	N	O	S	0	0	0
			2309	1459	396	448	6			
1	H	313	Total	C	N	O	S	0	0	0
			2324	1466	400	452	6			

- Molecule 2 is ASPARTIC ACID (three-letter code: ASP) (formula: C<sub>4</sub>H<sub>7</sub>NO<sub>4</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			9	4	1	4		
2	B	1	Total	C	N	O	0	0
			9	4	1	4		
2	C	1	Total	C	N	O	0	0
			9	4	1	4		
2	D	1	Total	C	N	O	0	0
			9	4	1	4		
2	E	1	Total	C	N	O	0	0
			9	4	1	4		
2	F	1	Total	C	N	O	0	0
			9	4	1	4		
2	G	1	Total	C	N	O	0	0
			9	4	1	4		
2	H	1	Total	C	N	O	0	0
			9	4	1	4		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	117	Total	O	0	0
			117	117		
3	B	86	Total	O	0	0
			86	86		
3	C	107	Total	O	0	0
			107	107		
3	D	108	Total	O	0	0
			108	108		

*Continued on next page...*

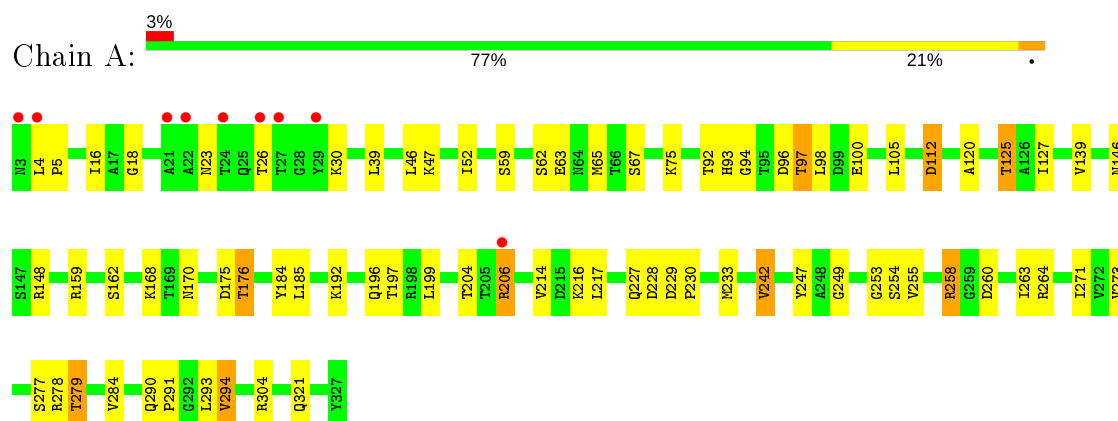
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	E	84	Total 84	O 84	0	0
3	F	68	Total 68	O 68	0	0
3	G	65	Total 65	O 65	0	0
3	H	73	Total 73	O 73	0	0

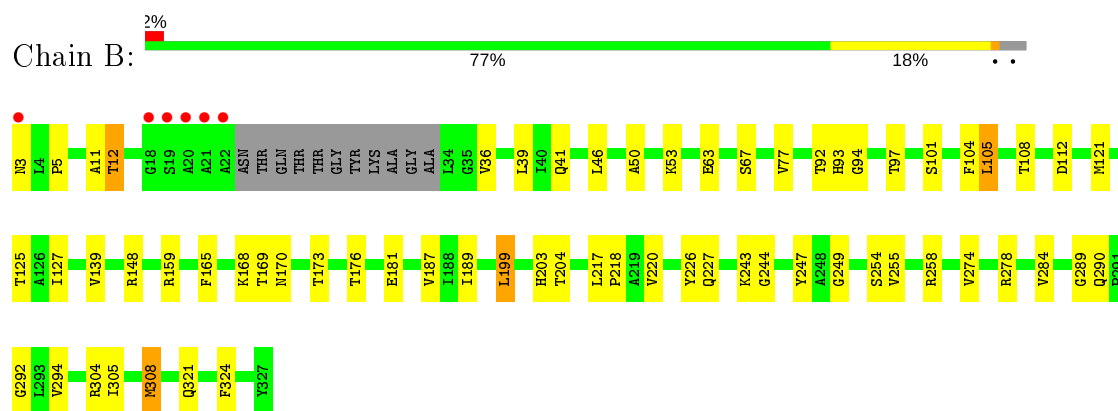
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

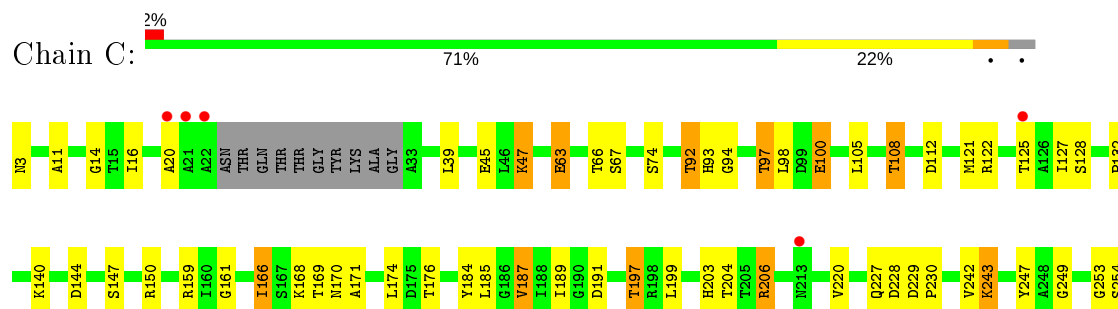
#### • Molecule 1: L-ASPARAGINASE



#### • Molecule 1: L-ASPARAGINASE

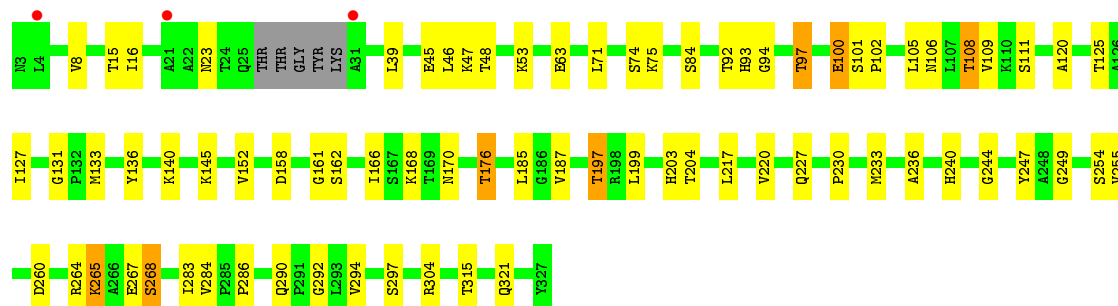
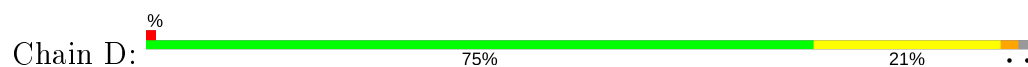


#### • Molecule 1: L-ASPARAGINASE

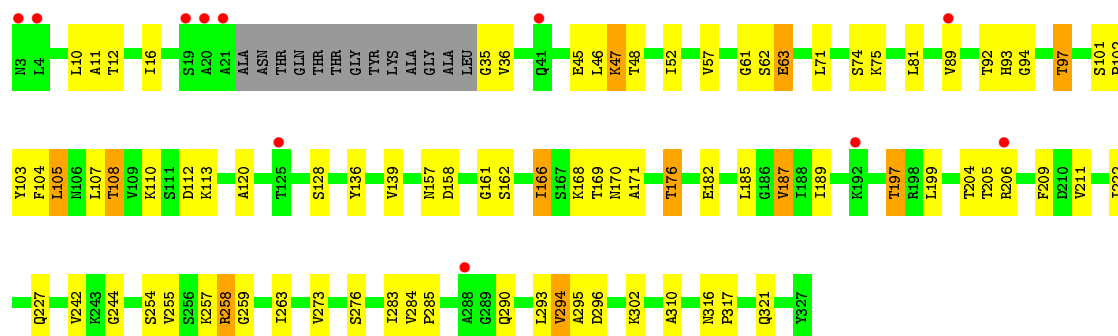




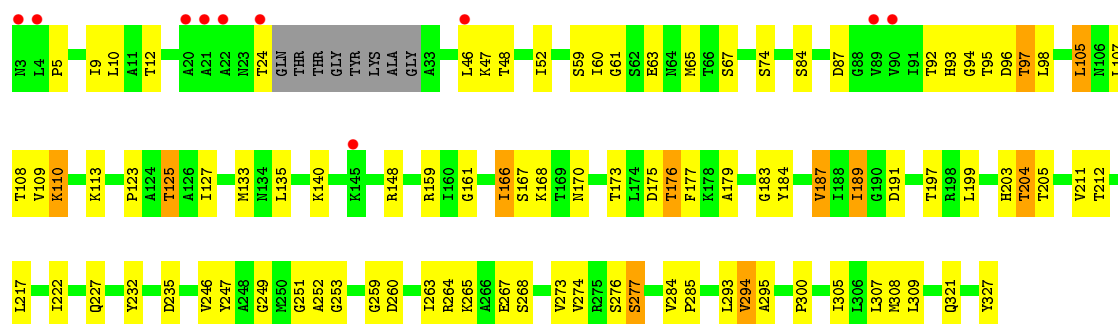
• Molecule 1: L-ASPARAGINASE



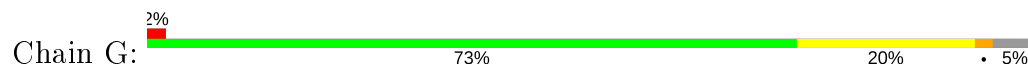
• Molecule 1: L-ASPARAGINASE



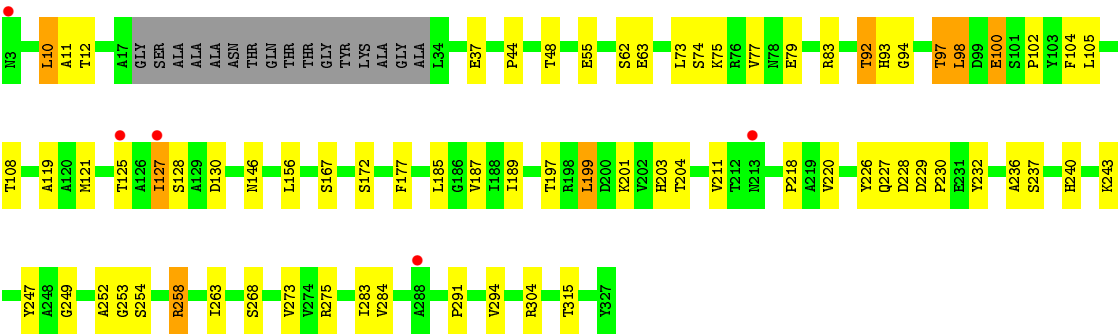
• Molecule 1: L-ASPARAGINASE



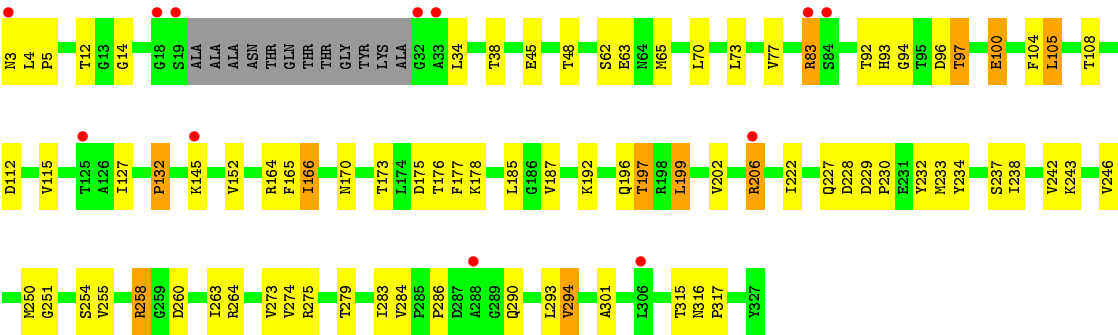
• Molecule 1: L-ASPARAGINASE







● Molecule 1: L-ASPARAGINASE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	73.65Å 135.63Å 250.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.50 19.92 – 2.50	Depositor EDS
% Data completeness (in resolution range)	96.1 (20.00-2.50) 96.0 (19.92-2.50)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.69 (at 2.50Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.188 , 0.266 0.191 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	44.5	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 42.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	19499	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	34.0	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 32.96 % 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 8.6345e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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](#)

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	0.47	0/2421	0.67	0/3288
1	B	0.48	0/2365	0.65	0/3211
1	C	0.46	0/2369	0.64	0/3217
1	D	0.46	0/2398	0.64	0/3257
1	E	0.45	0/2354	0.62	0/3196
1	F	0.49	1/2384 (0.0%)	0.62	0/3238
1	G	0.44	0/2343	0.61	0/3181
1	H	0.45	0/2358	0.61	0/3201
All	All	0.46	1/18992 (0.0%)	0.64	0/25789

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	24	THR	C-O	11.38	1.45	1.23

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	A	2386	0	2420	77	0
1	B	2331	0	2386	63	0
1	C	2335	0	2387	79	0
1	D	2364	0	2410	72	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2320	0	2375	81	0
1	F	2350	0	2400	80	0
1	G	2309	0	2371	58	0
1	H	2324	0	2375	67	0
2	A	9	0	3	1	0
2	B	9	0	3	0	0
2	C	9	0	3	2	0
2	D	9	0	3	1	0
2	E	9	0	3	0	0
2	F	9	0	3	0	0
2	G	9	0	3	1	0
2	H	9	0	3	1	0
3	A	117	0	0	10	0
3	B	86	0	0	3	0
3	C	107	0	0	4	0
3	D	108	0	0	6	0
3	E	84	0	0	6	0
3	F	68	0	0	4	0
3	G	65	0	0	5	0
3	H	73	0	0	6	0
All	All	19499	0	19148	511	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 13.

The worst 5 of 511 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:187:VAL:HB	3:D:2071:HOH:O	1.34	1.27
1:C:185:LEU:HA	1:C:197:THR:HG22	1.32	1.08
1:C:168:LYS:HG3	1:C:176:THR:HG21	1.30	1.06
1:E:263:ILE:HD12	1:E:273:VAL:HG11	1.41	0.99
1:B:97:THR:HG21	1:D:227:GLN:HE21	1.25	0.98

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	323/325 (99%)	308 (95%)	14 (4%)	1 (0%)	41	61
1	B	310/325 (95%)	294 (95%)	15 (5%)	1 (0%)	41	61
1	C	311/325 (96%)	299 (96%)	11 (4%)	1 (0%)	41	61
1	D	316/325 (97%)	303 (96%)	13 (4%)	0	100	100
1	E	308/325 (95%)	294 (96%)	12 (4%)	2 (1%)	25	43
1	F	313/325 (96%)	298 (95%)	13 (4%)	2 (1%)	25	43
1	G	305/325 (94%)	292 (96%)	13 (4%)	0	100	100
1	H	309/325 (95%)	290 (94%)	17 (6%)	2 (1%)	25	43
All	All	2495/2600 (96%)	2378 (95%)	108 (4%)	9 (0%)	34	54

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	5	PRO
1	F	204	THR
1	H	34	LEU
1	A	204	THR
1	C	20	ALA

### 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	253/260 (97%)	235 (93%)	18 (7%)	14	28

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	251/260 (96%)	239 (95%)	12 (5%)	25	48
1	C	251/260 (96%)	228 (91%)	23 (9%)	9	18
1	D	253/260 (97%)	239 (94%)	14 (6%)	21	41
1	E	251/260 (96%)	238 (95%)	13 (5%)	23	44
1	F	253/260 (97%)	230 (91%)	23 (9%)	9	18
1	G	252/260 (97%)	236 (94%)	16 (6%)	18	34
1	H	251/260 (96%)	231 (92%)	20 (8%)	12	23
All	All	2015/2080 (97%)	1876 (93%)	139 (7%)	15	30

5 of 139 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	197	THR
1	E	257	LYS
1	H	127	ILE
1	D	265	LYS
1	E	105	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 47 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	203	HIS
1	E	93	HIS
1	H	146	ASN
1	D	240	HIS
1	E	146	ASN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

8 ligands are modelled in this entry.

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 ⓘ

### 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<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	A	325/325 (100%)	-0.29	9 (2%) 53 56	22, 30, 36, 40	0
1	B	314/325 (96%)	-0.35	6 (1%) 66 69	21, 29, 36, 38	0
1	C	315/325 (96%)	-0.24	5 (1%) 72 74	22, 30, 36, 40	0
1	D	320/325 (98%)	-0.33	3 (0%) 84 86	20, 28, 36, 40	0
1	E	312/325 (96%)	-0.11	11 (3%) 44 47	31, 38, 51, 59	0
1	F	317/325 (97%)	-0.01	10 (3%) 47 51	31, 39, 46, 49	0
1	G	309/325 (95%)	-0.17	5 (1%) 72 74	31, 38, 51, 56	0
1	H	313/325 (96%)	-0.01	12 (3%) 40 43	31, 39, 51, 56	0
All	All	2525/2600 (97%)	-0.19	61 (2%) 59 62	20, 34, 46, 59	0

The worst 5 of 61 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	22	ALA	6.4
1	F	21	ALA	6.4
1	F	20	ALA	5.6
1	C	21	ALA	5.5
1	B	21	ALA	5.4

### 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 carbohydrates 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( $\text{\AA}^2$ )	Q<0.9
2	ASP	E	1328	9/9	0.86	0.24	60,62,62,63	0
2	ASP	H	1328	9/9	0.91	0.18	55,56,57,57	0
2	ASP	F	1328	9/9	0.92	0.21	53,54,56,56	0
2	ASP	G	1328	9/9	0.93	0.18	48,49,50,50	0
2	ASP	D	1328	9/9	0.94	0.17	53,54,55,55	0
2	ASP	C	1328	9/9	0.94	0.11	45,45,48,49	0
2	ASP	B	1328	9/9	0.95	0.13	35,38,38,39	0
2	ASP	A	1328	9/9	0.97	0.09	34,36,36,36	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.