



wwPDB X-ray Structure Validation Summary Report

Feb 28, 2014 – 10:13 PM GMT

PDB ID : 3LYC
Title : Crystal structure of Putative pectinase (YP_001304412.1) from Parabacteroides distasonis ATCC 8503 at 2.30 Å resolution
Authors : Joint Center for Structural Genomics (JCSG)
Deposited on : 2010-02-26
Resolution : 2.30 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

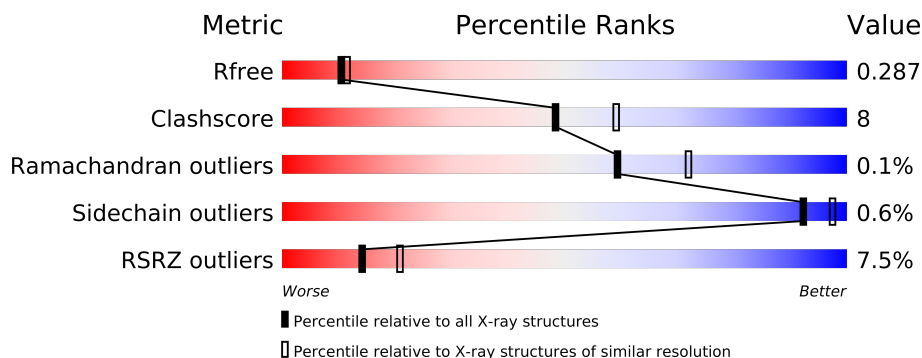
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.30 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}	66092	2929 (2.30-2.30)
Clashscore	79885	3679 (2.30-2.30)
Ramachandran outliers	78287	3642 (2.30-2.30)
Sidechain outliers	78261	3641 (2.30-2.30)
RSRZ outliers	66119	2930 (2.30-2.30)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	241	
1	B	241	
1	C	241	
1	D	241	
1	E	241	
1	F	241	
1	G	241	
1	H	241	
1	I	241	
1	J	241	
1	K	241	
1	L	241	
1	M	241	
1	N	241	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	O	241	
1	P	241	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 29713 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Putative pectinase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	241	Total	C	N	O	S	Se	0	7	0
			1831	1142	307	371	6	5			
1	B	237	Total	C	N	O	S	Se	0	3	0
			1785	1122	293	361	4	5			
1	C	241	Total	C	N	O	S	Se	0	8	0
			1831	1148	305	368	4	6			
1	D	240	Total	C	N	O	S	Se	0	4	0
			1780	1116	299	356	4	5			
1	E	241	Total	C	N	O	S	Se	0	5	0
			1836	1147	310	369	5	5			
1	F	239	Total	C	N	O	S	Se	0	4	0
			1782	1117	294	362	4	5			
1	G	241	Total	C	N	O	S	Se	0	5	0
			1811	1134	303	365	4	5			
1	H	241	Total	C	N	O	S	Se	0	6	0
			1822	1143	297	373	4	5			
1	I	241	Total	C	N	O	S	Se	0	5	0
			1811	1134	297	369	6	5			
1	J	238	Total	C	N	O	S	Se	0	7	0
			1809	1134	302	364	4	5			
1	K	241	Total	C	N	O	S	Se	0	5	0
			1816	1132	305	369	4	6			
1	L	235	Total	C	N	O	S	Se	0	6	0
			1746	1096	287	354	4	5			
1	M	241	Total	C	N	O	S	Se	0	10	0
			1862	1169	309	373	6	5			
1	N	238	Total	C	N	O	S	Se	0	7	0
			1824	1141	306	368	4	5			
1	O	240	Total	C	N	O	S	Se	0	7	0
			1836	1149	306	372	4	5			
1	P	232	Total	C	N	O	S	Se	0	1	0
			1681	1053	276	343	4	5			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	leader sequence	UNP A6LGH6
B	0	GLY	-	leader sequence	UNP A6LGH6
C	0	GLY	-	leader sequence	UNP A6LGH6
D	0	GLY	-	leader sequence	UNP A6LGH6
E	0	GLY	-	leader sequence	UNP A6LGH6
F	0	GLY	-	leader sequence	UNP A6LGH6
G	0	GLY	-	leader sequence	UNP A6LGH6
H	0	GLY	-	leader sequence	UNP A6LGH6
I	0	GLY	-	leader sequence	UNP A6LGH6
J	0	GLY	-	leader sequence	UNP A6LGH6
K	0	GLY	-	leader sequence	UNP A6LGH6
L	0	GLY	-	leader sequence	UNP A6LGH6
M	0	GLY	-	leader sequence	UNP A6LGH6
N	0	GLY	-	leader sequence	UNP A6LGH6
O	0	GLY	-	leader sequence	UNP A6LGH6
P	0	GLY	-	leader sequence	UNP A6LGH6

- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	90	Total O 90 90	0	0
2	B	38	Total O 38 38	0	0
2	C	50	Total O 50 50	0	0
2	D	21	Total O 21 21	0	0
2	E	97	Total O 97 97	0	0
2	F	46	Total O 46 46	0	0
2	G	64	Total O 64 64	0	0
2	H	24	Total O 24 24	0	0
2	I	84	Total O 84 84	0	0
2	J	50	Total O 50 50	0	0
2	K	55	Total O 55 55	0	0

Continued on next page...

Continued from previous page...

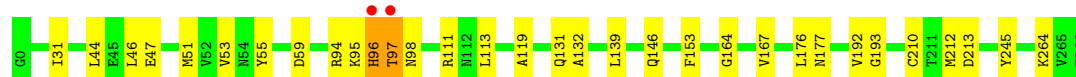
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	L	17	Total 17	O 17	0	0
2	M	88	Total 88	O 88	0	0
2	N	39	Total 39	O 39	0	0
2	O	63	Total 63	O 63	0	0
2	P	24	Total 24	O 24	0	0

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 errors 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.

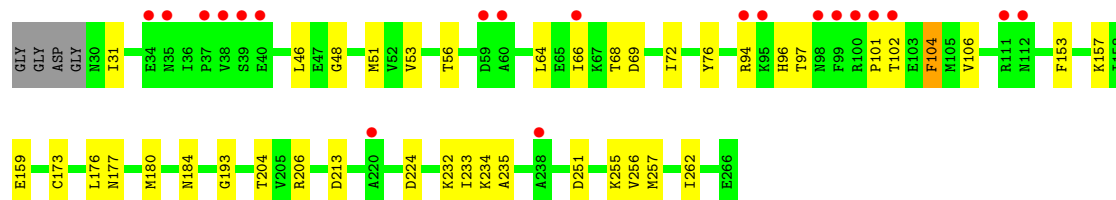
- Molecule 1: Putative pectinase

Chain A: 



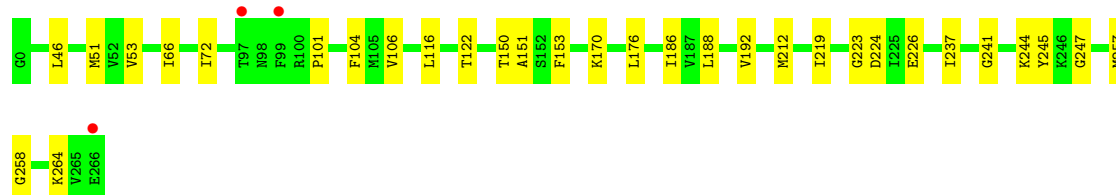
- Molecule 1: Putative pectinase

Chain B: 



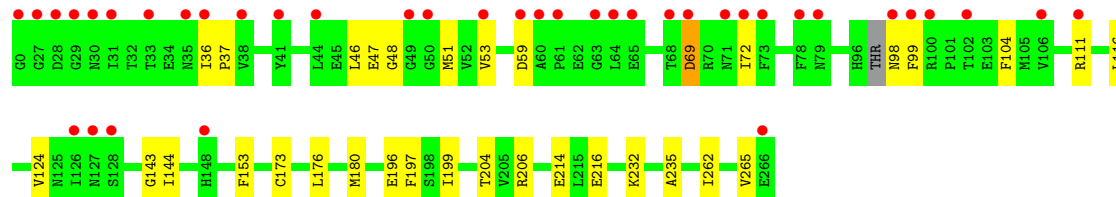
- Molecule 1: Putative pectinase

Chain C: 



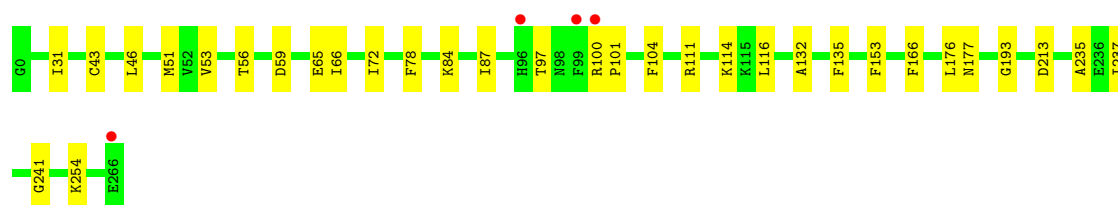
- Molecule 1: Putative pectinase

Chain D: 



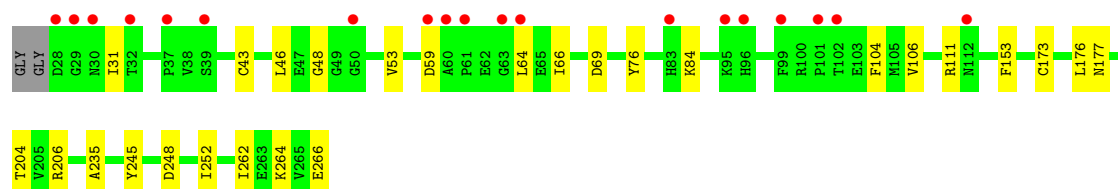
- Molecule 1: Putative pectinase

Chain E: 



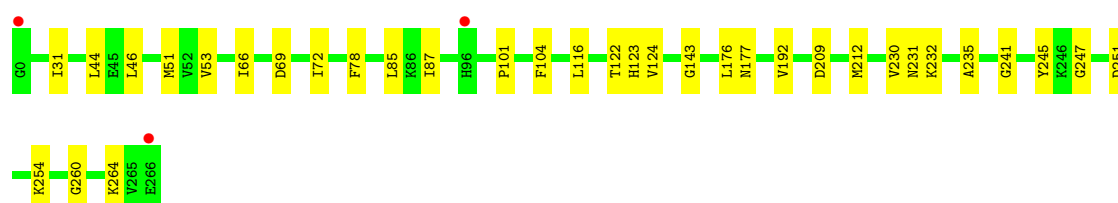
- Molecule 1: Putative pectinase

Chain F:



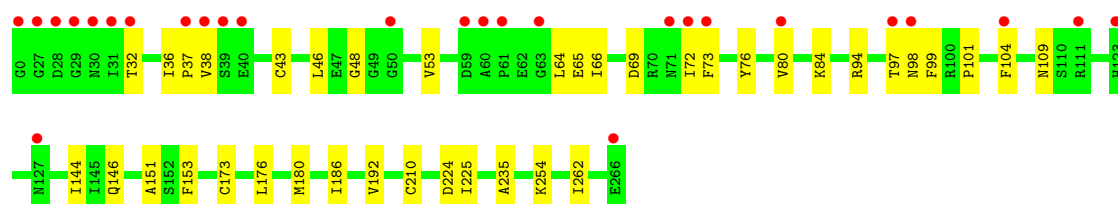
- Molecule 1: Putative pectinase

Chain G:



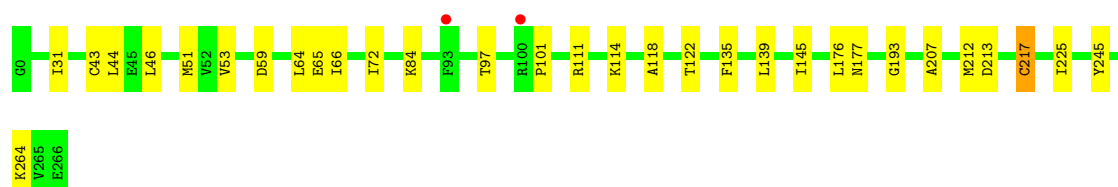
- Molecule 1: Putative pectinase

Chain H:



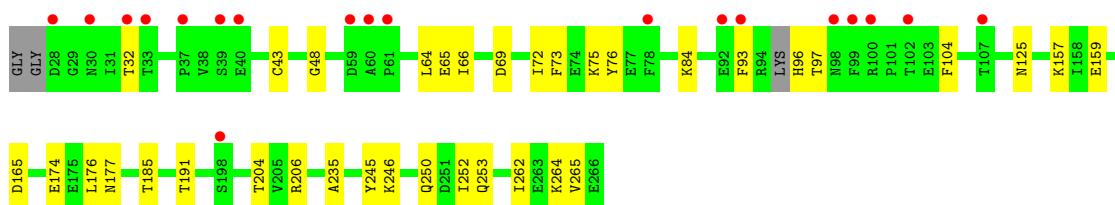
- Molecule 1: Putative pectinase

Chain I:



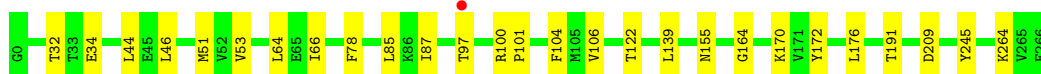
- Molecule 1: Putative pectinase

Chain J:



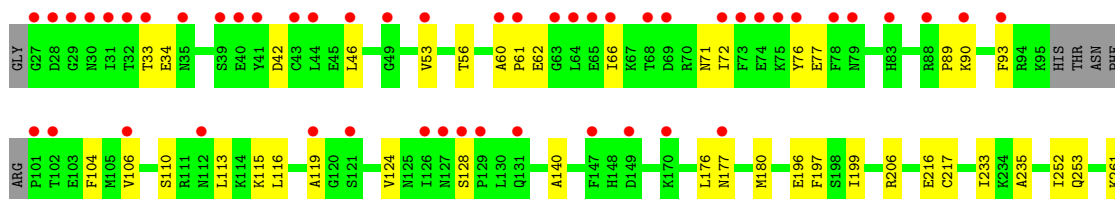
- Molecule 1: Putative pectinase

Chain K:



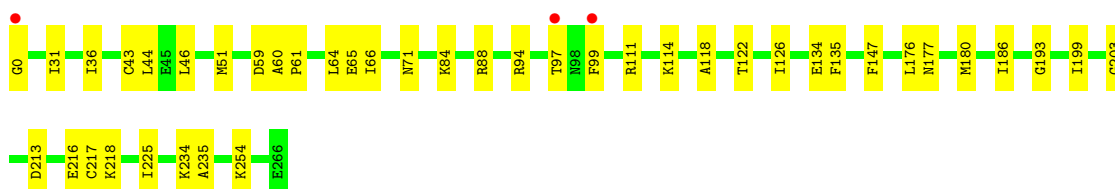
- Molecule 1: Putative pectinase

Chain L:



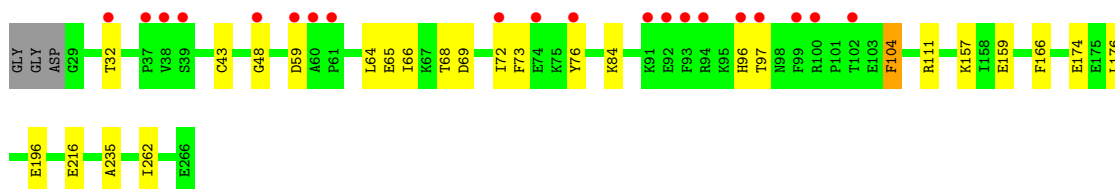
- Molecule 1: Putative pectinase

Chain M:



- Molecule 1: Putative pectinase

Chain N:



- Molecule 1: Putative pectinase

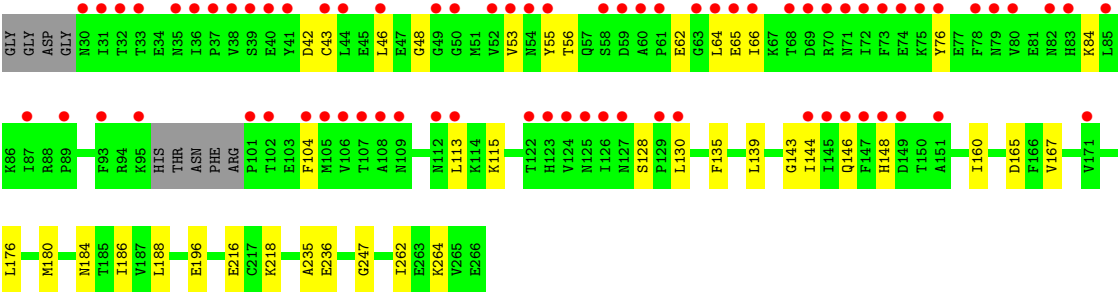
Chain O:





● Molecule 1: Putative pectinase

Chain P:



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	104.12Å 104.61Å 393.59Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.51 – 2.30 49.53 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.5 (49.51-2.30) 97.5 (49.53-2.30)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.78 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.244 , 0.280 0.250 , 0.287	Depositor DCC
R_{free} test set	9322 reflections (5.02%)	DCC
Wilson B-factor (Å ²)	30.4	Xtriage
Anisotropy	0.333	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 27.6	EDS
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	61 of 186629 reflections (0.033%)	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	29713	wwPDB-VP
Average B, all atoms (Å ²)	38.0	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 62.92 % 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 1.0466e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

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	0.64	0/1872	0.46	0/2510
1	B	0.56	0/1815	0.45	0/2435
1	C	0.58	0/1876	0.44	0/2515
1	D	0.51	0/1812	0.43	0/2431
1	E	0.66	0/1873	0.46	0/2511
1	F	0.54	0/1815	0.44	0/2439
1	G	0.58	0/1847	0.45	0/2477
1	H	0.53	0/1862	0.43	0/2500
1	I	0.66	2/1847 (0.1%)	0.45	0/2479
1	J	0.52	0/1848	0.43	0/2479
1	K	0.58	0/1851	0.45	0/2483
1	L	0.52	0/1783	0.43	0/2393
1	M	0.66	0/1915	0.46	0/2565
1	N	0.56	0/1868	0.43	0/2508
1	O	0.60	0/1879	0.45	0/2515
1	P	0.50	0/1702	0.43	0/2291
All	All	0.58	2/29465 (0.0%)	0.44	0/39531

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	217[A]	CYS	CB-SG	-5.15	1.73	1.81
1	I	217[B]	CYS	CB-SG	-5.15	1.73	1.81

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens

added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1831	0	1786	34	0
1	B	1785	0	1730	32	0
1	C	1831	0	1791	24	0
1	D	1780	0	1711	24	0
1	E	1836	0	1792	27	0
1	F	1782	0	1703	18	0
1	G	1811	0	1757	24	0
1	H	1822	0	1744	28	0
1	I	1811	0	1749	33	0
1	J	1809	0	1751	27	0
1	K	1816	0	1768	31	0
1	L	1746	0	1678	27	0
1	M	1862	0	1833	41	0
1	N	1824	0	1768	30	0
1	O	1836	0	1799	31	0
1	P	1681	0	1573	28	0
2	A	90	0	0	1	0
2	B	38	0	0	1	0
2	C	50	0	0	0	0
2	D	21	0	0	0	0
2	E	97	0	0	1	0
2	F	46	0	0	0	0
2	G	64	0	0	1	0
2	H	24	0	0	0	0
2	I	84	0	0	0	0
2	J	50	0	0	2	0
2	K	55	0	0	0	0
2	L	17	0	0	0	0
2	M	88	0	0	1	0
2	N	39	0	0	0	0
2	O	63	0	0	1	0
2	P	24	0	0	0	0
All	All	29713	0	27933	443	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 8.

The worst 5 of 443 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:K:46:LEU:HD23	1:K:51:MSE:CE	1.88	1.04
1:I:46:LEU:CD2	1:I:51:MSE:HE1	1.89	1.02
1:C:46:LEU:HD22	1:C:51:MSE:HE1	1.43	1.00
1:O:46:LEU:HB3	1:O:51:MSE:HE1	1.44	1.00
1:N:72:ILE:HD11	1:N:104:PHE:CZ	2.00	0.97

There are no symmetry-related clashes.

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	246/241 (102%)	236 (96%)	8 (3%)	2 (1%)	27	30
1	B	238/241 (99%)	224 (94%)	14 (6%)	0	100	100
1	C	247/241 (102%)	232 (94%)	15 (6%)	0	100	100
1	D	240/241 (100%)	231 (96%)	9 (4%)	0	100	100
1	E	244/241 (101%)	236 (97%)	8 (3%)	0	100	100
1	F	241/241 (100%)	231 (96%)	10 (4%)	0	100	100
1	G	244/241 (101%)	232 (95%)	12 (5%)	0	100	100
1	H	245/241 (102%)	234 (96%)	11 (4%)	0	100	100
1	I	244/241 (101%)	234 (96%)	10 (4%)	0	100	100
1	J	241/241 (100%)	231 (96%)	10 (4%)	0	100	100
1	K	244/241 (101%)	237 (97%)	7 (3%)	0	100	100
1	L	237/241 (98%)	223 (94%)	14 (6%)	0	100	100
1	M	249/241 (103%)	237 (95%)	12 (5%)	0	100	100
1	N	243/241 (101%)	231 (95%)	12 (5%)	0	100	100
1	O	243/241 (101%)	233 (96%)	10 (4%)	0	100	100
1	P	229/241 (95%)	216 (94%)	13 (6%)	0	100	100
All	All	3875/3856 (100%)	3698 (95%)	175 (4%)	2 (0%)	59	72

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	96	HIS
1	A	97	THR

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	A	194/190 (102%)	194 (100%)	0	100	100
1	B	186/190 (98%)	185 (100%)	1 (0%)	94	98
1	C	193/190 (102%)	192 (100%)	1 (0%)	94	98
1	D	181/190 (95%)	179 (99%)	2 (1%)	84	93
1	E	194/190 (102%)	193 (100%)	1 (0%)	94	98
1	F	184/190 (97%)	183 (100%)	1 (0%)	94	98
1	G	188/190 (99%)	187 (100%)	1 (0%)	94	98
1	H	188/190 (99%)	186 (99%)	2 (1%)	84	93
1	I	190/190 (100%)	190 (100%)	0	100	100
1	J	190/190 (100%)	188 (99%)	2 (1%)	84	93
1	K	192/190 (101%)	191 (100%)	1 (0%)	94	98
1	L	180/190 (95%)	178 (99%)	2 (1%)	84	93
1	M	200/190 (105%)	200 (100%)	0	100	100
1	N	194/190 (102%)	192 (99%)	2 (1%)	85	94
1	O	195/190 (103%)	195 (100%)	0	100	100
1	P	167/190 (88%)	166 (99%)	1 (1%)	92	97
All	All	3016/3040 (99%)	2999 (99%)	17 (1%)	92	97

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	104	PHE
1	H	109	ASN
1	L	104	PHE
1	G	69	ASP
1	N	104	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	253	GLN
1	H	109	ASN
1	J	83	HIS

5.3.3 RNA ⓘ

There are no RNA chains 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 carbohydrates in this entry.

5.6 Ligand geometry ⓘ

There are no ligands in this entry.

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, 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	241/241 (100%)	0.25	2 (0%) 83 90	22, 32, 44, 65	0
1	B	237/241 (98%)	0.69	20 (8%) 11 17	30, 41, 53, 72	0
1	C	241/241 (100%)	0.18	3 (1%) 75 83	24, 33, 44, 63	0
1	D	240/241 (99%)	0.93	39 (16%) 2 4	29, 42, 52, 69	0
1	E	241/241 (100%)	0.25	4 (1%) 67 76	23, 32, 44, 64	0
1	F	239/241 (99%)	0.63	19 (7%) 13 19	29, 42, 54, 68	0
1	G	241/241 (100%)	0.24	3 (1%) 75 83	24, 33, 44, 63	0
1	H	241/241 (100%)	0.81	27 (11%) 6 9	30, 42, 54, 65	0
1	I	241/241 (100%)	0.21	2 (0%) 83 90	23, 32, 44, 64	0
1	J	238/241 (98%)	0.65	19 (7%) 12 19	29, 42, 52, 64	0
1	K	241/241 (100%)	0.13	1 (0%) 90 95	23, 33, 44, 60	0
1	L	235/241 (97%)	1.11	50 (21%) 1 2	30, 42, 50, 62	0
1	M	241/241 (100%)	0.28	3 (1%) 75 83	23, 32, 44, 62	0
1	N	238/241 (98%)	0.69	20 (8%) 11 17	29, 42, 52, 64	0
1	O	240/241 (99%)	0.13	1 (0%) 90 95	24, 33, 44, 60	0
1	P	232/241 (96%)	1.42	73 (31%) 1 1	30, 42, 49, 60	0
All	All	3827/3856 (99%)	0.53	286 (7%) 14 21	22, 37, 50, 72	0

The worst 5 of 286 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	P	64	LEU	6.7
1	J	28	ASP	6.0
1	P	104	PHE	5.9
1	H	0	GLY	5.6
1	D	0	GLY	5.5

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.