



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 29, 2020 – 12:22 PM BST

PDB ID : 6JJK
Title : Crystal structure of the DegP dodecamer with a modulator
Authors : Cho, H.; Choi, Y.; Lee, H.H.; Kim, S.
Deposited on : 2019-02-26
Resolution : 3.60 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 ⓘ 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	:	2.13
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)	:	2.13

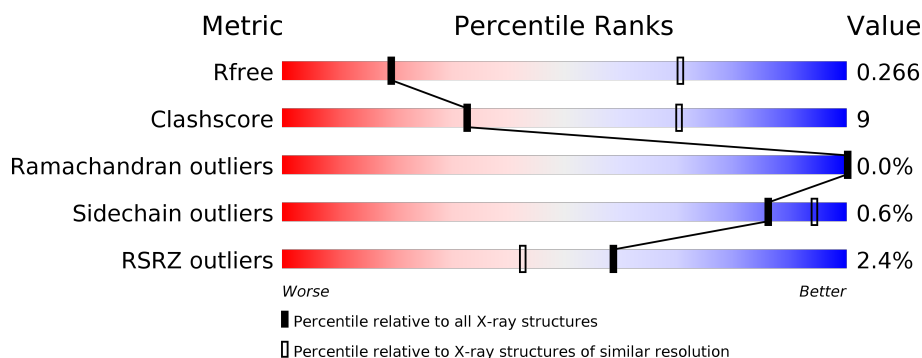
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.60 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.







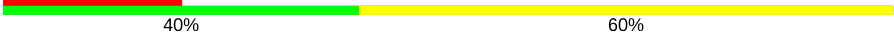
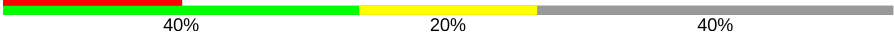

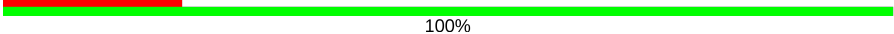

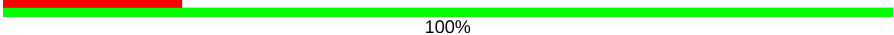


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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 ≥ 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	440	<div> <div>2%</div> <div> <div></div> <div>70%</div> <div>17%</div> <div>13%</div> </div> </div>
1	B	440	<div> <div>%</div> <div> <div></div> <div>73%</div> <div>15%</div> <div>11%</div> </div> </div>
1	C	440	<div> <div>2%</div> <div> <div></div> <div>72%</div> <div>15%</div> <div>13%</div> </div> </div>
1	D	440	<div> <div>5%</div> <div> <div></div> <div>67%</div> <div>21%</div> <div>12%</div> </div> </div>
1	E	440	<div> <div>%</div> <div> <div></div> <div>68%</div> <div>20%</div> <div>12%</div> </div> </div>
1	F	440	<div> <div>%</div> <div> <div></div> <div>66%</div> <div>20%</div> <div>14%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	5	
2	H	5	
2	I	5	
2	J	5	
2	K	5	
2	L	5	
2	M	5	
2	N	5	
2	O	5	
2	P	5	
2	Q	5	
2	R	5	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 16997 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 Periplasmic serine endoprotease DegP.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	381	Total	C	N	O	S	0	0	0
			2736	1703	482	538	13			
1	B	392	Total	C	N	O	S	0	0	0
			2820	1752	495	560	13			
1	C	382	Total	C	N	O	S	0	0	0
			2743	1709	482	539	13			
1	D	388	Total	C	N	O	S	0	0	0
			2784	1731	488	553	12			
1	E	388	Total	C	N	O	S	0	0	0
			2795	1737	491	554	13			
1	F	379	Total	C	N	O	S	0	0	0
			2732	1702	481	536	13			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	210	ALA	SER	engineered mutation	UNP P0C0V0
B	210	ALA	SER	engineered mutation	UNP P0C0V0
C	210	ALA	SER	engineered mutation	UNP P0C0V0
D	210	ALA	SER	engineered mutation	UNP P0C0V0
E	210	ALA	SER	engineered mutation	UNP P0C0V0
F	210	ALA	SER	engineered mutation	UNP P0C0V0

- Molecule 2 is a protein called CYS-TYR-TYR-LYS-ILE.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	G	5	Total	C	N	O	0	0	0
			28	18	5	5			
2	H	5	Total	C	N	O	0	0	0
			28	18	5	5			
2	I	5	Total	C	N	O	0	0	0
			35	24	5	6			

Continued on next page...

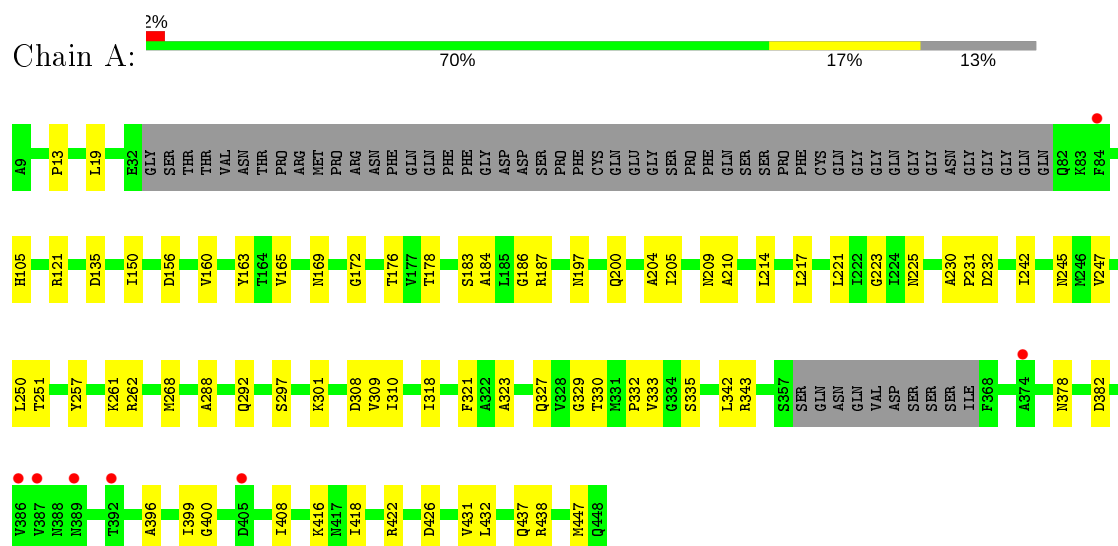
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	J	5	Total	C	N	O	S	0	0	0
			29	18	5	5	1			
2	K	5	Total	C	N	O	S	0	0	0
			36	24	5	6	1			
2	L	3	Total	C	N	O		0	0	0
			18	12	3	3				
2	M	5	Total	C	N	O	S	0	0	0
			47	33	6	7	1			
2	N	5	Total	C	N	O		0	0	0
			36	24	5	7				
2	O	5	Total	C	N	O		0	0	0
			25	15	5	5				
2	P	5	Total	C	N	O		0	0	0
			29	18	5	6				
2	Q	5	Total	C	N	O	S	0	0	0
			47	33	6	7	1			
2	R	5	Total	C	N	O		0	0	0
			29	18	5	6				

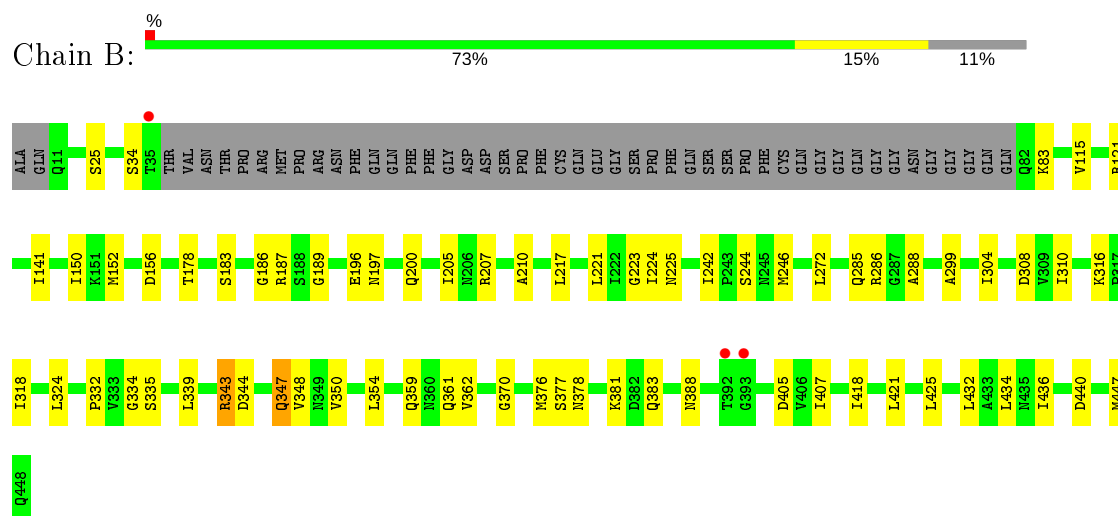
3 Residue-property plots [i](#)

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: Periplasmic serine endoprotease DegP



- Molecule 1: Periplasmic serine endoprotease DegP

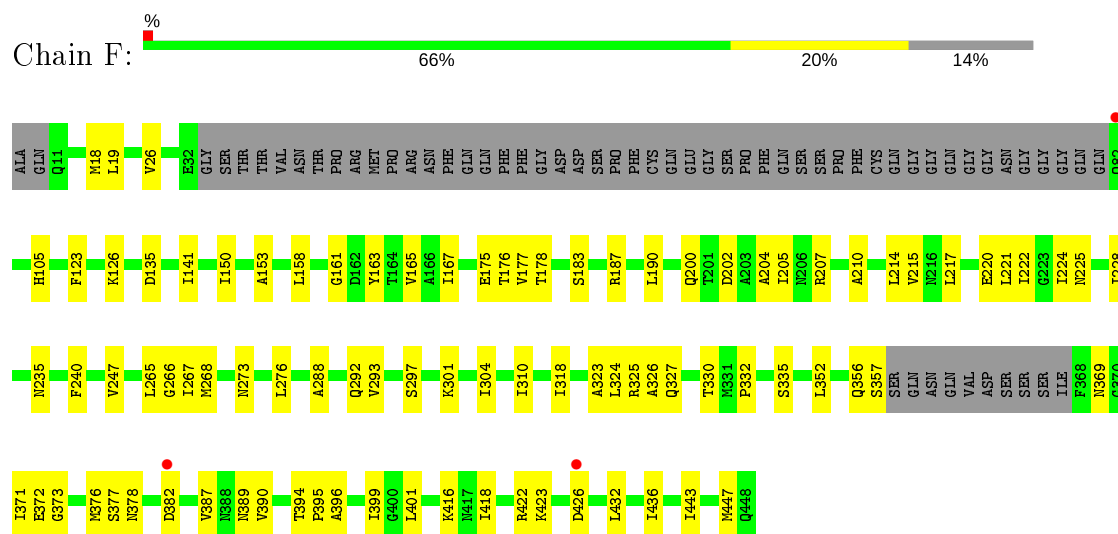


- Molecule 1: Periplasmic serine endoprotease DegP

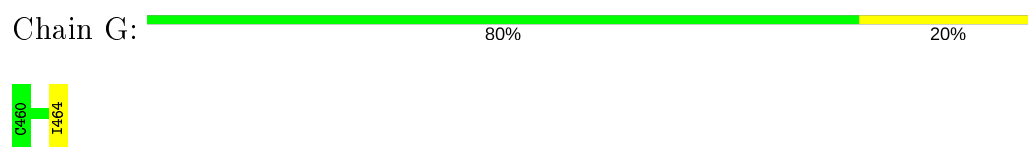




- Molecule 1: Periplasmic serine endoprotease DegP



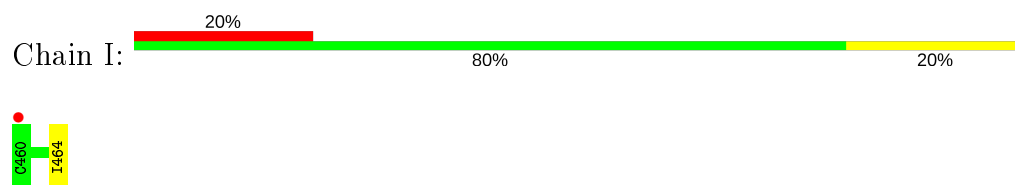
- Molecule 2: CYS-TYR-TYR-LYS-ILE



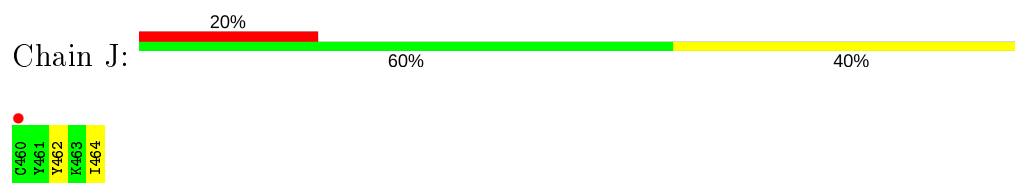
- Molecule 2: CYS-TYR-TYR-LYS-ILE



- Molecule 2: CYS-TYR-TYR-LYS-ILE



- Molecule 2: CYS-TYR-TYR-LYS-ILE

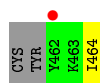


- Molecule 2: CYS-TYR-TYR-LYS-ILE

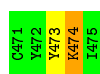




- Molecule 2: CYS-TYR-TYR-LYS-ILE



- Molecule 2: CYS-TYR-TYR-LYS-ILE



- Molecule 2: CYS-TYR-TYR-LYS-ILE



- Molecule 2: CYS-TYR-TYR-LYS-ILE

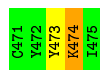


There are no outlier residues recorded for this chain.

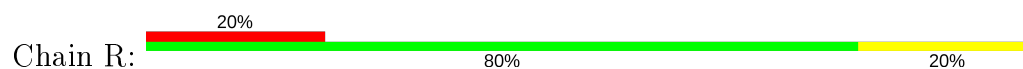
- Molecule 2: CYS-TYR-TYR-LYS-ILE



- Molecule 2: CYS-TYR-TYR-LYS-ILE



- Molecule 2: CYS-TYR-TYR-LYS-ILE



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	214.92Å 122.43Å 140.24Å 90.00° 117.72° 90.00°	Depositor
Resolution (Å)	35.48 – 3.60 35.48 – 3.60	Depositor EDS
% Data completeness (in resolution range)	99.1 (35.48-3.60) 99.1 (35.48-3.60)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.52 (at 3.56Å)	Xtriage
Refinement program	PHENIX 1.13_2998	Depositor
R, R_{free}	0.199 , 0.265 0.199 , 0.266	Depositor DCC
R_{free} test set	1843 reflections (4.96%)	wwPDB-VP
Wilson B-factor (Å ²)	126.0	Xtriage
Anisotropy	0.136	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 91.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.52$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.000 for 1/2*h-3/2*k,-1/2*h-1/2*k,-1/2*h +1/2*k-l 0.000 for 1/2*h+3/2*k,1/2*h-1/2*k,-1/2*h- 1/2*k-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	16997	wwPDB-VP
Average B, all atoms (Å ²)	136.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.74% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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	0.27	0/2759	0.48	0/3735
1	B	0.27	0/2845	0.49	0/3853
1	C	0.28	0/2766	0.48	0/3742
1	D	0.26	0/2808	0.49	0/3804
1	E	0.27	0/2819	0.48	0/3815
1	F	0.27	0/2755	0.48	0/3727
2	G	0.30	0/27	0.44	0/36
2	H	0.18	0/27	0.39	0/36
2	I	0.31	0/35	0.68	0/47
2	J	0.28	0/28	0.47	0/37
2	K	0.34	0/36	0.60	0/48
2	L	0.26	0/17	0.32	0/22
2	M	0.28	0/48	0.42	0/63
2	N	0.22	0/36	0.53	0/47
2	O	0.30	0/24	0.44	0/32
2	P	0.22	0/28	0.44	0/36
2	Q	0.33	0/48	0.46	0/63
2	R	0.22	0/28	0.44	0/36
All	All	0.27	0/17134	0.48	0/23179

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	A	2736	0	2768	43	0
1	B	2820	0	2849	45	0
1	C	2743	0	2793	41	0
1	D	2784	0	2807	57	0
1	E	2795	0	2831	58	0
1	F	2732	0	2782	62	0
2	G	28	0	18	1	0
2	H	28	0	18	2	0
2	I	35	0	25	1	0
2	J	29	0	21	3	0
2	K	36	0	28	5	0
2	L	18	0	14	3	0
2	M	47	0	46	3	0
2	N	36	0	25	0	0
2	O	25	0	9	0	0
2	P	29	0	18	0	0
2	Q	47	0	46	4	0
2	R	29	0	18	1	0
All	All	16997	0	17116	297	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 9.

The worst 5 of 297 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:334:GLY:H	1:D:354:LEU:HB2	1.43	0.82
1:C:275:GLU:HA	1:C:278:LYS:HD3	1.65	0.77
1:B:334:GLY:H	1:B:354:LEU:HB2	1.49	0.76
1:D:371:ILE:HA	1:D:395:PRO:HB2	1.69	0.74
1:E:210:ALA:HB2	2:K:464:ILE:HG13	1.69	0.74

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	375/440 (85%)	361 (96%)	14 (4%)	0	100	100
1	B	388/440 (88%)	366 (94%)	21 (5%)	1 (0%)	41	75
1	C	376/440 (86%)	359 (96%)	17 (4%)	0	100	100
1	D	382/440 (87%)	359 (94%)	23 (6%)	0	100	100
1	E	382/440 (87%)	361 (94%)	21 (6%)	0	100	100
1	F	373/440 (85%)	359 (96%)	14 (4%)	0	100	100
2	G	3/5 (60%)	3 (100%)	0	0	100	100
2	H	3/5 (60%)	3 (100%)	0	0	100	100
2	I	3/5 (60%)	2 (67%)	1 (33%)	0	100	100
2	J	3/5 (60%)	3 (100%)	0	0	100	100
2	K	3/5 (60%)	2 (67%)	1 (33%)	0	100	100
2	L	1/5 (20%)	1 (100%)	0	0	100	100
2	M	3/5 (60%)	3 (100%)	0	0	100	100
2	N	3/5 (60%)	3 (100%)	0	0	100	100
2	O	3/5 (60%)	3 (100%)	0	0	100	100
2	P	3/5 (60%)	3 (100%)	0	0	100	100
2	Q	3/5 (60%)	3 (100%)	0	0	100	100
2	R	3/5 (60%)	2 (67%)	1 (33%)	0	100	100
All	All	2310/2700 (86%)	2196 (95%)	113 (5%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	370	GLY

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	287/350 (82%)	286 (100%)	1 (0%)	92	97
1	B	300/350 (86%)	297 (99%)	3 (1%)	76	88
1	C	290/350 (83%)	289 (100%)	1 (0%)	92	97
1	D	295/350 (84%)	295 (100%)	0	100	100
1	E	298/350 (85%)	295 (99%)	3 (1%)	76	88
1	F	289/350 (83%)	289 (100%)	0	100	100
2	G	1/5 (20%)	1 (100%)	0	100	100
2	H	1/5 (20%)	1 (100%)	0	100	100
2	I	2/5 (40%)	2 (100%)	0	100	100
2	J	2/5 (40%)	2 (100%)	0	100	100
2	K	3/5 (60%)	3 (100%)	0	100	100
2	L	1/5 (20%)	1 (100%)	0	100	100
2	M	5/5 (100%)	4 (80%)	1 (20%)	1	8
2	N	2/5 (40%)	2 (100%)	0	100	100
2	P	1/5 (20%)	1 (100%)	0	100	100
2	Q	5/5 (100%)	4 (80%)	1 (20%)	1	8
2	R	1/5 (20%)	1 (100%)	0	100	100
All	All	1783/2155 (83%)	1773 (99%)	10 (1%)	86	94

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	84	PHE
1	E	14	SER
1	E	368	PHE
1	B	376	MET
1	E	207	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	437	GLN

5.3.3 RNA ⓘ

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

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	381/440 (86%)	-0.01	7 (1%) 68 53	78, 124, 172, 194	0
1	B	392/440 (89%)	-0.10	3 (0%) 86 75	86, 142, 187, 205	0
1	C	382/440 (86%)	0.07	10 (2%) 56 40	71, 122, 198, 215	0
1	D	388/440 (88%)	0.24	22 (5%) 23 14	93, 156, 210, 229	0
1	E	388/440 (88%)	0.09	5 (1%) 77 63	89, 140, 200, 224	0
1	F	379/440 (86%)	-0.02	3 (0%) 86 75	87, 130, 176, 187	0
2	G	5/5 (100%)	0.32	0 100 100	116, 136, 139, 140	0
2	H	5/5 (100%)	0.37	0 100 100	119, 138, 145, 150	0
2	I	5/5 (100%)	0.91	1 (20%) 1 0	107, 140, 159, 159	0
2	J	5/5 (100%)	0.37	1 (20%) 1 0	133, 155, 161, 172	0
2	K	5/5 (100%)	1.38	1 (20%) 1 0	124, 147, 160, 170	0
2	L	3/5 (60%)	1.30	1 (33%) 0 0	134, 134, 151, 159	0
2	M	5/5 (100%)	0.30	0 100 100	117, 130, 151, 163	0
2	N	5/5 (100%)	0.25	1 (20%) 1 0	150, 152, 161, 162	0
2	O	5/5 (100%)	-0.34	0 100 100	123, 135, 166, 173	0
2	P	5/5 (100%)	-0.10	1 (20%) 1 0	169, 171, 175, 177	0
2	Q	5/5 (100%)	-0.15	0 100 100	146, 152, 158, 159	0
2	R	5/5 (100%)	0.08	1 (20%) 1 0	140, 141, 159, 162	0
All	All	2368/2700 (87%)	0.05	57 (2%) 59 42	71, 135, 191, 229	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	373	GLY	5.1
2	K	460	CYS	4.9
1	D	391	LYS	4.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	390	VAL	4.4
1	D	362	VAL	4.2

6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.