wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 02:18 am BST

PDB ID	:	5X8L
Title	:	PD-L1 in complex with atezolizumab
Authors	:	Heo, Y.S.; Lee, H.T.
Deposited on	:	2017-03-03
Resolution	:	3.10 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

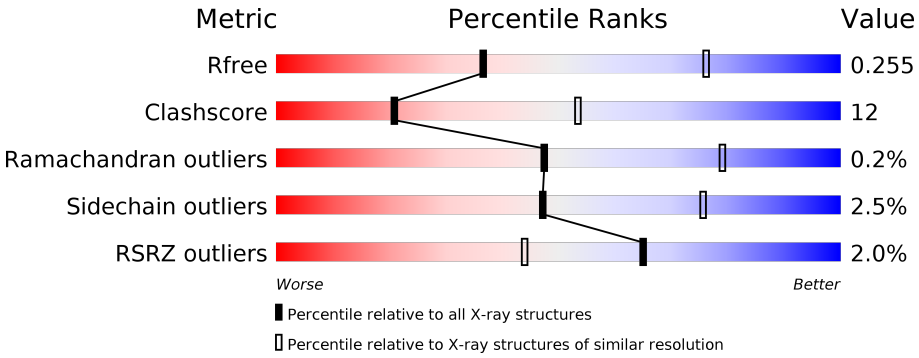
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1094 (3.10-3.10)
Clashscore	141614	1184 (3.10-3.10)
Ramachandran outliers	138981	1141 (3.10-3.10)
Sidechain outliers	138945	1141 (3.10-3.10)
RSRZ outliers	127900	1067 (3.10-3.10)

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 <=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	123	63% 26% • 8%
1	B	123	3% 73% 20% • 6%
1	C	123	69% 24% • 6%
1	D	123	76% 17% • 6%
1	E	123	76% 17% 7%
2	K	214	86% 14%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	L	214	2% 63% 28% 6%
2	M	214	3% 73% 24%
2	N	214	4% 67% 23% 5% 5%
2	O	214	6% 60% 36%
3	F	230	77% 15% 8%
3	G	230	% 72% 20% 8%
3	H	230	% 70% 20% 8%
3	J	230	3% 69% 21% 10%
3	S	230	3% 66% 25% 8%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20566 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 Programmed cell death 1 ligand 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	113	Total	C	N	O	S	0	0	0
			903	577	151	170	5			
1	B	116	Total	C	N	O	S	0	0	0
			928	593	155	175	5			
1	C	116	Total	C	N	O	S	0	0	0
			928	593	155	175	5			
1	D	116	Total	C	N	O	S	0	0	0
			928	593	155	175	5			
1	E	114	Total	C	N	O	S	0	0	0
			910	582	152	171	5			

There are 35 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	GLU	GLN	engineered mutation	UNP Q9NZQ7
A	135	HIS	-	expression tag	UNP Q9NZQ7
A	136	HIS	-	expression tag	UNP Q9NZQ7
A	137	HIS	-	expression tag	UNP Q9NZQ7
A	138	HIS	-	expression tag	UNP Q9NZQ7
A	139	HIS	-	expression tag	UNP Q9NZQ7
A	140	HIS	-	expression tag	UNP Q9NZQ7
B	47	GLU	GLN	engineered mutation	UNP Q9NZQ7
B	135	HIS	-	expression tag	UNP Q9NZQ7
B	136	HIS	-	expression tag	UNP Q9NZQ7
B	137	HIS	-	expression tag	UNP Q9NZQ7
B	138	HIS	-	expression tag	UNP Q9NZQ7
B	139	HIS	-	expression tag	UNP Q9NZQ7
B	140	HIS	-	expression tag	UNP Q9NZQ7
C	47	GLU	GLN	engineered mutation	UNP Q9NZQ7
C	135	HIS	-	expression tag	UNP Q9NZQ7
C	136	HIS	-	expression tag	UNP Q9NZQ7
C	137	HIS	-	expression tag	UNP Q9NZQ7
C	138	HIS	-	expression tag	UNP Q9NZQ7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	139	HIS	-	expression tag	UNP Q9NZQ7
C	140	HIS	-	expression tag	UNP Q9NZQ7
D	47	GLU	GLN	engineered mutation	UNP Q9NZQ7
D	135	HIS	-	expression tag	UNP Q9NZQ7
D	136	HIS	-	expression tag	UNP Q9NZQ7
D	137	HIS	-	expression tag	UNP Q9NZQ7
D	138	HIS	-	expression tag	UNP Q9NZQ7
D	139	HIS	-	expression tag	UNP Q9NZQ7
D	140	HIS	-	expression tag	UNP Q9NZQ7
E	47	GLU	GLN	engineered mutation	UNP Q9NZQ7
E	135	HIS	-	expression tag	UNP Q9NZQ7
E	136	HIS	-	expression tag	UNP Q9NZQ7
E	137	HIS	-	expression tag	UNP Q9NZQ7
E	138	HIS	-	expression tag	UNP Q9NZQ7
E	139	HIS	-	expression tag	UNP Q9NZQ7
E	140	HIS	-	expression tag	UNP Q9NZQ7

- Molecule 2 is a protein called atezolizumab light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	K	213	Total	C	N	O	S	0	0	0
			1638	1029	272	332	5			
2	L	202	Total	C	N	O	S	0	0	0
			1561	984	258	314	5			
2	M	213	Total	C	N	O	S	0	0	0
			1638	1029	272	332	5			
2	N	204	Total	C	N	O	S	0	0	0
			1574	991	260	318	5			
2	O	213	Total	C	N	O	S	0	0	0
			1638	1029	272	332	5			

- Molecule 3 is a protein called atezolizumab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	212	Total	C	N	O	S	0	0	0
			1594	1015	267	307	5			
3	G	211	Total	C	N	O	S	0	0	0
			1588	1012	266	305	5			
3	H	212	Total	C	N	O	S	0	0	0
			1592	1014	267	306	5			
3	S	211	Total	C	N	O	S	0	0	0
			1585	1007	266	307	5			

Continued on next page...

Continued from previous page...

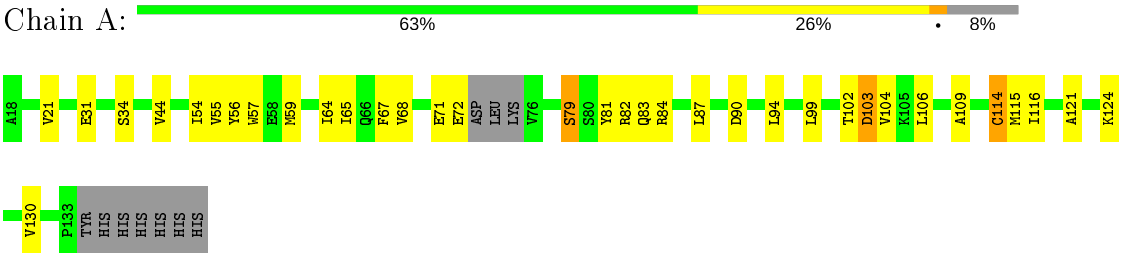
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	208	Total	C	N	O	S	0	0	0
			1561	992	262	302	5			

3 Residue-property plots (i)

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.

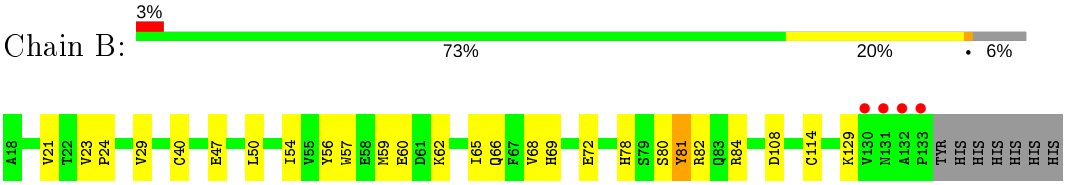
- Molecule 1: Programmed cell death 1 ligand 1

Chain A:



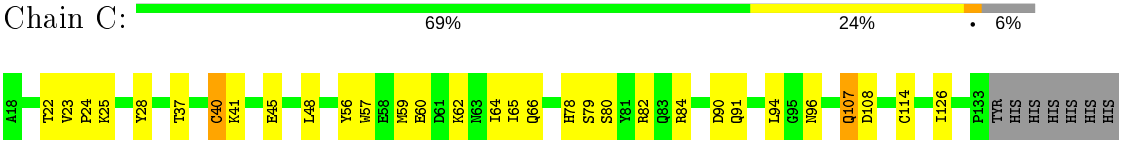
- Molecule 1: Programmed cell death 1 ligand 1

Chain B:



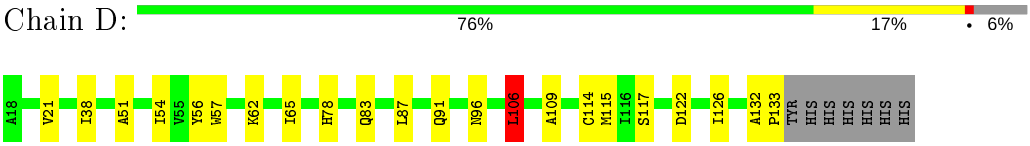
- Molecule 1: Programmed cell death 1 ligand 1

Chain C:



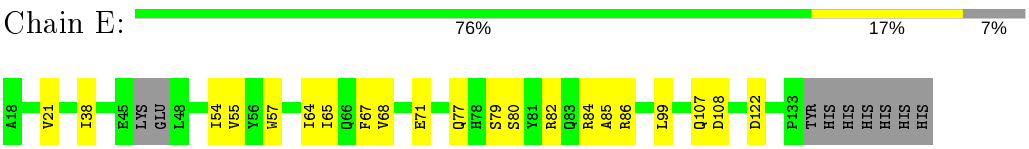
- Molecule 1: Programmed cell death 1 ligand 1

Chain D:



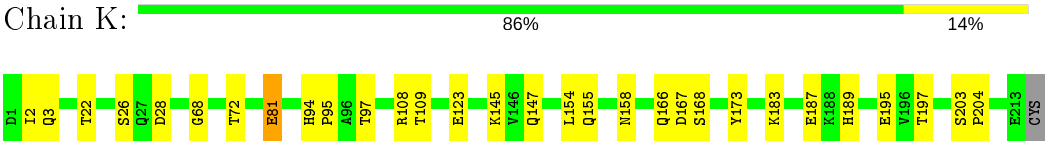
- Molecule 1: Programmed cell death 1 ligand 1

Chain E:



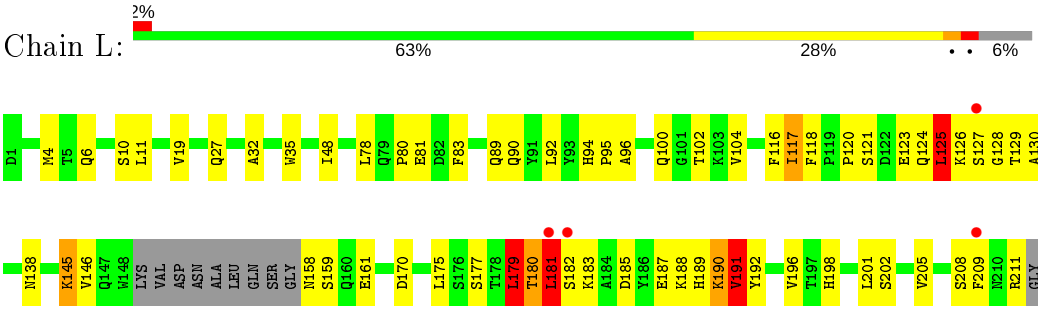
• Molecule 2: atezolizumab light chain

Chain K:



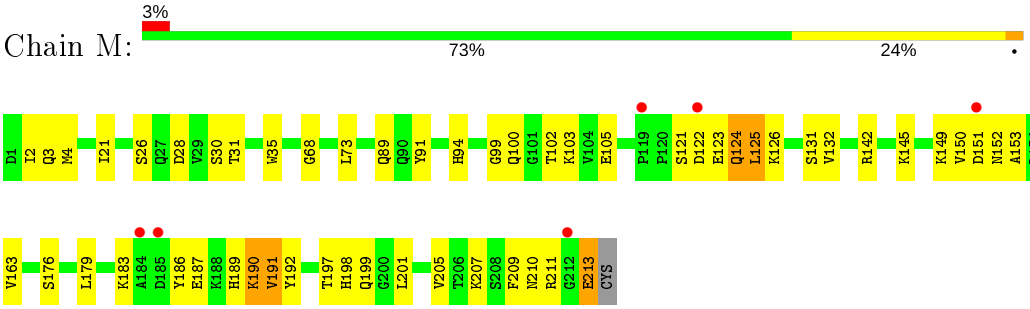
• Molecule 2: atezolizumab light chain

Chain L:



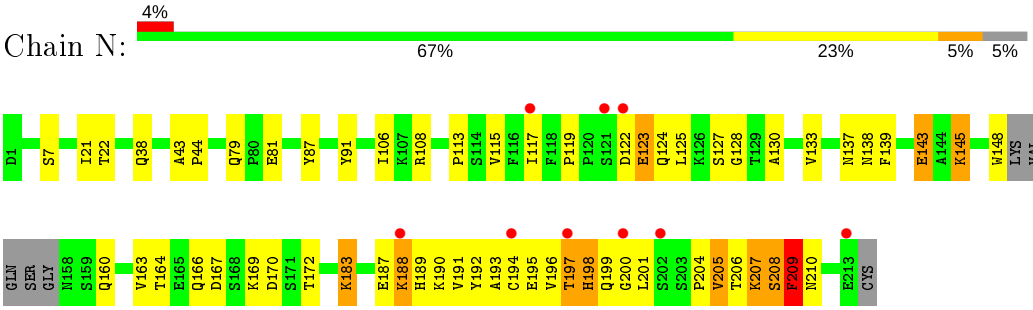
• Molecule 2: atezolizumab light chain

Chain M:



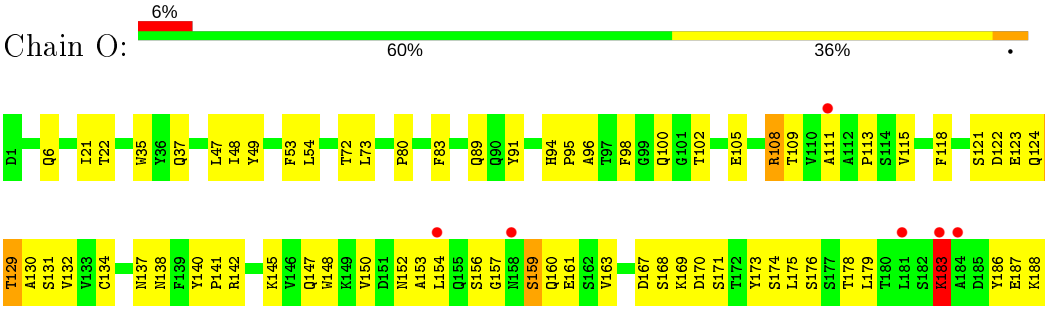
• Molecule 2: atezolizumab light chain

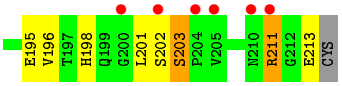
Chain N:



• Molecule 2: atezolizumab light chain

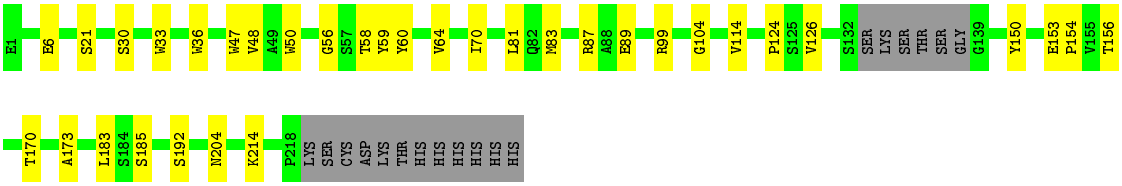
Chain O:





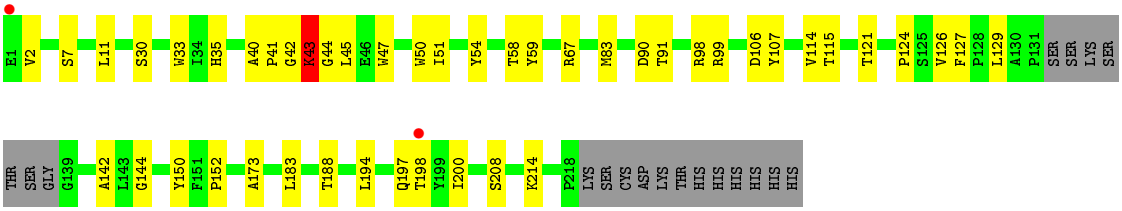
- Molecule 3: atezolizumab heavy chain

Chain F:



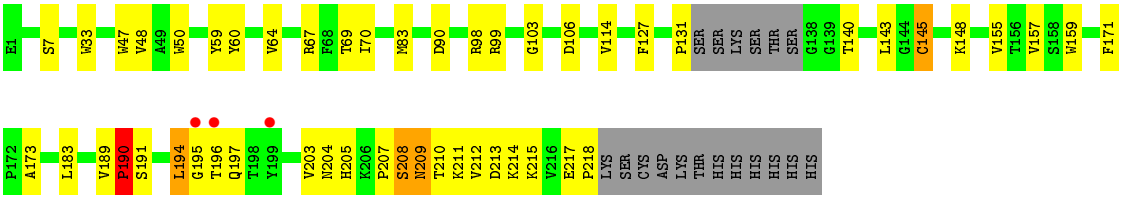
- Molecule 3: atezolizumab heavy chain

Chain G:



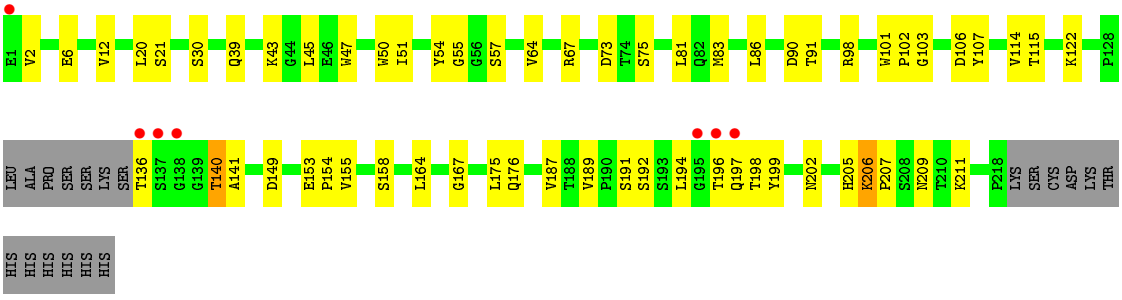
- Molecule 3: atezolizumab heavy chain

Chain H:



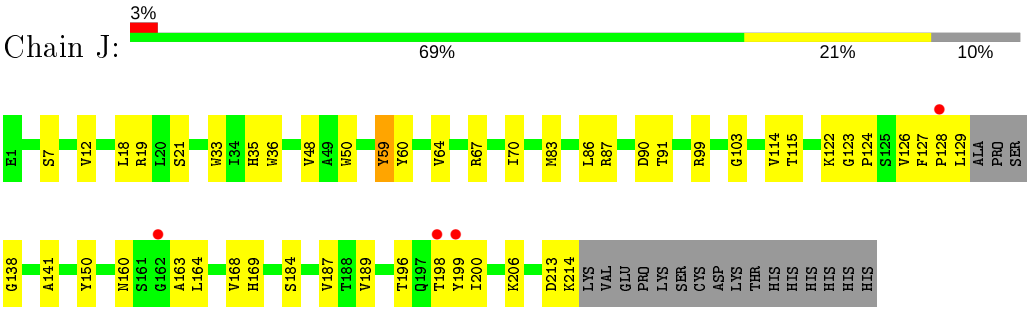
- Molecule 3: atezolizumab heavy chain

Chain S:



- Molecule 3: atezolizumab heavy chain

Chain J:


3%
69%
21%
10%
E1 S7 V12 L18 R19 L20 S21 W33 I34 H35 W36 V48 A49 W50 Y59 Y60 V64 R67 I70 M83 L86 R87 D90 T91 R99 G103 V114 T115 K122 G123 P124 S125 V126 F127 P128 L129 ALA PRO SER SER LYS SER T136 S137
G138 A141 Y150 N160 S161 G162 A163 L164 V168 H169 S184 V187 T188 V189 T196 Q197 T198 Y199 I200 K206 D213 K214 LYS VAL GLU PRO LYS SER CYS ASP LYS THR HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	92.20Å 169.79Å 206.10Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.92 – 3.10 29.92 – 3.10	Depositor EDS
% Data completeness (in resolution range)	97.0 (29.92-3.10) 97.0 (29.92-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.17 (at 3.11Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.208 , 0.256 0.210 , 0.255	Depositor DCC
R_{free} test set	2896 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	43.6	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 15.6	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	0.89	EDS
Total number of atoms	20566	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.53% 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	0.38	0/918	0.58	0/1240
1	B	0.38	0/944	0.59	0/1276
1	C	0.40	0/944	0.52	0/1276
1	D	0.48	0/944	0.67	2/1276 (0.2%)
1	E	0.36	0/925	0.54	0/1250
2	K	0.38	0/1675	0.55	0/2276
2	L	0.41	0/1597	0.75	4/2170 (0.2%)
2	M	0.40	0/1675	0.60	1/2276 (0.0%)
2	N	0.51	0/1610	0.75	0/2187
2	O	0.45	0/1675	0.70	2/2276 (0.1%)
3	F	0.39	0/1640	0.52	0/2241
3	G	0.39	0/1634	0.59	1/2233 (0.0%)
3	H	0.40	0/1638	0.57	0/2238
3	J	0.36	0/1605	0.55	0/2192
3	S	0.34	0/1630	0.55	0/2226
All	All	0.40	0/21054	0.61	10/28633 (0.0%)

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	A	0	1
1	D	0	1
2	K	0	1
2	L	0	4
2	M	0	2
2	N	0	6
2	O	0	5
3	G	0	1
3	H	0	5
3	J	0	1

Continued on next page...

Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
3	S	0	2
All	All	0	29

There are no bond length outliers.

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	G	43	LYS	CD-CE-NZ	-11.85	84.46	111.70
2	L	125	LEU	CA-CB-CG	10.55	139.57	115.30
2	L	181	LEU	CA-CB-CG	6.88	131.12	115.30
1	D	106	LEU	CB-CG-CD2	6.54	122.12	111.00
2	M	125	LEU	CA-CB-CG	-6.30	100.80	115.30

There are no chirality outliers.

5 of 29 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	102	THR	Peptide
2	K	189	HIS	Peptide
2	L	179	LEU	Peptide
2	L	181	LEU	Peptide
2	L	190	LYS	Peptide

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	903	0	903	25	0
1	B	928	0	932	16	0
1	C	928	0	932	28	0
1	D	928	0	932	14	0
1	E	910	0	912	14	0
2	K	1638	0	1588	18	0
2	L	1561	0	1514	56	0
2	M	1638	0	1588	47	0
2	N	1574	0	1523	62	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	O	1638	0	1588	75	0
3	F	1594	0	1538	24	0
3	G	1588	0	1533	35	0
3	H	1592	0	1536	38	0
3	J	1561	0	1501	38	0
3	S	1585	0	1525	42	0
All	All	20566	0	20045	495	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 12.

The worst 5 of 495 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)
3:S:140:THR:HA	3:S:191:SER:HB2	1.46	0.96
2:O:130:ALA:HB3	2:O:183:LYS:HD3	1.54	0.90
3:J:136:THR:HG22	3:J:138:GLY:H	1.39	0.86
1:A:82:ARG:HD2	1:A:83:GLN:H	1.41	0.85
3:G:67:ARG:NH2	3:G:90:ASP:OD2	2.11	0.84

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	109/123 (89%)	103 (94%)	6 (6%)	0	100 100
1	B	114/123 (93%)	109 (96%)	5 (4%)	0	100 100
1	C	114/123 (93%)	106 (93%)	7 (6%)	1 (1%)	17 52
1	D	114/123 (93%)	110 (96%)	4 (4%)	0	100 100
1	E	110/123 (89%)	109 (99%)	1 (1%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
2	K	211/214 (99%)	198 (94%)	13 (6%)	0	100 100
2	L	198/214 (92%)	180 (91%)	17 (9%)	1 (0%)	29 64
2	M	211/214 (99%)	199 (94%)	12 (6%)	0	100 100
2	N	200/214 (94%)	183 (92%)	16 (8%)	1 (0%)	29 64
2	O	211/214 (99%)	199 (94%)	12 (6%)	0	100 100
3	F	208/230 (90%)	199 (96%)	9 (4%)	0	100 100
3	G	207/230 (90%)	202 (98%)	5 (2%)	0	100 100
3	H	208/230 (90%)	203 (98%)	4 (2%)	1 (0%)	29 64
3	J	204/230 (89%)	200 (98%)	4 (2%)	0	100 100
3	S	207/230 (90%)	201 (97%)	6 (3%)	0	100 100
All	All	2626/2835 (93%)	2501 (95%)	121 (5%)	4 (0%)	47 79

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	45	GLU
2	N	123	GLU
3	H	209	ASN
2	L	191	VAL

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	98/108 (91%)	94 (96%)	4 (4%)	30 64
1	B	101/108 (94%)	100 (99%)	1 (1%)	76 90
1	C	101/108 (94%)	98 (97%)	3 (3%)	41 71
1	D	101/108 (94%)	99 (98%)	2 (2%)	55 80
1	E	99/108 (92%)	98 (99%)	1 (1%)	76 90
2	K	186/187 (100%)	185 (100%)	1 (0%)	88 94

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
2	L	178/187 (95%)	170 (96%)	8 (4%)	27 60
2	M	186/187 (100%)	181 (97%)	5 (3%)	44 74
2	N	179/187 (96%)	170 (95%)	9 (5%)	24 57
2	O	186/187 (100%)	178 (96%)	8 (4%)	29 62
3	F	174/191 (91%)	172 (99%)	2 (1%)	73 89
3	G	173/191 (91%)	170 (98%)	3 (2%)	60 83
3	H	173/191 (91%)	168 (97%)	5 (3%)	42 72
3	J	170/191 (89%)	166 (98%)	4 (2%)	49 76
3	S	173/191 (91%)	171 (99%)	2 (1%)	71 88
All	All	2278/2430 (94%)	2220 (98%)	58 (2%)	47 75

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	M	213	GLU
1	D	78	HIS
2	O	183	LYS
3	H	7	SER
3	H	189	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 14 such sidechains are listed below:

Mol	Chain	Res	Type
3	H	197	GLN
2	N	138	ASN
2	O	124	GLN
2	M	210	ASN
2	O	94	HIS

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

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, 95^{th} 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	113/123 (91%)	-0.25	0 100 100	22, 34, 72, 83	0
1	B	116/123 (94%)	-0.25	4 (3%) 45 24	20, 36, 57, 66	0
1	C	116/123 (94%)	-0.48	0 100 100	20, 28, 40, 51	0
1	D	116/123 (94%)	-0.43	0 100 100	21, 34, 48, 61	0
1	E	114/123 (92%)	-0.29	0 100 100	26, 36, 48, 70	0
2	K	213/214 (99%)	-0.44	0 100 100	14, 32, 57, 75	0
2	L	202/214 (94%)	-0.14	4 (1%) 65 44	22, 44, 111, 122	0
2	M	213/214 (99%)	-0.24	6 (2%) 53 30	14, 33, 90, 109	0
2	N	204/214 (95%)	0.05	9 (4%) 34 17	19, 51, 113, 120	0
2	O	213/214 (99%)	0.13	12 (5%) 24 11	20, 48, 106, 119	0
3	F	212/230 (92%)	-0.43	0 100 100	17, 31, 51, 65	0
3	G	211/230 (91%)	-0.33	2 (0%) 84 69	24, 33, 85, 112	0
3	H	212/230 (92%)	-0.21	3 (1%) 75 56	14, 32, 90, 136	0
3	J	208/230 (90%)	-0.19	6 (2%) 51 28	14, 32, 83, 121	0
3	S	211/230 (91%)	-0.12	7 (3%) 46 24	24, 43, 108, 136	0
All	All	2674/2835 (94%)	-0.22	53 (1%) 65 44	14, 35, 97, 136	0

The worst 5 of 53 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	J	198	THR	5.9
3	S	138	GLY	4.3
2	M	119	PRO	4.3
1	B	131	ASN	3.8
2	N	122	ASP	3.7

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.