wwPDB X-ray Structure Validation Summary Report (i)

Dec 22, 2022 – 09:19 am GMT

PDB ID	:	7OX2
Title	:	Fab 6E2: hIL-9 complex
Authors	:	De Vos, T.; Savvides, S.N.
Deposited on	:	2021-06-22
Resolution	:	3.34 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

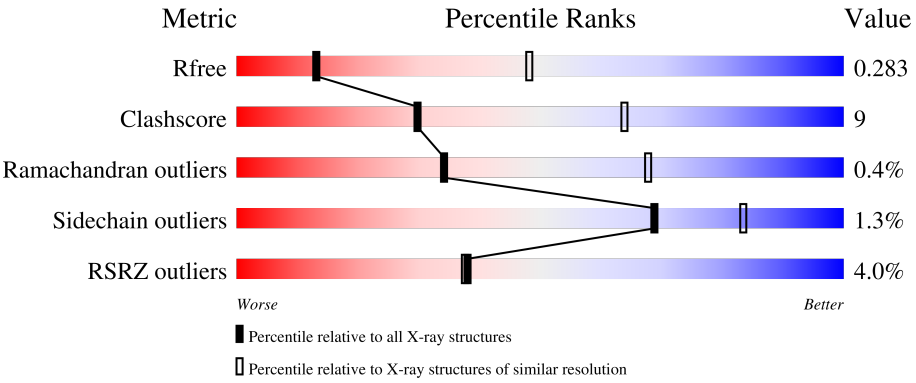
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.31.3
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.3

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 3.34 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.



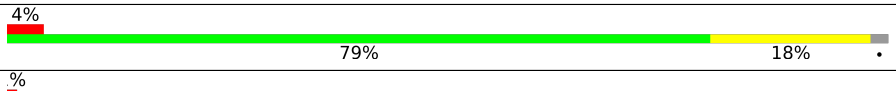
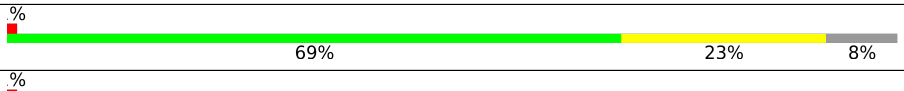
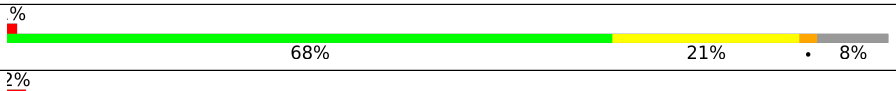
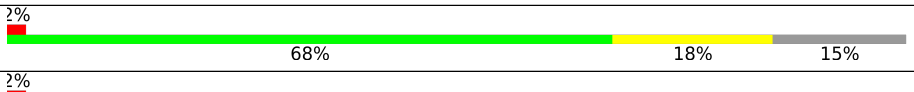
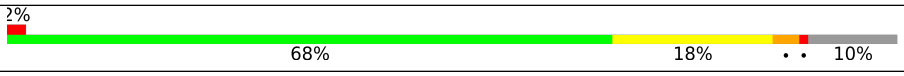
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1060 (3.38-3.30)
Clashscore	141614	1111 (3.38-3.30)
Ramachandran outliers	138981	1090 (3.38-3.30)
Sidechain outliers	138945	1089 (3.38-3.30)
RSRZ outliers	127900	1028 (3.38-3.30)

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 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% 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	220	81% 15%
1	B	220	4% 79% 17%
1	D	220	6% 79% 15% 5%
1	J	220	5% 74% 22% 5%
2	C	214	7% 85% 13%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	214	
2	K	214	
2	L	214	
3	M	130	
3	N	130	
3	O	130	
3	T	130	

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 31936 atoms, of which 15777 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 Heavy chain (Fab 6E2).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	212	Total	C	H	N	O	S	0	0	0
			3109	995	1539	256	314	5			
1	B	213	Total	C	H	N	O	S	0	0	0
			3121	998	1545	257	316	5			
1	D	210	Total	C	H	N	O	S	0	0	0
			3091	990	1531	254	311	5			
1	J	210	Total	C	H	N	O	S	0	0	0
			3085	988	1528	253	311	5			

- Molecule 2 is a protein called Light chain (Fab 6E2).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	L	209	Total	C	H	N	O	S	0	0	0
			3050	966	1493	263	324	4			
2	C	210	Total	C	H	N	O	S	0	0	0
			3060	969	1497	264	326	4			
2	E	209	Total	C	H	N	O	S	0	0	0
			3049	966	1492	263	324	4			
2	K	209	Total	C	H	N	O	S	0	0	0
			3048	966	1491	263	324	4			

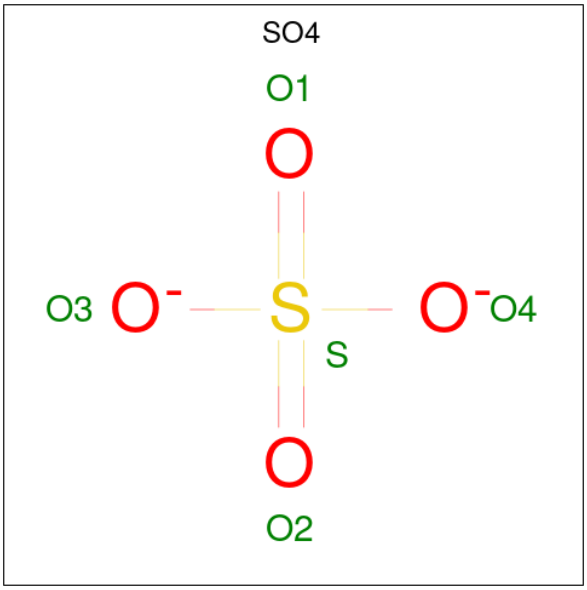
- Molecule 3 is a protein called Interleukin-9.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	T	117	Total	C	H	N	O	S	0	0	0
			1832	571	918	158	172	13			
3	O	111	Total	C	H	N	O	S	0	0	0
			1753	548	884	149	160	12			
3	M	120	Total	C	H	N	O	S	0	0	0
			1859	578	931	161	175	14			
3	N	119	Total	C	H	N	O	S	0	0	0
			1852	576	928	160	174	14			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
T	15	GLY	-	expression tag	UNP P15248
T	16	SER	-	expression tag	UNP P15248
T	17	HIS	-	expression tag	UNP P15248
T	18	MET	-	expression tag	UNP P15248
O	15	GLY	-	expression tag	UNP P15248
O	16	SER	-	expression tag	UNP P15248
O	17	HIS	-	expression tag	UNP P15248
O	18	MET	-	expression tag	UNP P15248
M	15	GLY	-	expression tag	UNP P15248
M	16	SER	-	expression tag	UNP P15248
M	17	HIS	-	expression tag	UNP P15248
M	18	MET	-	expression tag	UNP P15248
N	15	GLY	-	expression tag	UNP P15248
N	16	SER	-	expression tag	UNP P15248
N	17	HIS	-	expression tag	UNP P15248
N	18	MET	-	expression tag	UNP P15248

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	D	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	M	1	Total O S 5 4 1	0	0
4	M	1	Total O S 5 4 1	0	0

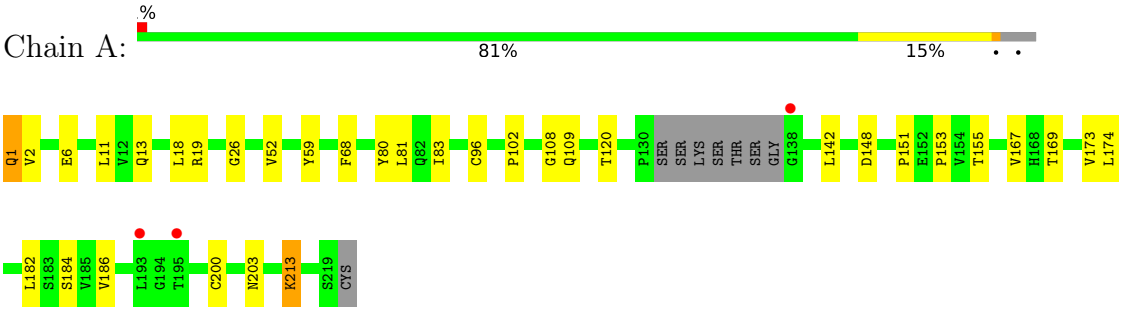
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	L	1	Total O 1 1	0	0
5	C	1	Total O 1 1	0	0

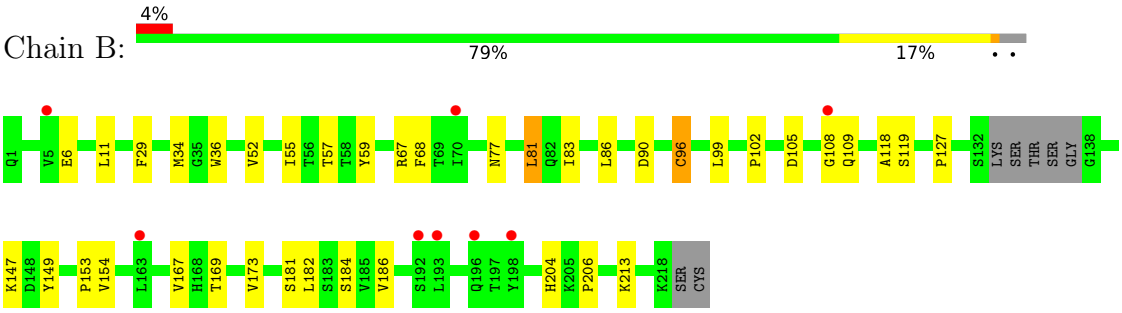
3 Residue-property plots (i)

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 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.

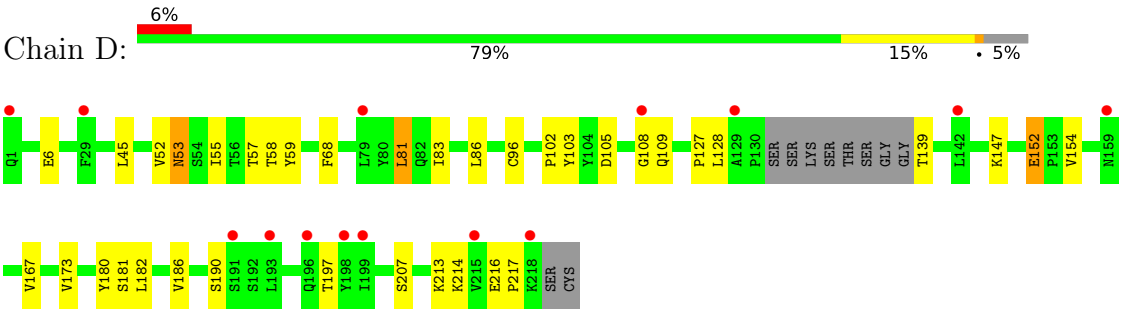
• Molecule 1: Heavy chain (Fab 6E2)



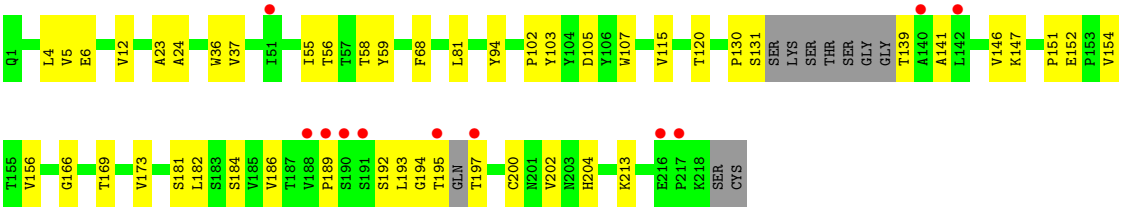
• Molecule 1: Heavy chain (Fab 6E2)



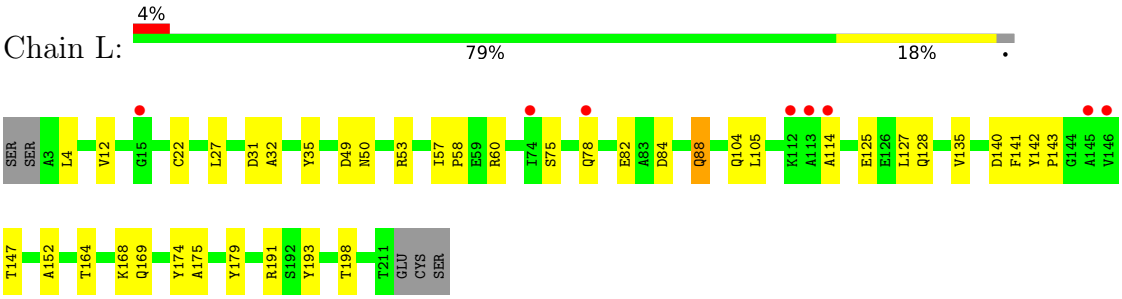
• Molecule 1: Heavy chain (Fab 6E2)



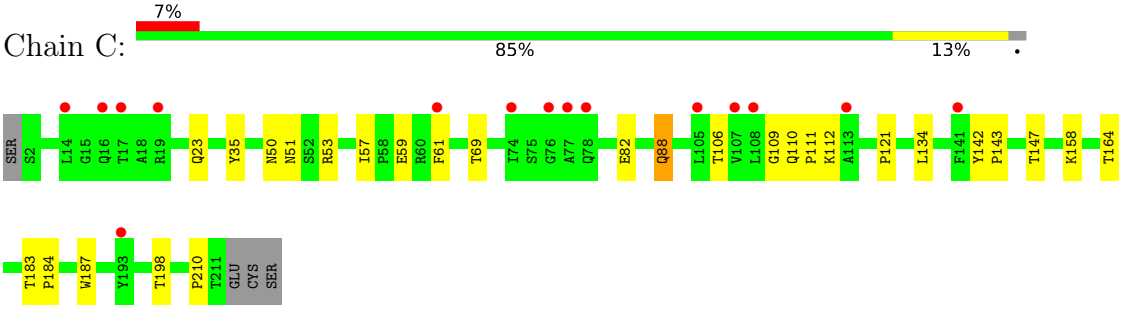
• Molecule 1: Heavy chain (Fab 6E2)



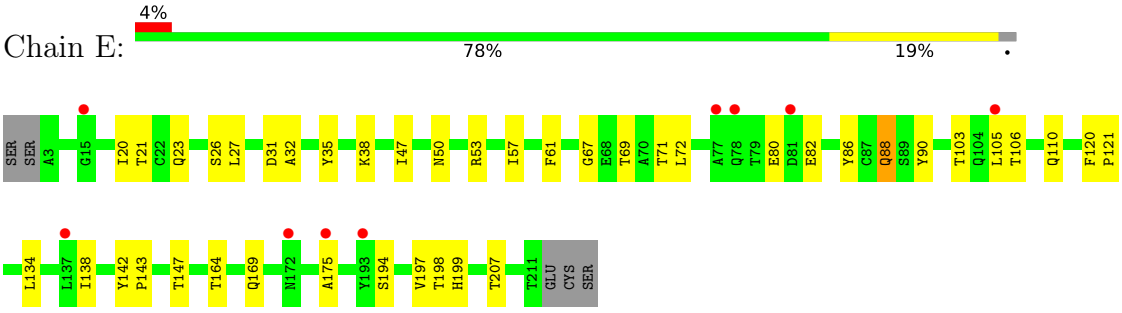
• Molecule 2: Light chain (Fab 6E2)



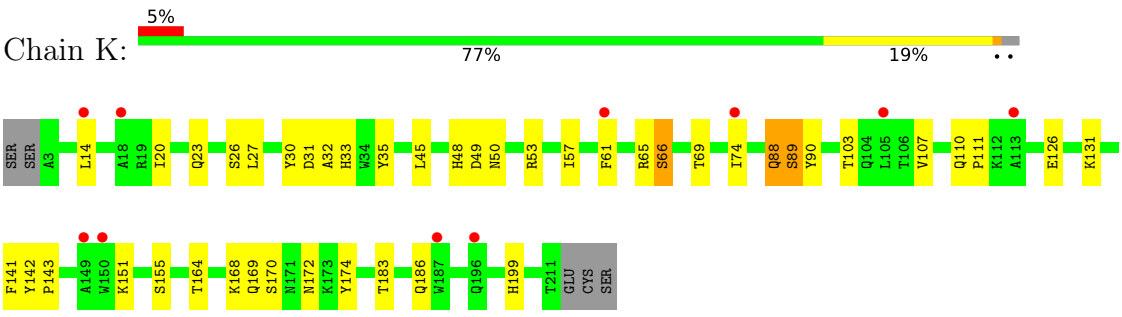
• Molecule 2: Light chain (Fab 6E2)



• Molecule 2: Light chain (Fab 6E2)

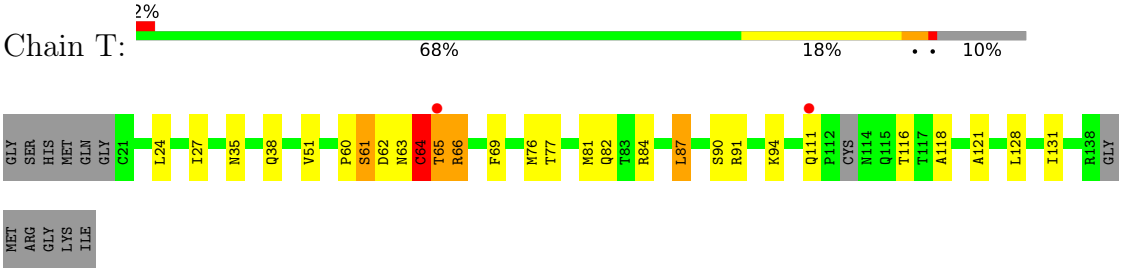


• Molecule 2: Light chain (Fab 6E2)



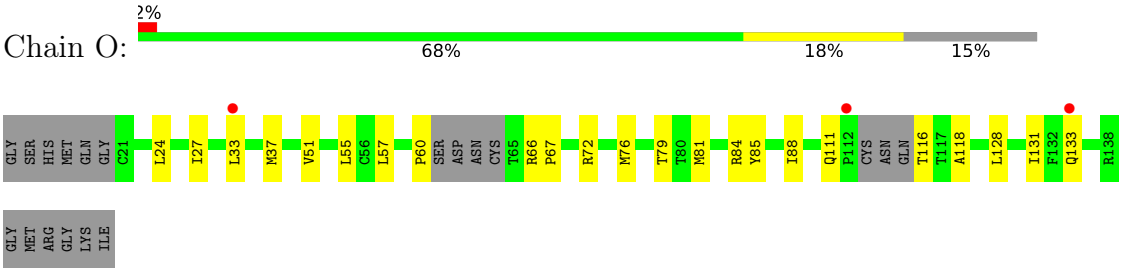
• Molecule 3: Interleukin-9

Chain T:



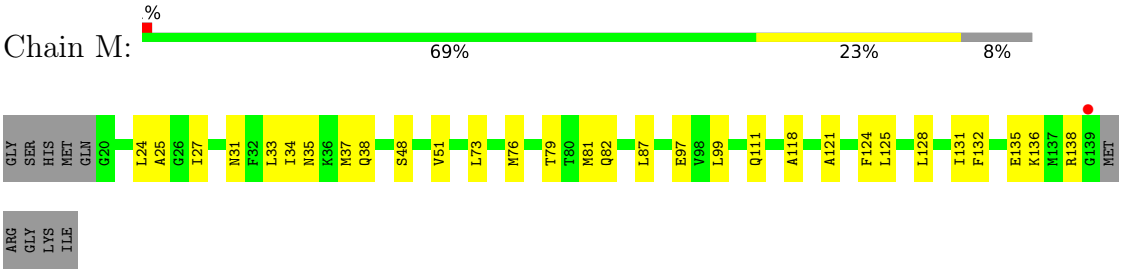
• Molecule 3: Interleukin-9

Chain O:



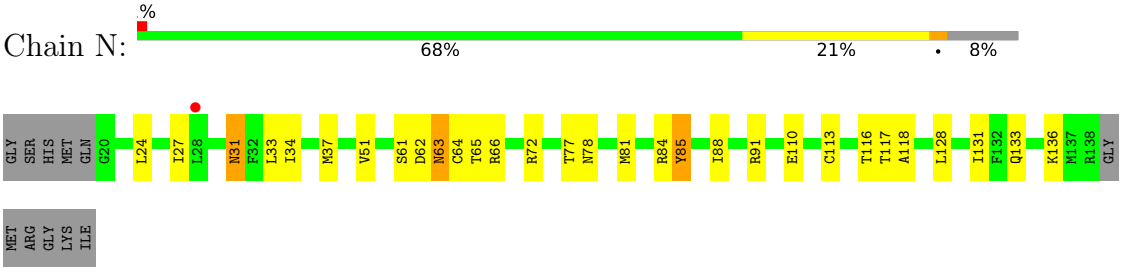
• Molecule 3: Interleukin-9

Chain M:



• Molecule 3: Interleukin-9

Chain N:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	69.36Å 94.36Å 107.33Å 82.54° 72.00° 80.79°	Depositor
Resolution (Å)	48.92 – 3.34 48.92 – 3.34	Depositor EDS
% Data completeness (in resolution range)	93.6 (48.92-3.34) 93.6 (48.92-3.34)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.22 (at 3.33Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.224 , 0.283 0.224 , 0.283	Depositor DCC
R_{free} test set	1728 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	78.5	Xtriage
Anisotropy	0.229	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 59.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	31936	wwPDB-VP
Average B, all atoms (Å^2)	86.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.39% 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

Bond lengths and bond angles in the following residue types are not validated in this section: SO4

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	0.31	0/1607	0.57	0/2194
1	B	0.29	0/1613	0.54	0/2202
1	D	0.29	0/1597	0.54	0/2181
1	J	0.34	0/1593	0.62	1/2174 (0.0%)
2	C	0.29	0/1599	0.55	0/2186
2	E	0.28	0/1593	0.52	0/2178
2	K	0.30	0/1593	0.56	0/2178
2	L	0.32	0/1593	0.57	0/2178
3	M	0.39	0/943	0.86	1/1271 (0.1%)
3	N	0.34	0/939	0.77	2/1266 (0.2%)
3	O	0.37	0/882	0.76	2/1186 (0.2%)
3	T	0.35	0/928	0.79	3/1250 (0.2%)
All	All	0.32	0/16480	0.62	9/22444 (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	D	0	1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	M	99	LEU	CB-CG-CD2	-6.69	99.62	111.00
3	T	66	ARG	NE-CZ-NH2	-6.38	117.11	120.30
3	N	85	TYR	CB-CG-CD2	-5.78	117.53	121.00
3	T	87	LEU	CB-CG-CD1	-5.63	101.44	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	J	152	GLU	C-N-CD	-5.44	108.64	120.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	152	GLU	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	1570	1539	1539	30	0
1	B	1576	1545	1544	33	2
1	D	1560	1531	1531	34	1
1	J	1557	1528	1527	32	1
2	C	1563	1497	1497	19	1
2	E	1557	1492	1492	35	0
2	K	1557	1491	1492	34	1
2	L	1557	1493	1492	32	2
3	M	928	931	931	27	0
3	N	924	928	928	26	0
3	O	869	884	884	27	0
3	T	914	918	919	25	0
4	A	5	0	0	0	0
4	B	5	0	0	1	0
4	D	5	0	0	0	0
4	M	10	0	0	2	0
5	C	1	0	0	0	0
5	L	1	0	0	0	0
All	All	16159	15777	15776	299	4

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 299 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:B:57:THR:OG1	3:O:133:GLN:OE1	1.67	1.12
3:M:111:GLN:OE1	2:K:66:SER:OG	1.68	1.10
1:J:58:THR:O	3:N:136:LYS:NZ	1.86	1.07
1:D:58:THR:O	3:M:136:LYS:NZ	1.98	0.96
3:N:66:ARG:NH2	3:N:113:CYS:SG	2.41	0.93

All (4) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:158:LYS:O	2:K:23:GLN:NE2[1_465]	2.10	0.10
2:L:128:GLN:NE2	1:B:119:SER:O[1_546]	2.11	0.09
1:D:207:SER:O	1:J:213:LYS:NZ[1_564]	2.13	0.07
2:L:128:GLN:HE22	1:B:119:SER:O[1_546]	1.55	0.05

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	208/220 (94%)	200 (96%)	6 (3%)	2 (1%)	15 49
1	B	209/220 (95%)	200 (96%)	8 (4%)	1 (0%)	29 63
1	D	206/220 (94%)	197 (96%)	8 (4%)	1 (0%)	29 63
1	J	204/220 (93%)	197 (97%)	6 (3%)	1 (0%)	29 63
2	C	208/214 (97%)	191 (92%)	17 (8%)	0	100 100
2	E	207/214 (97%)	197 (95%)	10 (5%)	0	100 100
2	K	207/214 (97%)	196 (95%)	11 (5%)	0	100 100
2	L	207/214 (97%)	197 (95%)	10 (5%)	0	100 100
3	M	118/130 (91%)	112 (95%)	6 (5%)	0	100 100
3	N	117/130 (90%)	112 (96%)	4 (3%)	1 (1%)	17 51

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
3	O	105/130 (81%)	103 (98%)	2 (2%)	0	100 100
3	T	113/130 (87%)	107 (95%)	3 (3%)	3 (3%)	5 28
All	All	2109/2256 (94%)	2009 (95%)	91 (4%)	9 (0%)	34 68

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	148	ASP
3	T	64	CYS
3	T	65	THR
1	J	194	GLY
3	N	62	ASP

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	177/184 (96%)	173 (98%)	4 (2%)	50 75
1	B	178/184 (97%)	175 (98%)	3 (2%)	60 80
1	D	176/184 (96%)	174 (99%)	2 (1%)	73 86
1	J	176/184 (96%)	173 (98%)	3 (2%)	60 80
2	C	175/179 (98%)	173 (99%)	2 (1%)	73 86
2	E	174/179 (97%)	173 (99%)	1 (1%)	86 92
2	K	174/179 (97%)	170 (98%)	4 (2%)	50 75
2	L	174/179 (97%)	172 (99%)	2 (1%)	73 86
3	M	110/118 (93%)	110 (100%)	0	100 100
3	N	110/118 (93%)	108 (98%)	2 (2%)	59 79
3	O	103/118 (87%)	103 (100%)	0	100 100
3	T	109/118 (92%)	108 (99%)	1 (1%)	78 88
All	All	1836/1924 (95%)	1812 (99%)	24 (1%)	69 83

5 of 24 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	88	GLN
1	J	200	CYS
1	J	193	LEU
2	K	66	SER
3	T	64	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
3	M	35	ASN
2	K	110	GLN
2	K	172	ASN
3	O	111	GLN
1	B	77	ASN

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)

5 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	SO4	M	201	-	4,4,4	0.15	0	6,6,6	0.10	0
4	SO4	B	301	-	4,4,4	0.15	0	6,6,6	0.18	0
4	SO4	D	301	-	4,4,4	0.15	0	6,6,6	0.23	0
4	SO4	M	202	-	4,4,4	0.18	0	6,6,6	0.14	0
4	SO4	A	301	-	4,4,4	0.14	0	6,6,6	0.09	0

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.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	M	201	SO4	1	0
4	B	301	SO4	1	0
4	M	202	SO4	1	0

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^{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	212/220 (96%)	0.29	3 (1%)	75	75	46, 67, 95, 115	0
1	B	213/220 (96%)	0.56	8 (3%)	40	39	47, 75, 104, 120	0
1	D	210/220 (95%)	0.64	14 (6%)	17	19	58, 83, 104, 111	0
1	J	210/220 (95%)	0.61	11 (5%)	27	27	49, 77, 110, 139	0
2	C	210/214 (98%)	0.57	15 (7%)	16	17	53, 84, 110, 113	0
2	E	209/214 (97%)	0.47	9 (4%)	35	36	55, 81, 98, 112	0
2	K	209/214 (97%)	0.45	10 (4%)	30	31	52, 76, 99, 111	0
2	L	209/214 (97%)	0.32	8 (3%)	40	39	45, 69, 95, 111	0
3	M	120/130 (92%)	0.33	1 (0%)	86	87	52, 76, 107, 134	0
3	N	119/130 (91%)	0.42	1 (0%)	86	87	53, 76, 109, 125	0
3	O	111/130 (85%)	0.40	3 (2%)	54	53	64, 86, 108, 133	0
3	T	117/130 (90%)	0.32	2 (1%)	70	69	51, 78, 117, 127	0
All	All	2149/2256 (95%)	0.46	85 (3%)	38	37	45, 77, 105, 139	0

The worst 5 of 85 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	129	ALA	5.0
2	L	145	ALA	4.5
2	E	81	ASP	4.2
2	K	196	GLN	4.1
2	C	61	PHE	4.0

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 (i)

There are no monosaccharides 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^{th} 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(Å^2)	Q<0.9
4	SO4	D	301	5/5	0.87	0.14	86,88,102,139	0
4	SO4	M	202	5/5	0.88	0.34	80,91,104,139	0
4	SO4	A	301	5/5	0.94	0.17	62,70,81,107	0
4	SO4	M	201	5/5	0.94	0.11	90,93,106,122	0
4	SO4	B	301	5/5	0.94	0.12	82,82,93,143	0

6.5 Other polymers (i)

There are no such residues in this entry.