



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 16, 2020 – 06:49 PM BST

PDB ID : 6URI
Title : HIV-1 Nef in complex with the CD4 cytoplasmic domain and the AP2 clathrin adaptor complex
Authors : Jia, X.; Kwon, Y.
Deposited on : 2019-10-23
Resolution : 3.00 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	:	2.13.1
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

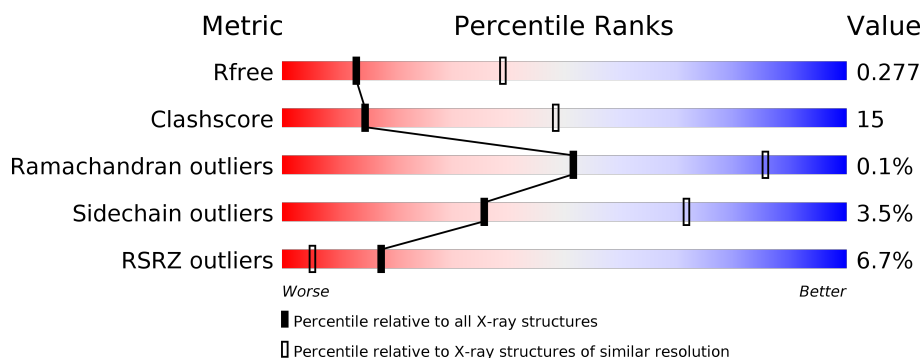
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (3.00-3.00)

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	641	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 69% 25% • 5% </div> </div>
2	S	142	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, green, yellow, orange, red);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 71% 26% • </div> </div>
3	N	186	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 3% 62% 24% • 12% </div> </div>
4	B	615	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 11% 45% 32% • 22% </div> </div>
5	M	135	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 12% 55% 33% • 11% </div> </div>
6	D	62	<div> <div style="width: 100%; height: 10px; background: linear-gradient(to right, red, orange, yellow, green, grey);"></div> <div style="display: flex; justify-content: space-between; margin-top: 5px;"> 5% 6% 13% • 77% </div> </div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12263 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 AP-2 complex subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	611	Total	C	N	O	S	0	0	0
			4823	3073	830	899	21			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	initiating methionine	UNP Q66HM2
A	-12	GLY	-	expression tag	UNP Q66HM2
A	-11	SER	-	expression tag	UNP Q66HM2
A	-10	SER	-	expression tag	UNP Q66HM2
A	-9	HIS	-	expression tag	UNP Q66HM2
A	-8	HIS	-	expression tag	UNP Q66HM2
A	-7	HIS	-	expression tag	UNP Q66HM2
A	-6	HIS	-	expression tag	UNP Q66HM2
A	-5	HIS	-	expression tag	UNP Q66HM2
A	-4	HIS	-	expression tag	UNP Q66HM2
A	-3	SER	-	expression tag	UNP Q66HM2
A	-2	GLN	-	expression tag	UNP Q66HM2
A	-1	ASP	-	expression tag	UNP Q66HM2
A	0	PRO	-	expression tag	UNP Q66HM2
A	622	GLU	-	expression tag	UNP Q66HM2
A	623	ASN	-	expression tag	UNP Q66HM2
A	624	LEU	-	expression tag	UNP Q66HM2
A	625	TYR	-	expression tag	UNP Q66HM2
A	626	PHE	-	expression tag	UNP Q66HM2
A	627	GLN	-	expression tag	UNP Q66HM2

- Molecule 2 is a protein called AP-2 complex subunit sigma.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	S	142	Total	C	N	O	S	0	0	0
			1202	778	202	215	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
S	45	ARG	LYS	conflict	UNP P53680

- Molecule 3 is a protein called Protein Nef.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	N	164	Total	C	N	O	S	0	0	0
			1346	861	233	248	4			

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
N	21	ALA	-	expression tag	UNP Q90VU7
N	22	GLY	-	expression tag	UNP Q90VU7
N	23	PHE	-	expression tag	UNP Q90VU7
N	24	SER	-	expression tag	UNP Q90VU7
N	25	MET	-	expression tag	UNP Q90VU7

- Molecule 4 is a protein called AP-2 complex subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	480	Total	C	N	O	S	0	0	0
			3797	2414	639	727	17			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-23	MET	-	initiating methionine	UNP P63010
B	-22	GLY	-	expression tag	UNP P63010
B	-21	SER	-	expression tag	UNP P63010
B	-20	SER	-	expression tag	UNP P63010
B	-19	HIS	-	expression tag	UNP P63010
B	-18	HIS	-	expression tag	UNP P63010
B	-17	HIS	-	expression tag	UNP P63010
B	-16	HIS	-	expression tag	UNP P63010
B	-15	HIS	-	expression tag	UNP P63010
B	-14	HIS	-	expression tag	UNP P63010
B	-13	SER	-	expression tag	UNP P63010
B	-12	GLN	-	expression tag	UNP P63010
B	-11	ASP	-	expression tag	UNP P63010
B	-10	PRO	-	expression tag	UNP P63010
B	-9	ASN	-	expression tag	UNP P63010

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	-8	SER	-	expression tag	UNP P63010
B	-7	SER	-	expression tag	UNP P63010
B	-6	SER	-	expression tag	UNP P63010
B	-5	ALA	-	expression tag	UNP P63010
B	-4	ARG	-	expression tag	UNP P63010
B	-3	LEU	-	expression tag	UNP P63010
B	-2	GLN	-	expression tag	UNP P63010
B	-1	VAL	-	expression tag	UNP P63010
B	0	ASP	-	expression tag	UNP P63010
B	16	SER	PHE	variant	UNP P63010
B	421	HIS	TYR	variant	UNP P63010
B	434	GLY	GLU	variant	UNP P63010

- Molecule 5 is a protein called AP-2 complex subunit mu.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	M	120	Total	C	N	O	S	0	0	0
			975	631	167	172	5			

- Molecule 6 is a protein called cDNA FLJ50658, highly similar to T-cell surface glycoprotein CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	D	14	Total	C	N	O	S	0	0	0
			118	73	24	20	1			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	358	GLY	-	expression tag	UNP B4DT49
D	359	VAL	-	expression tag	UNP B4DT49
D	360	ASP	-	expression tag	UNP B4DT49
D	361	GLY	-	expression tag	UNP B4DT49
D	362	SER	-	expression tag	UNP B4DT49
D	363	ASP	-	expression tag	UNP B4DT49
D	364	GLU	-	expression tag	UNP B4DT49
D	365	ALA	-	expression tag	UNP B4DT49
D	366	SER	-	expression tag	UNP B4DT49
D	367	GLU	-	expression tag	UNP B4DT49
D	368	LEU	-	expression tag	UNP B4DT49
D	369	ALA	-	expression tag	UNP B4DT49
D	370	CYS	-	expression tag	UNP B4DT49

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	371	PRO	-	expression tag	UNP B4DT49
D	372	THR	-	expression tag	UNP B4DT49
D	373	PRO	-	expression tag	UNP B4DT49
D	374	LYS	-	expression tag	UNP B4DT49
D	375	GLU	-	expression tag	UNP B4DT49
D	376	ASP	-	expression tag	UNP B4DT49
D	377	GLY	-	expression tag	UNP B4DT49
D	378	LEU	-	expression tag	UNP B4DT49
D	379	ALA	-	expression tag	UNP B4DT49
D	380	GLN	-	expression tag	UNP B4DT49
D	381	GLN	-	expression tag	UNP B4DT49
D	382	GLN	-	expression tag	UNP B4DT49
D	383	THR	-	expression tag	UNP B4DT49
D	384	GLN	-	expression tag	UNP B4DT49
D	385	LEU	-	expression tag	UNP B4DT49
D	386	ASN	-	expression tag	UNP B4DT49
D	387	LEU	-	expression tag	UNP B4DT49
D	388	ARG	-	expression tag	UNP B4DT49
D	389	GLY	-	expression tag	UNP B4DT49
D	390	SER	-	expression tag	UNP B4DT49
D	391	GLY	-	expression tag	UNP B4DT49
D	392	SER	-	expression tag	UNP B4DT49
D	393	GLY	-	expression tag	UNP B4DT49

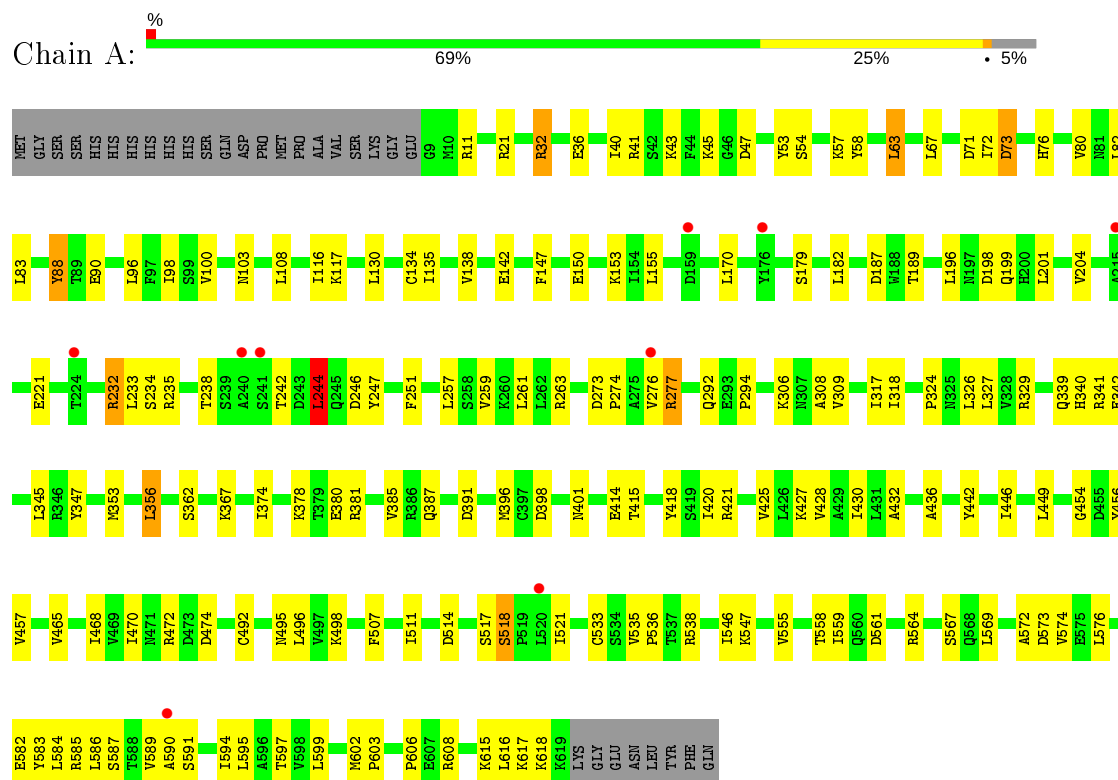
- Molecule 7 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total O 1 1	0	0
7	N	1	Total O 1 1	0	0

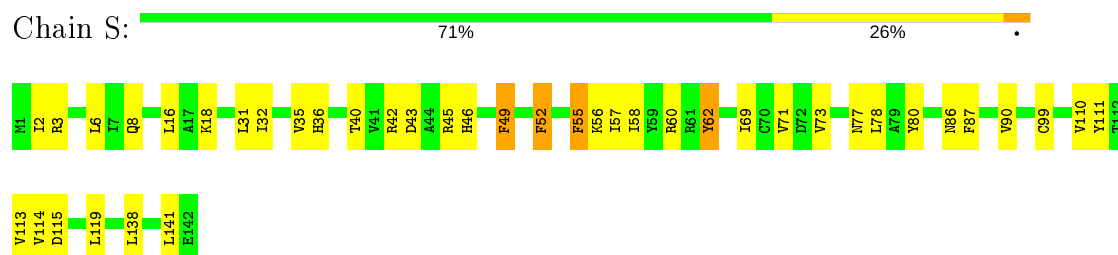
3 Residue-property plots [i](#)

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: AP-2 complex subunit alpha

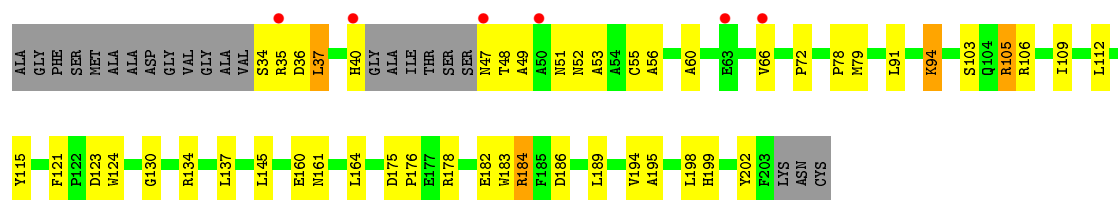


- Molecule 2: AP-2 complex subunit sigma

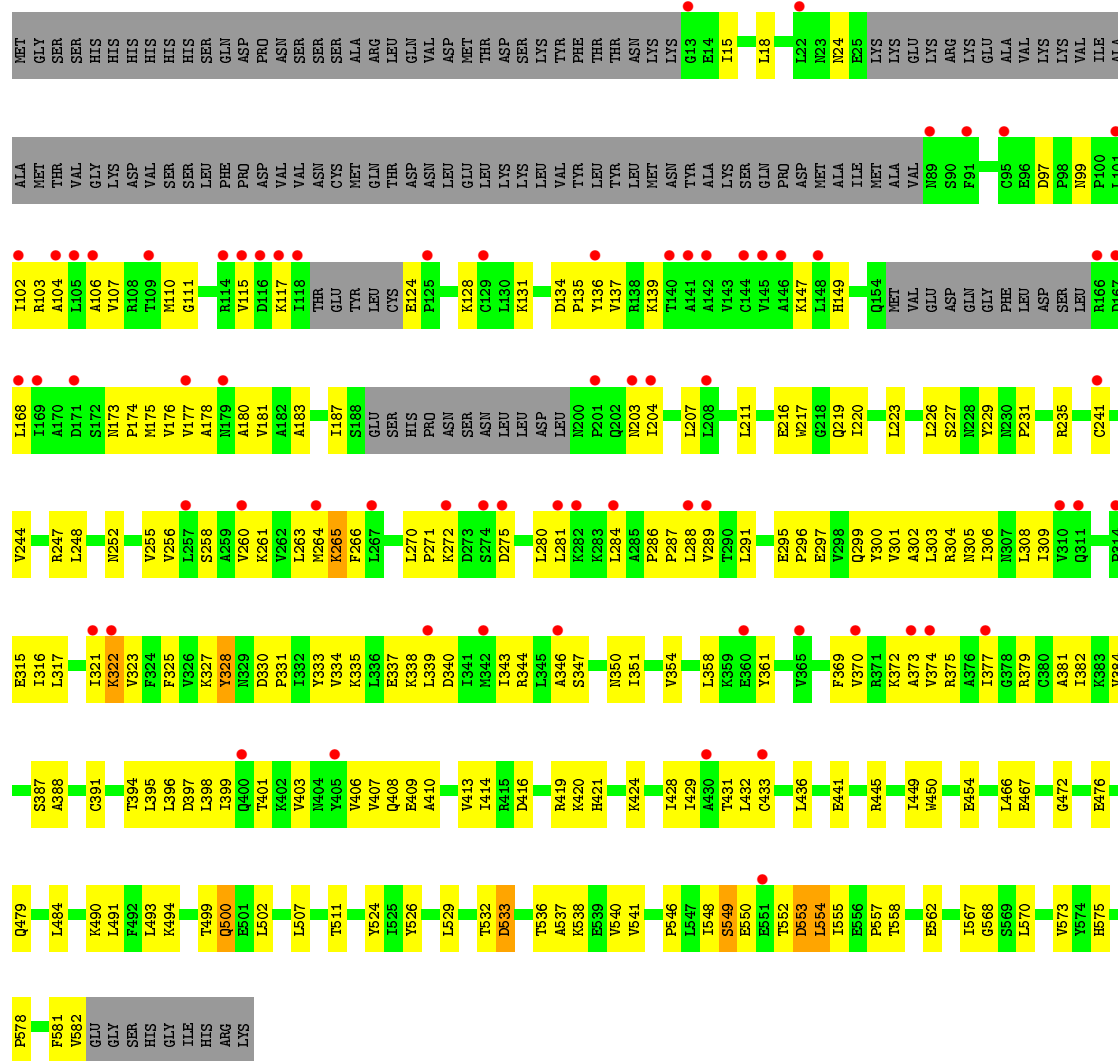


- Molecule 3: Protein Nef

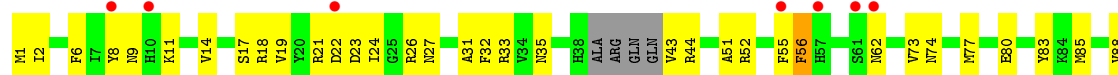


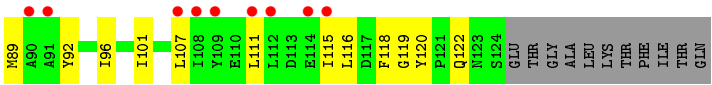


• Molecule 4: AP-2 complex subunit beta

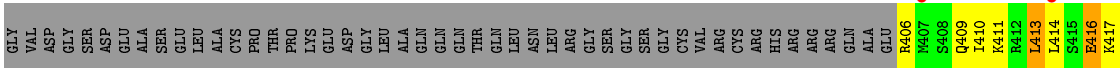


• Molecule 5: AP-2 complex subunit mu





- Molecule 6: cDNA FLJ50658, highly similar to T-cell surface glycoprotein CD4



4 Data and refinement statistics

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	109.15Å 109.15Å 178.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.82 – 3.00 48.82 – 3.00	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.82-3.00) 94.0 (48.82-3.00)	Depositor EDS
R_{merge}	0.29	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.34 (at 3.01Å)	Xtriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, R_{free}	0.241 , 0.277 0.241 , 0.277	Depositor DCC
R_{free} test set	1018 reflections (2.43%)	wwPDB-VP
Wilson B-factor (Å ²)	83.1	Xtriage
Anisotropy	0.106	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 67.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.054 for h,-k,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12263	wwPDB-VP
Average B, all atoms (Å ²)	118.0	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.*

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	0/4908	0.46	1/6653 (0.0%)
2	S	0.29	0/1226	0.51	0/1653
3	N	0.30	0/1390	0.53	1/1893 (0.1%)
4	B	0.32	0/3852	0.54	2/5226 (0.0%)
5	M	0.32	0/995	0.61	1/1343 (0.1%)
6	D	0.63	0/117	1.52	2/151 (1.3%)
All	All	0.30	0/12488	0.53	7/16919 (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
2	S	0	1
4	B	0	1
All	All	0	3

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	413	LEU	CB-CG-CD1	7.79	124.24	111.00
6	D	413	LEU	CB-CG-CD2	-7.64	98.01	111.00
1	A	244	LEU	CA-CB-CG	6.48	130.20	115.30
5	M	116	LEU	CA-CB-CG	5.86	128.77	115.30
3	N	37	LEU	CB-CG-CD1	5.39	120.16	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	242	THR	Peptide
4	B	549	SER	Peptide
2	S	62	TYR	Peptide

5.2 Too-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 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	4823	0	4936	107	0
2	S	1202	0	1195	30	0
3	N	1346	0	1269	54	0
4	B	3797	0	3891	169	0
5	M	975	0	973	37	0
6	D	118	0	137	27	0
7	A	1	0	0	0	0
7	N	1	0	0	0	0
All	All	12263	0	12401	377	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 15.

The worst 5 of 377 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:N:47:ASN:HB2	3:N:52:ASN:HB2	1.24	1.14
3:N:37:LEU:HA	6:D:416:GLU:HG2	1.48	0.95
3:N:121:PHE:HB3	6:D:414:LEU:HD21	1.48	0.94
1:A:446:ILE:HD11	1:A:465:VAL:HB	1.50	0.93
3:N:78:PRO:HG3	6:D:417:LYS:HB2	1.53	0.91

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	609/641 (95%)	595 (98%)	14 (2%)	0	100	100
2	S	140/142 (99%)	135 (96%)	5 (4%)	0	100	100
3	N	160/186 (86%)	154 (96%)	6 (4%)	0	100	100
4	B	470/615 (76%)	452 (96%)	17 (4%)	1 (0%)	47	82
5	M	116/135 (86%)	109 (94%)	6 (5%)	1 (1%)	17	55
6	D	12/62 (19%)	8 (67%)	4 (33%)	0	100	100
All	All	1507/1781 (85%)	1453 (96%)	52 (4%)	2 (0%)	51	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	B	550	GLU
5	M	62	ASN

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	536/562 (95%)	517 (96%)	19 (4%)	36	71
2	S	131/131 (100%)	128 (98%)	3 (2%)	50	80
3	N	143/156 (92%)	137 (96%)	6 (4%)	30	66
4	B	430/553 (78%)	416 (97%)	14 (3%)	38	73
5	M	105/117 (90%)	101 (96%)	4 (4%)	33	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
6	D	14/52 (27%)	13 (93%)	1 (7%)	14 46
All	All	1359/1571 (86%)	1312 (96%)	47 (4%)	36 71

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	S	55	PHE
3	N	183	TRP
5	M	22	ASP
3	N	55	CYS
3	N	184	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
3	N	40	HIS
3	N	199	HIS

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 ⓘ

There are no monosaccharides 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 [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	611/641 (95%)	-0.07	9 (1%) 73 46	56, 96, 135, 159	0
2	S	142/142 (100%)	-0.17	0 100 100	52, 80, 125, 148	0
3	N	164/186 (88%)	0.31	6 (3%) 41 17	66, 92, 187, 203	0
4	B	480/615 (78%)	0.72	69 (14%) 2 1	57, 155, 199, 248	0
5	M	120/135 (88%)	0.74	16 (13%) 3 1	120, 158, 191, 203	0
6	D	14/62 (22%)	0.99	3 (21%) 0 0	133, 178, 195, 240	0
All	All	1531/1781 (85%)	0.28	103 (6%) 17 5	52, 110, 185, 248	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	274	SER	8.5
3	N	40	HIS	7.3
4	B	118	ILE	6.2
4	B	281	LEU	6.0
4	B	365	VAL	5.7

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.