



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 25, 2024 – 01:57 AM EDT

PDB ID : 6HP3
Title : ARBITRIUM PEPTIDE RECEPTOR FROM SPBETA PHAGE
Authors : Marina, A.; Gallego del Sol, F.
Deposited on : 2018-09-19
Resolution : 2.70 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtrriage (Phenix) : 1.13
EDS : 2.37.1
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) : 2.37.1

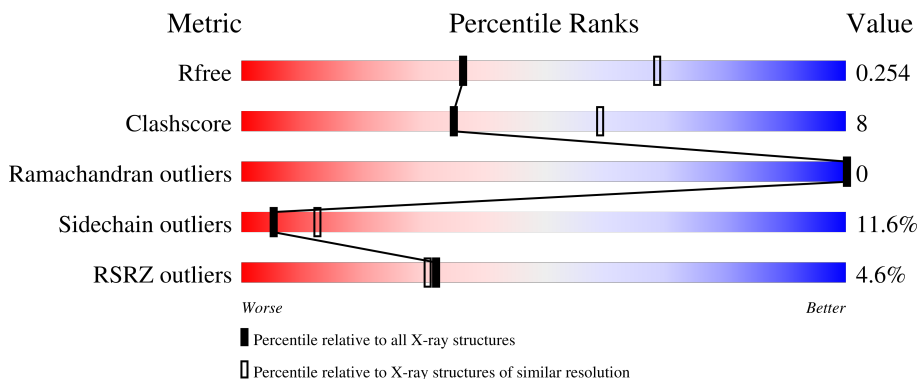
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.70 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.



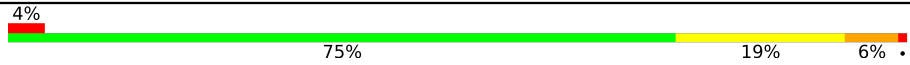

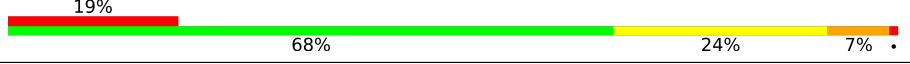
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2808 (2.70-2.70)
Clashscore	141614	3122 (2.70-2.70)
Ramachandran outliers	138981	3069 (2.70-2.70)
Sidechain outliers	138945	3069 (2.70-2.70)
RSRZ outliers	127900	2737 (2.70-2.70)

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 $\leq 5\%$. 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	386	 79% 14% 6%
1	B	386	 76% 17% 5% •
1	C	386	 77% 16% 6% •
1	D	386	 74% 19% 5% •
1	E	386	 75% 18% 7% •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	386	
1	G	386	
1	H	386	

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 25680 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 SPBc2 prophage-derived uncharacterized protein YopK.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	386	3183	2031	528	603	6	15	0	0	0
1	B	386	3183	2031	528	603	6	15	0	0	0
1	C	386	3183	2031	528	603	6	15	0	0	0
1	D	386	3183	2031	528	603	6	15	0	0	0
1	E	386	3188	2035	528	603	6	16	0	1	0
1	F	386	3183	2031	528	603	6	15	0	0	0
1	G	386	3183	2031	528	603	6	15	0	0	0
1	H	386	3183	2031	528	603	6	15	0	0	0

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	44	Total 44	O 44	0	0
2	B	27	Total 27	O 27	0	0
2	C	36	Total 36	O 36	0	0
2	D	29	Total 29	O 29	0	0
2	E	20	Total 20	O 20	0	0
2	F	21	Total 21	O 21	0	0

Continued on next page...

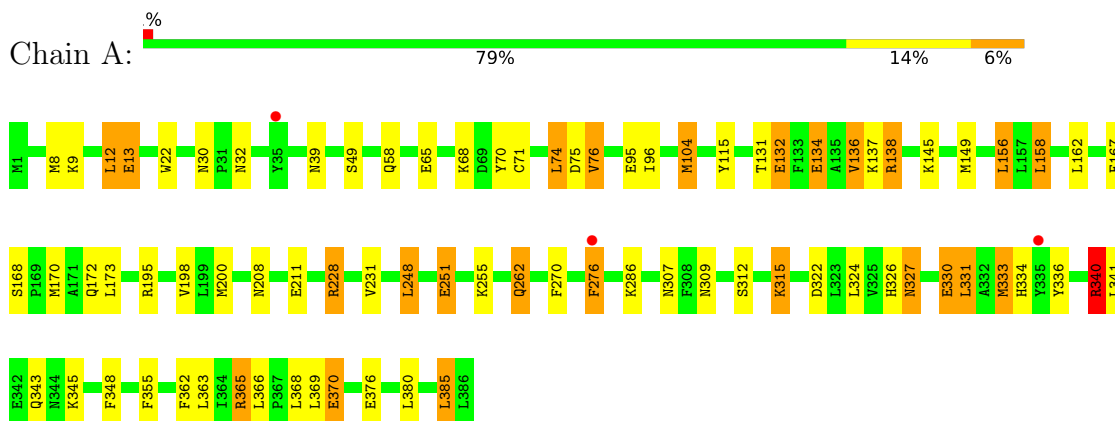
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	G	28	Total	O	0	0
			28	28		
2	H	6	Total	O	0	0
			6	6		

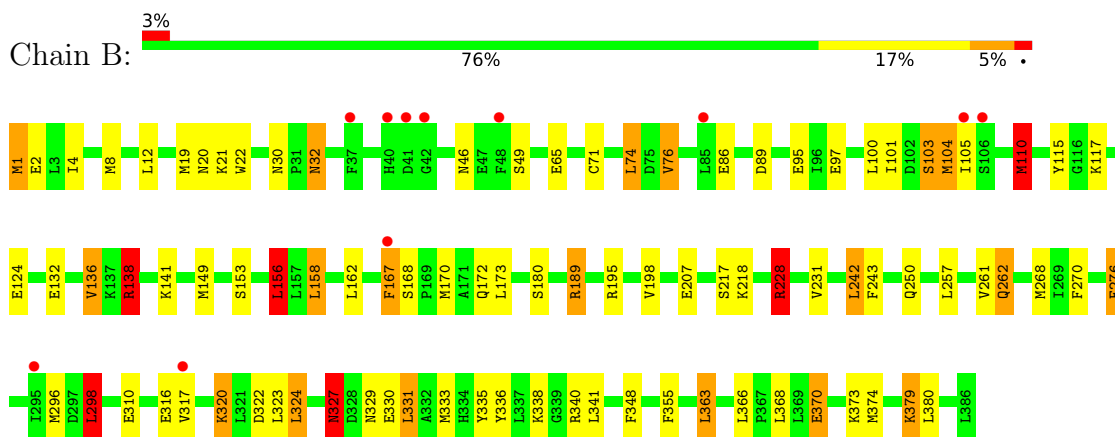
3 Residue-property plots [i](#)

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.

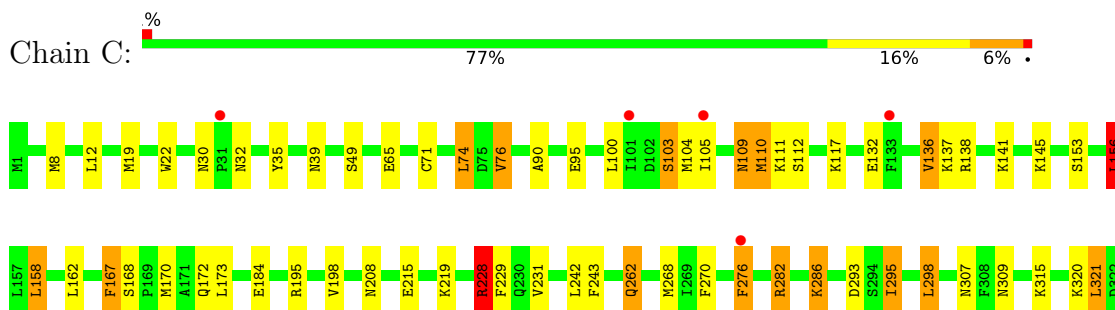
- Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK



- Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK

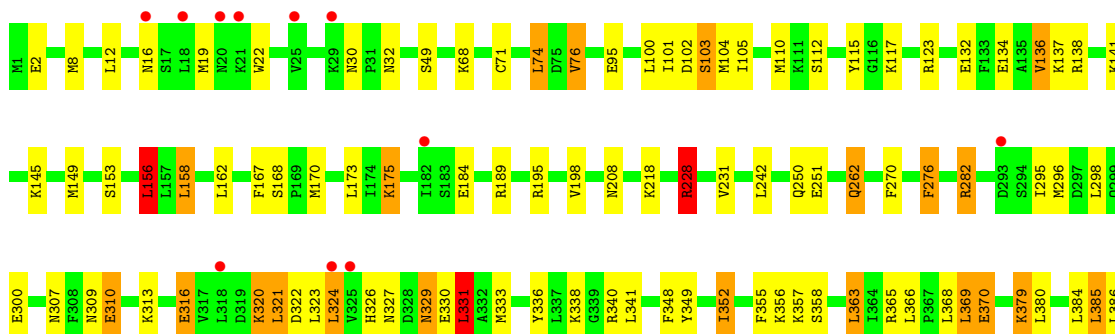
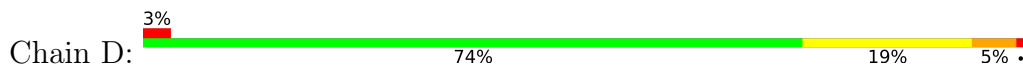


- Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK

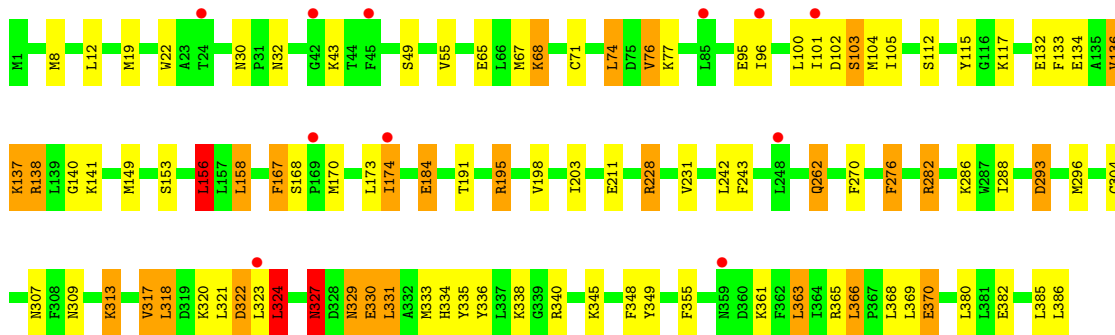
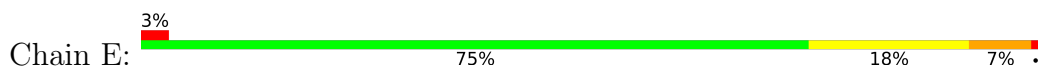




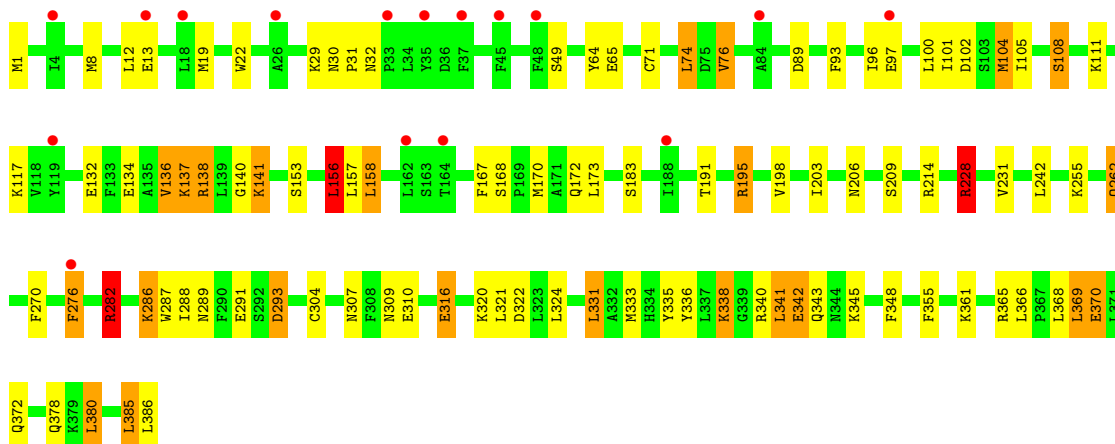
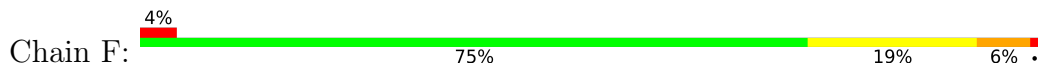
- Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK



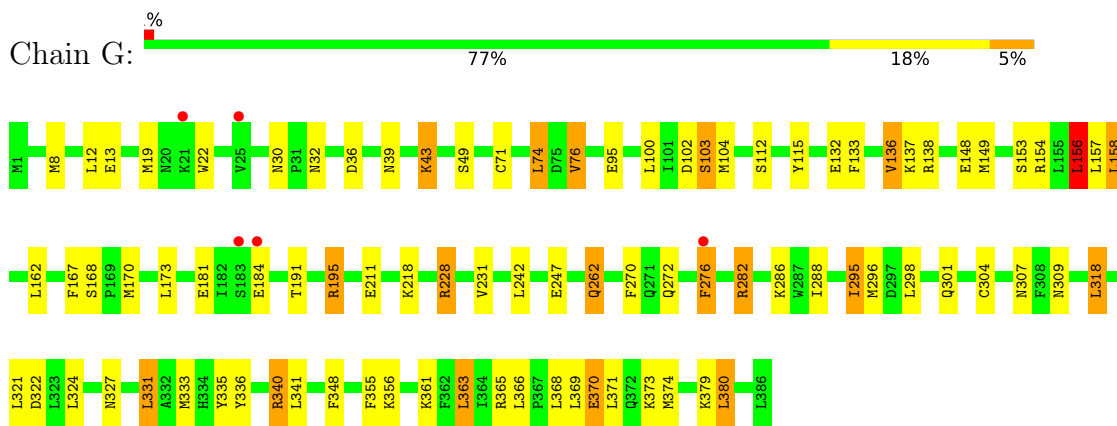
- Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK



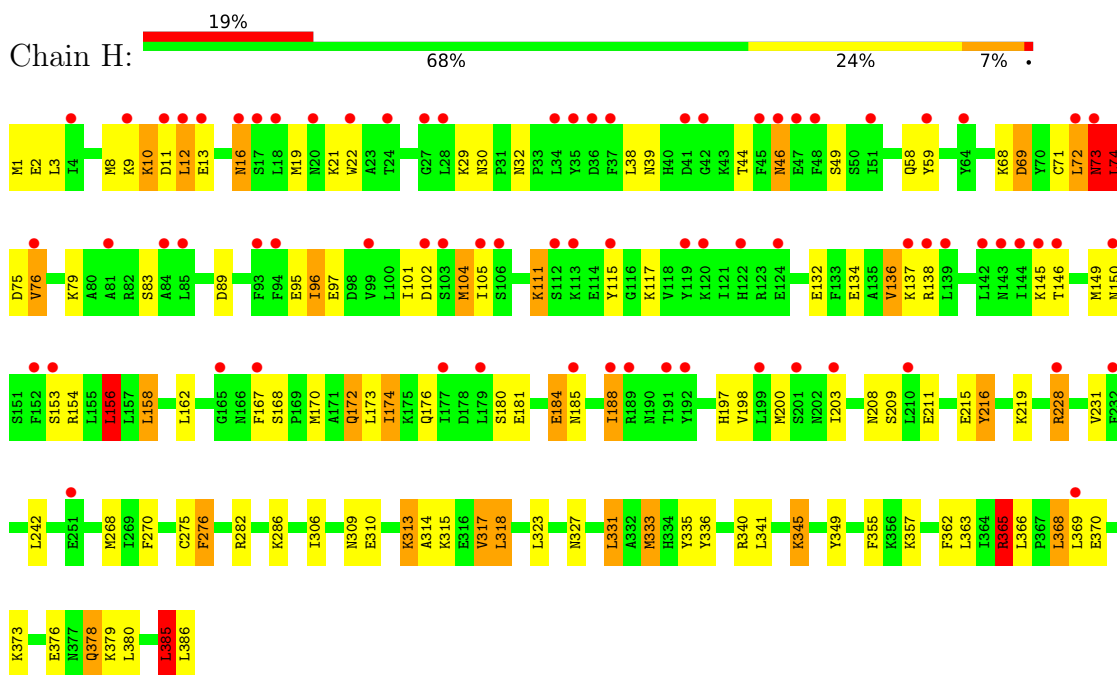
- Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK



- Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK



● Molecule 1: SPBc2 prophage-derived uncharacterized protein YopK



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	78.91Å 251.20Å 95.08Å 90.00° 90.89° 90.00°	Depositor
Resolution (Å)	125.60 – 2.70 95.07 – 2.70	Depositor EDS
% Data completeness (in resolution range)	99.9 (125.60-2.70) 99.9 (95.07-2.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.63 (at 2.69Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.219 , 0.250 0.226 , 0.254	Depositor DCC
R_{free} test set	4902 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å ²)	53.1	Xtrriage
Anisotropy	0.600	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 39.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.028 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	25680	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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	0.93	6/3226 (0.2%)	1.07	18/4315 (0.4%)
1	B	0.89	5/3226 (0.2%)	1.08	22/4315 (0.5%)
1	C	0.88	2/3226 (0.1%)	1.11	26/4315 (0.6%)
1	D	0.90	4/3226 (0.1%)	1.08	22/4315 (0.5%)
1	E	0.92	4/3234 (0.1%)	1.12	22/4325 (0.5%)
1	F	0.86	3/3226 (0.1%)	1.07	22/4315 (0.5%)
1	G	0.90	2/3226 (0.1%)	1.11	22/4315 (0.5%)
1	H	0.99	8/3226 (0.2%)	1.15	28/4315 (0.6%)
All	All	0.91	34/25816 (0.1%)	1.10	182/34530 (0.5%)

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	E	0	1
1	H	0	1
All	All	0	2

The worst 5 of 34 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	35	TYR	CG-CD1	8.72	1.50	1.39
1	H	13	GLU	CD-OE2	8.27	1.34	1.25
1	E	276	PHE	CG-CD1	7.69	1.50	1.38
1	H	16	ASN	CG-ND2	7.40	1.51	1.32
1	H	276	PHE	CG-CD1	7.31	1.49	1.38

The worst 5 of 182 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	228	ARG	NE-CZ-NH2	-11.86	114.37	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	E	102	ASP	CB-CG-OD2	-11.52	107.94	118.30
1	G	340	ARG	NE-CZ-NH2	11.10	125.85	120.30
1	E	321	LEU	CB-CG-CD2	10.11	128.19	111.00
1	H	156	LEU	CA-CB-CG	10.11	138.55	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	167	PHE	Sidechain
1	H	16	ASN	Sidechain

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	3183	0	3179	51	0
1	B	3183	0	3179	44	1
1	C	3183	0	3179	41	0
1	D	3183	0	3179	48	0
1	E	3188	0	3188	62	0
1	F	3183	0	3179	54	0
1	G	3183	0	3179	47	0
1	H	3183	0	3179	90	1
2	A	44	0	0	3	0
2	B	27	0	0	2	0
2	C	36	0	0	1	0
2	D	29	0	0	2	0
2	E	20	0	0	1	0
2	F	21	0	0	2	0
2	G	28	0	0	0	0
2	H	6	0	0	0	0
All	All	25680	0	25441	416	1

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 8.

The worst 5 of 416 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:H:170:MSE:O	1:H:174:ILE:HD11	1.40	1.21
1:A:167:PHE:HA	1:A:170:MSE:HE3	1.31	1.12
1:F:372:GLN:HE21	1:F:378:GLN:HG2	1.10	1.11
1:H:170:MSE:O	1:H:174:ILE:CD1	2.00	1.08
1:H:167:PHE:HA	1:H:170:MSE:HE3	1.34	1.06

All (1) 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)
1:B:46:ASN:ND2	1:H:378:GLN:OE1[1_455]	1.97	0.23

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	384/386 (100%)	378 (98%)	6 (2%)	0	100	100
1	B	384/386 (100%)	378 (98%)	6 (2%)	0	100	100
1	C	384/386 (100%)	376 (98%)	8 (2%)	0	100	100
1	D	384/386 (100%)	378 (98%)	6 (2%)	0	100	100
1	E	385/386 (100%)	377 (98%)	8 (2%)	0	100	100
1	F	384/386 (100%)	378 (98%)	6 (2%)	0	100	100
1	G	384/386 (100%)	377 (98%)	7 (2%)	0	100	100
1	H	384/386 (100%)	377 (98%)	7 (2%)	0	100	100
All	All	3073/3088 (100%)	3019 (98%)	54 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	360/345 (104%)	324 (90%)	36 (10%)	7	18
1	B	360/345 (104%)	316 (88%)	44 (12%)	5	11
1	C	360/345 (104%)	321 (89%)	39 (11%)	6	15
1	D	360/345 (104%)	312 (87%)	48 (13%)	4	9
1	E	361/345 (105%)	319 (88%)	42 (12%)	5	12
1	F	360/345 (104%)	316 (88%)	44 (12%)	5	11
1	G	360/345 (104%)	324 (90%)	36 (10%)	7	18
1	H	360/345 (104%)	314 (87%)	46 (13%)	4	10
All	All	2881/2760 (104%)	2546 (88%)	335 (12%)	5	12

5 of 335 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	242	LEU
1	G	370	GLU
1	F	310	GLU
1	G	102	ASP
1	H	104	MSE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 65 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	46	ASN
1	H	190	ASN
1	C	377	ASN
1	C	359	ASN
1	H	265	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](#)

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 [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	371/386 (96%)	0.08	3 (0%) 86 87	33, 51, 81, 100	0
1	B	371/386 (96%)	0.35	11 (2%) 50 51	26, 62, 106, 142	0
1	C	371/386 (96%)	0.22	5 (1%) 77 78	34, 57, 91, 117	0
1	D	371/386 (96%)	0.26	11 (2%) 50 51	34, 58, 99, 144	0
1	E	371/386 (96%)	0.40	11 (2%) 50 51	39, 65, 90, 108	0
1	F	371/386 (96%)	0.32	16 (4%) 35 33	37, 69, 109, 131	0
1	G	371/386 (96%)	0.20	5 (1%) 77 78	31, 57, 95, 114	0
1	H	371/386 (96%)	1.07	74 (19%) 1 0	41, 79, 135, 183	0
All	All	2968/3088 (96%)	0.36	136 (4%) 32 31	26, 61, 106, 183	0

The worst 5 of 136 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	13	GLU	8.8
1	H	16	ASN	7.4
1	H	12	LEU	6.8
1	H	81	ALA	6.3
1	H	85	LEU	6.3

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.