



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2023 – 09:42 AM EDT

PDB ID : 6WG6
Title : Crystal structure of human SMC1-SMC3 hinge domain heterodimer in north-open conformation
Authors : Shi, Z.B.; Yu, H.
Deposited on : 2020-04-04
Resolution : 3.54 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
Xtriage (Phenix) : 1.13
EDS : 2.35.1
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) : 2.35.1

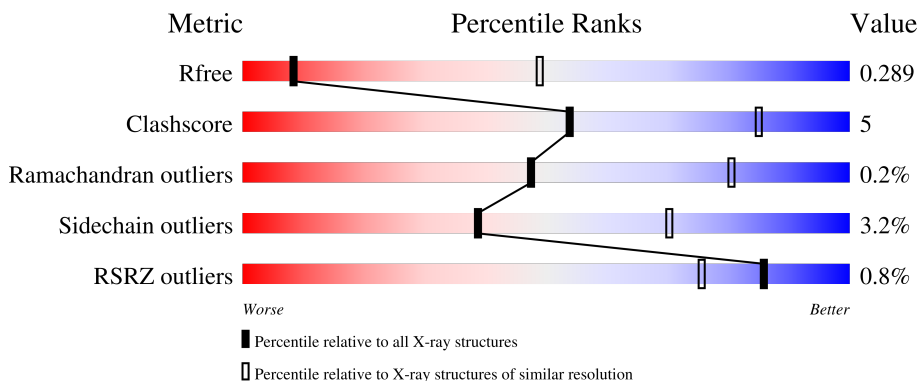
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.54 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	1028 (3.60-3.48)
Clashscore	141614	1109 (3.60-3.48)
Ramachandran outliers	138981	1073 (3.60-3.48)
Sidechain outliers	138945	1074 (3.60-3.48)
RSRZ outliers	127900	1079 (3.62-3.46)

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	233	 64% 14% 22%
1	C	233	 65% 12% 22%
1	E	233	 67% 9% 24%
1	G	233	 62% 14% 24%
1	I	233	 61% 12% 26%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	233	<p>6% 67% 8% 25%</p>
2	B	256	<p>75% 14% 8%</p>
2	D	256	<p>79% 12% 8%</p>
2	F	256	<p>77% 15% 8%</p>
2	H	256	<p>79% 13% 8%</p>
2	J	256	<p>79% 14% 7%</p>
2	L	256	<p>79% 14% 7%</p>
3	M	10	<p>20% 80%</p>
3	N	10	<p>20% 80%</p>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20038 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 Structural maintenance of chromosomes protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	182	1446	910	262	268	6	0	0	0
1	C	182	1446	910	262	268	6	0	0	0
1	E	178	1409	886	256	261	6	0	0	0
1	G	178	1409	886	256	261	6	0	0	0
1	I	173	1376	866	250	254	6	0	0	0
1	K	175	1390	875	253	256	6	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	470	GLY	-	expression tag	UNP G8JLG1
A	471	PRO	-	expression tag	UNP G8JLG1
C	470	GLY	-	expression tag	UNP G8JLG1
C	471	PRO	-	expression tag	UNP G8JLG1
E	470	GLY	-	expression tag	UNP G8JLG1
E	471	PRO	-	expression tag	UNP G8JLG1
G	470	GLY	-	expression tag	UNP G8JLG1
G	471	PRO	-	expression tag	UNP G8JLG1
I	470	GLY	-	expression tag	UNP G8JLG1
I	471	PRO	-	expression tag	UNP G8JLG1
K	470	GLY	-	expression tag	UNP G8JLG1
K	471	PRO	-	expression tag	UNP G8JLG1

- Molecule 2 is a protein called Structural maintenance of chromosomes protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	236	1908	1194	343	361	10	0	0	0
2	D	235	1902	1189	345	358	10	0	0	0
2	F	235	1899	1189	342	358	10	0	0	0
2	H	236	1908	1194	343	361	10	0	0	0
2	J	238	1927	1206	348	363	10	0	0	0
2	L	239	1938	1212	352	364	10	0	0	0

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	457	GLY	-	expression tag	UNP Q9UQE7
B	458	PRO	-	expression tag	UNP Q9UQE7
B	459	LEU	-	expression tag	UNP Q9UQE7
B	460	GLY	-	expression tag	UNP Q9UQE7
B	461	SER	-	expression tag	UNP Q9UQE7
B	462	GLY	-	expression tag	UNP Q9UQE7
B	463	ARG	-	expression tag	UNP Q9UQE7
B	464	PRO	-	expression tag	UNP Q9UQE7
D	457	GLY	-	expression tag	UNP Q9UQE7
D	458	PRO	-	expression tag	UNP Q9UQE7
D	459	LEU	-	expression tag	UNP Q9UQE7
D	460	GLY	-	expression tag	UNP Q9UQE7
D	461	SER	-	expression tag	UNP Q9UQE7
D	462	GLY	-	expression tag	UNP Q9UQE7
D	463	ARG	-	expression tag	UNP Q9UQE7
D	464	PRO	-	expression tag	UNP Q9UQE7
F	457	GLY	-	expression tag	UNP Q9UQE7
F	458	PRO	-	expression tag	UNP Q9UQE7
F	459	LEU	-	expression tag	UNP Q9UQE7
F	460	GLY	-	expression tag	UNP Q9UQE7
F	461	SER	-	expression tag	UNP Q9UQE7
F	462	GLY	-	expression tag	UNP Q9UQE7
F	463	ARG	-	expression tag	UNP Q9UQE7
F	464	PRO	-	expression tag	UNP Q9UQE7
H	457	GLY	-	expression tag	UNP Q9UQE7
H	458	PRO	-	expression tag	UNP Q9UQE7
H	459	LEU	-	expression tag	UNP Q9UQE7
H	460	GLY	-	expression tag	UNP Q9UQE7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
H	461	SER	-	expression tag	UNP Q9UQE7
H	462	GLY	-	expression tag	UNP Q9UQE7
H	463	ARG	-	expression tag	UNP Q9UQE7
H	464	PRO	-	expression tag	UNP Q9UQE7
J	457	GLY	-	expression tag	UNP Q9UQE7
J	458	PRO	-	expression tag	UNP Q9UQE7
J	459	LEU	-	expression tag	UNP Q9UQE7
J	460	GLY	-	expression tag	UNP Q9UQE7
J	461	SER	-	expression tag	UNP Q9UQE7
J	462	GLY	-	expression tag	UNP Q9UQE7
J	463	ARG	-	expression tag	UNP Q9UQE7
J	464	PRO	-	expression tag	UNP Q9UQE7
L	457	GLY	-	expression tag	UNP Q9UQE7
L	458	PRO	-	expression tag	UNP Q9UQE7
L	459	LEU	-	expression tag	UNP Q9UQE7
L	460	GLY	-	expression tag	UNP Q9UQE7
L	461	SER	-	expression tag	UNP Q9UQE7
L	462	GLY	-	expression tag	UNP Q9UQE7
L	463	ARG	-	expression tag	UNP Q9UQE7
L	464	PRO	-	expression tag	UNP Q9UQE7

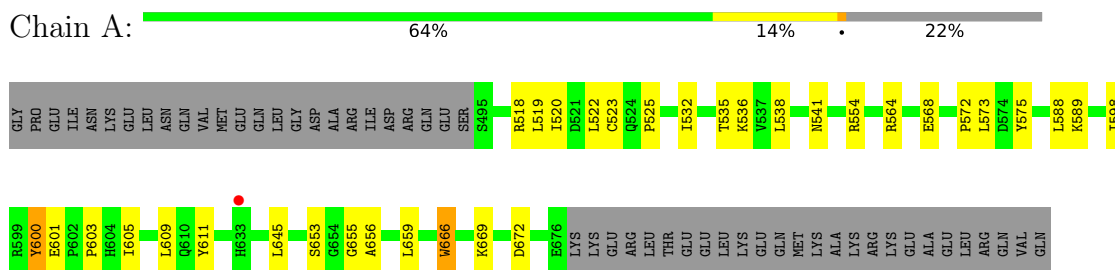
- Molecule 3 is a DNA chain called poly(dT).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
3	M	2	Total	C	N	O	P	0	0	0
			40	20	4	14	2			
3	N	2	Total	C	N	O	P	0	0	0
			40	20	4	14	2			

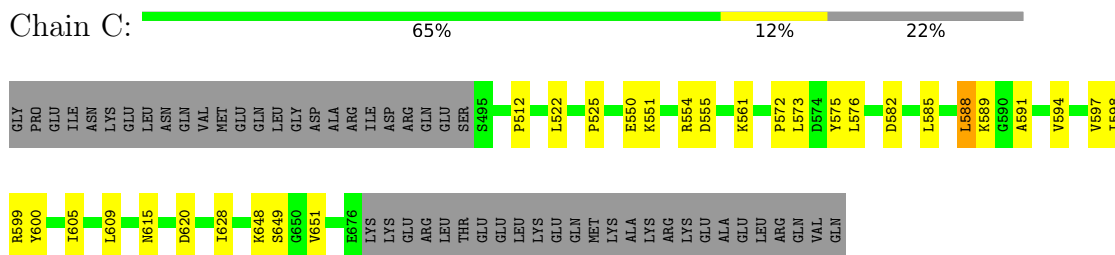
3 Residue-property plots

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.

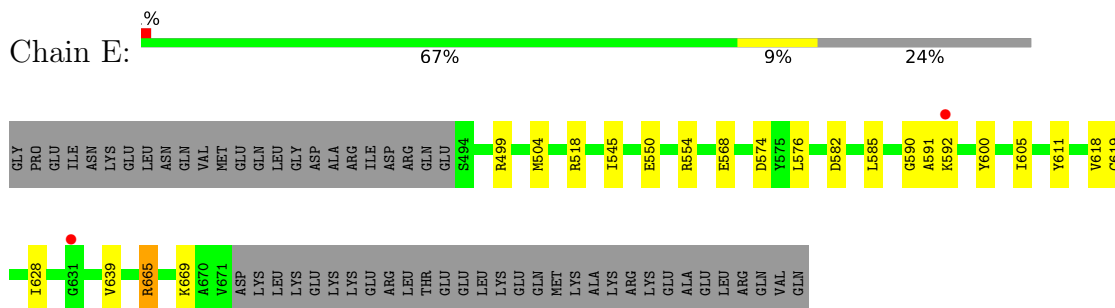
- Molecule 1: Structural maintenance of chromosomes protein



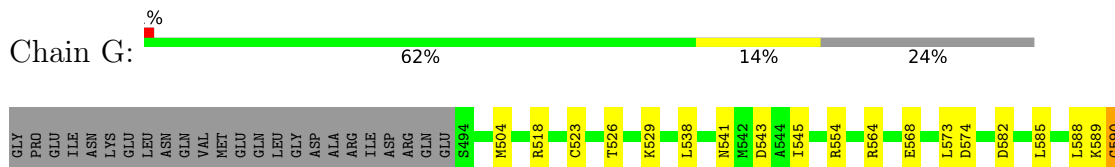
- Molecule 1: Structural maintenance of chromosomes protein

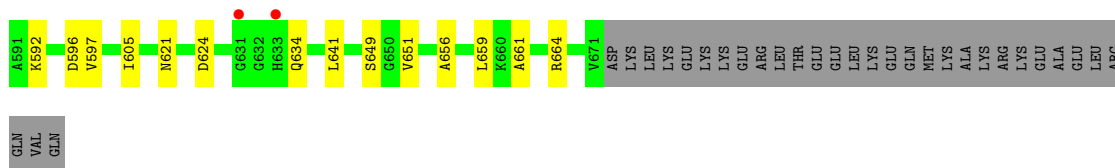


- Molecule 1: Structural maintenance of chromosomes protein

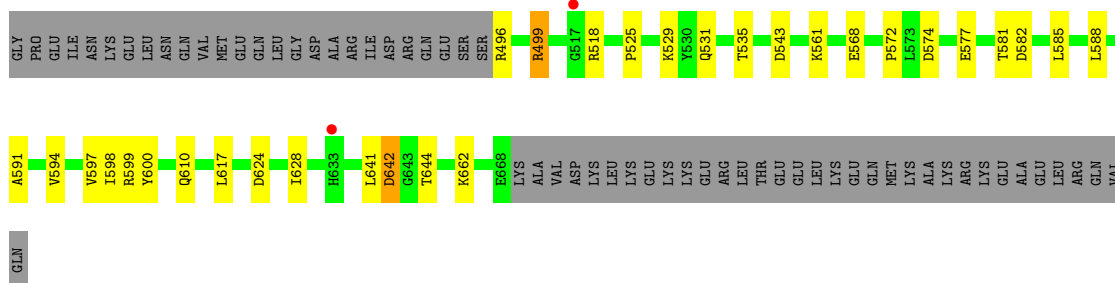


- Molecule 1: Structural maintenance of chromosomes protein

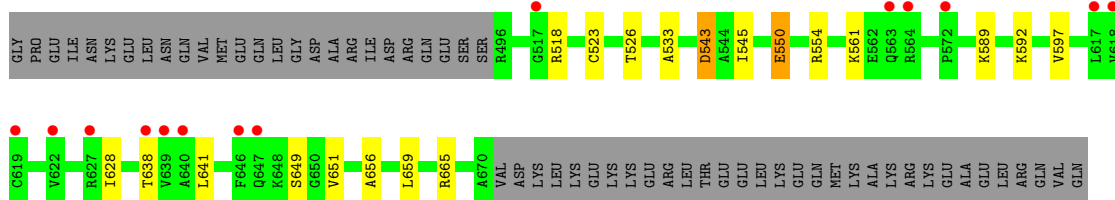




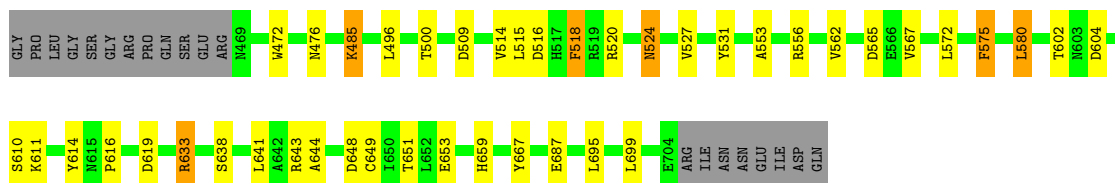
• Molecule 1: Structural maintenance of chromosomes protein



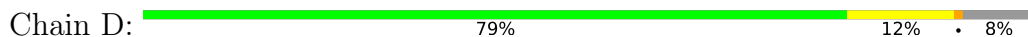
• Molecule 1: Structural maintenance of chromosomes protein




• Molecule 2: Structural maintenance of chromosomes protein 3

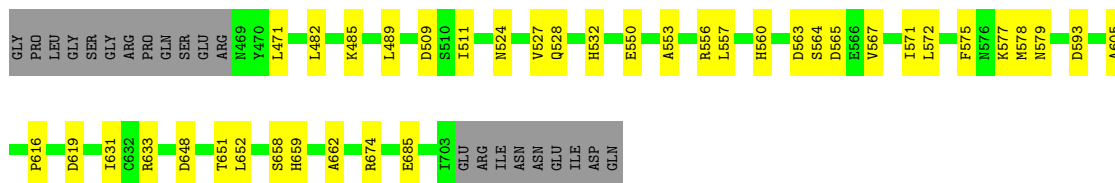


• Molecule 2: Structural maintenance of chromosomes protein 3




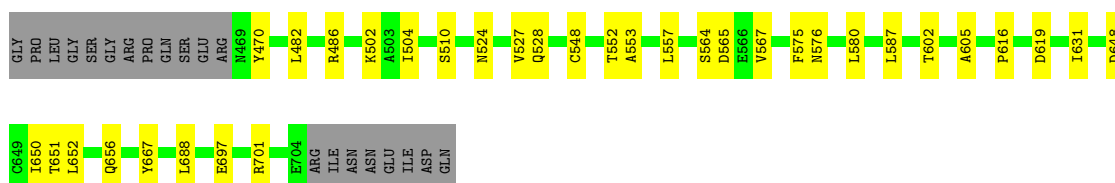
- Molecule 2: Structural maintenance of chromosomes protein 3

Chain F:  77% 15% 8%




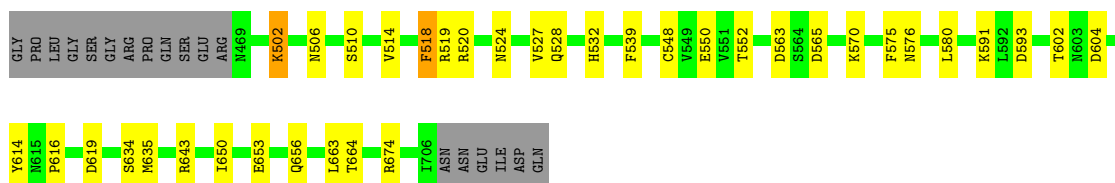
- Molecule 2: Structural maintenance of chromosomes protein 3

Chain H:  79% 13% 8%




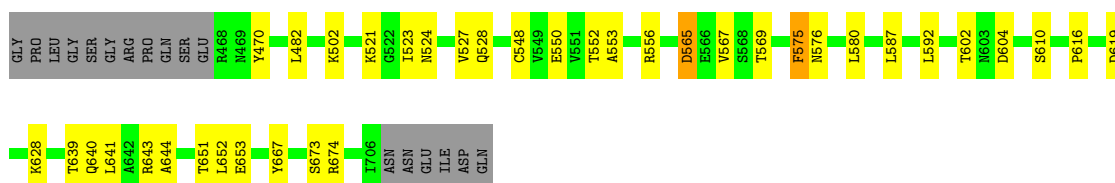
- Molecule 2: Structural maintenance of chromosomes protein 3

Chain J:  79% 14% 7%



- Molecule 2: Structural maintenance of chromosomes protein 3

Chain L:  79% 14% 7%



- Molecule 3: poly(dT)

Chain M:  20% 80%



- Molecule 3: poly(dT)

Chain N:  20% 80%

T3	DT	DT	DT	DT	DT	DT	DT	DT
T4	DT	DT	DT	DT	DT	DT	DT	DT

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	123.98Å 145.02Å 315.95Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.50 – 3.54 49.50 – 3.54	Depositor EDS
% Data completeness (in resolution range)	87.3 (49.50-3.54) 73.5 (49.50-3.54)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.69 (at 3.57Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.252 , 0.288 0.252 , 0.289	Depositor DCC
R_{free} test set	2000 reflections (3.26%)	wwPDB-VP
Wilson B-factor (Å ²)	44.1	Xtrriage
Anisotropy	0.495	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 29.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	20038	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 45.84 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2465e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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	0.29	0/1467	0.54	0/1968
1	C	0.27	0/1467	0.52	1/1968 (0.1%)
1	E	0.26	0/1430	0.47	0/1920
1	G	0.26	0/1430	0.50	0/1920
1	I	0.29	0/1397	0.53	0/1876
1	K	0.26	0/1411	0.49	0/1894
2	B	0.28	0/1939	0.50	1/2609 (0.0%)
2	D	0.27	0/1933	0.48	0/2600
2	F	0.27	0/1930	0.47	0/2597
2	H	0.27	0/1939	0.47	0/2609
2	J	0.26	0/1958	0.45	0/2634
2	L	0.26	0/1969	0.45	0/2648
3	M	0.87	0/43	1.60	0/64
3	N	1.07	0/43	1.91	1/64 (1.6%)
All	All	0.28	0/20356	0.50	3/27371 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	N	3	DT	O4'-C1'-N1	7.79	113.45	108.00
2	B	580	LEU	CA-CB-CG	5.65	128.30	115.30
1	C	588	LEU	CA-CB-CG	5.45	127.83	115.30

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	1446	0	1497	19	0
1	C	1446	0	1497	21	0
1	E	1409	0	1455	11	0
1	G	1409	0	1455	18	0
1	I	1376	0	1418	18	0
1	K	1390	0	1436	10	0
2	B	1908	0	1897	27	0
2	D	1902	0	1893	23	0
2	F	1899	0	1891	19	0
2	H	1908	0	1897	17	0
2	J	1927	0	1921	22	0
2	L	1938	0	1934	23	0
3	M	40	0	25	2	0
3	N	40	0	25	1	0
All	All	20038	0	20241	213	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 5.

The worst 5 of 213 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)
2:J:514:VAL:O	2:J:518:PHE:HB2	1.94	0.67
1:K:561:LYS:HA	2:L:667:TYR:HB2	1.77	0.65
2:H:486:ARG:HG2	2:H:688:LEU:HD11	1.78	0.65
2:D:532:HIS:NE2	2:D:563:ASP:OD1	2.31	0.64
2:L:575:PHE:HA	2:L:580:LEU:HD23	1.81	0.63

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	180/233 (77%)	171 (95%)	7 (4%)	2 (1%)	14	54
1	C	180/233 (77%)	170 (94%)	10 (6%)	0	100	100
1	E	176/233 (76%)	168 (96%)	7 (4%)	1 (1%)	25	65
1	G	176/233 (76%)	165 (94%)	10 (6%)	1 (1%)	25	65
1	I	171/233 (73%)	160 (94%)	11 (6%)	0	100	100
1	K	173/233 (74%)	163 (94%)	10 (6%)	0	100	100
2	B	234/256 (91%)	227 (97%)	6 (3%)	1 (0%)	34	71
2	D	233/256 (91%)	226 (97%)	6 (3%)	1 (0%)	34	71
2	F	233/256 (91%)	228 (98%)	5 (2%)	0	100	100
2	H	234/256 (91%)	225 (96%)	9 (4%)	0	100	100
2	J	236/256 (92%)	228 (97%)	8 (3%)	0	100	100
2	L	237/256 (93%)	231 (98%)	6 (2%)	0	100	100
All	All	2463/2934 (84%)	2362 (96%)	95 (4%)	6 (0%)	47	80

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	646	THR
1	A	601	GLU
2	B	524	ASN
1	A	655	GLY
1	G	590	GLY

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	155/201 (77%)	149 (96%)	6 (4%)	32	65
1	C	155/201 (77%)	153 (99%)	2 (1%)	69	87
1	E	151/201 (75%)	146 (97%)	5 (3%)	38	69

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	151/201 (75%)	146 (97%)	5 (3%)	38	69
1	I	147/201 (73%)	141 (96%)	6 (4%)	30	64
1	K	148/201 (74%)	143 (97%)	5 (3%)	37	69
2	B	206/223 (92%)	200 (97%)	6 (3%)	42	72
2	D	205/223 (92%)	198 (97%)	7 (3%)	37	69
2	F	205/223 (92%)	198 (97%)	7 (3%)	37	69
2	H	206/223 (92%)	200 (97%)	6 (3%)	42	72
2	J	208/223 (93%)	200 (96%)	8 (4%)	33	66
2	L	209/223 (94%)	203 (97%)	6 (3%)	42	72
All	All	2146/2544 (84%)	2077 (97%)	69 (3%)	39	70

5 of 69 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	J	593	ASP
1	K	543	ASP
2	L	565	ASP
1	E	600	TYR
1	E	592	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	531	GLN
2	D	528	GLN
2	F	493	GLN
2	F	560	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 [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	182/233 (78%)	-0.15	1 (0%) 91 84	22, 70, 117, 140	0
1	C	182/233 (78%)	-0.20	0 100 100	25, 66, 113, 132	0
1	E	178/233 (76%)	0.06	2 (1%) 80 69	40, 98, 138, 155	0
1	G	178/233 (76%)	0.07	2 (1%) 80 69	50, 100, 136, 147	0
1	I	173/233 (74%)	0.36	2 (1%) 79 67	73, 128, 160, 177	0
1	K	175/233 (75%)	0.63	14 (8%) 12 9	107, 146, 175, 189	0
2	B	236/256 (92%)	-0.35	0 100 100	26, 52, 108, 154	0
2	D	235/256 (91%)	-0.33	0 100 100	20, 49, 105, 141	0
2	F	235/256 (91%)	-0.40	0 100 100	22, 48, 90, 113	0
2	H	236/256 (92%)	-0.45	0 100 100	17, 47, 93, 118	0
2	J	238/256 (92%)	-0.30	0 100 100	41, 73, 114, 142	0
2	L	239/256 (93%)	-0.28	0 100 100	44, 79, 113, 143	0
3	M	2/10 (20%)	-0.28	0 100 100	86, 86, 86, 113	0
3	N	2/10 (20%)	-0.12	0 100 100	91, 91, 91, 120	0
All	All	2491/2954 (84%)	-0.15	21 (0%) 86 75	17, 72, 147, 189	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	I	517	GLY	5.2
1	I	633	HIS	3.8
1	K	646	PHE	3.1
1	K	564	ARG	3.1
1	K	619	CYS	2.9

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.