wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 09:41 PM EDT

PDB ID	:	5ME3
Title	:	Structure of the Scc2 C-terminus
Authors	:	Chao, W.C.H.; Singleton, M.R.
Deposited on	:	2016-11-14
Resolution	:	2.85 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 (i)) were used in the production of this report:

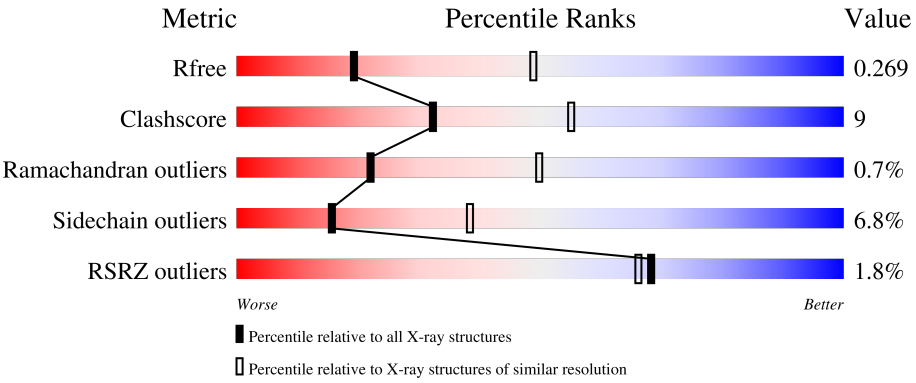
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.85 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	3168 (2.90-2.82)
Clashscore	141614	3438 (2.90-2.82)
Ramachandran outliers	138981	3348 (2.90-2.82)
Sidechain outliers	138945	3351 (2.90-2.82)
RSRZ outliers	127900	3103 (2.90-2.82)

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	1143	2% 66% 17% · 15%
1	B	1143	% 63% 19% · 16%
2	X	17	100%
3	Y	27	63% 37%
4	W	11	91% 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	Z	11	100%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 16017 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 Sister chromatid cohesion protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	973	Total	C	N	O	S	0	0	0
			7931	5085	1341	1460	45			
1	B	960	Total	C	N	O	S	0	0	0
			7811	5009	1316	1442	44			

There are 82 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	377	MET	-	initiating methionine	UNP Q750S2
A	1480	LYS	-	expression tag	UNP Q750S2
A	1481	SER	-	expression tag	UNP Q750S2
A	1482	SER	-	expression tag	UNP Q750S2
A	1483	ILE	-	expression tag	UNP Q750S2
A	1484	PRO	-	expression tag	UNP Q750S2
A	1485	GLU	-	expression tag	UNP Q750S2
A	1486	ASN	-	expression tag	UNP Q750S2
A	1487	LEU	-	expression tag	UNP Q750S2
A	1488	TYR	-	expression tag	UNP Q750S2
A	1489	PHE	-	expression tag	UNP Q750S2
A	1490	GLN	-	expression tag	UNP Q750S2
A	1491	SER	-	expression tag	UNP Q750S2
A	1492	TRP	-	expression tag	UNP Q750S2
A	1493	SER	-	expression tag	UNP Q750S2
A	1494	HIS	-	expression tag	UNP Q750S2
A	1495	PRO	-	expression tag	UNP Q750S2
A	1496	GLN	-	expression tag	UNP Q750S2
A	1497	PHE	-	expression tag	UNP Q750S2
A	1498	GLU	-	expression tag	UNP Q750S2
A	1499	LYS	-	expression tag	UNP Q750S2
A	1500	GLY	-	expression tag	UNP Q750S2
A	1501	GLY	-	expression tag	UNP Q750S2
A	1502	GLY	-	expression tag	UNP Q750S2
A	1503	SER	-	expression tag	UNP Q750S2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	1504	GLY	-	expression tag	UNP Q750S2
A	1505	GLY	-	expression tag	UNP Q750S2
A	1506	GLY	-	expression tag	UNP Q750S2
A	1507	SER	-	expression tag	UNP Q750S2
A	1508	GLY	-	expression tag	UNP Q750S2
A	1509	GLY	-	expression tag	UNP Q750S2
A	1510	GLY	-	expression tag	UNP Q750S2
A	1511	SER	-	expression tag	UNP Q750S2
A	1512	TRP	-	expression tag	UNP Q750S2
A	1513	SER	-	expression tag	UNP Q750S2
A	1514	HIS	-	expression tag	UNP Q750S2
A	1515	PRO	-	expression tag	UNP Q750S2
A	1516	GLN	-	expression tag	UNP Q750S2
A	1517	PHE	-	expression tag	UNP Q750S2
A	1518	GLU	-	expression tag	UNP Q750S2
A	1519	LYS	-	expression tag	UNP Q750S2
B	377	MET	-	initiating methionine	UNP Q750S2
B	1480	LYS	-	expression tag	UNP Q750S2
B	1481	SER	-	expression tag	UNP Q750S2
B	1482	SER	-	expression tag	UNP Q750S2
B	1483	ILE	-	expression tag	UNP Q750S2
B	1484	PRO	-	expression tag	UNP Q750S2
B	1485	GLU	-	expression tag	UNP Q750S2
B	1486	ASN	-	expression tag	UNP Q750S2
B	1487	LEU	-	expression tag	UNP Q750S2
B	1488	TYR	-	expression tag	UNP Q750S2
B	1489	PHE	-	expression tag	UNP Q750S2
B	1490	GLN	-	expression tag	UNP Q750S2
B	1491	SER	-	expression tag	UNP Q750S2
B	1492	TRP	-	expression tag	UNP Q750S2
B	1493	SER	-	expression tag	UNP Q750S2
B	1494	HIS	-	expression tag	UNP Q750S2
B	1495	PRO	-	expression tag	UNP Q750S2
B	1496	GLN	-	expression tag	UNP Q750S2
B	1497	PHE	-	expression tag	UNP Q750S2
B	1498	GLU	-	expression tag	UNP Q750S2
B	1499	LYS	-	expression tag	UNP Q750S2
B	1500	GLY	-	expression tag	UNP Q750S2
B	1501	GLY	-	expression tag	UNP Q750S2
B	1502	GLY	-	expression tag	UNP Q750S2
B	1503	SER	-	expression tag	UNP Q750S2
B	1504	GLY	-	expression tag	UNP Q750S2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1505	GLY	-	expression tag	UNP Q750S2
B	1506	GLY	-	expression tag	UNP Q750S2
B	1507	SER	-	expression tag	UNP Q750S2
B	1508	GLY	-	expression tag	UNP Q750S2
B	1509	GLY	-	expression tag	UNP Q750S2
B	1510	GLY	-	expression tag	UNP Q750S2
B	1511	SER	-	expression tag	UNP Q750S2
B	1512	TRP	-	expression tag	UNP Q750S2
B	1513	SER	-	expression tag	UNP Q750S2
B	1514	HIS	-	expression tag	UNP Q750S2
B	1515	PRO	-	expression tag	UNP Q750S2
B	1516	GLN	-	expression tag	UNP Q750S2
B	1517	PHE	-	expression tag	UNP Q750S2
B	1518	GLU	-	expression tag	UNP Q750S2
B	1519	LYS	-	expression tag	UNP Q750S2

- Molecule 2 is a protein called unassigned sequence of Scc2.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	X	17	Total	C	N	O	0	0	0
			85	51	17	17			

- Molecule 3 is a protein called Scc2 unassigned sequence.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	Y	17	Total	C	N	O	0	0	0
			85	51	17	17			

- Molecule 4 is a protein called Scc2 unassigned sequence.

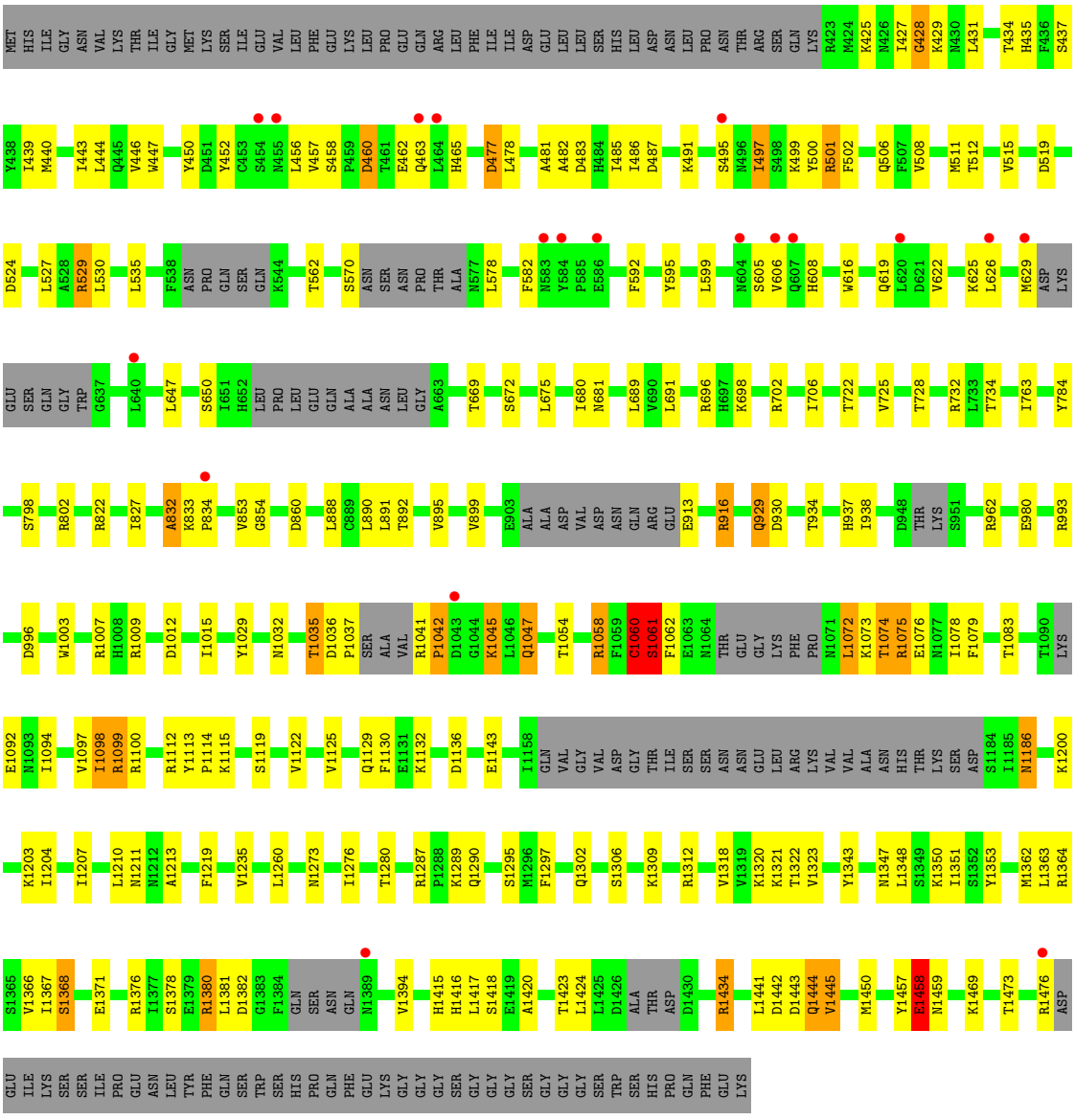
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
4	Z	11	Total	C	N	O	0	0	0
			55	33	11	11			
4	W	10	Total	C	N	O	0	0	0
			50	30	10	10			

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.

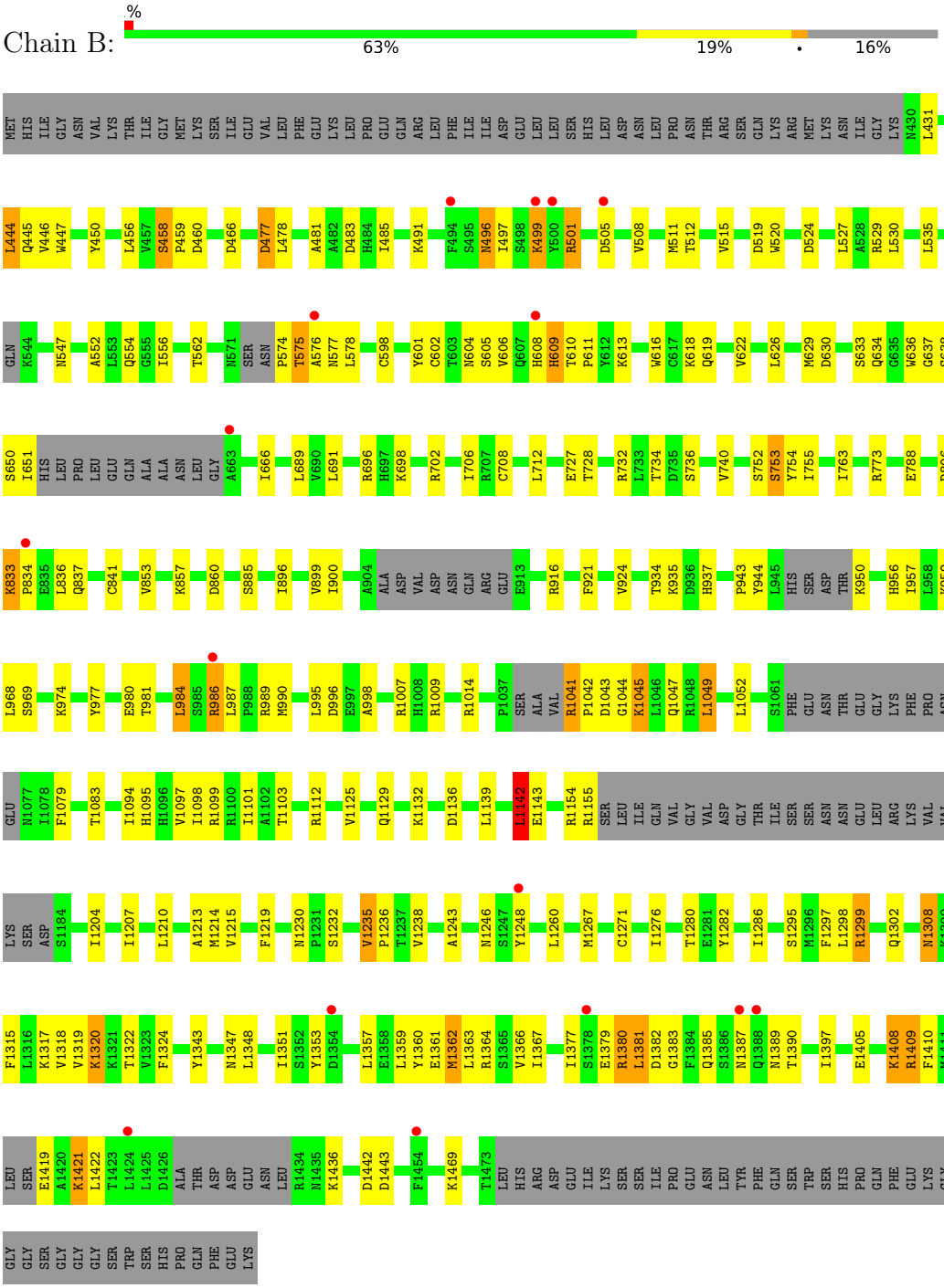
• Molecule 1: Sister chromatid cohesion protein 2

Chain A:



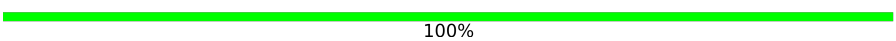
• Molecule 1: Sister chromatid cohesion protein 2

Chain B:



• Molecule 2: unassigned sequence of Scc2

Chain X:



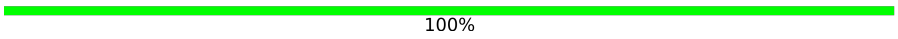
There are no outlier residues recorded for this chain.

• Molecule 3: Scc2 unassigned sequence

Chain Y:

• Molecule 4: Scc2 unassigned sequence

Chain Z:



There are no outlier residues recorded for this chain.

• Molecule 4: Scc2 unassigned sequence

Chain W:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	98.05Å 106.44Å 143.18Å 90.00° 104.52° 90.00°	Depositor
Resolution (Å)	48.81 – 2.85 49.68 – 2.85	Depositor EDS
% Data completeness (in resolution range)	98.7 (48.81-2.85) 98.7 (49.68-2.85)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.21 (at 2.86Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.217 , 0.269 0.217 , 0.269	Depositor DCC
R_{free} test set	3386 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å^2)	90.5	Xtriage
Anisotropy	0.071	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 63.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	16017	wwPDB-VP
Average B, all atoms (Å^2)	93.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% 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	0.27	0/8071	0.50	6/10895 (0.1%)
1	B	0.27	0/7952	0.51	2/10741 (0.0%)
All	All	0.27	0/16023	0.51	8/21636 (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	B	0	2
All	All	0	3

There are no bond length outliers.

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1142	LEU	CA-CB-CG	7.72	133.06	115.30
1	A	1458	GLU	C-N-CA	6.96	139.10	121.70
1	A	1441	LEU	CA-CB-CG	6.88	131.12	115.30
1	B	1049	LEU	CA-CB-CG	6.49	130.23	115.30
1	A	1444	GLN	C-N-CA	6.03	136.78	121.70

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	1060	CYS	Peptide
1	B	458	SER	Peptide
1	B	460	ASP	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	7931	0	8056	136	0
1	B	7811	0	7928	149	0
2	X	85	0	19	0	0
3	Y	85	0	19	0	0
4	W	50	0	12	0	0
4	Z	55	0	13	0	0
All	All	16017	0	16047	282	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 9.

The worst 5 of 282 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:A:1060:CYS:HA	1:A:1061:SER:HB2	1.39	1.01
1:B:1412:PHE:O	1:B:1416:HIS:HB2	1.70	0.91
1:B:1043:ASP:O	1:B:1045:LYS:N	2.07	0.86
1:A:1458:GLU:HB2	1:A:1459:ASN:HB2	1.61	0.83
1:B:501:ARG:NH1	1:B:505:ASP:OD2	2.13	0.80

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	947/1143 (83%)	901 (95%)	39 (4%)	7 (1%)	22 50

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	938/1143 (82%)	899 (96%)	33 (4%)	6 (1%)	25 53
All	All	1885/2286 (82%)	1800 (96%)	72 (4%)	13 (1%)	22 50

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1042	PRO
1	A	1061	SER
1	A	1445	VAL
1	B	1044	GLY
1	B	1385	GLN

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	906/1052 (86%)	842 (93%)	64 (7%)	14 36
1	B	891/1052 (85%)	833 (94%)	58 (6%)	17 41
All	All	1797/2104 (85%)	1675 (93%)	122 (7%)	16 38

5 of 122 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1380	ARG
1	B	1299	ARG
1	B	491	LYS
1	B	1295	SER
1	B	1409	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	619	GLN
1	A	929	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1459	ASN
1	B	445	GLN
1	B	1246	ASN

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

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	973/1143 (85%)	-0.10	19 (1%) 65 62	49, 80, 142, 192	0
1	B	960/1143 (83%)	-0.06	16 (1%) 70 68	57, 93, 142, 223	0
2	X	0/17	-	-	-	-
3	Y	0/27	-	-	-	-
4	W	0/11	-	-	-	-
4	Z	0/11	-	-	-	-
All	All	1933/2352 (82%)	-0.08	35 (1%) 68 66	49, 87, 142, 223	0

The worst 5 of 35 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1424	LEU	5.4
1	B	499	LYS	5.3
1	A	607	GLN	5.0
1	B	608	HIS	4.3
1	B	1387	ASN	4.2

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.