



wwPDB X-ray Structure Validation Summary Report ⓘ

May 15, 2020 – 08:27 pm BST

PDB ID : 5OQQ
Title : Crystal structure of the *S. cerevisiae* condensin Ycg1-Brn1 subcomplex
Authors : Kschonsak, M.; Hassler, M.; Haering, C.H.
Deposited on : 2017-08-14
Resolution : 2.79 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 : 2.11
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) : 2.11

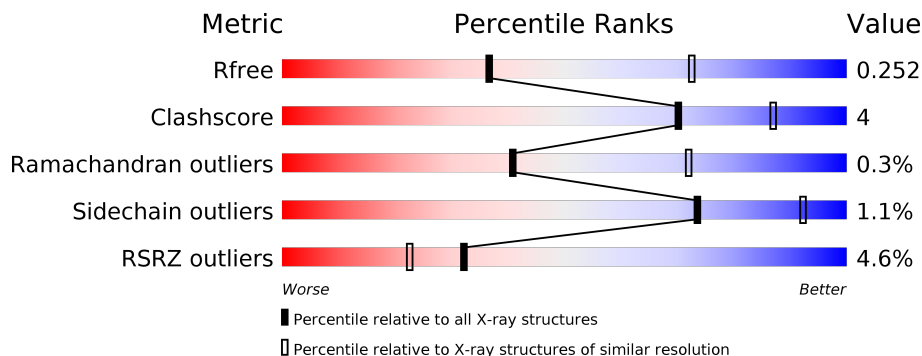
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.79 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	871	
1	B	871	
2	C	152	
2	D	152	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 28987 atoms, of which 14578 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 Condensin complex subunit 3.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	792	12884	4080	6506	1079	1190	29	0	0	0
1	B	824	13380	4230	6742	1123	1256	29	0	0	0

There are 116 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	5	MET	-	initiating methionine	UNP Q06680
A	?	-	GLN	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	LYS	deletion	UNP Q06680
A	?	-	ILE	deletion	UNP Q06680
A	?	-	LYS	deletion	UNP Q06680
A	?	-	SER	deletion	UNP Q06680
A	?	-	LYS	deletion	UNP Q06680
A	?	-	LYS	deletion	UNP Q06680
A	?	-	ILE	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	ARG	deletion	UNP Q06680
A	?	-	ARG	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	THR	deletion	UNP Q06680
A	?	-	SER	deletion	UNP Q06680
A	?	-	VAL	deletion	UNP Q06680
A	?	-	ASP	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	ASP	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLY	deletion	UNP Q06680
A	?	-	THR	deletion	UNP Q06680
A	?	-	HIS	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	ASP	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	VAL	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	ASP	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	ASP	deletion	UNP Q06680
A	?	-	ASP	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	ILE	deletion	UNP Q06680
A	?	-	SER	deletion	UNP Q06680
A	?	-	SER	deletion	UNP Q06680
A	?	-	PHE	deletion	UNP Q06680
A	?	-	HIS	deletion	UNP Q06680
A	?	-	SER	deletion	UNP Q06680
A	?	-	ALA	deletion	UNP Q06680
A	?	-	VAL	deletion	UNP Q06680
A	?	-	GLU	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	LEU	deletion	UNP Q06680
A	?	-	VAL	deletion	UNP Q06680
A	?	-	GLN	deletion	UNP Q06680
A	?	-	GLY	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	GLY	deletion	UNP Q06680
A	?	-	ASN	deletion	UNP Q06680
A	?	-	VAL	deletion	UNP Q06680
B	5	MET	-	initiating methionine	UNP Q06680
B	?	-	GLN	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	LYS	deletion	UNP Q06680
B	?	-	ILE	deletion	UNP Q06680
B	?	-	LYS	deletion	UNP Q06680
B	?	-	SER	deletion	UNP Q06680
B	?	-	LYS	deletion	UNP Q06680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	LYS	deletion	UNP Q06680
B	?	-	ILE	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	ARG	deletion	UNP Q06680
B	?	-	ARG	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	THR	deletion	UNP Q06680
B	?	-	SER	deletion	UNP Q06680
B	?	-	VAL	deletion	UNP Q06680
B	?	-	ASP	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	ASP	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	GLY	deletion	UNP Q06680
B	?	-	THR	deletion	UNP Q06680
B	?	-	HIS	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	ASP	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	VAL	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	ASP	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	ASP	deletion	UNP Q06680
B	?	-	ASP	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	ILE	deletion	UNP Q06680
B	?	-	SER	deletion	UNP Q06680
B	?	-	SER	deletion	UNP Q06680
B	?	-	PHE	deletion	UNP Q06680
B	?	-	HIS	deletion	UNP Q06680
B	?	-	SER	deletion	UNP Q06680
B	?	-	ALA	deletion	UNP Q06680
B	?	-	VAL	deletion	UNP Q06680
B	?	-	GLU	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	LEU	deletion	UNP Q06680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	VAL	deletion	UNP Q06680
B	?	-	GLN	deletion	UNP Q06680
B	?	-	GLY	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	GLY	deletion	UNP Q06680
B	?	-	ASN	deletion	UNP Q06680
B	?	-	VAL	deletion	UNP Q06680

- Molecule 2 is a protein called Condensin complex subunit 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	D	96	1625	528	799	147	148	3	0	0	0
2	C	64	1066	343	531	91	99	2	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	378	GLY	-	expression tag	UNP P38170
D	379	PRO	-	expression tag	UNP P38170
D	380	LEU	-	expression tag	UNP P38170
D	381	GLY	-	expression tag	UNP P38170
D	382	HIS	-	expression tag	UNP P38170
D	383	MET	-	expression tag	UNP P38170
C	378	GLY	-	expression tag	UNP P38170
C	379	PRO	-	expression tag	UNP P38170
C	380	LEU	-	expression tag	UNP P38170
C	381	GLY	-	expression tag	UNP P38170
C	382	HIS	-	expression tag	UNP P38170
C	383	MET	-	expression tag	UNP P38170

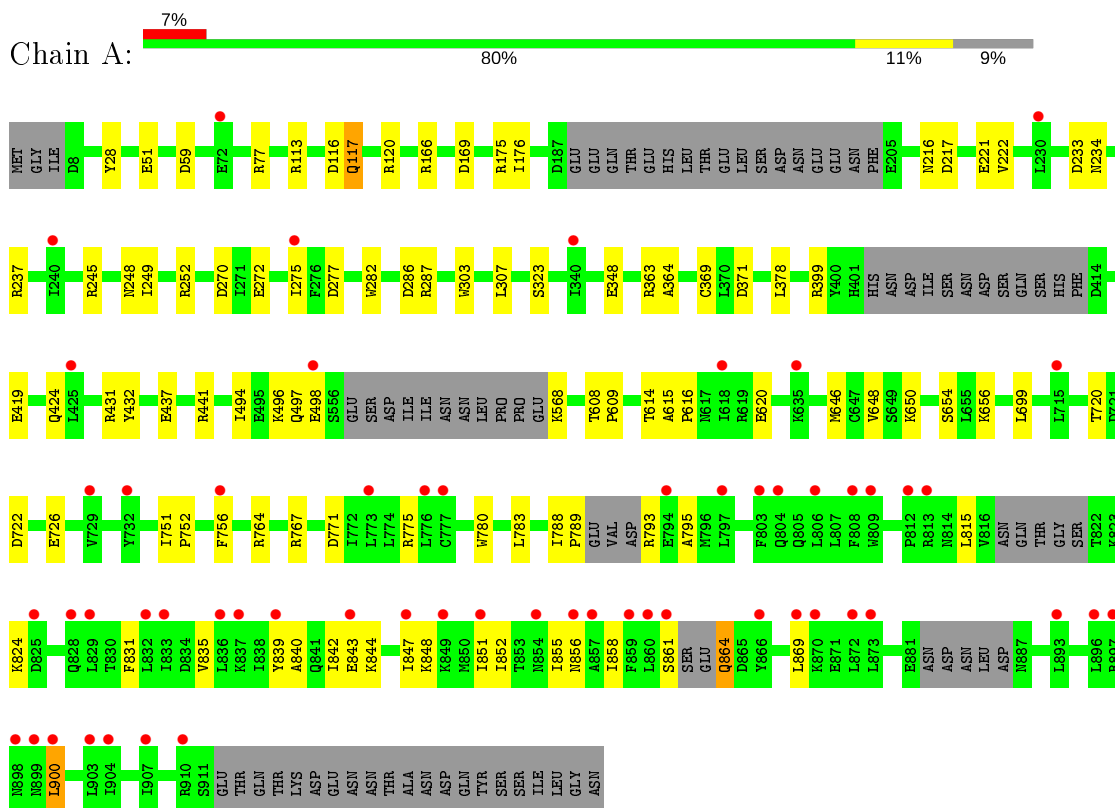
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	19	Total	O	0	0
			19	19		
3	B	9	Total	O	0	0
			9	9		
3	D	4	Total	O	0	0
			4	4		

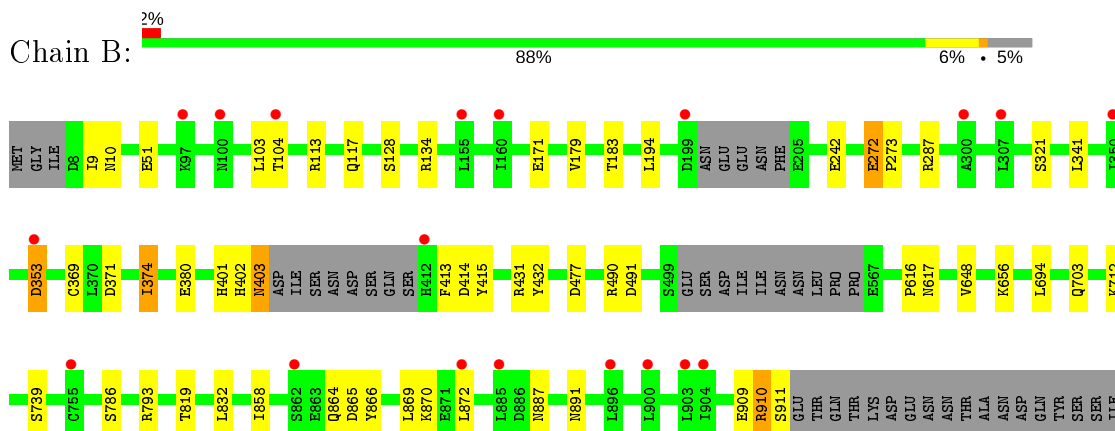
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: Condensin complex subunit 3



- Molecule 1: Condensin complex subunit 3



LEU
GLY
ASN

- Molecule 2: Condensin complex subunit 2



GLY	PRO	LEU	GLY	HIS	MET	S384	I385	F386	A392	E396	N399	R400	N401	W402	R405	V410	R411	N412	PHE	LYS	LYS	ALA	ALA	ASN	ASN	LEU	VAL	ASN	GLU	SER	SER	ASP	LEU	LEU	GLU	THR	THR	THR	THR	ILE	GLY	ASP	THR	THR	ASP	LYS	ASN	THR	THR	THR	ASP	LYS	SER	MET	ASP
-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

THR	LYS	LYS	HIS	LYS	GLN	LYS	VAL	L459	R460	R489	N489	L497	I509	T510	R511	L522	H525	ARG	LYS	HIS	THR
-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	-----	-----	-----	-----

- Molecule 2: Condensin complex subunit 2



GLY	PRO	LEU	THR	HIS	MET	SER	ILE	PHE	GLU	LYS	ASP	LEU	MET	ALA	TYR	PHE	ASP	GLU	ASN	LEU	ASN	ARG	ASN	ASN	TRP	ARG	GLY	ARG	GLU	HIS	TRP	LYS	VAL	VAL	ARG	ASN	LEU	VAL	ASN	LYS	GLU	SER	ASP	LEU	LEU	GLU	GLU	THR	THR	THR	THR	ILE	GLY	ASP	THR	THR
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ASP	LYS	ASN	THR	ASP	ASP	LYS	SER	MET	ASP	THR	LYS	LYS	LYS	HIS	LYS	GLN	LYS	LYS	W458	S477	K478	G479	K491	D501	S521	LEU	PHE	SER	HIS	ARG	LYS	HIS	THR
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	185.53Å 185.53Å 148.32Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	47.25 – 2.79 47.25 – 2.79	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.25-2.79) 100.0 (47.25-2.79)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.44 (at 2.77Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.205 , 0.249 0.206 , 0.252	Depositor DCC
R_{free} test set	1988 reflections (2.71%)	wwPDB-VP
Wilson B-factor (Å ²)	74.1	Xtrriage
Anisotropy	0.023	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 49.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.024 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	28987	wwPDB-VP
Average B, all atoms (Å ²)	88.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% 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	0.24	0/6472	0.40	0/8742
1	B	0.24	0/6740	0.40	0/9113
2	C	0.25	0/547	0.48	0/732
2	D	0.25	0/847	0.43	0/1133
All	All	0.24	0/14606	0.40	0/19720

There are no bond length outliers.

There are no bond angle outliers.

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	6378	6506	6531	60	0
1	B	6638	6742	6744	36	0
2	C	535	531	531	5	0
2	D	826	799	799	9	0
3	A	19	0	0	1	0
3	B	9	0	0	2	0
3	D	4	0	0	2	0
All	All	14409	14578	14605	103	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 4.

The worst 5 of 103 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:403:ASN:OD1	1:B:403:ASN:N	2.17	0.75
1:A:861:SER:O	1:A:864:GLN:NE2	2.21	0.73
1:A:793:ARG:HG2	1:A:795:ALA:H	1.54	0.72
1:A:287:ARG:NH1	2:C:501:ASP:OD1	2.24	0.70
1:A:722:ASP:OD1	1:A:764:ARG:NH1	2.25	0.70

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	776/871 (89%)	744 (96%)	32 (4%)	0	100	100
1	B	816/871 (94%)	788 (97%)	25 (3%)	3 (0%)	34	66
2	C	62/152 (41%)	56 (90%)	4 (6%)	2 (3%)	4	13
2	D	92/152 (60%)	85 (92%)	7 (8%)	0	100	100
All	All	1746/2046 (85%)	1673 (96%)	68 (4%)	5 (0%)	41	72

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	402	HIS
1	B	910	ARG
2	C	478	LYS
2	C	479	GLY
1	B	413	PHE

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	719/794 (91%)	712 (99%)	7 (1%)	76	93
1	B	750/794 (94%)	740 (99%)	10 (1%)	69	91
2	C	61/143 (43%)	61 (100%)	0	100	100
2	D	91/143 (64%)	90 (99%)	1 (1%)	73	92
All	All	1621/1874 (86%)	1603 (99%)	18 (1%)	73	92

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	171	GLU
1	B	272	GLU
1	B	403	ASN
1	A	900	LEU
1	B	117	GLN

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

Mol	Chain	Res	Type
1	A	864	GLN
1	B	117	GLN

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 carbohydrates 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	792/871 (90%)	0.51	58 (7%) 15 8	51, 79, 136, 256	0
1	B	824/871 (94%)	0.35	19 (2%) 60 51	46, 68, 114, 145	0
2	C	64/152 (42%)	0.19	1 (1%) 72 66	64, 88, 108, 122	0
2	D	96/152 (63%)	0.42	4 (4%) 36 26	50, 81, 122, 131	0
All	All	1776/2046 (86%)	0.42	82 (4%) 32 22	46, 74, 127, 256	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	866	TYR	5.5
1	A	860	LEU	5.5
1	A	900	LEU	5.1
1	B	885	LEU	4.8
1	A	833	ILE	4.4

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 carbohydrates in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.