W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 05:30 pm BST

PDB ID	:	5OQP
Title	:	Crystal structure of the S. cerevisiae condensin Ycg1-Brn1 subcomplex bound to DNA (crystal form I)
Authors	:	Kschonsak, M.; Hassler, M.; Haering, C.H.
Deposited on	:	2017-08-14
Resolution	:	2.98 Å(reported)

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

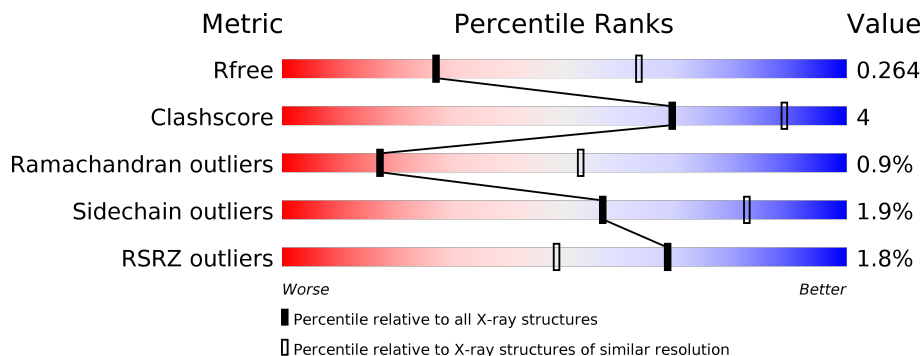|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 2.98 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	Percentile Ranks	Value
Rfree		0.264
Clashscore		4
Ramachandran outliers		0.9%
Sidechain outliers		1.9%
RSRZ outliers		1.8%

Worse   Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2754 (3.00-2.96)
Clashscore	141614	3103 (3.00-2.96)
Ramachandran outliers	138981	2993 (3.00-2.96)
Sidechain outliers	138945	2996 (3.00-2.96)
RSRZ outliers	127900	2644 (3.00-2.96)

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 <=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	869	85% 7% 7%
2	B	152	44% 17% • 37%
3	C	18	56% 44%
3	D	18	72% 28%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 15936 atoms, of which 7866 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
1	A	806	Total 13114	C 4150	H 6621	N 1099	O 1215	S 29	0	0	0

There are 58 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
A	?	-	GLY	deletion	UNP Q06680
A	?	-	THR	deletion	UNP Q06680

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
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

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	B	96	1637	527	813	147	147	3	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	378	GLY	-	expression tag	UNP P38170

Continued on next page...

Continued from previous page...

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

- Molecule 3 is a DNA chain called DNA (5'-D(*GP*AP*TP*GP*TP*GP*TP*AP*GP*CP* TP*AP*CP*AP*CP*AP*TP*C)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	C	18	Total 571	C 176	H 205	N 67	O 106	P 17	0	0	0
3	D	18	Total 571	C 176	H 205	N 67	O 106	P 17	0	0	0

- Molecule 4 is PENTAETHYLENE GLYCOL (three-letter code: 1PE) (formula: $C_{10}H_{22}O_6$).
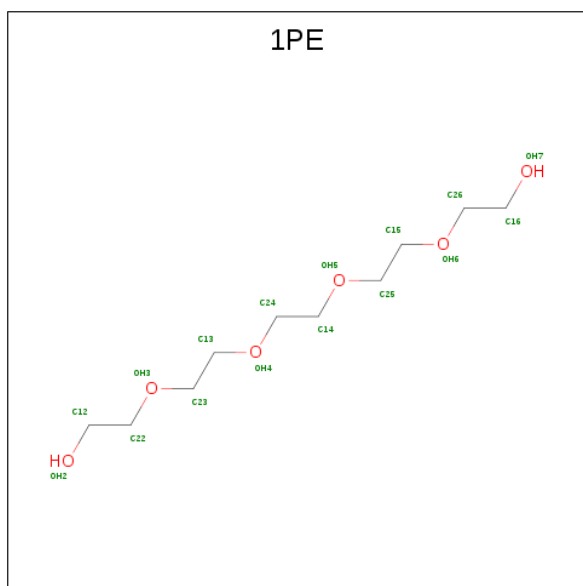
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total 38	C 10	H 22	O 6	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O

O3 O⁻ — S — O⁻ O4

S

O

O2
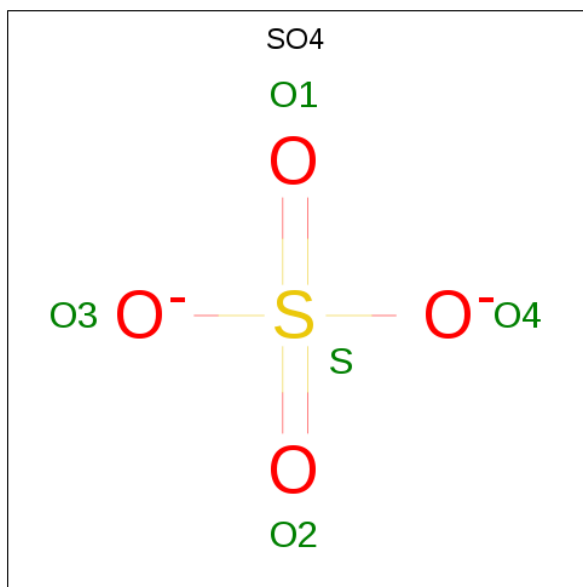
| Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		

3 Residue-property plots ⓘ

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.

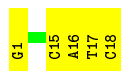- Molecule 1: Condensin complex subunit 3
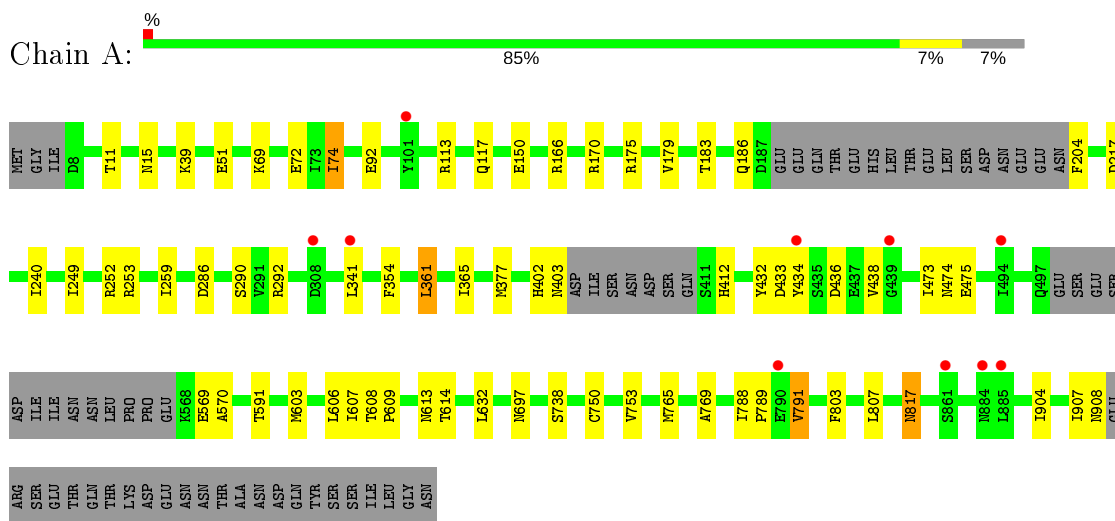
- Molecule 2: Condensin complex subunit 2
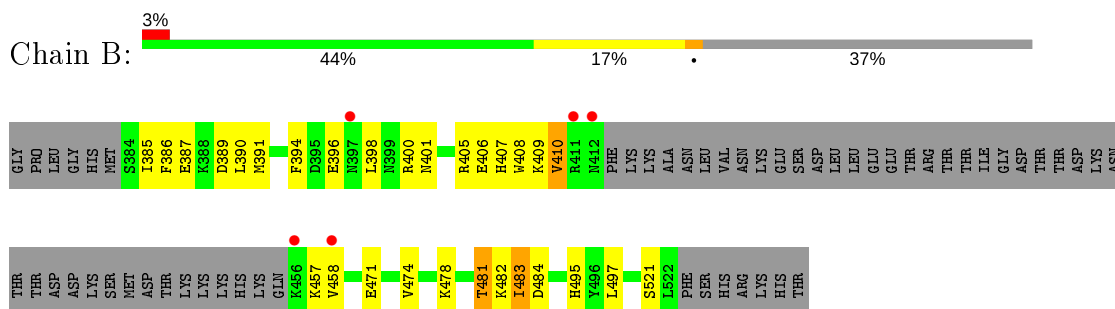
- Molecule 3: DNA (5'-D(*GP*AP*TP*GP*TP*GP*TP*AP*GP*CP*TP*AP*CP*AP*CP*AP*TP*C)-3')

- Molecule 3: DNA (5'-D(*GP*AP*TP*GP*TP*GP*TP*AP*GP*CP*TP*AP*CP*AP*CP*AP*TP*C)-3')

Chain D:

72%   28%

G1   C15 A16 T17 C18

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	80.14Å 114.83Å 155.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.17 – 2.98 47.17 – 2.98	Depositor EDS
% Data completeness (in resolution range)	100.0 (47.17-2.98) 91.6 (47.17-2.98)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.85 (at 2.96Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.218 , 0.262 0.219 , 0.264	Depositor DCC
R_{free} test set	2000 reflections (6.68%)	wwPDB-VP
Wilson B-factor (Å^2)	59.9	Xtriage
Anisotropy	0.451	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 44.0	EDS
L-test for twinning[2]	$< |L| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	15936	wwPDB-VP
Average B, all atoms (Å^2)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.05% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, 1PE

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	0/6595	0.40	0/8918
2	B	0.24	0/843	0.44	0/1126
3	C	0.56	0/410	0.97	0/631
3	D	0.51	0/410	0.94	0/631
All	All	0.28	0/8258	0.50	0/11306

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	6493	6621	6624	36	0
2	B	824	813	813	15	0
3	C	366	205	205	10	0
3	D	366	205	205	2	1
4	A	16	22	22	0	0
5	A	5	0	0	0	0
All	All	8070	7866	7869	58	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 4.

The worst 5 of 58 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:B:400:ARG:NH1	3:C:12:DA:OP1	2.12	0.82
1:A:907:ILE:O	1:A:908:ASN:OD1	2.02	0.78
1:A:286:ASP:O	1:A:292:ARG:NH1	2.21	0.72
2:B:478:LYS:O	2:B:481:THR:OG1	2.11	0.69
1:A:166:ARG:NH2	2:B:484:ASP:OD1	2.25	0.69

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)
3:D:1:DG:O5'	3:D:18:DC:OP2[3_554]	1.95	0.25

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	798/869 (92%)	768 (96%)	27 (3%)	3 (0%)	34	70
2	B	92/152 (60%)	76 (83%)	11 (12%)	5 (5%)	2	9
All	All	890/1021 (87%)	844 (95%)	38 (4%)	8 (1%)	17	53

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	483	ILE
2	B	521	SER
1	A	474	ASN
1	A	791	VAL
2	B	407	HIS

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	733/793 (92%)	721 (98%)	12 (2%)	62	85
2	B	91/143 (64%)	87 (96%)	4 (4%)	28	63
All	All	824/936 (88%)	808 (98%)	16 (2%)	57	82

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	606	LEU
1	A	632	LEU
2	B	394	PHE
1	A	432	TYR
2	B	398	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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)

2 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z| > 2$	Counts	RMSZ	$\#|Z| > 2$
4	1PE	A	1001	-	15,15,15	0.52	0	14,14,14	0.34	0
5	SO4	A	1002	-	4,4,4	0.14	0	6,6,6	0.06	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	1PE	A	1001	-	-	<mark>5/13/13/13</mark>	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	1001	1PE	OH4-C13-C23-OH3
4	A	1001	1PE	OH5-C14-C24-OH4
4	A	1001	1PE	OH7-C16-C26-OH6
4	A	1001	1PE	C24-C14-OH5-C25
4	A	1001	1PE	OH6-C15-C25-OH5

There are no ring outliers.

No monomer is involved in short contacts.

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	806/869 (92%)	0.12	10 (1%)	79	61	28, 55, 98, 166	0
2	B	96/152 (63%)	0.50	5 (5%)	27	16	34, 78, 136, 173	0
3	C	18/18 (100%)	0.43	2 (11%)	5	3	100, 127, 167, 179	0
3	D	18/18 (100%)	0.14	0	100	100	86, 118, 150, 158	0
All	All	938/1057 (88%)	0.17	17 (1%)	68	48	28, 58, 117, 179	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	397	ASN	4.3
2	B	411	ARG	4.2
1	A	790	GLU	3.6
1	A	101	TYR	3.5
3	C	6	DG	3.3

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 carbohydrates in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	1PE	A	1001	16/16	0.92	0.23	46,67,84,94	0
5	SO4	A	1002	5/5	0.98	0.13	33,37,45,46	0

6.5 Other polymers (i)

There are no such residues in this entry.