wwPDB X-ray Structure Validation Summary Report (i)

Jun 17, 2024 – 06:00 AM EDT

PDB ID	:	5OQR
Title	:	Crystal structure of the S. pombe condensin Cnd3-Cnd2 subcomplex
Authors	:	Kschonsak, M.; Hassler, M.; Haering, C.H.
Deposited on	:	2017-08-14
Resolution	:	2.61 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:

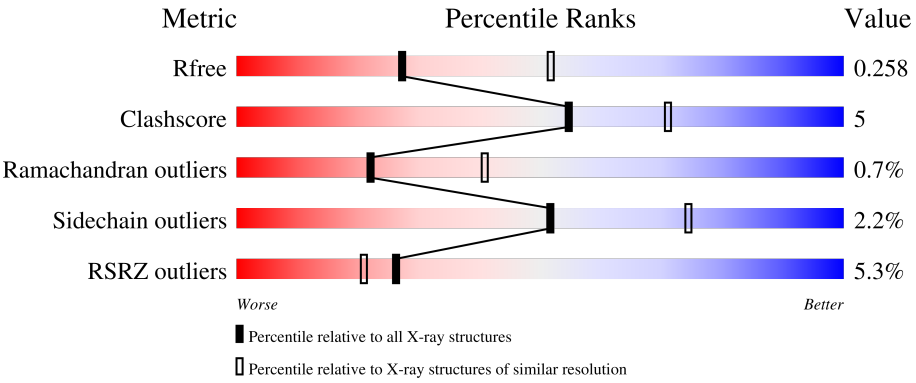
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.61 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	3797 (2.64-2.60)
Clashscore	141614	4168 (2.64-2.60)
Ramachandran outliers	138981	4093 (2.64-2.60)
Sidechain outliers	138945	4093 (2.64-2.60)
RSRZ outliers	127900	3731 (2.64-2.60)

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	788	2% 85% 9% • 5%
1	B	788	6% 77% 14% • 8%
2	C	135	3% 30% 6% • 61%
2	D	135	13% 43% 7% • 50%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 26152 atoms, of which 13120 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	751	Total	C	H	N	O	S	0	0	0
			12231	3873	6155	1029	1145	29			
1	B	726	Total	C	H	N	O	S	0	0	0
			11862	3761	5967	996	1109	29			

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	GLN	deletion	UNP Q10429
A	?	-	THR	deletion	UNP Q10429
A	?	-	GLU	deletion	UNP Q10429
A	?	-	GLU	deletion	UNP Q10429
A	?	-	GLN	deletion	UNP Q10429
A	?	-	GLY	deletion	UNP Q10429
A	?	-	ASN	deletion	UNP Q10429
A	?	-	SER	deletion	UNP Q10429
A	?	-	ASN	deletion	UNP Q10429
A	?	-	ALA	deletion	UNP Q10429
A	?	-	PRO	deletion	UNP Q10429
A	?	-	GLU	deletion	UNP Q10429
A	?	-	LEU	deletion	UNP Q10429
A	?	-	ASN	deletion	UNP Q10429
A	?	-	LYS	deletion	UNP Q10429
A	?	-	ASN	deletion	UNP Q10429
A	?	-	ASP	deletion	UNP Q10429
A	?	-	TYR	deletion	UNP Q10429
A	?	-	GLU	deletion	UNP Q10429
A	?	-	GLY	deletion	UNP Q10429
A	?	-	GLU	deletion	UNP Q10429
A	?	-	GLU	deletion	UNP Q10429
A	?	-	ILE	deletion	UNP Q10429
A	?	-	THR	deletion	UNP Q10429
A	?	-	VAL	deletion	UNP Q10429

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	SER	deletion	UNP Q10429
A	?	-	GLN	deletion	UNP Q10429
A	?	-	LYS	deletion	UNP Q10429
A	?	-	SER	deletion	UNP Q10429
A	?	-	PRO	deletion	UNP Q10429
A	?	-	SER	deletion	UNP Q10429
A	?	-	PRO	deletion	UNP Q10429
A	?	-	SER	deletion	UNP Q10429
A	?	-	LEU	deletion	UNP Q10429
A	?	-	PRO	deletion	UNP Q10429
B	?	-	GLN	deletion	UNP Q10429
B	?	-	THR	deletion	UNP Q10429
B	?	-	GLU	deletion	UNP Q10429
B	?	-	GLU	deletion	UNP Q10429
B	?	-	GLN	deletion	UNP Q10429
B	?	-	GLY	deletion	UNP Q10429
B	?	-	ASN	deletion	UNP Q10429
B	?	-	SER	deletion	UNP Q10429
B	?	-	ASN	deletion	UNP Q10429
B	?	-	ALA	deletion	UNP Q10429
B	?	-	PRO	deletion	UNP Q10429
B	?	-	GLU	deletion	UNP Q10429
B	?	-	LEU	deletion	UNP Q10429
B	?	-	ASN	deletion	UNP Q10429
B	?	-	LYS	deletion	UNP Q10429
B	?	-	ASN	deletion	UNP Q10429
B	?	-	ASP	deletion	UNP Q10429
B	?	-	TYR	deletion	UNP Q10429
B	?	-	GLU	deletion	UNP Q10429
B	?	-	GLY	deletion	UNP Q10429
B	?	-	GLU	deletion	UNP Q10429
B	?	-	GLU	deletion	UNP Q10429
B	?	-	ILE	deletion	UNP Q10429
B	?	-	THR	deletion	UNP Q10429
B	?	-	VAL	deletion	UNP Q10429
B	?	-	SER	deletion	UNP Q10429
B	?	-	GLN	deletion	UNP Q10429
B	?	-	LYS	deletion	UNP Q10429
B	?	-	SER	deletion	UNP Q10429
B	?	-	PRO	deletion	UNP Q10429
B	?	-	SER	deletion	UNP Q10429
B	?	-	PRO	deletion	UNP Q10429

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	?	-	SER	deletion	UNP Q10429
B	?	-	LEU	deletion	UNP Q10429
B	?	-	PRO	deletion	UNP Q10429

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

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	C	52	Total	C	H	N	O	S	0	0	0
			878	287	434	75	80	2			
2	D	68	Total	C	H	N	O	S	0	0	0
			1134	369	564	95	103	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	410	GLY	-	expression tag	UNP Q9Y7R3
C	411	PRO	-	expression tag	UNP Q9Y7R3
C	412	LEU	-	expression tag	UNP Q9Y7R3
C	413	GLY	-	expression tag	UNP Q9Y7R3
C	414	HIS	-	expression tag	UNP Q9Y7R3
C	415	MET	-	expression tag	UNP Q9Y7R3
D	410	GLY	-	expression tag	UNP Q9Y7R3
D	411	PRO	-	expression tag	UNP Q9Y7R3
D	412	LEU	-	expression tag	UNP Q9Y7R3
D	413	GLY	-	expression tag	UNP Q9Y7R3
D	414	HIS	-	expression tag	UNP Q9Y7R3
D	415	MET	-	expression tag	UNP Q9Y7R3

- Molecule 3 is water.

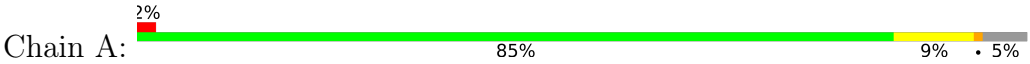
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	24	Total	O	0	0
			24	24		
3	B	22	Total	O	0	0
			22	22		
3	D	1	Total	O	0	0
			1	1		

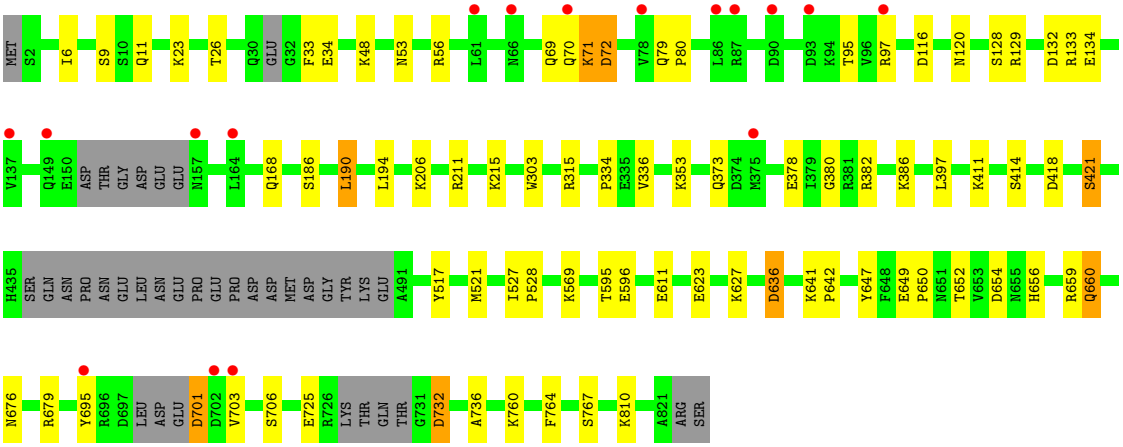
3 Residue-property plots (i)

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: Condensin complex subunit 3

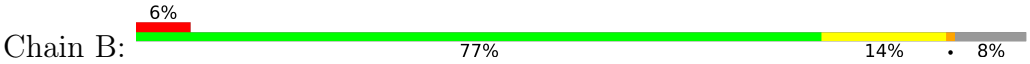
Chain A:

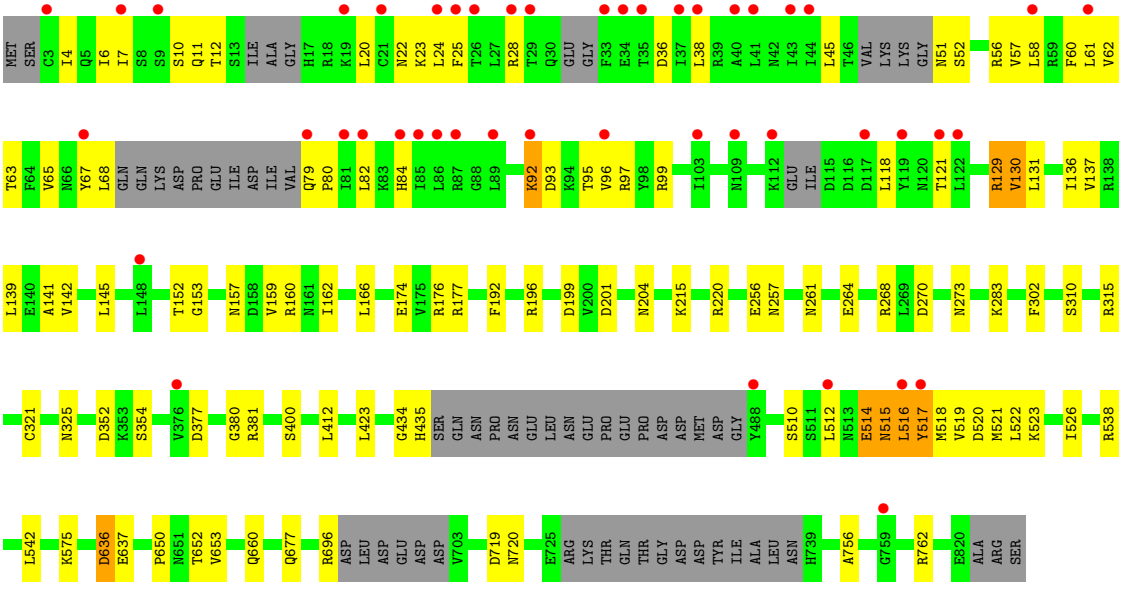




• Molecule 1: Condensin complex subunit 3

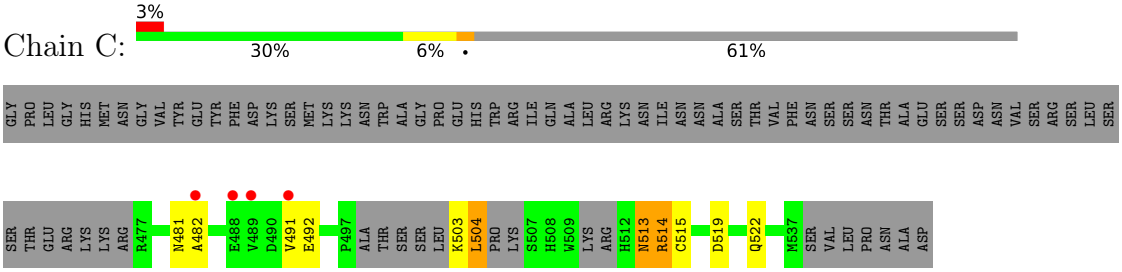
Chain B:





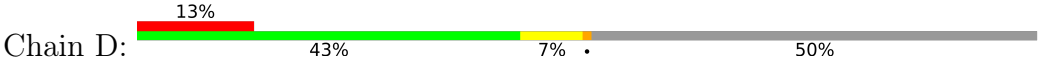
• Molecule 2: Condensin complex subunit 2

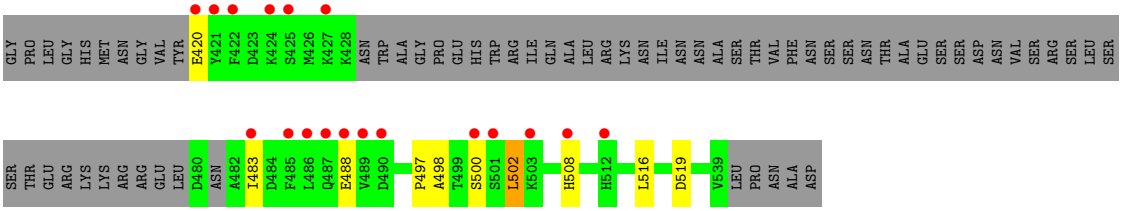
Chain C:



• Molecule 2: Condensin complex subunit 2

Chain D:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	88.41Å 142.11Å 176.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.06 – 2.61 49.06 – 2.61	Depositor EDS
% Data completeness (in resolution range)	99.5 (49.06-2.61) 96.8 (49.06-2.61)	Depositor EDS
R_{merge}	0.21	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.38 (at 2.61Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.223 , 0.258 0.223 , 0.258	Depositor DCC
R_{free} test set	2000 reflections (2.92%)	wwPDB-VP
Wilson B-factor (Å^2)	53.2	Xtriage
Anisotropy	0.367	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.1	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	0.94	EDS
Total number of atoms	26152	wwPDB-VP
Average B, all atoms (Å^2)	82.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.67% 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	0.25	0/6169	0.40	0/8335
1	B	0.25	0/5984	0.43	0/8080
2	C	0.24	0/452	0.44	0/606
2	D	0.26	0/582	0.46	0/779
All	All	0.25	0/13187	0.42	0/17800

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

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	6076	6155	6165	45	4
1	B	5895	5967	5974	81	1
2	C	444	434	435	11	0
2	D	570	564	574	10	0
3	A	24	0	0	2	0
3	B	22	0	0	5	0
3	D	1	0	0	0	0
All	All	13032	13120	13148	136	4

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 136 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:220:ARG:NH2	1:B:257:ASN:O	2.08	0.87
1:B:22:ASN:ND2	1:B:67:TYR:OH	2.08	0.86
2:D:500:SER:HB2	2:D:502:LEU:HD21	1.60	0.83
1:A:129:ARG:NH2	1:A:132:ASP:OD1	2.12	0.82
1:A:168:GLN:OE1	2:C:514:ARG:NH1	2.13	0.81

All (4) 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:A:34:GLU:OE2	1:A:569:LYS:NZ[2_665]	2.11	0.09
1:A:386:LYS:NZ	1:A:596:GLU:OE2[4_456]	2.13	0.07
1:A:353:LYS:HZ3	1:B:510:SER:OG[3_656]	1.54	0.06
1:A:386:LYS:HZ1	1:A:596:GLU:OE2[4_456]	1.55	0.05

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	739/788 (94%)	720 (97%)	15 (2%)	4 (0%)	29	50
1	B	708/788 (90%)	671 (95%)	32 (4%)	5 (1%)	22	41
2	C	44/135 (33%)	40 (91%)	3 (7%)	1 (2%)	6	10
2	D	63/135 (47%)	51 (81%)	11 (18%)	1 (2%)	9	18
All	All	1554/1846 (84%)	1482 (95%)	61 (4%)	11 (1%)	22	41

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	736	ALA
1	B	92	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	515	ASN
1	B	516	LEU
2	C	514	ARG

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	686/721 (95%)	670 (98%)	16 (2%)	50	73
1	B	667/721 (92%)	654 (98%)	13 (2%)	57	78
2	C	50/123 (41%)	48 (96%)	2 (4%)	31	55
2	D	65/123 (53%)	63 (97%)	2 (3%)	40	65
All	All	1468/1688 (87%)	1435 (98%)	33 (2%)	52	74

5 of 33 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	696	ARG
2	C	504	LEU
2	D	502	LEU
1	A	701	ASP
1	A	660	GLN

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

Mol	Chain	Res	Type
1	B	22	ASN
1	B	704	GLN
2	C	481	ASN
2	D	513	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	751/788 (95%)	0.17	17 (2%) 60 55	26, 52, 123, 160	0
1	B	726/788 (92%)	0.44	46 (6%) 20 15	30, 66, 161, 210	0
2	C	52/135 (38%)	0.74	4 (7%) 13 10	33, 114, 176, 194	0
2	D	68/135 (50%)	1.45	18 (26%) 0 0	39, 126, 199, 248	0
All	All	1597/1846 (86%)	0.37	85 (5%) 26 21	26, 63, 156, 248	0

The worst 5 of 85 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	40	ALA	6.6
1	B	7	ILE	6.1
1	B	33	PHE	6.1
2	D	421	TYR	6.1
1	B	79	GLN	5.9

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.