wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 06:50 AM GMT

PDB ID	:	2YIN
Title	:	STRUCTURE OF THE COMPLEX BETWEEN DOCK2 AND RAC1.
Authors	:	Kulkarni, K.A.; Yang, J.; Zhang, Z.; Barford, D.
Deposited on	:	2011-05-16
Resolution	:	2.70 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

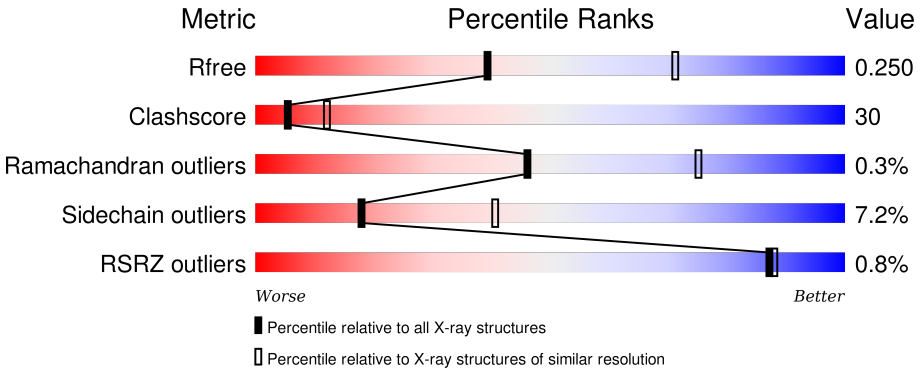
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.70 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}	91344	2103 (2.70-2.70)
Clashscore	102246	2422 (2.70-2.70)
Ramachandran outliers	100387	2382 (2.70-2.70)
Sidechain outliers	100360	2382 (2.70-2.70)
RSRZ outliers	91569	2107 (2.70-2.70)

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. 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	436	56% 35% • 6%
1	B	436	53% 36% 5% 6%
2	C	196	57% 31% • 8%
2	D	196	53% 39% • 8%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9616 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 DEDICATOR OF CYTOKINESIS PROTEIN 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	411	Total	C	N	O	S	0	0	0
			3379	2175	550	632	22			
1	B	410	Total	C	N	O	S	0	0	0
			3338	2154	534	628	22			

There are 10 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1187	GLY	-	EXPRESSION TAG	UNP Q92608
A	1188	GLU	-	EXPRESSION TAG	UNP Q92608
A	1189	CYS	-	EXPRESSION TAG	UNP Q92608
A	1190	GLY	-	EXPRESSION TAG	UNP Q92608
A	1191	ASP	-	EXPRESSION TAG	UNP Q92608
B	1187	GLY	-	EXPRESSION TAG	UNP Q92608
B	1188	GLU	-	EXPRESSION TAG	UNP Q92608
B	1189	CYS	-	EXPRESSION TAG	UNP Q92608
B	1190	GLY	-	EXPRESSION TAG	UNP Q92608
B	1191	ASP	-	EXPRESSION TAG	UNP Q92608

- Molecule 2 is a protein called RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	180	Total	C	N	O	S	0	0	0
			1399	898	231	262	8			
2	D	181	Total	C	N	O	S	0	0	0
			1409	906	231	264	8			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-18	MET	-	EXPRESSION TAG	UNP P63000
C	-17	ALA	-	EXPRESSION TAG	UNP P63000

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	-16	HIS	-	EXPRESSION TAG	UNP P63000
C	-15	HIS	-	EXPRESSION TAG	UNP P63000
C	-14	HIS	-	EXPRESSION TAG	UNP P63000
C	-13	HIS	-	EXPRESSION TAG	UNP P63000
C	-12	HIS	-	EXPRESSION TAG	UNP P63000
C	-11	HIS	-	EXPRESSION TAG	UNP P63000
C	-10	SER	-	EXPRESSION TAG	UNP P63000
C	-9	SER	-	EXPRESSION TAG	UNP P63000
C	-8	GLY	-	EXPRESSION TAG	UNP P63000
C	-7	LEU	-	EXPRESSION TAG	UNP P63000
C	-6	GLU	-	EXPRESSION TAG	UNP P63000
C	-5	VAL	-	EXPRESSION TAG	UNP P63000
C	-4	LEU	-	EXPRESSION TAG	UNP P63000
C	-3	PHE	-	EXPRESSION TAG	UNP P63000
C	-2	GLN	-	EXPRESSION TAG	UNP P63000
C	-1	GLY	-	EXPRESSION TAG	UNP P63000
C	0	THR	-	EXPRESSION TAG	UNP P63000
D	-18	MET	-	EXPRESSION TAG	UNP P63000
D	-17	ALA	-	EXPRESSION TAG	UNP P63000
D	-16	HIS	-	EXPRESSION TAG	UNP P63000
D	-15	HIS	-	EXPRESSION TAG	UNP P63000
D	-14	HIS	-	EXPRESSION TAG	UNP P63000
D	-13	HIS	-	EXPRESSION TAG	UNP P63000
D	-12	HIS	-	EXPRESSION TAG	UNP P63000
D	-11	HIS	-	EXPRESSION TAG	UNP P63000
D	-10	SER	-	EXPRESSION TAG	UNP P63000
D	-9	SER	-	EXPRESSION TAG	UNP P63000
D	-8	GLY	-	EXPRESSION TAG	UNP P63000
D	-7	LEU	-	EXPRESSION TAG	UNP P63000
D	-6	GLU	-	EXPRESSION TAG	UNP P63000
D	-5	VAL	-	EXPRESSION TAG	UNP P63000
D	-4	LEU	-	EXPRESSION TAG	UNP P63000
D	-3	PHE	-	EXPRESSION TAG	UNP P63000
D	-2	GLN	-	EXPRESSION TAG	UNP P63000
D	-1	GLY	-	EXPRESSION TAG	UNP P63000
D	0	THR	-	EXPRESSION TAG	UNP P63000

- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	35	Total O 35 35	0	0

Continued on next page...

Continued from previous page...

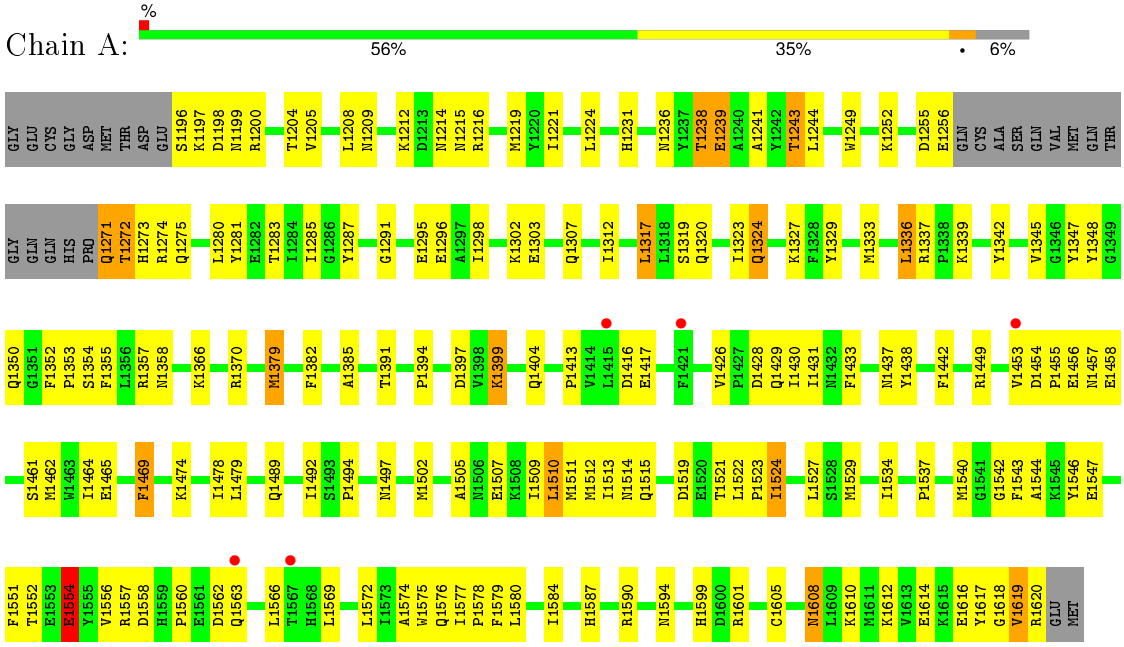
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	27	Total O 27 27	0	0
3	C	13	Total O 13 13	0	0
3	D	16	Total O 16 16	0	0

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 errors 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.

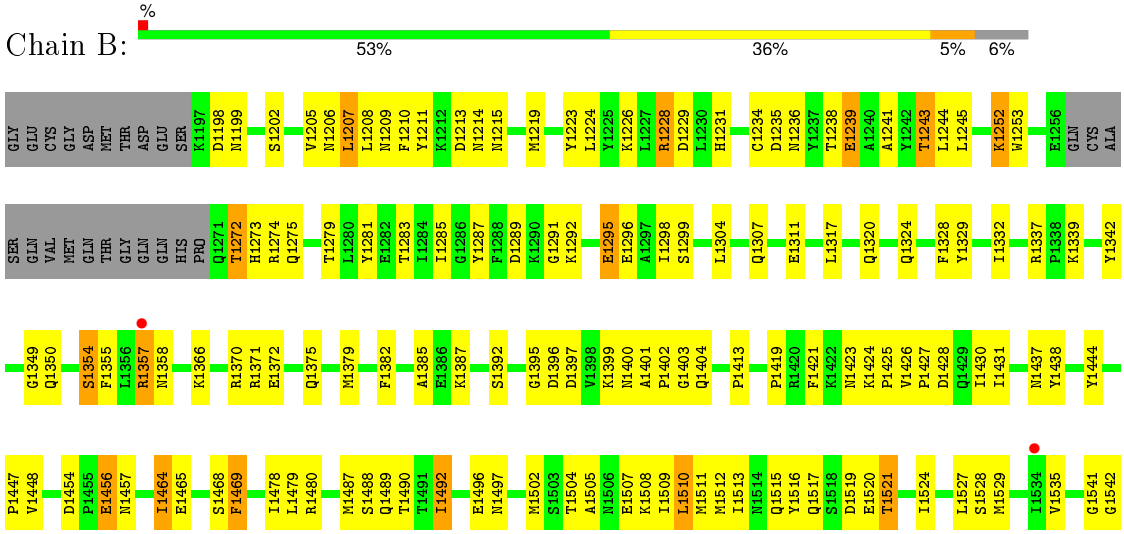
• Molecule 1: DEDICATOR OF CYTOKINESIS PROTEIN 2

Chain A:



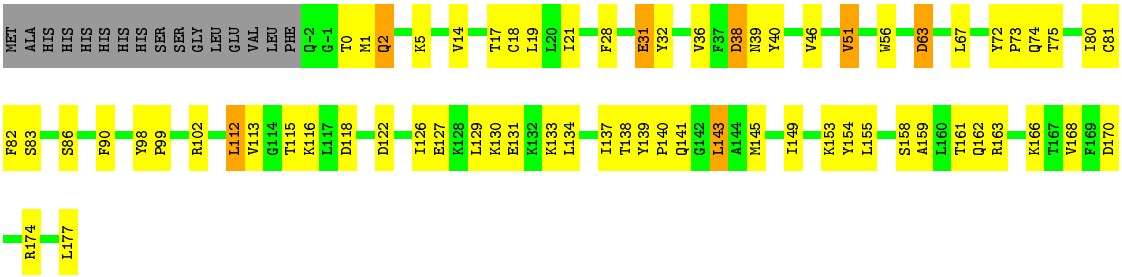
• Molecule 1: DEDICATOR OF CYTOKINESIS PROTEIN 2

Chain B:



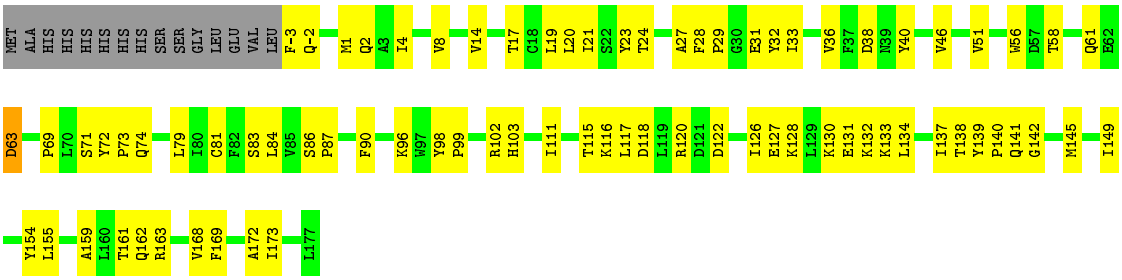
• Molecule 2: RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1

Chain C:



• Molecule 2: RAS-RELATED C3 BOTULINUM TOXIN SUBSTRATE 1

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	68.51Å 98.61Å 130.13Å 90.00° 99.64° 90.00°	Depositor
Resolution (Å)	64.37 – 2.70 64.37 – 2.70	Depositor EDS
% Data completeness (in resolution range)	92.7 (64.37-2.70) 92.8 (64.37-2.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.72 (at 2.69Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.203 , 0.255 0.203 , 0.250	Depositor DCC
R_{free} test set	2217 reflections (5.35%)	DCC
Wilson B-factor (Å^2)	63.5	Xtriage
Anisotropy	0.404	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 61.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.46$, $< L^2 > = 0.29$	Xtriage
Outliers	0 of 43676 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	9616	wwPDB-VP
Average B, all atoms (Å^2)	68.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.58% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.75	2/3467 (0.1%)	0.72	3/4695 (0.1%)
1	B	0.72	0/3426	0.72	1/4648 (0.0%)
2	C	0.64	0/1429	0.71	1/1944 (0.1%)
2	D	0.56	0/1440	0.68	1/1960 (0.1%)
All	All	0.70	2/9762 (0.0%)	0.71	6/13247 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1554	GLU	CB-CG	-10.56	1.32	1.52
1	A	1554	GLU	CD-OE2	8.09	1.34	1.25

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1419	PRO	CB-CA-C	-7.99	92.02	112.00
1	A	1554	GLU	CG-CD-OE2	-7.62	103.06	118.30
2	C	38	ASP	N-CA-C	7.17	130.35	111.00
2	D	38	ASP	N-CA-C	6.30	128.01	111.00
1	A	1370	ARG	NE-CZ-NH1	5.72	123.16	120.30

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	3379	0	3224	201	0
1	B	3338	0	3144	218	0
2	C	1399	0	1417	80	0
2	D	1409	0	1416	87	0
3	A	35	0	0	3	0
3	B	27	0	0	2	0
3	C	13	0	0	3	0
3	D	16	0	0	0	0
All	All	9616	0	9201	557	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 30.

The worst 5 of 557 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:1426:VAL:CG1	1:A:1430:ILE:HD11	1.58	1.31
1:A:1298:ILE:HD13	1:A:1329:TYR:CE1	1.68	1.27
1:B:1298:ILE:HD13	1:B:1329:TYR:CE1	1.71	1.23
1:A:1560:PRO:O	1:A:1563:GLN:HG3	1.41	1.20
1:B:1252:LYS:HD3	1:B:1253:TRP:N	1.56	1.20

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	407/436 (93%)	397 (98%)	9 (2%)	1 (0%)	52 80
1	B	406/436 (93%)	393 (97%)	12 (3%)	1 (0%)	52 80
2	C	178/196 (91%)	173 (97%)	5 (3%)	0	100 100
2	D	179/196 (91%)	173 (97%)	5 (3%)	1 (1%)	30 59

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	1170/1264 (93%)	1136 (97%)	31 (3%)	3 (0%)	46 75

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	120	ARG
1	A	1312	ILE
1	B	1560	PRO

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	362/398 (91%)	339 (94%)	23 (6%)	22 47
1	B	352/398 (88%)	315 (90%)	37 (10%)	8 19
2	C	154/169 (91%)	144 (94%)	10 (6%)	21 46
2	D	154/169 (91%)	150 (97%)	4 (3%)	54 83
All	All	1022/1134 (90%)	948 (93%)	74 (7%)	18 41

5 of 74 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	1295	GLU
1	B	1357	ARG
2	C	143	LEU
1	B	1299	SER
1	B	1317	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 49 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	1215	ASN
1	B	1350	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	D	-2	GLN
1	B	1307	GLN
1	B	1358	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 carbohydrates 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	411/436 (94%)	0.29	5 (1%) 81 81	33, 68, 107, 127	0
1	B	410/436 (94%)	0.20	5 (1%) 81 81	37, 68, 106, 123	0
2	C	180/196 (91%)	0.11	0 100 100	42, 63, 88, 105	0
2	D	181/196 (92%)	0.11	0 100 100	44, 62, 89, 107	0
All	All	1182/1264 (93%)	0.20	10 (0%) 87 88	33, 66, 102, 127	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1421	PHE	3.2
1	A	1567	THR	3.0
1	A	1453	VAL	2.8
1	A	1415	LEU	2.5
1	A	1563	GLN	2.4

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 carbohydrates 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.