WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Feb 14, 2017 – 01:05 pm GMT

PDB ID	:	4JQ5
Title	:	Crystal structure of the human Nup49CCS2+3* coiled-coil segment
Authors	:	Stuwe, T.; Bley, C.J.; Mayo, D.J.; Hoelz, A.
Deposited on	:	2013-03-20
Resolution	:	2.19 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 (i)) were used in the production of this report:

|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

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 2.19 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		0.291
Clashscore		4
Ramachandran outliers		0
Sidechain outliers		6.7%
RSRZ outliers		10.2%

Worse   *Better*

▌Percentile relative to all X-ray structures

▯Percentile relative to X-ray structures of similar resolution
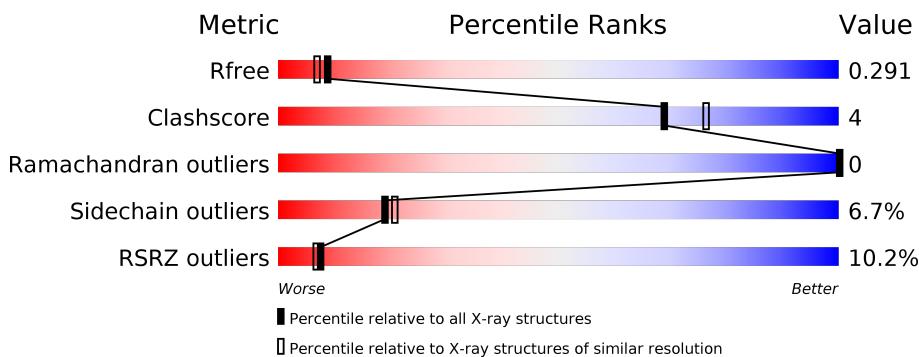
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	4002 (2.20-2.20)
Clashscore	112137	4730 (2.20-2.20)
Ramachandran outliers	110173	4656 (2.20-2.20)
Sidechain outliers	110143	4657 (2.20-2.20)
RSRZ outliers	101464	4033 (2.20-2.20)

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	86	12% / 81% / 17%
1	B	86	7% / 85% / 14%
1	C	86	15% / 76% / 21%
1	D	86	8% / 86% / 13%
1	E	86	12% / 79% / 21%
1	F	86	8% / 83% / 17%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	G	86	15% / 90% / 10%
1	H	86	6% / 84% / 15%
1	I	86	13% / 79% / 20%
1	J	86	9% / 85% / 13%
1	K	86	8% / 83% / 17%
1	L	86	9% / 87% / 13%
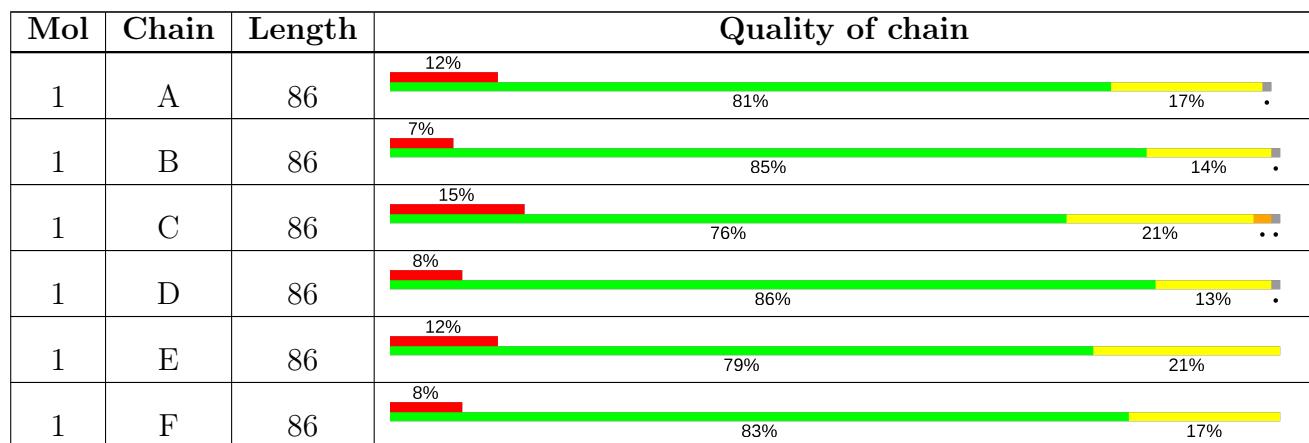
2 Entry composition ⓘ

There is only 1 type of molecule in this entry. The entry contains 17026 atoms, of which 8500 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 Nucleoporin p58/p45.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	85	Total	C	H	N	O	S	0	0	0
			1413	452	706	122	130	3			
1	B	85	Total	C	H	N	O	S	0	0	0
			1413	452	706	122	130	3			
1	C	85	Total	C	H	N	O	S	0	0	0
			1413	452	706	122	130	3			
1	D	85	Total	C	H	N	O	S	0	0	0
			1413	452	706	122	130	3			
1	E	86	Total	C	H	N	O	S	0	0	0
			1423	455	710	123	132	3			
1	F	86	Total	C	H	N	O	S	0	0	0
			1423	455	710	123	132	3			
1	G	86	Total	C	H	N	O	S	0	0	0
			1423	455	710	123	132	3			
1	H	85	Total	C	H	N	O	S	0	0	0
			1413	452	706	122	130	3			
1	I	86	Total	C	H	N	O	S	0	0	0
			1423	455	710	123	132	3			
1	J	86	Total	C	H	N	O	S	0	0	0
			1423	455	710	123	132	3			
1	K	86	Total	C	H	N	O	S	0	0	0
			1423	455	710	123	132	3			
1	L	86	Total	C	H	N	O	S	0	0	0
			1423	455	710	123	132	3			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
B	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
C	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
D	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
E	326	SER	-	EXPRESSION TAG	UNP Q9BVL2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
G	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
H	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
I	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
J	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
K	326	SER	-	EXPRESSION TAG	UNP Q9BVL2
L	326	SER	-	EXPRESSION TAG	UNP Q9BVL2

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: Nucleoporin p58/p45
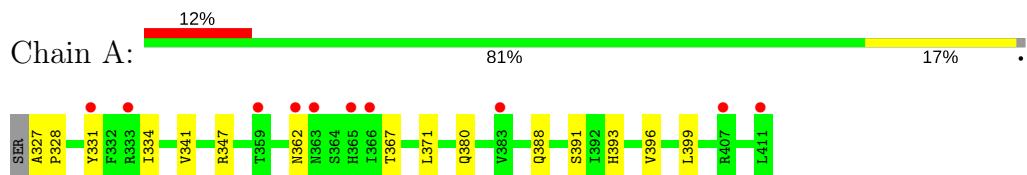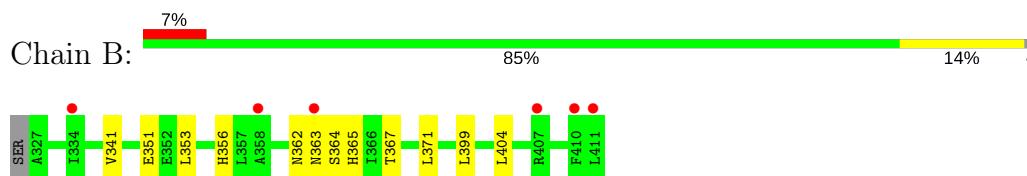
Chain A:

• Molecule 1: Nucleoporin p58/p45

Chain B:

• Molecule 1: Nucleoporin p58/p45

Chain C:

• Molecule 1: Nucleoporin p58/p45

Chain D:

• Molecule 1: Nucleoporin p58/p45

Chain E:

• Molecule 1: Nucleoporin p58/p45

Chain F:

• Molecule 1: Nucleoporin p58/p45

Chain G:

• Molecule 1: Nucleoporin p58/p45

Chain H:

• Molecule 1: Nucleoporin p58/p45

Chain I:

• Molecule 1: Nucleoporin p58/p45

Chain J:

• Molecule 1: Nucleoporin p58/p45

Chain K:

• Molecule 1: Nucleoporin p58/p45

Chain L:
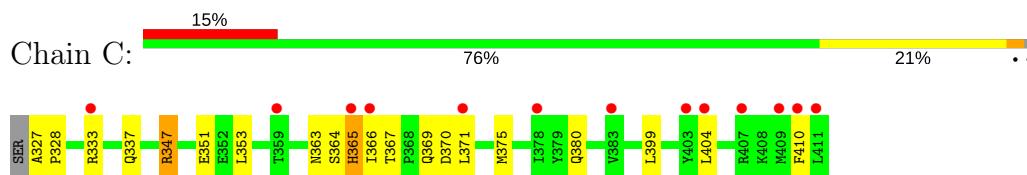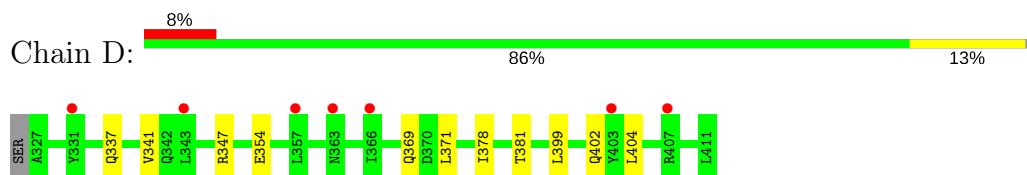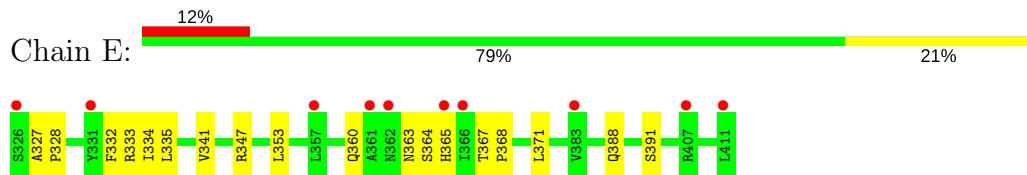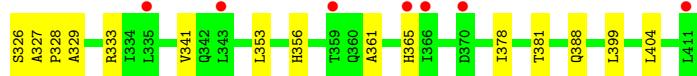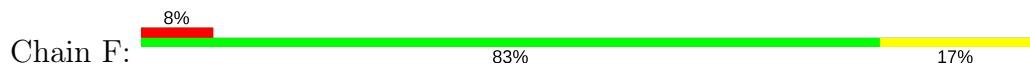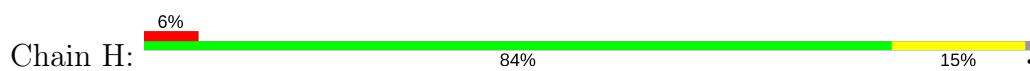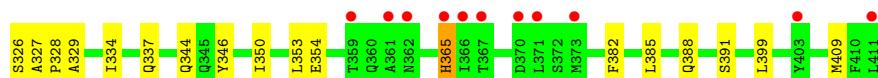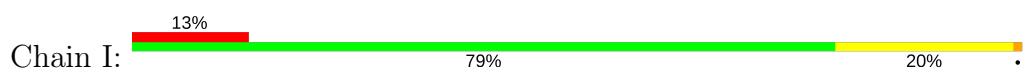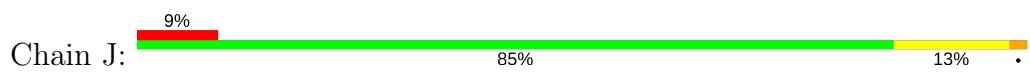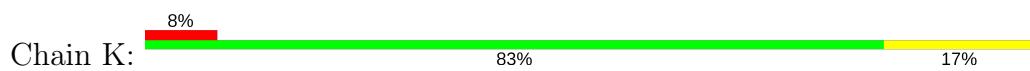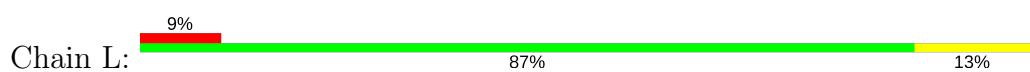
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	63.82Å 77.10Å 77.20Å 120.04° 90.12° 89.96°	Depositor
Resolution (Å)	19.83 – 2.19 19.83 – 2.19	Depositor EDS
% Data completeness (in resolution range)	97.0 (19.83-2.19) 75.2 (19.83-2.19)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.60 (at 2.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8_1069)	Depositor
R, R_{free}	0.285 , 0.293 0.284 , 0.291	Depositor DCC
R_{free} test set	1992 reflections (3.16%)	DCC
Wilson B-factor (Å^2)	42.1	Xtriage
Anisotropy	0.573	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 22.3	EDS
L-test for twinning[2]	$< |L| > = 0.53, < L^2 > = 0.37$	Xtriage
Estimated twinning fraction	0.458 for h,-l,k+l 0.458 for h,k+l,-k 0.457 for h,-k-l,k 0.457 for h,l,-k-l 0.487 for h,-k,-l 0.458 for -h,k,-k-l 0.460 for -h,l,k 0.477 for -h,-k-l,l 0.477 for -h,k+l,-l 0.459 for -h,-k,k+l 0.459 for -h,-l,-k	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	17026	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.06% 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	0.23	0/720	0.36	0/972
1	B	0.23	0/720	0.37	0/972
1	C	0.23	0/720	0.35	0/972
1	D	0.24	0/720	0.38	0/972
1	E	0.22	0/726	0.34	0/980
1	F	0.23	0/726	0.36	0/980
1	G	0.23	0/726	0.35	0/980
1	H	0.23	0/720	0.35	0/972
1	I	0.22	0/726	0.34	0/980
1	J	0.23	0/726	0.37	0/980
1	K	0.23	0/726	0.34	0/980
1	L	0.22	0/726	0.35	0/980
All	All	0.23	0/8682	0.36	0/11720

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	707	706	702	8	0
1	B	707	706	702	3	0
1	C	707	706	702	6	0
1	D	707	706	702	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	713	710	707	11	0
1	F	713	710	707	10	0
1	G	713	710	707	5	0
1	H	707	706	702	6	0
1	I	713	710	707	13	0
1	J	713	710	707	9	0
1	K	713	710	707	11	0
1	L	713	710	707	10	0
All	All	8526	8500	8459	73	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 73 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:L:326:SER:OG	1:L:329:ALA:CB	2.34	0.75
1:L:326:SER:OG	1:L:329:ALA:HB3	1.89	0.72
1:B:362:ASN:O	1:C:333:ARG:NE	2.24	0.70
1:E:332:PHE:O	1:E:335:LEU:N	2.26	0.69
1:K:358:ALA:O	1:K:362:ASN:ND2	2.30	0.65

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	83/86 (96%)	81 (98%)	2 (2%)	0	100	100
1	B	83/86 (96%)	80 (96%)	3 (4%)	0	100	100
1	C	83/86 (96%)	79 (95%)	4 (5%)	0	100	100

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	83/86 (96%)	80 (96%)	3 (4%)	0	100	100
1	E	84/86 (98%)	75 (89%)	9 (11%)	0	100	100
1	F	84/86 (98%)	84 (100%)	0	0	100	100
1	G	84/86 (98%)	83 (99%)	1 (1%)	0	100	100
1	H	83/86 (96%)	78 (94%)	5 (6%)	0	100	100
1	I	84/86 (98%)	79 (94%)	5 (6%)	0	100	100
1	J	84/86 (98%)	81 (96%)	3 (4%)	0	100	100
1	K	84/86 (98%)	82 (98%)	2 (2%)	0	100	100
1	L	84/86 (98%)	82 (98%)	2 (2%)	0	100	100
All	All	1003/1032 (97%)	964 (96%)	39 (4%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	76/77 (99%)	71 (93%)	5 (7%)	19	21
1	B	76/77 (99%)	68 (90%)	8 (10%)	8	7
1	C	76/77 (99%)	63 (83%)	13 (17%)	2	2
1	D	76/77 (99%)	69 (91%)	7 (9%)	11	10
1	E	77/77 (100%)	72 (94%)	5 (6%)	20	22
1	F	77/77 (100%)	72 (94%)	5 (6%)	20	22
1	G	77/77 (100%)	74 (96%)	3 (4%)	37	46
1	H	76/77 (99%)	71 (93%)	5 (7%)	19	21
1	I	77/77 (100%)	72 (94%)	5 (6%)	20	22
1	J	77/77 (100%)	74 (96%)	3 (4%)	37	46
1	K	77/77 (100%)	75 (97%)	2 (3%)	51	64
1	L	77/77 (100%)	76 (99%)	1 (1%)	73	85

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	919/924 (100%)	857 (93%)	62 (7%)	19 21

5 of 62 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	369	GLN
1	E	353	LEU
1	J	381	THR
1	D	402	GLN
1	E	365	HIS

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

Mol	Chain	Res	Type
1	E	342	GLN
1	E	365	HIS

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 ⓘ

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, 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(Å²)	Q<0.9
1	A	85/86 (98%)	1.09	10 (11%)	5	4	28, 50, 86, 93	0
1	B	85/86 (98%)	0.92	6 (7%)	17	16	28, 49, 81, 94	0
1	C	85/86 (98%)	1.14	13 (15%)	2	2	28, 49, 87, 93	0
1	D	85/86 (98%)	0.94	7 (8%)	12	11	28, 49, 79, 86	0
1	E	86/86 (100%)	1.12	10 (11%)	5	5	28, 50, 86, 95	0
1	F	86/86 (100%)	0.96	7 (8%)	13	12	29, 50, 83, 95	0
1	G	86/86 (100%)	1.16	13 (15%)	3	2	27, 51, 82, 97	0
1	H	85/86 (98%)	1.01	5 (5%)	23	22	29, 50, 86, 93	0
1	I	86/86 (100%)	1.03	11 (12%)	4	4	29, 49, 84, 93	0
1	J	86/86 (100%)	0.89	8 (9%)	9	8	27, 49, 80, 89	0
1	K	86/86 (100%)	1.02	7 (8%)	13	12	28, 48, 84, 92	0
1	L	86/86 (100%)	1.01	8 (9%)	9	8	27, 48, 81, 93	0
All	All	1027/1032 (99%)	1.02	105 (10%)	7	7	27, 50, 85, 97	0

The worst 5 of 105 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	366	ILE	7.7
1	E	411	LEU	6.7
1	K	326	SER	6.3
1	J	326	SER	5.5
1	C	411	LEU	5.4

6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

WORLDWIDE
wwPDB
PROTEIN DATA BANK

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.