



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 10, 2025 – 02:34 PM JST

PDB ID : 9J0A
Title : Complex structure of ANKRD11/STAG2/RAD21
Authors : Liu, H.; Cai, Q.; Zhang, M.
Deposited on : 2024-08-02
Resolution : 3.30 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

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

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.41.2

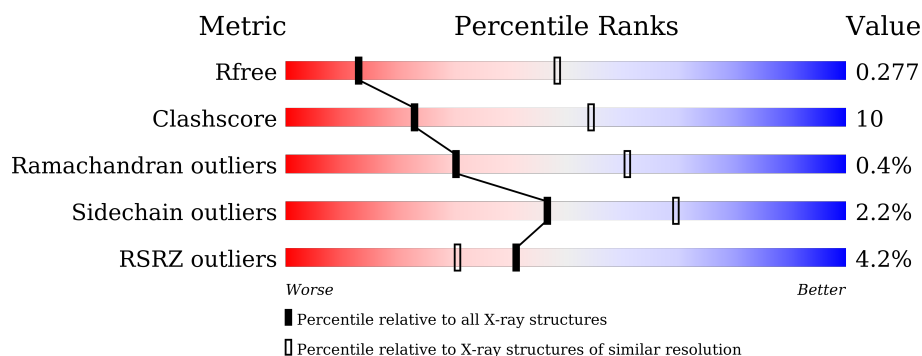
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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}	164625	1085 (3.32-3.28)
Clashscore	180529	1128 (3.32-3.28)
Ramachandran outliers	177936	1125 (3.32-3.28)
Sidechain outliers	177891	1124 (3.32-3.28)
RSRZ outliers	164620	1085 (3.32-3.28)

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 $\leq 5\%$. 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	985	<div> <div>5%</div> <div>68%</div> <div>25%</div> <div>6%</div> </div>
1	D	985	<div> <div>3%</div> <div>71%</div> <div>23%</div> <div>.</div> </div>
2	B	143	<div> <div>%</div> <div>41%</div> <div>10%</div> <div>.</div> <div>48%</div> </div>
2	E	143	<div> <div>44%</div> <div>8%</div> <div>48%</div> </div>
3	C	41	<div> <div>7%</div> <div>51%</div> <div>7%</div> <div>41%</div> </div>
3	F	41	<div> <div>7%</div> <div>37%</div> <div>20%</div> <div>.</div> <div>41%</div> </div>

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 16806 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 Cohesin subunit SA-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	926	Total	C	N	O	S	0	0	0
			7541	4800	1264	1423	54			
1	D	941	Total	C	N	O	S	0	0	0
			7635	4861	1279	1439	56			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	76	GLY	-	expression tag	UNP O35638
A	77	PRO	-	expression tag	UNP O35638
A	78	GLY	-	expression tag	UNP O35638
A	79	SER	-	expression tag	UNP O35638
D	76	GLY	-	expression tag	UNP O35638
D	77	PRO	-	expression tag	UNP O35638
D	78	GLY	-	expression tag	UNP O35638
D	79	SER	-	expression tag	UNP O35638

- Molecule 2 is a protein called 64-kDa C-terminal product.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	74	Total	C	N	O	S	0	0	0
			601	391	101	106	3			
2	E	74	Total	C	N	O	S	0	0	0
			601	391	101	106	3			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	278	GLY	-	expression tag	UNP Q61550
B	279	PRO	-	expression tag	UNP Q61550
B	280	GLY	-	expression tag	UNP Q61550
B	281	SER	-	expression tag	UNP Q61550

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	278	GLY	-	expression tag	UNP Q61550
E	279	PRO	-	expression tag	UNP Q61550
E	280	GLY	-	expression tag	UNP Q61550
E	281	SER	-	expression tag	UNP Q61550

- Molecule 3 is a protein called Ankyrin repeat domain-containing protein 11.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	24	Total	C	N	O	0	0	0
			207	127	32	48			
3	F	24	Total	C	N	O	0	0	0
			207	127	32	48			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	217	GLY	-	expression tag	UNP E9Q4F7
C	218	PRO	-	expression tag	UNP E9Q4F7
C	219	GLY	-	expression tag	UNP E9Q4F7
C	220	SER	-	expression tag	UNP E9Q4F7
F	217	GLY	-	expression tag	UNP E9Q4F7
F	218	PRO	-	expression tag	UNP E9Q4F7
F	219	GLY	-	expression tag	UNP E9Q4F7
F	220	SER	-	expression tag	UNP E9Q4F7

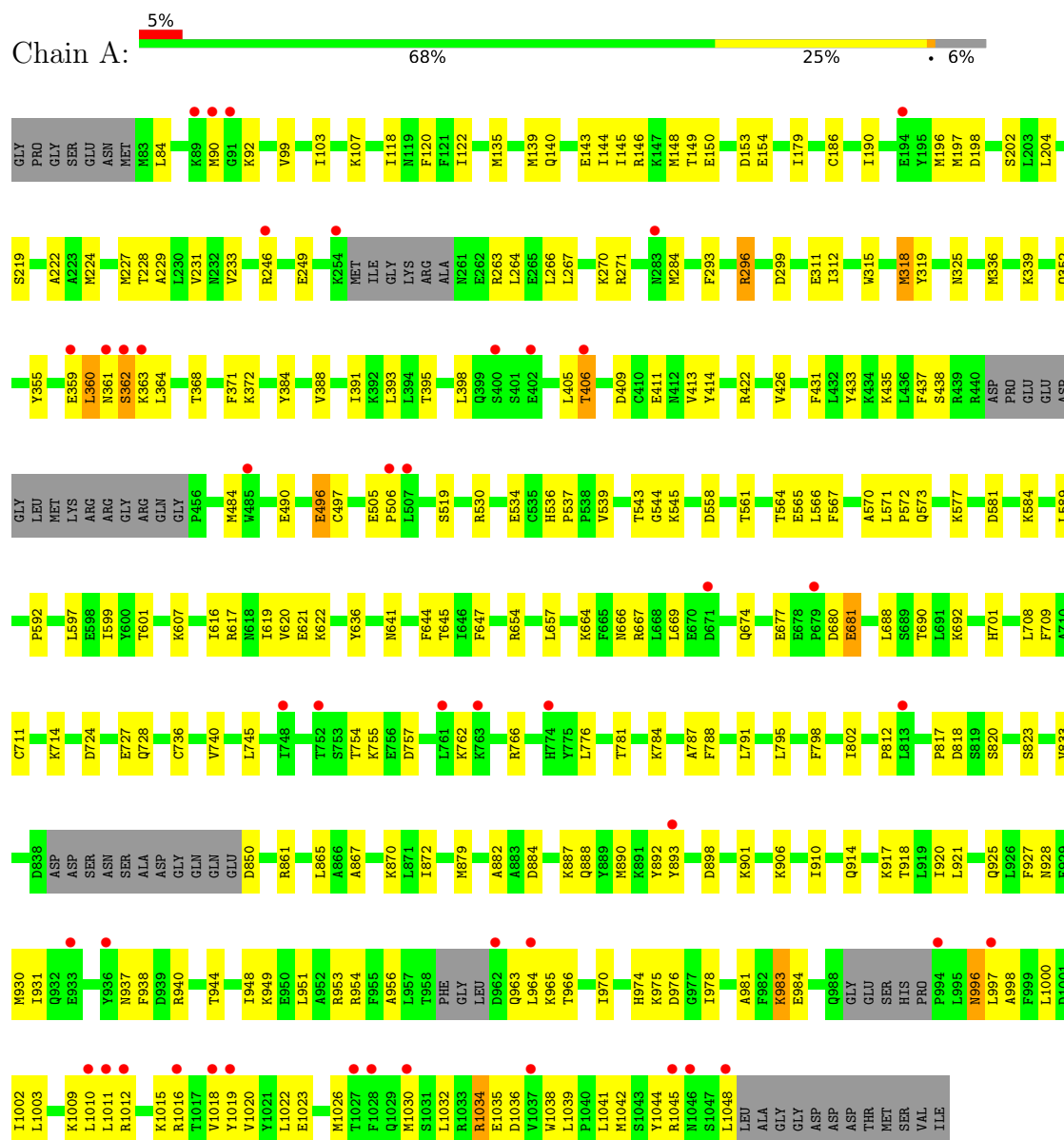
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	6	Total	O	0	0
			6	6		
4	C	1	Total	O	0	0
			1	1		
4	D	6	Total	O	0	0
			6	6		
4	E	1	Total	O	0	0
			1	1		

3 Residue-property plots

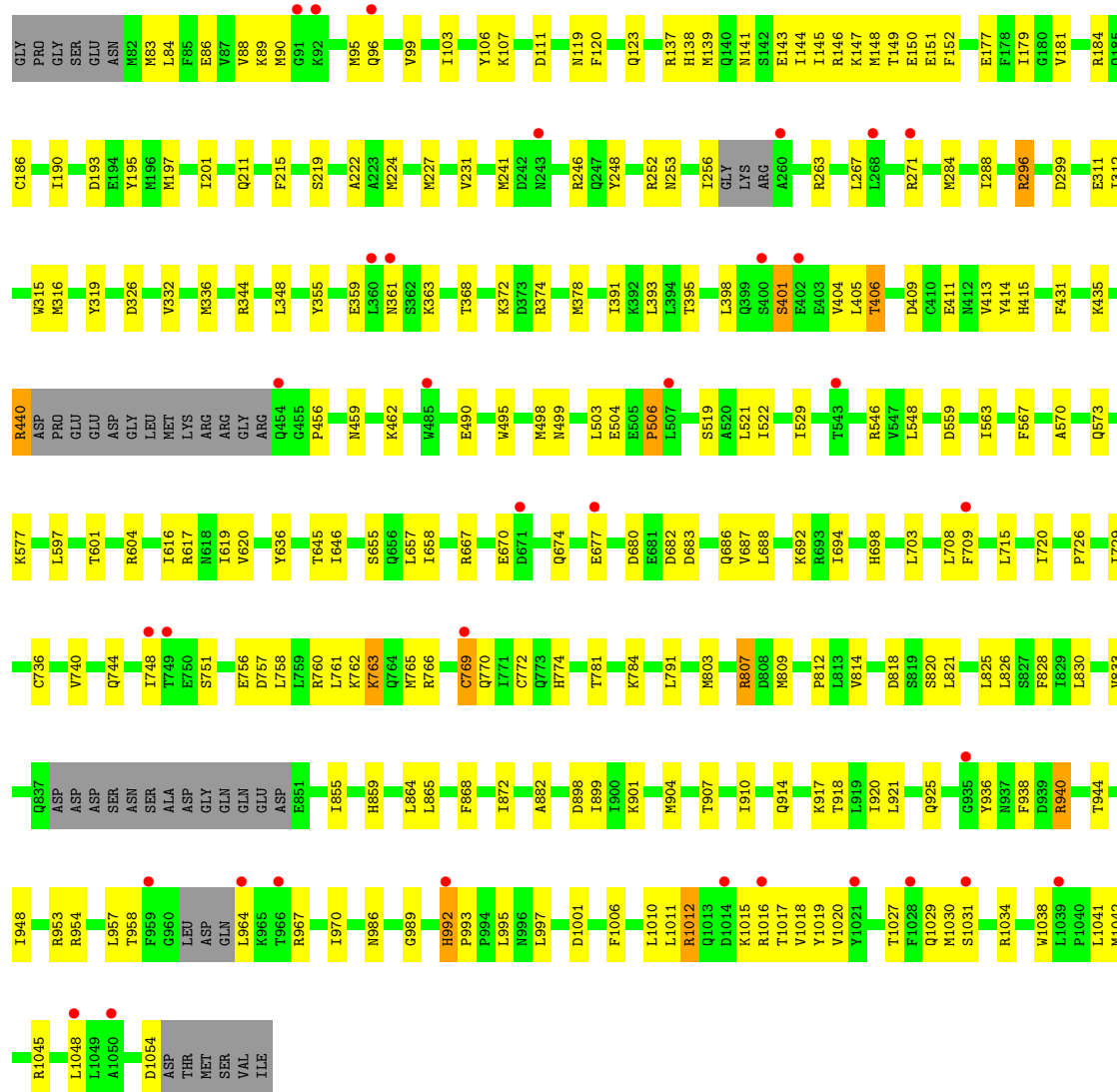
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Cohesin subunit SA-2

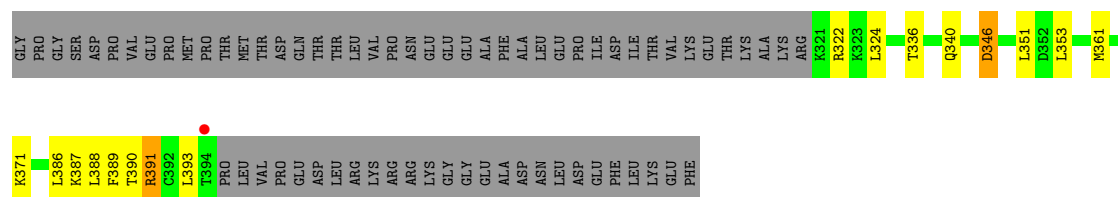
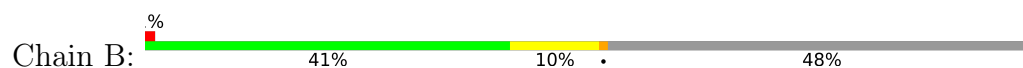


• Molecule 1: Cohesin subunit SA-2

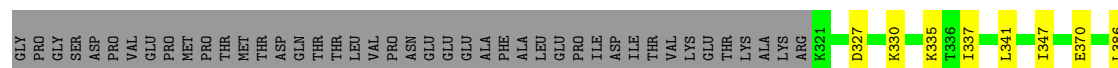
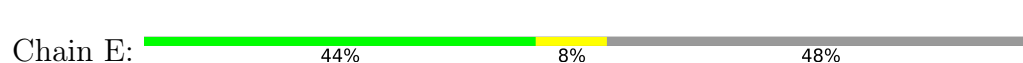




• Molecule 2: 64-kDa C-terminal product



• Molecule 2: 64-kDa C-terminal product

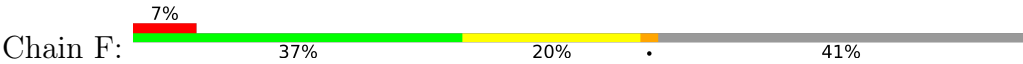




● Molecule 3: Ankyrin repeat domain-containing protein 11



● Molecule 3: Ankyrin repeat domain-containing protein 11



4 Data and refinement statistics

Property	Value	Source
Space group	I 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	280.54Å 48.77Å 290.91Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	31.88 – 3.30 31.88 – 3.30	Depositor EDS
% Data completeness (in resolution range)	99.8 (31.88-3.30) 99.8 (31.88-3.30)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.56 (at 3.33Å)	Xtriage
Refinement program	PHENIX (1.20.1_4487: ???)	Depositor
R, R_{free}	0.233 , 0.279 0.233 , 0.277	Depositor DCC
R_{free} test set	3190 reflections (5.24%)	wwPDB-VP
Wilson B-factor (Å ²)	84.8	Xtriage
Anisotropy	0.431	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 78.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.040 for -l,k,h 0.034 for -l,-k,-h 0.046 for h,-k,-l	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	16806	wwPDB-VP
Average B, all atoms (Å ²)	96.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.34% 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

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	0.27	0/7671	0.49	0/10338
1	D	0.27	0/7768	0.48	0/10471
2	B	0.25	0/612	0.51	0/826
2	E	0.24	0/612	0.50	0/826
3	C	0.25	0/211	0.46	0/285
3	F	0.24	0/211	0.47	0/285
All	All	0.27	0/17085	0.48	0/23031

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	D	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	440	ARG	Sidechain
1	D	992	HIS	Peptide

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	7541	0	7539	174	0
1	D	7635	0	7633	153	1
2	B	601	0	643	14	0
2	E	601	0	643	6	0
3	C	207	0	177	3	0
3	F	207	0	177	5	1
4	A	6	0	0	0	0
4	C	1	0	0	0	0
4	D	6	0	0	0	0
4	E	1	0	0	0	0
All	All	16806	0	16812	341	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 10.

The worst 5 of 341 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:938:PHE:O	1:A:940:ARG:NH2	2.13	0.81
2:B:351:LEU:HD22	2:B:353:LEU:HD12	1.65	0.79
1:D:141:ASN:ND2	3:F:242:ASP:OD1	2.17	0.77
1:A:325:ASN:HA	1:A:363:LYS:HG3	1.66	0.77
1:A:246:ARG:HH21	1:A:249:GLU:HB3	1.50	0.76

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)
1:D:246:ARG:NH1	3:F:227:GLU:OE2[4_445]	2.17	0.03

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	914/985 (93%)	876 (96%)	32 (4%)	6 (1%)	19	50
1	D	931/985 (94%)	900 (97%)	29 (3%)	2 (0%)	44	71
2	B	72/143 (50%)	70 (97%)	2 (3%)	0	100	100
2	E	72/143 (50%)	71 (99%)	1 (1%)	0	100	100
3	C	22/41 (54%)	19 (86%)	3 (14%)	0	100	100
3	F	22/41 (54%)	16 (73%)	6 (27%)	0	100	100
All	All	2033/2338 (87%)	1952 (96%)	73 (4%)	8 (0%)	30	61

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	406	THR
1	A	965	LYS
1	D	406	THR
1	A	360	LEU
1	A	362	SER

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	837/883 (95%)	815 (97%)	22 (3%)	41	66
1	D	845/883 (96%)	832 (98%)	13 (2%)	60	77
2	B	69/130 (53%)	67 (97%)	2 (3%)	37	63
2	E	69/130 (53%)	67 (97%)	2 (3%)	37	63
3	C	24/37 (65%)	24 (100%)	0	100	100
3	F	24/37 (65%)	22 (92%)	2 (8%)	9	31
All	All	1868/2100 (89%)	1827 (98%)	41 (2%)	47	69

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	401	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	1001	ASP
1	D	604	ARG
1	D	769	CYS
2	E	327	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	474	HIS
1	D	536	HIS
1	D	770	GLN
1	D	587	ASN
1	D	415	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 oligosaccharides 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	926/985 (94%)	0.24	45 (4%) 36 28	51, 88, 160, 197	0
1	D	941/985 (95%)	0.25	34 (3%) 46 34	47, 96, 151, 220	0
2	B	74/143 (51%)	-0.16	1 (1%) 73 60	53, 77, 101, 131	0
2	E	74/143 (51%)	-0.19	0 100 100	54, 79, 106, 124	0
3	C	24/41 (58%)	1.09	3 (12%) 9 9	70, 109, 135, 171	0
3	F	24/41 (58%)	0.76	3 (12%) 9 9	73, 105, 124, 135	0
All	All	2063/2338 (88%)	0.23	86 (4%) 41 31	47, 91, 153, 220	0

The worst 5 of 86 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	362	SER	5.2
1	A	997	LEU	4.8
3	C	241	ASP	4.6
1	D	964	LEU	3.9
1	A	1019	TYR	3.6

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 monosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.