



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 20, 2025 – 12:24 pm GMT

PDB ID : 9FH8
Title : Crystal structure of the SPD-2 domain of *Apis dorsata* CEP192
Authors : van Breugel, M.
Deposited on : 2024-05-26
Resolution : 3.50 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
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.41.5

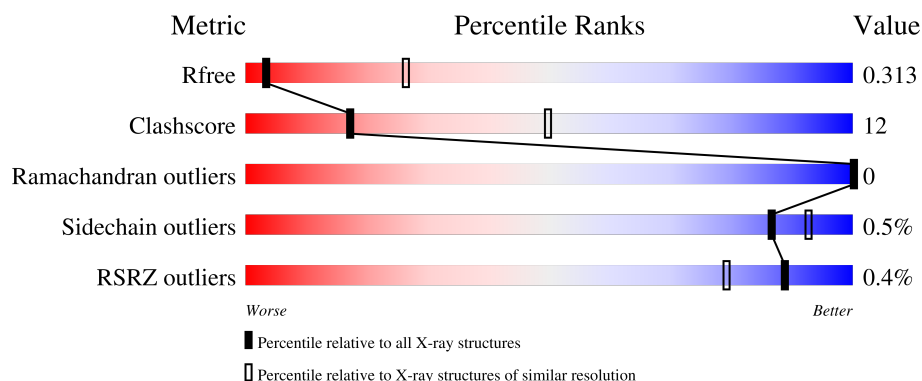
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.50 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	1094 (3.56-3.44)
Clashscore	180529	1045 (3.54-3.46)
Ramachandran outliers	177936	1032 (3.54-3.46)
Sidechain outliers	177891	1033 (3.54-3.46)
RSRZ outliers	164620	1093 (3.56-3.44)

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	341	 57% 28% 16%
1	B	341	 52% 23% 24%
1	C	341	 62% 15% 23%
1	D	341	 55% 18% 26%
1	E	341	 % 49% 26% 24%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	341	 48% 21% 31%
1	G	341	 50% 24% 26%
1	H	341	 61% 23% 16%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 16742 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.

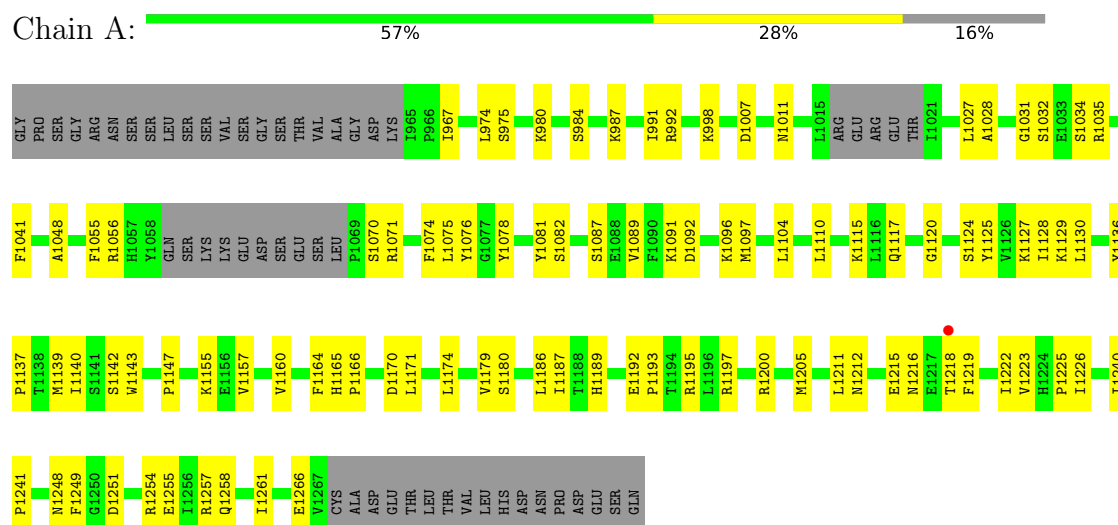
- Molecule 1 is a protein called Centrosomal protein 192 (CEP192).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	288	Total	C	N	O	S	Se	0	0	0
			2305	1469	402	425	3	6			
1	B	259	Total	C	N	O	S	Se	0	0	0
			2066	1321	352	384	3	6			
1	C	262	Total	C	N	O	S	Se	0	0	0
			2085	1327	358	390	4	6			
1	D	251	Total	C	N	O	S	Se	0	0	0
			2003	1277	342	374	4	6			
1	E	258	Total	C	N	O	S	Se	0	0	0
			2066	1321	353	383	3	6			
1	F	236	Total	C	N	O	S	Se	0	0	0
			1891	1206	326	350	3	6			
1	G	253	Total	C	N	O	S	Se	0	0	0
			2024	1289	350	375	4	6			
1	H	287	Total	C	N	O	S	Se	0	0	0
			2302	1466	402	425	3	6			

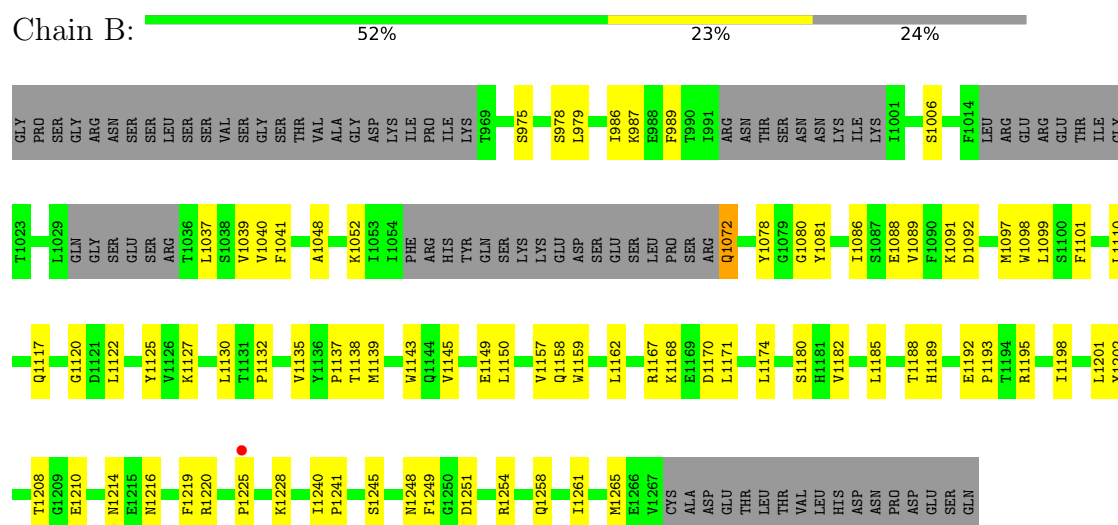
3 Residue-property plots

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: Centrosomal protein 192 (CEP192)

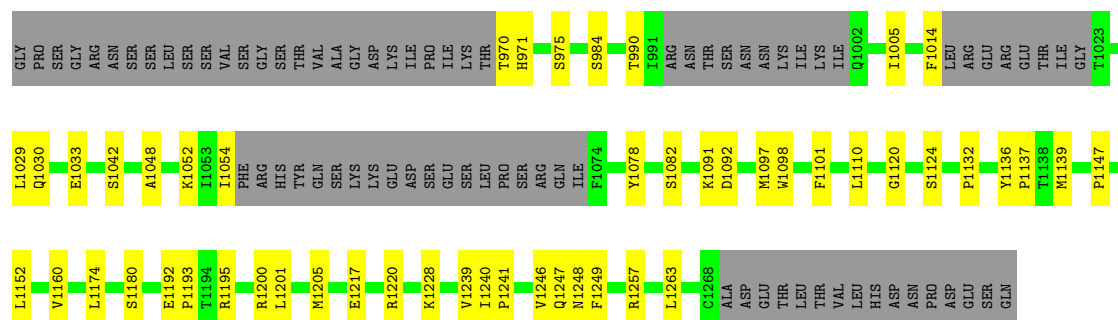


• Molecule 1: Centrosomal protein 192 (CEP192)

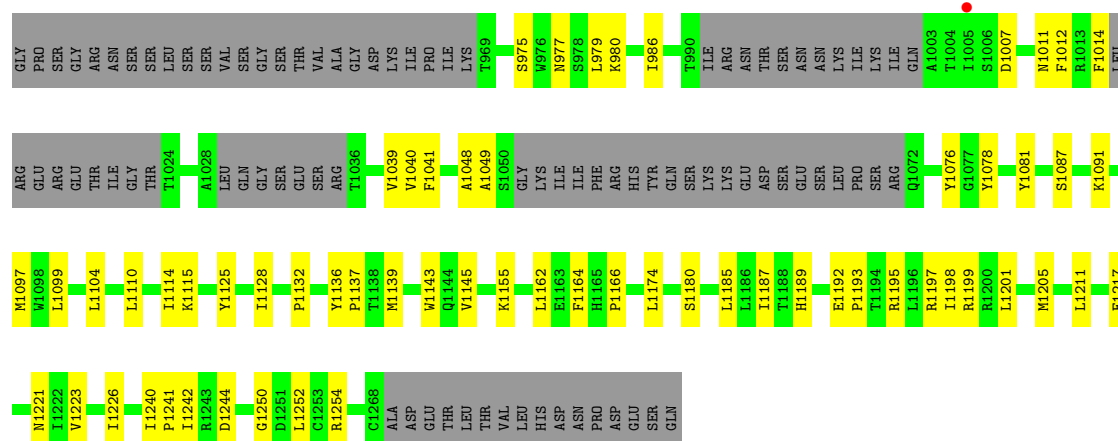


• Molecule 1: Centrosomal protein 192 (CEP192)

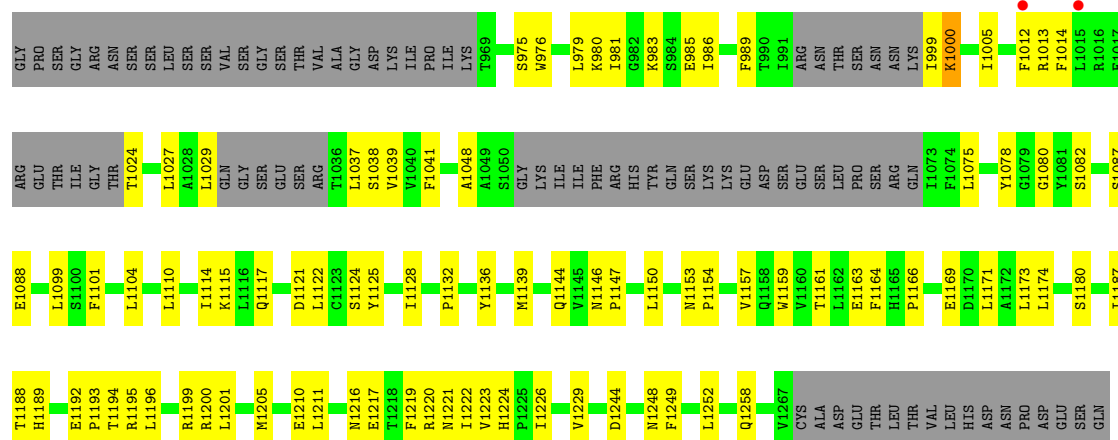




- Molecule 1: Centrosomal protein 192 (CEP192)

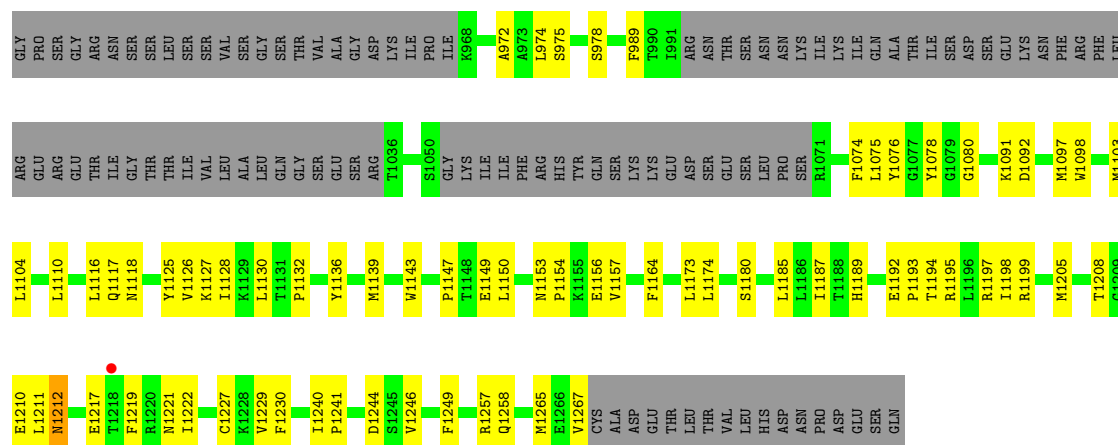


- Molecule 1: Centrosomal protein 192 (CEP192)



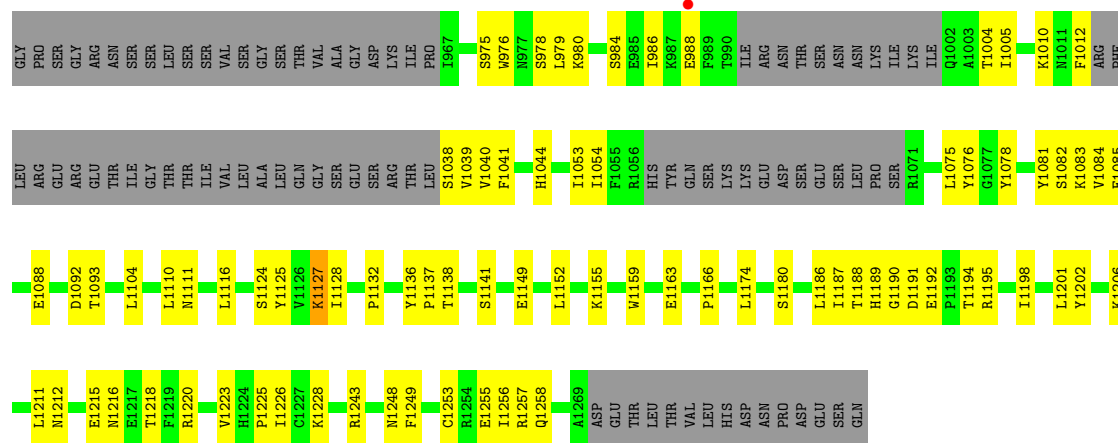
- Molecule 1: Centrosomal protein 192 (CEP192)





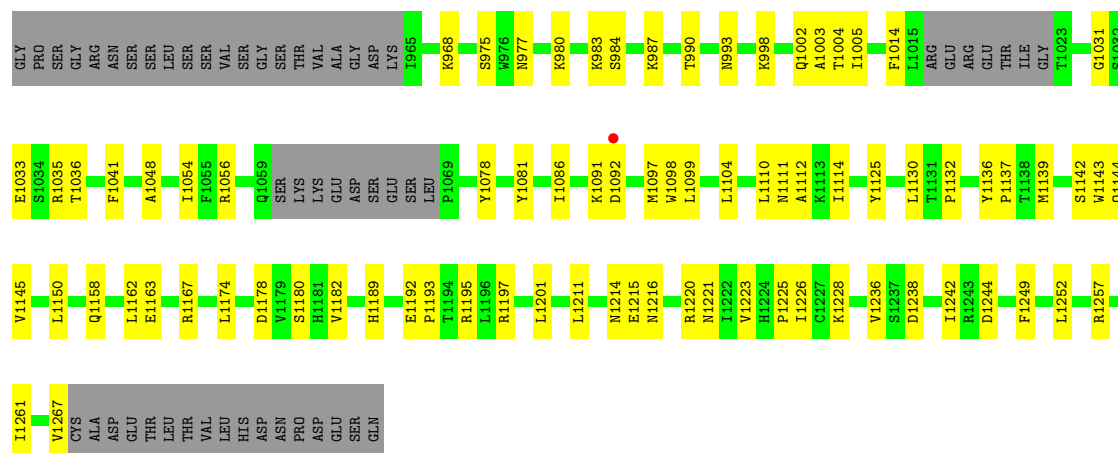
• Molecule 1: Centrosomal protein 192 (CEP192)

Chain G: 50% 24% 26%



• Molecule 1: Centrosomal protein 192 (CEP192)

Chain H: 61% 23% 16%



4 Data and refinement statistics

Property	Value	Source
Space group	I 41	Depositor
Cell constants a, b, c, α , β , γ	183.30Å 183.30Å 190.04Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.29 – 3.50 43.29 – 3.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (43.29-3.50) 99.9 (43.29-3.50)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.77 (at 3.48Å)	Xtriage
Refinement program	PHENIX (1.17.1_3660: ???)	Depositor
R, R_{free}	0.287 , 0.313 0.287 , 0.313	Depositor DCC
R_{free} test set	2026 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å ²)	119.5	Xtriage
Anisotropy	0.121	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 57.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.005 for l,-k,h 0.007 for -l,-k,-h 0.012 for -h,-l,-k 0.005 for -h,l,k 0.032 for -h,k,-l	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	16742	wwPDB-VP
Average B, all atoms (Å ²)	116.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.29% 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 [i](#)

5.1 Standard geometry [i](#)

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	0/2343	0.48	0/3150
1	B	0.27	0/2097	0.49	0/2819
1	C	0.25	0/2117	0.48	0/2845
1	D	0.25	0/2034	0.49	1/2734 (0.0%)
1	E	0.27	0/2097	0.52	0/2818
1	F	0.26	0/1922	0.48	0/2584
1	G	0.26	0/2056	0.50	0/2760
1	H	0.26	0/2340	0.50	0/3146
All	All	0.26	0/17006	0.49	1/22856 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	1211	LEU	CA-CB-CG	5.33	127.56	115.30

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	2305	0	2354	64	0
1	B	2066	0	2101	52	0
1	C	2085	0	2110	34	0
1	D	2003	0	2020	47	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	E	2066	0	2102	72	0
1	F	1891	0	1918	51	0
1	G	2024	0	2048	58	0
1	H	2302	0	2348	54	0
All	All	16742	0	17001	419	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 12.

The worst 5 of 419 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:G:1125:TYR:HE2	1:G:1226:ILE:HD11	1.41	0.83
1:B:986:ILE:HG22	1:B:1040:VAL:HG12	1.63	0.79
1:H:1002:GLN:HB3	1:H:1056:ARG:HB2	1.66	0.76
1:H:1033:GLU:HB3	1:H:1035:ARG:HH12	1.49	0.75
1:H:1192:GLU:OE2	1:H:1195:ARG:NH2	2.21	0.73

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	282/341 (83%)	271 (96%)	11 (4%)	0	100	100
1	B	249/341 (73%)	243 (98%)	6 (2%)	0	100	100
1	C	254/341 (74%)	246 (97%)	8 (3%)	0	100	100
1	D	241/341 (71%)	233 (97%)	8 (3%)	0	100	100
1	E	248/341 (73%)	235 (95%)	13 (5%)	0	100	100
1	F	230/341 (67%)	216 (94%)	14 (6%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	245/341 (72%)	227 (93%)	18 (7%)	0	100	100
1	H	281/341 (82%)	270 (96%)	11 (4%)	0	100	100
All	All	2030/2728 (74%)	1941 (96%)	89 (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	262/303 (86%)	262 (100%)	0	100	100
1	B	235/303 (78%)	233 (99%)	2 (1%)	75	86
1	C	237/303 (78%)	237 (100%)	0	100	100
1	D	228/303 (75%)	228 (100%)	0	100	100
1	E	235/303 (78%)	234 (100%)	1 (0%)	89	95
1	F	215/303 (71%)	213 (99%)	2 (1%)	75	86
1	G	229/303 (76%)	228 (100%)	1 (0%)	89	95
1	H	262/303 (86%)	259 (99%)	3 (1%)	70	83
All	All	1903/2424 (78%)	1894 (100%)	9 (0%)	86	93

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	993	ASN
1	H	1221	ASN
1	F	1118	ASN
1	F	1212	ASN
1	G	1127	LYS

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	1216	ASN
1	F	1221	ASN
1	G	1258	GLN
1	F	1258	GLN
1	C	1247	GLN

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	282/341 (82%)	-0.23	1 (0%) 89 79	64, 106, 140, 171	0
1	B	253/341 (74%)	-0.05	1 (0%) 89 79	72, 117, 150, 182	0
1	C	256/341 (75%)	-0.14	0 100 100	69, 112, 154, 168	0
1	D	245/341 (71%)	-0.06	1 (0%) 89 79	73, 117, 160, 191	0
1	E	252/341 (73%)	-0.11	2 (0%) 82 67	71, 117, 166, 171	0
1	F	230/341 (67%)	-0.05	1 (0%) 89 79	76, 126, 165, 173	0
1	G	247/341 (72%)	-0.12	1 (0%) 89 79	73, 120, 162, 174	0
1	H	281/341 (82%)	-0.17	1 (0%) 89 79	71, 108, 154, 172	0
All	All	2046/2728 (75%)	-0.12	8 (0%) 89 79	64, 115, 159, 191	0

The worst 5 of 8 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	1015	LEU	2.7
1	G	988	GLU	2.5
1	H	1092	ASP	2.4
1	A	1218	THR	2.3
1	B	1225	PRO	2.1

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.