Full wwPDB X-ray Structure Validation Report ⓘ

Jan 14, 2024 – 05:10 am GMT

PDB ID	:	6RRK
Title	:	Crystal structure of the central region of human cohesin subunit STAG1 in complex with RAD21 peptide
Authors	:	Newman, J.A.; katis, V.L.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.; Bountra, C.; Gileadi, O.
Deposited on	:	2019-05-20
Resolution	:	3.17 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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:

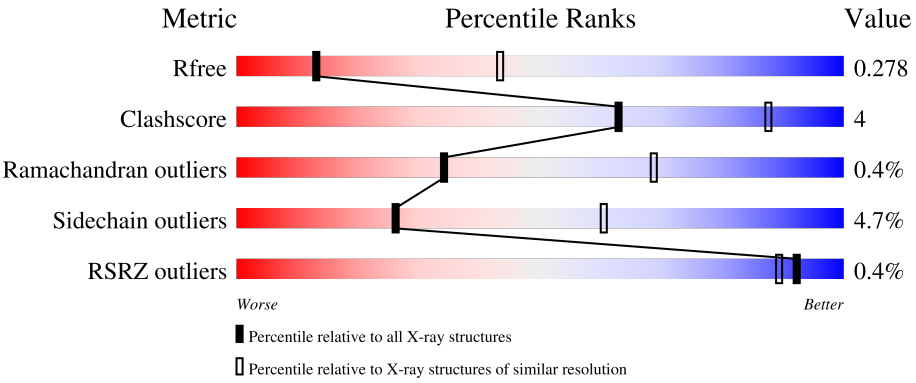
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.17 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	1467 (3.20-3.16)
Clashscore	141614	1599 (3.20-3.16)
Ramachandran outliers	138981	1574 (3.20-3.16)
Sidechain outliers	138945	1573 (3.20-3.16)
RSRZ outliers	127900	1423 (3.20-3.16)

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	459	% 79% 15% 6%
1	B	459	79% 15% • 6%
2	C	40	50% 18% 32%
2	D	40	68% 32%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 7289 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-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	431	Total	C	N	O	S	0	0	0
			3419	2179	566	651	23			
1	A	430	Total	C	N	O	S	0	0	0
			3426	2181	569	653	23			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	457	SER	-	expression tag	UNP Q8WVM7
B	458	MET	-	expression tag	UNP Q8WVM7
A	457	SER	-	expression tag	UNP Q8WVM7
A	458	MET	-	expression tag	UNP Q8WVM7

- Molecule 2 is a protein called Double-strand-break repair protein rad21 homolog.

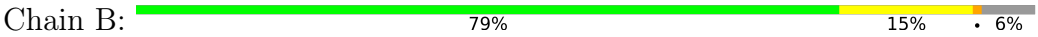
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	27	Total	C	N	O	S	0	0	0
			222	147	39	35	1			
2	D	27	Total	C	N	O	S	0	0	0
			222	147	39	35	1			

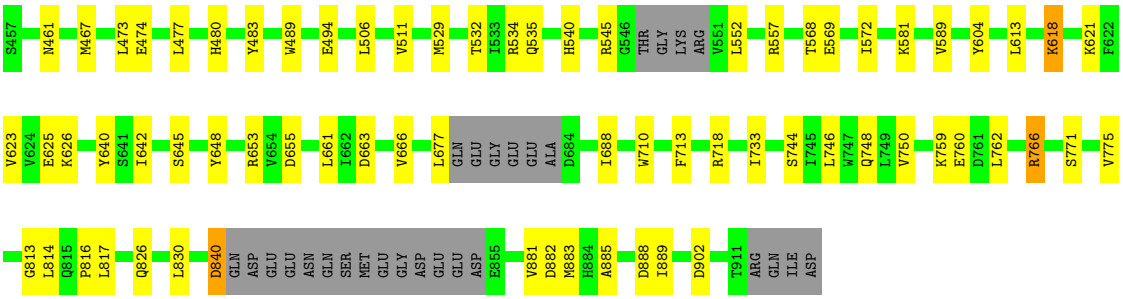
3 Residue-property plots (i)

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: Cohesin subunit SA-1

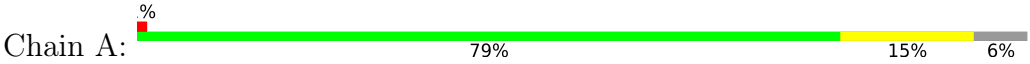
Chain B:

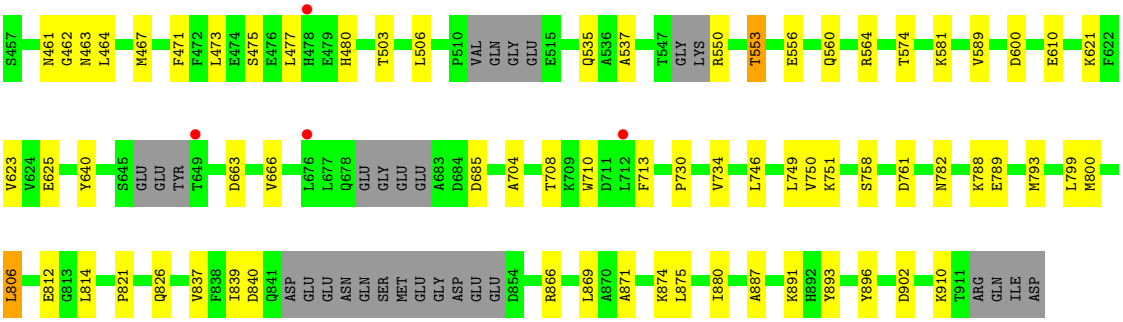




• Molecule 1: Cohesin subunit SA-1

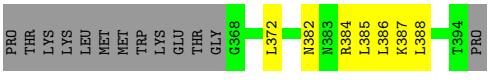
Chain A:





• Molecule 2: Double-strand-break repair protein rad21 homolog

Chain C:



• Molecule 2: Double-strand-break repair protein rad21 homolog

Chain D:


PRO
THR
LYS
LYS
LEU
MET
MET
TRP
LYS
GLU
THR
GLY
G368
T394
PRO

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	74.33Å 86.29Å 118.58Å 90.00° 95.36° 90.00°	Depositor
Resolution (Å)	60.23 – 3.17 60.23 – 3.17	Depositor EDS
% Data completeness (in resolution range)	96.4 (60.23-3.17) 96.4 (60.23-3.17)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.81 (at 3.19Å)	Xtriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.219 , 0.278 0.219 , 0.278	Depositor DCC
R_{free} test set	1180 reflections (4.79%)	wwPDB-VP
Wilson B-factor (Å^2)	99.6	Xtriage
Anisotropy	0.343	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 64.9	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	0.95	EDS
Total number of atoms	7289	wwPDB-VP
Average B, all atoms (Å^2)	121.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.47% 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	0.24	0/3485	0.39	0/4713
1	B	0.24	0/3480	0.39	0/4709
2	C	0.24	0/227	0.37	0/307
2	D	0.24	0/227	0.35	0/307
All	All	0.24	0/7419	0.39	0/10036

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	3426	0	3384	32	0
1	B	3419	0	3371	30	0
2	C	222	0	237	5	0
2	D	222	0	237	0	0
All	All	7289	0	7229	63	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.

All (63) 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:477:LEU:HB3	1:B:480:HIS:HB2	1.76	0.67
1:A:589:VAL:HG11	1:A:623:VAL:HG11	1.82	0.62
1:B:589:VAL:HG11	1:B:623:VAL:HG11	1.83	0.59
1:A:871:ALA:HA	1:A:874:LYS:HE2	1.85	0.59
1:B:688:ILE:HG23	1:B:733:ILE:HD11	1.85	0.58
1:A:866:ARG:NH2	1:A:902:ASP:OD2	2.37	0.58
1:B:677:LEU:HD11	1:B:718:ARG:HH22	1.68	0.57
1:A:471:PHE:O	1:A:475:SER:OG	2.20	0.56
1:A:564:ARG:NH2	1:A:600:ASP:OD1	2.38	0.56
1:A:839:ILE:O	1:A:896:TYR:OH	2.21	0.55
1:A:506:LEU:O	1:A:581:LYS:NZ	2.31	0.54
1:B:480:HIS:HA	1:B:483:TYR:HD2	1.72	0.54
1:B:613:LEU:HD23	1:B:653:ARG:HD2	1.89	0.54
1:A:746:LEU:HB3	2:C:385:LEU:HD22	1.90	0.53
1:A:902:ASP:OD1	1:A:902:ASP:N	2.41	0.53
1:B:572:ILE:HD12	1:B:604:TYR:CZ	2.43	0.53
1:B:552:LEU:HB2	1:B:557:ARG:HH21	1.73	0.53
1:A:461:ASN:O	1:A:464:LEU:N	2.41	0.53
1:B:529:MET:O	1:B:532:THR:OG1	2.23	0.53
1:A:821:PRO:O	1:A:826:GLN:NE2	2.40	0.52
1:B:814:LEU:HD22	1:B:817:LEU:HD12	1.92	0.51
1:B:568:THR:O	1:B:572:ILE:HG12	2.10	0.51
1:A:799:LEU:HD22	1:A:821:PRO:HG3	1.92	0.51
2:C:384:ARG:HA	2:C:387:LYS:HD2	1.93	0.51
1:A:704:ALA:HA	2:C:382:ASN:HB2	1.93	0.49
1:B:642:ILE:O	1:B:645:SER:OG	2.20	0.49
1:B:885:ALA:O	1:B:889:ILE:HG13	2.13	0.48
1:B:506:LEU:O	1:B:581:LYS:NZ	2.38	0.47
1:A:749:LEU:HD23	2:C:388:LEU:HD13	1.97	0.47
1:A:666:VAL:HG21	1:A:710:TRP:CD1	2.51	0.46
1:A:553:THR:OG1	1:A:556:GLU:OE1	2.26	0.46
1:A:887:ALA:O	1:A:891:LYS:HG2	2.15	0.46
1:B:766:ARG:NE	1:B:816:PRO:O	2.30	0.45
1:A:874:LYS:NZ	2:C:372:LEU:O	2.50	0.45
1:B:621:LYS:HG3	1:B:661:LEU:HD13	1.99	0.45
1:B:744:SER:O	1:B:748:GLN:HG3	2.17	0.45
1:B:489:TRP:CD2	1:B:535:GLN:HG2	2.52	0.45
1:B:746:LEU:O	1:B:750:VAL:HG23	2.17	0.44
1:B:534:ARG:NH2	1:B:540:HIS:O	2.47	0.44
1:A:758:SER:HB3	1:A:761:ASP:OD1	2.16	0.44
1:A:461:ASN:O	1:A:463:ASN:N	2.51	0.44
1:B:666:VAL:HG21	1:B:710:TRP:CD1	2.53	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:762:LEU:O	1:B:766:ARG:HB2	2.18	0.43
1:B:771:SER:O	1:B:775:VAL:HG23	2.18	0.43
1:A:537:ALA:O	1:A:560:GLN:NE2	2.47	0.43
1:A:806:LEU:HD11	1:A:814:LEU:HD13	2.01	0.43
1:A:789:GLU:O	1:A:793:MET:HG3	2.19	0.43
1:B:473:LEU:HD12	1:B:473:LEU:HA	1.87	0.43
1:A:503:THR:HG21	1:A:574:THR:HG21	2.01	0.42
1:A:782:ASN:O	1:A:788:LYS:HE3	2.18	0.42
1:B:621:LYS:O	1:B:625:GLU:HG3	2.19	0.42
1:A:621:LYS:O	1:A:625:GLU:HG3	2.19	0.42
1:B:762:LEU:HD21	1:B:813:GLY:O	2.19	0.42
1:A:708:THR:HG21	1:A:751:LYS:NZ	2.34	0.42
1:A:746:LEU:O	1:A:750:VAL:HG23	2.20	0.42
1:A:837:VAL:HG11	1:A:869:LEU:HB2	2.01	0.42
1:B:840:ASP:OD1	1:B:840:ASP:N	2.44	0.41
1:A:621:LYS:NZ	1:A:625:GLU:OE2	2.51	0.41
1:B:618:LYS:HE3	1:B:618:LYS:HB3	1.80	0.41
1:A:730:PRO:O	1:A:734:VAL:HG23	2.20	0.41
1:B:826:GLN:HG2	1:B:881:VAL:HG12	2.02	0.41
1:B:483:TYR:HB3	1:B:545:ARG:NH2	2.36	0.41
1:A:875:LEU:HD22	1:A:880:ILE:HG21	2.03	0.40

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	418/459 (91%)	404 (97%)	13 (3%)	1 (0%)	47 78
1	B	423/459 (92%)	400 (95%)	20 (5%)	3 (1%)	22 60
2	C	25/40 (62%)	24 (96%)	1 (4%)	0	100 100
2	D	25/40 (62%)	24 (96%)	1 (4%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	891/998 (89%)	852 (96%)	35 (4%)	4 (0%)	34 69

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	511	VAL
1	A	462	GLY
1	B	648	TYR
1	B	461	ASN

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	382/415 (92%)	364 (95%)	18 (5%)	26 60
1	B	379/415 (91%)	360 (95%)	19 (5%)	24 58
2	C	25/37 (68%)	24 (96%)	1 (4%)	31 64
2	D	25/37 (68%)	25 (100%)	0	100 100
All	All	811/904 (90%)	773 (95%)	38 (5%)	26 60

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	467	MET
1	B	474	GLU
1	B	494	GLU
1	B	569	GLU
1	B	618	LYS
1	B	626	LYS
1	B	640	TYR
1	B	655	ASP
1	B	663	ASP
1	B	713	PHE
1	B	759	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	760	GLU
1	B	766	ARG
1	B	830	LEU
1	B	840	ASP
1	B	882	ASP
1	B	883	MET
1	B	888	ASP
1	B	902	ASP
1	A	467	MET
1	A	473	LEU
1	A	477	LEU
1	A	480	HIS
1	A	535	GLN
1	A	550	ARG
1	A	553	THR
1	A	610	GLU
1	A	640	TYR
1	A	663	ASP
1	A	685	ASP
1	A	713	PHE
1	A	800	MET
1	A	806	LEU
1	A	812	GLU
1	A	840	ASP
1	A	893	TYR
1	A	910	LYS
2	C	386	LEU

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

Mol	Chain	Res	Type
1	B	493	GLN
1	B	777	GLN

5.3.3 RNA

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 monosaccharides 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, 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	430/459 (93%)	0.02	4 (0%) 84 75	75, 120, 166, 184	0
1	B	431/459 (93%)	-0.08	0 100 100	65, 116, 163, 193	0
2	C	27/40 (67%)	0.01	0 100 100	95, 119, 148, 153	0
2	D	27/40 (67%)	0.03	0 100 100	100, 124, 144, 155	0
All	All	915/998 (91%)	-0.03	4 (0%) 92 89	65, 118, 165, 193	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	676	LEU	2.7
1	A	649	THR	2.6
1	A	478	HIS	2.2
1	A	712	LEU	2.1

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.