



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 5, 2024 – 04:12 pm BST

PDB ID : 8RO8
Title : Human cohesin SMC1A-HD(longCC-EQ)/RAD21-C complex - Apo closed
P-loop conformation
Authors : Vitoria Gomes, M.; Romier, C.
Deposited on : 2024-01-11
Resolution : 1.90 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 : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.002 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.38.2

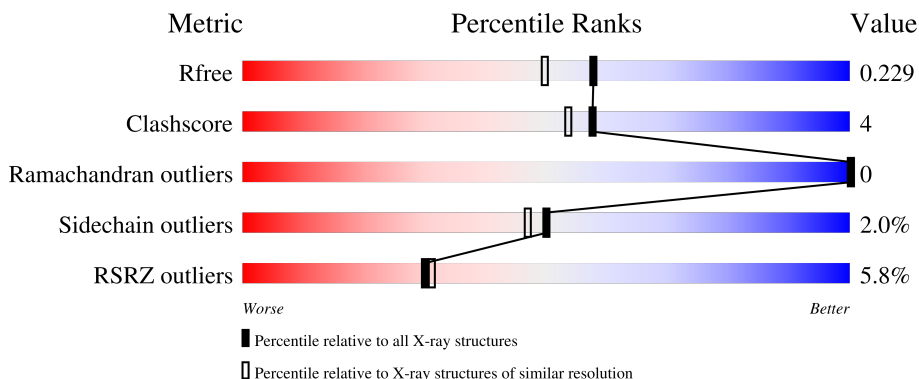
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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	456	 5% 67% 9% 23%
2	B	81	 2% 72% 6% 22%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3420 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 Structural maintenance of chromosomes protein 1A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	350	2788	1783	472	524	9	0	6	0

There are 15 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	978	GLU	-	linker	UNP Q14683
A	979	SER	-	linker	UNP Q14683
A	980	SER	-	linker	UNP Q14683
A	981	LYS	-	linker	UNP Q14683
A	982	HIS	-	linker	UNP Q14683
A	983	PRO	-	linker	UNP Q14683
A	984	THR	-	linker	UNP Q14683
A	985	SER	-	linker	UNP Q14683
A	986	LEU	-	linker	UNP Q14683
A	987	VAL	-	linker	UNP Q14683
A	988	PRO	-	linker	UNP Q14683
A	989	ARG	-	linker	UNP Q14683
A	990	GLY	-	linker	UNP Q14683
A	991	SER	-	linker	UNP Q14683
A	1157	GLN	GLU	engineered mutation	UNP Q14683

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	63	500	320	85	94	1	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	557	MET	-	initiating methionine	UNP O60216

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	630	GLY	-	expression tag	UNP O60216
B	631	SER	-	expression tag	UNP O60216
B	632	LEU	-	expression tag	UNP O60216
B	633	GLU	-	expression tag	UNP O60216
B	634	VAL	-	expression tag	UNP O60216
B	635	LEU	-	expression tag	UNP O60216
B	636	PHE	-	expression tag	UNP O60216
B	637	GLN	-	expression tag	UNP O60216

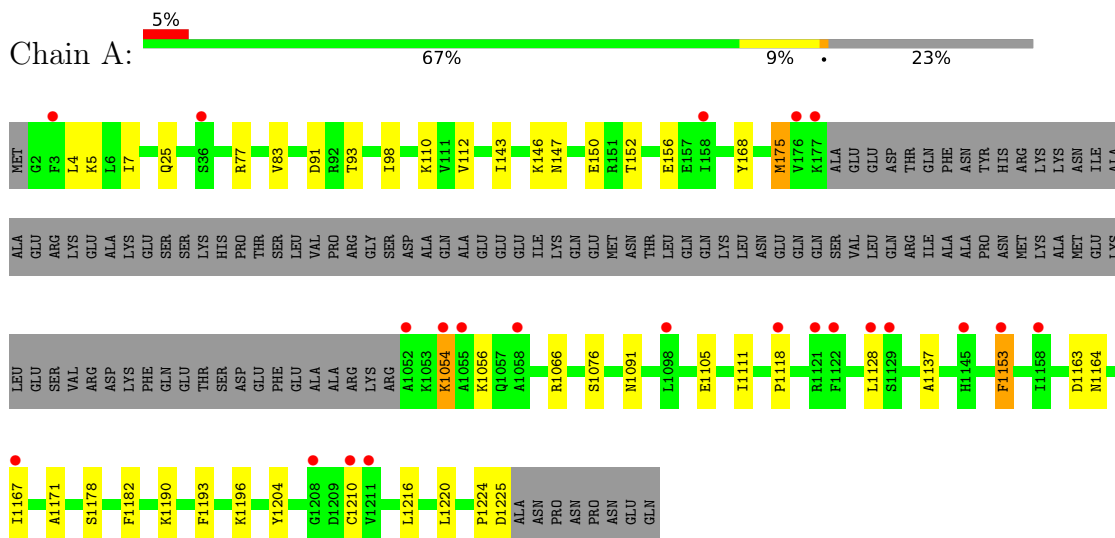
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	96	Total O 96 96	0	0
3	B	36	Total O 36 36	0	0

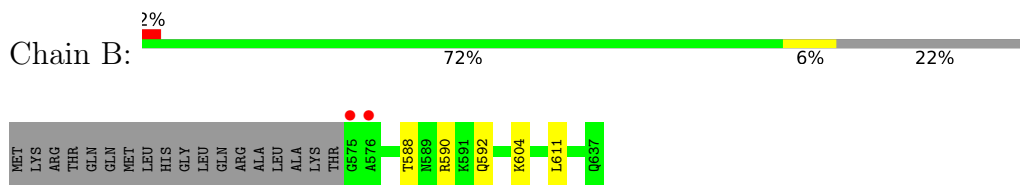
3 Residue-property plots [i](#)

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: Structural maintenance of chromosomes protein 1A



- Molecule 2: 64-kDa C-terminal product



4 Data and refinement statistics i

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	158.55Å 67.07Å 51.63Å 90.00° 92.32° 90.00°	Depositor
Resolution (Å)	44.06 – 1.90 44.06 – 1.90	Depositor EDS
% Data completeness (in resolution range)	99.4 (44.06-1.90) 99.7 (44.06-1.90)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.48 (at 1.91Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.195 , 0.231 0.194 , 0.229	Depositor DCC
R_{free} test set	2092 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å ²)	38.2	Xtrriage
Anisotropy	0.661	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 54.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.026 for -h,-k,l	Xtrriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3420	wwPDB-VP
Average B, all atoms (Å ²)	62.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.16% 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	0.34	0/2857	0.52	0/3844
2	B	0.33	0/509	0.52	0/686
All	All	0.34	0/3366	0.52	0/4530

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 [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	2788	0	2823	28	0
2	B	500	0	504	4	0
3	A	96	0	0	4	1
3	B	36	0	0	0	0
All	All	3420	0	3327	29	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 4.

All (29) 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:1091:ASN:HD22	1:A:1118:PRO:HB3	1.45	0.80

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1216:LEU:HD11	2:B:611:LEU:HD13	1.75	0.68
1:A:1220[B]:LEU:HG	2:B:590:ARG:HG2	1.81	0.62
1:A:77:ARG:NH1	3:A:1303:HOH:O	2.33	0.61
1:A:83:VAL:HG22	1:A:93[A]:THR:HG22	1.83	0.60
1:A:168:TYR:OH	1:A:1105:GLU:OE2	2.18	0.59
1:A:156:GLU:OE2	1:A:1066:ARG:NH2	2.37	0.58
1:A:1171:ALA:HB1	1:A:1196:LYS:HG3	1.85	0.57
1:A:175:MET:HG3	1:A:1056:LYS:HA	1.85	0.57
1:A:143:ILE:HD12	1:A:146:LYS:HD2	1.90	0.54
1:A:1091:ASN:ND2	1:A:1118:PRO:HB3	2.20	0.54
1:A:143:ILE:HG23	1:A:1137:ALA:HB1	1.90	0.54
1:A:1167:ILE:HD11	1:A:1193:PHE:HB2	1.91	0.53
1:A:25[A]:GLN:OE1	3:A:1301:HOH:O	2.19	0.49
1:A:1178:SER:HA	1:A:1182:PHE:O	2.13	0.48
1:A:1164:ASN:OD1	1:A:1164:ASN:N	2.45	0.48
1:A:25[A]:GLN:NE2	3:A:1304:HOH:O	2.36	0.47
1:A:1054:LYS:HE3	1:A:1054:LYS:HB2	1.60	0.46
1:A:5:LYS:NZ	1:A:91:ASP:OD2	2.36	0.45
1:A:4:LEU:HD21	1:A:7:ILE:HG13	1.99	0.43
1:A:77:ARG:HA	1:A:98:ILE:O	2.18	0.43
1:A:1190:LYS:NZ	3:A:1313:HOH:O	2.52	0.42
1:A:1204:TYR:CD2	2:B:604:LYS:HG2	2.54	0.42
1:A:1224:PRO:O	1:A:1225:ASP:HB2	2.20	0.42
1:A:152:THR:OG1	1:A:1111:ILE:HG12	2.20	0.41
1:A:1153:PHE:HD1	1:A:1153:PHE:HA	1.74	0.41
1:A:110:LYS:HE2	1:A:112:VAL:HG12	2.02	0.41
1:A:147:ASN:OD1	1:A:150:GLU:N	2.47	0.41
2:B:588:THR:HB	2:B:592:GLN:HB2	2.04	0.40

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)
3:A:1304:HOH:O	3:A:1385:HOH:O[4_445]	2.18	0.02

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	352/456 (77%)	343 (97%)	9 (3%)	0	100	100
2	B	61/81 (75%)	61 (100%)	0	0	100	100
All	All	413/537 (77%)	404 (98%)	9 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	301/388 (78%)	294 (98%)	7 (2%)	45	41
2	B	54/69 (78%)	54 (100%)	0	100	100
All	All	355/457 (78%)	348 (98%)	7 (2%)	50	47

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	175	MET
1	A	1054	LYS
1	A	1076	SER
1	A	1128	LEU
1	A	1153	PHE
1	A	1163	ASP
1	A	1210	CYS

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

Mol	Chain	Res	Type
1	A	40	ASN
1	A	1091	ASN

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	350/456 (76%)	0.58	22 (6%) 27 28	27, 60, 110, 139	6 (1%)
2	B	63/81 (77%)	0.25	2 (3%) 50 52	38, 49, 73, 95	0
All	All	413/537 (76%)	0.53	24 (5%) 30 31	27, 57, 109, 139	6 (1%)

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1052	ALA	5.1
2	B	575	GLY	3.5
1	A	1208	GLY	3.3
1	A	1158	ILE	3.3
1	A	36	SER	3.0
1	A	176	VAL	3.0
1	A	1129	SER	2.7
1	A	1122	PHE	2.7
1	A	1210	CYS	2.7
1	A	3	PHE	2.5
1	A	1054	LYS	2.4
1	A	177	LYS	2.3
1	A	1145[A]	HIS	2.3
1	A	1058	ALA	2.2
1	A	1167	ILE	2.2
1	A	1121	ARG	2.2
1	A	1211	VAL	2.2
1	A	1055	ALA	2.2
1	A	1098	LEU	2.2
1	A	1128	LEU	2.1
1	A	1153	PHE	2.1
1	A	1118	PRO	2.1
1	A	158	ILE	2.1
2	B	576	ALA	2.0

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.