



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 5, 2024 – 04:14 pm BST

PDB ID : 8RO7
Title : Human cohesin SMC1A-HD(longCC)/RAD21-C complex - Open/closed
P-loop conformation
Authors : Shaik, T.B.; Romier, C.
Deposited on : 2024-01-11
Resolution : 2.09 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

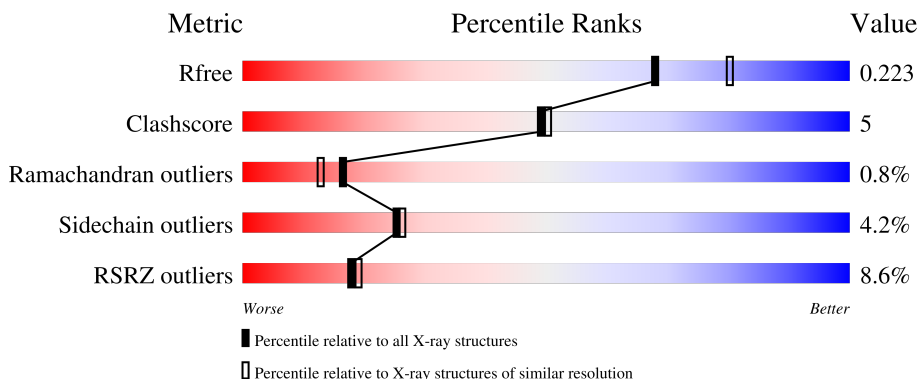
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


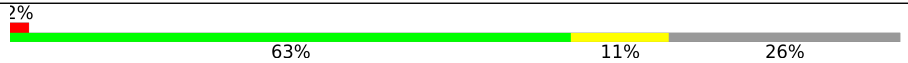
The reported resolution of this entry is 2.09 Å.

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	6234 (2.10-2.10)
Clashscore	180529	6893 (2.10-2.10)
Ramachandran outliers	177936	6839 (2.10-2.10)
Sidechain outliers	177891	6840 (2.10-2.10)
RSRZ outliers	164620	6234 (2.10-2.10)

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	
2	B	81	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 3188 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 (SMC1A).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	335	2654	1693	448	505	8	0	4	0

- 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	60	482	310	81	90	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
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

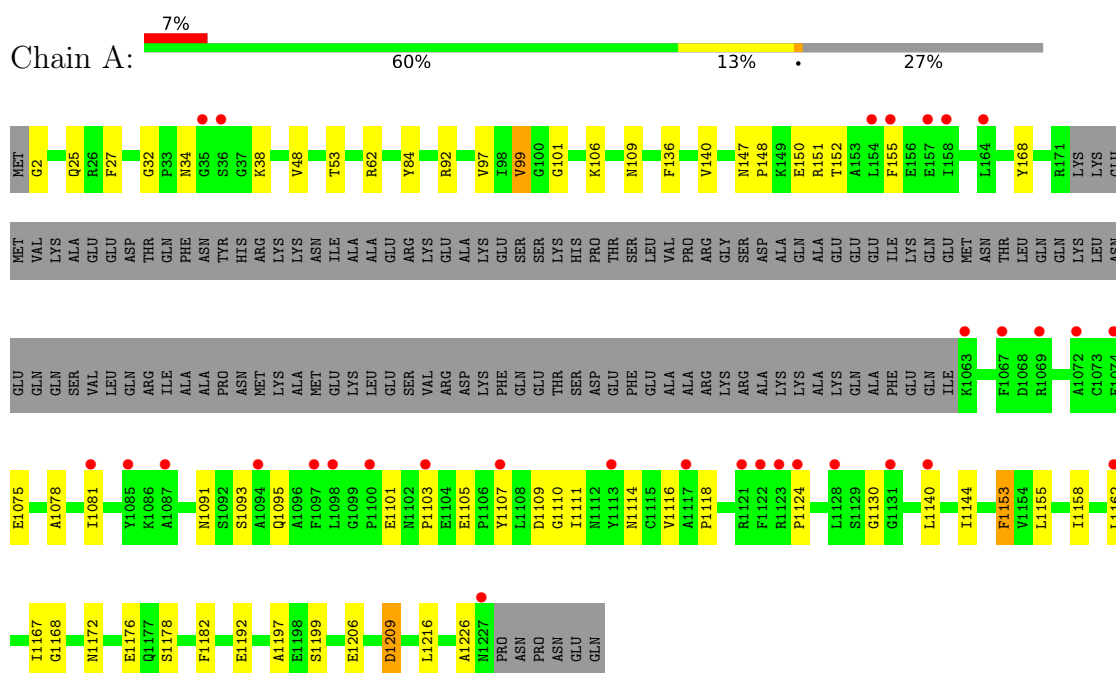
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	42	Total	O	0	0
			42	42		
3	B	10	Total	O	0	0
			10	10		

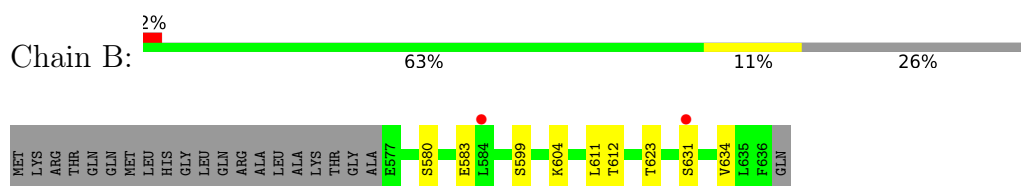
3 Residue-property plots

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: Structural maintenance of chromosomes protein 1A (SMC1A)



- 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	187.60Å 64.59Å 47.13Å 90.00° 103.97° 90.00°	Depositor
Resolution (Å)	32.30 – 2.09 32.30 – 2.09	Depositor EDS
% Data completeness (in resolution range)	98.9 (32.30-2.09) 99.0 (32.30-2.09)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.28 (at 2.10Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.199 , 0.223 0.199 , 0.223	Depositor DCC
R_{free} test set	1605 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	53.9	Xtrriage
Anisotropy	0.318	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 71.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.030 for -h-2*1,-k,l	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3188	wwPDB-VP
Average B, all atoms (Å ²)	87.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.67% 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	0.43	1/2707 (0.0%)	0.56	0/3647
2	B	0.44	0/491	0.54	0/662
All	All	0.43	1/3198 (0.0%)	0.56	0/4309

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	48	VAL	CB-CG2	5.51	1.64	1.52

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	2654	0	2648	31	1
2	B	482	0	488	4	1
3	A	42	0	0	2	0
3	B	10	0	0	0	0
All	All	3188	0	3136	33	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 5.

All (33) 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:1168:GLY:O	1:A:1172:ASN:ND2	2.20	0.72
1:A:148:PRO:HG3	1:A:151:ARG:HH21	1.55	0.71
1:A:25:GLN:NE2	3:A:1302:HOH:O	2.26	0.69
1:A:1216:LEU:HD11	2:B:611:LEU:HD13	1.79	0.64
1:A:1209:ASP:N	1:A:1209:ASP:OD1	2.30	0.63
1:A:62:ARG:NH2	3:A:1301:HOH:O	2.32	0.61
2:B:631:SER:HB2	2:B:634:VAL:HG23	1.84	0.60
1:A:1109:ASP:OD1	1:A:1110:GLY:N	2.32	0.57
1:A:62:ARG:HH21	1:A:101:GLY:H	1.54	0.55
1:A:2:GLY:HA2	1:A:84:TYR:OH	2.07	0.55
1:A:1172:ASN:O	1:A:1176:GLU:HG2	2.08	0.53
1:A:1091:ASN:ND2	1:A:1093:SER:HB3	2.23	0.52
1:A:1162:LEU:HD13	1:A:1167:ILE:HA	1.93	0.51
1:A:1206:GLU:OE2	2:B:604:LYS:NZ	2.36	0.50
1:A:1114:ASN:HA	1:A:1124:PRO:HA	1.94	0.50
1:A:1105:GLU:OE1	1:A:1107:TYR:OH	2.22	0.49
1:A:1167:ILE:HD13	1:A:1192:GLU:HB2	1.96	0.48
1:A:1178:SER:HA	1:A:1182:PHE:O	2.13	0.48
1:A:2:GLY:HA2	1:A:84:TYR:CZ	2.51	0.46
1:A:152:THR:OG1	1:A:1111:ILE:HG12	2.14	0.46
1:A:147:ASN:O	1:A:150:GLU:HG2	2.17	0.45
1:A:1153[A]:PHE:CD2	1:A:1155:LEU:HD21	2.52	0.44
1:A:97:VAL:HG12	1:A:99:VAL:HG22	1.99	0.44
1:A:151:ARG:O	1:A:155:PHE:HD1	1.99	0.44
1:A:1140:LEU:O	1:A:1144:ILE:HG13	2.19	0.43
1:A:1101:GLU:O	1:A:1103:PRO:HD3	2.18	0.43
2:B:631:SER:O	2:B:634:VAL:HB	2.19	0.43
1:A:1078:ALA:O	1:A:1081:ILE:HG22	2.18	0.42
1:A:168:TYR:OH	1:A:1107:TYR:OH	2.08	0.42
1:A:106:LYS:HE2	1:A:109:ASN:HA	2.02	0.42
1:A:1095:GLN:O	1:A:1116:VAL:N	2.52	0.41
1:A:32:GLY:N	1:A:38:LYS:HD2	2.36	0.41
1:A:27:PHE:CD2	1:A:1197:ALA:HA	2.56	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)
1:A:34:ASN:ND2	2:B:583:GLU:OE1[4_455]	2.09	0.11

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	335/456 (74%)	316 (94%)	16 (5%)	3 (1%)	14	11
2	B	58/81 (72%)	56 (97%)	2 (3%)	0	100	100
All	All	393/537 (73%)	372 (95%)	18 (5%)	3 (1%)	16	13

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1118	PRO
1	A	1226	ALA
1	A	1130	GLY

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	284/388 (73%)	273 (96%)	11 (4%)	27	29
2	B	53/69 (77%)	49 (92%)	4 (8%)	11	9
All	All	337/457 (74%)	322 (96%)	15 (4%)	25	24

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	53	THR
1	A	92	ARG
1	A	99	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	136	PHE
1	A	140	VAL
1	A	1075	GLU
1	A	1153[A]	PHE
1	A	1153[B]	PHE
1	A	1158	ILE
1	A	1199	SER
1	A	1209	ASP
2	B	580	SER
2	B	599	SER
2	B	612	THR
2	B	623	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

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(Å ²)	Q<0.9
1	A	335/456 (73%)	0.66	32 (9%) 15 16	26, 77, 165, 188	4 (1%)
2	B	60/81 (74%)	0.50	2 (3%) 49 51	50, 66, 109, 135	0
All	All	395/537 (73%)	0.64	34 (8%) 18 19	26, 75, 164, 188	4 (1%)

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	36[A]	SER	9.0
1	A	35[A]	GLY	4.5
1	A	1085	TYR	3.2
1	A	1121	ARG	3.2
1	A	1107	TYR	3.1
1	A	1124	PRO	3.1
1	A	1122	PHE	3.0
1	A	1097	PHE	2.9
1	A	1087	ALA	2.9
1	A	1063	LYS	2.8
1	A	1098	LEU	2.7
1	A	158	ILE	2.7
1	A	1069	ARG	2.6
1	A	1100	PRO	2.5
1	A	1117	ALA	2.4
1	A	1103	PRO	2.4
1	A	1131	GLY	2.4
1	A	1140	LEU	2.3
1	A	1227	ASN	2.3
1	A	1074	PHE	2.3
1	A	1162	LEU	2.3
1	A	155	PHE	2.3
1	A	1067	PHE	2.3
1	A	164	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	157	GLU	2.2
1	A	1123	ARG	2.2
1	A	1128	LEU	2.2
1	A	1072	ALA	2.2
2	B	631	SER	2.1
2	B	584	LEU	2.1
1	A	1081	ILE	2.1
1	A	154	LEU	2.1
1	A	1094	ALA	2.0
1	A	1113	TYR	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.