



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 09:06 PM GMT

PDB ID : 4UX3
Title : cohesin Smc3-HD:Scc1-N complex from yeast
Authors : Gligoris, T.G.; Nasmyth, K.; Lowe, J.
Deposited on : 2014-08-18
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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

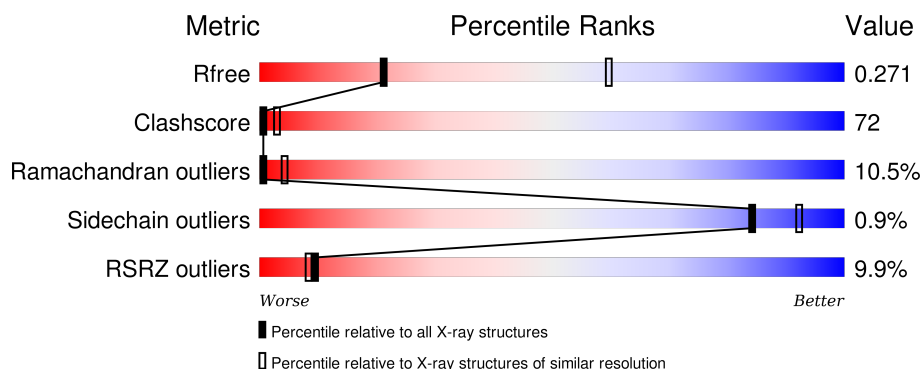
1 Overall quality at a glance

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}	91344	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	543	
2	B	121	

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	AGS	A	9999	-	-	X	-

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 4098 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 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	436	Total	C	N	O	S	0	0	0
			3521	2211	620	678	12			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	EXPRESSION TAG	UNP P47037
A	1	ALA	-	EXPRESSION TAG	UNP P47037
A	262	SER	-	LINKER	UNP P47037
A	263	SER	-	LINKER	UNP P47037
A	264	LYS	-	LINKER	UNP P47037
A	265	HIS	-	LINKER	UNP P47037
A	266	PRO	-	LINKER	UNP P47037
A	267	THR	-	LINKER	UNP P47037
A	268	SER	-	LINKER	UNP P47037
A	269	LEU	-	LINKER	UNP P47037
A	270	VAL	-	LINKER	UNP P47037
A	271	PRO	-	LINKER	UNP P47037
A	272	ARG	-	LINKER	UNP P47037
A	273	GLY	-	LINKER	UNP P47037
A	1231	TRP	-	EXPRESSION TAG	UNP P47037
A	1232	SER	-	EXPRESSION TAG	UNP P47037
A	1233	HIS	-	EXPRESSION TAG	UNP P47037
A	1234	PRO	-	EXPRESSION TAG	UNP P47037
A	1235	GLN	-	EXPRESSION TAG	UNP P47037
A	1236	PHE	-	EXPRESSION TAG	UNP P47037
A	1237	GLU	-	EXPRESSION TAG	UNP P47037
A	1238	LYS	-	EXPRESSION TAG	UNP P47037

- Molecule 2 is a protein called MITOTIC CHROMOSOME DETERMINANT-RELATED PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	71	Total	C	N	O	S	0	0	0
			545	339	90	113	3			

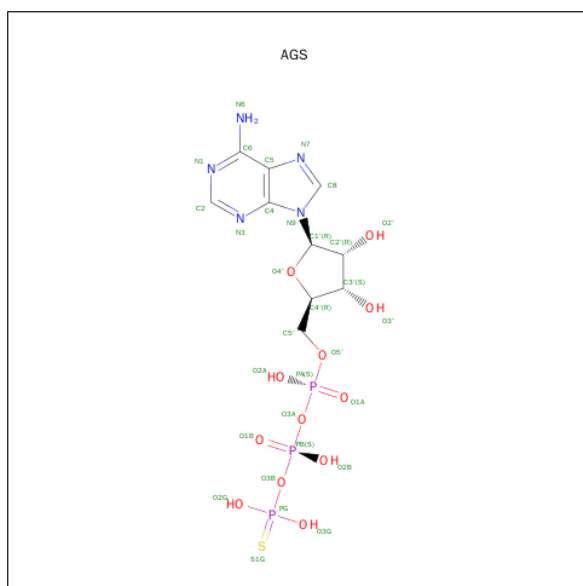
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	116	HIS	-	EXPRESSION TAG	UNP A6ZXW3
B	117	HIS	-	EXPRESSION TAG	UNP A6ZXW3
B	118	HIS	-	EXPRESSION TAG	UNP A6ZXW3
B	119	HIS	-	EXPRESSION TAG	UNP A6ZXW3
B	120	HIS	-	EXPRESSION TAG	UNP A6ZXW3
B	121	HIS	-	EXPRESSION TAG	UNP A6ZXW3

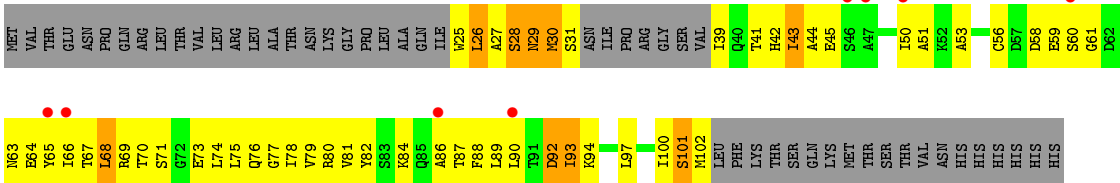
- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Mg	0	0
			1	1		

- Molecule 4 is PHOSPHOTHIOPHOSPHORIC ACID-ADENYLATE ESTER (three-letter code: AGS) (formula: C₁₀H₁₆N₅O₁₂P₃S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
4	A	1	Total	C	N	O	P	S	0	0
			31	10	5	12	3	1		



4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	73.15Å 94.80Å 284.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 3.30 48.82 – 3.00	Depositor EDS
% Data completeness (in resolution range)	99.8 (30.00-3.30) 82.2 (48.82-3.00)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.79 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.274 , 0.329 0.279 , 0.271	Depositor DCC
R_{free} test set	758 reflections (5.62%)	DCC
Wilson B-factor (Å ²)	123.5	Xtriage
Anisotropy	0.460	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 142.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	2 of 18925 reflections (0.011%)	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	4098	wwPDB-VP
Average B, all atoms (Å ²)	182.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.86% of the height of the origin peak. No significant pseudotranslation is detected.*

¹ Intensities estimated from amplitudes.

² Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MG, AGS

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	0.57	3/3559 (0.1%)	0.88	5/4765 (0.1%)
2	B	0.53	0/549	0.81	1/738 (0.1%)
All	All	0.56	3/4108 (0.1%)	0.87	6/5503 (0.1%)

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	A	0	7
2	B	0	1
All	All	0	8

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	147	PRO	CA-C	8.27	1.69	1.52
1	A	121	ARG	CG-CD	8.04	1.72	1.51
1	A	121	ARG	NE-CZ	-7.12	1.23	1.33

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	121	ARG	NE-CZ-NH2	-16.99	111.80	120.30
1	A	121	ARG	NE-CZ-NH1	8.43	124.52	120.30
1	A	150	LYS	N-CA-C	-7.87	89.76	111.00
1	A	151	ILE	CB-CA-C	5.92	123.44	111.60
1	A	154	LEU	CB-CG-CD2	-5.67	101.36	111.00

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	121	ARG	Sidechain
1	A	138	SER	Peptide
1	A	146	VAL	Peptide
1	A	148	GLN	Peptide
1	A	2	TYR	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	3521	0	3581	507	1
2	B	545	0	545	102	0
3	A	1	0	0	0	0
4	A	31	0	12	13	0
All	All	4098	0	4138	596	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 72.

The worst 5 of 596 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)
2:B:65:TYR:CZ	2:B:66:ILE:HG12	1.31	1.60
1:A:22:ASN:HB2	1:A:23:PHE:CG	1.47	1.50
1:A:22:ASN:HB2	1:A:23:PHE:CD1	1.64	1.30
1:A:116:TYR:CE2	1:A:125:LYS:HG2	1.67	1.30
2:B:65:TYR:CE2	2:B:66:ILE:HG12	1.69	1.25

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:236:ASP:OD1	1:A:236:ASP:OD1[4_5310]	1.81	0.39

5.3 Torsion angles

5.3.1 Protein backbone

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	428/543 (79%)	287 (67%)	96 (22%)	45 (10%)	1	4
2	B	67/121 (55%)	36 (54%)	24 (36%)	7 (10%)	1	4
All	All	495/664 (74%)	323 (65%)	120 (24%)	52 (10%)	1	4

5 of 52 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	26	HIS
1	A	120	ASP
1	A	122	ASN
1	A	152	VAL
1	A	153	ALA

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	392/488 (80%)	390 (100%)	2 (0%)	92	95
2	B	61/107 (57%)	59 (97%)	2 (3%)	45	78
All	All	453/595 (76%)	449 (99%)	4 (1%)	84	92

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	156	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	246	MET
2	B	30	MET
2	B	39	ILE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	148	GLN
1	A	156	ASN
1	A	1120	GLN
1	A	144	ASN
1	A	226	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	AGS	A	9999	3	24,33,33	2.10	4 (16%)	28,52,52	2.70	4 (14%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	AGS	A	9999	3	-	0/15/38/38	0/3/3/3

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	9999	AGS	PG-O2G	-2.34	1.47	1.55
4	A	9999	AGS	PG-O3G	2.06	1.62	1.55
4	A	9999	AGS	C5-C4	3.51	1.48	1.40
4	A	9999	AGS	PG-S1G	8.30	2.06	1.90

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	9999	AGS	PA-O3A-PB	-9.47	106.14	132.73
4	A	9999	AGS	N3-C2-N1	-6.93	123.59	128.89
4	A	9999	AGS	PB-O3B-PG	-5.73	113.44	132.67
4	A	9999	AGS	C4-C5-N7	-3.35	106.40	109.48

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	9999	AGS	13	0

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	436/543 (80%)	0.51	42 (9%)	10 9	73, 177, 240, 307	0
2	B	71/121 (58%)	0.65	8 (11%)	7 5	116, 174, 280, 295	0
All	All	507/664 (76%)	0.53	50 (9%)	9 8	73, 176, 249, 307	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	111	LEU	7.6
1	A	235	TYR	6.0
2	B	47	ALA	5.3
1	A	29	VAL	4.4
1	A	112	LYS	4.2

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 carbohydrates in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled ‘Q< 0.9’ lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	AGS	A	9999	31/31	0.96	0.23	-0.71	114,148,160,166	0
3	MG	A	8999	1/1	0.96	0.52	-	118,118,118,118	0

6.5 Other polymers [i](#)

There are no such residues in this entry.