wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 09:05 PM GMT

PDB ID	:	4UVJ
Title	:	Cohesin subunit Scc3 from yeast, 674-1072
Authors	:	Roig, M.B.; Nasmyth, K.; Lowe, J.
Deposited on	:	2014-08-06
Resolution	:	2.10 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

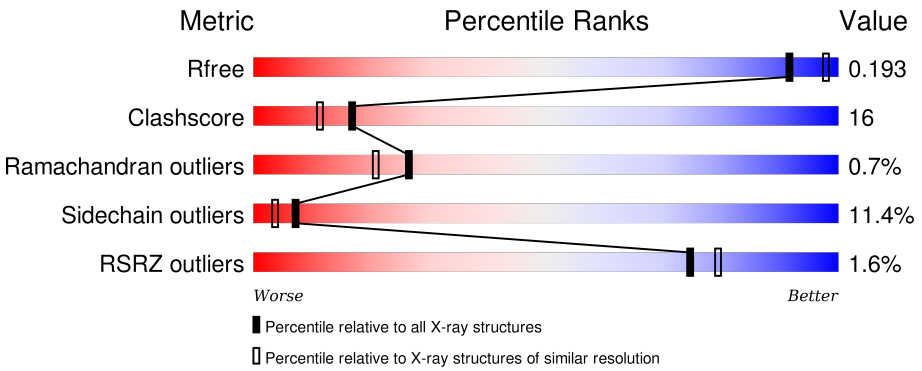
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)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

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	3939 (2.10-2.10)
Clashscore	102246	4460 (2.10-2.10)
Ramachandran outliers	100387	4413 (2.10-2.10)
Sidechain outliers	100360	4414 (2.10-2.10)
RSRZ outliers	91569	3948 (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 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 <=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	406	2% 59% 26% 5% 10%
1	B	406	% 66% 20% • • 10%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6403 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 SCC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	364	3001	1965	477	548	11	0	0	0
1	B	364	3001	1965	477	548	11	0	0	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	673	MET	-	EXPRESSION TAG	UNP P40541
A	1073	HIS	-	EXPRESSION TAG	UNP P40541
A	1074	HIS	-	EXPRESSION TAG	UNP P40541
A	1075	HIS	-	EXPRESSION TAG	UNP P40541
A	1076	HIS	-	EXPRESSION TAG	UNP P40541
A	1077	HIS	-	EXPRESSION TAG	UNP P40541
A	1078	HIS	-	EXPRESSION TAG	UNP P40541
B	673	MET	-	EXPRESSION TAG	UNP P40541
B	1073	HIS	-	EXPRESSION TAG	UNP P40541
B	1074	HIS	-	EXPRESSION TAG	UNP P40541
B	1075	HIS	-	EXPRESSION TAG	UNP P40541
B	1076	HIS	-	EXPRESSION TAG	UNP P40541
B	1077	HIS	-	EXPRESSION TAG	UNP P40541
B	1078	HIS	-	EXPRESSION TAG	UNP P40541

- Molecule 2 is water.

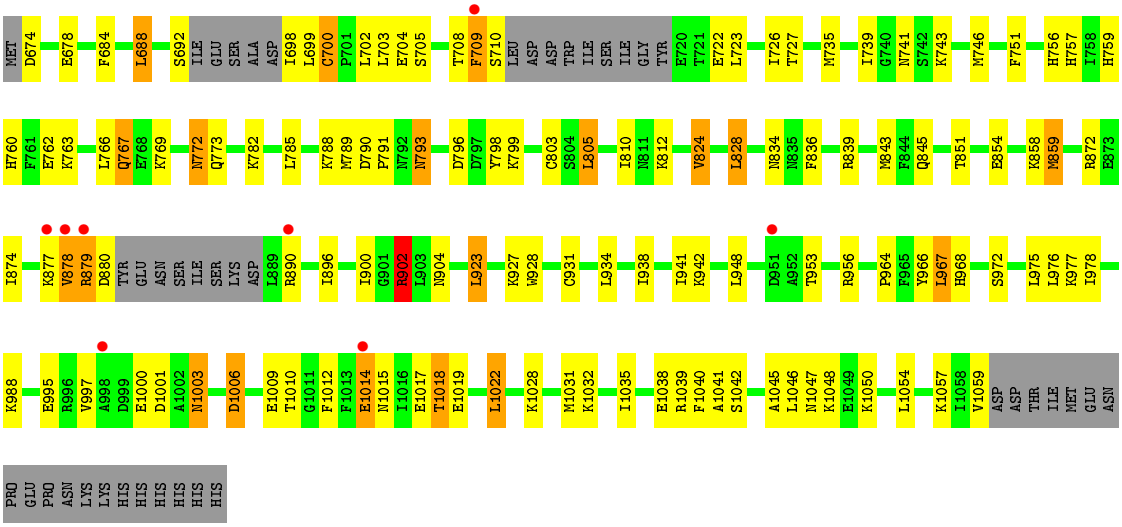
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	192	192	192	0	0
2	B	209	209	209	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

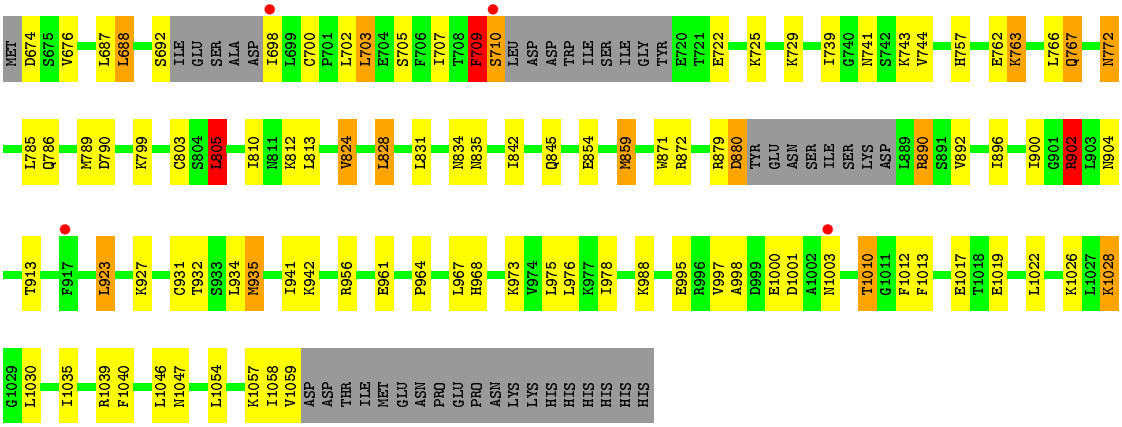
- Molecule 1: COHESIN SUBUNIT SCC3

Chain A:



- Molecule 1: COHESIN SUBUNIT SCC3

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	56.79Å 58.02Å 80.18Å 80.68° 82.00° 63.99°	Depositor
Resolution (Å)	50.00 – 2.10 41.34 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.5 (50.00-2.10) 93.8 (41.34-2.10)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.84 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.8.0073	Depositor
R, R_{free}	0.186 , 0.247 0.189 , 0.193	Depositor DCC
R_{free} test set	2598 reflections (5.31%)	DCC
Wilson B-factor (Å^2)	25.3	Xtriage
Anisotropy	0.207	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 45.2	EDS
Estimated twinning fraction	0.024 for -k,-h,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 51500 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	6403	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.79% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.76	0/3060	0.91	5/4129 (0.1%)
1	B	0.79	0/3060	0.93	4/4129 (0.1%)
All	All	0.78	0/6120	0.92	9/8258 (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	1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	935	MET	CG-SD-CE	-10.74	83.01	100.20
1	A	859	MET	CG-SD-CE	-6.31	90.11	100.20
1	B	805	LEU	CB-CG-CD2	6.21	121.55	111.00
1	B	902	ARG	NE-CZ-NH2	-6.04	117.28	120.30
1	A	923	LEU	CB-CG-CD2	5.77	120.81	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	877	LYS	Peptide

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	3001	0	3069	117	0
1	B	3001	0	3069	85	0
2	A	192	0	0	32	0
2	B	209	0	0	29	1
All	All	6403	0	6138	199	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 16.

The worst 5 of 199 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:803:CYS:HB2	2:A:2069:HOH:O	1.28	1.30
1:B:803:CYS:HB2	2:B:2075:HOH:O	1.32	1.25
1:A:810:ILE:HG23	1:A:859:MET:CE	1.72	1.19
1:B:834:ASN:ND2	1:B:902:ARG:HH22	1.42	1.16
1:A:785:LEU:HD23	1:A:789:MET:HE2	1.26	1.15

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)
2:B:2053:HOH:O	2:B:2203:HOH:O[1_655]	1.83	0.37

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	356/406 (88%)	344 (97%)	10 (3%)	2 (1%)	30	24
1	B	356/406 (88%)	344 (97%)	9 (2%)	3 (1%)	24	17
All	All	712/812 (88%)	688 (97%)	19 (3%)	5 (1%)	26	21

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	709	PHE
1	B	709	PHE
1	B	998	ALA
1	B	1001	ASP
1	A	879	ARG

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	337/377 (89%)	293 (87%)	44 (13%)	5	2
1	B	337/377 (89%)	304 (90%)	33 (10%)	10	6
All	All	674/754 (89%)	597 (89%)	77 (11%)	7	4

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1019	GLU
1	A	1057	LYS
1	B	1028	LYS
1	A	1022	LEU
1	A	1040	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 37 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1005	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	741	ASN
1	B	970	ASN
1	A	1015	ASN
1	A	1047	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)

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	364/406 (89%)	-0.07	8 (2%)	65	71	14, 32, 66, 100	0
1	B	364/406 (89%)	-0.22	4 (1%)	82	86	15, 29, 58, 86	0
All	All	728/812 (89%)	-0.14	12 (1%)	74	79	14, 30, 64, 100	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	709	PHE	5.8
1	B	698	ILE	4.8
1	B	710	SER	3.8
1	B	917	PHE	3.5
1	A	890	ARG	3.2

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)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.