Full wwPDB X-ray Structure Validation Report ⓘ

May 21, 2020 – 01:55 pm BST

PDB ID	:	5T8V
Title	:	Chaetomium thermophilum cohesin loader SCC2, C-terminal fragment
Authors	:	Tomchick, D.R.; Yu, H.; Kikuchi, S.; Ouyang, Z.; Borek, D.; Otwinowski, Z.
Deposited on	:	2016-09-08
Resolution	:	2.80 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 following versions of software and data (see references ⓘ) were used in the production of this report:

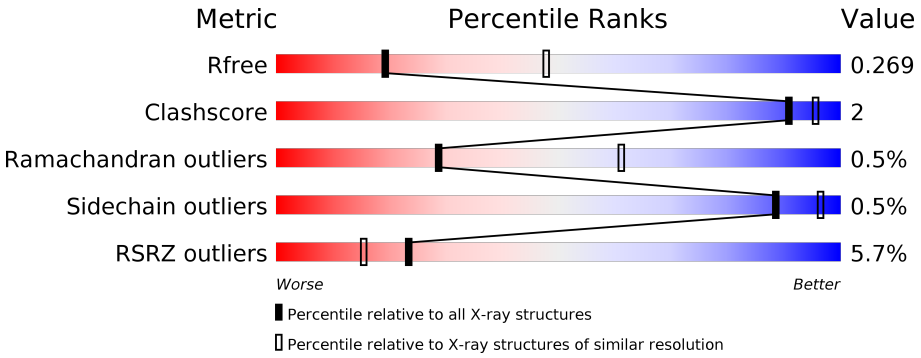
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 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	1456	5% 85% 5% 10%

2 Entry composition

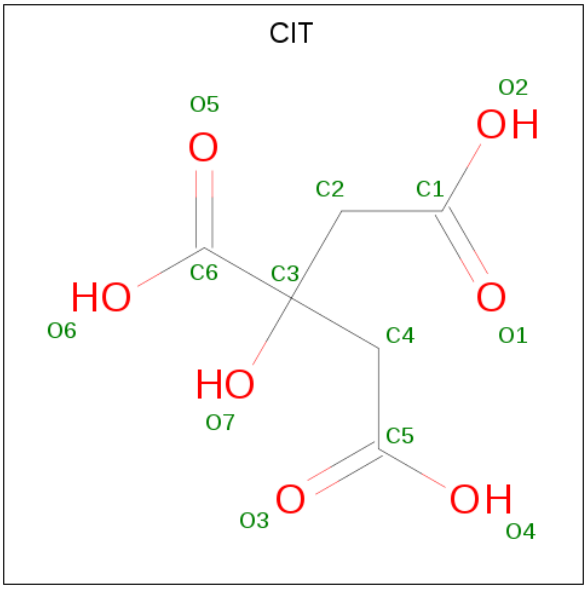
There are 2 unique types of molecules in this entry. The entry contains 20912 atoms, of which 10586 are hydrogens and 0 are deuteriums.

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 Putative uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	1315	20894	6527	10581	1787	1941	58	0	0	0

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



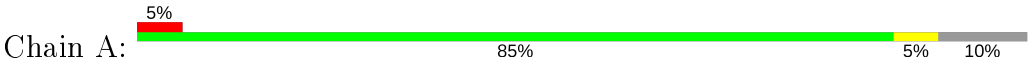
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	A	1	18	6	5	7	0	0

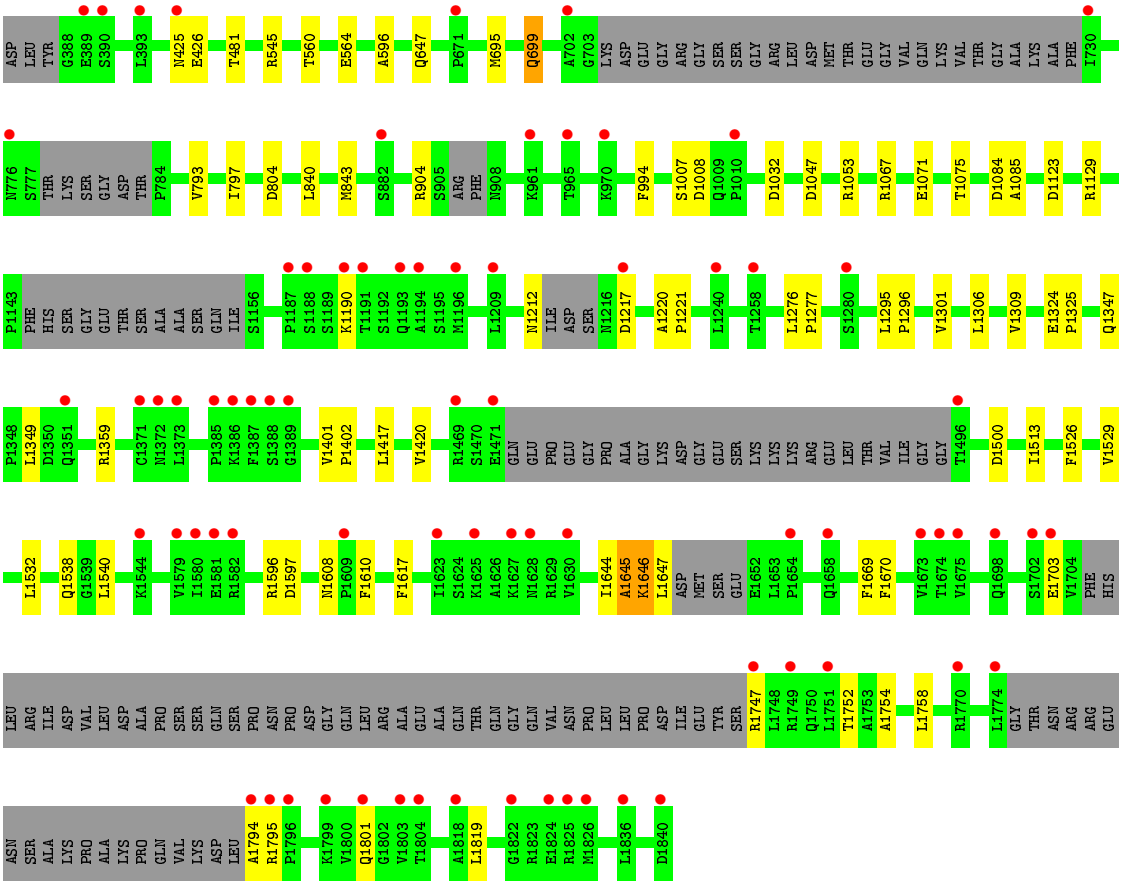
3 Residue-property plots (i)

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

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	66.73Å 88.78Å 160.30Å 90.00° 93.17° 90.00°	Depositor
Resolution (Å)	36.24 – 2.80 36.24 – 2.80	Depositor EDS
% Data completeness (in resolution range)	95.5 (36.24-2.80) 95.5 (36.24-2.80)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.35 (at 2.81Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.228 , 0.267 0.229 , 0.269	Depositor DCC
R_{free} test set	2004 reflections (4.52%)	wwPDB-VP
Wilson B-factor (Å^2)	44.6	Xtriage
Anisotropy	0.051	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 37.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	20912	wwPDB-VP
Average B, all atoms (Å^2)	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.70% 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

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

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	0/10464	0.39	0/14145

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	10313	10581	10581	35	0
2	A	13	5	5	0	0
All	All	10326	10586	10586	35	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 2.

All (35) 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:804:ASP:OD2	1:A:904:ARG:NH2	2.28	0.66
1:A:1212:ASN:ND2	1:A:1212:ASN:O	2.35	0.60

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:1301:VAL:O	1:A:1359:ARG:NH2	2.34	0.60
1:A:1123:ASP:O	1:A:1129:ARG:NE	2.38	0.55
1:A:1752:THR:HB	1:A:1819:LEU:HD21	1.89	0.55
1:A:560:THR:O	1:A:564:GLU:N	2.41	0.54
1:A:1084:ASP:OD1	1:A:1085:ALA:N	2.41	0.53
1:A:1644:ILE:O	1:A:1645:ALA:CB	2.61	0.49
1:A:1220:ALA:HB1	1:A:1221:PRO:HD2	1.96	0.48
1:A:695:MET:O	1:A:699:GLN:HG3	2.15	0.47
1:A:1306:LEU:HA	1:A:1309:VAL:HG12	1.96	0.47
1:A:425:ASN:OD1	1:A:426:GLU:N	2.46	0.46
1:A:1401:VAL:N	1:A:1402:PRO:CD	2.79	0.46
1:A:1071:GLU:O	1:A:1075:THR:HG23	2.16	0.46
1:A:1047:ASP:O	1:A:1053:ARG:NH1	2.47	0.45
1:A:1596:ARG:NH1	1:A:1597:ASP:OD1	2.49	0.45
1:A:1703:GLU:OE1	1:A:1747:ARG:NE	2.48	0.45
1:A:1526:PHE:HA	1:A:1529:VAL:HG12	1.99	0.44
1:A:1295:LEU:N	1:A:1296:PRO:HD2	2.32	0.44
1:A:1754:ALA:O	1:A:1758:LEU:HD12	2.18	0.43
1:A:1646:LYS:O	1:A:1647:LEU:C	2.56	0.43
1:A:1032:ASP:OD2	1:A:1067:ARG:NH1	2.51	0.43
1:A:596:ALA:HB2	1:A:647:GLN:OE1	2.18	0.43
1:A:840:LEU:HD23	1:A:843:MET:HE3	2.01	0.43
1:A:793:VAL:O	1:A:797:ILE:HG12	2.19	0.43
1:A:1324:GLU:N	1:A:1325:PRO:HD2	2.33	0.42
1:A:1417:LEU:HA	1:A:1420:VAL:HG12	2.01	0.42
1:A:1500:ASP:OD1	1:A:1500:ASP:N	2.53	0.42
1:A:1276:LEU:N	1:A:1277:PRO:HD2	2.36	0.41
1:A:1538:GLN:HB2	1:A:1540:LEU:HD12	2.02	0.41
1:A:1347:GLN:O	1:A:1349:LEU:N	2.46	0.41
1:A:1513:ILE:HG22	1:A:1532:LEU:HD13	2.03	0.40
1:A:1794:ALA:O	1:A:1795:ARG:HB2	2.21	0.40
1:A:1007:SER:OG	1:A:1008:ASP:N	2.54	0.40
1:A:481:THR:O	1:A:545:ARG:NH2	2.54	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	1295/1456 (89%)	1207 (93%)	82 (6%)	6 (0%)	29 61

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1645	ALA
1	A	1190	LYS
1	A	1646	LYS
1	A	994	PHE
1	A	1608	ASN
1	A	1801	GLN

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	1146/1262 (91%)	1140 (100%)	6 (0%)	88 96

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	699	GLN
1	A	1217	ASP
1	A	1610	PHE
1	A	1617	PHE
1	A	1669	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	1670	PHE

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

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

1 ligand is modelled in this entry.

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	$\#\|Z\| > 2$
2	CIT	A	2001	-	3,12,12	1.36	0	3,17,17	2.22	1 (33%)

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
2	CIT	A	2001	-	-	1/6/16/16	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	2001	CIT	C3-C2-C1	-3.11	110.01	114.98

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	2001	CIT	O7-C3-C4-C5

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	1315/1456 (90%)	0.45	75 (5%)	23 15	9, 50, 102, 162	0

All (75) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	882	SER	7.2
1	A	1749	ARG	6.8
1	A	1190	LYS	5.5
1	A	1803	VAL	5.2
1	A	1188	SER	5.1
1	A	1804	THR	5.1
1	A	1822	GLY	5.0
1	A	1193	GLN	4.9
1	A	1774	LEU	4.9
1	A	1581	GLU	4.3
1	A	1191	THR	4.3
1	A	730	ILE	4.2
1	A	1795	ARG	4.2
1	A	1351	GLN	4.2
1	A	1824	GLU	3.8
1	A	1799	LYS	3.8
1	A	1751	LEU	3.7
1	A	1372	ASN	3.7
1	A	1194	ALA	3.6
1	A	1471	GLU	3.5
1	A	1496	THR	3.5
1	A	1673	VAL	3.4
1	A	1469	ARG	3.4
1	A	1825	ARG	3.3
1	A	1387	PHE	3.2
1	A	1628	ASN	2.9
1	A	671	PRO	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	970	LYS	2.8
1	A	1389	GLY	2.8
1	A	702	ALA	2.7
1	A	1010	PRO	2.7
1	A	393	LEU	2.7
1	A	1609	PRO	2.7
1	A	1796	PRO	2.7
1	A	1187	PRO	2.7
1	A	1654	PRO	2.6
1	A	1658	GLN	2.6
1	A	1801	GLN	2.6
1	A	1840	ASP	2.5
1	A	965	THR	2.5
1	A	1582	ARG	2.5
1	A	1627	LYS	2.5
1	A	1794	ALA	2.5
1	A	1625	LYS	2.4
1	A	1240	LEU	2.4
1	A	389	GLU	2.4
1	A	1386	LYS	2.4
1	A	1623	ILE	2.4
1	A	390	SER	2.4
1	A	1217	ASP	2.4
1	A	1580	ILE	2.4
1	A	1770	ARG	2.4
1	A	961	LYS	2.3
1	A	1702	SER	2.3
1	A	1258	THR	2.3
1	A	1371	CYS	2.3
1	A	1674	THR	2.3
1	A	1818	ALA	2.3
1	A	1836	LEU	2.2
1	A	1579	VAL	2.2
1	A	1826	MET	2.2
1	A	1209	LEU	2.2
1	A	1385	PRO	2.2
1	A	1544	LYS	2.2
1	A	1196	MET	2.2
1	A	1373	LEU	2.2
1	A	1280	SER	2.1
1	A	1698	GLN	2.1
1	A	1630	VAL	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1675	VAL	2.1
1	A	1388	SER	2.1
1	A	1747	ARG	2.1
1	A	425	ASN	2.0
1	A	776	ASN	2.0
1	A	1703	GLU	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 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. The B-factors column lists the minimum, median, 95^{th} 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	B-factors(Å^2)	Q<0.9
2	CIT	A	2001	13/13	0.78	0.31	90,93,114,114	0

6.5 Other polymers (i)

There are no such residues in this entry.