



Full wwPDB X-ray Structure Validation Report ⓘ

Jan 30, 2024 – 01:06 PM EST

PDB ID : 1G1K
Title : COHESIN MODULE FROM THE CELLULOSOME OF CLOSTRIDIUM CELLULOLYTICUM
Authors : Spinelli, S.; Fierobe, H.-P.; Belaich, A.; Belaich, J.-P.; Henrissat, B.; Cambilau, C.
Deposited on : 2000-10-12
Resolution : 2.00 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
Xtrriage (Phenix) : **NOT EXECUTED**
EDS : **NOT EXECUTED**
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

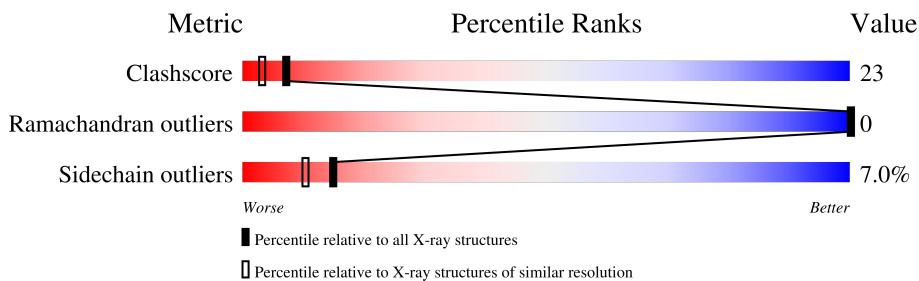
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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))
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)

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 $\leq 5\%$.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	143	 72% 24% .
1	B	143	 68% 25% 6% .

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2280 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 SCAFFOLDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	143	1028	652	164	209	3	8	0	0
1	B	143	1026	651	163	209	3	9	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	-	cloning artifact	UNP Q45996
B	2	ALA	-	cloning artifact	UNP Q45996

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	106	106	106	0	0
2	B	120	120	120	0	0

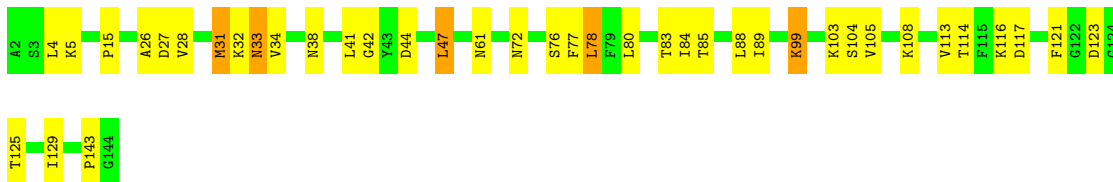
3 Residue-property plots [i](#)

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. 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. 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.

Note EDS was not executed.

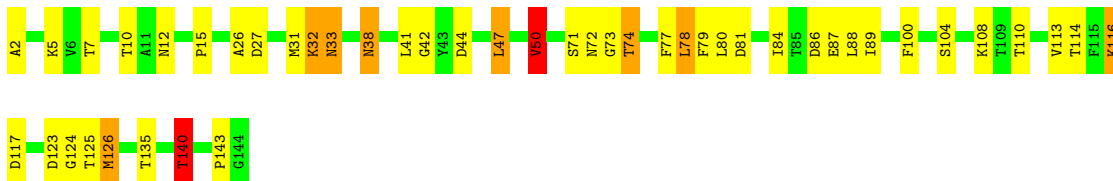
- Molecule 1: SCAFFOLDING PROTEIN

Chain A:  72% 24%



- Molecule 1: SCAFFOLDING PROTEIN

Chain B:  68% 25% 6%



4 Data and refinement statistics

Xtrriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	56.50Å 57.25Å 105.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.00	Depositor
% Data completeness (in resolution range)	(Not available) (20.00-2.00)	Depositor
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.230 , 0.270	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	2280	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

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	0.82	3/1043 (0.3%)	0.93	4/1419 (0.3%)
1	B	0.85	4/1041 (0.4%)	1.12	9/1415 (0.6%)
All	All	0.84	7/2084 (0.3%)	1.03	13/2834 (0.5%)

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	33	ASN	C-N	-8.92	1.13	1.34
1	A	32	LYS	C-N	-8.06	1.15	1.34
1	A	33	ASN	C-O	6.89	1.36	1.23
1	B	50	VAL	CB-CG2	5.97	1.65	1.52
1	A	33	ASN	N-CA	-5.42	1.35	1.46
1	B	140	THR	CA-CB	5.40	1.67	1.53
1	B	124	GLY	C-O	5.21	1.31	1.23

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	50	VAL	CA-CB-CG1	8.98	124.37	110.90
1	B	50	VAL	CA-CB-CG2	8.79	124.09	110.90
1	B	32	LYS	C-N-CA	8.45	142.83	121.70
1	B	126	MET	CG-SD-CE	-7.10	88.84	100.20
1	B	126	MET	CB-CA-C	-6.98	96.44	110.40
1	B	126	MET	CB-CG-SD	6.58	132.14	112.40
1	A	72	ASN	C-N-CA	-6.46	108.73	122.30
1	B	71	SER	N-CA-CB	6.32	119.98	110.50
1	A	113	VAL	N-CA-C	-5.40	96.41	111.00
1	B	113	VAL	N-CA-C	-5.40	96.42	111.00
1	B	33	ASN	O-C-N	-5.26	114.28	122.70
1	A	33	ASN	CA-C-O	5.21	131.05	120.10
1	A	31	MET	CG-SD-CE	5.03	108.25	100.20

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	1028	0	1040	53	0
1	B	1026	0	1035	49	0
2	A	106	0	0	9	0
2	B	120	0	0	5	0
All	All	2280	0	2075	92	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 23.

All (92) 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:80:LEU:HD12	1:B:78:LEU:CD1	1.40	1.47
1:A:80:LEU:HD12	1:B:78:LEU:HD12	1.30	1.13
1:A:80:LEU:CD1	1:B:78:LEU:CD1	2.32	1.08
1:A:4:LEU:HB2	1:A:129:ILE:HD11	1.41	1.01
1:A:80:LEU:HD12	1:B:78:LEU:HD11	1.02	1.01
1:A:80:LEU:CD1	1:B:78:LEU:HD11	1.90	1.01
1:A:33:ASN:O	1:A:123:ASP:HA	1.61	1.01
1:A:99:LYS:HE3	2:A:217:HOH:O	1.60	1.00
1:A:80:LEU:CD1	1:B:78:LEU:HD12	1.96	0.94
1:A:38:ASN:HB3	2:A:220:HOH:O	1.68	0.93
1:B:87:GLU:OE2	2:B:249:HOH:O	1.89	0.89
1:A:117:ASP:HB3	1:B:117:ASP:HB3	1.56	0.86
1:A:129:ILE:HB	2:A:190:HOH:O	1.77	0.84
1:B:33:ASN:O	1:B:123:ASP:HA	1.77	0.83
1:B:78:LEU:C	1:B:78:LEU:HD13	1.99	0.82
1:A:99:LYS:HE2	2:A:249:HOH:O	1.81	0.80
1:B:50:VAL:HG12	1:B:100:PHE:CE1	2.17	0.79
1:A:38:ASN:OD1	1:B:38:ASN:ND2	2.18	0.76
1:A:99:LYS:CE	2:A:217:HOH:O	2.28	0.75
1:B:110:THR:OG1	1:B:140:THR:HB	1.89	0.72

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:42:GLY:CA	1:B:116:LYS:HD2	2.20	0.72
1:A:4:LEU:CB	1:A:129:ILE:HD11	2.20	0.71
1:A:77:PHE:C	1:A:78:LEU:HD12	2.11	0.71
1:B:42:GLY:HA3	1:B:116:LYS:HD2	1.73	0.70
1:A:121:PHE:O	1:A:129:ILE:HG22	1.91	0.69
1:A:33:ASN:O	1:A:123:ASP:CA	2.40	0.68
1:A:85:THR:HB	2:A:227:HOH:O	1.94	0.66
1:A:121:PHE:HB2	1:A:129:ILE:CG2	2.26	0.65
1:A:108:LYS:HA	1:A:143:PRO:HD3	1.79	0.64
1:A:4:LEU:HB2	1:A:129:ILE:CD1	2.21	0.63
1:A:5:LYS:HB2	1:A:26:ALA:HB3	1.80	0.63
1:B:74:THR:HB	1:B:116:LYS:NZ	2.13	0.63
1:B:140:THR:HG22	2:B:230:HOH:O	1.98	0.63
1:A:121:PHE:CB	1:A:129:ILE:HG21	2.30	0.61
1:B:5:LYS:HE2	1:B:27:ASP:OD2	2.01	0.61
1:B:78:LEU:HD13	1:B:79:PHE:N	2.16	0.60
1:A:78:LEU:HD12	1:A:78:LEU:N	2.17	0.60
1:A:34:VAL:HB	1:A:89:ILE:HB	1.85	0.59
1:A:42:GLY:HA3	1:A:116:LYS:HD2	1.82	0.59
1:A:105:VAL:HG23	1:A:143:PRO:HG3	1.86	0.58
1:B:78:LEU:CD1	1:B:78:LEU:C	2.71	0.57
1:B:5:LYS:HB2	1:B:26:ALA:HB3	1.86	0.57
1:A:84:ILE:HD12	1:A:84:ILE:N	2.20	0.56
1:B:50:VAL:HG12	1:B:100:PHE:HE1	1.67	0.55
1:B:41:LEU:HA	1:B:114:THR:O	2.08	0.54
1:A:129:ILE:HG23	1:A:129:ILE:O	2.08	0.53
1:A:83:THR:C	1:A:84:ILE:HD12	2.30	0.52
1:A:121:PHE:HB3	1:A:129:ILE:HG21	1.91	0.52
1:A:121:PHE:CB	1:A:129:ILE:CG2	2.87	0.52
1:A:129:ILE:HD12	2:A:190:HOH:O	2.09	0.52
1:A:42:GLY:HA3	1:A:116:LYS:CD	2.39	0.52
1:A:76:SER:OG	1:A:78:LEU:HD11	2.10	0.52
1:A:121:PHE:HB2	1:A:129:ILE:HG21	1.88	0.52
1:B:10:THR:HG23	2:B:230:HOH:O	2.11	0.50
1:B:125:THR:O	1:B:126:MET:HB2	2.11	0.50
1:B:15:PRO:HB3	1:B:104:SER:HA	1.93	0.50
1:B:140:THR:CG2	2:B:230:HOH:O	2.59	0.50
1:A:38:ASN:HA	1:A:77:PHE:O	2.11	0.49
1:B:74:THR:CG2	1:B:116:LYS:HZ3	2.26	0.49
1:B:38:ASN:HA	1:B:77:PHE:O	2.13	0.48
1:B:74:THR:HB	1:B:116:LYS:HZ3	1.77	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:7:THR:HG23	1:B:135:THR:HB	1.95	0.48
1:B:31:MET:O	1:B:32:LYS:HB2	2.12	0.48
1:B:108:LYS:HA	1:B:143:PRO:HD2	1.94	0.47
1:A:84:ILE:N	1:A:84:ILE:CD1	2.78	0.47
1:B:80:LEU:HD23	1:B:81:ASP:N	2.29	0.47
1:B:74:THR:HB	1:B:116:LYS:HZ2	1.79	0.46
1:A:28:VAL:HG13	1:A:34:VAL:HG23	1.98	0.46
1:A:4:LEU:CB	1:A:129:ILE:CD1	2.90	0.46
1:A:117:ASP:HB3	1:B:117:ASP:CB	2.39	0.46
1:B:50:VAL:HG12	1:B:100:PHE:CD1	2.51	0.46
1:A:80:LEU:HD11	1:B:78:LEU:HD12	1.93	0.45
1:B:44:ASP:OD1	1:B:47:LEU:CD2	2.64	0.45
1:B:33:ASN:CB	1:B:89:ILE:O	2.65	0.44
1:A:44:ASP:OD1	1:A:47:LEU:HD22	2.18	0.44
1:B:74:THR:CG2	1:B:116:LYS:NZ	2.80	0.44
1:B:44:ASP:CG	1:B:47:LEU:CD2	2.85	0.44
1:B:42:GLY:HA2	1:B:73:GLY:O	2.18	0.44
1:A:38:ASN:CB	2:A:220:HOH:O	2.46	0.44
1:A:15:PRO:HB3	1:A:103:LYS:O	2.17	0.43
1:A:41:LEU:HA	1:A:114:THR:O	2.17	0.43
1:A:27:ASP:O	1:A:31:MET:HG2	2.18	0.43
1:B:2:ALA:N	2:B:227:HOH:O	2.51	0.43
1:B:125:THR:O	1:B:125:THR:OG1	2.35	0.42
1:B:84:ILE:HG22	1:B:86:ASP:OD1	2.19	0.42
1:A:78:LEU:N	1:A:78:LEU:CD1	2.82	0.42
1:B:33:ASN:O	1:B:123:ASP:CA	2.60	0.41
1:B:80:LEU:HD23	1:B:80:LEU:C	2.40	0.41
1:A:42:GLY:CA	1:A:116:LYS:HD3	2.50	0.41
1:A:15:PRO:HB3	1:A:104:SER:HA	2.02	0.41
1:A:84:ILE:HD13	2:A:180:HOH:O	2.22	0.40
1:B:80:LEU:C	1:B:80:LEU:CD2	2.90	0.40

There are no symmetry-related clashes.

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	141/143 (99%)	136 (96%)	5 (4%)	0	100	100
1	B	141/143 (99%)	137 (97%)	4 (3%)	0	100	100
All	All	282/286 (99%)	273 (97%)	9 (3%)	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	115/115 (100%)	109 (95%)	6 (5%)	23	19
1	B	114/115 (99%)	104 (91%)	10 (9%)	10	6
All	All	229/230 (100%)	213 (93%)	16 (7%)	15	10

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	47	LEU
1	A	61	ASN
1	A	78	LEU
1	A	88	LEU
1	A	99	LYS
1	A	125	THR
1	B	12	ASN
1	B	38	ASN
1	B	47	LEU
1	B	50	VAL
1	B	72	ASN
1	B	74	THR
1	B	78	LEU
1	B	88	LEU
1	B	116	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	140	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	61	ASN
1	A	65	ASN
1	B	65	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 monosaccharides 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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1
1	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	32:LYS	C	33:ASN	N	1.15
1	B	33:ASN	C	34:VAL	N	1.13

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

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

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.