



wwPDB EM Validation Summary Report ⓘ

Feb 20, 2025 – 08:12 AM EST

PDB ID : 9B8V
EMDB ID : EMD-44359
Title : AlphaFold2 informed cryo-EM model of the E. coli cellulose synthase Bc-sAG3B6 complex
Authors : Verma, P.; Zimmer, J.
Deposited on : 2024-04-01
Resolution : 4.00 Å(reported)

This is a wwPDB EM Validation Summary 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/EMValidationReportHelp>
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:

EMDB validation analysis : 0.0.1.dev117
MolProbity : 4.02b-467
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
MapQ : **FAILED**
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.41.4

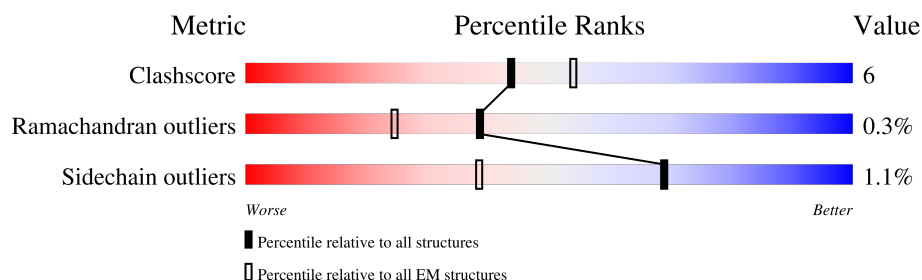
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

ELECTRON MICROSCOPY

The reported resolution of this entry is 4.00 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)	EM structures (#Entries)
Clashscore	210492	15764
Ramachandran outliers	207382	16835
Sidechain outliers	206894	16415

The table below summarises the geometric issues observed across the polymeric chains and their fit to the map. The red, orange, yellow and green segments of the 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 $\leq 5\%$.

Mol	Chain	Length	Quality of chain
1	A	763	77% 7% 15%
1	F	763	80% 5% 15%
1	G	763	78% 8% 13%
1	H	763	78% 7% 15%
1	I	763	79% 6% 15%
1	J	763	79% 8% 13%
2	B	567	14% 10% 76%
2	C	567	16% 9% 75%
2	D	567	15% 10% 75%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	E	887	 79%15%• 5%

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 70366 atoms, of which 29654 are hydrogens and 0 are deuteriums.

In the tables below, 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 Cellulose synthase operon protein B.

Mol	Chain	Residues	Atoms						AltConf	Trace
1	G	664	Total	C	H	N	O	S	0	0
			9920	3316	4730	887	964	23		
1	A	645	Total	C	H	N	O	S	0	0
			9952	3171	4953	851	953	24		
1	F	647	Total	C	H	N	O	S	0	0
			9971	3177	4960	853	957	24		
1	H	645	Total	C	H	N	O	S	0	0
			9938	3168	4942	850	954	24		
1	I	646	Total	C	H	N	O	S	0	0
			9957	3174	4953	851	955	24		
1	J	664	Total	C	H	N	O	S	0	0
			10263	3265	5116	879	979	24		

There are 54 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	17	ALA	-	expression tag	UNP P37652
G	18	TRP	-	expression tag	UNP P37652
G	19	SER	-	expression tag	UNP P37652
G	20	HIS	-	expression tag	UNP P37652
G	21	PRO	-	expression tag	UNP P37652
G	22	GLN	-	expression tag	UNP P37652
G	23	PHE	-	expression tag	UNP P37652
G	24	GLU	-	expression tag	UNP P37652
G	25	LYS	-	expression tag	UNP P37652
A	17	ALA	-	expression tag	UNP P37652
A	18	TRP	-	expression tag	UNP P37652
A	19	SER	-	expression tag	UNP P37652
A	20	HIS	-	expression tag	UNP P37652
A	21	PRO	-	expression tag	UNP P37652
A	22	GLN	-	expression tag	UNP P37652
A	23	PHE	-	expression tag	UNP P37652
A	24	GLU	-	expression tag	UNP P37652
A	25	LYS	-	expression tag	UNP P37652

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	17	ALA	-	expression tag	UNP P37652
F	18	TRP	-	expression tag	UNP P37652
F	19	SER	-	expression tag	UNP P37652
F	20	HIS	-	expression tag	UNP P37652
F	21	PRO	-	expression tag	UNP P37652
F	22	GLN	-	expression tag	UNP P37652
F	23	PHE	-	expression tag	UNP P37652
F	24	GLU	-	expression tag	UNP P37652
F	25	LYS	-	expression tag	UNP P37652
H	17	ALA	-	expression tag	UNP P37652
H	18	TRP	-	expression tag	UNP P37652
H	19	SER	-	expression tag	UNP P37652
H	20	HIS	-	expression tag	UNP P37652
H	21	PRO	-	expression tag	UNP P37652
H	22	GLN	-	expression tag	UNP P37652
H	23	PHE	-	expression tag	UNP P37652
H	24	GLU	-	expression tag	UNP P37652
H	25	LYS	-	expression tag	UNP P37652
I	17	ALA	-	expression tag	UNP P37652
I	18	TRP	-	expression tag	UNP P37652
I	19	SER	-	expression tag	UNP P37652
I	20	HIS	-	expression tag	UNP P37652
I	21	PRO	-	expression tag	UNP P37652
I	22	GLN	-	expression tag	UNP P37652
I	23	PHE	-	expression tag	UNP P37652
I	24	GLU	-	expression tag	UNP P37652
I	25	LYS	-	expression tag	UNP P37652
J	17	ALA	-	expression tag	UNP P37652
J	18	TRP	-	expression tag	UNP P37652
J	19	SER	-	expression tag	UNP P37652
J	20	HIS	-	expression tag	UNP P37652
J	21	PRO	-	expression tag	UNP P37652
J	22	GLN	-	expression tag	UNP P37652
J	23	PHE	-	expression tag	UNP P37652
J	24	GLU	-	expression tag	UNP P37652
J	25	LYS	-	expression tag	UNP P37652

- Molecule 2 is a protein called Cellulose biosynthesis protein BcsG.

Mol	Chain	Residues	Atoms					AltConf	Trace
2	B	137	Total	C	N	O	S	0	0
			1136	781	182	170	3		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					AltConf	Trace
2	C	142	Total	C	N	O	S	0	0
			1193	823	190	177	3		
2	D	144	Total	C	N	O	S	0	0
			1205	829	192	181	3		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	560	ASP	-	expression tag	UNP P37659
B	561	TYR	-	expression tag	UNP P37659
B	562	LYS	-	expression tag	UNP P37659
B	563	ASP	-	expression tag	UNP P37659
B	564	ASP	-	expression tag	UNP P37659
B	565	ASP	-	expression tag	UNP P37659
B	566	ASP	-	expression tag	UNP P37659
B	567	LYS	-	expression tag	UNP P37659
C	560	ASP	-	expression tag	UNP P37659
C	561	TYR	-	expression tag	UNP P37659
C	562	LYS	-	expression tag	UNP P37659
C	563	ASP	-	expression tag	UNP P37659
C	564	ASP	-	expression tag	UNP P37659
C	565	ASP	-	expression tag	UNP P37659
C	566	ASP	-	expression tag	UNP P37659
C	567	LYS	-	expression tag	UNP P37659
D	560	ASP	-	expression tag	UNP P37659
D	561	TYR	-	expression tag	UNP P37659
D	562	LYS	-	expression tag	UNP P37659
D	563	ASP	-	expression tag	UNP P37659
D	564	ASP	-	expression tag	UNP P37659
D	565	ASP	-	expression tag	UNP P37659
D	566	ASP	-	expression tag	UNP P37659
D	567	LYS	-	expression tag	UNP P37659

- Molecule 3 is a protein called Cellulose synthase catalytic subunit [UDP-forming].

Mol	Chain	Residues	Atoms					AltConf	Trace
3	E	845	Total	C	N	O	S	0	0
			6831	4460	1189	1148	34		

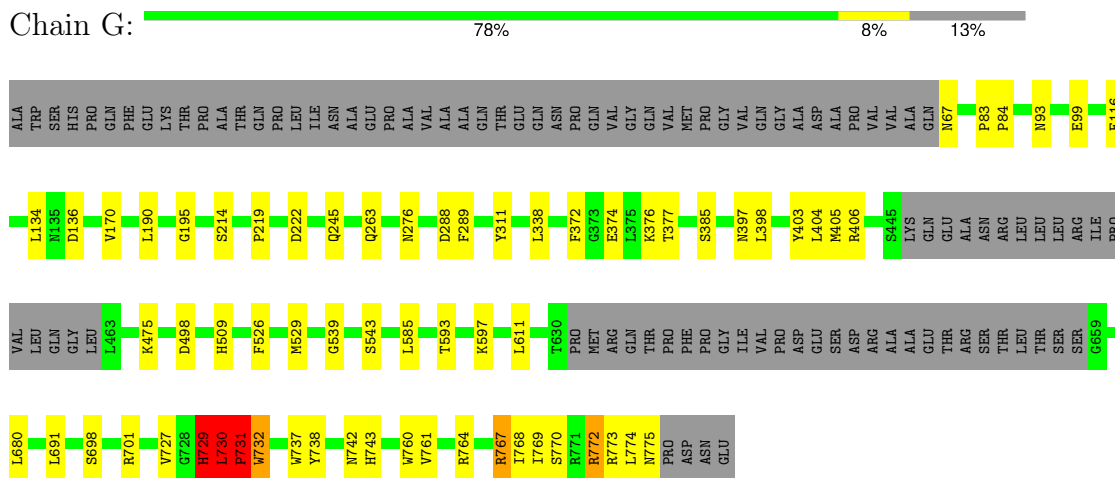
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	MET	-	initiating methionine	UNP P37653
E	1	GLY	-	cloning artifact	UNP P37653
E	873	HIS	-	expression tag	UNP P37653
E	874	HIS	-	expression tag	UNP P37653
E	875	HIS	-	expression tag	UNP P37653
E	876	HIS	-	expression tag	UNP P37653
E	877	HIS	-	expression tag	UNP P37653
E	878	HIS	-	expression tag	UNP P37653
E	879	LEU	-	expression tag	UNP P37653
E	880	GLU	-	expression tag	UNP P37653
E	881	HIS	-	expression tag	UNP P37653
E	882	HIS	-	expression tag	UNP P37653
E	883	HIS	-	expression tag	UNP P37653
E	884	HIS	-	expression tag	UNP P37653
E	885	HIS	-	expression tag	UNP P37653
E	886	HIS	-	expression tag	UNP P37653

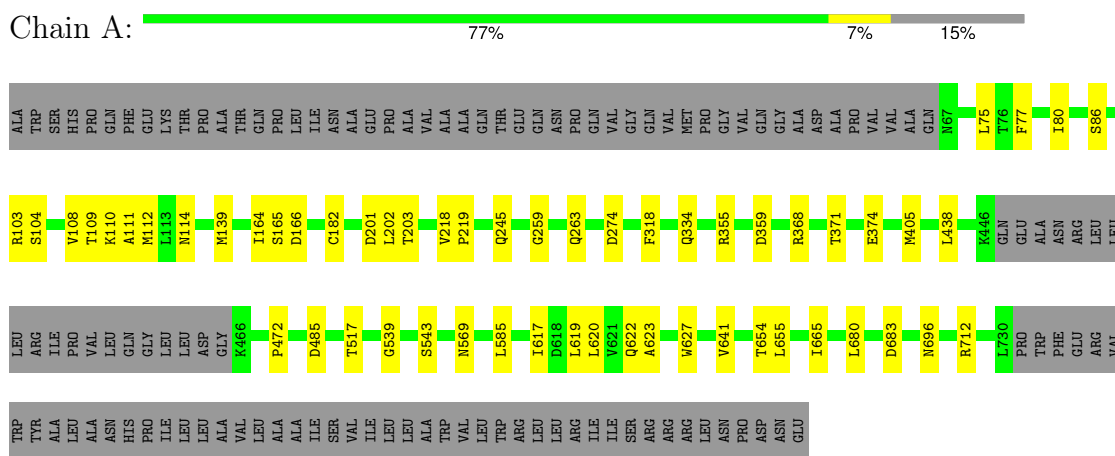
3 Residue-property plots

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. 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.

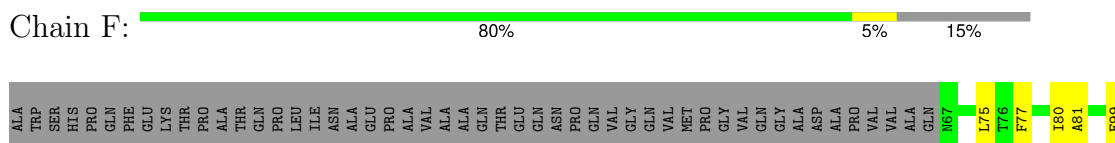
• Molecule 1: Cellulose synthase operon protein B

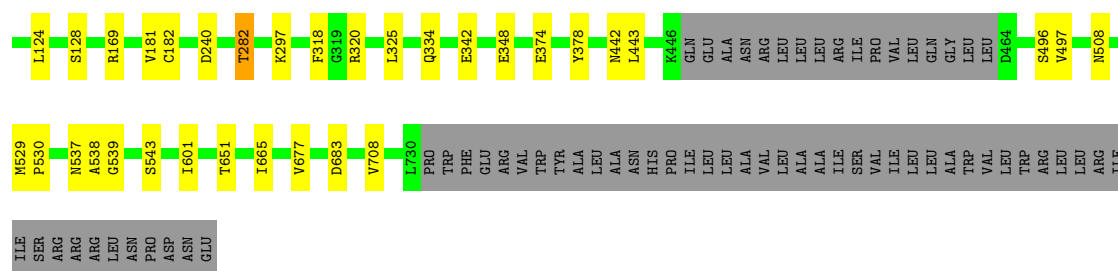


• Molecule 1: Cellulose synthase operon protein B



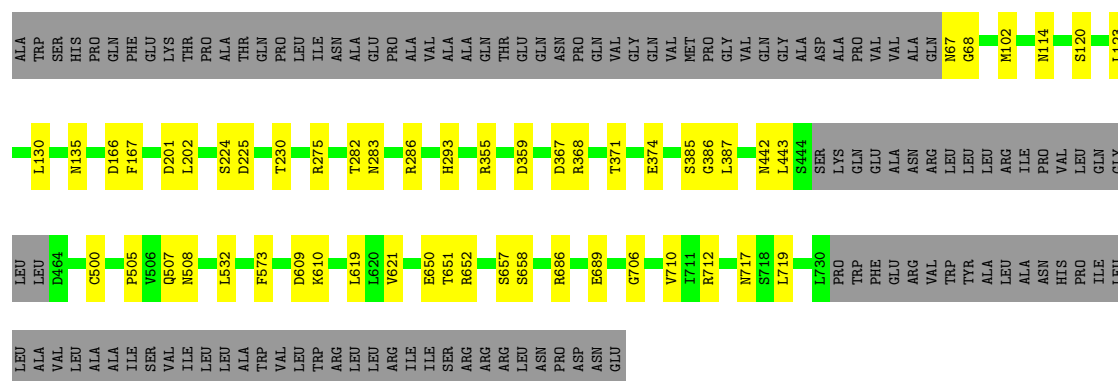
• Molecule 1: Cellulose synthase operon protein B





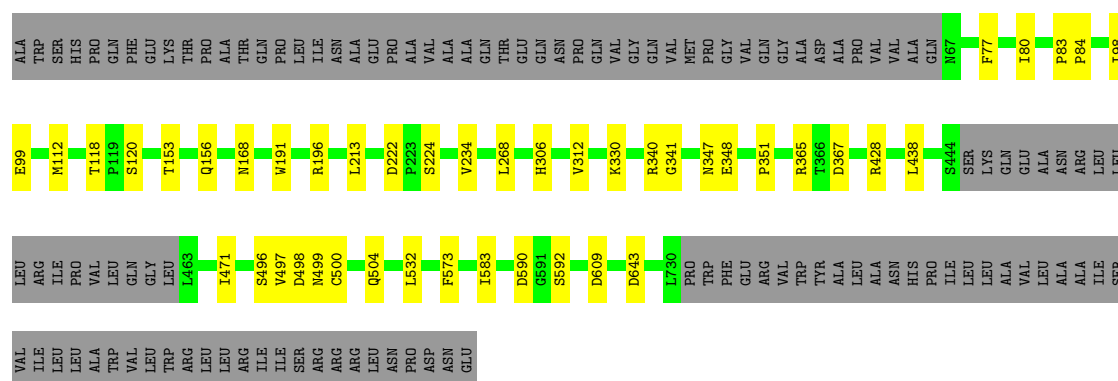
• Molecule 1: Cellulose synthase operon protein B

Chain H: 78% 7% 15%



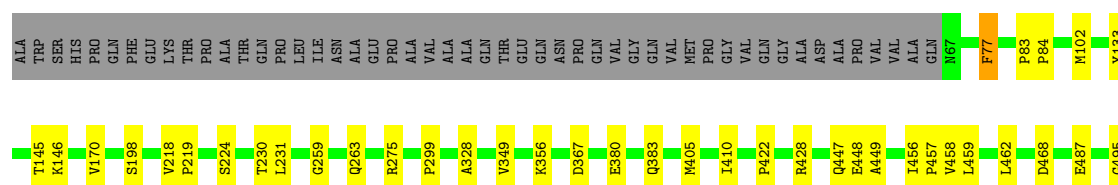
• Molecule 1: Cellulose synthase operon protein B

Chain I: 79% 6% 15%

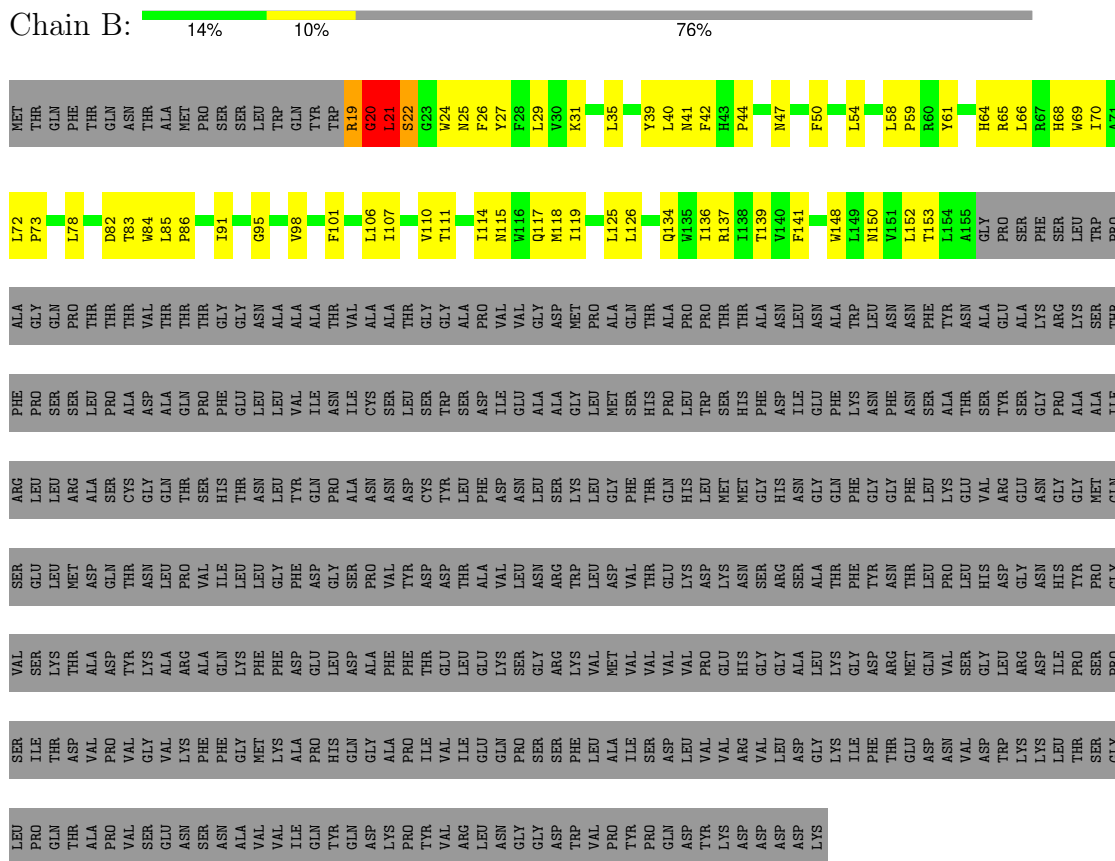


• Molecule 1: Cellulose synthase operon protein B

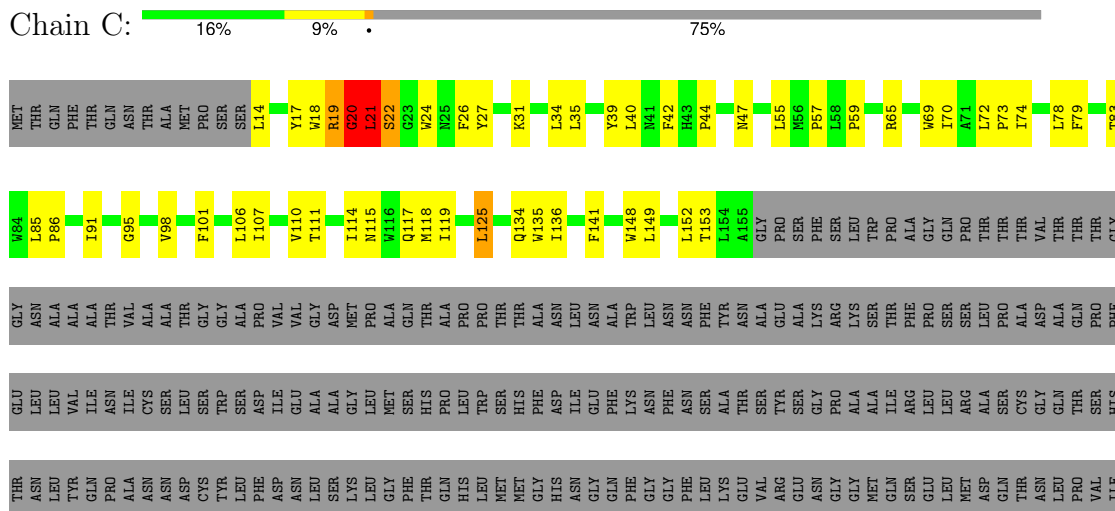
Chain J: 79% 8% 13%

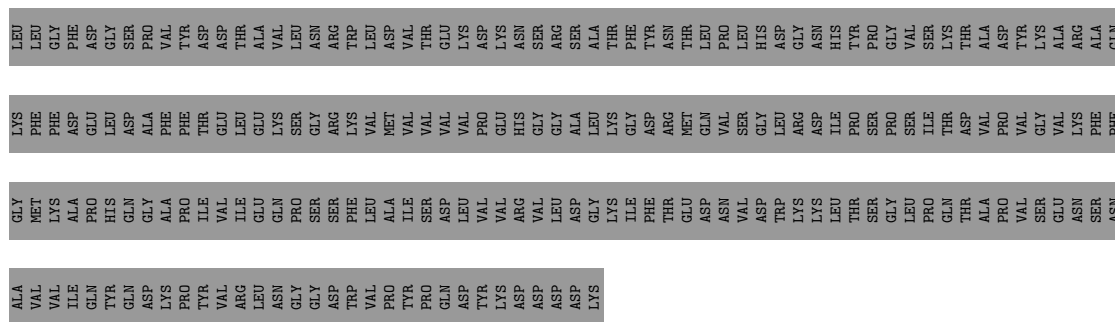


- Molecule 2: Cellulose biosynthesis protein BcsG

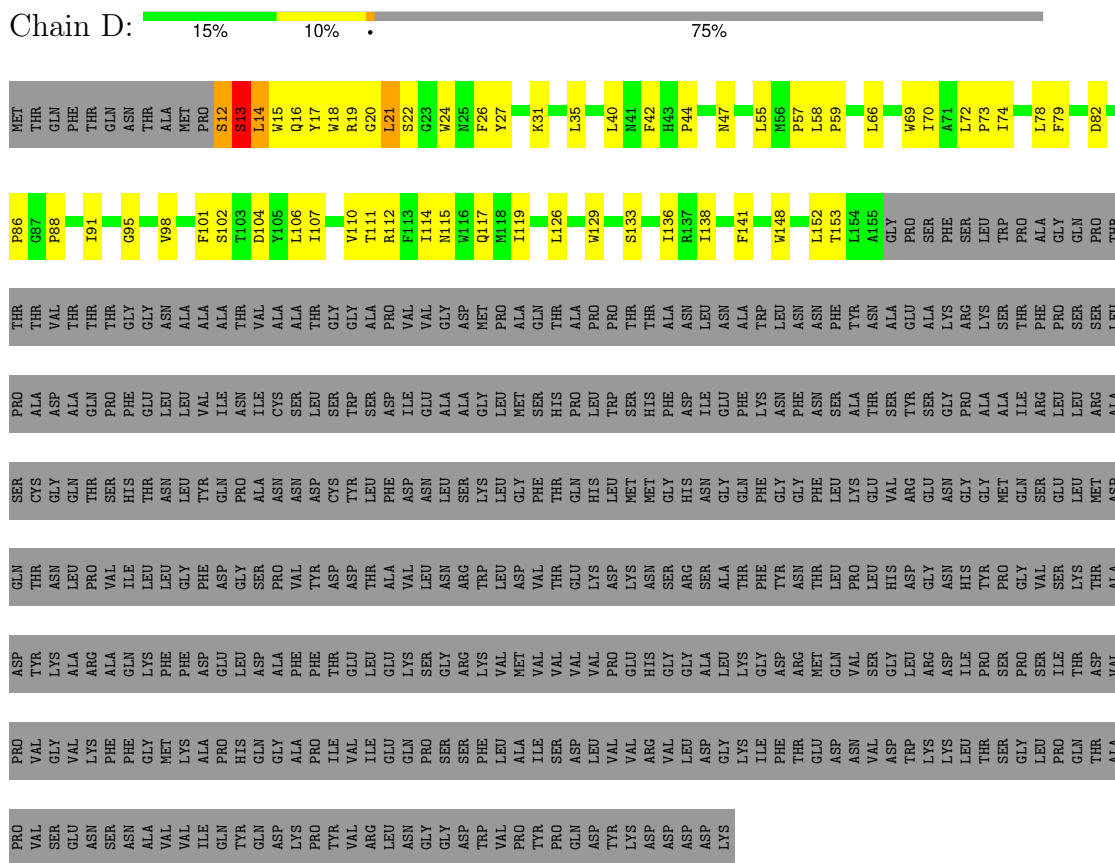


- Molecule 2: Cellulose biosynthesis protein BcsG

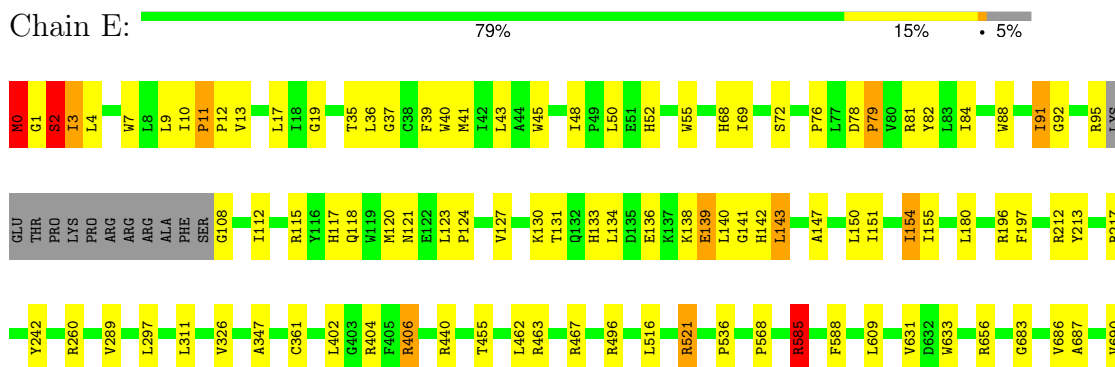




- Molecule 2: Cellulose biosynthesis protein BcsG



- Molecule 3: Cellulose synthase catalytic subunit [UDP-forming]





4 Experimental information

Property	Value	Source
EM reconstruction method	SINGLE PARTICLE	Depositor
Imposed symmetry	POINT, Not provided	
Number of particles used	39522	Depositor
Resolution determination method	FSC 0.143 CUT-OFF	Depositor
CTF correction method	PHASE FLIPPING AND AMPLITUDE CORRECTION	Depositor
Microscope	FEI TITAN KRIOS	Depositor
Voltage (kV)	300	Depositor
Electron dose ($e^-/\text{\AA}^2$)	51	Depositor
Minimum defocus (nm)	1000	Depositor
Maximum defocus (nm)	2500	Depositor
Magnification	Not provided	
Image detector	GATAN K3 BIOQUANTUM (6k x 4k)	Depositor

5 Model quality

5.1 Standard geometry

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	0.25	0/5113	0.44	0/6959
1	F	0.25	0/5125	0.43	0/6975
1	G	0.29	0/5312	0.47	1/7233 (0.0%)
1	H	0.25	0/5110	0.43	0/6956
1	I	0.25	0/5118	0.42	0/6967
1	J	0.25	0/5263	0.43	0/7164
2	B	0.92	4/1180 (0.3%)	0.76	5/1617 (0.3%)
2	C	0.89	3/1242 (0.2%)	0.72	3/1704 (0.2%)
2	D	0.85	5/1254 (0.4%)	0.86	4/1720 (0.2%)
3	E	0.70	5/7028 (0.1%)	0.92	21/9560 (0.2%)
All	All	0.45	17/41745 (0.0%)	0.58	34/56855 (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	G	0	4
2	B	0	3
2	C	0	3
2	D	0	8
3	E	0	6
All	All	0	24

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	59	PRO	N-CD	-9.09	1.35	1.47
2	D	86	PRO	N-CD	-8.89	1.35	1.47
2	B	85	LEU	C-N	8.78	1.50	1.34
2	C	86	PRO	N-CD	-8.67	1.35	1.47
2	B	86	PRO	N-CD	-8.07	1.36	1.47

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	13	SER	O-C-N	-15.61	97.73	122.70
2	D	12	SER	O-C-N	-14.51	99.48	122.70
3	E	2	SER	O-C-N	-13.61	100.92	122.70
3	E	463	ARG	NE-CZ-NH1	10.37	125.48	120.30
3	E	521	ARG	NE-CZ-NH1	9.81	125.20	120.30

There are no chirality outliers.

5 of 24 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	19	ARG	Mainchain
1	G	729	HIS	Mainchain
1	G	731	PRO	Peptide
1	G	772	ARG	Sidechain
1	G	773	ARG	Sidechain

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	4999	4953	4951	38	0
1	F	5011	4960	4958	22	0
1	G	5190	4730	5179	55	0
1	H	4996	4942	4940	28	0
1	I	5004	4953	4951	27	0
1	J	5147	5116	5114	38	0
2	B	1136	0	1154	58	0
2	C	1193	0	1201	55	0
2	D	1205	0	1212	64	0
3	E	6831	0	6917	153	0
All	All	40712	29654	40577	510	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 6.

The worst 5 of 510 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)
3:E:10:ILE:CD1	3:E:12:PRO:HG2	1.65	1.25
1:G:730:LEU:HB3	1:G:731:PRO:HD2	1.23	1.19
2:D:40:LEU:HD21	2:D:82:ASP:HB3	1.22	1.14
2:B:40:LEU:HD21	2:B:82:ASP:HB3	1.29	1.08
2:C:44:PRO:HG2	2:C:153:THR:HG21	1.36	1.07

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 PDB entries followed by that with respect to all EM entries.

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	641/763 (84%)	611 (95%)	30 (5%)	0	100	100
1	F	643/763 (84%)	622 (97%)	21 (3%)	0	100	100
1	G	658/763 (86%)	631 (96%)	23 (4%)	4 (1%)	22	58
1	H	641/763 (84%)	620 (97%)	21 (3%)	0	100	100
1	I	642/763 (84%)	625 (97%)	17 (3%)	0	100	100
1	J	662/763 (87%)	636 (96%)	26 (4%)	0	100	100
2	B	135/567 (24%)	131 (97%)	2 (2%)	2 (2%)	8	40
2	C	140/567 (25%)	136 (97%)	1 (1%)	3 (2%)	5	33
2	D	142/567 (25%)	136 (96%)	5 (4%)	1 (1%)	19	55
3	E	841/887 (95%)	816 (97%)	18 (2%)	7 (1%)	16	53
All	All	5145/7166 (72%)	4964 (96%)	164 (3%)	17 (0%)	38	71

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	730	LEU
1	G	732	TRP
2	D	14	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	E	3	ILE
3	E	139	GLU

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 PDB entries followed by that with respect to all EM entries.

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	557/657 (85%)	553 (99%)	4 (1%)	81	86
1	F	558/657 (85%)	554 (99%)	4 (1%)	81	86
1	G	574/657 (87%)	566 (99%)	8 (1%)	62	75
1	H	556/657 (85%)	549 (99%)	7 (1%)	65	77
1	I	557/657 (85%)	550 (99%)	7 (1%)	65	77
1	J	573/657 (87%)	568 (99%)	5 (1%)	75	83
2	B	118/480 (25%)	118 (100%)	0	100	100
2	C	123/480 (26%)	122 (99%)	1 (1%)	79	84
2	D	125/480 (26%)	125 (100%)	0	100	100
3	E	729/767 (95%)	718 (98%)	11 (2%)	60	75
All	All	4470/6149 (73%)	4423 (99%)	47 (1%)	69	80

5 of 47 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	130	LEU
1	I	77	PHE
1	H	202	LEU
1	H	500	CYS
1	I	153	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	293	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	F	490	ASN
1	J	508	ASN
1	H	393	ASN
2	C	134	GLN

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 Map visualisation

This section contains visualisations of the EMDB entry EMD-44359. These allow visual inspection of the internal detail of the map and identification of artifacts.

No raw map or half-maps were deposited for this entry and therefore no images, graphs, etc. pertaining to the raw map can be shown.

6.1 Orthogonal projections

This section was not generated.

6.2 Central slices

This section was not generated.

6.3 Largest variance slices

This section was not generated.

6.4 Orthogonal standard-deviation projections (False-color)

This section was not generated.

6.5 Orthogonal surface views

This section was not generated.

6.6 Mask visualisation

This section was not generated. No masks/segmentation were deposited.

7 Map analysis ⓘ

This section contains the results of statistical analysis of the map.

7.1 Map-value distribution ⓘ

This section was not generated.

7.2 Volume estimate versus contour level ⓘ

This section was not generated.

7.3 Rotationally averaged power spectrum ⓘ

This section was not generated. The rotationally averaged power spectrum had issues being displayed.

8 Fourier-Shell correlation

This section was not generated. No FSC curve or half-maps provided.

9 Map-model fit

This section was not generated.