



# Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 03:52 pm BST

PDB ID : 2REE  
Title : Crystal structure of the loading GNATL domain of CurA from *Lyngbya majuscula*  
Authors : Geders, T.W.; Smith, J.L.  
Deposited on : 2007-09-26  
Resolution : 1.95 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](mailto: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

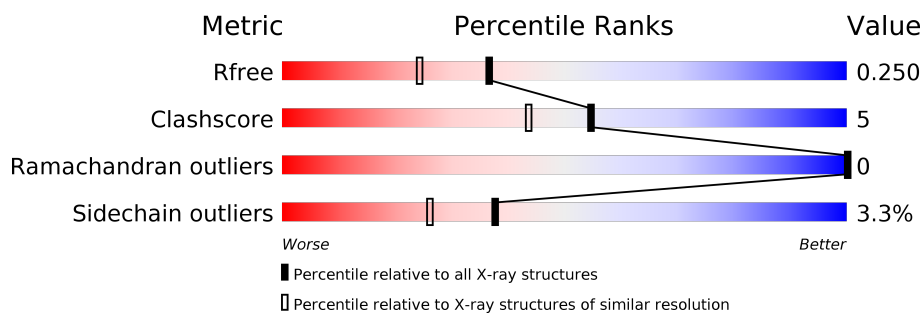
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.95 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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)

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  $\geq 3$ , 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	224	
1	B	224	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 3925 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.

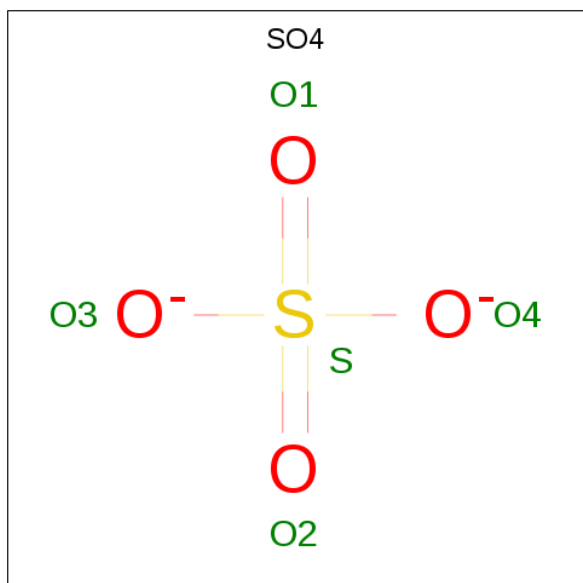
- Molecule 1 is a protein called CurA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	199	Total	C	N	O	S	0	17	0
			1731	1107	295	319	10			
1	B	199	Total	C	N	O	S	0	13	0
			1701	1085	290	317	9			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	216	SER	-	EXPRESSION TAG	UNP Q6DNF2
A	217	ASN	-	EXPRESSION TAG	UNP Q6DNF2
A	218	ALA	-	EXPRESSION TAG	UNP Q6DNF2
B	216	SER	-	EXPRESSION TAG	UNP Q6DNF2
B	217	ASN	-	EXPRESSION TAG	UNP Q6DNF2
B	218	ALA	-	EXPRESSION TAG	UNP Q6DNF2

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 10 8 2	0	1
2	B	1	Total O S 10 8 2	0	1

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0


- Molecule 5 is water.

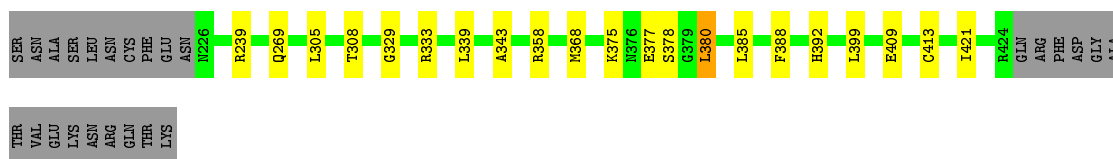
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	243	Total O 245 245	0	2
5	B	214	Total O 215 215	0	1

### 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. 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.

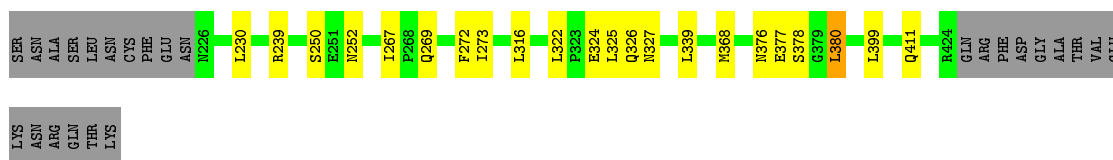
#### • Molecule 1: CurA

Chain A: 



#### • Molecule 1: CurA

Chain B: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.46 Å 91.46 Å 138.88 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	43.40 – 1.95 43.44 – 1.95	Depositor EDS
% Data completeness (in resolution range)	99.6 (43.40-1.95) 99.5 (43.44-1.95)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.54 (at 1.95 Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.172 , 0.211 0.220 , 0.250	Depositor DCC
$R_{free}$ test set	2501 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	31.4	Xtriage
Anisotropy	0.465	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 59.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.022 for -h,-k,l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3925	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	29.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.22% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: GOL, SO4, CL

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	0.63	0/1795	0.65	0/2432
1	B	0.55	0/1756	0.66	0/2385
All	All	0.59	0/3551	0.65	0/4817

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	1731	0	1762	14	0
1	B	1701	0	1715	22	0
2	A	10	0	0	0	0
2	B	10	0	0	1	0
3	A	1	0	0	0	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
5	A	245	0	0	4	0
5	B	215	0	0	3	0
All	All	3925	0	3493	36	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 5.

All (36) 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:B:376[B]:ASN:OD1	1:B:378[B]:SER:HB2	1.51	1.11
1:B:411:GLN:NE2	5:B:793:HOH:O	1.86	1.06
1:B:269[B]:GLN:H	1:B:269[B]:GLN:HE21	1.24	0.86
1:B:378[B]:SER:HB3	1:B:380:LEU:HD22	1.69	0.74
1:B:269[B]:GLN:H	1:B:269[B]:GLN:NE2	1.85	0.74
1:B:376[B]:ASN:OD1	1:B:378[B]:SER:CB	2.38	0.67
1:B:377[A]:GLU:H	1:B:377[A]:GLU:CD	1.97	0.66
1:A:399:LEU:HB3	1:A:413[B]:CYS:SG	2.37	0.65
1:A:375[A]:LYS:O	5:A:889:HOH:O	2.15	0.64
1:B:368:MET:HG3	1:B:399:LEU:HD21	1.84	0.59
1:A:308[A]:THR:HG21	5:A:791:HOH:O	2.02	0.59
1:B:326:GLN:HG2	2:B:502[A]:SO4:O3	2.03	0.59
1:B:376[B]:ASN:CG	1:B:378[B]:SER:HB2	2.23	0.58
1:B:324:GLU:HG3	1:B:325:LEU:HD12	1.86	0.57
1:B:378[B]:SER:CB	5:B:747:HOH:O	2.52	0.57
1:A:343:ALA:HA	1:A:421:ILE:HD12	1.87	0.56
1:A:358[A]:ARG:HH21	1:A:385[A]:LEU:CD2	2.18	0.56
1:A:308[A]:THR:HG22	5:A:916:HOH:O	2.07	0.54
1:A:358[A]:ARG:HH21	1:A:385[A]:LEU:HD23	1.74	0.53
1:B:250:SER:OG	1:B:252[B]:ASN:ND2	2.43	0.51
1:B:378[B]:SER:HB3	5:B:747:HOH:O	2.11	0.51
1:B:376[B]:ASN:C	1:B:378[B]:SER:H	2.14	0.50
1:A:409:GLU:OE1	5:A:915:HOH:O	2.20	0.47
1:B:378[B]:SER:HB2	1:B:380:LEU:HB2	1.96	0.47
1:A:368[A]:MET:HG2	1:A:399:LEU:HD21	1.98	0.45
1:B:376[B]:ASN:C	1:B:378[B]:SER:N	2.69	0.44
1:B:252[B]:ASN:H	1:B:252[B]:ASN:HD22	1.64	0.44
1:A:377[A]:GLU:H	1:A:377[A]:GLU:CD	2.21	0.44
1:A:378[B]:SER:HB3	1:A:380:LEU:HD22	1.99	0.44
1:B:230:LEU:HD11	1:B:272:PHE:HB3	2.00	0.43
1:A:388:PHE:O	1:A:392:HIS:HD2	2.03	0.42
1:A:329:GLY:O	1:A:333:ARG:HD3	2.19	0.42
1:A:399:LEU:HD22	1:A:413[B]:CYS:SG	2.59	0.41
1:B:376[B]:ASN:OD1	1:B:378[B]:SER:N	2.52	0.41
1:B:239:ARG:HD2	1:B:239:ARG:HA	1.90	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	214/224 (96%)	210 (98%)	4 (2%)	0	100	100
1	B	210/224 (94%)	204 (97%)	6 (3%)	0	100	100
All	All	424/448 (95%)	414 (98%)	10 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	198/204 (97%)	192 (97%)	6 (3%)	41	30
1	B	194/204 (95%)	187 (96%)	7 (4%)	35	23
All	All	392/408 (96%)	379 (97%)	13 (3%)	38	26

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	239	ARG
1	A	269[A]	GLN
1	A	269[B]	GLN
1	A	305	LEU
1	A	339	LEU
1	A	380	LEU
1	B	267	ILE
1	B	273	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	316	LEU
1	B	322	LEU
1	B	327	ASN
1	B	339	LEU
1	B	380	LEU

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

Mol	Chain	Res	Type
1	A	328	GLN
1	A	411	GLN
1	B	411	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	$\# Z  > 2$
2	SO4	B	502[B]	-	4,4,4	0.12	0	6,6,6	0.19	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	GOL	B	602	-	5,5,5	0.25	0	5,5,5	0.97	0
2	SO4	A	501[B]	-	4,4,4	0.19	0	6,6,6	0.18	0
2	SO4	B	502[A]	-	4,4,4	0.22	0	6,6,6	0.18	0
4	GOL	A	601	-	5,5,5	0.35	0	5,5,5	0.86	0
2	SO4	A	501[A]	-	4,4,4	0.23	0	6,6,6	0.35	0

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	GOL	B	602	-	-	3/4/4/4	-
4	GOL	A	601	-	-	0/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	602	GOL	C1-C2-C3-O3
4	B	602	GOL	O2-C2-C3-O3
4	B	602	GOL	O1-C1-C2-C3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	502[A]	SO4	1	0

## 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 ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.3 Carbohydrates ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.4 Ligands ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.

### 6.5 Other polymers ⓘ

Unable to reproduce the depositors R factor - this section is therefore empty.