wwPDB X-ray Structure Validation Summary Report (i)

Oct 6, 2021 – 05:09 am BST

PDB ID	:	7BME
Title	:	Crystal structure of a R18W mutant of the DNA-binding protein RemA from Geobacillus thermodenitrificans
Authors	:	Altegoer, F.; Mrusek, D.; Bange, G.
Deposited on	:	2021-01-20
Resolution	:	2.60 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

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

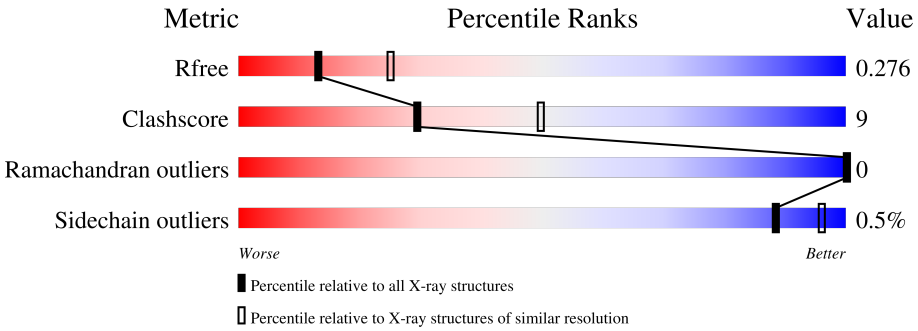
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.60 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)

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 <=5%

Mol	Chain	Length	Quality of chain
1	A	94	61% 19% 20%
1	B	94	66% 15% • 18%
1	C	94	67% 13% 20%
1	D	94	65% 13% • 21%
1	E	94	68% 14% 18%
1	F	94	67% 14% 19%
1	G	94	53% 27% 20%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4051 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 Putative regulatory protein GTNG_1019.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	75	Total	C	N	O	S	0	0	0
			579	364	107	106	2			
1	B	77	Total	C	N	O	S	0	0	0
			589	373	106	107	3			
1	C	75	Total	C	N	O	S	0	0	0
			577	365	104	107	1			
1	D	74	Total	C	N	O	S	0	0	0
			559	353	100	105	1			
1	E	77	Total	C	N	O	S	0	0	0
			593	375	106	109	3			
1	F	76	Total	C	N	O	S	0	0	0
			585	370	105	108	2			
1	G	75	Total	C	N	O	S	0	0	0
			562	353	104	103	2			

There are 63 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-6	MET	-	initiating methionine	UNP A4IM41
A	-5	GLY	-	expression tag	UNP A4IM41
A	-4	HIS	-	expression tag	UNP A4IM41
A	-3	HIS	-	expression tag	UNP A4IM41
A	-2	HIS	-	expression tag	UNP A4IM41
A	-1	HIS	-	expression tag	UNP A4IM41
A	0	HIS	-	expression tag	UNP A4IM41
A	1	HIS	-	expression tag	UNP A4IM41
A	18	TRP	ARG	engineered mutation	UNP A4IM41
B	-6	MET	-	initiating methionine	UNP A4IM41
B	-5	GLY	-	expression tag	UNP A4IM41
B	-4	HIS	-	expression tag	UNP A4IM41
B	-3	HIS	-	expression tag	UNP A4IM41
B	-2	HIS	-	expression tag	UNP A4IM41
B	-1	HIS	-	expression tag	UNP A4IM41

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	0	HIS	-	expression tag	UNP A4IM41
B	1	HIS	-	expression tag	UNP A4IM41
B	18	TRP	ARG	engineered mutation	UNP A4IM41
C	-6	MET	-	initiating methionine	UNP A4IM41
C	-5	GLY	-	expression tag	UNP A4IM41
C	-4	HIS	-	expression tag	UNP A4IM41
C	-3	HIS	-	expression tag	UNP A4IM41
C	-2	HIS	-	expression tag	UNP A4IM41
C	-1	HIS	-	expression tag	UNP A4IM41
C	0	HIS	-	expression tag	UNP A4IM41
C	1	HIS	-	expression tag	UNP A4IM41
C	18	TRP	ARG	engineered mutation	UNP A4IM41
D	-6	MET	-	initiating methionine	UNP A4IM41
D	-5	GLY	-	expression tag	UNP A4IM41
D	-4	HIS	-	expression tag	UNP A4IM41
D	-3	HIS	-	expression tag	UNP A4IM41
D	-2	HIS	-	expression tag	UNP A4IM41
D	-1	HIS	-	expression tag	UNP A4IM41
D	0	HIS	-	expression tag	UNP A4IM41
D	1	HIS	-	expression tag	UNP A4IM41
D	18	TRP	ARG	engineered mutation	UNP A4IM41
E	-6	MET	-	initiating methionine	UNP A4IM41
E	-5	GLY	-	expression tag	UNP A4IM41
E	-4	HIS	-	expression tag	UNP A4IM41
E	-3	HIS	-	expression tag	UNP A4IM41
E	-2	HIS	-	expression tag	UNP A4IM41
E	-1	HIS	-	expression tag	UNP A4IM41
E	0	HIS	-	expression tag	UNP A4IM41
E	1	HIS	-	expression tag	UNP A4IM41
E	18	TRP	ARG	engineered mutation	UNP A4IM41
F	-6	MET	-	initiating methionine	UNP A4IM41
F	-5	GLY	-	expression tag	UNP A4IM41
F	-4	HIS	-	expression tag	UNP A4IM41
F	-3	HIS	-	expression tag	UNP A4IM41
F	-2	HIS	-	expression tag	UNP A4IM41
F	-1	HIS	-	expression tag	UNP A4IM41
F	0	HIS	-	expression tag	UNP A4IM41
F	1	HIS	-	expression tag	UNP A4IM41
F	18	TRP	ARG	engineered mutation	UNP A4IM41
G	-6	MET	-	initiating methionine	UNP A4IM41
G	-5	GLY	-	expression tag	UNP A4IM41
G	-4	HIS	-	expression tag	UNP A4IM41

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	-3	HIS	-	expression tag	UNP A4IM41
G	-2	HIS	-	expression tag	UNP A4IM41
G	-1	HIS	-	expression tag	UNP A4IM41
G	0	HIS	-	expression tag	UNP A4IM41
G	1	HIS	-	expression tag	UNP A4IM41
G	18	TRP	ARG	engineered mutation	UNP A4IM41

- Molecule 2 is water.

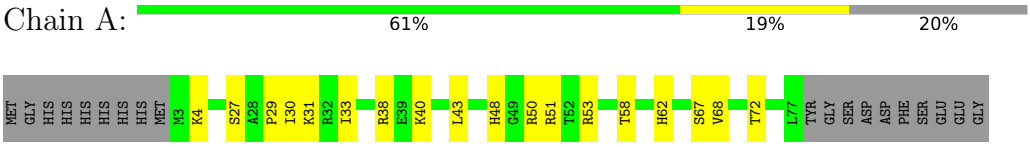
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	2	Total O 2 2	0	0
2	C	3	Total O 3 3	0	0
2	E	2	Total O 2 2	0	0

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.

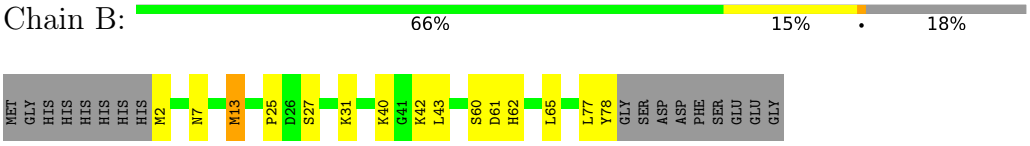
• Molecule 1: Putative regulatory protein GTNG_1019

Chain A:



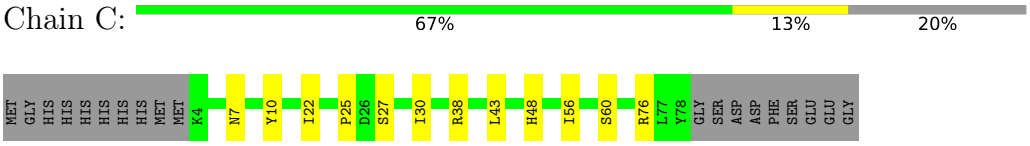
• Molecule 1: Putative regulatory protein GTNG_1019

Chain B:



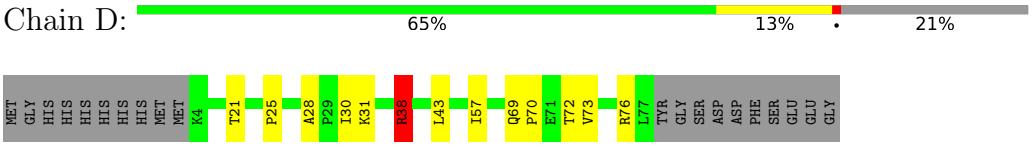
• Molecule 1: Putative regulatory protein GTNG_1019

Chain C:



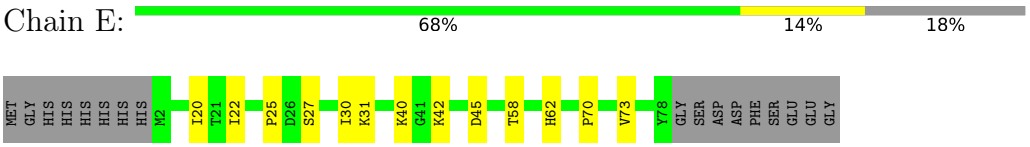
• Molecule 1: Putative regulatory protein GTNG_1019

Chain D:



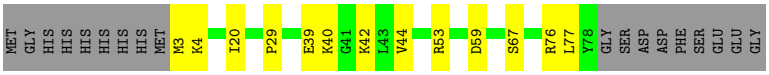
• Molecule 1: Putative regulatory protein GTNG_1019

Chain E:



• Molecule 1: Putative regulatory protein GTNG_1019

Chain F:



• Molecule 1: Putative regulatory protein GTNG_1019

Chain G:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	103.89Å 116.97Å 114.14Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.28 – 2.60 47.28 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.6 (47.28-2.60) 99.6 (47.28-2.60)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.30 (at 2.61Å)	Xtriage
Refinement program	PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.216 , 0.276 0.216 , 0.276	Depositor DCC
R_{free} test set	1089 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	76.7	Xtriage
Anisotropy	0.268	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4051	wwPDB-VP
Average B, all atoms (Å^2)	88.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.09% 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

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	0.41	0/587	0.64	1/795 (0.1%)
1	B	0.36	0/598	0.66	1/811 (0.1%)
1	C	0.38	0/586	0.58	0/796
1	D	0.39	0/567	0.96	2/771 (0.3%)
1	E	0.40	0/602	0.66	0/816
1	F	0.45	1/594 (0.2%)	0.68	1/806 (0.1%)
1	G	0.37	0/569	0.60	0/772
All	All	0.40	1/4103 (0.0%)	0.69	5/5567 (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	D	0	1
1	G	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	39	GLU	CB-CG	-5.09	1.42	1.52

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	38	ARG	NE-CZ-NH1	-17.71	111.45	120.30
1	D	38	ARG	CD-NE-CZ	11.69	139.96	123.60
1	A	31	LYS	CD-CE-NZ	-7.14	95.27	111.70
1	F	76	ARG	NE-CZ-NH1	-5.77	117.41	120.30
1	B	13	MET	CG-SD-CE	-5.49	91.42	100.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	38	ARG	Sidechain
1	G	51	ARG	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	579	0	605	15	0
1	B	589	0	608	10	0
1	C	577	0	594	10	0
1	D	559	0	574	7	0
1	E	593	0	612	15	0
1	F	585	0	603	8	0
1	G	562	0	583	17	0
2	A	2	0	0	0	0
2	C	3	0	0	0	0
2	E	2	0	0	0	0
All	All	4051	0	4179	78	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 9.

The worst 5 of 78 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:E:25:PRO:HA	1:E:30:ILE:HD11	1.44	1.00
1:F:42:LYS:NZ	1:G:61:ASP:OD2	2.14	0.81
1:B:7:ASN:HA	1:B:13:MET:HE1	1.64	0.77
1:E:25:PRO:HA	1:E:30:ILE:CD1	2.16	0.74
1:E:40:LYS:O	1:E:40:LYS:HD3	1.95	0.66

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	73/94 (78%)	70 (96%)	3 (4%)	0	100	100
1	B	75/94 (80%)	71 (95%)	4 (5%)	0	100	100
1	C	73/94 (78%)	71 (97%)	2 (3%)	0	100	100
1	D	72/94 (77%)	71 (99%)	1 (1%)	0	100	100
1	E	75/94 (80%)	74 (99%)	1 (1%)	0	100	100
1	F	74/94 (79%)	74 (100%)	0	0	100	100
1	G	73/94 (78%)	66 (90%)	7 (10%)	0	100	100
All	All	515/658 (78%)	497 (96%)	18 (4%)	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	64/80 (80%)	64 (100%)	0	100	100
1	B	64/80 (80%)	63 (98%)	1 (2%)	62	82
1	C	63/80 (79%)	63 (100%)	0	100	100
1	D	61/80 (76%)	60 (98%)	1 (2%)	62	82
1	E	65/80 (81%)	65 (100%)	0	100	100
1	F	64/80 (80%)	64 (100%)	0	100	100
1	G	61/80 (76%)	61 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	442/560 (79%)	440 (100%)	2 (0%)	88 96

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	78	TYR
1	D	31	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	75	ASN
1	D	75	ASN
1	E	75	ASN
1	F	35	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 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)

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.