



Full wwPDB X-ray Structure Validation Report ⓘ

May 14, 2020 – 07:07 pm BST

PDB ID : 3NPI
Title : Crystal structure of a TetR family regulatory protein (DIP1788) from CORYNEBACTERIUM DIPHTHERIAE at 2.96 Å resolution
Authors : Joint Center for Structural Genomics (JCSG)
Deposited on : 2010-06-28
Resolution : 2.96 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

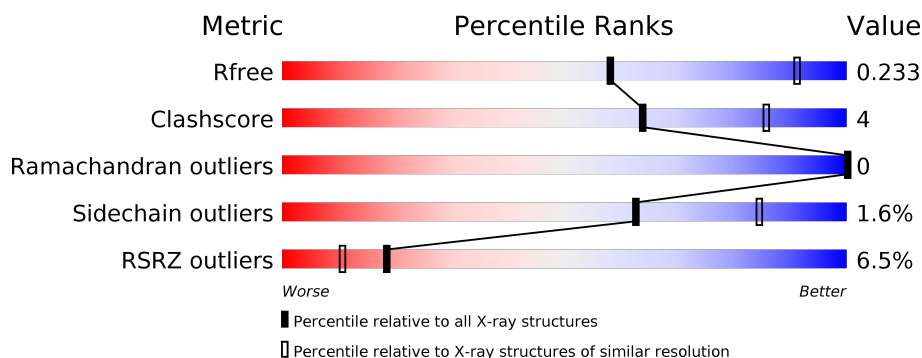
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.96 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	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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 $\leq 5\%$. 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	251	<div> <div>6%</div> <div> <div></div> <div>76%</div> <div>9%</div> <div>15%</div> </div> </div>
1	B	251	<div> <div>4%</div> <div> <div></div> <div>74%</div> <div>10%</div> <div>16%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	SO4	A	258	-	-	-	X

2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 3211 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 TetR family regulatory protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	213	Total	C	N	O	S	Se	0	0	0
			1577	988	277	300	2	10			
1	B	210	Total	C	N	O	S	Se	0	1	0
			1594	1002	277	303	2	10			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	leader sequence	UNP Q6NFU8
B	0	GLY	-	leader sequence	UNP Q6NFU8

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



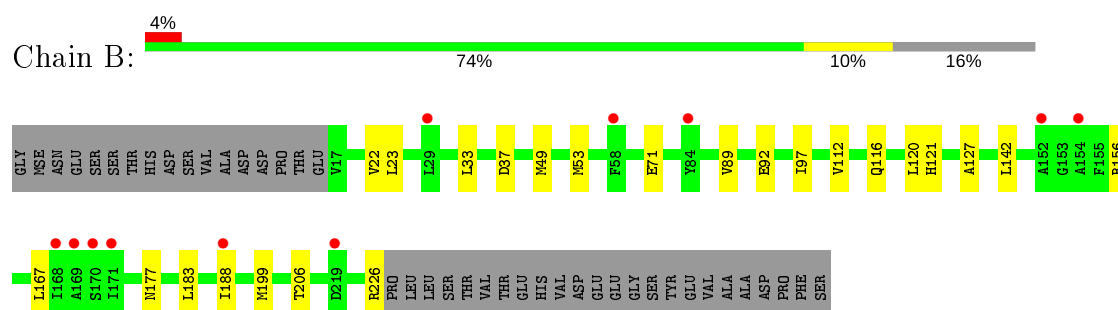
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 1: TetR family regulatory protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α , β , γ	137.24Å 137.24Å 54.73Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.71 – 2.96 29.71 – 2.96	Depositor EDS
% Data completeness (in resolution range)	(Not available) (29.71-2.96) 99.8 (29.71-2.96)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.65 (at 2.95Å)	Xtriage
Refinement program	BUSTER-TNT BUSTER 2.8.0, BUSTER 2.8.0	Depositor
R, R_{free}	0.201 , 0.222 0.211 , 0.233	Depositor DCC
R_{free} test set	609 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å ²)	87.6	Xtriage
Anisotropy	0.173	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 70.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.037 for -h,-k,l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	3211	wwPDB-VP
Average B, all atoms (Å ²)	97.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.03% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: SO4

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	0.47	0/1592	0.19	0/2142
1	B	0.48	0/1611	0.20	0/2164
All	All	0.47	0/3203	0.19	0/4306

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	1577	0	1487	16	0
1	B	1594	0	1539	19	0
2	A	20	0	0	0	0
2	B	20	0	0	0	0
All	All	3211	0	3026	26	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 4.

All (26) 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:120:LEU:HD12	1:B:120:LEU:HD12	1.56	0.88
1:A:119:ASN:HB3	1:B:177:ASN:O	1.97	0.64
1:A:127:ALA:HB1	1:B:183:LEU:HD23	1.85	0.57
1:B:112:VAL:O	1:B:116[B]:GLN:HG2	2.06	0.55
1:A:117:MSE:SE	1:B:120:LEU:HD13	2.58	0.53
1:B:49:MSE:HE2	1:B:53:MSE:HG3	1.92	0.52
1:A:49:MSE:HE2	1:A:53:MSE:HG3	1.93	0.50
1:A:206:THR:HG23	1:B:206:THR:HG23	1.93	0.50
1:A:188:ILE:HG12	1:B:226:ARG:HB3	1.93	0.49
1:A:120:LEU:HD11	1:B:116[A]:GLN:HG3	1.95	0.48
1:B:167:LEU:HD21	1:B:206:THR:HG22	1.95	0.48
1:A:179:ARG:HG3	1:A:191:MSE:HG3	1.96	0.48
1:B:22:VAL:HG21	1:B:49:MSE:SE	2.63	0.48
1:A:183:LEU:HD23	1:B:127:ALA:HB1	1.96	0.48
1:A:167:LEU:HD21	1:A:206:THR:HG22	1.94	0.47
1:A:33:LEU:HB3	1:A:37:ASP:HB2	1.99	0.45
1:A:199:MSE:HA	1:A:199:MSE:HE2	1.99	0.45
1:B:199:MSE:HA	1:B:199:MSE:HE2	1.98	0.45
1:B:116[B]:GLN:HG3	1:B:177:ASN:HD21	1.81	0.45
1:B:33:LEU:HB3	1:B:37:ASP:HB2	1.99	0.44
1:B:89:VAL:HG23	1:B:92:GLU:HB2	1.98	0.44
1:B:23:LEU:HD21	1:B:71:GLU:HB2	2.00	0.44
1:A:22:VAL:HG21	1:A:49:MSE:SE	2.68	0.44
1:B:97:ILE:HG21	1:B:142:LEU:HD13	2.00	0.42
1:A:93:GLY:O	1:A:97:ILE:HG12	2.21	0.41
1:A:226:ARG:HB3	1:B:188:ILE:HG12	2.04	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	211/251 (84%)	207 (98%)	4 (2%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	209/251 (83%)	201 (96%)	8 (4%)	0	100	100
All	All	420/502 (84%)	408 (97%)	12 (3%)	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	155/202 (77%)	152 (98%)	3 (2%)	57	81
1	B	162/202 (80%)	160 (99%)	2 (1%)	71	88
All	All	317/404 (78%)	312 (98%)	5 (2%)	62	84

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	52	ARG
1	A	121	HIS
1	A	156	ARG
1	B	121	HIS
1	B	156	ARG

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 [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

8 ligands are 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	SO4	B	254	-	4,4,4	0.17	0	6,6,6	0.06	0
2	SO4	A	255	-	4,4,4	0.24	0	6,6,6	0.04	0
2	SO4	B	251	-	4,4,4	0.23	0	6,6,6	0.04	0
2	SO4	A	258	-	4,4,4	0.17	0	6,6,6	0.05	0
2	SO4	B	256	-	4,4,4	0.16	0	6,6,6	0.05	0
2	SO4	A	257	-	4,4,4	0.16	0	6,6,6	0.05	0
2	SO4	B	253	-	4,4,4	0.20	0	6,6,6	0.05	0
2	SO4	A	252	-	4,4,4	0.17	0	6,6,6	0.05	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

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, 95th 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	203/251 (80%)	0.40	15 (7%) 14 8	70, 106, 143, 170	0
1	B	200/251 (79%)	0.40	11 (5%) 25 15	59, 86, 117, 132	0
All	All	403/502 (80%)	0.40	26 (6%) 18 11	59, 97, 133, 170	0

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	86	ALA	5.8
1	A	170	SER	3.9
1	B	170	SER	3.4
1	A	227	PRO	3.4
1	A	15	THR	3.3
1	B	152	ALA	3.0
1	A	171	ILE	3.0
1	B	84	TYR	2.8
1	B	171	ILE	2.8
1	A	88	ALA	2.8
1	B	219	ASP	2.8
1	A	217	SER	2.7
1	B	58	PHE	2.6
1	A	107	GLN	2.5
1	A	16	GLU	2.5
1	A	166	THR	2.5
1	B	29	LEU	2.5
1	A	167	LEU	2.4
1	A	106	VAL	2.2
1	A	168	ILE	2.2
1	B	188	ILE	2.2
1	B	154	ALA	2.1
1	B	169	ALA	2.1
1	A	169	ALA	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	116	GLN	2.1
1	B	168	ILE	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, 95th 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	SO4	A	258	5/5	0.77	0.53	199,203,204,204	0
2	SO4	B	256	5/5	0.84	0.27	155,160,161,161	0
2	SO4	A	257	5/5	0.85	0.42	163,167,168,168	0
2	SO4	A	255	5/5	0.87	0.17	158,162,163,164	0
2	SO4	B	253	5/5	0.89	0.20	149,153,154,155	0
2	SO4	B	254	5/5	0.90	0.18	143,147,148,149	0
2	SO4	B	251	5/5	0.91	0.21	135,140,140,141	0
2	SO4	A	252	5/5	0.92	0.34	134,138,139,140	0

6.5 Other polymers [i](#)

There are no such residues in this entry.