



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 5, 2026 – 04:23 PM UTC

PDB ID : 3PG5 / pdb_00003pg5
Title : Crystal structure of protein DIP2308 from Corynebacterium diphtheriae, Northeast Structural Genomics Consortium Target Cdr78
Authors : Forouhar, F.; Lew, S.; Seetharaman, J.; Lee, D.; Ciccocanti, C.; Sahdev, S.; Nair, R.; Rost, B.; Acton, T.B.; Xiao, R.; Everett, J.K.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on : 2010-10-30
Resolution : 3.30 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 ⓘ 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 : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.49

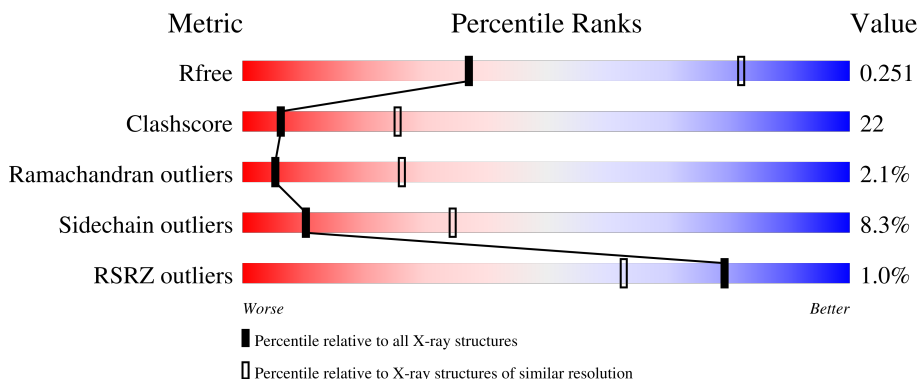
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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}	180053	1169 (3.32-3.28)
Clashscore	190562	1209 (3.32-3.28)
Ramachandran outliers	187476	1188 (3.32-3.28)
Sidechain outliers	187428	1187 (3.32-3.28)
RSRZ outliers	180081	1169 (3.32-3.28)

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 $\leq 5\%$. 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	361	 46% 32% 5% 17%
1	B	361	 42% 36% 5% 17%
1	C	361	 50% 33% 5% 14%
1	D	361	 47% 34% 5% 16%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 9613 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	298	2358	1495	412	442	3	6	0	0	0
1	B	301	2377	1506	416	446	3	6	0	0	0
1	C	312	2468	1566	430	463	3	6	0	0	0
1	D	302	2389	1516	418	446	3	6	0	0	0

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	354	LEU	-	expression tag	UNP Q6NEG6
A	355	GLU	-	expression tag	UNP Q6NEG6
A	356	HIS	-	expression tag	UNP Q6NEG6
A	357	HIS	-	expression tag	UNP Q6NEG6
A	358	HIS	-	expression tag	UNP Q6NEG6
A	359	HIS	-	expression tag	UNP Q6NEG6
A	360	HIS	-	expression tag	UNP Q6NEG6
A	361	HIS	-	expression tag	UNP Q6NEG6
B	354	LEU	-	expression tag	UNP Q6NEG6
B	355	GLU	-	expression tag	UNP Q6NEG6
B	356	HIS	-	expression tag	UNP Q6NEG6
B	357	HIS	-	expression tag	UNP Q6NEG6
B	358	HIS	-	expression tag	UNP Q6NEG6
B	359	HIS	-	expression tag	UNP Q6NEG6
B	360	HIS	-	expression tag	UNP Q6NEG6
B	361	HIS	-	expression tag	UNP Q6NEG6
C	354	LEU	-	expression tag	UNP Q6NEG6
C	355	GLU	-	expression tag	UNP Q6NEG6
C	356	HIS	-	expression tag	UNP Q6NEG6
C	357	HIS	-	expression tag	UNP Q6NEG6
C	358	HIS	-	expression tag	UNP Q6NEG6

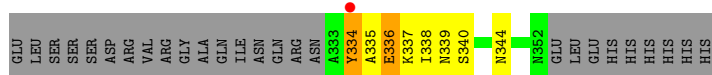
Continued on next page...

Continued from previous page...

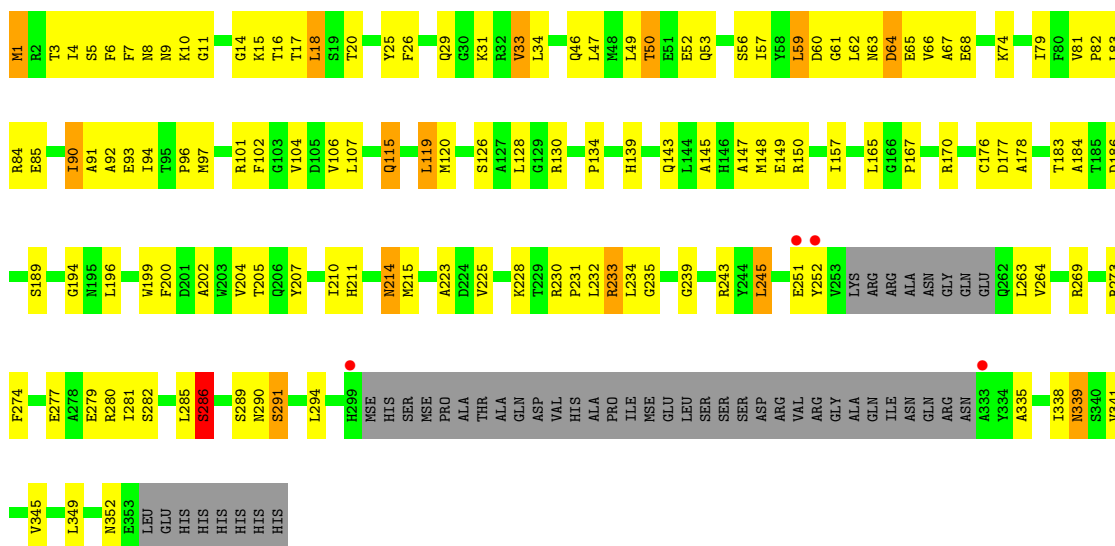
Chain	Residue	Modelled	Actual	Comment	Reference
C	359	HIS	-	expression tag	UNP Q6NEG6
C	360	HIS	-	expression tag	UNP Q6NEG6
C	361	HIS	-	expression tag	UNP Q6NEG6
D	354	LEU	-	expression tag	UNP Q6NEG6
D	355	GLU	-	expression tag	UNP Q6NEG6
D	356	HIS	-	expression tag	UNP Q6NEG6
D	357	HIS	-	expression tag	UNP Q6NEG6
D	358	HIS	-	expression tag	UNP Q6NEG6
D	359	HIS	-	expression tag	UNP Q6NEG6
D	360	HIS	-	expression tag	UNP Q6NEG6
D	361	HIS	-	expression tag	UNP Q6NEG6

- Molecule 2 is water.

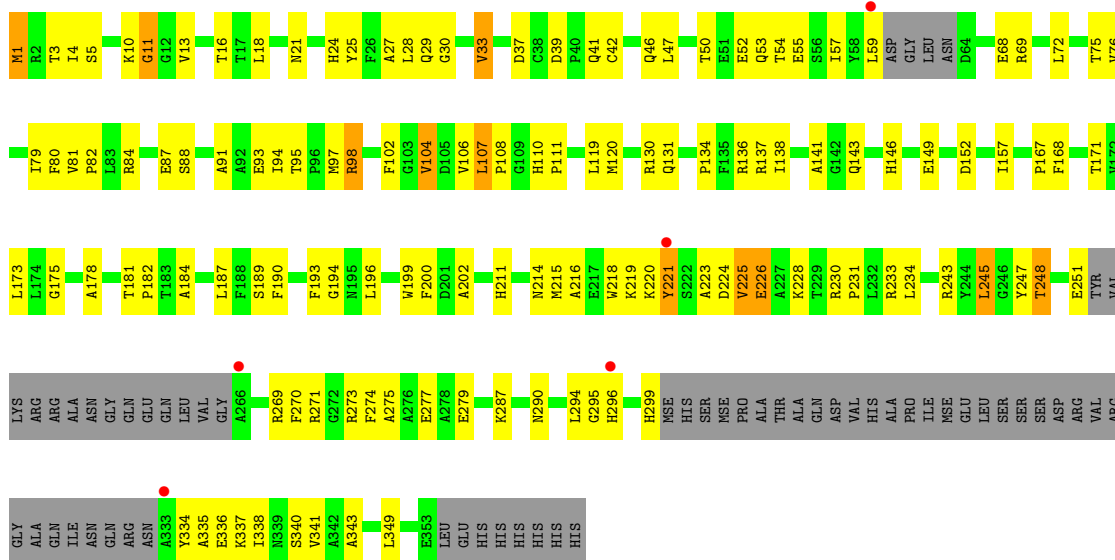
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	7	Total O 7 7	0	0
2	B	7	Total O 7 7	0	0
2	C	2	Total O 2 2	0	0
2	D	5	Total O 5 5	0	0



• Molecule 1: Uncharacterized protein



• Molecule 1: Uncharacterized protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 4 21 2	Depositor
Cell constants a, b, c, α , β , γ	199.96Å 199.96Å 106.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 3.30 20.00 – 3.30	Depositor EDS
% Data completeness (in resolution range)	88.5 (20.00-3.30) 93.9 (20.00-3.30)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	0.19	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.71 (at 3.31Å)	Xtrriage
Refinement program	CNS 1.2 & XtalView	Depositor
R, R_{free}	0.187 , 0.235 0.204 , 0.251	Depositor DCC
R_{free} test set	6105 reflections (9.85%)	wwPDB-VP
Wilson B-factor (Å ²)	60.6	Xtrriage
Anisotropy	0.231	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 42.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	9613	wwPDB-VP
Average B, all atoms (Å ²)	56.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.77% 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 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	0.60	1/2408 (0.0%)	0.89	1/3252 (0.0%)
1	B	0.56	1/2428 (0.0%)	0.91	5/3281 (0.2%)
1	C	0.56	0/2522	0.91	4/3409 (0.1%)
1	D	0.54	0/2441	0.87	2/3297 (0.1%)
All	All	0.56	2/9799 (0.0%)	0.90	12/13239 (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	A	0	1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	81	VAL	CA-CB	-6.25	1.51	1.54
1	B	81	VAL	CA-CB	-5.27	1.51	1.54

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	290	ASN	CB-CA-C	6.88	121.03	109.48
1	D	290	ASN	CB-CA-C	6.84	121.03	109.75
1	B	290	ASN	CB-CA-C	6.81	121.04	110.14
1	C	290	ASN	O-C-N	-6.39	115.11	123.16
1	C	290	ASN	CB-CA-C	6.38	121.06	109.37

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	154	TYR	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	2358	0	2277	95	0
1	B	2377	0	2296	126	0
1	C	2468	0	2378	103	0
1	D	2389	0	2300	105	0
2	A	7	0	0	0	0
2	B	7	0	0	1	0
2	C	2	0	0	0	0
2	D	5	0	0	0	0
All	All	9613	0	9251	418	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 22.

The worst 5 of 418 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:D:1:MSE:HG3	1:D:149:GLU:HB2	1.43	1.00
1:B:33:VAL:HG13	1:B:104:VAL:HG23	1.52	0.92
1:A:219:LYS:HG3	1:A:225:VAL:HG11	1.51	0.91
1:B:178:ALA:HB1	1:B:245:LEU:HD13	1.57	0.85
1:B:8:ASN:HB3	1:B:15:LYS:HD3	1.55	0.85

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	290/361 (80%)	255 (88%)	31 (11%)	4 (1%)	9	33
1	B	295/361 (82%)	255 (86%)	32 (11%)	8 (3%)	4	22
1	C	306/361 (85%)	270 (88%)	30 (10%)	6 (2%)	6	27
1	D	294/361 (81%)	247 (84%)	40 (14%)	7 (2%)	4	24
All	All	1185/1444 (82%)	1027 (87%)	133 (11%)	25 (2%)	5	26

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	233	ARG
1	C	11	GLY
1	C	286	SER
1	B	11	GLY
1	B	294	LEU

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	246/290 (85%)	220 (89%)	26 (11%)	6	24
1	B	248/290 (86%)	226 (91%)	22 (9%)	9	32
1	C	257/290 (89%)	241 (94%)	16 (6%)	16	45
1	D	248/290 (86%)	229 (92%)	19 (8%)	12	37
All	All	999/1160 (86%)	916 (92%)	83 (8%)	10	34

5 of 83 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	115	GLN
1	D	106	VAL
1	C	214	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	18	LEU
1	D	245	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	339	ASN
1	D	131	GLN
1	C	344	ASN
1	D	43	ASN
1	D	169	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 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 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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	292/361 (80%)	-0.29	1 (0%) 90 82	19, 47, 80, 98	0
1	B	295/361 (81%)	-0.18	2 (0%) 84 70	32, 54, 86, 121	0
1	C	306/361 (84%)	-0.15	4 (1%) 75 56	32, 54, 88, 106	0
1	D	296/361 (81%)	-0.14	5 (1%) 69 50	30, 55, 97, 107	0
All	All	1189/1444 (82%)	-0.19	12 (1%) 79 63	19, 53, 90, 121	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	333	ALA	3.7
1	C	252	TYR	3.0
1	D	333	ALA	3.0
1	D	221	TYR	2.9
1	A	333	ALA	2.8

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 oligosaccharides in this entry.

6.4 Ligands [i](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.