



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 25, 2017 – 10:47 am GMT

PDB ID : 5M1T
Title : PaMucR Phosphodiesterase, c-di-GMP complex
Authors : Hutchin, A.; Tews, I.; Walsh, M.A.
Deposited on : 2016-10-10
Resolution : 2.27 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

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : recalc29047
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : recalc29047

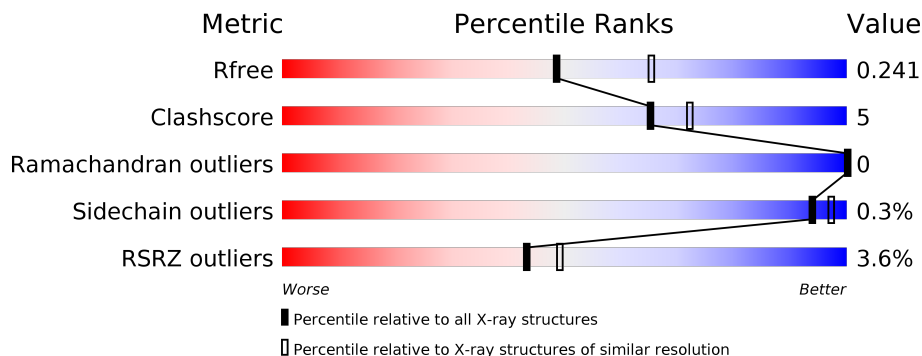
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION



The reported resolution of this entry is 2.27 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}	100719	5609 (2.30-2.26)
Clashscore	112137	6364 (2.30-2.26)
Ramachandran outliers	110173	6281 (2.30-2.26)
Sidechain outliers	110143	6281 (2.30-2.26)
RSRZ outliers	101464	5639 (2.30-2.26)

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. 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	281	
1	B	281	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3871 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 MucR Phosphodiesterase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	247	Total	C	N	O	S	0	0	0
			1913	1216	337	353	7			
1	B	225	Total	C	N	O	S	0	0	0
			1650	1050	287	306	7			

There are 42 discrepancies between the modelled and reference sequences:

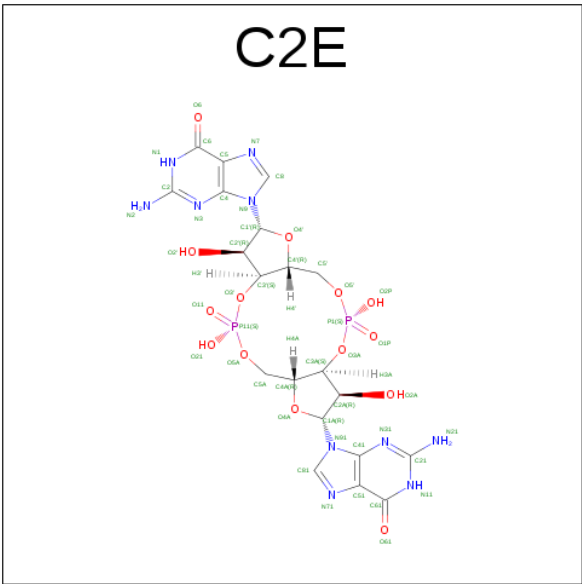
Chain	Residue	Modelled	Actual	Comment	Reference
A	405	MET	-	initiating methionine	UNP Q9I310
A	406	GLY	-	expression tag	UNP Q9I310
A	407	SER	-	expression tag	UNP Q9I310
A	408	SER	-	expression tag	UNP Q9I310
A	409	HIS	-	expression tag	UNP Q9I310
A	410	HIS	-	expression tag	UNP Q9I310
A	411	HIS	-	expression tag	UNP Q9I310
A	412	HIS	-	expression tag	UNP Q9I310
A	413	HIS	-	expression tag	UNP Q9I310
A	414	HIS	-	expression tag	UNP Q9I310
A	415	SER	-	expression tag	UNP Q9I310
A	416	SER	-	expression tag	UNP Q9I310
A	417	GLY	-	expression tag	UNP Q9I310
A	418	LEU	-	expression tag	UNP Q9I310
A	419	VAL	-	expression tag	UNP Q9I310
A	420	PRO	-	expression tag	UNP Q9I310
A	421	ARG	-	expression tag	UNP Q9I310
A	422	GLY	-	expression tag	UNP Q9I310
A	423	SER	-	expression tag	UNP Q9I310
A	424	HIS	-	expression tag	UNP Q9I310
A	425	MET	-	expression tag	UNP Q9I310
B	405	MET	-	initiating methionine	UNP Q9I310
B	406	GLY	-	expression tag	UNP Q9I310
B	407	SER	-	expression tag	UNP Q9I310
B	408	SER	-	expression tag	UNP Q9I310

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	409	HIS	-	expression tag	UNP Q9I310
B	410	HIS	-	expression tag	UNP Q9I310
B	411	HIS	-	expression tag	UNP Q9I310
B	412	HIS	-	expression tag	UNP Q9I310
B	413	HIS	-	expression tag	UNP Q9I310
B	414	HIS	-	expression tag	UNP Q9I310
B	415	SER	-	expression tag	UNP Q9I310
B	416	SER	-	expression tag	UNP Q9I310
B	417	GLY	-	expression tag	UNP Q9I310
B	418	LEU	-	expression tag	UNP Q9I310
B	419	VAL	-	expression tag	UNP Q9I310
B	420	PRO	-	expression tag	UNP Q9I310
B	421	ARG	-	expression tag	UNP Q9I310
B	422	GLY	-	expression tag	UNP Q9I310
B	423	SER	-	expression tag	UNP Q9I310
B	424	HIS	-	expression tag	UNP Q9I310
B	425	MET	-	expression tag	UNP Q9I310

- Molecule 2 is 9,9'-[(2R,3R,3aS,5S,7aR,9R,10R,10aS,12S,14aR)-3,5,10,12-tetrahydroxy-5,12-dioxidoctahydro-2H,7H-difuro[3,2-d:3',2'-j][1,3,7,9,2,8]tetraoxadiphosphacyclododecine-2,9-diyl]bis(2-amino-1,9-dihydro-6H-purin-6-one) (three-letter code: C2E) (formula: C₂₀H₂₄N₁₀O₁₄P₂).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			46	20	10	14	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	B	1	Total	C	N	O	P	0	0
			46	20	10	14	2		

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Mg	0	0
			2	2		
3	A	2	Total	Mg	0	0
			2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	131	Total	O	0	0
			131	131		
4	B	81	Total	O	0	0
			81	81		

4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	46.37Å 116.11Å 52.14Å 90.00° 102.52° 90.00°	Depositor
Resolution (Å)	45.27 – 2.27 45.27 – 2.27	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.27-2.27) 99.9 (45.27-2.27)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.17 (at 2.27Å)	Xtriage
Refinement program	REFMAC 5.8.0155	Depositor
R, R_{free}	0.195 , 0.238 0.203 , 0.241	Depositor DCC
R_{free} test set	1249 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	23.1	Xtriage
Anisotropy	0.449	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 38.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3871	wwPDB-VP
Average B, all atoms (Å ²)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.90% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, C2E

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	0/1948	0.76	2/2646 (0.1%)
1	B	0.53	0/1674	0.68	0/2275
All	All	0.57	0/3622	0.73	2/4921 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	549	ARG	NE-CZ-NH1	7.74	124.17	120.30
1	A	549	ARG	NE-CZ-NH2	-6.49	117.05	120.30

There are no chirality outliers.

There are no planarity outliers.

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	1913	0	1939	16	0
1	B	1650	0	1632	18	0
2	A	46	0	22	0	0
2	B	46	0	22	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	131	0	0	5	2

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	81	0	0	10	0
All	All	3871	0	3615	34	2

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.

The worst 5 of 34 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:512:MET:HE3	1:A:525:ILE:HB	1.50	0.91
1:A:619:GLU:HB2	1:A:626:ASP:OD2	1.73	0.88
1:B:506:ASP:CB	4:B:854:HOH:O	2.22	0.87
1:B:536:HIS:ND1	4:B:801:HOH:O	2.14	0.80
1:A:512:MET:CE	1:A:525:ILE:HB	2.11	0.79

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:A:911:HOH:O	4:A:925:HOH:O[1_455]	1.81	0.39
4:A:919:HOH:O	4:A:920:HOH:O[1_655]	1.88	0.32

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	245/281 (87%)	241 (98%)	4 (2%)	0	100	100
1	B	219/281 (78%)	213 (97%)	6 (3%)	0	100	100
All	All	464/562 (83%)	454 (98%)	10 (2%)	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	204/231 (88%)	204 (100%)	0	100	100
1	B	165/231 (71%)	164 (99%)	1 (1%)	89	94
All	All	369/462 (80%)	368 (100%)	1 (0%)	94	97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	523	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 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	C2E	A	701	3	44,52,52	1.08	4 (9%)	50,82,82	1.94	13 (26%)
2	C2E	B	701	3	44,52,52	1.08	4 (9%)	50,82,82	1.97	11 (22%)

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
2	C2E	A	701	3	-	0/22/62/62	0/6/7/7
2	C2E	B	701	3	-	0/22/62/62	0/6/7/7

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	701	C2E	C51-C41	2.43	1.46	1.40
2	A	701	C2E	C5-C4	2.59	1.46	1.40
2	A	701	C2E	C51-C41	2.60	1.46	1.40
2	B	701	C2E	C6-C5	2.92	1.46	1.41
2	A	701	C2E	C6-C5	3.03	1.47	1.41

The worst 5 of 24 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	701	C2E	C6-C5-C4	-4.75	116.12	120.84
2	A	701	C2E	C61-C51-C41	-4.40	116.47	120.84
2	A	701	C2E	C5-C6-N1	-4.32	117.33	123.48
2	A	701	C2E	C51-C61-N11	-3.92	117.90	123.48
2	B	701	C2E	C61-C51-C41	-3.86	117.01	120.84

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 [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	247/281 (87%)	-0.37	1 (0%) 92 94	20, 29, 43, 52	0
1	B	225/281 (80%)	0.22	16 (7%) 17 21	21, 40, 74, 87	0
All	All	472/562 (83%)	-0.09	17 (3%) 43 49	20, 31, 69, 87	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	482	ILE	4.4
1	B	499	GLN	3.7
1	B	450	LEU	3.6
1	B	481	LEU	3.3
1	B	514	LEU	3.3

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(\AA^2)	Q<0.9
2	C2E	A	701	46/46	0.97	0.10	-0.52	20,23,27,28	0
2	C2E	B	701	46/46	0.97	0.10	-0.54	26,31,34,40	0
3	MG	A	703	1/1	0.94	0.09	-0.96	26,26,26,26	0
3	MG	A	702	1/1	0.99	0.06	-2.85	13,13,13,13	0
3	MG	B	703	1/1	0.99	0.06	-4.35	16,16,16,16	0
3	MG	B	702	1/1	0.96	0.07	-	31,31,31,31	0

6.5 Other polymers [i](#)

There are no such residues in this entry.