



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 26, 2016 – 10:42 PM EDT

PDB ID : 5GO3
Title : Crystal structure of a di-nucleotide cyclase Vibrio mutant
Authors : Ming, Z.H.; Wang, W.; Xie, Y.C.; Chen, Y.C.; Yan, L.M.; Lou, Z.Y.
Deposited on : 2016-07-26
Resolution : 2.20 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	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20027939
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
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)	:	rb-20027939

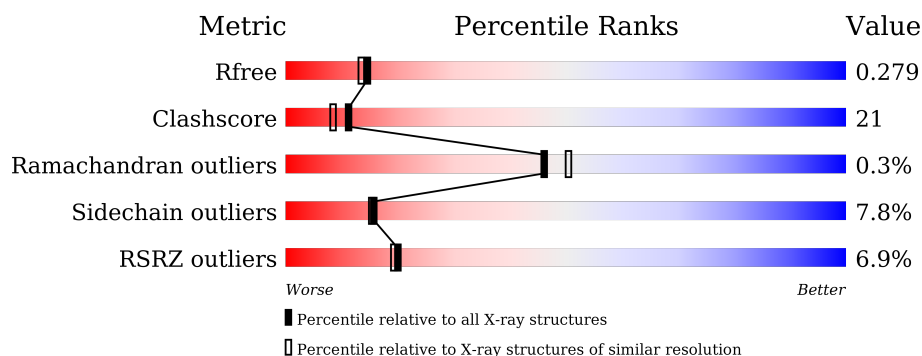
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	91344	3774 (2.20-2.20)
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (2.20-2.20)

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	447	<div> <div>7%</div> <div> <div></div> <div>50%</div> <div>26%</div> <div>5%</div> <div>19%</div> </div> </div>
1	B	447	<div> <div>4%</div> <div> <div></div> <div>56%</div> <div>21%</div> <div>•</div> <div>19%</div> </div> </div>

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6104 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 Cyclic GMP-AMP synthase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	362	Total	C	N	O	S	0	0	0
			2904	1837	503	548	16			
1	B	364	Total	C	N	O	S	0	0	0
			2922	1849	506	551	16			

There are 26 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP Q9KVG7
A	-3	PRO	-	expression tag	UNP Q9KVG7
A	-2	LEU	-	expression tag	UNP Q9KVG7
A	-1	GLY	-	expression tag	UNP Q9KVG7
A	0	SER	-	expression tag	UNP Q9KVG7
A	131	ASN	ASP	engineered mutation	UNP Q9KVG7
A	133	ASN	ASP	engineered mutation	UNP Q9KVG7
A	437	HIS	-	expression tag	UNP Q9KVG7
A	438	HIS	-	expression tag	UNP Q9KVG7
A	439	HIS	-	expression tag	UNP Q9KVG7
A	440	HIS	-	expression tag	UNP Q9KVG7
A	441	HIS	-	expression tag	UNP Q9KVG7
A	442	HIS	-	expression tag	UNP Q9KVG7
B	-4	GLY	-	expression tag	UNP Q9KVG7
B	-3	PRO	-	expression tag	UNP Q9KVG7
B	-2	LEU	-	expression tag	UNP Q9KVG7
B	-1	GLY	-	expression tag	UNP Q9KVG7
B	0	SER	-	expression tag	UNP Q9KVG7
B	131	ASN	ASP	engineered mutation	UNP Q9KVG7
B	133	ASN	ASP	engineered mutation	UNP Q9KVG7
B	437	HIS	-	expression tag	UNP Q9KVG7
B	438	HIS	-	expression tag	UNP Q9KVG7
B	439	HIS	-	expression tag	UNP Q9KVG7
B	440	HIS	-	expression tag	UNP Q9KVG7
B	441	HIS	-	expression tag	UNP Q9KVG7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	442	HIS	-	expression tag	UNP Q9KVG7

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	126	Total 126	O 126	0	0
2	B	152	Total 152	O 152	0	0

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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. A red dot above a residue indicates a poor fit to the electron density ($\text{RSRZ} > 2$). 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.

Chain A:

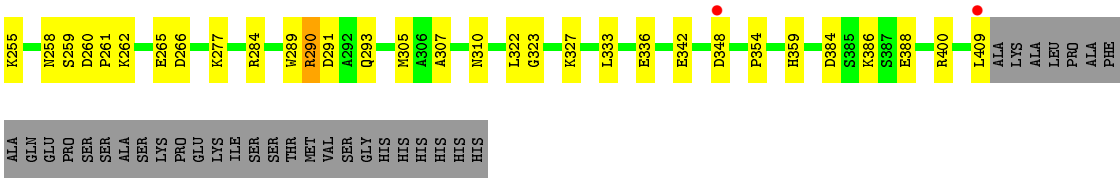
7% 50% 26% 5% 19%

Amino Acid	Count
GLY	1
PRO	1
LEU	1
GLY	1
SER	1
MET	1
ARG	1
M3	1
T4	1
W5	1
T12	1
L18	1
L22	1
V23	1
L24	1
T25	1
E28	1
R36	1
K37	1
I39	1
R40	1
L41	1
R42	1
T43	1
R44	1
K55	1
A56	1
V57	1
K58	1
K59	1
S60	1
A61	1
L62	1
I66	1
I67	1
Q68	1
E69	1
S72	1
T73	1
T74	1
Q75	1
I76	1
K77	1
H78	1
L79	1
V87	1
L90	1
W92	1
E98	1
R99	1
D100	1
E101	1
F102	1
L103	1
G104	1
L105	1
T106	1
P107	1
R108	1
F109	1
W110	1
T111	1
G112	1
G113	1
SER	1
F115	1
Q116	1
L117	1
D118	1
T119	1
L120	1
M121	1
P126	1
G127	1
Q128	1
E129	1
M130	1
M131	1
T132	1
M138	1
P139	1
M140	1
P141	1
L142	1
F143	1
E144	1
S145	1
E146	1
P147	1
K148	1
T149	1
VAL	1
ALA	1
H150	1
H151	1
S152	1
L153	1
L154	1
THR	1
I155	1
L156	1
ASP	1
SER	1
GLU	1
THR	1
TYR	1
GLU	1
LEU	1
E168	1
A160	1
S161	1
A167	1
E168	1
P200	1
ASP	1
GLU	1
PHE	1
GLN	1
LYS	1
ILE	1
ALA	1
LEU	1
GLU	1
ALA	1
ASN	1
ARG	1
SER	1
PHE	1
VAL	1
LYS	1
GLY	1
ILE	1
ALA	1
PHE	1
GLU	1
SER	1
TYR	1
VAL	1
ALA	1
ASP	1
SER	1
ILE	1
THR	1
ASP	1
SER	1
GLU	1
THR	1
TYR	1
GLU	1
LEU	1
P241	1
S242	1
E243	1
K244	1
V245	1
A248	1
L249	1
R250	1
E251	1
GLY	1
ASP	1
R254	1
K255	1
W256	1
T257	1
N258	1
S259	1
D260	1
P261	1
K262	1
E265	1
D266	1
W267	1
S271	1
R274	1
H278	1
L279	1
R280	1
K281	1
V282	1
C283	1
R284	1
K287	1
A288	1
V289	1
R290	1
D291	1
A292	1
Q293	1
L307	1
L322	1
G323	1
E324	1
K327	1
L333	1
E336	1
P344	1
D345	1
S246	1
T347	1
D348	1
E349	1
K350	1
P351	1
P354	1
P355	1
S356	1
H359	1
S378	1
E381	1
S382	1
A383	1
D384	1
S385	1
E388	1
N389	1
E405	1
V408	1
L409	1
ALA	1
LYS	1
ALA	1
LEU	1
PRO	1
ALA	1
PHE	1
ALA	1

Chain B:

4% 56% 21% 19%

GLY P107 R108 F109 W110 T111 G112 S113 F115 Q116 Y117 D118 L119 L120 R121 A122 P123 L124 F124 G127 Q128 E129 M130 M138 L139 M140 P141 I142 E146 P147 K148 A149 L152 L153 L154 I155 L156 A160 S161 L162 K163 K164 L165 K177 R182 I183 K184 L185 A186 A187 E188 F189 T190 H191 I192 M196 Y197 A198 I199 K200 K201 ASP GLU PHE GLN LYS GLN ILE ALA LEU GLU ALA ASP ARG SER PHE VAL LYS GLY ALA ILE PHE GLU SER TRP VAL ASP ASP SER GLU TRP TRP TRP GLU L240 D241 S242 E243 N244 I245 N246 E251 GLY ASP E252



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	74.75Å 46.27Å 121.79Å 90.00° 99.82° 90.00°	Depositor
Resolution (Å)	43.17 – 2.20 43.17 – 2.20	Depositor EDS
% Data completeness (in resolution range)	99.2 (43.17-2.20) 95.0 (43.17-2.20)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.71 (at 2.20Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.228 , 0.292 0.222 , 0.279	Depositor DCC
R_{free} test set	2019 reflections (5.04%)	DCC
Wilson B-factor (Å ²)	27.3	Xtriage
Anisotropy	0.153	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 42.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6104	wwPDB-VP
Average B, all atoms (Å ²)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.57 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 6.6827e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹ Intensities estimated from amplitudes.

² 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 [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	0.48	1/2961 (0.0%)	0.84	16/3991 (0.4%)
1	B	0.44	0/2979	0.96	21/4014 (0.5%)
All	All	0.46	1/5940 (0.0%)	0.90	37/8005 (0.5%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	107	PRO	N-CD	13.36	1.66	1.47

The worst 5 of 37 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	61	ALA	N-CA-CB	-16.81	86.56	110.10
1	A	141	PRO	N-CA-C	15.45	152.26	112.10
1	B	107	PRO	CB-CA-C	-13.04	79.39	112.00
1	B	60	SER	CB-CA-C	-12.65	86.06	110.10
1	B	3	MET	N-CA-CB	-12.64	87.85	110.60

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	2904	0	2899	137	0
1	B	2922	0	2926	105	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	A	126	0	0	6	0
2	B	152	0	0	8	0
All	All	6104	0	5825	241	1

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 21.

The worst 5 of 241 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:385:SER:HB3	1:A:388:GLU:OE1	1.22	1.29
1:A:245:VAL:HG12	1:A:259:SER:N	1.53	1.21
1:B:290:ARG:NH1	1:B:291:ASP:OD1	1.73	1.20
1:A:385:SER:CB	1:A:388:GLU:OE1	2.02	1.08
1:B:24:LEU:HD11	1:B:130:MET:HE1	1.38	1.04

All (1) 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)
1:B:85:ARG:NH2	1:B:127:GLY:O[2_444]	1.90	0.30

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	354/447 (79%)	340 (96%)	12 (3%)	2 (1%)	30	29
1	B	356/447 (80%)	345 (97%)	11 (3%)	0	100	100
All	All	710/894 (79%)	685 (96%)	23 (3%)	2 (0%)	46	50

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	118	ASP
1	A	67	ILE

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	320/391 (82%)	296 (92%)	24 (8%)	17	17
1	B	323/391 (83%)	297 (92%)	26 (8%)	15	15
All	All	643/782 (82%)	593 (92%)	50 (8%)	16	15

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	378	SER
1	B	42	ARG
1	B	254	ARG
1	A	384	ASP
1	B	18	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	244	ASN
1	A	246	ASN
1	A	359	HIS
1	B	244	ASN
1	B	359	HIS

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](#)

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	362/447 (80%)	0.46	32 (8%) 12 11	15, 31, 56, 68	0
1	B	364/447 (81%)	0.29	18 (4%) 33 33	14, 27, 47, 63	0
All	All	726/894 (81%)	0.38	50 (6%) 20 19	14, 29, 53, 68	0

The worst 5 of 50 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	244	ASN	9.1
1	A	245	VAL	8.2
1	A	244	ASN	7.1
1	B	3	MET	6.7
1	A	3	MET	6.0

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.