



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 12:36 am BST

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](mailto: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  
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

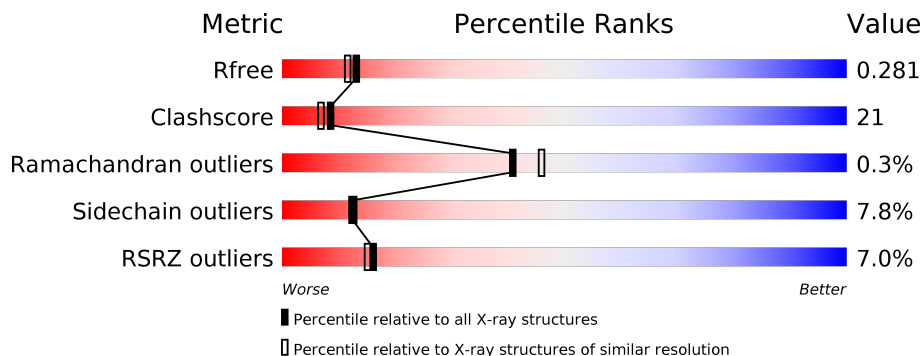
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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))
$R_{free}$	130704	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (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  $\geq 3$ , 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	447	 7% 50% 26% 5% 19%
1	B	447	 4% 56% 21% • 19%

## 2 Entry composition i

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
			Total	C	N	O	S			
1	A	362	2904	1837	503	548	16	0	0	0
1	B	364	2922	1849	506	551	16	0	0	0

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

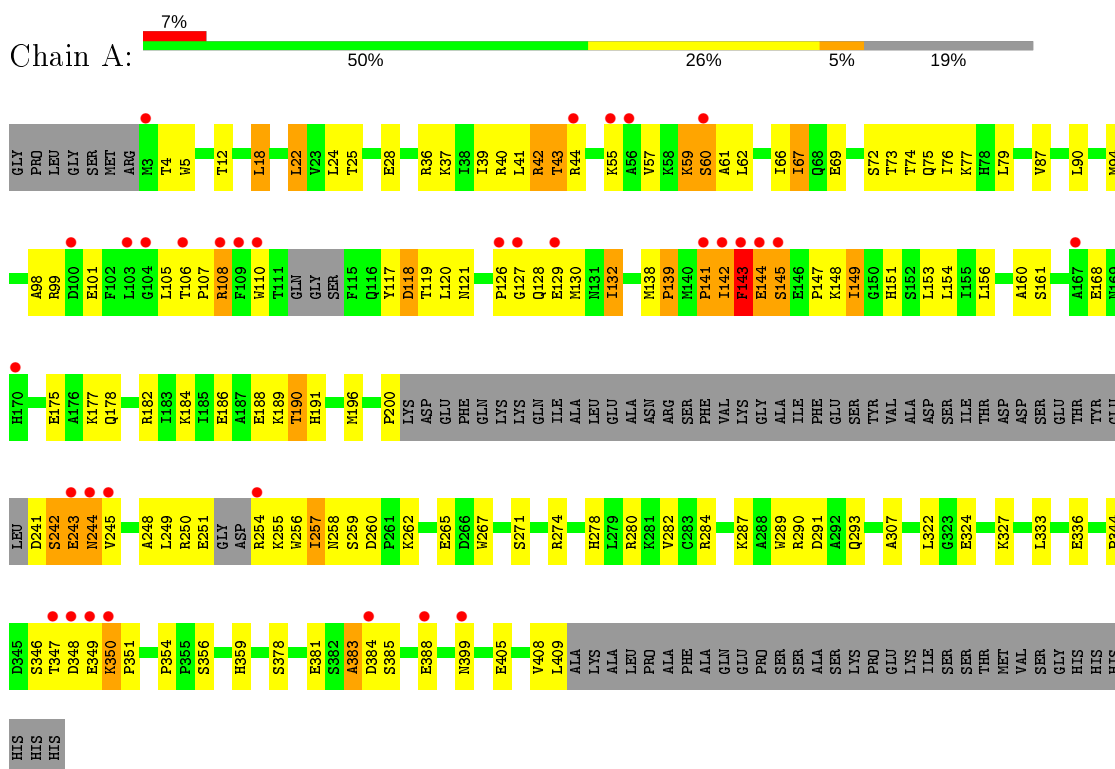
- Molecule 2 is water.

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

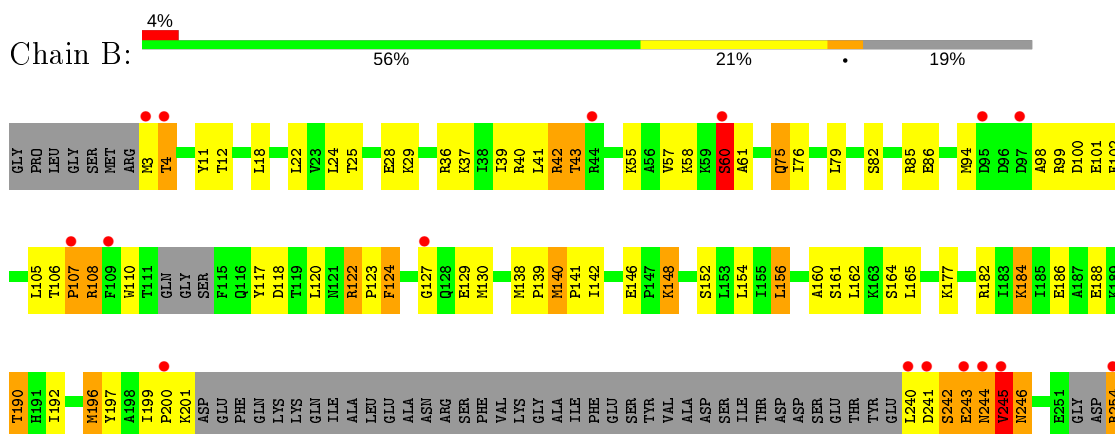
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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 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 ( $RSRZ > 2$ ). 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.

- Molecule 1: Cyclic GMP-AMP synthase



- Molecule 1: Cyclic GMP-AMP synthase



K255	ALA
M258	GLN
S259	PRO
D260	SER
P261	SER
K262	ALA
E265	SER
D266	LYS
K277	PRO
R284	GLU
W289	ILE
R290	SER
D291	THR
A292	MET
Q293	VAL
M305	SER
A306	GLY
A307	HIS
N310	HIS
L322	HIS
G323	HIS
K327	HIS
L333	HIS
E336	HIS
E342	HIS
D348	HIS
P354	HIS
H359	HIS
D384	HIS
S385	HIS
K386	HIS
S387	HIS
E388	HIS
R400	HIS
L409	HIS
ALA	ALA
LYS	LYS
ALA	ALA
LEU	LEU
PRO	PRO
ALA	ALA
PHE	PHE

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.24 (at 2.20Å)	Xtrriage
Refinement program	PHENIX 1.7.1_743	Depositor
R, $R_{free}$	0.228 , 0.292 0.225 , 0.281	Depositor DCC
$R_{free}$ test set	2118 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	26.8	Xtrriage
Anisotropy	0.144	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 42.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	6104	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 34.58 % 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 6.6676e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>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%)	25	26
1	B	356/447 (80%)	345 (97%)	11 (3%)	0	100	100
All	All	710/894 (79%)	685 (96%)	23 (3%)	2 (0%)	41	46

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%)	13	14
1	B	323/391 (83%)	297 (92%)	26 (8%)	12	12
All	All	643/782 (82%)	593 (92%)	50 (8%)	12	13

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q < 0.9
1	A	362/447 (80%)	0.46	33 (9%) 9 8	15, 31, 56, 68	0
1	B	364/447 (81%)	0.29	18 (4%) 29 28	14, 27, 47, 63	0
All	All	726/894 (81%)	0.38	51 (7%) 16 15	14, 29, 53, 68	0

The worst 5 of 51 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	244	ASN	9.0
1	A	245	VAL	8.1
1	A	244	ASN	7.1
1	B	3	MET	6.8
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.