



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:51 PM GMT

PDB ID : 1HI8  
Title : RNA DEPENDENT RNA POLYMERASE FROM DSRNA BACTERIO-  
PHAGE PHI6  
Authors : Grimes, J.M.; Butcher, S.J.; Makeyev, E.V.; Bamford, D.H.; Stuart, D.I.  
Deposited on : 2001-01-03  
Resolution : 2.50 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  
<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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
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) : trunk26865

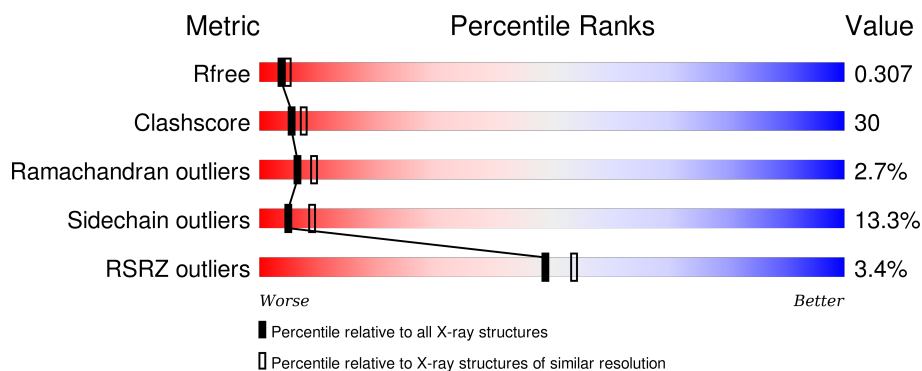
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.50 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	3553 (2.50-2.50)
Clashscore	102246	4242 (2.50-2.50)
Ramachandran outliers	100387	4156 (2.50-2.50)
Sidechain outliers	100360	4158 (2.50-2.50)
RSRZ outliers	91569	3562 (2.50-2.50)

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. 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	664	 2% 53% 38% 9%
1	B	664	 4% 52% 39% 9%

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	MG	A	665	-	-	-	X
2	MG	B	665	-	-	-	X

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 10532 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 P2 PROTEIN.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	664	Total	C	N	O	S	Se	0	0	0
			5265	3342	914	977	7	25			
1	B	664	Total	C	N	O	S	Se	0	0	0
			5265	3342	914	977	7	25			

There are 50 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	16	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	87	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	90	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	95	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	134	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	160	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	181	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	195	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	226	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	273	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	284	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	347	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	401	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	406	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	413	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	429	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	456	MSE	ILE	SEE REMARK 999	UNP P11124
A	474	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	486	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	521	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	554	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	590	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	605	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	646	MSE	MET	MODIFIED RESIDUE	UNP P11124
A	662	MSE	MET	MODIFIED RESIDUE	UNP P11124

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	16	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	87	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	90	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	95	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	134	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	160	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	181	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	195	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	226	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	273	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	284	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	347	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	401	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	406	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	413	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	429	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	456	MSE	ILE	SEE REMARK 999	UNP P11124
B	474	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	486	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	521	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	554	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	590	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	605	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	646	MSE	MET	MODIFIED RESIDUE	UNP P11124
B	662	MSE	MET	MODIFIED RESIDUE	UNP P11124

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Mg 1 1	0	0
2	A	1	Total Mg 1 1	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	110.00Å 110.00Å 159.30Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.50 47.63 – 2.50	Depositor EDS
% Data completeness (in resolution range)	88.0 (50.00-2.50) 88.2 (47.63-2.50)	Depositor EDS
$R_{merge}$	0.90	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.18 (at 2.51Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.280 , 0.316 0.275 , 0.307	Depositor DCC
$R_{free}$ test set	3339 reflections (5.07%)	DCC
Wilson B-factor (Å <sup>2</sup> )	41.2	Xtriage
Anisotropy	0.643	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 52.1	EDS
Estimated twinning fraction	0.006 for -h,-k,l 0.029 for h,-h-k,-l 0.016 for -k,-h,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 65827 reflections	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	10532	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.94% of the height of the origin peak. No significant pseudotranslation is detected.*

<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, 0.375 respectively for untwinned datasets, and 0.333, 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

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	0.74	0/5371	0.96	14/7222 (0.2%)
1	B	0.74	0/5371	0.96	14/7222 (0.2%)
All	All	0.74	0/10742	0.96	28/14444 (0.2%)

There are no bond length outliers.

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	376	HIS	N-CA-C	-8.18	88.93	111.00
1	A	376	HIS	N-CA-C	-8.17	88.93	111.00
1	B	467	GLY	N-CA-C	-7.80	93.61	113.10
1	A	467	GLY	N-CA-C	-7.79	93.64	113.10
1	A	626	ASN	N-CA-C	-6.95	92.23	111.00

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	5265	0	5165	309	0
1	B	5265	0	5165	311	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	10532	0	10330	620	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 30.

The worst 5 of 620 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:B:55:ASN:HD21	1:B:59:ARG:HD2	0.98	1.14
1:A:55:ASN:HD21	1:A:59:ARG:HD2	0.98	1.14
1:A:22:ILE:HD12	1:A:22:ILE:H	1.15	1.12
1:B:537:ARG:HD2	1:B:537:ARG:H	1.11	1.12
1:A:537:ARG:HD2	1:A:537:ARG:H	1.11	1.10

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	662/664 (100%)	586 (88%)	58 (9%)	18 (3%)	6	9
1	B	662/664 (100%)	586 (88%)	58 (9%)	18 (3%)	6	9
All	All	1324/1328 (100%)	1172 (88%)	116 (9%)	36 (3%)	6	9

5 of 36 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	91	ASN
1	A	149	SER
1	A	210	ALA
1	A	349	TYR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	375	GLY

### 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	557/532 (105%)	483 (87%)	74 (13%)	5	9
1	B	557/532 (105%)	483 (87%)	74 (13%)	5	9
All	All	1114/1064 (105%)	966 (87%)	148 (13%)	5	9

5 of 148 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	556	ASP
1	B	55	ASN
1	B	506	ARG
1	A	587	ARG
1	B	16	MSE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 44 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	533	GLN
1	B	55	ASN
1	B	525	GLN
1	A	608	GLN
1	B	21	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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 [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	639/664 (96%)	0.07	16 (2%) 61 65	18, 40, 74, 138	0
1	B	639/664 (96%)	0.16	27 (4%) 40 45	18, 40, 74, 138	0
All	All	1278/1328 (96%)	0.12	43 (3%) 49 54	18, 40, 74, 138	0

The worst 5 of 43 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	603	ALA	9.2
1	B	607	ARG	8.5
1	B	608	GLN	8.2
1	B	611	LEU	7.7
1	B	535	GLY	7.0

### 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. 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
2	MG	B	665	1/1	0.84	0.32	8.80	42,42,42,42	0
2	MG	A	665	1/1	0.87	0.41	7.72	42,42,42,42	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.