



Full wwPDB X-ray Structure Validation Report ⓘ

Sep 22, 2018 – 11:56 AM EDT

PDB ID : 6B1S
Title : Hydrogen Bonding Complementary, not size complementarity is key in the formation of the double helix
Authors : Singh, I.; Georgiadis, M.M.
Deposited on : 2017-09-18
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : **FAILED**
Mogul : 1.7.3 (157068), CSD as539be (2018)
Xtrriage (Phenix) : 1.13
EDS : rb-20031172
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031172

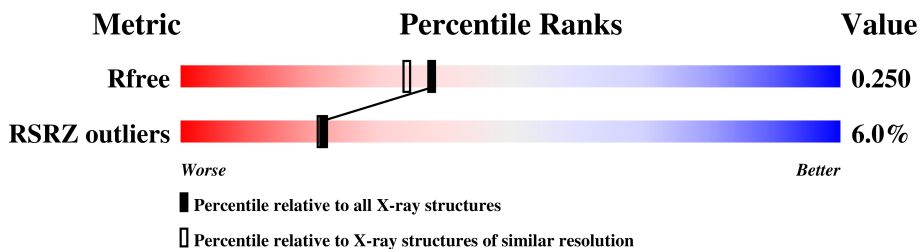
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	111664	7193 (2.00-2.00)
RSRZ outliers	108989	7011 (2.00-2.00)

MolProbity failed to run properly - the sequence quality summary graphics cannot be shown.

2 Entry composition i

There are 3 unique types of molecules in this entry. The entry contains 9147 atoms, of which 4291 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 Reverse transcriptase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	A	245	Total	C	H	N	O	S	0	0	0
			3938	1261	1978	343	350	6			
1	B	246	Total	C	H	N	O	S	0	0	0
			3964	1272	1987	344	354	7			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	GLY	-	expression tag	UNP P03355
A	21	SER	-	expression tag	UNP P03355
A	22	HIS	-	expression tag	UNP P03355
A	23	MET	-	expression tag	UNP P03355
B	20	GLY	-	expression tag	UNP P03355
B	21	SER	-	expression tag	UNP P03355
B	22	HIS	-	expression tag	UNP P03355
B	23	MET	-	expression tag	UNP P03355

- Molecule 2 is a DNA chain called DNA (5'-D(*CP*TP*TP*AP*TP*AP*(CGY)P*(CGY)P*TP*TP*TP*AP*TP*AP*AP*G)-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	P			
2	E	16	Total	C	H	N	O	P	0	0	0
			490	159	163	55	98	15			
2	F	16	Total	C	H	N	O	P	0	0	0
			490	159	163	55	98	15			

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	147	Total	O	0	0
			147	147		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	117	Total 117	O 117	0	0
3	E	1	Total 1	O 1	0	0

MolProbity failed to run properly - this section is therefore empty.

3 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	53.12Å 95.80Å 144.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.74 – 2.00 35.68 – 2.00	Depositor EDS
% Data completeness (in resolution range)	99.9 (14.74-2.00) 99.9 (35.68-2.00)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.11 (at 2.00Å)	Xtrriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.217 , 0.251 0.217 , 0.250	Depositor DCC
R_{free} test set	2468 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å ²)	27.8	Xtrriage
Anisotropy	0.130	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.44 , 51.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	9147	wwPDB-VP
Average B, all atoms (Å ²)	42.0	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 19.98 % 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 9.6163e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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.

4 Model quality [i](#)

4.1 Standard geometry [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.2 Too-close contacts [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.2 Protein sidechains [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.3.3 RNA [i](#)

MolProbity failed to run properly - this section is therefore empty.

4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

4 non-standard protein/DNA/RNA residues are modelled in this entry.

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	CGY	E	7	2	17,23,24	1.28	3 (17%)	22,33,36	1.74	5 (22%)
2	CGY	E	8	2	17,23,24	1.25	3 (17%)	22,33,36	1.38	3 (13%)
2	CGY	F	7	2	17,23,24	1.26	2 (11%)	22,33,36	1.57	5 (22%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	CGY	F	8	2	17,23,24	1.23	3 (17%)	22,33,36	2.00	6 (27%)

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	CGY	E	7	2	-	0/7/25/26	0/2/2/2
2	CGY	E	8	2	-	0/7/25/26	0/2/2/2
2	CGY	F	7	2	-	0/7/25/26	0/2/2/2
2	CGY	F	8	2	-	0/7/25/26	0/2/2/2

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	7	CGY	C2-N1	-2.76	1.27	1.34
2	E	7	CGY	C2-N1	-2.75	1.27	1.34
2	E	8	CGY	C2-N1	-2.74	1.27	1.34
2	F	8	CGY	C2-N1	-2.72	1.27	1.34
2	E	7	CGY	C4-N4	-2.70	1.27	1.34
2	F	7	CGY	C4-N4	-2.67	1.27	1.34
2	F	8	CGY	C4-N4	-2.67	1.27	1.34
2	E	8	CGY	C4-N4	-2.66	1.27	1.34
2	E	8	CGY	C5-C4	-2.21	1.38	1.41
2	F	8	CGY	C5-C4	-2.14	1.38	1.41
2	E	7	CGY	C5-C4	-2.09	1.38	1.41

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	8	CGY	C5-C4-N4	-2.86	118.23	123.07
2	E	7	CGY	C5-C4-N4	-2.82	118.30	123.07
2	F	7	CGY	C5-C4-N4	-2.79	118.36	123.07
2	E	8	CGY	C5-C4-N4	-2.77	118.39	123.07
2	F	7	CGY	C4-N3-C2	-2.62	116.89	119.80
2	F	7	CGY	C4'-O4'-C1'	-2.57	103.46	108.88
2	F	8	CGY	C4-N3-C2	-2.38	117.15	119.80
2	E	7	CGY	C4-N3-C2	-2.33	117.21	119.80
2	E	8	CGY	C4-N3-C2	-2.24	117.31	119.80
2	F	8	CGY	C2'-C1'-C1	-2.21	111.49	114.71
2	F	8	CGY	C6-C5-N5	2.01	121.12	118.65

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	8	CGY	O4'-C1'-C2'	2.23	107.10	103.66
2	E	7	CGY	C6-C5-N5	2.35	121.54	118.65
2	F	7	CGY	O4'-C1'-C1	2.76	112.60	109.76
2	F	7	CGY	C3'-C2'-C1'	2.76	105.29	102.69
2	E	7	CGY	O4'-C1'-C1	3.24	113.09	109.76
2	E	8	CGY	C3'-C2'-C1'	3.35	105.83	102.69
2	E	7	CGY	C3'-C2'-C1'	3.92	106.37	102.69
2	F	8	CGY	C3'-C2'-C1'	6.53	108.82	102.69

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

4.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

4.6 Ligand geometry [i](#)

There are no ligands in this entry.

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data

5.1 Protein, DNA and RNA chains

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	245/259 (94%)	0.61	10 (4%) 37 37	16, 31, 60, 84	0
1	B	246/259 (94%)	0.51	12 (4%) 29 29	18, 34, 62, 87	0
2	E	14/16 (87%)	1.43	6 (42%) 0 0	35, 56, 74, 83	0
2	F	14/16 (87%)	1.34	3 (21%) 1 0	37, 57, 77, 81	0
All	All	519/550 (94%)	0.61	31 (5%) 22 21	16, 34, 65, 87	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	234	LEU	5.3
1	A	100	PRO	4.2
1	B	231	THR	3.6
2	E	13	DT	3.5
1	B	26	LEU	3.4
1	A	179	ILE	3.3
2	E	12	DA	3.3
1	B	264	LYS	3.2
1	B	101	VAL	3.2
1	B	110	ARG	3.2
1	A	64	TYR	3.2
1	B	145	TRP	3.1
1	B	99	LEU	3.1
2	F	13	DT	2.9
1	A	274	LYS	2.8
1	A	145	TRP	2.8
1	B	29	PHE	2.7
2	F	9	DT	2.7
2	E	11	DT	2.6
1	B	64	TYR	2.6
1	A	101	VAL	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	E	14	DA	2.5
1	B	100	PRO	2.4
2	F	12	DA	2.4
1	B	109	TYR	2.2
2	E	9	DT	2.2
1	B	62	LYS	2.2
1	A	99	LEU	2.1
2	E	10	DT	2.1
1	A	163	THR	2.0
1	A	25	TRP	2.0

5.2 Non-standard residues in protein, DNA, RNA chains [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. 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	B-factors(Å ²)	Q<0.9
2	CGY	E	8	22/23	0.77	0.23	62,68,74,76	0
2	CGY	F	7	22/23	0.78	0.29	57,72,82,90	0
2	CGY	F	8	22/23	0.79	0.27	58,69,83,86	0
2	CGY	E	7	22/23	0.80	0.26	52,65,74,75	0

5.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.4 Ligands [i](#)

There are no ligands in this entry.

5.5 Other polymers [i](#)

There are no such residues in this entry.