



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 05:31 AM GMT

PDB ID : 2R2T
Title : d(ATTAGTTAACTAAAT) complexed with MMLV RT catalytic fragment
Authors : Goodwin, K.D.; Lewis, M.A.; Long, E.C.; Georgiadis, M.M.
Deposited on : 2007-08-27
Resolution : 2.00 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 (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

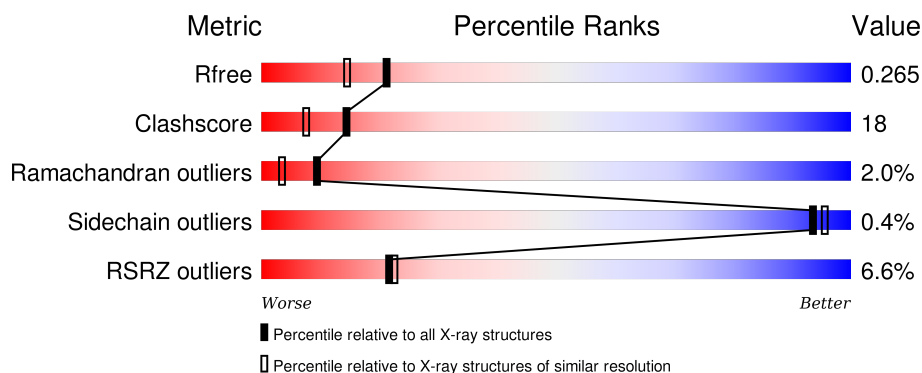
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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}	91344	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	B	8	<div style="width: 100%; background-color: yellow;"></div> 100%
2	G	8	<div style="width: 100%; background-color: green; position: relative;"> 13% 88% </div>
3	A	255	<div style="width: 100%; background-color: green; position: relative;"> 7% 75% 24% • </div>

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 2542 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 DNA chain called DNA (5'-D(*DAP*DTP*DTP*DTP*DAP*DGP*DTP*D T)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	8	Total	C	N	O	P	0	0	0
			161	80	25	49	7			

- Molecule 2 is a DNA chain called DNA (5'-D(P*DAP*DAP*DCP*DTP*DAP*DAP*DAP*DT)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	G	8	Total	C	N	O	P	0	0	0
			164	79	32	45	8			

- Molecule 3 is a protein called Reverse transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	255	Total	C	N	O	S	0	0	0
			2041	1311	356	367	7			

- Molecule 4 is water.

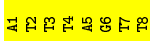
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	171	Total	O	0	0
			171	171		
4	B	2	Total	O	0	0
			2	2		
4	G	3	Total	O	0	0
			3	3		

3 Residue-property plots

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 (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.

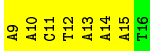
- Molecule 1: DNA (5'-D(*DAP*DTP*DTP*DTP*DAP*DGP*DTP*DT)-3')

Chain B: 



A1 T2 T3 T4 A5 G6 T7 T8

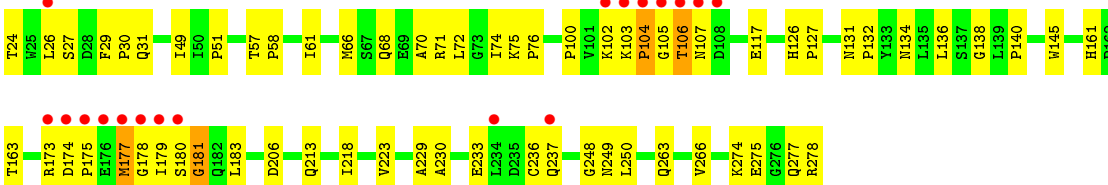
- Molecule 2: DNA (5'-D(P*DAP*DAP*DCP*DTP*DAP*DAP*DAP*DT)-3')

Chain G: 


A9 A10 C11 T12 A13 A14 A15 T16

- Molecule 3: Reverse transcriptase

Chain A: 


T24 W25 L26 S27 D28 F29 P30 Q31 I49 I50 P51 T57 P58 I61 M66 S67 Q68 E69 A70 R71 L72 G73 I74 K75 P76 P100 Y101 K102 K103 P104 G105 T106 M107 D108 E117 H126 P127 M131 P132 Y133 M134 L135 L136 S137 G138 L139 P140 W145 H161 P162
T163 R173 D174 P175 E176 M177 G178 I179 S180 G181 Q182 L183 D206 Q213 I218 V223 A229 A230 E233 L234 D235 C236 Q237 G248 N249 L250 Q263 V266 K274 E275 G276 Q277 R278

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	54.64Å 145.80Å 46.88Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.00 43.72 – 2.00	Depositor EDS
% Data completeness (in resolution range)	92.3 (50.00-2.00) 92.4 (43.72-2.00)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.42 (at 2.00Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.224 , 0.264 0.225 , 0.265	Depositor DCC
R_{free} test set	1166 reflections (4.84%)	DCC
Wilson B-factor (Å ²)	29.1	Xtriage
Anisotropy	0.276	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 49.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 25116 reflections (0.004%)	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	2542	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.18% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	B	0.23	0/179	0.64	0/275
2	G	0.23	0/184	0.68	0/281
3	A	0.32	0/2097	0.59	0/2858
All	All	0.30	0/2460	0.60	0/3414

There are no bond length outliers.

There are no bond angle outliers.

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	B	161	0	95	17	1
2	G	164	0	91	16	1
3	A	2041	0	2056	50	0
4	A	171	0	0	3	0
4	B	2	0	0	0	0
4	G	3	0	0	0	0
All	All	2542	0	2242	83	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 18.

The worst 5 of 83 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)
2:G:14:DA:H2''	2:G:15:DA:H5'	1.40	1.03
3:A:233:GLU:HG2	3:A:237:GLN:HE21	1.32	0.93
1:B:3:DT:H2''	1:B:4:DT:H5'	1.51	0.91
1:B:5:DA:H2''	1:B:6:DG:H5''	1.60	0.83
2:G:12:DT:H2''	2:G:13:DA:H5'	1.67	0.77

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:8:DT:O3'	2:G:9:DA:P[2_765]	1.61	0.59

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
3	A	253/255 (99%)	241 (95%)	7 (3%)	5 (2%)	9 3

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	106	THR
3	A	223	VAL
3	A	104	PRO
3	A	177	MET
3	A	181	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	
3	A	224/224 (100%)	223 (100%)	1 (0%)	93	95

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

Mol	Chain	Res	Type
3	A	177	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	161	HIS
3	A	213	GLN
3	A	238	GLN
3	A	144	GLN
3	A	237	GLN

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 ⓘ

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	B	8/8 (100%)	0.48	0	100	100	46, 61, 75, 79	0
2	G	8/8 (100%)	0.79	0	100	100	33, 74, 85, 85	0
3	A	255/255 (100%)	0.56	18 (7%)	19	20	18, 32, 67, 85	0
All	All	271/271 (100%)	0.56	18 (6%)	22	22	18, 33, 74, 85	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	106	THR	10.5
3	A	105	GLY	9.7
3	A	179	ILE	8.2
3	A	178	GLY	7.3
3	A	180	SER	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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.