



# Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 09:21 PM GMT

PDB ID : 3NYB  
Title : Structure and function of the polymerase core of TRAMP, a RNA surveillance complex  
Authors : Reinisch, K.M.; Hamill, S.  
Deposited on : 2010-07-14  
Resolution : 2.70 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

---

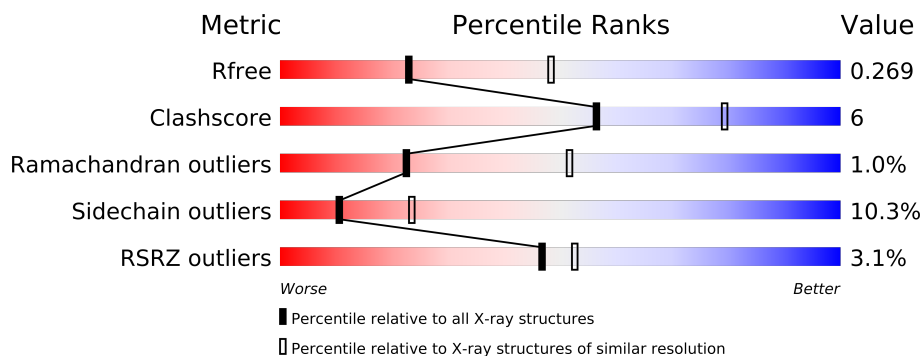
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.15 2013  
Xtriage (Phenix) : dev-1323  
EDS : stable22639  
Percentile statistics : 21963  
Refmac : 5.8.0049  
CCP4 : 6.3.0 (Settle)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)  
Validation Pipeline (wwPDB-VP) : stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.70 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))
$R_{free}$	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	323	
2	B	83	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 3139 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Poly(A) RNA polymerase protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	323	Total	C	N	O	S	0	0	0
			2577	1658	439	476	4			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	159	GLY	-	EXPRESSION TAG	UNP P53632
A	160	SER	-	EXPRESSION TAG	UNP P53632
A	293	ALA	ASP	ENGINEERED MUTATION	UNP P53632

- Molecule 2 is a protein called Protein AIR2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	64	Total	C	N	O	S	0	0	0
			502	309	94	93	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	116	GLY	-	EXPRESSION TAG	UNP Q12476
B	117	SER	-	EXPRESSION TAG	UNP Q12476

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	2	Total	Zn	0	0
			2	2		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	56	Total 56	O 56	0	0
4	B	2	Total 2	O 2	0	0

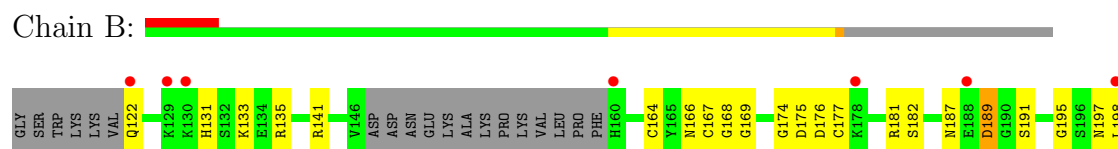
### 3 Residue-property plots

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: Poly(A) RNA polymerase protein 2



- Molecule 2: Protein AIR2



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.72Å 92.72Å 103.63Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.66 – 2.70 37.44 – 2.61	Depositor EDS
% Data completeness (in resolution range)	96.8 (24.66-2.70) 95.5 (37.44-2.61)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.24 (at 2.61Å)	Xtriage
Refinement program	CNS1.2 and PHENIX (phenix.refine: 1.5.2)	Depositor
R, $R_{free}$	0.194 , 0.258 0.210 , 0.269	Depositor DCC
$R_{free}$ test set	979 reflections (6.94%)	DCC
Wilson B-factor (Å <sup>2</sup> )	52.2	Xtriage
Anisotropy	0.004	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 37.6	EDS
Estimated twinning fraction	0.031 for -h,-k,l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 16163 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	3139	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.29% of the height of the origin peak. No significant pseudotranslation is detected.*

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: ZN

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	0.36	0/2640	0.55	0/3577
2	B	0.30	0/512	0.50	0/683
All	All	0.35	0/3152	0.54	0/4260

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Close contacts

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2577	0	0	14	0
2	B	502	0	0	8	0
3	B	2	0	0	0	0
4	A	56	0	0	2	0
4	B	2	0	0	0	0
All	All	3139	0	0	20	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 6.

All (20) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:368:PRO:O	1:A:372:LEU:CD1	2.29	0.80
1:A:329:LEU:O	1:A:332:ARG:O	2.11	0.68
1:A:367:ASP:N	1:A:367:ASP:OD1	2.36	0.59
2:B:131:HIS:CE1	2:B:189:ASP:OD2	2.61	0.53
1:A:187:VAL:O	1:A:191:SER:OG	2.27	0.53
1:A:386:ASN:ND2	2:B:191:SER:OG	2.43	0.51
1:A:178:TRP:CH2	1:A:472:LYS:O	2.63	0.51
2:B:164:CYS:SG	2:B:174:GLY:CA	3.01	0.49
2:B:166:ASN:ND2	2:B:182:SER:N	2.61	0.48
1:A:432:ASN:OD1	1:A:432:ASN:O	2.32	0.48
1:A:178:TRP:CZ2	4:A:41:HOH:O	2.68	0.47
2:B:167:CYS:O	2:B:169:GLY:N	2.48	0.46
1:A:381:GLU:OE2	2:B:141:ARG:NH2	2.50	0.45
1:A:243:SER:O	1:A:245:LEU:N	2.50	0.44
1:A:268:THR:OG1	1:A:269:GLU:N	2.51	0.44
1:A:358:HIS:O	1:A:361:ILE:N	2.52	0.43
2:B:187:ASN:CG	2:B:187:ASN:O	2.56	0.43
1:A:466:PHE:CE2	1:A:469:ARG:NH1	2.87	0.42
2:B:195:GLY:C	2:B:197:ASN:N	2.73	0.42
1:A:178:TRP:NE1	4:A:41:HOH:O	2.37	0.41

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	321/323 (99%)	299 (93%)	20 (6%)	2 (1%)	33	66
2	B	60/83 (72%)	50 (83%)	8 (13%)	2 (3%)	6	13
All	All	381/406 (94%)	349 (92%)	28 (7%)	4 (1%)	22	51

All (4) Ramachandran outliers are listed below:



Mol	Chain	Res	Type
1	A	244	GLU
1	A	276	ALA
2	B	135	ARG
2	B	168	GLY

### 5.3.2 Protein sidechains ⓘ

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	283/283 (100%)	256 (90%)	27 (10%)	12	28
2	B	56/73 (77%)	48 (86%)	8 (14%)	5	12
All	All	339/356 (95%)	304 (90%)	35 (10%)	10	23

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	178	TRP
1	A	196	GLU
1	A	237	ILE
1	A	240	VAL
1	A	243	SER
1	A	259	SER
1	A	261	LEU
1	A	281	ILE
1	A	300	ASN
1	A	302	ILE
1	A	337	VAL
1	A	355	LEU
1	A	356	HIS
1	A	362	ILE
1	A	367	ASP
1	A	374	VAL
1	A	376	LEU
1	A	396	SER
1	A	413	GLN
1	A	421	LEU
1	A	430	SER
1	A	431	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	433	ILE
1	A	434	SER
1	A	439	ASN
1	A	445	LYS
1	A	480	LYS
2	B	122	GLN
2	B	133	LYS
2	B	175	ASP
2	B	176	ASP
2	B	177	CYS
2	B	181	ARG
2	B	189	ASP
2	B	198	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

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.

## 5.7 Other polymers ⓘ

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 ⓘ

### 6.1 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, 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	323/323 (100%)	-0.35	5 (1%) 70 75	23, 50, 95, 127	0
2	B	64/83 (77%)	0.49	7 (10%) 6 6	55, 83, 124, 140	0
All	All	387/406 (95%)	-0.21	12 (3%) 47 52	23, 56, 109, 140	0

All (12) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	122	GLN	4.2
1	A	159	GLY	3.9
2	B	130	LYS	3.5
2	B	178	LYS	3.5
2	B	129	LYS	3.3
2	B	188	GLU	2.8
1	A	466	PHE	2.8
1	A	411	ALA	2.4
1	A	410	SER	2.3
2	B	160	HIS	2.1
2	B	198	LEU	2.1
1	A	248	LYS	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains ⓘ

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands

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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	ZN	B	2148	1/1	0.05	-1.27	82,82,82,82	0
3	ZN	B	2147	1/1	0.07	-1.85	121,121,121,121	0

## 6.5 Other polymers

There are no such residues in this entry.