



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 28, 2014 – 10:51 PM GMT

PDB ID : 3TBI
Title : Crystal structure of T4 gp33 bound to E. coli RNAP beta-flap domain
Authors : Twist, K.A.F.; Campbell, E.A.; Darst, S.A.
Deposited on : 2011-08-06
Resolution : 3.00 Å(reported)

This is a full wwPDB validation 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/ValidationPDFNotes.html>

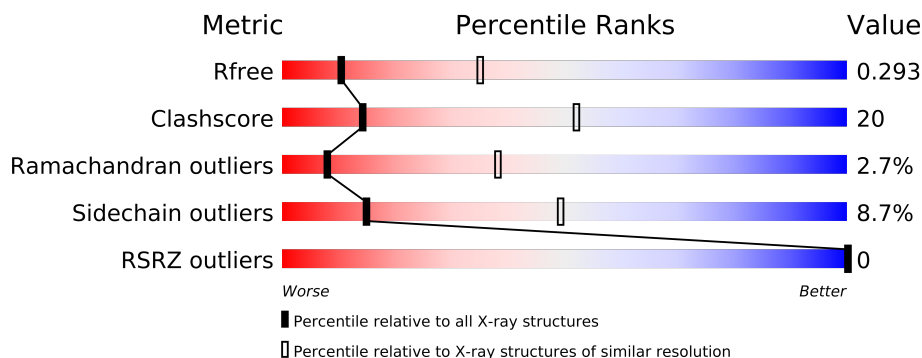
The following versions of software and data (see [references](#)) were used in the production of this report:

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 3.00 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}	66092	1216 (3.00-3.00)
Clashscore	79885	1594 (3.00-3.00)
Ramachandran outliers	78287	1537 (3.00-3.00)
Sidechain outliers	78261	1540 (3.00-3.00)
RSRZ outliers	66119	1217 (3.00-3.00)

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. 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	115	
2	B	228	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2331 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 RNA polymerase-associated protein Gp33.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	71	Total	C	N	O	S	Se	0	0	0
			558	356	89	110	1	2			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP P13338
A	-1	PRO	-	EXPRESSION TAG	UNP P13338
A	0	HIS	-	EXPRESSION TAG	UNP P13338

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	B	228	Total	C	N	O	S	Se	3	0	0
			1773	1115	303	352	1	2			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	830	MSE	-	EXPRESSION TAG	UNP P0A8V2

4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α , β , γ	55.22Å 112.25Å 164.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.00 – 3.00 29.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.9 (29.00-3.00) 99.2 (29.00-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.90 (at 3.00Å)	Xtriage
Refinement program	REFMAC 5.6.0107	Depositor
R, R_{free}	0.251 , 0.295 0.257 , 0.293	Depositor DCC
R_{free} test set	500 reflections (4.98%)	DCC
Wilson B-factor (Å ²)	69.4	Xtriage
Anisotropy	0.035	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.28 , 53.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 10551 reflections	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	2331	wwPDB-VP
Average B, all atoms (Å ²)	112.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 8.22% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.28	0/564	0.44	0/757
2	B	0.36	1/1789 (0.1%)	0.43	0/2404
All	All	0.34	1/2353 (0.0%)	0.43	0/3161

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	954	LYS	CG-CD	-6.22	1.31	1.52

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	558	0	17	17	1
2	B	1773	0	0	30	0
All	All	2331	0	17	46	1

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 20.

All (46) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:43:ILE:O	1:A:46:LYS:CG	2.23	0.87
2:B:949:GLU:O	2:B:953:LEU:CD2	2.34	0.76
2:B:915:ASP:OD2	2:B:919:ARG:NH2	2.19	0.75
2:B:975:ILE:CD1	2:B:997:TRP:CE3	2.74	0.70
1:A:41:GLY:O	1:A:44:ILE:CG2	2.40	0.68
2:B:896:THR:CG2	2:B:899:GLU:CD	2.62	0.68
2:B:886:LYS:NZ	2:B:916:SER:OG	2.33	0.61
2:B:1030:GLU:OE2	2:B:1033:ARG:NH2	2.33	0.61
2:B:892:GLU:O	2:B:894:GLN:N	2.34	0.61
1:A:55:TYR:O	1:A:59:THR:CG2	2.49	0.61
2:B:851:THR:OG1	2:B:853:ASP:N	2.34	0.60
2:B:985:GLU:O	2:B:986:ALA:C	2.41	0.58
2:B:915:ASP:OD1	2:B:915:ASP:C	2.42	0.58
2:B:896:THR:CG2	2:B:899:GLU:CG	2.83	0.57
2:B:864:LYS:CA	2:B:864:LYS:CE	2.84	0.55
1:A:66:ASN:CB	1:A:68:ILE:CD1	2.86	0.54
1:A:66:ASN:N	1:A:66:ASN:ND2	2.57	0.52
2:B:946:LEU:O	2:B:950:GLU:N	2.42	0.52
1:A:38:GLN:O	1:A:42:PHE:CD2	2.64	0.51
1:A:44:ILE:CG2	1:A:45:GLU:N	2.74	0.51
2:B:1003:THR:O	2:B:1004:ASP:C	2.48	0.51
1:A:82:ILE:O	1:A:85:ILE:N	2.44	0.51
1:A:92:GLU:CB	1:A:94:LEU:CD2	2.89	0.50
2:B:955:GLN:O	2:B:956:ALA:C	2.50	0.50
1:A:37:ARG:O	1:A:38:GLN:C	2.49	0.50
2:B:892:GLU:O	2:B:893:THR:C	2.51	0.49
1:A:98:SER:CB	2:B:853:ASP:OD2	2.61	0.49
2:B:997:TRP:N	2:B:997:TRP:CD1	2.84	0.46
2:B:892:GLU:O	2:B:895:LEU:N	2.49	0.45
1:A:35:LEU:O	1:A:36:ASP:C	2.55	0.45
1:A:35:LEU:O	1:A:37:ARG:N	2.50	0.44
1:A:37:ARG:O	1:A:41:GLY:N	2.51	0.44
2:B:975:ILE:O	2:B:979:LEU:N	2.51	0.44
2:B:864:LYS:N	2:B:864:LYS:CE	2.81	0.44
2:B:891:GLY:O	2:B:892:GLU:C	2.56	0.43
2:B:975:ILE:CD1	2:B:976:ARG:N	2.81	0.43
1:A:54:SER:O	1:A:55:TYR:C	2.57	0.42
2:B:888:THR:O	2:B:914:LYS:N	2.53	0.42
2:B:976:ARG:O	2:B:977:ALA:C	2.56	0.42
2:B:975:ILE:CG1	2:B:976:ARG:N	2.82	0.42
2:B:985:GLU:O	2:B:988:LYS:N	2.53	0.41
2:B:996:ARG:NH1	2:B:999:GLU:CG	2.83	0.41
2:B:955:GLN:O	2:B:957:LYS:N	2.53	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:44:ILE:O	1:A:47:MSE:N	2.54	0.41
2:B:868:SER:OG	2:B:942:ASP:OD1	2.39	0.41
1:A:34:LEU:O	1:A:35:LEU:C	2.60	0.40

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	Distance(Å)	Clash(Å)
1:A:67:SER:OG	1:A:67:SER:OG[2_455]	1.84	0.36

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	69/115 (60%)	56 (81%)	10 (14%)	3 (4%)	4	23
2	B	226/228 (99%)	190 (84%)	31 (14%)	5 (2%)	10	45
All	All	295/343 (86%)	246 (83%)	41 (14%)	8 (3%)	8	38

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	893	THR
2	B	986	ALA
1	A	36	ASP
2	B	892	GLU
1	A	35	LEU
2	B	1005	GLU
1	A	58	ALA
2	B	1004	ASP

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	61/102 (60%)	56 (92%)	5 (8%)	17	52
2	B	193/191 (101%)	176 (91%)	17 (9%)	14	48
All	All	254/293 (87%)	232 (91%)	22 (9%)	15	49

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	ASP
1	A	56	LEU
1	A	59	THR
1	A	66	ASN
1	A	70	GLU
2	B	830	MSE
2	B	836	LEU
2	B	851	THR
2	B	864	LYS
2	B	868	SER
2	B	876	GLU
2	B	887	VAL
2	B	896	THR
2	B	908	GLU
2	B	951	MSE
2	B	967	LEU
2	B	992	LEU
2	B	1004	ASP
2	B	1006	GLU
2	B	1029	LEU
2	B	1038	GLN
2	B	1046	VAL

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 ⓘ

There are no ligands in this entry.

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, 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	71/115 (61%)	-0.19	0 100 100	62, 121, 208, 230	0
2	B	228/228 (100%)	-0.25	0 100 100	50, 98, 195, 230	1 (0%)
All	All	299/343 (87%)	-0.23	0 100 100	50, 103, 199, 230	1 (0%)

There are no RSRZ outliers to report.

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.