



wwPDB X-ray Structure Validation Summary Report ⓘ

May 22, 2020 – 03:18 pm BST

PDB ID : 2ATW
Title : Structure of a Mycobacterium tuberculosis NusA-RNA complex
Authors : Beuth, B.; Pennell, S.; Arnvig, K.B.; Martin, S.R.; Taylor, I.A.
Deposited on : 2005-08-26
Resolution : 2.25 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

<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 : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

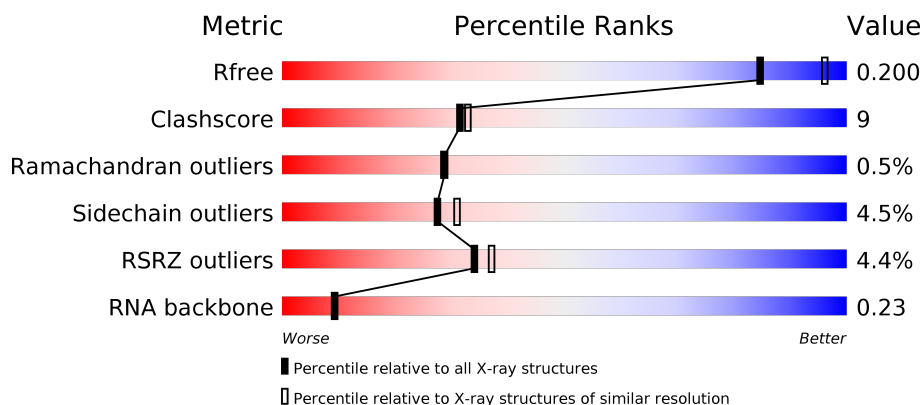
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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}	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)
RNA backbone	3102	1016 (2.66-1.86)

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 respectively. 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	12	
1	D	12	
2	A	251	
2	C	251	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 4083 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 RNA chain called ribosomal RNA (5'- AGAACUCAAUAG -3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	12	Total	C	N	O	P	0	0	0
			255	116	50	78	11			
1	D	12	Total	C	N	O	P	0	0	0
			255	116	50	78	11			

- Molecule 2 is a protein called Transcription elongation protein nusA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	223	Total	C	N	O	S	0	0	0
			1676	1037	320	315	4			
2	C	225	Total	C	N	O	S	0	0	0
			1689	1044	322	319	4			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	348	LEU	-	CLONING ARTIFACT	UNP P0A5M2
A	349	GLU	-	CLONING ARTIFACT	UNP P0A5M2
A	350	HIS	-	EXPRESSION TAG	UNP P0A5M2
A	351	HIS	-	EXPRESSION TAG	UNP P0A5M2
A	352	HIS	-	EXPRESSION TAG	UNP P0A5M2
A	353	HIS	-	EXPRESSION TAG	UNP P0A5M2
A	354	HIS	-	EXPRESSION TAG	UNP P0A5M2
A	355	HIS	-	EXPRESSION TAG	UNP P0A5M2
C	348	LEU	-	CLONING ARTIFACT	UNP P0A5M2
C	349	GLU	-	CLONING ARTIFACT	UNP P0A5M2
C	350	HIS	-	EXPRESSION TAG	UNP P0A5M2
C	351	HIS	-	EXPRESSION TAG	UNP P0A5M2
C	352	HIS	-	EXPRESSION TAG	UNP P0A5M2
C	353	HIS	-	EXPRESSION TAG	UNP P0A5M2
C	354	HIS	-	EXPRESSION TAG	UNP P0A5M2
C	355	HIS	-	EXPRESSION TAG	UNP P0A5M2

- Molecule 3 is water.

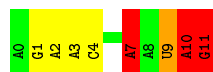
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	11	Total 11	O 11	0	0
3	D	5	Total 5	O 5	0	0
3	A	102	Total 102	O 102	0	0
3	C	90	Total 90	O 90	0	0

3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes 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: ribosomal RNA (5'- AGAACUCAAUAG -3')

Chain B: 



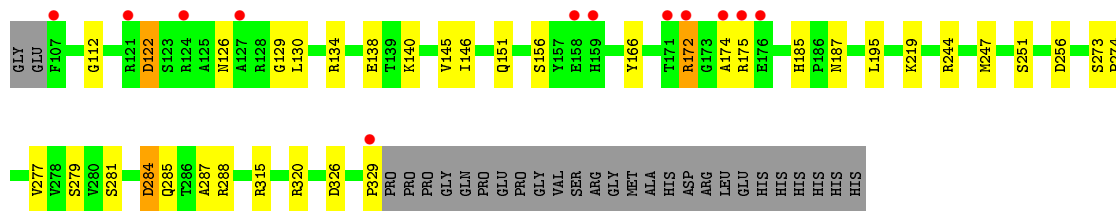
- Molecule 1: ribosomal RNA (5'- AGAACUCAAUAG -3')

Chain D: 




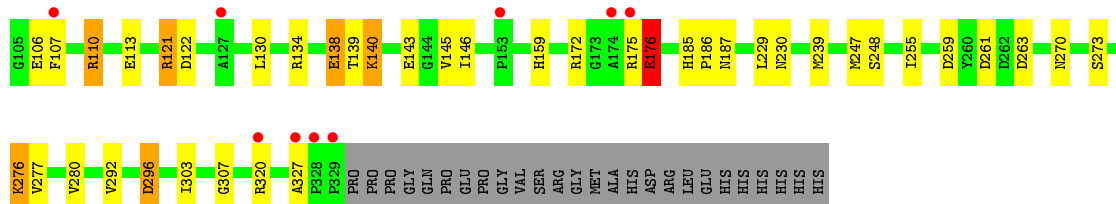
- Molecule 2: Transcription elongation protein nusA

Chain A: 



- Molecule 2: Transcription elongation protein nusA

Chain C: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	64.56 Å 89.54 Å 100.89 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.00 – 2.25 14.99 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.4 (15.00-2.25) 99.4 (14.99-2.25)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.28 (at 2.24 Å)	Xtriage
Refinement program	REFMAC 5.0	Depositor
R, R_{free}	0.206 , 0.273 0.210 , 0.200	Depositor DCC
R_{free} test set	1402 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å ²)	32.7	Xtriage
Anisotropy	0.014	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 35.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	4083	wwPDB-VP
Average B, all atoms (Å ²)	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.41% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	1.42	2/286 (0.7%)	2.08	11/444 (2.5%)
1	D	1.31	2/286 (0.7%)	2.01	9/444 (2.0%)
2	A	0.77	0/1698	0.91	3/2300 (0.1%)
2	C	0.70	0/1711	0.94	5/2317 (0.2%)
All	All	0.86	4/3981 (0.1%)	1.18	28/5505 (0.5%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	3	A	N7-C5	-6.42	1.35	1.39
1	D	10	A	C5-C6	-5.63	1.35	1.41
1	B	1	G	C2-N3	5.54	1.37	1.32
1	D	0	A	N7-C5	-5.46	1.35	1.39

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	10	A	O4'-C1'-N9	-10.65	99.68	108.20
1	B	2	A	O4'-C1'-N9	9.98	116.19	108.20
1	D	2	A	O4'-C1'-N9	7.68	114.34	108.20
1	B	3	A	O4'-C1'-N9	-7.34	102.32	108.20
2	C	263	ASP	CB-CG-OD2	7.09	124.68	118.30

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	255	0	132	4	0
1	D	255	0	132	2	0
2	A	1676	0	1717	30	1
2	C	1689	0	1726	32	0
3	A	102	0	0	14	0
3	B	11	0	0	1	0
3	C	90	0	0	12	1
3	D	5	0	0	0	0
All	All	4083	0	3707	68	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 9.

The worst 5 of 68 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:A:329:PRO:HD2	3:A:2021:HOH:O	1.61	0.99
2:A:185:HIS:HD2	2:A:187:ASN:H	1.11	0.94
2:A:138:GLU:HG3	3:A:2039:HOH:O	1.68	0.93
2:C:185:HIS:HD2	2:C:187:ASN:H	1.15	0.90
2:C:113:GLU:HG3	3:C:1035:HOH:O	1.73	0.88

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)
2:A:140:LYS:NZ	3:C:1128:HOH:O[1_554]	2.18	0.02

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	
2	A	221/251 (88%)	210 (95%)	10 (4%)	1 (0%)	29	29
2	C	223/251 (89%)	216 (97%)	6 (3%)	1 (0%)	34	37
All	All	444/502 (88%)	426 (96%)	16 (4%)	2 (0%)	29	29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	174	ALA
2	C	176	GLU

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	
2	A	176/199 (88%)	170 (97%)	6 (3%)	37	45
2	C	177/199 (89%)	167 (94%)	10 (6%)	21	21
All	All	353/398 (89%)	337 (96%)	16 (4%)	27	31

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	121	ARG
2	C	138	GLU
2	C	239	MET
2	C	110	ARG
2	C	248	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
2	A	187	ASN
2	C	285	GLN
2	C	185	HIS
2	A	185	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	161	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	B	11/12 (91%)	3 (27%)	0
1	D	11/12 (91%)	3 (27%)	2 (18%)
All	All	22/24 (91%)	6 (27%)	2 (9%)

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	B	7	A
1	B	10	A
1	B	11	G
1	D	7	A
1	D	10	A

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	D	9	U
1	D	10	A

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 [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, 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(Å²)	Q<0.9
1	B	12/12 (100%)	-0.45	0	100	100	30, 32, 37, 38	0
1	D	12/12 (100%)	-0.35	0	100	100	28, 36, 42, 50	0
2	A	223/251 (88%)	0.12	12 (5%)	25	28	19, 28, 51, 64	0
2	C	225/251 (89%)	0.17	9 (4%)	38	40	21, 31, 50, 70	0
All	All	472/526 (89%)	0.12	21 (4%)	34	37	19, 30, 51, 70	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	C	329	PRO	9.8
2	C	175	ARG	7.3
2	C	328	PRO	7.2
2	A	329	PRO	5.5
2	C	327	ALA	5.2

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.