Full wwPDB X-ray Structure Validation Report ⓘ

Oct 1, 2021 – 03:12 am BST

PDB ID	:	7NLH
Title	:	S. cerevisiae Ty1 p22 restriction factor, Gag CA-CTD, AUG1 variant
Authors	:	Cottee, M.A.; Taylor, I.A.
Deposited on	:	2021-02-22
Resolution	:	2.80 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:

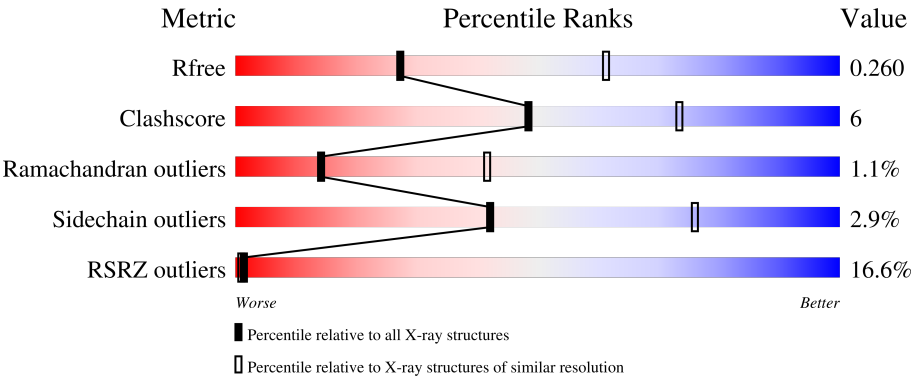
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 of 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 <=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	A	116	11% 66% 11% • 22%
1	B	116	8% 72% 6% • 22%
1	C	116	20% 64% 13% • 22%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2189 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 protein called Ty1 Gag p22.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	90	Total	C	N	O	S	0	0	0
			724	451	130	140	3			
1	B	91	Total	C	N	O	S	0	1	0
			741	464	132	142	3			
1	C	90	Total	C	N	O	S	0	0	0
			724	451	130	140	3			

There are 27 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	356	PRO	-	expression tag	UNP P08405
A	357	LEU	-	expression tag	UNP P08405
A	358	GLU	-	expression tag	UNP P08405
A	359	HIS	-	expression tag	UNP P08405
A	360	HIS	-	expression tag	UNP P08405
A	361	HIS	-	expression tag	UNP P08405
A	362	HIS	-	expression tag	UNP P08405
A	363	HIS	-	expression tag	UNP P08405
A	364	HIS	-	expression tag	UNP P08405
B	356	PRO	-	expression tag	UNP P08405
B	357	LEU	-	expression tag	UNP P08405
B	358	GLU	-	expression tag	UNP P08405
B	359	HIS	-	expression tag	UNP P08405
B	360	HIS	-	expression tag	UNP P08405
B	361	HIS	-	expression tag	UNP P08405
B	362	HIS	-	expression tag	UNP P08405
B	363	HIS	-	expression tag	UNP P08405
B	364	HIS	-	expression tag	UNP P08405
C	356	PRO	-	expression tag	UNP P08405
C	357	LEU	-	expression tag	UNP P08405
C	358	GLU	-	expression tag	UNP P08405
C	359	HIS	-	expression tag	UNP P08405
C	360	HIS	-	expression tag	UNP P08405

Continued on next page...

Continued from previous page...

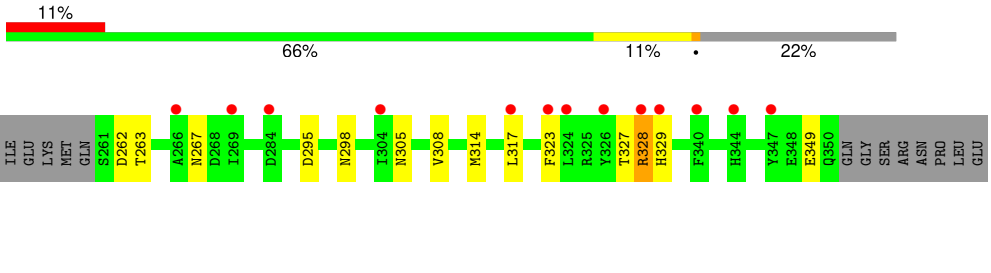
Chain	Residue	Modelled	Actual	Comment	Reference
C	361	HIS	-	expression tag	UNP P08405
C	362	HIS	-	expression tag	UNP P08405
C	363	HIS	-	expression tag	UNP P08405
C	364	HIS	-	expression tag	UNP P08405

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

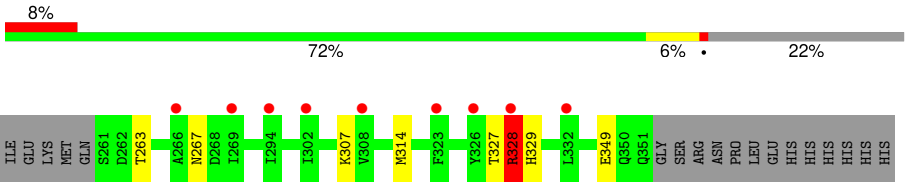
• Molecule 1: Ty1 Gag p22

Chain A:



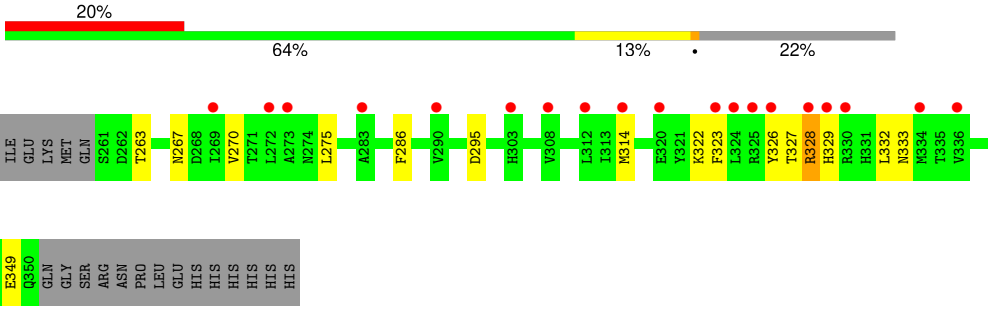
• Molecule 1: Ty1 Gag p22

Chain B:



• Molecule 1: Ty1 Gag p22

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	282.12Å 282.12Å 39.97Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	141.46 – 2.80 141.06 – 2.80	Depositor EDS
% Data completeness (in resolution range)	100.0 (141.46-2.80) 100.0 (141.06-2.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.32 (at 2.82Å)	Xtriage
Refinement program	REFMAC 5.8.0258, PHENIX 1.14_3260	Depositor
R, R_{free}	0.256 , 0.264 0.248 , 0.260	Depositor DCC
R_{free} test set	1259 reflections (5.28%)	wwPDB-VP
Wilson B-factor (Å^2)	98.7	Xtriage
Anisotropy	0.074	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	2189	wwPDB-VP
Average B, all atoms (Å^2)	161.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.64% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

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	0.64	0/735	0.74	0/995
1	B	0.65	0/756	0.76	0/1023
1	C	0.65	0/735	0.74	0/995
All	All	0.65	0/2226	0.74	0/3013

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

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	A	724	0	706	9	0
1	B	741	0	723	8	0
1	C	724	0	706	12	0
All	All	2189	0	2135	26	0

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 6.

All (26) 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)
1:A:327:THR:O	1:A:329:HIS:N	2.29	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:327:THR:O	1:C:329:HIS:N	2.29	0.66
1:B:327:THR:O	1:B:329:HIS:N	2.29	0.65
1:B:328:ARG:O	1:B:328:ARG:HD2	1.96	0.65
1:A:263:THR:HG23	1:A:267:ASN:ND2	2.24	0.52
1:A:305:ASN:OD1	1:A:308:VAL:HG12	2.09	0.51
1:C:263:THR:HG23	1:C:267:ASN:ND2	2.26	0.51
1:A:262:ASP:O	1:B:307:LYS:HE2	2.13	0.49
1:B:263:THR:HG23	1:B:267:ASN:ND2	2.27	0.49
1:B:328:ARG:HG3	1:B:329:HIS:N	2.29	0.47
1:C:267:ASN:O	1:C:270:VAL:HG12	2.15	0.46
1:B:314:MET:SD	1:B:328:ARG:HB2	2.56	0.45
1:C:332:LEU:HD23	1:C:333:ASN:H	1.81	0.45
1:A:314:MET:SD	1:A:328:ARG:HB2	2.59	0.43
1:A:317:LEU:HD12	1:A:317:LEU:HA	1.93	0.43
1:B:328:ARG:HD2	1:B:328:ARG:C	2.39	0.42
1:C:314:MET:SD	1:C:328:ARG:HB2	2.59	0.42
1:C:327:THR:C	1:C:329:HIS:H	2.24	0.42
1:C:332:LEU:HD23	1:C:333:ASN:N	2.35	0.42
1:A:323:PHE:CE1	1:C:323:PHE:HB2	2.55	0.41
1:A:323:PHE:CZ	1:C:323:PHE:HB2	2.55	0.41
1:A:327:THR:C	1:A:329:HIS:H	2.24	0.41
1:B:327:THR:C	1:B:329:HIS:H	2.23	0.41
1:C:275:LEU:HD22	1:C:286:PHE:HE1	1.86	0.41
1:C:275:LEU:HD22	1:C:286:PHE:CE1	2.56	0.41
1:C:322:LYS:HE3	1:C:326:TYR:OH	2.21	0.41

There are no symmetry-related clashes.

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
1	A	88/116 (76%)	84 (96%)	3 (3%)	1 (1%)	14 41

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	90/116 (78%)	86 (96%)	3 (3%)	1 (1%)	14 41
1	C	88/116 (76%)	83 (94%)	4 (4%)	1 (1%)	14 41
All	All	266/348 (76%)	253 (95%)	10 (4%)	3 (1%)	14 41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	328	ARG
1	B	328	ARG
1	C	328	ARG

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
1	A	79/104 (76%)	76 (96%)	3 (4%)	33 67
1	B	81/104 (78%)	79 (98%)	2 (2%)	47 80
1	C	79/104 (76%)	77 (98%)	2 (2%)	47 80
All	All	239/312 (77%)	232 (97%)	7 (3%)	42 76

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	295	ASP
1	A	298	ASN
1	A	349	GLU
1	B	328	ARG
1	B	349	GLU
1	C	295	ASP
1	C	349	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	300	ASN

5.3.3 RNA

There are no RNA molecules 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 monosaccharides 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, 95^{th} 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	90/116 (77%)	0.99	13 (14%) 2 1	113, 140, 185, 215	0
1	B	91/116 (78%)	0.91	9 (9%) 7 4	96, 125, 157, 167	0
1	C	90/116 (77%)	1.18	23 (25%) 0 0	173, 207, 242, 258	0
All	All	271/348 (77%)	1.02	45 (16%) 1 1	96, 150, 234, 258	0

All (45) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	323[A]	PHE	5.2
1	C	336	VAL	4.7
1	C	320	GLU	4.2
1	C	340	PHE	3.9
1	C	303	HIS	3.4
1	C	325	ARG	3.4
1	C	328	ARG	3.3
1	C	330	ARG	3.3
1	C	323	PHE	3.2
1	A	323	PHE	3.2
1	C	312	LEU	3.1
1	C	314	MET	3.1
1	A	328	ARG	3.0
1	B	302	ILE	3.0
1	A	326	TYR	2.8
1	B	308	VAL	2.8
1	A	347	TYR	2.7
1	A	344	HIS	2.7
1	B	328	ARG	2.7
1	A	329	HIS	2.7
1	C	269	ILE	2.6
1	A	340	PHE	2.6
1	C	329	HIS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	290	VAL	2.6
1	C	272	LEU	2.5
1	B	294	ILE	2.5
1	C	343	ILE	2.5
1	B	266	ALA	2.5
1	A	269	ILE	2.4
1	C	308	VAL	2.4
1	A	317	LEU	2.3
1	C	283	ALA	2.3
1	B	269	ILE	2.2
1	C	326	TYR	2.2
1	C	347	TYR	2.2
1	A	324	LEU	2.2
1	C	273	ALA	2.2
1	B	326	TYR	2.1
1	C	346	ILE	2.1
1	C	334	MET	2.1
1	C	324	LEU	2.1
1	A	284	ASP	2.1
1	A	304	ILE	2.0
1	A	266	ALA	2.0
1	B	332	LEU	2.0

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 monosaccharides 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.