Full wwPDB X-ray Structure Validation Report ⓘ

Jun 24, 2024 – 09:01 AM EDT

PDB ID	:	5NVR
Title	:	Crystal structure of the Rif1 N-terminal domain (RIF1-NTD) from Saccharomyces cerevisiae
Authors	:	Bunker, R.D.; Shi, T.; Thoma, N.H.
Deposited on	:	2017-05-04
Resolution	:	3.95 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

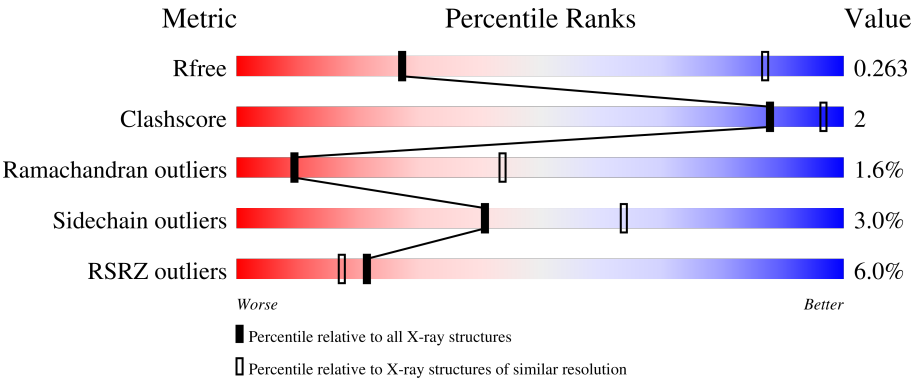
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 3.95 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	1025 (4.22-3.70)
Clashscore	141614	1085 (4.22-3.70)
Ramachandran outliers	138981	1047 (4.22-3.70)
Sidechain outliers	138945	1039 (4.22-3.70)
RSRZ outliers	127900	1013 (4.28-3.64)

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	1110	6% 88% 8%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 17642 atoms, of which 8951 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 Telomere length regulator protein RIF1.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S	Se			
1	A	1072	17642	5634	8951	1426	1597	12	22	0	0	0

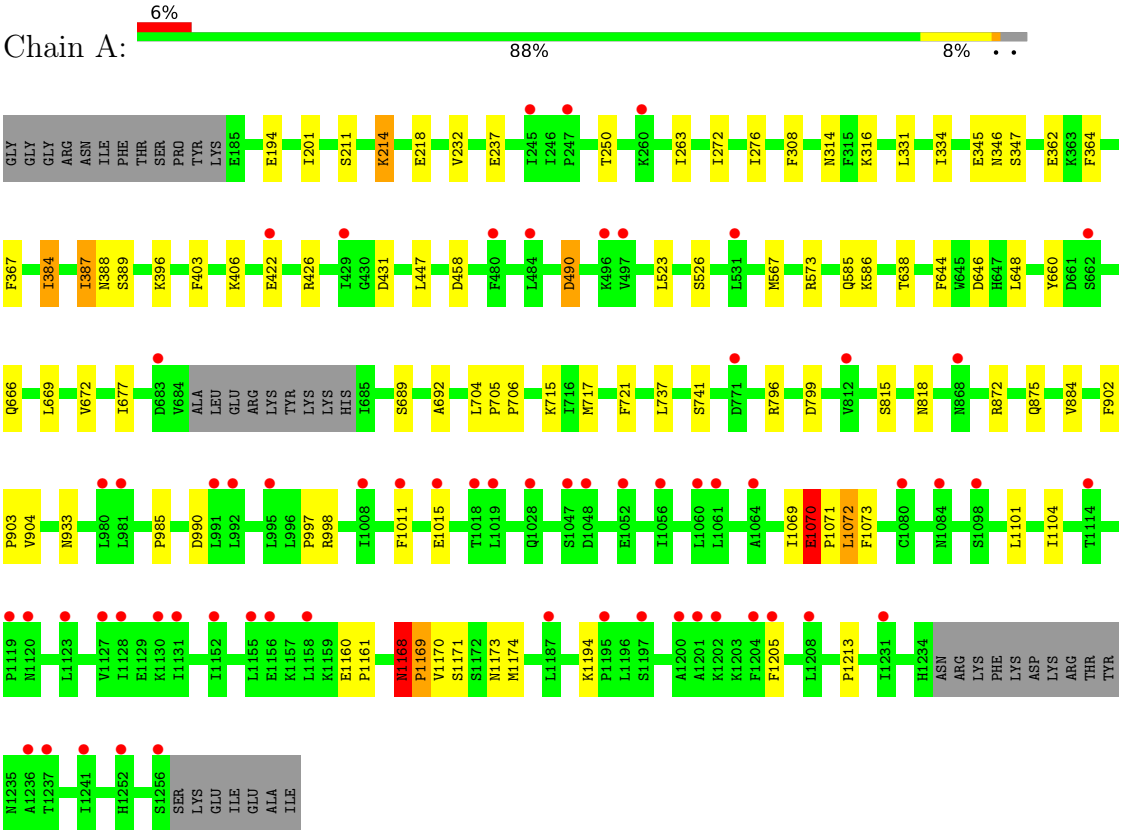
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	173	GLY	-	expression tag	UNP P29539
A	174	GLY	-	expression tag	UNP P29539
A	175	GLY	-	expression tag	UNP P29539
A	176	ARG	-	expression tag	UNP P29539

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: Telomere length regulator protein RIF1



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	203.57Å 203.57Å 197.72Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.30 – 3.95 49.29 – 3.94	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.30-3.95) 99.9 (49.29-3.94)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.71 (at 4.00Å)	Xtriage
Refinement program	PHENIX DEV_2439	Depositor
R, R_{free}	0.212 , 0.261 0.216 , 0.263	Depositor DCC
R_{free} test set	1132 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	182.9	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 200.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.44$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	17642	wwPDB-VP
Average B, all atoms (Å^2)	241.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.48% 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	0.25	1/8840 (0.0%)	0.40	0/11936

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	1194	LYS	C-N	5.86	1.45	1.34

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	8691	8951	8950	36	0
All	All	8691	8951	8950	36	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 2.

All (36) 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:422:GLU:OE1	1:A:426:ARG:NH2	2.24	0.70
1:A:1015:GLU:OE1	1:A:1015:GLU:N	2.32	0.62

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:362:GLU:OE1	1:A:406:LYS:NZ	2.36	0.58
1:A:237:GLU:N	1:A:237:GLU:OE1	2.36	0.57
1:A:737:LEU:O	1:A:741:SER:OG	2.19	0.56
1:A:1070:GLU:H	1:A:1071:PRO:HD2	1.72	0.54
1:A:387:ILE:HG23	1:A:388:ASN:H	1.73	0.53
1:A:1168:ASN:N	1:A:1169:PRO:CD	2.72	0.53
1:A:872:ARG:O	1:A:875:GLN:HG3	2.09	0.52
1:A:1173:ASN:OD1	1:A:1174:MSE:N	2.44	0.50
1:A:660:TYR:O	1:A:666:GLN:NE2	2.43	0.50
1:A:997:PRO:HB2	1:A:998:ARG:HA	1.95	0.49
1:A:1070:GLU:H	1:A:1071:PRO:CD	2.26	0.49
1:A:1072:LEU:H	1:A:1072:LEU:HD13	1.79	0.48
1:A:717:MSE:HE1	1:A:741:SER:HB3	1.97	0.47
1:A:705:PRO:HB2	1:A:706:PRO:HD2	1.96	0.46
1:A:717:MSE:HE3	1:A:721:PHE:CE1	2.51	0.46
1:A:646:ASP:OD2	1:A:715:LYS:NZ	2.48	0.46
1:A:737:LEU:O	1:A:741:SER:CB	2.64	0.45
1:A:447:LEU:HD23	1:A:526:SER:HB3	1.98	0.44
1:A:345:GLU:O	1:A:347:SER:N	2.50	0.44
1:A:997:PRO:CB	1:A:998:ARG:HA	2.48	0.43
1:A:431:ASP:OD1	1:A:431:ASP:N	2.51	0.43
1:A:214:LYS:HA	1:A:263:ILE:HG21	2.01	0.43
1:A:644:PHE:O	1:A:648:LEU:HB3	2.18	0.43
1:A:573:ARG:NH1	1:A:692:ALA:O	2.52	0.41
1:A:669:LEU:O	1:A:672:VAL:HB	2.20	0.41
1:A:1160:GLU:HB2	1:A:1161:PRO:HD3	2.01	0.41
1:A:384:ILE:HG12	1:A:396:LYS:HZ2	1.85	0.41
1:A:815:SER:OG	1:A:818:ASN:ND2	2.44	0.41
1:A:331:LEU:O	1:A:334:ILE:HG22	2.20	0.41
1:A:490:ASP:OD1	1:A:490:ASP:N	2.53	0.41
1:A:1169:PRO:O	1:A:1171:SER:N	2.54	0.41
1:A:796:ARG:HG3	1:A:799:ASP:HB2	2.03	0.41
1:A:272:ILE:HG23	1:A:276:ILE:HD12	2.02	0.41
1:A:211:SER:O	1:A:214:LYS:HB3	2.21	0.40

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	1066/1110 (96%)	972 (91%)	77 (7%)	17 (2%)	9 44

All (17) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1070	GLU
1	A	346	ASN
1	A	904	VAL
1	A	1069	ILE
1	A	1170	VAL
1	A	903	PRO
1	A	1169	PRO
1	A	387	ILE
1	A	458	ASP
1	A	704	LEU
1	A	985	PRO
1	A	1213	PRO
1	A	201	ILE
1	A	523	LEU
1	A	384	ILE
1	A	677	ILE
1	A	1168	ASN

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	1007/1018 (99%)	977 (97%)	30 (3%)	41 64

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	194	GLU
1	A	214	LYS
1	A	218	GLU
1	A	232	VAL
1	A	250	THR
1	A	308	PHE
1	A	314	ASN
1	A	316	LYS
1	A	364	PHE
1	A	367	PHE
1	A	389	SER
1	A	403	PHE
1	A	490	ASP
1	A	567	MSE
1	A	585	GLN
1	A	586	LYS
1	A	638	THR
1	A	689	SER
1	A	884	VAL
1	A	902	PHE
1	A	933	ASN
1	A	990	ASP
1	A	1011	PHE
1	A	1070	GLU
1	A	1072	LEU
1	A	1073	PHE
1	A	1101	LEU
1	A	1104	ILE
1	A	1168	ASN
1	A	1205	PHE

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

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 monosaccharides 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

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	1050/1110 (94%)	0.34	63 (6%) 21 17	112, 211, 333, 470	0

All (63) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1156	GLU	6.9
1	A	1152	ILE	5.4
1	A	496	LYS	4.8
1	A	1155	LEU	4.7
1	A	1204	PHE	4.2
1	A	1195	PRO	4.1
1	A	247	PRO	4.0
1	A	1201	ALA	3.9
1	A	812	VAL	3.8
1	A	1015	GLU	3.8
1	A	1114	THR	3.8
1	A	1158	LEU	3.7
1	A	1202	LYS	3.7
1	A	1084	ASN	3.5
1	A	429	ILE	3.4
1	A	683	ASP	3.4
1	A	1098	SER	3.3
1	A	992	LEU	3.2
1	A	422	GLU	3.1
1	A	480	PHE	3.1
1	A	1241	ILE	3.1
1	A	1018	THR	3.0
1	A	1231	ILE	3.0
1	A	981	LEU	3.0
1	A	1056	ILE	3.0
1	A	1197	SER	2.9
1	A	1120	ASN	2.9

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1256	SER	2.9
1	A	1130	LYS	2.9
1	A	1052	GLU	2.9
1	A	1119	PRO	2.9
1	A	484	LEU	2.9
1	A	1236	ALA	2.9
1	A	260	LYS	2.8
1	A	1047	SER	2.8
1	A	1011	PHE	2.8
1	A	1208	LEU	2.8
1	A	1252	HIS	2.7
1	A	1080	CYS	2.7
1	A	1048	ASP	2.6
1	A	1127	VAL	2.6
1	A	1019	LEU	2.6
1	A	771	ASP	2.5
1	A	868	ASN	2.5
1	A	497	VAL	2.5
1	A	1205	PHE	2.5
1	A	1200	ALA	2.5
1	A	1123	LEU	2.5
1	A	1128	ILE	2.4
1	A	1237	THR	2.3
1	A	995	LEU	2.3
1	A	245	ILE	2.3
1	A	980	LEU	2.3
1	A	1131	ILE	2.2
1	A	991	LEU	2.1
1	A	1061	LEU	2.1
1	A	531	LEU	2.1
1	A	1060	LEU	2.1
1	A	1187	LEU	2.1
1	A	1028	GLN	2.1
1	A	662	SER	2.1
1	A	1008	ILE	2.0
1	A	1064	ALA	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.