W O R L D W I D E

wwPDB

P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 07:06 AM GMT

PDB ID	:	2ZII
Title	:	Crystal Structure of Yeast Vps74-N-term Truncation Variant
Authors	:	Schmitz, K.R.; Li, S.; Setty, T.G.; Ferguson, K.M.
Deposited on	:	2008-02-18
Resolution	:	3.05 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

http://wwpdb.org/validation/2016/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:

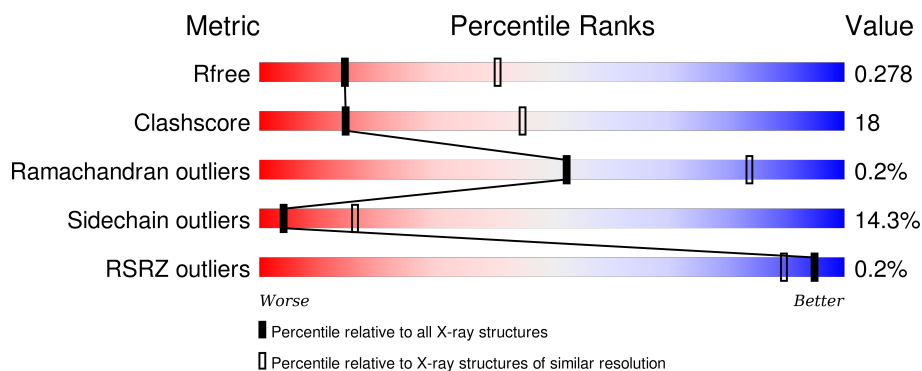|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.05 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	Percentile Ranks	Value
Rfree		0.278
Clashscore		18
Ramachandran outliers		0.2%
Sidechain outliers		14.3%
RSRZ outliers		0.2%

Worse   *Better*

■ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	91344	1191 (3.10-3.02)
Clashscore	102246	1303 (3.10-3.02)
Ramachandran outliers	100387	1254 (3.10-3.02)
Sidechain outliers	100360	1254 (3.10-3.02)
RSRZ outliers	91569	1197 (3.10-3.02)

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. 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	288	67% 24% 7% •
1	B	288	59% 32% 7% •
1	C	288	61% 28% 8% •
1	D	288	68% 24% 7% •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8590 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 Vacuolar protein sorting-associated protein 74.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	281	Total	C	N	O	S	0	0	0
			2149	1365	356	416	12			
1	B	281	Total	C	N	O	S	0	0	0
			2126	1348	356	411	11			
1	C	279	Total	C	N	O	S	0	0	0
			2125	1351	355	408	11			
1	D	283	Total	C	N	O	S	0	0	0
			2188	1391	371	414	12			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	58	GLY	-	EXPRESSION TAG	UNP Q06385
A	59	SER	-	EXPRESSION TAG	UNP Q06385
B	58	GLY	-	EXPRESSION TAG	UNP Q06385
B	59	SER	-	EXPRESSION TAG	UNP Q06385
C	58	GLY	-	EXPRESSION TAG	UNP Q06385
C	59	SER	-	EXPRESSION TAG	UNP Q06385
D	58	GLY	-	EXPRESSION TAG	UNP Q06385
D	59	SER	-	EXPRESSION TAG	UNP Q06385

- Molecule 2 is water.

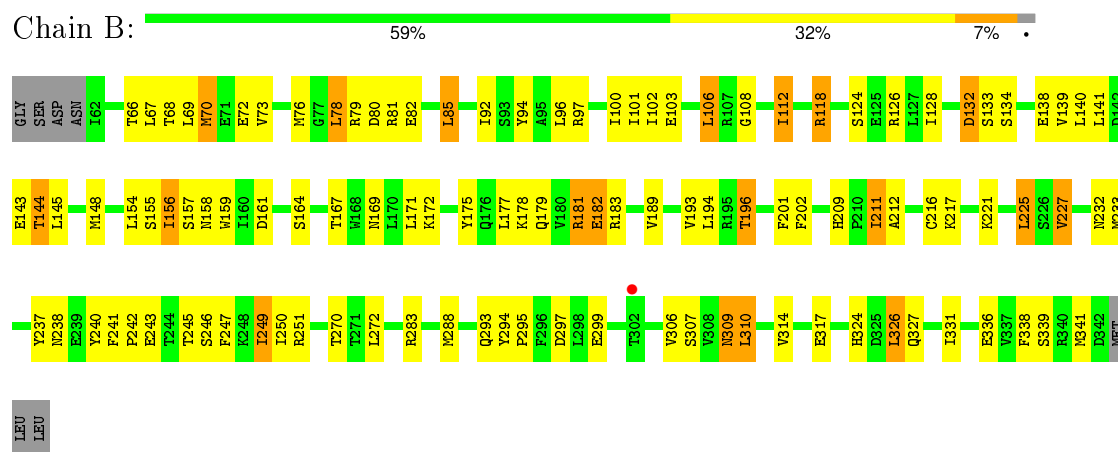| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	O	0	0
			1	1		
2	D	1	Total	O	0	0
			1	1		

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.
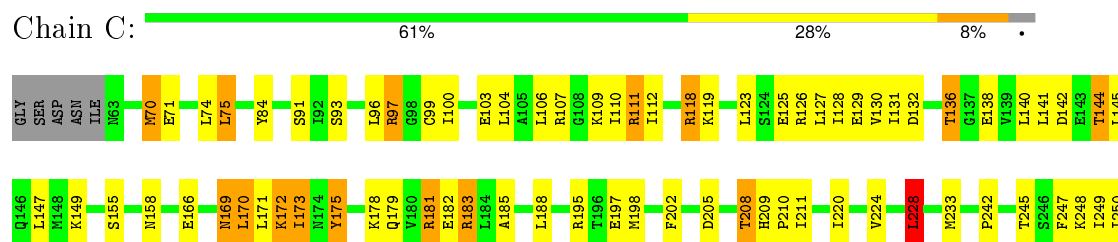
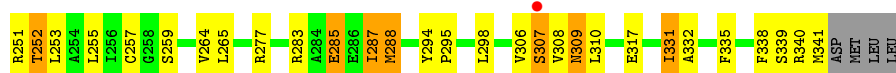• Molecule 1: Vacuolar protein sorting-associated protein 74
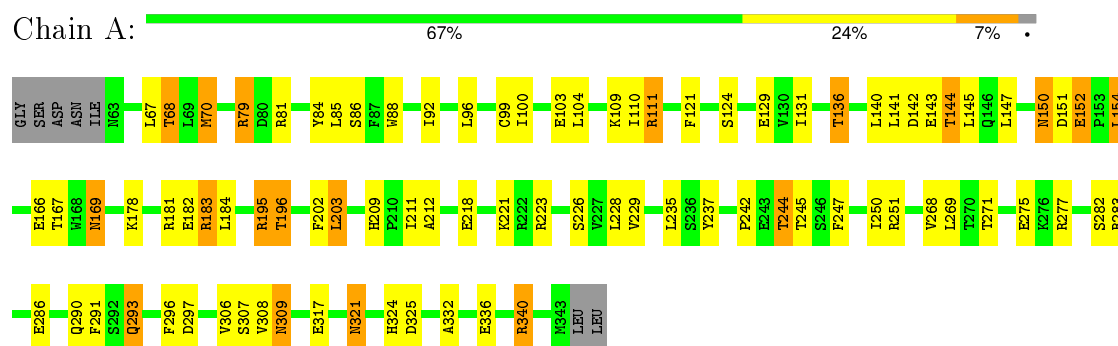
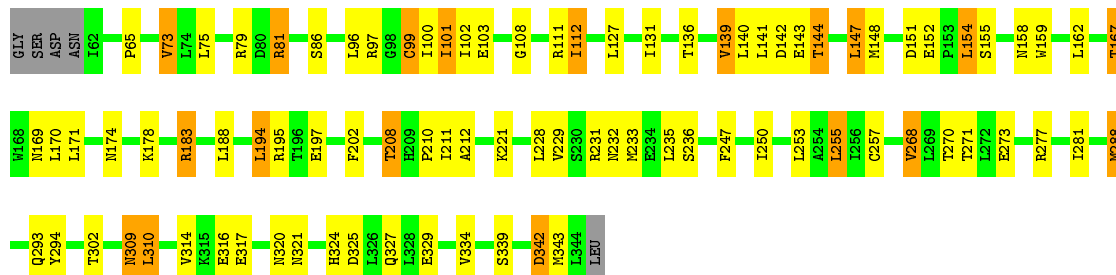• Molecule 1: Vacuolar protein sorting-associated protein 74

• Molecule 1: Vacuolar protein sorting-associated protein 74

R251 T252 L253 A254 L255 I256 C257 G258 S259 V264 L265 R277 R283 A284 E285 E286 I287 M288 Y294 P295 L298 V306 S307 V308 N309 L310 E317 I331 A332 F335 F338 S339 R340 M341 ASP MET LEU LEU

• Molecule 1: Vacuolar protein sorting-associated protein 74

Chain D:

68%   24%   7%   •

GLY SER ASP ASN I62 P65 V73 L74 L75 R79 D80 R81 S86 L96 R97 G98 C99 I100 I101 I102 E103 G108 R111 I112 L127 I131 T136 V139 L140 L141 D142 E143 T144 L147 M148 D151 E152 P153 L154 S155 N158 W159 L162 T167

W168 N169 L170 L171 N174 K178 R183 L188 L194 R195 T196 E197 F202 T208 H209 P210 I211 A212 K221 L228 V229 S230 R231 N232 M233 E234 L235 S236 F247 I250 L253 A254 L255 I256 C257 V268 L269 T270 T271 L272 E273 R277 I281 M288

Q293 Y294 T302 N309 L310 V314 K315 E316 E317 N320 N321 H324 D325 L326 Q327 L328 E329 V334 S339 D342 M343 L344 LEU

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	104.08Å 104.08Å 292.76Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.10 – 3.05 49.10 – 3.05	Depositor EDS
% Data completeness (in resolution range)	99.7 (49.10-3.05) 99.8 (49.10-3.05)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.25 (at 3.07Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.220 , 0.281 0.219 , 0.278	Depositor DCC
R_{free} test set	1790 reflections (5.26%)	DCC
Wilson B-factor (Å^2)	57.8	Xtriage
Anisotropy	0.283	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 59.8	EDS
Estimated twinning fraction	0.013 for -h,-k,l	Xtriage
L-test for twinning[2]	$< |L| > = 0.51, < L^2 > = 0.35$	Xtriage
Outliers	1 of 35793 reflections (0.003%)	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	8590	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 22.79 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 5.3113e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.67	0/2180	0.82	2/2959 (0.1%)
1	B	0.58	0/2157	0.76	0/2934
1	C	0.63	0/2156	0.85	6/2929 (0.2%)
1	D	0.68	1/2218 (0.0%)	0.82	2/3005 (0.1%)
All	All	0.64	1/8711 (0.0%)	0.81	10/11827 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	99	CYS	CB-SG	-5.99	1.72	1.81

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	228	LEU	CA-CB-CG	8.54	134.93	115.30
1	A	235	LEU	CA-CB-CG	5.83	128.70	115.30
1	A	154	LEU	CB-CG-CD2	-5.82	101.10	111.00
1	C	111	ARG	NE-CZ-NH1	5.71	123.16	120.30
1	C	170	LEU	CB-CA-C	-5.59	99.59	110.20

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	2149	0	2079	73	0
1	B	2126	0	2026	96	0
1	C	2125	0	2049	79	0
1	D	2188	0	2162	62	0
2	A	1	0	0	0	0
2	D	1	0	0	0	0
All	All	8590	0	8316	300	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 18.

The worst 5 of 300 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:B:106:LEU:HD13	1:B:331:ILE:HD11	1.35	1.05
1:A:70:MET:H	1:A:70:MET:HE2	1.18	1.05
1:A:242:PRO:HD2	1:A:245:THR:CG2	1.87	1.04
1:C:233:MET:HE1	1:C:310:LEU:HD13	1.45	0.97
1:C:198:MET:HE1	1:C:205:ASP:HB3	1.50	0.93

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	279/288 (97%)	258 (92%)	21 (8%)	0	100	100
1	B	279/288 (97%)	250 (90%)	28 (10%)	1 (0%)	39	74
1	C	277/288 (96%)	253 (91%)	23 (8%)	1 (0%)	39	74
1	D	281/288 (98%)	256 (91%)	25 (9%)	0	100	100
All	All	1116/1152 (97%)	1017 (91%)	97 (9%)	2 (0%)	52	84

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	270	THR
1	C	173	ILE

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	226/260 (87%)	197 (87%)	29 (13%)	5	20
1	B	219/260 (84%)	188 (86%)	31 (14%)	4	17
1	C	221/260 (85%)	186 (84%)	35 (16%)	3	12
1	D	233/260 (90%)	199 (85%)	34 (15%)	4	15
All	All	899/1040 (86%)	770 (86%)	129 (14%)	4	16

5 of 129 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	326	LEU
1	C	144	THR
1	D	268	VAL
1	C	70	MET
1	C	106	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 25 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	324	HIS
1	C	146	GLN
1	D	309	ASN
1	B	327	GLN
1	C	158	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 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 (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(Å^2)	Q<0.9
1	A	281/288 (97%)	-0.46	0	100	100	25, 38, 53, 68	0
1	B	281/288 (97%)	-0.23	1 (0%)	93	84	32, 49, 65, 70	0
1	C	279/288 (96%)	-0.29	1 (0%)	93	84	31, 46, 63, 66	0
1	D	283/288 (98%)	-0.36	0	100	100	22, 40, 56, 66	0
All	All	1124/1152 (97%)	-0.33	2 (0%)	95	89	22, 44, 62, 70	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	302	THR	2.5
1	C	307	SER	2.3

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.