Full wwPDB X-ray Structure Validation Report ⓘ

Nov 25, 2024 – 05:26 PM EST

PDB ID	:	2Y7O
Title	:	Structure of N-terminal domain of Candida albicans als9-2 - G299W mutant
Authors	:	Salgado, P.S.; Burchell, L.; Cota, E.
Deposited on	:	2011-01-31
Resolution	:	1.75 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:

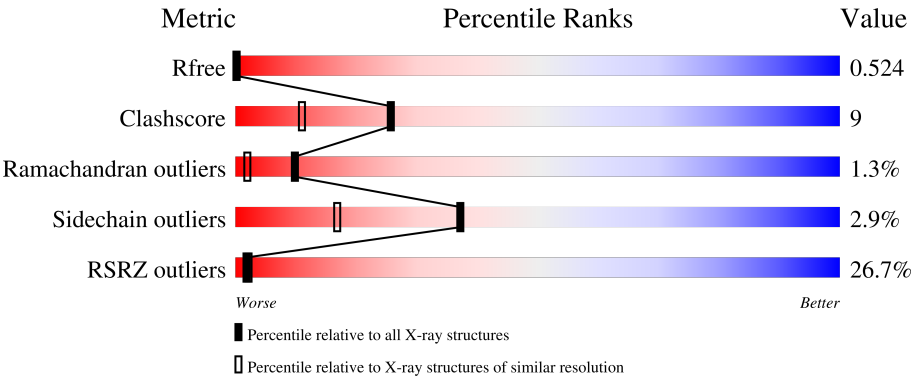
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.21
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.004 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.40

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 1.75 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}	164625	2888 (1.76-1.76)
Clashscore	180529	3097 (1.76-1.76)
Ramachandran outliers	177936	3072 (1.76-1.76)
Sidechain outliers	177891	3072 (1.76-1.76)
RSRZ outliers	164620	2887 (1.76-1.76)

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%

Mol	Chain	Length	Quality of chain
1	A	312	27% 88% 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 2637 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 AGGLUTININ-LIKE ALS9 PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	311	2362	1497	365	490	10	1	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MET	-	expression tag	UNP Q5A8T1
A	51	THR	ASN	conflict	UNP Q5A8T1
A	212	VAL	ILE	conflict	UNP Q5A8T1
A	296	TRP	GLY	engineered mutation	UNP Q5A8T1

- Molecule 2 is water.

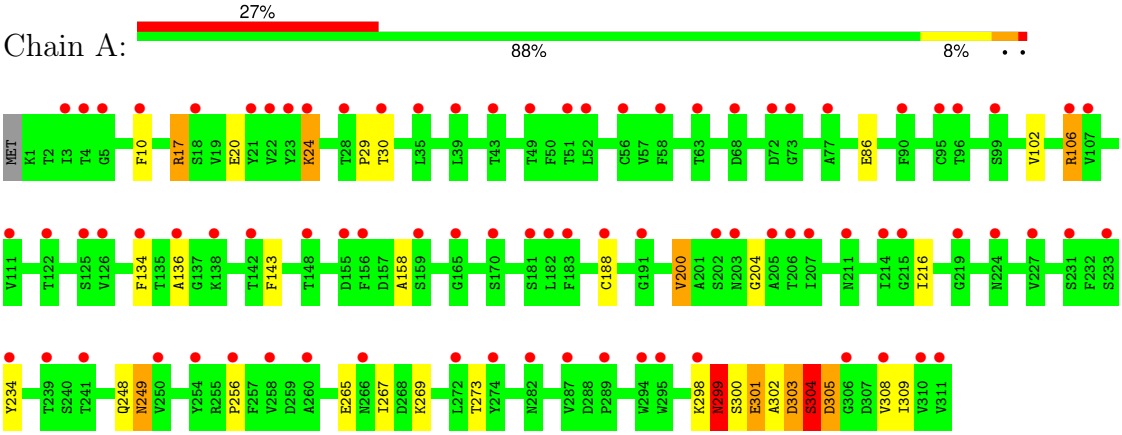
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	275	275	275	0	0

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. 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. 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: AGGLUTININ-LIKE ALS9 PROTEIN

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.37Å 68.31Å 89.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.51 – 1.75 36.51 – 1.75	Depositor EDS
% Data completeness (in resolution range)	100.0 (36.51-1.75) 52.7 (36.51-1.75)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.39 (at 1.26Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.195 , 0.223 0.546 , 0.524	Depositor DCC
R_{free} test set	1289 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	4.4	Xtriage
Anisotropy	2.187	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 245.6	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.68	EDS
Total number of atoms	2637	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.32% 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	1.74	14/2420 (0.6%)	1.19	4/3314 (0.1%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	106	ARG	CZ-NH1	53.99	2.03	1.33
1	A	106	ARG	CZ-NH2	11.09	1.47	1.33
1	A	106	ARG	NE-CZ	9.77	1.45	1.33
1	A	301	GLU	CD-OE2	8.36	1.34	1.25
1	A	86	GLU	CD-OE1	6.43	1.32	1.25
1	A	234	TYR	CE1-CZ	-6.27	1.30	1.38
1	A	304	SER	C-O	6.18	1.35	1.23
1	A	24	LYS	CE-NZ	6.07	1.64	1.49
1	A	301	GLU	CD-OE1	5.94	1.32	1.25
1	A	265	GLU	CD-OE1	5.88	1.32	1.25
1	A	134	PHE	CE1-CZ	5.82	1.48	1.37
1	A	304	SER	CB-OG	5.76	1.49	1.42
1	A	143	PHE	CE2-CZ	5.58	1.48	1.37
1	A	188	CYS	CB-SG	-5.37	1.73	1.81

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	106	ARG	NE-CZ-NH2	-27.27	106.67	120.30
1	A	265	GLU	C-N-CA	-7.58	102.76	121.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	106	ARG	NE-CZ-NH1	6.64	123.62	120.30
1	A	308	VAL	CG1-CB-CG2	-5.14	102.68	110.90

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	106	ARG	Sidechain

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	A	2362	0	2231	43	0
2	A	275	0	0	1	0
All	All	2637	0	2231	43	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 9.

All (43) 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:298:LYS:HA	1:A:299:ASN:CB	1.47	1.31
1:A:299:ASN:N	1:A:301:GLU:OE1	1.79	1.16
1:A:298:LYS:HA	1:A:299:ASN:HB2	1.18	1.14
1:A:301:GLU:O	1:A:302:ALA:HB3	1.39	1.10
1:A:298:LYS:HA	1:A:299:ASN:HB3	1.30	1.08
1:A:299:ASN:H	1:A:301:GLU:CD	1.61	1.02
1:A:298:LYS:CA	1:A:299:ASN:CB	2.38	1.02
1:A:303:ASP:CG	1:A:304:SER:H	1.66	0.97
1:A:298:LYS:NZ	1:A:300:SER:HB2	1.81	0.96
1:A:301:GLU:O	1:A:302:ALA:CB	2.12	0.93
1:A:298:LYS:CA	1:A:299:ASN:HB2	2.00	0.90
1:A:298:LYS:HG2	1:A:300:SER:N	1.87	0.90

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:299:ASN:O	1:A:299:ASN:ND2	2.07	0.88
1:A:298:LYS:HZ3	1:A:300:SER:HB2	1.39	0.83
1:A:298:LYS:HZ3	1:A:300:SER:CB	1.94	0.81
1:A:17:ARG:HB3	2:A:2019:HOH:O	1.80	0.80
1:A:298:LYS:HG2	1:A:300:SER:H	1.45	0.79
1:A:303:ASP:CG	1:A:304:SER:N	2.40	0.75
1:A:300:SER:H	1:A:301:GLU:HA	1.52	0.74
1:A:298:LYS:NZ	1:A:300:SER:CB	2.50	0.74
1:A:298:LYS:C	1:A:301:GLU:OE1	2.28	0.71
1:A:298:LYS:CA	1:A:299:ASN:HB3	2.14	0.71
1:A:298:LYS:HD2	1:A:309:ILE:HD11	1.76	0.66
1:A:300:SER:N	1:A:301:GLU:HA	2.17	0.59
1:A:299:ASN:HD22	1:A:299:ASN:C	1.98	0.56
1:A:298:LYS:O	1:A:301:GLU:OE1	2.25	0.55
1:A:299:ASN:N	1:A:301:GLU:CD	2.44	0.55
1:A:303:ASP:OD2	1:A:304:SER:N	2.28	0.53
1:A:204:GLY:HA3	1:A:267:ILE:HD13	1.89	0.53
1:A:200:VAL:HG13	1:A:273:THR:HB	1.91	0.52
1:A:136:ALA:HB2	1:A:158:ALA:HB2	1.96	0.48
1:A:248:GLN:O	1:A:249:ASN:O	2.32	0.47
1:A:24:LYS:HB3	1:A:24:LYS:HE3	1.85	0.44
1:A:216:ILE:HG22	1:A:256:PRO:HA	1.99	0.44
1:A:298:LYS:HZ1	1:A:303:ASP:CB	2.31	0.44
1:A:136:ALA:CB	1:A:158:ALA:HB2	2.47	0.43
1:A:298:LYS:HZ1	1:A:303:ASP:HB3	1.83	0.43
1:A:298:LYS:CG	1:A:299:ASN:HB3	2.48	0.43
1:A:267:ILE:HD12	1:A:269:LYS:O	2.19	0.43
1:A:298:LYS:HZ2	1:A:300:SER:HB2	1.74	0.42
1:A:29:PRO:HA	1:A:30:THR:HA	1.88	0.41
1:A:298:LYS:NZ	1:A:303:ASP:HB2	2.36	0.41
1:A:248:GLN:O	1:A:249:ASN:C	2.57	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	309/312 (99%)	295 (96%)	10 (3%)	4 (1%)	10 2

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	249	ASN
1	A	299	ASN
1	A	305	ASP
1	A	303	ASP

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	272/274 (99%)	264 (97%)	8 (3%)	37 17

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	10	PHE
1	A	17	ARG
1	A	20	GLU
1	A	102	VAL
1	A	200	VAL
1	A	299	ASN
1	A	304	SER
1	A	305	ASP

Sometimes 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 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 oligosaccharides 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

Warning: The R factor obtained from EDS is 0.416, which does not match the depositor's R factor of 0.19545. Please interpret the results in this section carefully.

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	311/312 (99%)	1.42	83 (26%) 2 2	12, 22, 39, 63	4 (1%)

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	256	PRO	4.9
1	A	224	ASN	4.7
1	A	219	GLY	4.4
1	A	107	VAL	4.2
1	A	188	CYS	4.1
1	A	99	SER	3.8
1	A	215	GLY	3.7
1	A	295	TRP	3.7
1	A	231	SER	3.7
1	A	126	VAL	3.6
1	A	136	ALA	3.6
1	A	191	GLY	3.5
1	A	106	ARG	3.5
1	A	170	SER	3.4
1	A	111	VAL	3.4
1	A	56	CYS	3.3
1	A	22	VAL	3.3
1	A	207	ILE	3.2
1	A	181	SER	3.2
1	A	287	VAL	3.2
1	A	49	THR	3.2
1	A	3	ILE	3.2
1	A	90	PHE	3.0
1	A	24	LYS	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	254	TYR	3.0
1	A	39	LEU	2.9
1	A	310	VAL	2.9
1	A	95	CYS	2.9
1	A	298	LYS	2.9
1	A	234	TYR	2.9
1	A	241	THR	2.8
1	A	239	THR	2.7
1	A	156	PHE	2.7
1	A	18	SER	2.6
1	A	122	THR	2.6
1	A	272	LEU	2.6
1	A	10	PHE	2.6
1	A	35	LEU	2.5
1	A	274	TYR	2.5
1	A	138	LYS	2.5
1	A	306	GLY	2.5
1	A	77	ALA	2.5
1	A	73	GLY	2.5
1	A	58	PHE	2.5
1	A	182	LEU	2.5
1	A	202	SER	2.4
1	A	159	SER	2.4
1	A	134	PHE	2.4
1	A	165	GLY	2.4
1	A	148	THR	2.4
1	A	289	PRO	2.4
1	A	308	VAL	2.4
1	A	63	THR	2.3
1	A	23	TYR	2.3
1	A	282	ASN	2.3
1	A	206	THR	2.3
1	A	4	THR	2.3
1	A	28	THR	2.3
1	A	311	VAL	2.3
1	A	125	SER	2.3
1	A	294	TRP	2.2
1	A	155	ASP	2.2
1	A	205	ALA	2.2
1	A	266	ASN	2.2
1	A	233	SER	2.2
1	A	203	ASN	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	183	PHE	2.2
1	A	96	THR	2.2
1	A	250	VAL	2.2
1	A	52	LEU	2.1
1	A	43	THR	2.1
1	A	211	ASN	2.1
1	A	51	THR	2.1
1	A	5	GLY	2.1
1	A	227	VAL	2.1
1	A	260	ALA	2.0
1	A	142	THR	2.0
1	A	68	ASP	2.0
1	A	72	ASP	2.0
1	A	21	TYR	2.0
1	A	258	VAL	2.0
1	A	30	THR	2.0
1	A	214	ILE	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.