Full wwPDB X-ray Structure Validation Report ⓘ

Jun 17, 2024 – 05:44 AM EDT

PDB ID	:	3KT4
Title	:	Crystal structure of Tpa1 from Saccharomyces cerevisiae, a component of the messenger ribonucleoprotein complex
Authors	:	Kim, H.S.; Kim, H.L.; Kim, K.H.; Kim, D.J.; Lee, S.J.; Yoon, J.Y.; Yoon, H.J.; Lee, H.Y.; Park, S.B.; Kim, S.-J.; Lee, J.Y.; Suh, S.W.
Deposited on	:	2009-11-24
Resolution	:	2.73 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:

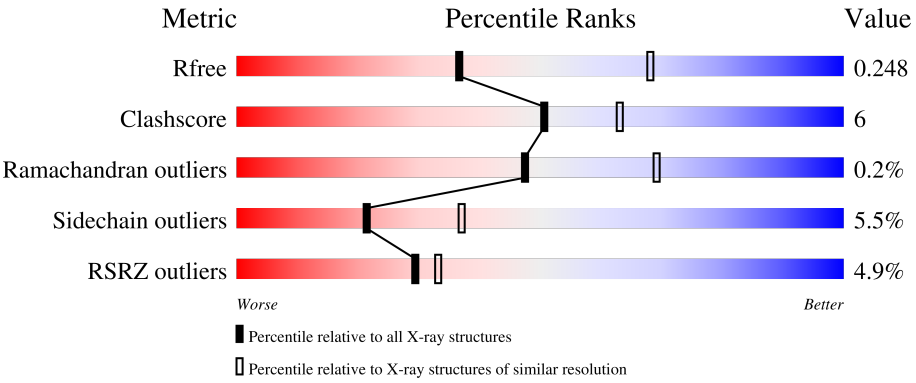
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, 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 2.73 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	1271 (2.76-2.72)
Clashscore	141614	1322 (2.76-2.72)
Ramachandran outliers	138981	1297 (2.76-2.72)
Sidechain outliers	138945	1298 (2.76-2.72)
RSRZ outliers	127900	1243 (2.76-2.72)

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	633	4% 74% 12% • 11%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4803 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 PKHD-type hydroxylase TPA1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	563	Total	C	N	O	S	Se	0	0	0
			4603	2953	768	869	6	7			

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	MSE	-	EXPRESSION TAG	UNP P40032
A	645	LEU	-	EXPRESSION TAG	UNP P40032
A	646	GLU	-	EXPRESSION TAG	UNP P40032
A	647	HIS	-	EXPRESSION TAG	UNP P40032
A	648	HIS	-	EXPRESSION TAG	UNP P40032
A	649	HIS	-	EXPRESSION TAG	UNP P40032
A	650	HIS	-	EXPRESSION TAG	UNP P40032
A	651	HIS	-	EXPRESSION TAG	UNP P40032
A	652	HIS	-	EXPRESSION TAG	UNP P40032

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	Fe	0	0
			1	1		

- Molecule 3 is water.

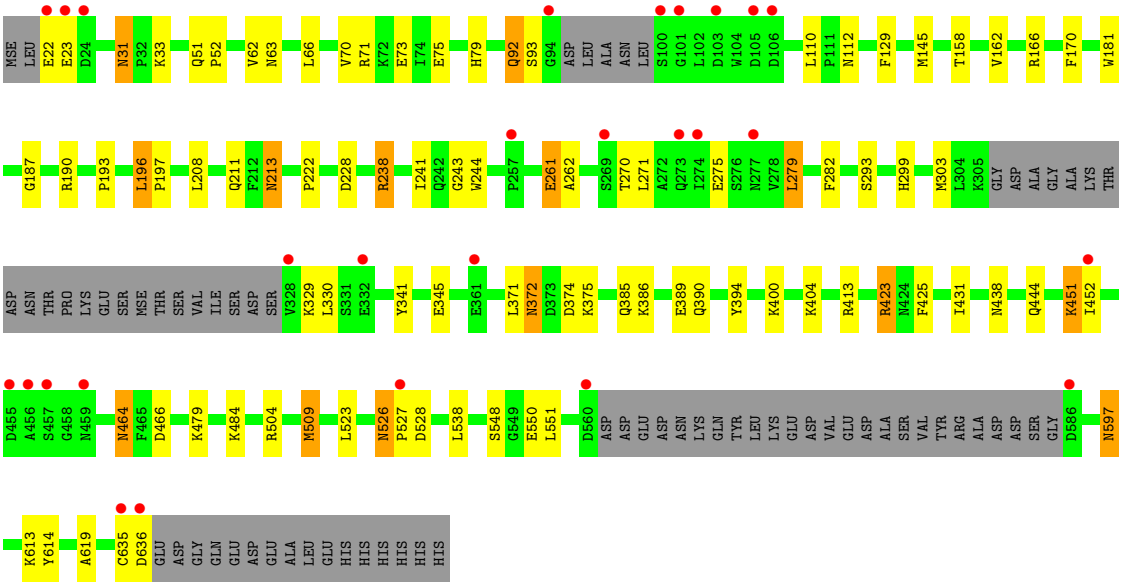
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	199	Total	O	0	0
			199	199		

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: PKHD-type hydroxylase TPA1

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	136.29Å 136.29Å 83.28Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 2.73 19.94 – 2.73	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-2.73) 99.8 (19.94-2.73)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	6.61 (at 2.75Å)	Xtriage
Refinement program	REFMAC 5.5.0072	Depositor
R, R_{free}	0.190 , 0.247 0.187 , 0.248	Depositor DCC
R_{free} test set	1222 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	40.0	Xtriage
Anisotropy	0.130	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 40.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.023 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4803	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.79% 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

Bond lengths and bond angles in the following residue types are not validated in this section: FE

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	0.50	0/4708	0.61	2/6349 (0.0%)

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

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	196	LEU	CA-CB-CG	6.22	129.60	115.30
1	A	330	LEU	CA-CB-CG	5.42	127.76	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	509	MSE	Peptide

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	4603	0	4529	52	0
2	A	1	0	0	0	0
3	A	199	0	0	11	0
All	All	4803	0	4529	52	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 (52) 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:181:TRP:H	1:A:211:GLN:HE22	1.20	0.89
1:A:444:GLN:HG3	3:A:983:HOH:O	1.76	0.86
1:A:464:ASN:HD22	1:A:466:ASP:H	1.32	0.77
1:A:548:SER:OG	1:A:550:GLU:HG2	1.84	0.77
1:A:371:LEU:H	1:A:597:ASN:HD21	1.36	0.73
1:A:526:ASN:HD21	1:A:528:ASP:HB2	1.55	0.70
1:A:526:ASN:HA	3:A:972:HOH:O	1.91	0.69
1:A:526:ASN:C	1:A:526:ASN:HD22	1.95	0.69
1:A:92:GLN:HG2	3:A:936:HOH:O	1.91	0.69
1:A:62:VAL:H	1:A:213:ASN:HD21	1.43	0.65
1:A:423:ARG:HG2	1:A:425:PHE:CZ	2.31	0.65
1:A:166:ARG:HD3	1:A:244:TRP:CD1	2.32	0.64
1:A:79:HIS:HB2	3:A:932:HOH:O	1.99	0.63
1:A:145:MSE:HE3	1:A:241:ILE:HG21	1.82	0.62
1:A:413:ARG:HD3	1:A:509:MSE:HE3	1.80	0.62
1:A:394:TYR:HA	3:A:992:HOH:O	2.00	0.60
1:A:345:GLU:HB2	3:A:945:HOH:O	2.02	0.59
1:A:31:ASN:HD22	1:A:33:LYS:H	1.55	0.55
1:A:73:GLU:HB3	1:A:110:LEU:HD23	1.88	0.55
1:A:22:GLU:HA	3:A:990:HOH:O	2.08	0.52
1:A:509:MSE:SE	3:A:992:HOH:O	2.76	0.52
1:A:63:ASN:HD22	1:A:66:LEU:H	1.59	0.51
1:A:162:VAL:HG21	1:A:222:PRO:HA	1.92	0.51
1:A:464:ASN:ND2	1:A:466:ASP:H	2.06	0.50
1:A:635:CYS:O	1:A:636:ASP:HB2	2.15	0.47
1:A:190:ARG:NH1	1:A:228:ASP:OD1	2.48	0.47
1:A:451:LYS:HD3	1:A:451:LYS:HA	1.80	0.47
1:A:282:PHE:HD1	1:A:431:ILE:HG21	1.80	0.46
1:A:613:LYS:HE3	1:A:614:TYR:HB3	1.96	0.46
1:A:275:GLU:HA	1:A:279:LEU:HB2	1.97	0.46
1:A:293:SER:HB2	3:A:919:HOH:O	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:187:GLY:O	1:A:238:ARG:NH2	2.45	0.45
1:A:372:ASN:C	1:A:372:ASN:HD22	2.19	0.45
1:A:385:GLN:NE2	1:A:389:GLU:OE1	2.46	0.45
1:A:509:MSE:HE2	3:A:992:HOH:O	2.16	0.45
1:A:526:ASN:HD22	1:A:527:PRO:N	2.15	0.45
1:A:371:LEU:H	1:A:597:ASN:ND2	2.08	0.44
1:A:341:TYR:CZ	1:A:375:LYS:HE2	2.53	0.44
1:A:479:LYS:O	1:A:484:LYS:NZ	2.51	0.44
1:A:181:TRP:N	1:A:211:GLN:HE22	2.01	0.43
1:A:261:GLU:HG3	1:A:262:ALA:N	2.33	0.43
1:A:66:LEU:O	1:A:70:VAL:HG23	2.19	0.42
1:A:404:LYS:HE3	1:A:423:ARG:HB2	2.02	0.42
1:A:73:GLU:OE2	1:A:112:ASN:ND2	2.48	0.42
1:A:52:PRO:HB2	1:A:193:PRO:HA	2.02	0.42
1:A:170:PHE:CZ	1:A:243:GLY:HA3	2.55	0.42
1:A:386:LYS:HA	1:A:390:GLN:HG3	2.02	0.41
1:A:71:ARG:O	1:A:75:GLU:HG3	2.20	0.41
1:A:93:SER:HA	3:A:958:HOH:O	2.19	0.41
1:A:196:LEU:HB3	1:A:197:PRO:HD2	2.02	0.41
1:A:372:ASN:ND2	1:A:375:LYS:H	2.17	0.41
1:A:551:LEU:HD13	1:A:619:ALA:HA	2.02	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	555/633 (88%)	536 (97%)	18 (3%)	1 (0%)	47 69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	23	GLU

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	513/564 (91%)	485 (94%)	28 (6%)	21 37

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	ASN
1	A	51	GLN
1	A	92	GLN
1	A	129	PHE
1	A	158	THR
1	A	208	LEU
1	A	213	ASN
1	A	238	ARG
1	A	261	GLU
1	A	270	THR
1	A	271	LEU
1	A	279	LEU
1	A	299	HIS
1	A	303	MSE
1	A	329	LYS
1	A	372	ASN
1	A	374	ASP
1	A	400	LYS
1	A	423	ARG
1	A	438	ASN
1	A	451	LYS
1	A	452	ILE
1	A	464	ASN
1	A	504	ARG
1	A	523	LEU
1	A	526	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	538	LEU
1	A	597	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (17) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	31	ASN
1	A	40	GLN
1	A	51	GLN
1	A	54	ASN
1	A	63	ASN
1	A	211	GLN
1	A	213	ASN
1	A	249	GLN
1	A	295	HIS
1	A	299	HIS
1	A	372	ASN
1	A	390	GLN
1	A	438	ASN
1	A	464	ASN
1	A	499	GLN
1	A	526	ASN
1	A	597	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

Of 1 ligands modelled in this entry, 1 is monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	556/633 (87%)	-0.03	27 (4%) 29 33	21, 34, 60, 76	0

All (27) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	328	VAL	6.6
1	A	94	GLY	4.9
1	A	23	GLU	4.7
1	A	100	SER	4.0
1	A	636	ASP	3.9
1	A	586	ASP	3.7
1	A	459	ASN	3.6
1	A	22	GLU	3.6
1	A	455	ASP	3.5
1	A	456	ALA	3.2
1	A	277	ASN	3.0
1	A	24	ASP	3.0
1	A	457	SER	2.9
1	A	635	CYS	2.9
1	A	527	PRO	2.8
1	A	105	ASP	2.8
1	A	101	GLY	2.7
1	A	332	GLU	2.5
1	A	103	ASP	2.5
1	A	106	ASP	2.5
1	A	560	ASP	2.4
1	A	269	SER	2.3
1	A	273	GLN	2.2
1	A	257	PRO	2.1
1	A	361	GLU	2.1
1	A	274	ILE	2.1
1	A	452	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)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	FE	A	701	1/1	0.94	0.06	45,45,45,45	0

6.5 Other polymers (i)

There are no such residues in this entry.