Full wwPDB X-ray Structure Validation Report ⓘ

Apr 21, 2025 – 06:09 PM EDT

PDB ID	:	9O0H / pdb_00009o0h
Title	:	The ubiquitin-associated domain of human thirty-eight negative kinase 1, fused to the 3TEL crystallization chaperone via a 2-glycine linker
Authors	:	Samarawickrama, P.; Pedroza Romo, M.J.; Ludlow, K.; Keliiliki, A.; Doukov, T.; Moody, J.D.
Deposited on	:	2025-04-02
Resolution	:	2.24 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:

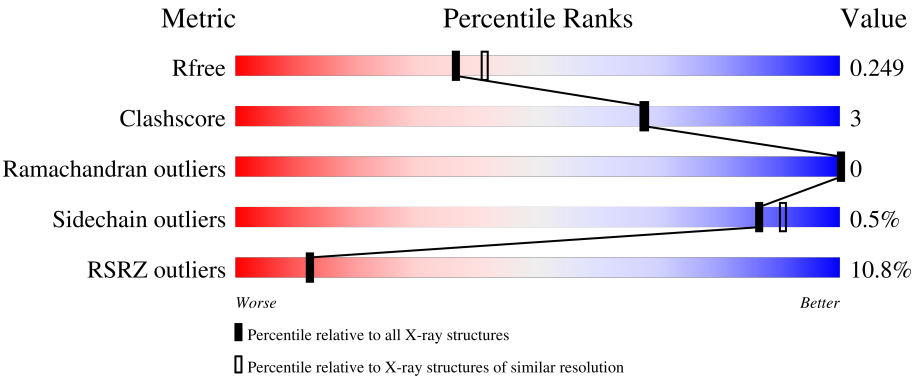
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

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 2.24 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	3139 (2.26-2.22)
Clashscore	180529	3381 (2.26-2.22)
Ramachandran outliers	177936	3334 (2.26-2.22)
Sidechain outliers	177891	3335 (2.26-2.22)
RSRZ outliers	164620	3138 (2.26-2.22)

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	B	325	10%, 86%, 8%, 6%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2470 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 Transcription factor ETV6,Non-receptor tyrosine-protein kinase TNK1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	306	Total	C	N	O	S	0	0	0
			2381	1550	415	412	4			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	GLY	-	expression tag	UNP P41212
B	67	GLU	VAL	engineered mutation	UNP P41212
B	80	PRO	-	linker	UNP P41212
B	81	GLY	-	linker	UNP P41212
B	82	GLY	-	linker	UNP P41212
B	83	GLY	-	linker	UNP P41212
B	84	GLY	-	linker	UNP P41212
B	85	SER	-	linker	UNP P41212
B	86	THR	-	linker	UNP P41212
B	165	PRO	-	linker	UNP P41212
B	166	GLY	-	linker	UNP P41212
B	167	GLY	-	linker	UNP P41212
B	168	GLY	-	linker	UNP P41212
B	169	GLY	-	linker	UNP P41212
B	170	SER	-	linker	UNP P41212
B	171	THR	-	linker	UNP P41212
B	247	GLY	-	linker	UNP P41212
B	248	GLY	-	linker	UNP P41212
B	269	ALA	CYS	engineered mutation	UNP Q13470
B	303	ALA	CYS	engineered mutation	UNP Q13470

- Molecule 2 is water.

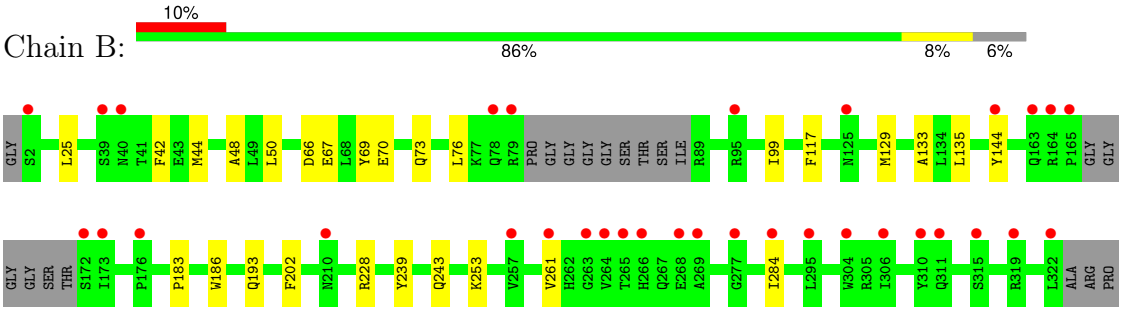
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	89	Total	O	0	0
			89	89		

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: Transcription factor ETV6,Non-receptor tyrosine-protein kinase TNK1

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	55.15Å 65.54Å 112.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.20 – 2.24 42.20 – 2.24	Depositor EDS
% Data completeness (in resolution range)	93.7 (42.20-2.24) 82.3 (42.20-2.24)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	0.62 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.224 , 0.249 0.224 , 0.249	Depositor DCC
R_{free} test set	1056 reflections (5.18%)	wwPDB-VP
Wilson B-factor (Å^2)	32.1	Xtriage
Anisotropy	0.687	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 49.2	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	2470	wwPDB-VP
Average B, all atoms (Å^2)	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.36% 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	B	0.24	0/2444	0.45	0/3321

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	B	2381	0	2225	16	0
2	B	89	0	0	2	0
All	All	2470	0	2225	16	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 3.

All (16) 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:133:ALA:HB2	1:B:202:PHE:HB3	1.87	0.57
1:B:48:ALA:HB2	1:B:117:PHE:HB3	1.86	0.56
1:B:50:LEU:HD22	1:B:76:LEU:HD11	1.93	0.51
1:B:66:ASP:O	1:B:70:GLU:HG2	2.14	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:129:MET:HB2	1:B:133:ALA:HB3	1.96	0.48
1:B:99:ILE:HD13	1:B:135:LEU:HD12	1.95	0.47
1:B:261:VAL:HG23	1:B:284:ILE:HG23	1.96	0.47
1:B:243:GLN:NE2	2:B:403:HOH:O	2.42	0.46
1:B:66:ASP:OD1	1:B:67:GLU:N	2.50	0.44
1:B:253:LYS:NZ	2:B:408:HOH:O	2.52	0.42
1:B:44:MET:HB2	1:B:48:ALA:HB3	2.02	0.42
1:B:183:PRO:HA	1:B:186:TRP:CD2	2.55	0.42
1:B:69:TYR:O	1:B:73:GLN:HG2	2.19	0.42
1:B:144:TYR:CZ	1:B:228:ARG:HD3	2.55	0.41
1:B:25:LEU:HD11	1:B:42:PHE:HB2	2.02	0.41
1:B:239:TYR:CZ	1:B:243:GLN:HG3	2.56	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	B	300/325 (92%)	295 (98%)	5 (2%)	0	100 100

There are no Ramachandran outliers to report.

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	B	222/283 (78%)	221 (100%)	1 (0%)	86 90

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	193	GLN

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 oligosaccharides 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	B	306/325 (94%)	0.65	33 (10%)	12	12	32, 46, 84, 112	0

All (33) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	39	SER	4.5
1	B	165	PRO	4.2
1	B	40	ASN	3.7
1	B	266	HIS	3.6
1	B	277	GLY	3.6
1	B	265	THR	3.6
1	B	125	ASN	3.5
1	B	164	ARG	3.3
1	B	2	SER	3.3
1	B	176	PRO	3.3
1	B	295	LEU	3.3
1	B	264	VAL	3.3
1	B	144	TYR	3.1
1	B	172	SER	3.1
1	B	210	ASN	2.8
1	B	310	TYR	2.8
1	B	304	TRP	2.7
1	B	319	ARG	2.6
1	B	268	GLU	2.5
1	B	311	GLN	2.5
1	B	95	ARG	2.5
1	B	261	VAL	2.4
1	B	257	VAL	2.3
1	B	263	GLY	2.3
1	B	78	GLN	2.2
1	B	173	ILE	2.2
1	B	163	GLN	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	322	LEU	2.2
1	B	284	ILE	2.1
1	B	315	SER	2.1
1	B	269	ALA	2.1
1	B	79	ARG	2.0
1	B	306	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.