wwPDB X-ray Structure Validation Summary Report (i)

Feb 5, 2024 – 01:46 AM EST

PDB ID	:	1T9H
Title	:	The crystal structure of YloQ, a circularly permuted GTPase.
Authors	:	Levdikov, V.M.; Blagova, E.V.; Brannigan, J.A.; Cladiere, L.; Antson, A.A.; Isupov, M.N.; Seror, S.J.; Wilkinson, A.J.
Deposited on	:	2004-05-17
Resolution	:	1.60 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

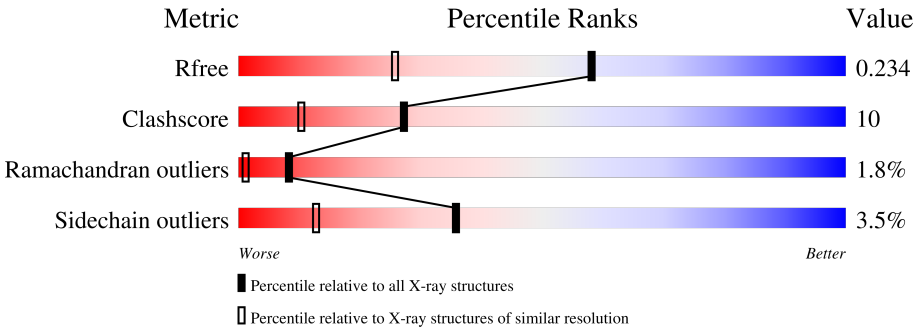
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.60 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	3398 (1.60-1.60)
Clashscore	141614	3665 (1.60-1.60)
Ramachandran outliers	138981	3564 (1.60-1.60)
Sidechain outliers	138945	3563 (1.60-1.60)

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	307	78% 12% • • 7%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	IUM	A	401	-	-	X	-
2	IUM	A	402	-	-	X	-
2	IUM	A	403	-	-	X	-
2	IUM	A	404	-	-	X	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	IUM	A	405	-	-	X	-
2	IUM	A	406	-	-	X	-
2	IUM	A	407	-	-	X	-
2	IUM	A	408	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 2666 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.

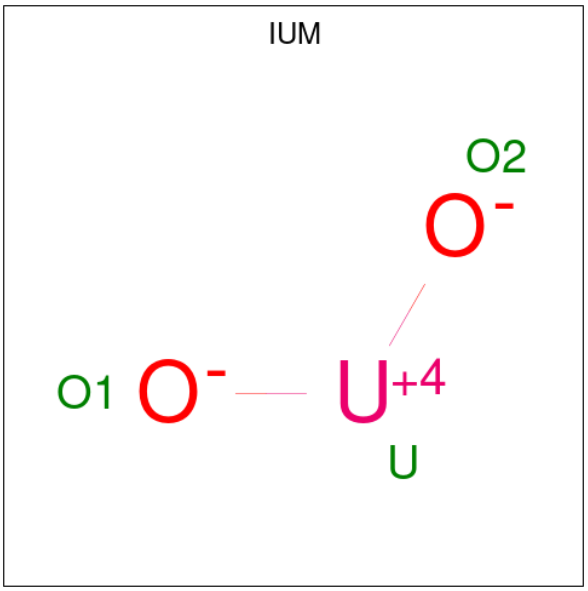
- Molecule 1 is a protein called Probable GTPase engC.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	287	2322	1464	395	451	12	0	5	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-8	MET	-	initiating methionine	UNP O34530
A	-7	ALA	-	expression tag	UNP O34530
A	-6	ARG	-	expression tag	UNP O34530
A	-5	HIS	-	expression tag	UNP O34530
A	-4	HIS	-	expression tag	UNP O34530
A	-3	HIS	-	expression tag	UNP O34530
A	-2	HIS	-	expression tag	UNP O34530
A	-1	HIS	-	expression tag	UNP O34530
A	0	HIS	-	expression tag	UNP O34530

- Molecule 2 is URANYL (VI) ION (three-letter code: IUM) (formula: O_2U).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O U 3 2 1	0	0
2	A	1	Total O U 3 2 1	0	0
2	A	1	Total O U 3 2 1	0	0
2	A	1	Total O U 3 2 1	0	0
2	A	1	Total O U 3 2 1	0	0
2	A	1	Total O U 3 2 1	0	0
2	A	1	Total O U 3 2 1	0	0
2	A	1	Total O U 3 2 1	0	0

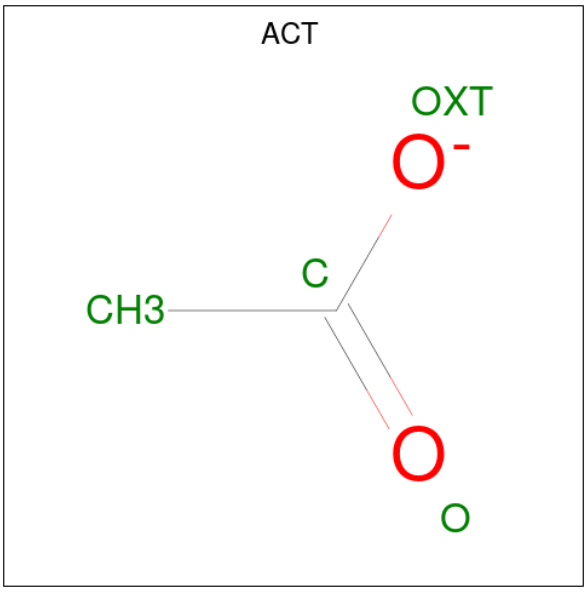
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Zn 1 1	0	0

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	4	Total Ca 4 4	0	0

- Molecule 5 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 4 2 2	0	0
5	A	1	Total C O 4 2 2	0	0
5	A	1	Total C O 4 2 2	0	0

- Molecule 6 is water.

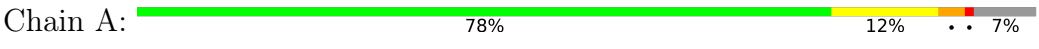
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	303	Total O 303 303	0	0

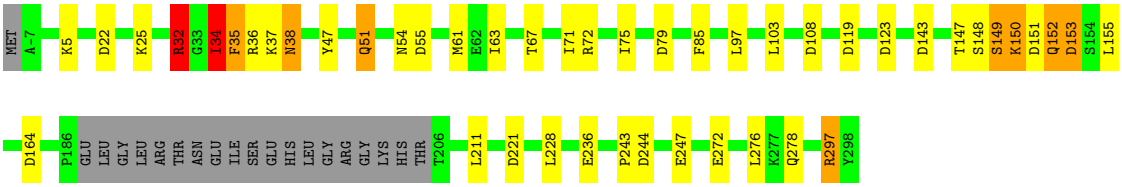
3 Residue-property plots (i)

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: Probable GTPase engC

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	47.97Å 47.97Å 279.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.50 – 1.60 20.03 – 1.60	Depositor EDS
% Data completeness (in resolution range)	100.0 (20.50-1.60) 99.4 (20.03-1.60)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.03 (at 1.60Å)	Xtriage
Refinement program	REFMAC 5.1.9999	Depositor
R, R_{free}	0.144 , 0.178 0.208 , 0.234	Depositor DCC
R_{free} test set	2243 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	14.3	Xtriage
Anisotropy	0.269	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 57.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.45$, $< L^2 > = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	2666	wwPDB-VP
Average B, all atoms (Å^2)	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.88% 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: IUM, ZN, ACT, CA

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	0.44	0/2394	0.81	12/3231 (0.4%)

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.

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	32	ARG	NE-CZ-NH1	6.91	123.75	120.30
1	A	119	ASP	CB-CG-OD2	6.25	123.92	118.30
1	A	108	ASP	CB-CG-OD1	6.14	123.83	118.30
1	A	153	ASP	CB-CG-OD2	5.99	123.69	118.30
1	A	221	ASP	CB-CG-OD2	5.48	123.23	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	34	ILE	Peptide

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	2322	0	2274	31	1
2	A	24	0	0	16	0
3	A	1	0	0	0	0
4	A	4	0	0	0	0
5	A	12	0	8	0	0
6	A	303	0	0	2	2
All	All	2666	0	2282	47	2

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 10.

The worst 5 of 47 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:152:GLN:O	1:A:155:LEU:HD12	1.75	0.87
1:A:147:THR:HB	1:A:152:GLN:HB2	1.73	0.70
1:A:150:LYS:HG2	1:A:150:LYS:O	1.91	0.70
1:A:34:ILE:O	1:A:36:ARG:N	2.22	0.70
1:A:147:THR:CB	1:A:152:GLN:HB2	2.22	0.70

All (2) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:A:433:HOH:O	6:A:552:HOH:O[1_545]	1.93	0.27
1:A:272:GLU:OE2	6:A:433:HOH:O[1_565]	2.12	0.08

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	288/307 (94%)	275 (96%)	8 (3%)	5 (2%)	9 1

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	34	ILE
1	A	35	PHE
1	A	38	ASN
1	A	149	SER
1	A	150	LYS

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	261/273 (96%)	252 (97%)	9 (3%)	37 13

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	153	ASP
1	A	297	ARG
1	A	51	GLN
1	A	54	ASN
1	A	149	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	54	ASN
1	A	152	GLN

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 16 ligands modelled in this entry, 5 are monoatomic - leaving 11 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
2	IUM	A	406	-	0,2,2	-	-	-		
2	IUM	A	402	-	0,2,2	-	-	-		
2	IUM	A	401	-	0,2,2	-	-	-		
2	IUM	A	405	-	0,2,2	-	-	-		
2	IUM	A	407	-	0,2,2	-	-	-		
5	ACT	A	416	-	3,3,3	0.99	0	3,3,3	1.57	1 (33%)
2	IUM	A	404	-	0,2,2	-	-	-		
2	IUM	A	403	-	0,2,2	-	-	-		
5	ACT	A	418	-	3,3,3	0.72	0	3,3,3	1.35	0
2	IUM	A	408	-	0,2,2	-	-	-		
5	ACT	A	417	-	3,3,3	0.84	0	3,3,3	0.81	0

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
5	A	416	ACT	OXT-C-O	-2.11	114.29	122.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	406	IUM	2	0
2	A	402	IUM	2	0
2	A	401	IUM	2	0
2	A	405	IUM	2	0
2	A	407	IUM	2	0
2	A	404	IUM	2	0
2	A	403	IUM	2	0
2	A	408	IUM	2	0

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.