WORLDWIDE PDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 17, 2023 – 06:57 AM EDT

PDB ID	:	2DWQ
Title	:	Thermus thermophilus YchF GTP-binding protein
Authors	:	Kukimoto-Niino, M.; Murayama, K.; Shorouzu, M.; Kuramitsu, S.; Yokoyama, S.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)
Deposited on	:	2006-08-16
Resolution	:	2.95 Å(reported)

The following versions of software and data (see references ⓘ) were used in the production of this report:

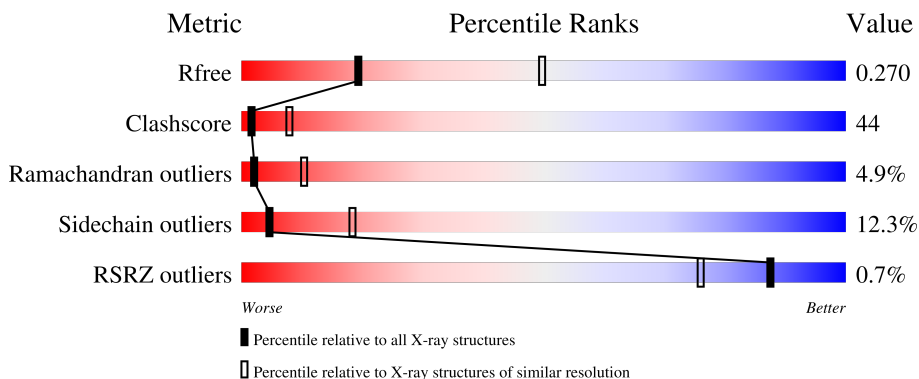|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.95 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		0.270
Clashscore		44
Ramachandran outliers		4.9%
Sidechain outliers		12.3%
RSRZ outliers		0.7%

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}	130704	3104 (3.00-2.92)
Clashscore	141614	3462 (3.00-2.92)
Ramachandran outliers	138981	3340 (3.00-2.92)
Sidechain outliers	138945	3343 (3.00-2.92)
RSRZ outliers	127900	2986 (3.00-2.92)

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	368	37% 49% 10% •
1	B	368	.% 29% 51% 11% • 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5410 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 GTP-binding protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	354	Total	C	N	O	S	0	0	0
			2748	1739	496	510	3			
1	B	337	Total	C	N	O	S	0	0	0
			2631	1668	475	486	2			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	22	Total	O	0	0
			22	22		
2	B	9	Total	O	0	0
			9	9		

3 Residue-property plots ⓘ

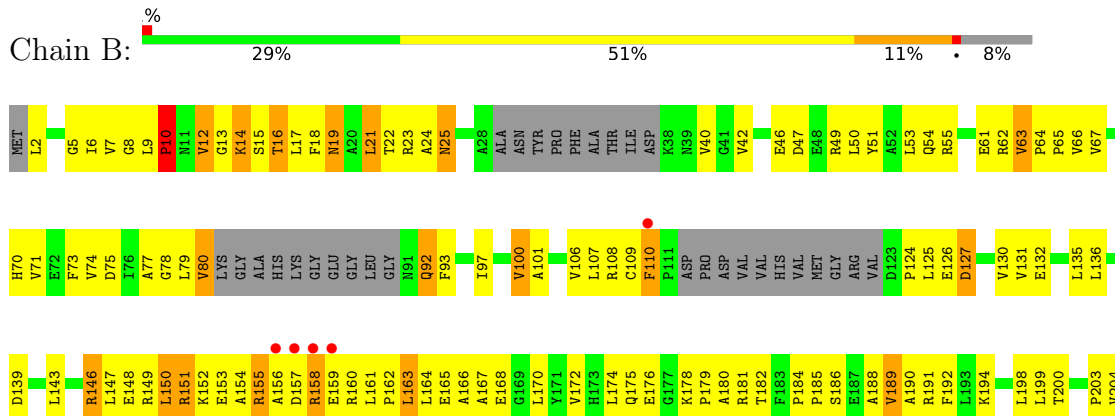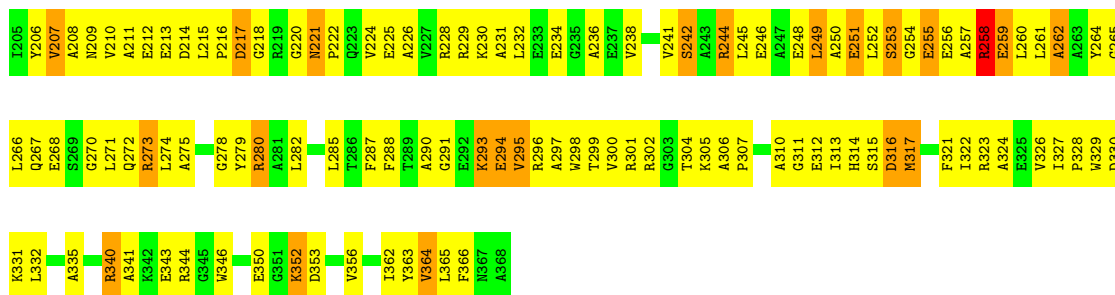These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: GTP-binding protein

Chain A:

- Molecule 1: GTP-binding protein

Chain B:

I205 Y206 V207 A208 N209 V210 A211 E212 E213 D214 L215 P216 D217 G218 R219 G220 N221 P222 Q223 V224 E225 A226 V227 R228 R229 K230 L232 E233 E234 G235 A236 E237 V238 V241 S242 A243 R244 L245 E246 A247 E248 L249 A250 E251 L252 S253 G254 E255 E256 A257 R258 E259 L260 L261 A262 A263 Y264 G265

L266 Q267 E268 S269 G270 L271 Q272 R273 L274 A275 G278 Y279 R280 A281 L282 L285 T286 F287 F288 T289 A290 G291 E292 K293 E294 V295 R296 A297 W298 T299 V300 R301 R302 G303 T304 K305 A306 P307 A310 G311 E312 I313 H314 S315 D316 M317 F321 I322 R323 A324 E325 V326 I327 P328 W329 D330

K331 L332 A335 R340 A341 K342 E343 R344 G345 W346 E350 G351 K352 D353 V356 I362 Y363 V364 L365 F366 N367 A368

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	52.94Å 54.36Å 92.58Å 80.11° 76.46° 73.12°	Depositor
Resolution (Å)	37.04 – 2.95 37.03 – 2.95	Depositor EDS
% Data completeness (in resolution range)	94.2 (37.04-2.95) 94.2 (37.03-2.95)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.12	Depositor
$<I/\sigma(I)>$ [1]	3.36 (at 2.95Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.208 , 0.281 0.200 , 0.270	Depositor DCC
R_{free} test set	1872 reflections (9.86%)	wwPDB-VP
Wilson B-factor (Å^2)	62.8	Xtriage
Anisotropy	0.507	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.29 , 65.2	EDS
L-test for twinning[2]	$<|L|> = 0.50, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	5410	wwPDB-VP
Average B, all atoms (Å^2)	60.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.69% 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	0.95	1/2795 (0.0%)	1.09	4/3787 (0.1%)
1	B	0.76	0/2675	0.99	1/3625 (0.0%)
All	All	0.86	1/5470 (0.0%)	1.04	5/7412 (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 (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	144	GLU	CG-CD	5.05	1.59	1.51

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	84	HIS	N-CA-C	6.34	128.13	111.00
1	B	280	ARG	NE-CZ-NH1	-6.22	117.19	120.30
1	A	86	GLY	N-CA-C	-5.79	98.61	113.10
1	A	100	VAL	CG1-CB-CG2	-5.36	102.32	110.90
1	A	62	ARG	N-CA-C	-5.00	97.50	111.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	51	TYR	Sidechain

WORLDWIDE wwPDB PROTEIN DATA BANK

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	2748	0	2784	220	0
1	B	2631	0	2666	274	0
2	A	22	0	0	6	0
2	B	9	0	0	2	0
All	All	5410	0	5450	481	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 44.

The worst 5 of 481 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:23:ARG:HE	1:B:251:GLU:HG2	0.95	1.08
1:A:16:THR:HB	1:A:243:ALA:HB1	1.37	1.05
1:A:302:ARG:HH11	1:A:302:ARG:HB2	1.18	1.04
1:A:244:ARG:HG3	1:A:244:ARG:HH11	1.24	1.02
1:B:258:ARG:HE	1:B:258:ARG:HA	1.24	1.02

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	348/368 (95%)	286 (82%)	48 (14%)	14 (4%)	3	14
1	B	329/368 (89%)	264 (80%)	46 (14%)	19 (6%)	1	7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	677/736 (92%)	550 (81%)	94 (14%)	33 (5%)	2	10

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	81	LYS
1	A	82	GLY
1	A	189	VAL
1	A	352	LYS
1	B	10	PRO

5.3.2   Protein sidechains ⓘ

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	277/289 (96%)	246 (89%)	31 (11%)	6	22
1	B	266/289 (92%)	230 (86%)	36 (14%)	4	15
All	All	543/578 (94%)	476 (88%)	67 (12%)	4	18

5 of 67 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	279	TYR
1	B	293	LYS
1	B	352	LYS
1	A	288	PHE
1	A	287	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	11	ASN
1	B	19	ASN
1	B	54	GLN

Continued on next page...

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Continued from previous page...

Mol	Chain	Res	Type
1	A	357	GLN
1	A	92	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 ⓘ

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 ⓘ

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(Å²)	Q<0.9
1	A	354/368 (96%)	-0.59	0	100	100	14, 49, 90, 119	1 (0%)
1	B	337/368 (91%)	-0.31	5 (1%)	73	57	16, 65, 108, 145	1 (0%)
All	All	691/736 (93%)	-0.45	5 (0%)	87	76	14, 56, 100, 145	2 (0%)

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	159	GLU	2.9
1	B	158	ARG	2.7
1	B	156	ALA	2.4
1	B	157	ASP	2.2
1	B	110	PHE	2.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no monosaccharides in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.