wwPDB X-ray Structure Validation Summary Report (i)

Oct 22, 2023 – 09:40 PM EDT

PDB ID	:	3FHO
Title	:	Structure of S. pombe Dbp5
Authors	:	Cheng, Z.; Song, H.
Deposited on	:	2008-12-09
Resolution	:	2.80 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 (1)) were used in the production of this report:

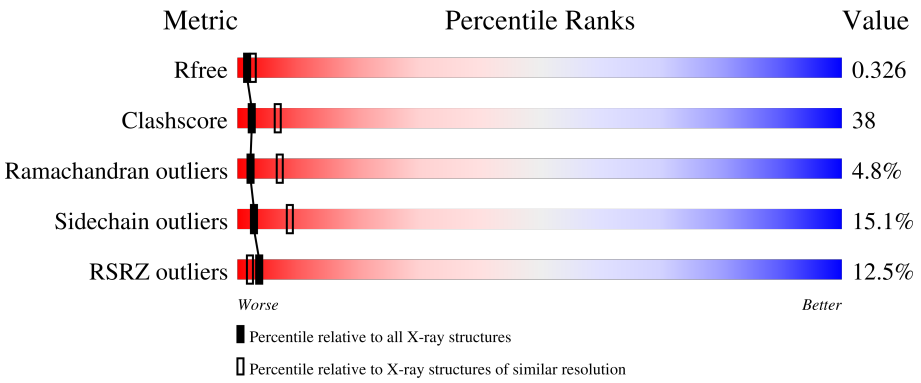
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation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

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	508	8% 26% 28% 6% 40%
1	B	508	7% 28% 27% 5% 40%

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 4771 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 ATP-dependent RNA helicase dbp5.

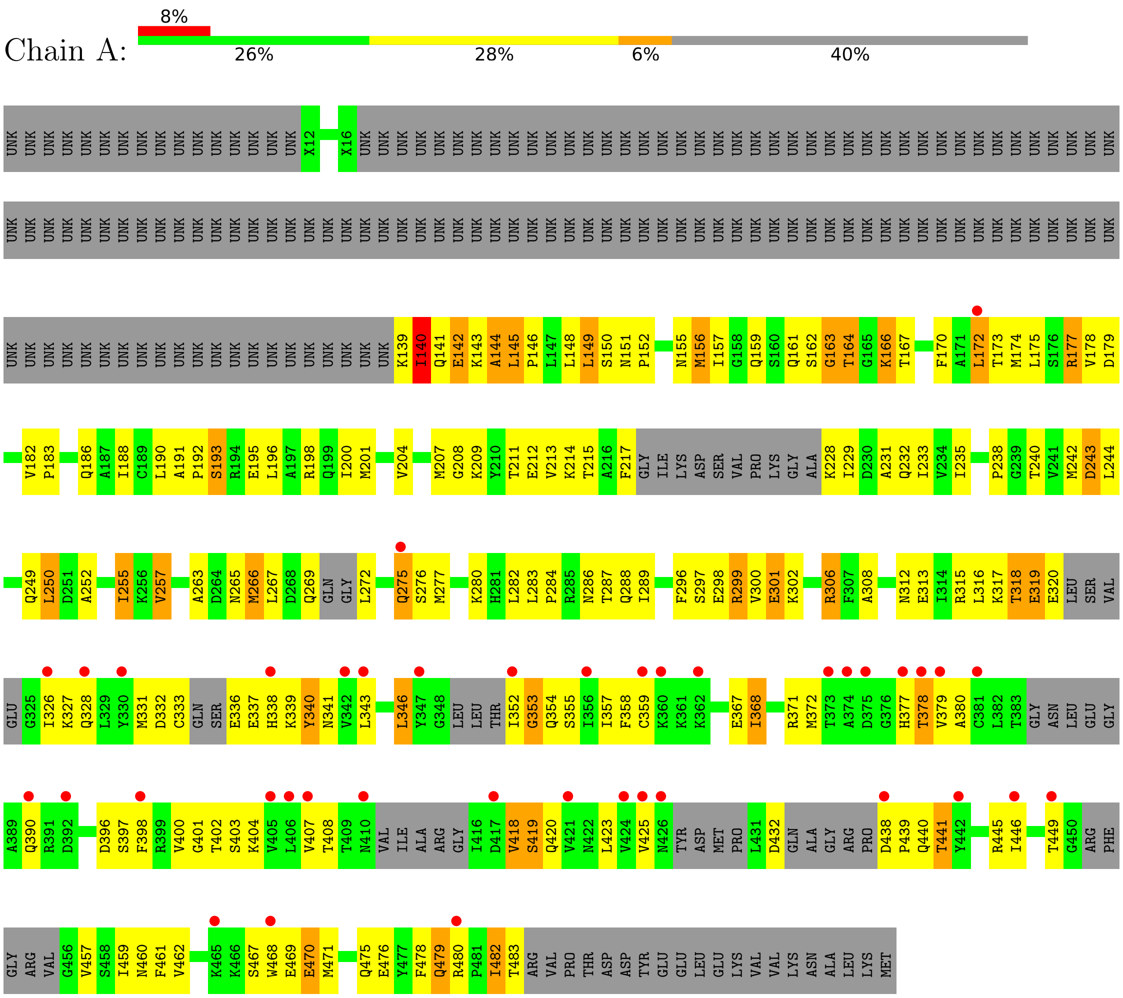
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	305	2381	1493	420	452	16	0	0	0
			Total	C	N	O	S			
1	B	306	2390	1497	422	455	16	0	0	0

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.

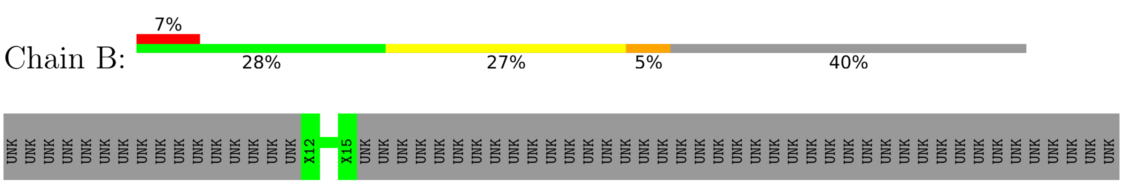
- Molecule 1: ATP-dependent RNA helicase dbp5

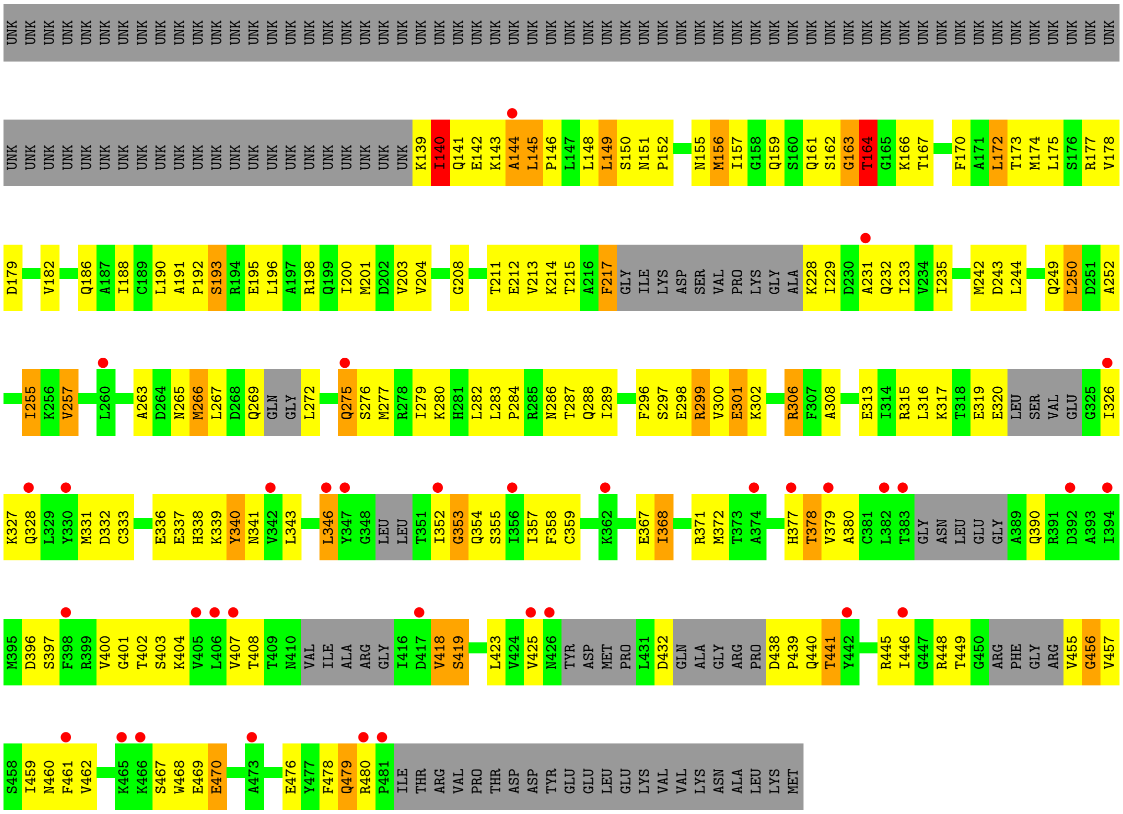
Chain A:



- Molecule 1: ATP-dependent RNA helicase dbp5

Chain B:




UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK
UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK UNK K139 I140 Q141 E142 K143 A144 L145 P146 L147 L148 L149 S150 N151 P152 N155 M156 I157 G158 Q159 S160 Q161 S162 G163 T164 G165 K166 T167 F170 A171 L172 T173 M174 L175 S176 R177 V178
D179 V182 Q186 A187 I188 C189 L190 A191 P192 S193 R194 E195 L196 A197 R198 Q199 I200 M201 D202 V203 V204 G208 T211 E212 V213 K214 T215 A216 F217 GLY ILE LYS ASP SER VAL PRO LYS GLY ALA K228 I229 D230 A231 Q232 I233 V234 I235 M242 D243 L244 Q249 L250 D251 A252
I255 K256 V257 L260 A263 D264 N265 M266 L267 D268 Q269 GLN GLY L272 Q275 S276 M277 R278 I279 K280 H281 L282 L283 P284 R285 N286 T287 Q288 I289 F296 S297 E298 R299 V300 E301 K302 R306 F307 A308 E313 I314 R315 L316 K317 T318 E319 E320 LEU SER VAL GLU G325 I326
K327 Q328 L329 Y330 M331 D332 C333 E336 E337 H338 K339 Y340 N341 V342 L343 L346 Y347 G348 LEU LEU T351 I352 G353 Q354 S355 I356 I357 F358 C359 K362 E367 I368 R371 M372 T373 A374 H377 T378 V379 A380 C381 L382 T383 GLY ASN LEU GLU GLY A389 Q390 R391 D392 A393 I394
M395 D396 S397 F398 R399 V400 G401 T402 S403 K404 V405 L406 V407 T408 T409 N410 VAL ILE ALA ARG GLY I416 D417 V418 S419 L423 V424 V425 N426 TYR ASP MET PRO L431 D432 GLN ALA GLY ARG PRO D438 P439 Q440 T441 Y442 R445 I446 G447 R448 T449 G450 ARG PHE GLY ARG V455 G456 V457
S458 I459 N460 F461 V462 K465 K466 S467 W468 E469 E470 A473 E476 Y477 F478 Q479 R480 P481 ILE THR ARG VAL PRO THR ASP ASP TYR GLU GLU LEU GLU LYS VAL VAL LYS ASN ALA LEU LYS MET

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	108.74Å 144.04Å 79.14Å 90.00° 89.96° 90.00°	Depositor
Resolution (Å)	20.00 – 2.80 44.80 – 2.80	Depositor EDS
% Data completeness (in resolution range)	98.6 (20.00-2.80) 98.2 (44.80-2.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.65 (at 2.81Å)	Xtriage
Refinement program	REFMAC 5.4.0077	Depositor
R, R_{free}	0.290 , 0.327 0.291 , 0.326	Depositor DCC
R_{free} test set	1492 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	62.0	Xtriage
Anisotropy	0.953	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 69.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.458 for -h,-k,l	Xtriage
F_o,F_c correlation	0.89	EDS
Total number of atoms	4771	wwPDB-VP
Average B, all atoms (Å^2)	16.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% 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 (i)

5.1 Standard geometry (i)

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	0.52	0/2380	0.69	0/3196
1	B	0.52	0/2395	0.68	0/3218
All	All	0.52	0/4775	0.69	0/6414

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	B	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	164	THR	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	2381	0	2392	191	0
1	B	2390	0	2402	176	0
All	All	4771	0	4794	364	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 38.

The worst 5 of 364 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:140:ILE:HG23	1:A:145:LEU:HD12	1.20	1.18
1:B:340:TYR:CE2	1:B:367:GLU:OE2	2.02	1.13
1:A:201:MET:CE	1:A:217:PHE:HB2	1.80	1.11
1:A:340:TYR:CE2	1:A:367:GLU:OE2	2.04	1.10
1:A:482:ILE:HD13	1:A:482:ILE:H	1.12	1.10

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	278/508 (55%)	231 (83%)	33 (12%)	14 (5%)	2 6
1	B	282/508 (56%)	238 (84%)	31 (11%)	13 (5%)	2 7
All	All	560/1016 (55%)	469 (84%)	64 (11%)	27 (5%)	2 7

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	145	LEU
1	A	319	GLU
1	B	145	LEU
1	A	390	GLN
1	B	150	SER

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	260/315 (82%)	219 (84%)	41 (16%)	2 8
1	B	262/315 (83%)	224 (86%)	38 (14%)	3 9
All	All	522/630 (83%)	443 (85%)	79 (15%)	3 9

5 of 79 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	275	GLN
1	B	408	THR
1	B	282	LEU
1	B	337	GLU
1	B	460	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 15 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	151	ASN
1	B	426	ASN
1	B	159	GLN
1	B	460	ASN
1	B	275	GLN

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 monosaccharides 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 (i)

6.1 Protein, DNA and RNA chains (i)

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(Å^2)	Q<0.9
1	A	300/508 (59%)	1.02	40 (13%)	3	2	5, 16, 21, 60	0
1	B	302/508 (59%)	0.92	35 (11%)	4	2	5, 16, 21, 64	0
All	All	602/1016 (59%)	0.97	75 (12%)	3	2	5, 16, 21, 64	0

The worst 5 of 75 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	352	ILE	7.4
1	A	417	ASP	6.7
1	A	377	HIS	6.0
1	B	326	ILE	5.9
1	A	379	VAL	5.9

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.