



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 13, 2017 – 04:05 am GMT

PDB ID : 2P6N  
Title : Human DEAD-box RNA helicase DDX41, helicase domain  
Authors : Karlberg, T.; Ogg, D.; Arrowsmith, C.H.; Berglund, H.; Busam, R.D.; Collins, R.; Dahlgren, L.G.; Edwards, A.; Flodin, S.; Flores, A.; Graslund, S.; Hallberg, B.M.; Hammarstrom, M.; Johansson, I.; Kotenyova, T.; Lehtio, L.; Moche, M.; Nordlund, P.; Nyman, T.; Persson, C.; Sagemark, J.; Stenmark, P.; Sundstrom, M.; Thorsell, A.G.; Van Den Berg, S.; Weigelt, J.; Holmberg-Schiavone, L.; Structural Genomics Consortium (SGC)  
Deposited on : 2007-03-19  
Resolution : 2.60 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

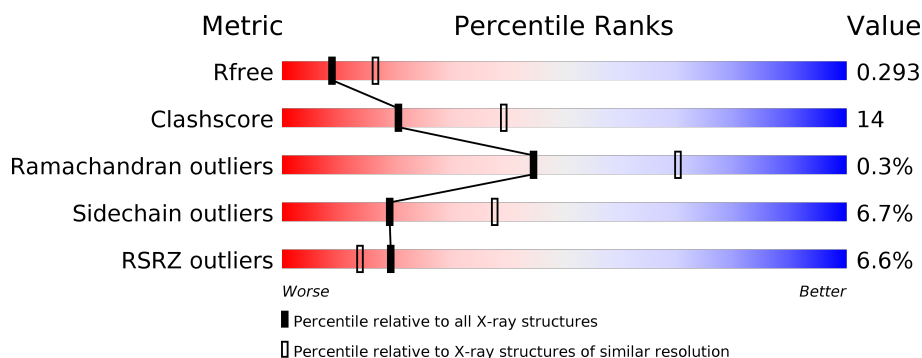
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	191	<div> <div>6%</div> <div> <div></div> <div>59%</div> <div>25%</div> <div>•</div> <div>16%</div> </div> </div>
1	B	191	<div> <div>5%</div> <div> <div></div> <div>59%</div> <div>21%</div> <div>•</div> <div>17%</div> </div> </div>

## 2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2481 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 DDX41.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	160	Total	C	N	O	S	0	0	0
			1252	802	206	237	7			
1	B	158	Total	C	N	O	S	0	0	0
			1229	788	202	233	6			

There are 48 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	379	MET	-	CLONING ARTIFACT	UNP Q9UJV9
A	380	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
A	381	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
A	382	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
A	383	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
A	384	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
A	385	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
A	386	SER	-	CLONING ARTIFACT	UNP Q9UJV9
A	387	SER	-	CLONING ARTIFACT	UNP Q9UJV9
A	388	GLY	-	CLONING ARTIFACT	UNP Q9UJV9
A	389	VAL	-	CLONING ARTIFACT	UNP Q9UJV9
A	390	ASP	-	CLONING ARTIFACT	UNP Q9UJV9
A	391	LEU	-	CLONING ARTIFACT	UNP Q9UJV9
A	392	GLY	-	CLONING ARTIFACT	UNP Q9UJV9
A	393	THR	-	CLONING ARTIFACT	UNP Q9UJV9
A	394	GLU	-	CLONING ARTIFACT	UNP Q9UJV9
A	395	ASN	-	CLONING ARTIFACT	UNP Q9UJV9
A	396	LEU	-	CLONING ARTIFACT	UNP Q9UJV9
A	397	TYR	-	CLONING ARTIFACT	UNP Q9UJV9
A	398	PHE	-	CLONING ARTIFACT	UNP Q9UJV9
A	399	GLN	-	CLONING ARTIFACT	UNP Q9UJV9
A	400	SER	-	CLONING ARTIFACT	UNP Q9UJV9
A	401	MET	-	CLONING ARTIFACT	UNP Q9UJV9
A	525	CYS	ARG	ENGINEERED	UNP Q9UJV9
B	379	MET	-	CLONING ARTIFACT	UNP Q9UJV9

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	380	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
B	381	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
B	382	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
B	383	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
B	384	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
B	385	HIS	-	CLONING ARTIFACT	UNP Q9UJV9
B	386	SER	-	CLONING ARTIFACT	UNP Q9UJV9
B	387	SER	-	CLONING ARTIFACT	UNP Q9UJV9
B	388	GLY	-	CLONING ARTIFACT	UNP Q9UJV9
B	389	VAL	-	CLONING ARTIFACT	UNP Q9UJV9
B	390	ASP	-	CLONING ARTIFACT	UNP Q9UJV9
B	391	LEU	-	CLONING ARTIFACT	UNP Q9UJV9
B	392	GLY	-	CLONING ARTIFACT	UNP Q9UJV9
B	393	THR	-	CLONING ARTIFACT	UNP Q9UJV9
B	394	GLU	-	CLONING ARTIFACT	UNP Q9UJV9
B	395	ASN	-	CLONING ARTIFACT	UNP Q9UJV9
B	396	LEU	-	CLONING ARTIFACT	UNP Q9UJV9
B	397	TYR	-	CLONING ARTIFACT	UNP Q9UJV9
B	398	PHE	-	CLONING ARTIFACT	UNP Q9UJV9
B	399	GLN	-	CLONING ARTIFACT	UNP Q9UJV9
B	400	SER	-	CLONING ARTIFACT	UNP Q9UJV9
B	401	MET	-	CLONING ARTIFACT	UNP Q9UJV9
B	525	CYS	ARG	ENGINEERED	UNP Q9UJV9



- Molecule 1: ATP-dependent RNA helicase DDX41



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.01Å 68.01Å 305.60Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.72 – 2.60 29.72 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (29.72-2.60) 100.0 (29.72-2.50)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	0.03	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	7.57 (at 2.51Å)	Xtriage
Refinement program	REFMAC 5.3.0032	Depositor
R, $R_{free}$	0.243 , 0.294 0.247 , 0.293	Depositor DCC
$R_{free}$ test set	693 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	43.1	Xtriage
Anisotropy	0.476	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 70.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.38$ , $\langle L^2 \rangle = 0.21$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2481	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 49.87 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 7.0118e-05. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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	0.75	0/1271	0.77	0/1718
1	B	0.72	0/1247	0.76	0/1686
All	All	0.74	0/2518	0.76	0/3404

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	A	1252	0	1284	40	0
1	B	1229	0	1259	32	0
All	All	2481	0	2543	70	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 14.

The worst 5 of 70 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:462:ILE:HG22	1:B:492:ALA:HB1	1.43	0.99
1:A:546:MET:HE2	1:A:566:LEU:HD23	1.45	0.98

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:425:LEU:HD11	1:B:455:LYS:HG2	1.45	0.98
1:A:431:THR:HG22	1:A:432:PRO:O	1.65	0.96
1:B:407:ASP:OD2	1:B:529:THR:HG22	1.75	0.85

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	156/191 (82%)	146 (94%)	10 (6%)	0	100	100
1	B	154/191 (81%)	148 (96%)	5 (3%)	1 (1%)	28	53
All	All	310/382 (81%)	294 (95%)	15 (5%)	1 (0%)	44	70

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	562	VAL

### 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	137/161 (85%)	130 (95%)	7 (5%)	28	52
1	B	133/161 (83%)	122 (92%)	11 (8%)	13	25

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
All	All	270/322 (84%)	252 (93%)	18 (7%)	19	38

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	425	LEU
1	B	431	THR
1	B	526	SER
1	A	564	GLN
1	B	413	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	410	GLN
1	A	567	HIS
1	B	410	GLN
1	B	449	HIS

### 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 carbohydrates 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 ⓘ

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	160/191 (83%)	0.47	11 (6%) 18 12	51, 57, 65, 67	0
1	B	158/191 (82%)	0.38	10 (6%) 21 15	51, 57, 64, 66	0
All	All	318/382 (83%)	0.42	21 (6%) 19 14	51, 57, 64, 67	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	525	CYS	11.0
1	A	406	LEU	6.8
1	A	407	ASP	4.4
1	A	500	ALA	4.1
1	B	476	GLU	3.8

### 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 carbohydrates 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.