



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 23, 2024 – 01:11 AM EDT

PDB ID : 6I3P  
Title : Crystal structure of DEAH-box ATPase Prp22 with bound ssRNA  
Authors : Hamann, F.; Ficner, R.  
Deposited on : 2018-11-07  
Resolution : 2.75 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

<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)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

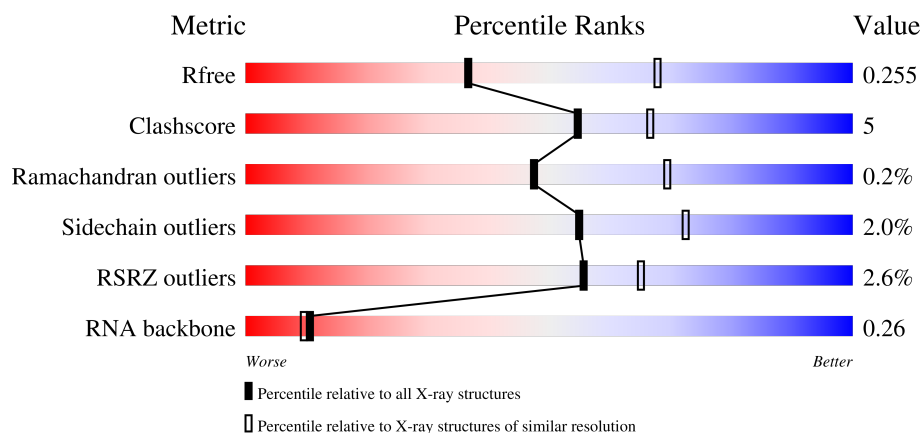
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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.





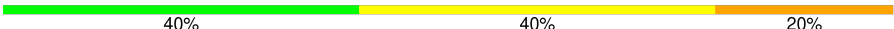

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)
RNA backbone	3102	1060 (3.02-2.50)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	677	<div> <div>3%</div> <div>78%</div> <div>14%</div> <div>8%</div> </div>
1	B	677	<div> <div>3%</div> <div>81%</div> <div>10%</div> <div>8%</div> </div>
1	C	677	<div> <div>2%</div> <div>82%</div> <div>10%</div> <div>7%</div> </div>
1	D	677	<div> <div>2%</div> <div>83%</div> <div>10%</div> <div>7%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	E	10	 60% 40%
2	F	10	 60% 30% 10%
2	G	10	 40% 40% 20%
2	H	10	 50% 50%

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 20540 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.

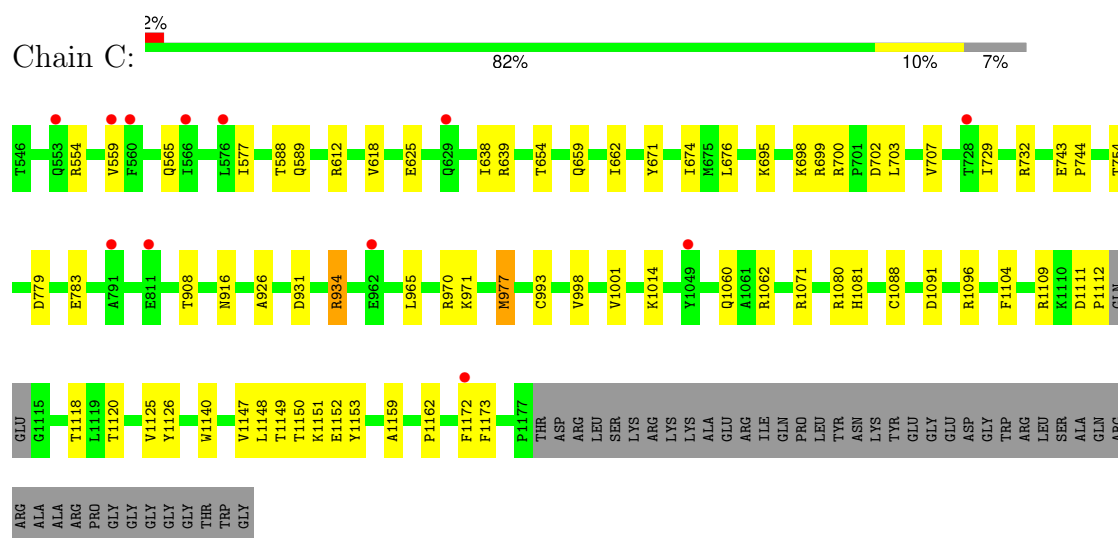
- Molecule 1 is a protein called Putative pre-mRNA splicing factor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	626	Total	C	N	O	S	0	0	0
			4934	3147	836	922	29			
1	B	623	Total	C	N	O	S	0	0	0
			4917	3135	833	920	29			
1	C	630	Total	C	N	O	S	0	0	0
			4976	3171	844	931	30			
1	D	630	Total	C	N	O	S	0	0	0
			4977	3172	844	931	30			

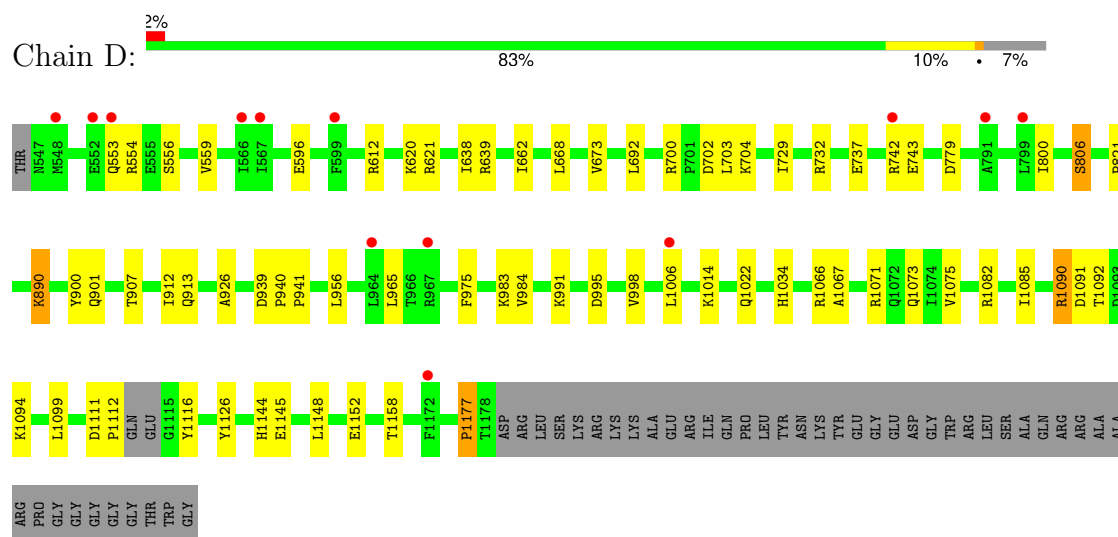
- Molecule 2 is a RNA chain called RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	10	Total	C	N	O	P	0	0	0
			184	81	18	75	10			
2	F	10	Total	C	N	O	P	0	0	0
			184	81	18	75	10			
2	G	10	Total	C	N	O	P	0	0	0
			184	81	18	75	10			
2	H	10	Total	C	N	O	P	0	0	0
			184	81	18	75	10			





- Molecule 1: Putative pre-mRNA splicing factor



- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')



- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')





- Molecule 2: RNA (5'-R(P\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*UP\*U)-3')

Chain H:  50% 50%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	139.79Å 140.50Å 159.55Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.24 – 2.75 49.24 – 2.75	Depositor EDS
% Data completeness (in resolution range)	99.9 (49.24-2.75) 100.0 (49.24-2.75)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.66 (at 2.77Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, $R_{free}$	0.222 , 0.253 0.224 , 0.255	Depositor DCC
$R_{free}$ test set	4107 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	72.2	Xtriage
Anisotropy	0.459	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 45.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	20540	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	78.0	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 38.45 % 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 3.6788e-04. 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 [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	0/5036	0.71	0/6828
1	B	0.52	0/5017	0.73	0/6798
1	C	0.52	0/5080	0.71	0/6887
1	D	0.50	0/5081	0.70	0/6889
2	E	0.62	1/201 (0.5%)	1.19	0/309
2	F	0.85	2/201 (1.0%)	1.09	0/309
2	G	0.62	0/201	1.20	0/309
2	H	0.76	1/201 (0.5%)	1.17	0/309
All	All	0.53	4/21018 (0.0%)	0.74	0/28638

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	8	U	O3'-P	-5.92	1.54	1.61
2	F	1	U	O3'-P	-5.65	1.54	1.61
2	H	3	U	O3'-P	-5.58	1.54	1.61
2	E	1	U	O3'-P	-5.11	1.55	1.61

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 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	4934	0	4957	61	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	4917	0	4941	53	0
1	C	4976	0	5005	50	0
1	D	4977	0	5009	59	0
2	E	184	0	90	1	0
2	F	184	0	90	1	0
2	G	184	0	90	3	0
2	H	184	0	90	3	0
All	All	20540	0	20272	214	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 5.

The worst 5 of 214 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:D:662:ILE:CD1	1:D:668:LEU:HG	1.72	1.19
1:C:700:ARG:NH1	1:C:703:LEU:HB2	1.58	1.18
1:A:921:ILE:HD11	1:A:946:MET:CE	1.83	1.09
1:C:977:MET:CE	1:C:1001:VAL:HG11	1.91	1.00
1:D:662:ILE:HD11	1:D:668:LEU:HG	1.40	1.00

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	622/677 (92%)	605 (97%)	16 (3%)	1 (0%)	47 69
1	B	617/677 (91%)	598 (97%)	18 (3%)	1 (0%)	47 69
1	C	626/677 (92%)	608 (97%)	18 (3%)	0	100 100
1	D	626/677 (92%)	607 (97%)	17 (3%)	2 (0%)	41 60

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	2491/2708 (92%)	2418 (97%)	69 (3%)	4 (0%)	47 69

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	1091	ASP
1	D	1177	PRO
1	A	842	TYR
1	B	643	VAL

### 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	539/581 (93%)	528 (98%)	11 (2%)	55 72
1	B	537/581 (92%)	527 (98%)	10 (2%)	57 73
1	C	545/581 (94%)	533 (98%)	12 (2%)	52 70
1	D	545/581 (94%)	535 (98%)	10 (2%)	59 75
All	All	2166/2324 (93%)	2123 (98%)	43 (2%)	55 72

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	934	ARG
1	D	890	LYS
1	C	970	ARG
1	C	1172	PHE
1	D	975	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 17 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	1072	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	1106	ASN
1	B	659	GLN
1	B	680	HIS
1	C	573	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	E	8/10 (80%)	2 (25%)	0
2	F	8/10 (80%)	2 (25%)	1 (12%)
2	G	8/10 (80%)	4 (50%)	1 (12%)
2	H	8/10 (80%)	2 (25%)	0
All	All	32/40 (80%)	10 (31%)	2 (6%)

5 of 10 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	E	5	U
2	E	9	U
2	F	5	U
2	F	9	U
2	G	2	U

All (2) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	F	9	U
2	G	2	U

## 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 [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	626/677 (92%)	0.40	17 (2%) 54 63	53, 77, 109, 152	0
1	B	623/677 (92%)	0.37	23 (3%) 41 49	54, 74, 102, 128	0
1	C	630/677 (93%)	0.28	12 (1%) 66 75	53, 73, 102, 123	0
1	D	630/677 (93%)	0.25	13 (2%) 63 72	53, 75, 107, 133	0
2	E	10/10 (100%)	0.59	0 100 100	64, 84, 110, 116	0
2	F	10/10 (100%)	0.44	0 100 100	63, 73, 86, 101	0
2	G	10/10 (100%)	0.28	0 100 100	63, 79, 96, 107	0
2	H	10/10 (100%)	0.20	0 100 100	67, 84, 104, 107	0
All	All	2549/2748 (92%)	0.32	65 (2%) 56 65	53, 75, 106, 152	0

The worst 5 of 65 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	822	PRO	4.7
1	A	550	ILE	3.8
1	C	728	THR	3.6
1	B	602	TYR	3.6
1	A	566	ILE	3.5

### 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.