



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 04:09 pm GMT

PDB ID : 4DB2  
Title : Mss116p DEAD-box helicase domain 2 bound to an RNA duplex  
Authors : Mallam, A.L.; Del Campo, M.; Gilman, B.D.; Sidote, D.J.; Lambowitz, A.  
Deposited on : 2012-01-13  
Resolution : 3.16 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

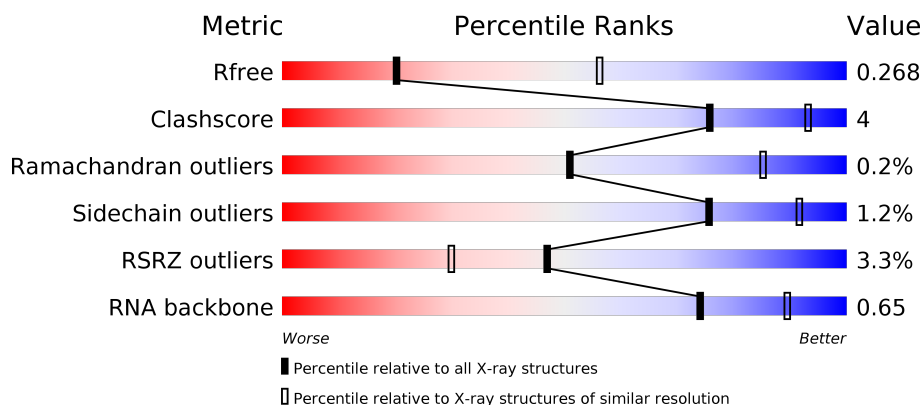
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.16 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	1259 (3.20-3.12)
Clashscore	112137	1397 (3.20-3.12)
Ramachandran outliers	110173	1368 (3.20-3.12)
Sidechain outliers	110143	1367 (3.20-3.12)
RSRZ outliers	101464	1264 (3.20-3.12)
RNA backbone	2435	1000 (3.52-2.80)

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 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	257	<div> <div>92%</div> <div>7%</div> <div>•</div> </div>
1	B	257	<div> <div>93%</div> <div>7%</div> </div>
1	C	257	<div> <div>88%</div> <div>11%</div> <div>•</div> </div>
1	D	257	<div> <div>84%</div> <div>8%</div> <div>8%</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	E	14	<div><div></div><div>79%21%</div></div>
2	F	14	<div><div></div><div>100%</div></div>
2	G	14	<div><div></div><div>43%29%21%7%</div></div>
2	H	14	<div><div></div><div>57%29%7%7%</div></div>
2	I	14	<div><div></div><div>64%21%14%</div></div>
2	J	14	<div><div></div><div>100%</div></div>

## 2 Entry composition [i](#)

There are 2 unique types of molecules in this entry. The entry contains 17989 atoms, of which 8512 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 MSS116, mitochondrial.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	256	Total	C	H	N	O	S	0	0	0
			3842	1246	1902	314	374	6			
1	B	256	Total	C	H	N	O	S	0	0	0
			3912	1258	1950	326	372	6			
1	C	256	Total	C	H	N	O	S	0	0	0
			3892	1261	1925	322	378	6			
1	D	237	Total	C	H	N	O	S	0	0	0
			3662	1188	1815	296	358	5			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	341	SER	-	EXPRESSION TAG	UNP P15424
A	597	TYR	-	EXPRESSION TAG	UNP P15424
B	341	SER	-	EXPRESSION TAG	UNP P15424
B	597	TYR	-	EXPRESSION TAG	UNP P15424
C	341	SER	-	EXPRESSION TAG	UNP P15424
C	597	TYR	-	EXPRESSION TAG	UNP P15424
D	341	SER	-	EXPRESSION TAG	UNP P15424
D	597	TYR	-	EXPRESSION TAG	UNP P15424

- Molecule 2 is a RNA chain called 5'-R(\*GP\*GP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*CP\*GP\*C P\*CP\*C)-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	E	14	Total	C	H	N	O	P	0	0	0
			454	133	156	56	96	13			
2	F	14	Total	C	H	N	O	P	0	0	0
			454	133	156	56	96	13			
2	G	13	Total	C	H	N	O	P	0	0	0
			411	120	140	51	88	12			
2	H	14	Total	C	H	N	O	P	0	0	0
			454	133	156	56	96	13			

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	I	14	Total	C	H	N	O	P	0	0	0
			454	133	156	56	96	13			
2	J	14	Total	C	H	N	O	P	0	0	0
			454	133	156	56	96	13			

### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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.

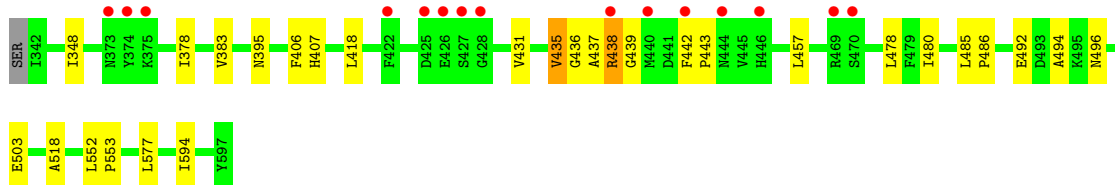
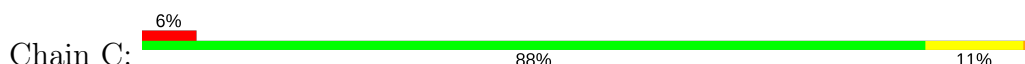
- Molecule 1: ATP-dependent RNA helicase MSS116, mitochondrial



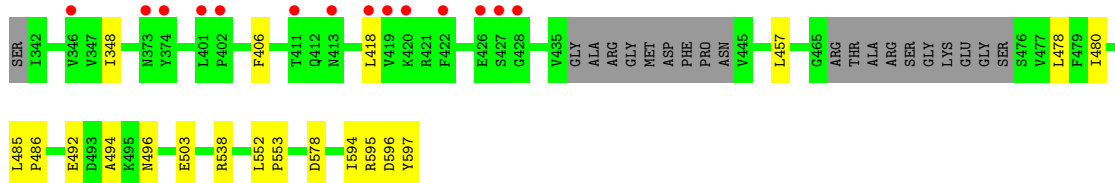
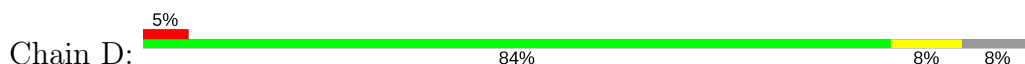
- Molecule 1: ATP-dependent RNA helicase MSS116, mitochondrial



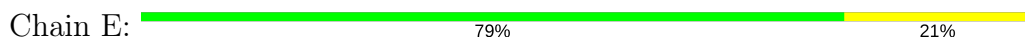
- Molecule 1: ATP-dependent RNA helicase MSS116, mitochondrial



- Molecule 1: ATP-dependent RNA helicase MSS116, mitochondrial



- Molecule 2: 5'-R(\*GP\*GP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*CP\*C)-3'





- Molecule 2: 5'-R(\*GP\*GP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*CP\*C)-3'

Chain F: 100%

There are no outlier residues recorded for this chain.

- Molecule 2: 5'-R(\*GP\*GP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*CP\*C)-3'

Chain G: 43% 29% 21% 7%



- Molecule 2: 5'-R(\*GP\*GP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*CP\*C)-3'

Chain H: 57% 29% 7% 7%



- Molecule 2: 5'-R(\*GP\*GP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*CP\*C)-3'

Chain I: 64% 21% 14%



- Molecule 2: 5'-R(\*GP\*GP\*GP\*CP\*GP\*GP\*GP\*CP\*CP\*CP\*GP\*CP\*CP\*C)-3'

Chain J: 100%

There are no outlier residues recorded for this chain.

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	160.53Å 88.42Å 121.24Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.62 – 3.16 42.62 – 3.16	Depositor EDS
% Data completeness (in resolution range)	99.2 (42.62-3.16) 87.6 (42.62-3.16)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.41 (at 3.19Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.224 , 0.268 0.218 , 0.268	Depositor DCC
$R_{free}$ test set	1517 reflections (5.04%)	DCC
Wilson B-factor (Å <sup>2</sup> )	65.7	Xtriage
Anisotropy	0.346	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 40.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	17989	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	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 28.01 % 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 1.9860e-03. 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	0.22	0/1974	0.39	0/2677
1	B	0.22	0/1996	0.40	0/2701
1	C	0.22	0/2002	0.39	0/2711
1	D	0.22	0/1878	0.39	0/2543
2	E	0.25	0/332	0.62	0/517
2	F	0.24	0/332	0.64	0/517
2	G	0.37	0/302	0.90	0/471
2	H	0.31	0/332	0.87	2/517 (0.4%)
2	I	0.29	0/332	0.65	0/517
2	J	0.25	0/332	0.65	0/517
All	All	0.23	0/9812	0.49	2/13688 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	7	G	C4-C5-N7	5.46	112.98	110.80
2	H	7	G	N9-C4-C5	-5.37	103.25	105.40

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	1940	1902	1896	12	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1962	1950	1944	9	0
1	C	1967	1925	1921	17	0
1	D	1847	1815	1811	10	0
2	E	298	156	156	2	0
2	F	298	156	156	0	0
2	G	271	140	141	4	0
2	H	298	156	156	4	0
2	I	298	156	156	7	0
2	J	298	156	156	0	0
All	All	9477	8512	8493	62	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 4.

The worst 5 of 62 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)
2:I:3:G:H2'	2:I:4:G:H5'	1.67	0.76
2:G:7:G:O6	2:H:7:G:O6	2.06	0.72
2:I:9:C:H2'	2:I:10:C:H5''	1.75	0.68
2:H:12:C:H2'	2:H:13:C:H5''	1.79	0.63
2:I:3:G:H2'	2:I:4:G:C5'	2.32	0.59

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	254/257 (99%)	245 (96%)	9 (4%)	0	100	100
1	B	254/257 (99%)	248 (98%)	6 (2%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	254/257 (99%)	243 (96%)	10 (4%)	1 (0%)	38	76
1	D	231/257 (90%)	222 (96%)	8 (4%)	1 (0%)	38	76
All	All	993/1028 (97%)	958 (96%)	33 (3%)	2 (0%)	51	85

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	438	ARG
1	D	596	ASP

### 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	208/233 (89%)	206 (99%)	2 (1%)	80	93
1	B	211/233 (91%)	208 (99%)	3 (1%)	71	90
1	C	209/233 (90%)	207 (99%)	2 (1%)	80	93
1	D	202/233 (87%)	199 (98%)	3 (2%)	70	90
All	All	830/932 (89%)	820 (99%)	10 (1%)	75	91

5 of 10 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	578	ASP
1	C	435	VAL
1	D	503	GLU
1	B	503	GLU
1	C	503	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	395	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	E	13/14 (92%)	0	0
2	F	13/14 (92%)	0	0
2	G	12/14 (85%)	6 (50%)	0
2	H	13/14 (92%)	3 (23%)	0
2	I	13/14 (92%)	2 (15%)	0
2	J	13/14 (92%)	0	0
All	All	77/84 (91%)	11 (14%)	0

5 of 11 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	G	2	G
2	G	9	C
2	G	10	C
2	G	11	G
2	G	12	C

There are no RNA pucker outliers to report.

## 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	256/257 (99%)	-0.02	3 (1%) 79 66	40, 63, 106, 148	0
1	B	256/257 (99%)	0.07	4 (1%) 72 58	41, 67, 122, 134	0
1	C	256/257 (99%)	0.18	15 (5%) 23 12	41, 74, 116, 172	0
1	D	237/257 (92%)	0.30	14 (5%) 23 12	46, 74, 116, 154	0
2	E	14/14 (100%)	-0.04	0 100 100	48, 59, 67, 71	0
2	F	14/14 (100%)	-0.01	0 100 100	49, 53, 69, 78	0
2	G	13/14 (92%)	0.92	0 100 100	82, 124, 156, 168	0
2	H	14/14 (100%)	0.34	0 100 100	104, 139, 149, 154	0
2	I	14/14 (100%)	-0.00	0 100 100	52, 63, 86, 106	0
2	J	14/14 (100%)	0.17	0 100 100	53, 58, 99, 110	0
All	All	1088/1112 (97%)	0.14	36 (3%) 47 30	40, 69, 123, 172	0

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	469	ARG	5.8
1	B	582	LEU	5.6
1	D	426	GLU	5.0
1	D	427	SER	4.8
1	B	572	VAL	4.3

### 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.