



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2018 – 09:01 am GMT

PDB ID : 1E7K  
Title : Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment  
Authors : Vidovic, I.; Nottrott, S.; Harthmuth, K.; Luhrmann, R.; Ficner, R.  
Deposited on : 2000-08-29  
Resolution : 2.90 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Xtriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30686

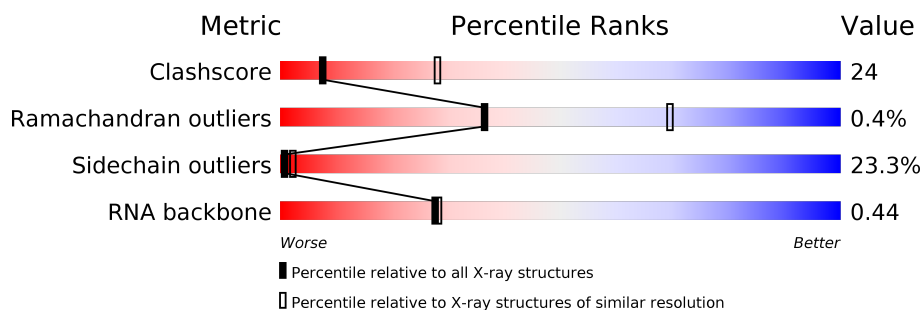
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
Clashscore	122078	1924 (2.90-2.90)
Ramachandran outliers	120005	1884 (2.90-2.90)
Sidechain outliers	119972	1886 (2.90-2.90)
RNA backbone	2633	1057 (3.20-2.60)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	128	
1	B	128	
2	C	22	
2	D	22	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2655 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 15.5 KD RNA BINDING PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	125	Total	C	N	O	S	0	0	0
			968	611	172	180	5			
1	B	123	Total	C	N	O	S	0	0	0
			955	604	170	176	5			

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

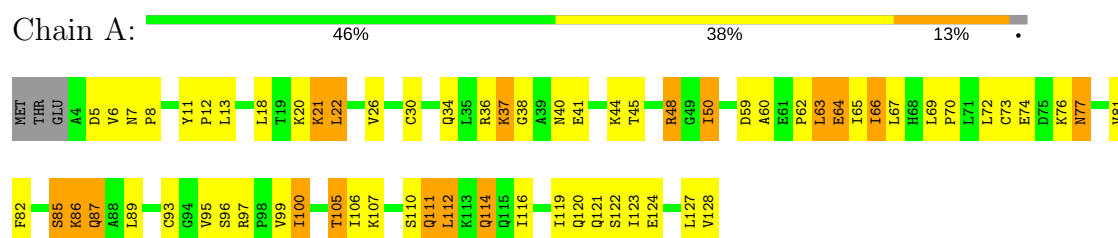
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	17	Total	C	N	O	P	0	0	0
			366	164	72	114	16			
2	D	17	Total	C	N	O	P	0	0	0
			366	164	72	114	16			

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

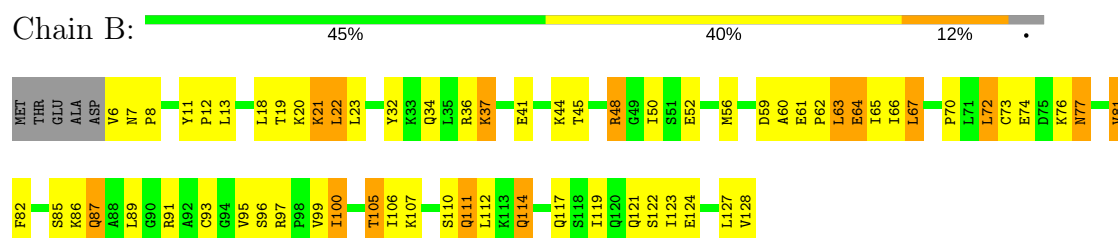
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are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

Note EDS was not executed.

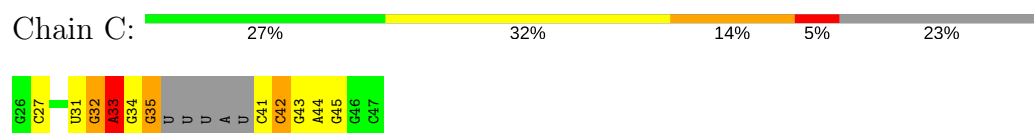
#### • Molecule 1: 15.5 KD RNA BINDING PROTEIN



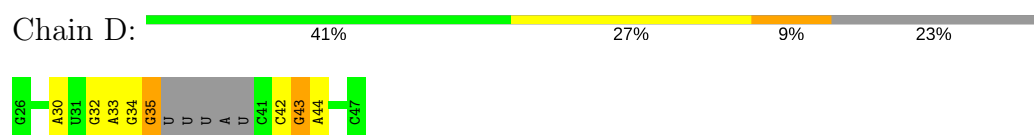
#### • Molecule 1: 15.5 KD RNA BINDING PROTEIN



#### • Molecule 2: RNA (5'-R(\*GP\*CP\*CP\*AP\*AP\*UP\*GP\*AP\*GP\*GP\*UP\*UP\*UP\* AP\*UP\*C P\*CP\*GP\*AP\*GP\*G\*C(-3'))



#### • Molecule 2: RNA (5'-R(\*GP\*CP\*CP\*AP\*AP\*UP\*GP\*AP\*GP\*GP\*UP\*UP\*UP\* AP\*UP\*C P\*CP\*GP\*AP\*GP\*G\*C(-3'))



## 4 Data and refinement statistics

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	45.76Å 55.29Å 146.74Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.90	Depositor
% Data completeness (in resolution range)	93.0 (30.00-2.90)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
Refinement program	X-PLOR 3.851	Depositor
R, $R_{free}$	0.216 , 0.302	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	2655	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

## 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	0.39	0/980	0.60	0/1323
1	B	0.41	0/967	0.61	0/1305
2	C	0.50	0/409	0.81	1/635 (0.2%)
2	D	0.46	0/409	0.77	0/635
All	All	0.43	0/2765	0.67	1/3898 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	33	A	C5'-C4'-C3'	-5.21	107.66	116.00

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	968	0	1017	53	1
1	B	955	0	1008	53	0
2	C	366	0	189	10	1
2	D	366	0	189	4	0
All	All	2655	0	2403	118	1

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 24.

The worst 5 of 118 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:63:LEU:HD13	1:A:67:LEU:HD21	1.41	1.02
1:B:95:VAL:HG12	1:B:97:ARG:H	1.27	0.96
1:B:63:LEU:HD13	1:B:67:LEU:HD21	1.45	0.96
1:A:95:VAL:HG12	1:A:97:ARG:H	1.34	0.90
2:C:35:G:H1	2:C:41:C:H42	1.16	0.90

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:121:GLN:OE1	2:C:27:C:O2'[1_655]	2.07	0.13

## 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	123/128 (96%)	115 (94%)	8 (6%)	0	100	100
1	B	121/128 (94%)	116 (96%)	4 (3%)	1 (1%)	21	54
All	All	244/256 (95%)	231 (95%)	12 (5%)	1 (0%)	36	69

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	67	LEU

### 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	108/111 (97%)	80 (74%)	28 (26%)	0	2
1	B	107/111 (96%)	85 (79%)	22 (21%)	1	4
All	All	215/222 (97%)	165 (77%)	50 (23%)	1	2

5 of 50 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	107	LYS
1	A	124	GLU
1	B	111	GLN
1	A	111	GLN
1	A	114	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	114	GLN
1	B	7	ASN
1	B	77	ASN
1	A	87	GLN
1	B	47	ASN

### 5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	15/22 (68%)	4 (26%)	0
2	D	15/22 (68%)	4 (26%)	1 (6%)
All	All	30/44 (68%)	8 (26%)	1 (3%)

5 of 8 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	32	G
2	C	33	A
2	C	35	G
2	C	42	C

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
2	D	35	G

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
2	D	43	G

## 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 [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates [i](#)

EDS was not executed - this section is therefore empty.

### 6.4 Ligands [i](#)

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers [i](#)

EDS was not executed - this section is therefore empty.