



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 05:41 am GMT

PDB ID : 1HVU  
Title : HUMAN IMMUNODEFICIENCY VIRUS TYPE 1 REVERSE TRANSCRIPTASE COMPLEXED WITH A 33-BASE NUCLEOTIDE RNA PSEUDO-KNOT  
Authors : Jaeger, J.; Restle, T.; Steitz, T.A.  
Deposited on : 1998-06-30  
Resolution : 4.75 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

<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  
Xtrriage (Phenix) : **NOT EXECUTED**  
EDS : **NOT EXECUTED**  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : recalc28949

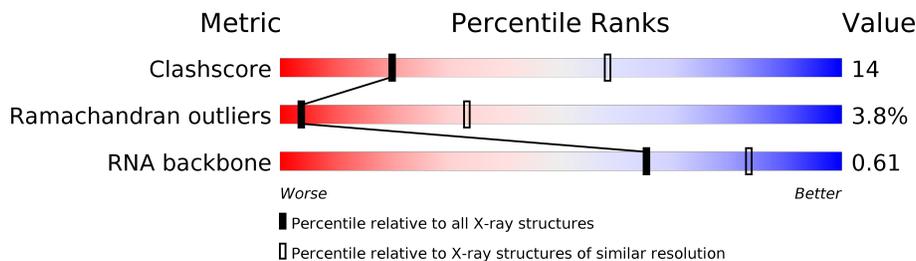
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.75 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))
Clashscore	112137	1063 (5.82-3.70)
Ramachandran outliers	110173	1000 (5.82-3.70)
RNA backbone	2435	1040 (6.46-2.90)

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	C	30	 20% 60% 13% 7%
1	F	30	 20% 60% 13% 7%
1	I	30	 20% 60% 13% 7%
1	L	30	 20% 60% 13% 7%
2	A	554	 87% 11% .
2	D	554	 86% 11% .
2	G	554	 86% 12% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	J	554	 86% 11%
3	B	423	 80% 13% 5%
3	E	423	 81% 13% 5%
3	H	423	 81% 13% 5%
3	K	423	 81% 13% 5%

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 21416 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 RNA chain called RNA (33 NUCLEOTIDE RNA PSEUDOKNOT).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	30	638	287	116	206	29	0	0	0
1	F	30	638	287	116	206	29	0	0	0
1	I	30	638	287	116	206	29	0	0	0
1	L	30	638	287	116	206	29	0	0	0

- Molecule 2 is a protein called PROTEIN (HIV-1 REVERSE TRANSCRIPTASE).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	A	554	2737	1629	554	554	0	0	0
2	D	554	2737	1629	554	554	0	0	0
2	G	554	2737	1629	554	554	0	0	0
2	J	554	2737	1629	554	554	0	0	0

- Molecule 3 is a protein called PROTEIN (HIV-1 REVERSE TRANSCRIPTASE).

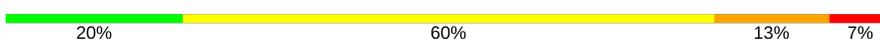
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	B	400	1979	1179	400	400	0	0	0
3	E	400	1979	1179	400	400	0	0	0
3	H	400	1979	1179	400	400	0	0	0
3	K	400	1979	1179	400	400	0	0	0

### 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: RNA (33 NUCLEOTIDE RNA PSEUDOKNOT)

Chain C: 



- Molecule 1: RNA (33 NUCLEOTIDE RNA PSEUDOKNOT)

Chain F: 



- Molecule 1: RNA (33 NUCLEOTIDE RNA PSEUDOKNOT)

Chain I: 



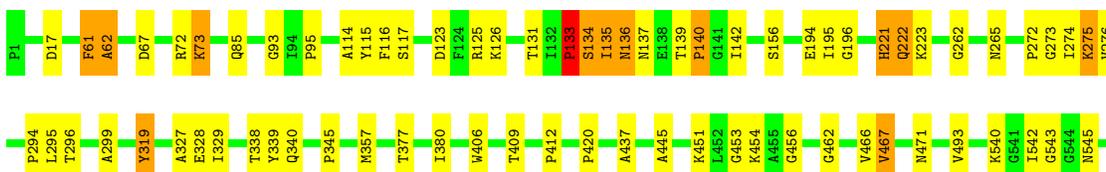
- Molecule 1: RNA (33 NUCLEOTIDE RNA PSEUDOKNOT)

Chain L: 



- Molecule 2: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

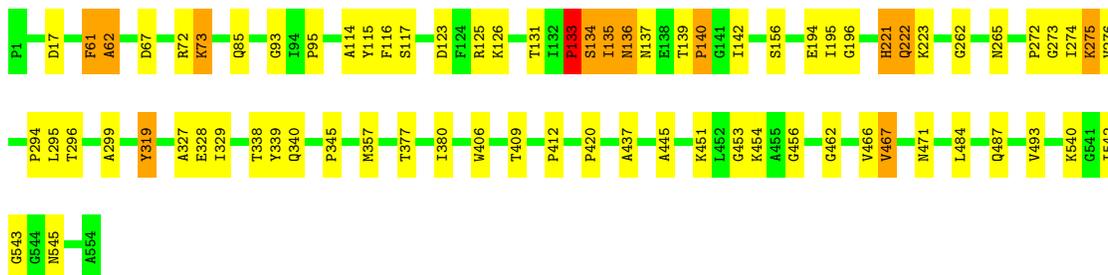
Chain A: 



A554

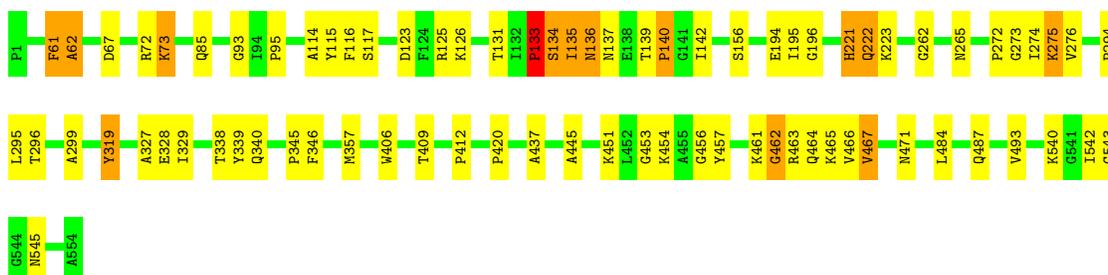
• Molecule 2: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

Chain D: 86% 11%



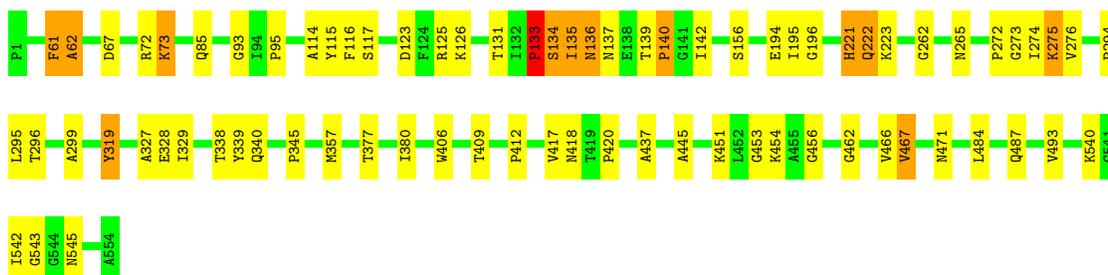
• Molecule 2: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

Chain G: 86% 12%



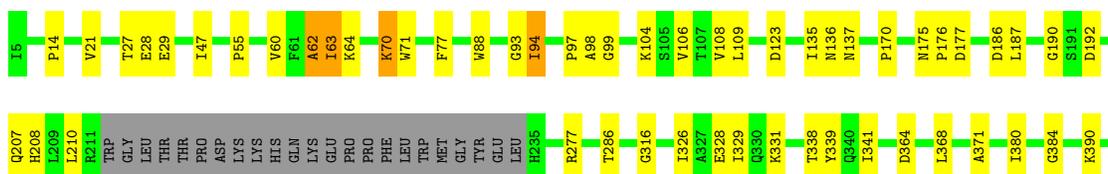
• Molecule 2: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

Chain J: 86% 11%



• Molecule 3: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

Chain B: 80% 13% 5%





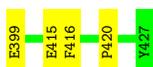
- Molecule 3: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

Chain E: 81% 13% • 5%



- Molecule 3: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

Chain H: 81% 13% • 5%



- Molecule 3: PROTEIN (HIV-1 REVERSE TRANSCRIPTASE)

Chain K: 81% 13% • 5%



## 4 Data and refinement statistics

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

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	166.70Å 169.20Å 331.40Å 90.00° 104.50° 90.00°	Depositor
Resolution (Å)	30.00 – 4.75	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-4.75)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.06	Depositor
Refinement program	X-PLOR 3.853	Depositor
R, $R_{free}$	0.340 , 0.413	Depositor
Estimated twinning fraction	No twinning to report.	Xtrriage
Total number of atoms	21416	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	52.0	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	C	1.30	4/714 (0.6%)	0.97	5/1111 (0.5%)
1	F	1.30	4/714 (0.6%)	0.97	5/1111 (0.5%)
1	I	1.30	4/714 (0.6%)	0.97	5/1111 (0.5%)
1	L	1.30	4/714 (0.6%)	0.97	5/1111 (0.5%)
2	A	0.71	4/2736 (0.1%)	1.00	9/3809 (0.2%)
2	D	0.71	4/2736 (0.1%)	1.00	9/3809 (0.2%)
2	G	0.71	4/2736 (0.1%)	1.00	9/3809 (0.2%)
2	J	0.71	4/2736 (0.1%)	1.00	9/3809 (0.2%)
3	B	0.99	3/1977 (0.2%)	1.05	6/2752 (0.2%)
3	E	0.99	3/1977 (0.2%)	1.05	6/2752 (0.2%)
3	H	0.99	3/1977 (0.2%)	1.05	6/2752 (0.2%)
3	K	0.99	3/1977 (0.2%)	1.05	6/2752 (0.2%)
All	All	0.91	44/21708 (0.2%)	1.01	80/30688 (0.3%)

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
2	A	0	1
2	D	0	1
2	G	0	1
2	J	0	1
All	All	0	4

The worst 5 of 44 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	26	A	O3'-P	27.49	1.94	1.61
1	C	26	A	O3'-P	27.49	1.94	1.61
1	F	26	A	O3'-P	27.49	1.94	1.61
1	I	26	A	O3'-P	27.45	1.94	1.61

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	70	LYS	C-N	24.64	1.90	1.34

The worst 5 of 80 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	K	64	LYS	C-N-CA	12.11	151.97	121.70
3	B	64	LYS	C-N-CA	12.10	151.96	121.70
3	E	64	LYS	C-N-CA	12.10	151.95	121.70
3	H	64	LYS	C-N-CA	12.09	151.92	121.70
1	C	27	A	O5'-P-OP1	10.45	123.24	110.70

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	319	TYR	Mainchain
2	D	319	TYR	Mainchain
2	G	319	TYR	Mainchain
2	J	319	TYR	Mainchain

## 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	C	638	0	323	38	0
1	F	638	0	323	39	0
1	I	638	0	323	37	0
1	L	638	0	323	37	0
2	A	2737	0	1181	48	0
2	D	2737	0	1181	51	0
2	G	2737	0	1181	49	30
2	J	2737	0	1181	51	3
3	B	1979	0	832	35	0
3	E	1979	0	832	34	0
3	H	1979	0	832	34	0
3	K	1979	0	832	34	0
All	All	21416	0	9344	436	30

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 436 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:J:62:ALA:CB	2:J:73:LYS:HA	1.14	1.62
2:A:62:ALA:CB	2:A:73:LYS:HA	1.14	1.59
2:D:62:ALA:CB	2:D:73:LYS:HA	1.14	1.57
2:G:62:ALA:CB	2:G:73:LYS:HA	1.14	1.55
3:E:62:ALA:C	3:E:63:ILE:N	1.70	1.44

The worst 5 of 30 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)
2:G:463:ARG:N	2:G:465:LYS:N[2_656]	0.58	1.62
2:G:463:ARG:C	2:G:464:GLN:N[2_656]	0.70	1.50
2:G:464:GLN:N	2:G:464:GLN:N[2_656]	0.73	1.47
2:G:463:ARG:N	2:G:464:GLN:C[2_656]	0.89	1.31
2:G:463:ARG:C	2:G:464:GLN:CA[2_656]	1.01	1.19

## 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	
2	A	552/554 (100%)	449 (81%)	79 (14%)	24 (4%)	3	31
2	D	552/554 (100%)	449 (81%)	79 (14%)	24 (4%)	3	31
2	G	552/554 (100%)	449 (81%)	79 (14%)	24 (4%)	3	31
2	J	552/554 (100%)	450 (82%)	78 (14%)	24 (4%)	3	31
3	B	396/423 (94%)	322 (81%)	62 (16%)	12 (3%)	5	38
3	E	396/423 (94%)	322 (81%)	62 (16%)	12 (3%)	5	38

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	H	396/423 (94%)	322 (81%)	62 (16%)	12 (3%)	5	38
3	K	396/423 (94%)	322 (81%)	62 (16%)	12 (3%)	5	38
All	All	3792/3908 (97%)	3085 (81%)	563 (15%)	144 (4%)	4	33

5 of 144 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	62	ALA
2	A	134	SER
2	A	135	ILE
2	A	136	ASN
2	A	140	PRO

### 5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	C	29/30 (96%)	6 (20%)	0
1	F	29/30 (96%)	6 (20%)	0
1	I	29/30 (96%)	6 (20%)	0
1	L	29/30 (96%)	6 (20%)	0
All	All	116/120 (96%)	24 (20%)	0

5 of 24 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	C	12	U
1	C	25	A
1	C	26	A
1	C	27	A
1	C	28	A

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 [i](#)

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

### 6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

### 6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

### 6.4 Ligands

EDS was not executed - this section is therefore empty.

### 6.5 Other polymers

EDS was not executed - this section is therefore empty.