



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 18, 2018 – 11:39 am GMT

PDB ID : 2RF2  
Title : HIV reverse transcriptase in complex with inhibitor 7e (NNRTI)  
Authors : Yan, Y.; Prasad, S.  
Deposited on : 2007-09-27  
Resolution : 2.40 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  
Mogul : 1.7.3 (157068), CSD as539be (2018)  
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

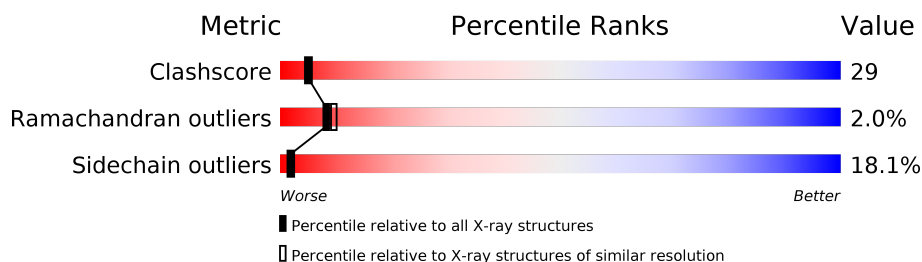
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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	3953 (2.40-2.40)
Ramachandran outliers	120005	3894 (2.40-2.40)
Sidechain outliers	119972	3895 (2.40-2.40)

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	563	
2	B	443	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8593 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 Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 2.7.7.7) (EC 3.1.26.4) (p66 RT).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	554	Total	C	N	O	S	0	0	0
			4510	2915	754	833	8			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MET	-	EXPRESSION TAG	UNP P04585
A	-1	ASN	-	EXPRESSION TAG	UNP P04585
A	0	SER	-	EXPRESSION TAG	UNP P04585

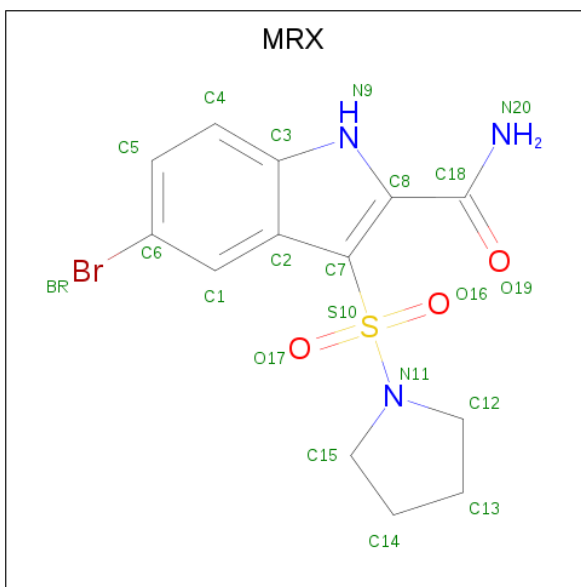
- Molecule 2 is a protein called Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 2.7.7.7) (EC 3.1.26.4) (p66 RT).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	405	Total	C	N	O	S	0	0	0
			3352	2182	555	609	6			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	-2	MET	-	EXPRESSION TAG	UNP P04585
B	-1	ASN	-	EXPRESSION TAG	UNP P04585
B	0	SER	-	EXPRESSION TAG	UNP P04585

- Molecule 3 is 5-bromo-3-(pyrrolidin-1-ylsulfonyl)-1H-indole-2-carboxamide (three-letter code: MRX) (formula: C<sub>13</sub>H<sub>14</sub>BrN<sub>3</sub>O<sub>3</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	A	1	Total	Br	C	N	O	S	0	0
			21	1	13	3	3	1		

- Molecule 4 is water.

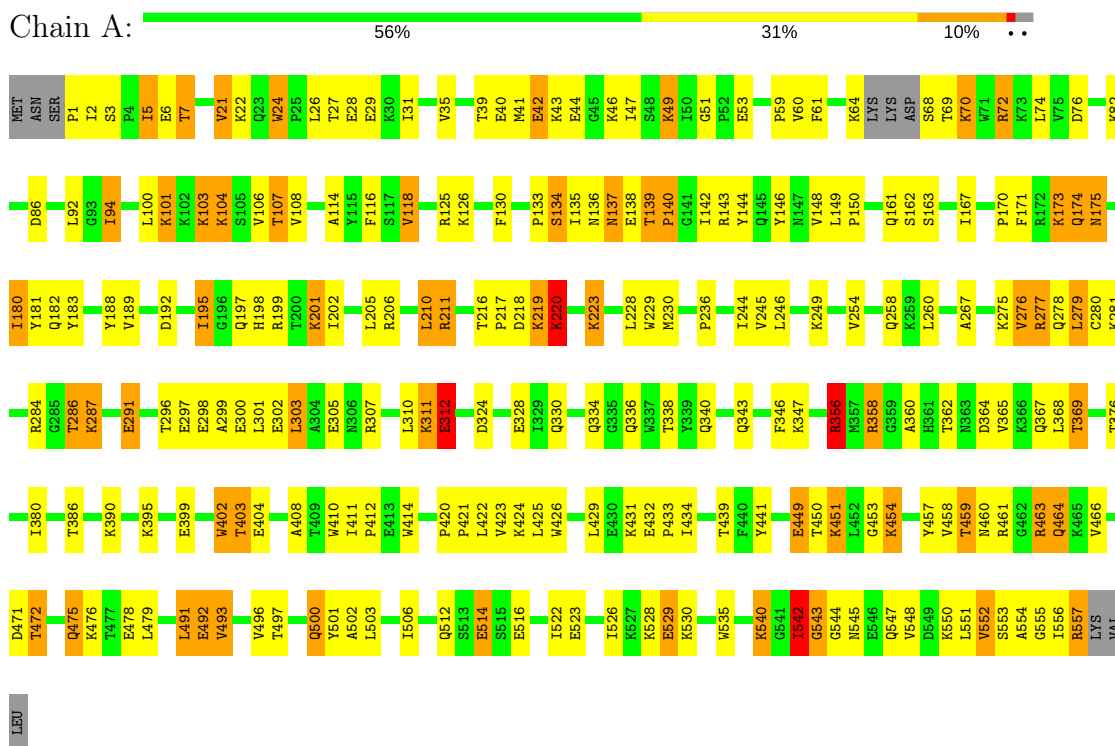
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	413	Total	O	0	0
			413	413		
4	B	297	Total	O	0	0
			297	297		

### 3 Residue-property plots

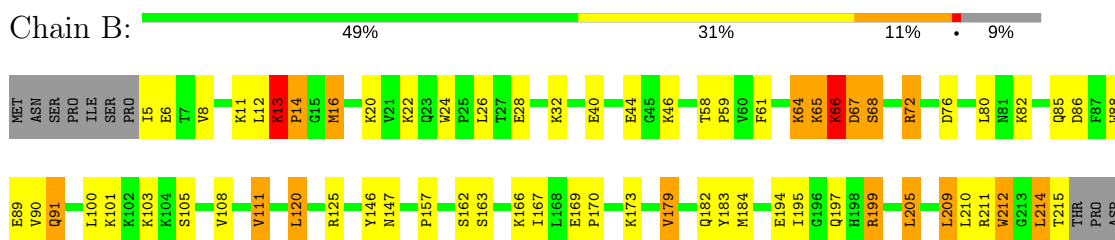
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 2.7.7.7) (EC 3.1.26.4) (p66 RT)



- Molecule 2: Reverse transcriptase/ribonuclease H (EC 2.7.7.49) (EC 2.7.7.7) (EC 3.1.26.4) (p66 RT)



LYS	K281	G352	GLU
LYS	L282	K353	PRO
HIS	L283	Y354	ILE
GLN	R284	A355	VAL
LYS	G285	R356	GLY
GLU	T286	MET	ALA
PRO	K287	ARG	GLU
PRO	A288	GLY	THR
PHE	L289	ALA	PHE
LEU	T290	H361	
TRP		T362	
MET	L293	N363	
G231	P294	D364	
Y232	L295	V365	
E233	T296	K366	
	E297	Q367	
	E298	L368	
P236	A299	T369	
D237	E300	E370	
K238	L301	A371	
Y239	E302	V372	
T240	L303	Q373	
Y241	A304	K374	
Q242	E305	I375	
P243	N306	T376	
I244		T377	
T245	I309	E378	
L246	L310		
	K311	K385	
K249	E312		
D250	P313	K388	
S251	V314	F389	
V252	H315	K390	
T253	Q316		
V254	V317	E396	
N255	Y318		
D256	Y319	T400	
T257	D320	W401	
Q258	P321		
K259		Y405	
L260	D324	W406	
V261	L325	Q407	
G262		A408	
K263	E328	T409	
L264		W410	
N265	Q334	E415	
W266	G335		
A267	Q336	P420	
S268	W337	F421	
Q269		L422	
I270	I341	V423	
Y271	Y342	K424	
P272	Q343	L425	
G273	E344	W426	
I274	P345	Y427	
K275	F346	Q428	
V276	K347	LEU	
R277	N348	GLU	
Q278	L349	LYS	
L279			
C280			

## 4 Data and refinement statistics

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

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.29Å 155.81Å 155.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	15.77 – 2.40	Depositor
% Data completeness (in resolution range)	99.2 (15.77-2.40)	Depositor
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.11	Depositor
Refinement program	BUSTER-TNT 2.1.1	Depositor
R, $R_{free}$	0.185 , 0.258	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	8593	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: MRX

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	0.62	0/4625	0.79	1/6282 (0.0%)
2	B	0.62	0/3446	0.82	4/4682 (0.1%)
All	All	0.62	0/8071	0.80	5/10964 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	13	LYS	C-N-CD	-16.25	84.85	120.60
2	B	271	TYR	C-N-CD	-7.60	103.88	120.60
2	B	120	LEU	CA-CB-CG	5.89	128.84	115.30
2	B	244	ILE	N-CA-C	-5.33	96.60	111.00
1	A	86	ASP	CB-CG-OD2	5.01	122.81	118.30

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	4510	0	4558	258	0
2	B	3352	0	3380	204	0
3	A	21	0	14	3	0
4	A	413	0	0	18	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	297	0	0	12	0
All	All	8593	0	7952	452	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 29.

The worst 5 of 452 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:64:LYS:HE2	1:A:69:THR:HG23	1.21	1.17
2:B:253:THR:HG22	2:B:256:ASP:H	1.07	1.10
2:B:422:LEU:HA	2:B:425:LEU:HD21	1.35	1.08
1:A:463:ARG:HG2	1:A:463:ARG:HH11	1.19	1.06
2:B:246:LEU:HD11	2:B:310:LEU:HD23	1.39	1.03

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	548/563 (97%)	510 (93%)	27 (5%)	11 (2%)	8	9
2	B	399/443 (90%)	372 (93%)	19 (5%)	8 (2%)	8	9
All	All	947/1006 (94%)	882 (93%)	46 (5%)	19 (2%)	8	9

5 of 19 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	139	THR
1	A	356	ARG
1	A	543	GLY

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	14	PRO
2	B	241	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	494/503 (98%)	402 (81%)	92 (19%)	2	2
2	B	369/403 (92%)	305 (83%)	64 (17%)	2	2
All	All	863/906 (95%)	707 (82%)	156 (18%)	2	2

5 of 156 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	461	ARG
1	A	542	ILE
2	B	336	GLN
1	A	464	GLN
1	A	496	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 21 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	509	GLN
1	A	547	GLN
2	B	306	ASN
1	A	507	GLN
2	B	336	GLN

### 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](#)

1 ligand is modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z  > 2$	Counts	RMSZ	# $ Z  > 2$
3	MRX	A	561	-	21,23,23	3.01	8 (38%)	26,35,35	3.92	14 (53%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	MRX	A	561	-	-	0/6/23/23	0/3/3/3

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	561	MRX	C8-C18	-3.20	1.42	1.51
3	A	561	MRX	C5-C6	2.15	1.42	1.38
3	A	561	MRX	O19-C18	2.31	1.28	1.24
3	A	561	MRX	BR-C6	2.67	1.95	1.90
3	A	561	MRX	C1-C6	3.06	1.41	1.36

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	561	MRX	C12-N11-C15	-7.97	101.00	110.56
3	A	561	MRX	O17-S10-N11	-7.00	100.19	106.69
3	A	561	MRX	O16-S10-N11	-5.17	101.89	106.69
3	A	561	MRX	O19-C18-N20	-5.01	115.31	122.60
3	A	561	MRX	C5-C6-C1	-4.85	117.06	121.96

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	561	MRX	3	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	297:GLU	C	298:GLU	N	2.65

## 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.